

Climate Drift in the CMIP5 Models

Author:

Sen Gupta, Alexander; Jourdain, Nicolas; Brown, Jaclyn N.; Monselesan, Didier

Publication details:

Journal of Climate

v. 26

Chapter No. 21

pp. 8597-8615

0894-8755 (ISSN)

Publication Date:

2013

Publisher DOI:

<http://dx.doi.org/10.1175/JCLI-D-12-00521.1>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/53679> in <https://unsworks.unsw.edu.au> on 2024-03-29

Climate Drift in the CMIP5 Models*

ALEXANDER SEN GUPTA

Climate Change Research Centre, and ARC Centre of Excellence for Climate Systems Science, University of New South Wales, Sydney, New South Wales, Australia

NICOLAS C. JOURDAIN

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

JACLYN N. BROWN AND DIDIER MONSELESAN

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

(Manuscript received 26 July 2012, in final form 19 May 2013)

ABSTRACT

Climate models often exhibit spurious long-term changes independent of either internal variability or changes to external forcing. Such changes, referred to as model “drift,” may distort the estimate of forced change in transient climate simulations. The importance of drift is examined in comparison to historical trends over recent decades in the Coupled Model Intercomparison Project (CMIP). Comparison based on a selection of metrics suggests a significant overall reduction in the magnitude of drift from phase 3 of CMIP (CMIP3) to phase 5 of CMIP (CMIP5). The direction of both ocean and atmospheric drift is systematically biased in some models introducing statistically significant drift in globally averaged metrics. Nevertheless, for most models globally averaged drift remains weak compared to the associated forced trends and is often smaller than the difference between trends derived from different ensemble members or the error introduced by the aliasing of natural variability. An exception to this is metrics that include the deep ocean (e.g., steric sea level) where drift can dominate in forced simulations. In such circumstances drift must be corrected for using information from concurrent control experiments. Many CMIP5 models now include ocean biogeochemistry. Like physical models, biogeochemical models generally undergo long spinup integrations to minimize drift. Nevertheless, based on a limited subset of models, it is found that drift is an important consideration and must be accounted for. For properties or regions where drift is important, the drift correction method must be carefully considered. The use of a drift estimate based on the full control time series is recommended to minimize the contamination of the drift estimate by internal variability.

1. Introduction

Model drift refers to spurious long-term changes in general circulation models that are unrelated to either changes in external forcing or internal low-frequency

variability. Drift can be caused by a number of factors. For example, a simulation’s initial state may not be in dynamical balance with the representation of physics in the model; “coupling shock” may occur during the coupling of model components resulting in discontinuities in surface fluxes (e.g., Rahmstorf 1995) or numerical errors may exist in the model that mean that heat or moisture is not fully conserved (e.g., Lucarini and Ragone 2011; Liepert and Previdi 2012). In these cases, a model may drift from its initial state toward a quasi-steady state over some period of time (although in the case of nonconserved heat or water a steady solution may not be attainable). The time scale over which the climate system adjusts will be determined by the time it

*Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-12-00521.s1>.

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

takes for anomalies to be advected or mixed through the ocean, which may be many thousands of years [e.g., Peacock and Maltrud (2006); the adjustment of the atmosphere and land surface is many orders of magnitude faster]. Given the complexity and resolution of modern climate models, spinup periods of thousands of years are prohibitive given the available computational resources and the requirement for numerous transient simulations [e.g., as part of phase 5 of the Coupled Model Intercomparison Project (CMIP5); Taylor et al. (2012); see the appendix for a complete list of model names and expansions]. Instead, models are generally spun up for a few hundred years (although multimillennium spinups and complex multistage spinups are sometimes performed; Table 1). As a result, externally forced climate model experiments (e.g., where changes are made to greenhouse gases, aerosols, ozone, or insolation) are undertaken in models that are often not fully equilibrated and may exhibit changes that are associated with the adjustment process, in addition to any changes that are directly related to external forcing or internal variability.

A primary use of climate models is to help us understand how and why changes in external forcing drive changes to the climate system, both in the past (hindcasts) or the future (projections). However, drift can contaminate the externally forced signal, masking the resulting climate change. It is therefore necessary to understand how large drift is in comparison to any forced signal, under what circumstances drift may be neglected, and where drift cannot be neglected, how best to correct for that drift.

The relative importance of drift has been recently assessed by Sen Gupta et al. (2012, hereafter SG12) for models from phase 3 of the Coupled Model Intercomparison Project (CMIP3) that were used to provide projections for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (Solomon et al. 2007) and prior to that by Covey et al. (2006) for the suite of models from phase 2+ of the Coupled Model Intercomparison Project (CMIP2+). Here, we extend this work to examine the selected models taking part in the latest CMIP5 intercomparison that will be used to inform the IPCC Fifth Assessment Report. We assess to what extent, and where, drift continues to be important in these latest model runs.

SG12 put forward some general conclusions with regard to the CMIP3 suite of models. These include:

- Drift shows little systematic directional bias either from region to region or from model to model. As a result, drift generally becomes less important (compared to any forced trend) for larger regions or when considering averages across multiple models.

- Given that drift affects the full ocean, while forced changes, at least over the historical period, are usually confined to the upper few hundred meters (except in high-latitude regions), drift generally dominates any forced signal below 1–2 km. As such, any examination of subsurface changes or depth-integrated changes (e.g., steric sea level) must pay particular attention to drift and the method used for the correction of drift. SG12 gave examples of drift correction reversing the sign of both regional and globally averaged steric sea level rise.
- Even though the adjustment time scale of the atmosphere is fast, as the ocean is coupled to the atmosphere, if surface ocean properties drift then atmospheric properties will also drift.

The aim of this study is to document the scale of drift in a selection of physical and biogeochemical properties and provide advice on under what circumstances it is important to correct for drift and under what circumstances drift can be neglected.

2. Method

In our examination of climate drift we use output from the CMIP5 (Taylor et al. 2012) and CMIP3 (Meehl et al. 2007) initiatives of the World Climate Research Programme to bring together output from an unprecedented array of climate and Earth system models (ESMs).

To assess the importance of model drift with respect to forced trends in the CMIP5 models, we examine the period 1945–2005 from the historical simulations (which are generally forced by historical observations of greenhouse gases, aerosols, ozone, and insolation). The 1945–2005 period is chosen as it is primarily over the latter half of the twentieth century that a forced signal has become distinguishable from multidecadal- to centennial-scale climate variability (e.g., Hegerl et al. 1996). We use a longer period than SG12 (who examined 1950–2000) as the CMIP5 historical simulations are integrated to at least 2005 (as opposed to 2000 for CMIP3), and by using the longer period the error associated with the aliasing of internal variability in the calculation of trends is reduced.

The drift corresponding to this historical period is estimated using a number of different methods.

- The “150-yr linear drift” uses a 150-yr linear trend (1900–2050) bracketing the historical (1945–2005) period. We also examine a shorter 100-yr period. This method was used in SG12 and in other studies (e.g., Downes et al. 2010; Sen Gupta et al. 2009).
- The “quadratic or cubic drift” uses a quadratic or cubic polynomial fit to the full available control time

TABLE 1. Available information on model spinup. See appendix for complete expansions of model names.

Model	Spinup information	Reference
ACCESS1.0	—	Bureau of Meteorology (2010)
BCC-CSM1.1	The initial condition is derived from a 200-yr control run with external forcings fixed at 1850.	http://q.cmp5.ceda.ac.uk/
CanESM2	—	Chylek et al. (2011)
CCSM4	After coupling, model integrated for few 100 yr with adjustments made to sea ice albedo and relative humidity threshold to achieve radiative balance at the top of atmosphere. Model was then spun up for a further few 100 yr followed by >1000-yr control simulation.	Gent et al. (2011)
CNRM-CM5	The 200-yr coupled spinup simulation from World Ocean Database 2005 ocean state at rest. Ten historical ensemble simulations were initialized at 50-yr intervals after the spinup (as such the last ensemble had the equivalent of a 700-yr spinup).	Voldoire et al. (2013)
CSIRO Mk3.6.0	The 260-yr spinup run to allow the model to equilibrate after the inclusion of a new interactive aerosol scheme from previous model version (Mk3.5.0).	Rotstayn et al. (2010, 2012)
EC-EARTH	—	Hazeleger et al. (2010)
GFDL-CM3	Coupled spinup of ~2000 yr.	S. Griffies (2011, personal communication); Donner et al. (2011)
GFDL-ESM2G	Started from year 1601 of a spinup run.	http://q.cmp5.ceda.ac.uk/
GFDL-ESM2M	Started from year 2401 of a spinup run.	http://q.cmp5.ceda.ac.uk/
GISS-E2-H	—	Schmidt et al. (2006)
GISS-E2-R	—	Schmidt et al. (2006)
HadGEM2-CC	Ocean model initialized from earlier model version and integrated for 400 yr. Coupled spinup length not specified.	Martin et al. (2011); Collins et al. (2011)
HadGEM2-ES	Initial condition for control obtained as final point of long run (~200 yr) with preindustrial conditions.	http://q.cmp5.ceda.ac.uk/ ; Volodin et al. (2010)
INM-CM4.0	—	
IPSL-CM5A-LR	Four stages: (i) 2500-yr spinup of ocean (without carbon cycle), (ii) 600-yr spinup with carbon cycle, (iii) few thousand years off line spinup of ocean and land carbon cycle models (using output from 600-yr spinup above) to equilibrate biogeochemistry, and (iv) 400-yr fully coupled.	Dufresne et al. (2013)
MIROC-ESM	Multistage spinup using coupled and standalone models for a few thousand years followed by 180-yr coupled spinup in low-resolution model version, followed by 100-yr online spinup of full-resolution model.	Watanabe et al. (2011)
MIROC-ESM-CHEM	Offline integration of atmospheric chemistry model for ~40 yr. Chemistry spinup was coupled to MIROC-ESM at year 20 of piControl and integrated for 4 yr to reach equilibrium.	Watanabe et al. (2011)
MIROC5	Ocean and atmosphere spun up independently (530 yr for ocean). Followed by coupled spinup of ~1000 yr (while parameter values tuned to achieve a realistic climate).	Watanabe et al. (2010)
MPI-ESM-LR	—	Raddatz et al. (2007)
MPI-ESM-P	—	Raddatz et al. (2007)
MRI-CGCM3	—	Yukimoto et al. (2011)
NorESM1-M	Model spinup of 700 yr with 1850 forcing conditions, from prescribed initial conditions.	http://q.cmp5.ceda.ac.uk/

series. Cubic or quadratic drift curves have been used to remove drift in previous studies examining sea level change (e.g., Gregory et al. 2001, 2006; Ammann et al. 2007; Gleckler et al. 2012). The effective drift corresponding to the 1945–2005 period is the linear trend of the quadratic or cubic curve over this period.

- The “full linear drift” uses a linear trend based on the entire available control time series (e.g., Wang 2013).

The drift estimated using one of the above methods can be subtracted from the 1945–2005 historical trend to obtain an estimate of the true forced response. This method assumes that the drift component that is present in the preindustrial control remains relatively unmodified in the associated transient simulation.

Simulated forced trends usually become stronger in the future (as a result of increased greenhouse forcing),

while drift might be expected to diminish (moving closer toward equilibration; although we demonstrate below that this may not necessarily be the case). As such, our analysis represents a worst-case limit for the importance of drift relative to the forced signal; the importance of drift will generally be reduced when considering climate projections resulting from the presence of much larger magnitude forced changes.

Historical and control data were annually averaged prior calculating linear trends or other least squares fits. Metadata information (including the historical simulation branch time) was used to temporally align the control simulation to the historical simulations for all ensemble members. Multimodel means or medians are calculated by first averaging across ensemble members (where available) for individual models and then averaging across ensemble means. The significance of linear trends are tested using a Student's *t* test (at 90% level) where the degrees of freedom are adjusted to account for lag-1 autocorrelation (Santer et al. 2000 and references therein). We note however that the estimation of statistical significance may be inflated by the presence of long-term persistence (e.g., Cohn and Lins 2005). Model variable and ensemble information is provided in Table 2 with tracking information for model, variable, and ensemble combinations provided in Table S1 of the supplemental material.

There is a large spread in the length of the control time series made available by the different modeling groups, ranging from 100 years for the high-resolution MIROC4h model to over 1000 years. The median available control simulation length has increased from 375 years for CMIP3 (based on 22 models) to 500 years for CMIP5 (based on 30 models). A simulation length of 500 years (after a suitable spinup period) is the minimum recommended control simulation length for inclusion in CMIP5 (Taylor et al. 2012). The length of the spinup period also varies substantially across the models (see discussion).

3. Results

a. Sea surface temperature

Figure 1a shows the temporal evolution of globally averaged SST for preindustrial control simulations, filtered using a cubic polynomial. It is clear that for SST the low-frequency evolution is generally nonlinear. In some models (e.g., MIROC5 or CNRM-CM3), it would seem that the rate of change is decreasing with time, as we might expect for a model moving toward a steady state. However, this is not the case for many models, some of which exhibit an increasing rate of change over time (e.g., GFDL-CM3), while others show more complex

TABLE 2. Number of historical ensemble members used for each model variable [sea surface temperature (Tos), steric sea level (Zossga), precipitation (Pr), sea ice concentration (Sic), IntDIC, and O₂] and length of available piControl simulation (right column). Only a single ensemble member is available for each piControl simulation.

Model	Tos	Zossga	Pr	Sic	IntDIC	O ₂	piControl length (yr)
ACCESS1.0	1	1	1	1			500
BCC-CSM1.1	3	3	3				400
BCC-CSM1.1(m)		3					500
CCSM4	4	5	6				501
CMCC-CM		1					500
CMCC-CMS		1					330
CNRM-CM5	10	7	10	10			850
CSIRO Mk3.6.0	10		10	10			500
CanESM2	5	5	5	1			996
EC-EARTH	10						451
GFDL-CM3	1		5	5			500
GFDL-ESM2G		1	3				500
GFDL-ESM2M		1		1			500
GISS-E2-H	5		5	1			480
GISS-E2-R	10		4	1			849
HadGEM2-CC	1	1	4		1		240
HadGEM2-ES	2	4		1	1	1	576
INM-CM4.0	1		1	1			500
IPSL-CM5A-LR	4	6	4	2	1	1	1000
IPSL-CM5A-MR		3					300
MIROC-ESM		3	3	1			255
MIROC-ESM-CHEM		1	1				531
MIROC4h		3	1	1			100
MIROC5	3	5	3	1			670
MPI-ESM-LR	3	3			1		1000
MPI-ESM-MR		3					1000
MPI-ESM-P	2	2					1156
MRI-CGCM3	3	3	3	1			500
NorESM1-M	3	3	3	3			252
NorESM1-ME		1	1	1			501

behavior. Figures 2a–c show examples of annually averaged SST control time series for three models that have accelerating, decelerating, and an oscillatory rate of change over the duration of the control. Similar to the problem of separating externally forced trends from internal climate variability either in observations or transient climate simulations, it is not possible to categorically determine what component of the control simulation evolution is drift (and so would be coherent across coincident portions of forced simulations initialized from the control run) and what part is internal variability (and so would not be coherent in coincident forced simulations). In principle, it should be possible to separate drift from internal variability with a perturbed initial state ensemble of control runs (in the same way historical simulation ensembles are used to isolate forced trends from internal variability); however, such ensembles

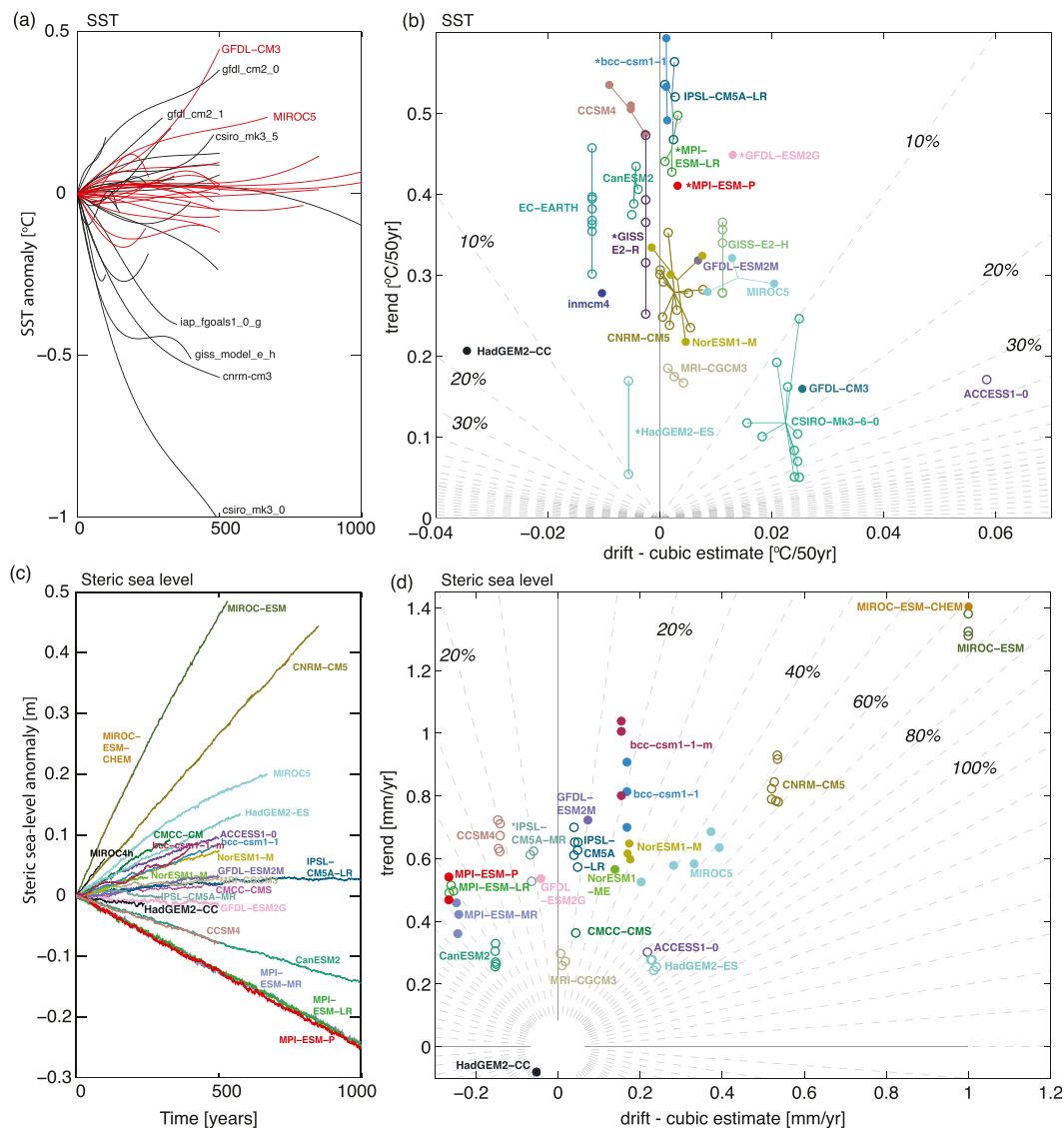


FIG. 1. (a) Cubic polynomial fit to globally averaged SST control time series for CMIP3 (black) and CMIP5 (red), relative to start point; (b) associated CMIP5 1945–2005 historical simulation trend vs associated drift estimate (using cubic drift estimate); (c) annually averaged steric sea level control time series for CMIP5 models, relative to start point; and (d) associated CMIP5 1945–2005 historical simulation trend vs associated drift estimate (using cubic fit). Ensemble members are shown in same color. The asterisk next to a model indicates the models for which the full linear drift estimation is not significant at 90% level (all linear drifts are significant for sea level).

are not available for CMIP5 and are probably impractical to achieve.

A comparison of the control and historical integrations for the Institute of Atmospheric Physics (IAP) model from CMIP3, for example (Fig. 2c), might suggest that the initial downward trend followed by a weaker upward trend evident in the first 150 years of the control simulation may also form part of the transient response in the historical simulation. However, this similarity may simply be coincidental, and the low-frequency changes in the control simulation may just be internal variability

that is not coherent across the historical simulation. Understanding the temporal evolution of the drift is important as the assumed structure can make a large difference when adjusting for drift. Figures 2d–f show the size of the estimated drift correction that would be associated with a concurrent 60-yr period of forced simulation (for different periods along the control time series) based on the 100-yr linear, 150-yr linear, quadratic, cubic, and full control linear drift estimation techniques (see methods). Similar correction techniques to these have been used in previous studies. In general,

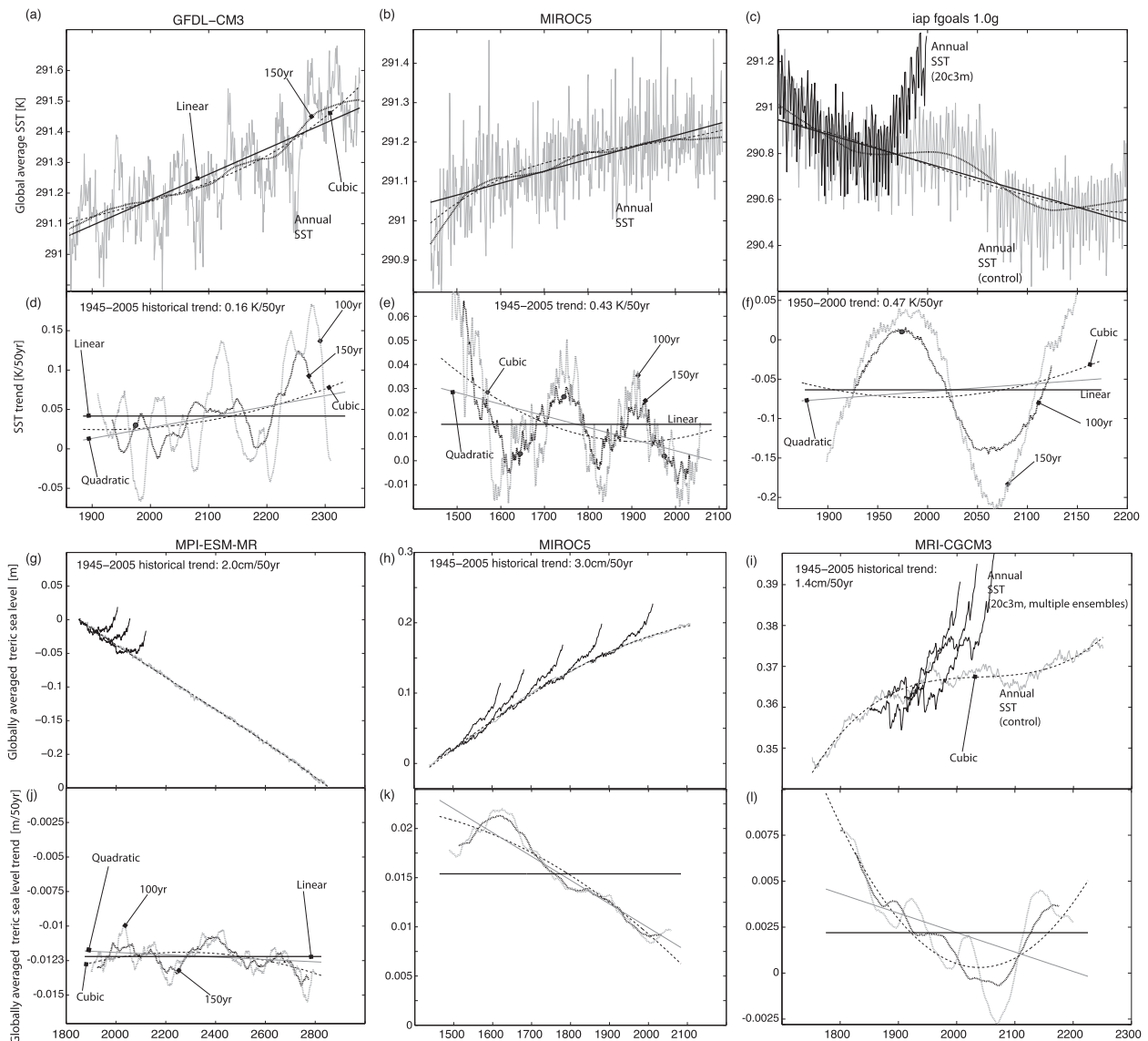


FIG. 2. (a)–(c) Annual averaged SST from preindustrial control for selected models [(c) also shows historical SST]. Superimposed are linear trend, cubic fit, and 150-yr filtered time series. (d)–(f) Moving trends based on linear (black solid), quadratic (gray solid), and cubic (black dashed) fits, and 100- (thick black dashed) and 150-yr (thick gray dashed) moving averages for corresponding models. Dark gray circles indicate year 1975 for historical simulation ensemble members. (g)–(i) Annual averaged steric sea level from control and historical simulations for selected models, cubic fit superimposed. (j)–(l) Moving trends based on linear, quadratic, and cubic fits and 100- and 150-yr moving averages for corresponding models.

the methods listed above exhibit higher-to-lower temporal variability. Based on the GFDL model, for example (Fig. 2d), the full control linear drift estimate would result in an approximate $0.045 \text{ K (50 yr)}^{-1}$ reduction of a historical or projection simulation trend (irrespective of the time period under consideration), whereas the 100-yr linear drift estimate would result in a correction anywhere from an approximate $0.05 \text{ K (50 yr)}^{-1}$ increase to an approximate $0.175 \text{ K (50 yr)}^{-1}$ decrease in the forced simulation trend, depending on the time

period. To put this into context, the 1945–2005 historical simulation trend for this GFDL model is approximately $0.16 \text{ K (50 yr)}^{-1}$. Without knowing the structure of the drift we cannot say which of these drift estimation methods is more valid.

Based on linear trends through the full globally averaged SST control time series, 15 of the 22 models have statistically significant drifts (at a 90% level, where lag-1 autocorrelation has been accounted for). Despite this, the importance of drift is relatively small for most

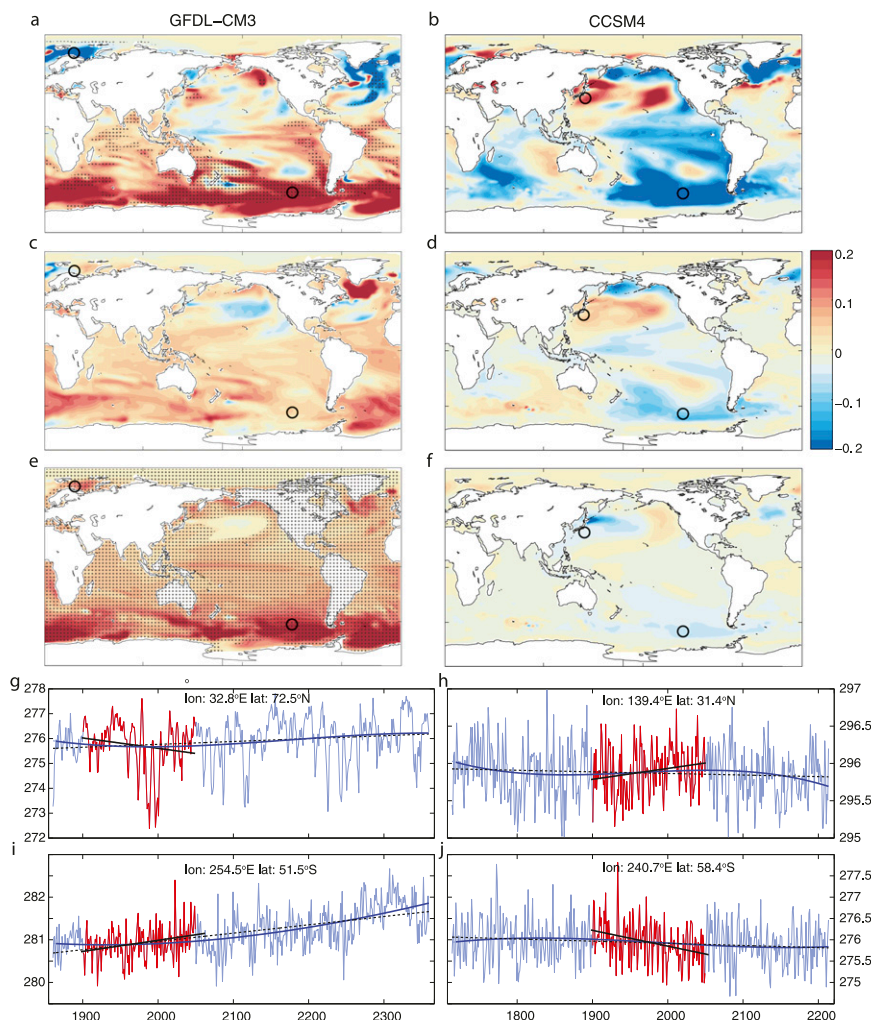


FIG. 3. SST drift in the preindustrial control [$^{\circ}\text{C} (100 \text{ yr})^{-1}$; 1945–2005] estimated using (a),(b) 150-yr linear trend (1900–2050); (c),(d) cubic fit to full control time series; (e),(f) linear trend over full control time series; and (g)–(j) annual time series, cubic fit, linear 150-yr, and full time series fit for selected regions (indicated by circles on maps) for (left) GFDL-CM3 and (right) CCSM4. Mottling indicates regions of significant trend (90% level) using a t test using effective degrees of freedom estimated from the time series lag-1 autocorrelation (following Mitchell et al. 1966).

models. Figure 1b shows the percentage error associated with drift for the 1945–2005 historical simulation trends. Given the very low-frequency variability inherent in the SST control time series (Fig. 1a), we have chosen to use a cubic drift estimate. For most models the estimated drift makes up less than 10% of the historical trend. Moreover, for models with multiple ensemble members the drift is generally considerably smaller than the spread in historical trends across the members. In a few models/ensemble members the drift is more important, although this also relates to the relatively weak historical trends in these models. For example, in the ACCESS1.0 model drift accounts for over 30% of

the historical trend (based on a cubic drift estimate). Similar results are found if a full linear drift estimate were used instead of a cubic estimate (figure not shown), although the details for individual models do change (e.g., drift importance reduces to about 20% for ACCESS1.0, while that for GFDL-CM3 increases to about 25%).

Figure 3 shows spatially resolved SST drift estimates for GFDL-CM3, which has a relatively large globally averaged SST drift and CCSM4 whose globally averaged drift is relatively weak (Fig. 1b). Using the 150-yr linear trend as the drift proxy, we find that both models have regions of strong drift. Using the longer-term (cubic and

full linear) estimates we see that the GFDL-CM3 drift remains large in many regions while it becomes much weaker in the CCSM4, suggesting that even when using a relatively long 150-yr window to calculate drift, large values exist as a result of the aliasing of internal variability. This is clearly seen in Figs. 3d–f, which show gridpoint drift time series at different locations. At some locations the sign of the drift estimate is reversed when using different drift methods. While drift may indeed contain centennial-scale variability, we would suggest that avoiding the often large trends associated with internal variability aliasing is desirable. The full time series metrics are a more conservative choice that is less likely to be contaminated by internal variability not associated with the drift and is probably a safer choice, particularly at local scales where internal variability will be large.

A comparison of CMIP3 and CMIP5 full linear drift estimates suggests a significant reduction in the magnitude of SST drift. The multimodel mean of globally averaged SST drift magnitudes (i.e., the absolute value is taken after globally averaging the local drift) for CMIP5 is $0.02^{\circ}\text{C} (100\text{ yr})^{-1}$ compared to $0.06^{\circ}\text{C} (100\text{ yr})^{-1}$ for CMIP3. Alternatively, by taking the absolute value of the drift at each grid cell prior to globally averaging and taking a multimodel mean, we have a measure of the typical strength of local drift. In this case, the multimodel mean has reduced from $0.09^{\circ}\text{C} (100\text{ yr})^{-1}$ for CMIP3 to $0.02^{\circ}\text{C} (100\text{ yr})^{-1}$ for CMIP5. This improvement is clear from Fig. 4 that shows the area-weighted frequency distribution of SST drift values for CMIP3 and CMIP5. It is also evident that while many of the distributions are centered around zero (i.e., there is no systematic drift direction from one location to another), many models do have an SST drift that is biased toward warming or cooling. Indeed, it is the models that have large areas of consistent drift direction that have the largest globally averaged drift magnitudes. Across different models positive and negative biases tend to cancel out and consequently the effect of drift on the multimodel mean becomes negligible.

While the direction of drift tends to be nonsystematic across different models, there appear to be common regions where the magnitude of drift is larger or smaller.

Figure 5 shows the multimodel mean drift magnitude (i.e., the absolute value of the local drift is taken prior to averaging across models). Previous studies (e.g., Rahmstorf 1995; Cai and Chu 1996; Cai and Gordon 1999) have noted that drift in ocean temperatures are often relatively large in the region of vigorous convective mixing in the Southern Ocean and North Atlantic. These convective regions are sensitive to small changes in the vertical density structure of the ocean. As a result

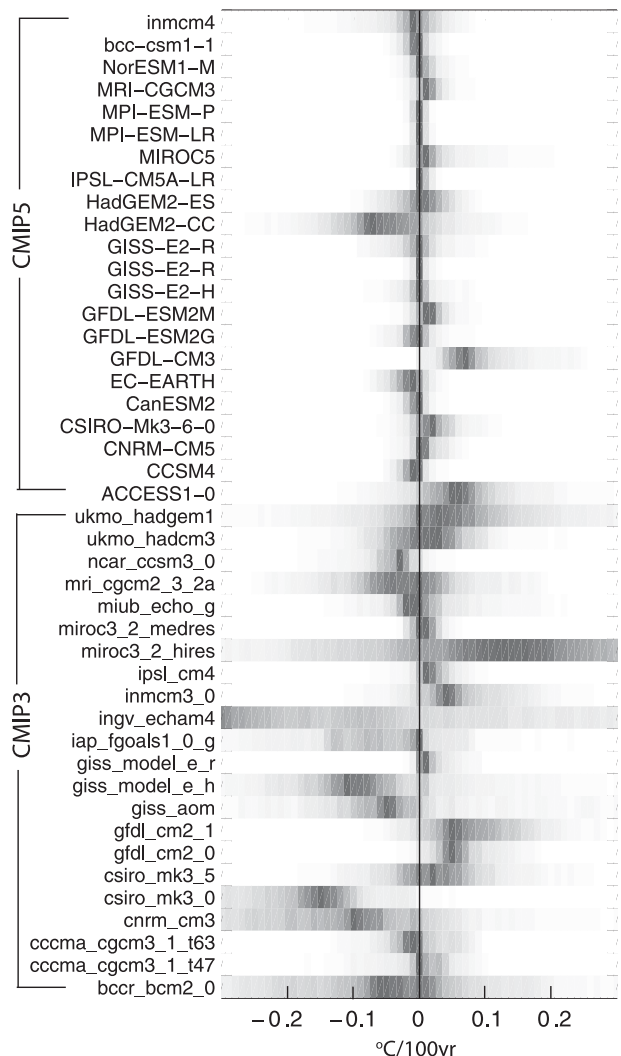


FIG. 4. Area-weighted probability distribution function of local SST drift (based on full linear drift estimate; $^{\circ}\text{C century}^{-1}$; normalized by maximum value) for CMIP5 and CMIP3 preindustrial controls.

drift related changes can become large in these regions. Consistent with the CMIP3 (see SG12), we find that for the CMIP5 models the relative importance of SST drift is smallest in the tropics with the strongest drift magnitudes in the North Atlantic and Pacific and throughout the Southern Ocean. In the Northern Hemisphere, the drift-to-trend ratio (not shown) is largest between 40° and 60°N because of enhanced drift magnitudes. In the Southern Hemisphere, drift is particularly important south of about 50°S because of both the enhanced midlatitude drift magnitudes and relatively weak forced trends. Projections suggest that warming of the Southern Ocean will remain small over the twenty-first century, compared to the global-mean warming, with no significant warming in some regions (Sen Gupta et al.

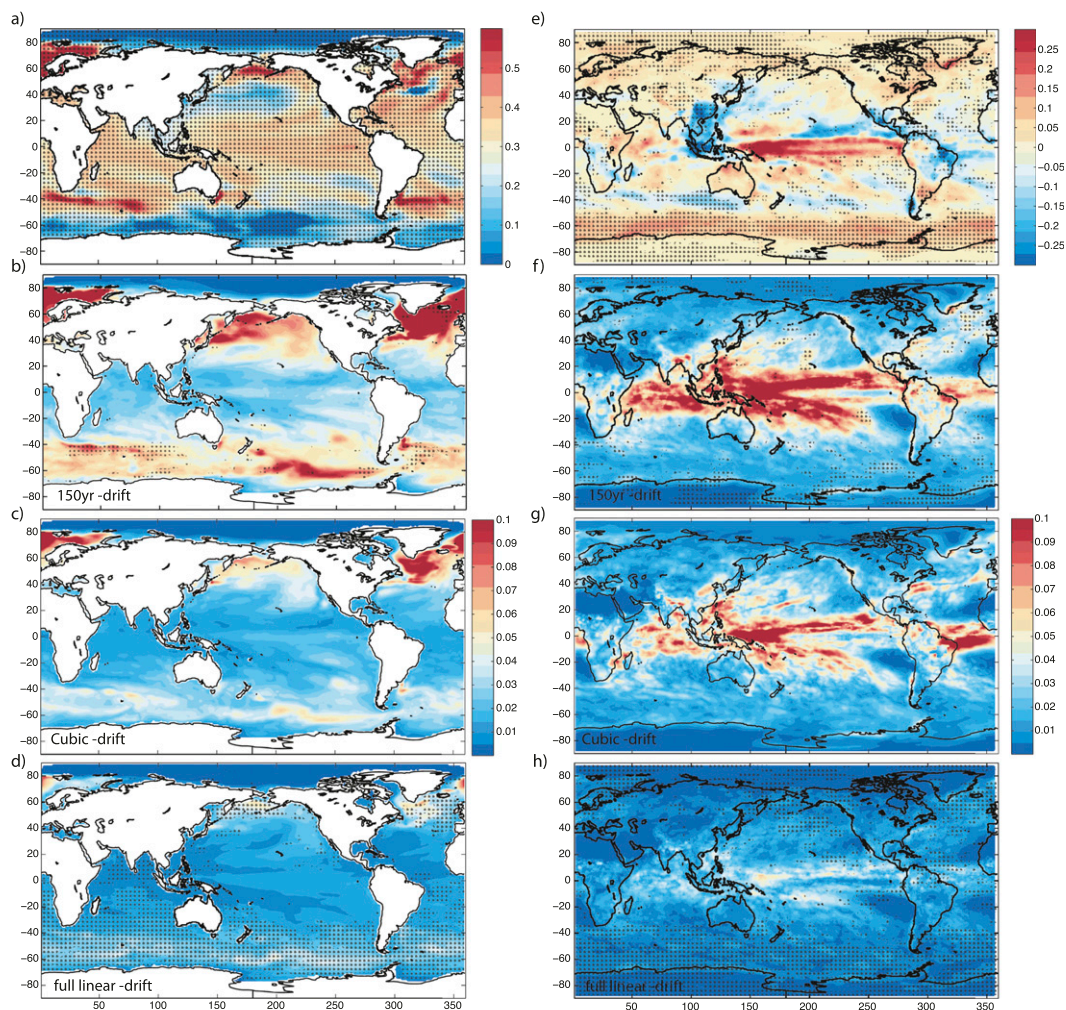


FIG. 5. (a) Multimodel mean 1945–2005 SST trend and corresponding multimodel mean local drift magnitude using (b) 150-yr linear trend, (c) full time series cubic fit, and (d) full time series linear fit. (e)–(h) As in (a)–(d), but for precipitation. Mottling indicates regions where (a),(b),(d) 50% or (e),(f),(h) 33% of the models have significant trends (at 90% level).

2009). Drift is therefore likely to remain an important consideration for this region. The 150-yr CMIP5 drift estimates shown in Fig. 5b can be compared to Fig. 4b of SG12 for CMIP3. Again it is apparent that in most regions the magnitude of drift is considerably reduced in the new generation of models. While local drift magnitudes based on the full linear method (Fig. 5d) are considerably weaker compared to the 150-yr linear method (Fig. 5b), the longer time period used to calculate the trends means that a much larger surface area shows statistically significant drift, particularly in those regions of high drift magnitudes, as described above.

b. Precipitation

A strong relationship exists between full linear drift in globally averaged SST and precipitation across the

models (Fig. 6a), which suggest that atmospheric drift can be largely explained by long-term change in SST. The majority of models have a globally averaged drift magnitude that is less than 10% of the historical trend (Fig. 7). Although the MIROC4h model has the largest drift, it is not statistically significant as the available control is only 100 years in duration. The largest drift associated error (where both drift and trend are significant) is associated with the two MIROC-ESM models, where the drift makes up over 30% of the historical trend.

Figures 5f–h show the multimodel mean local drift magnitudes for precipitation. As with SST, drift estimates are usually considerably reduced using the full control methods and in particular the linear estimate. As noted by SG12 for the CMIP3 models, drift magnitudes

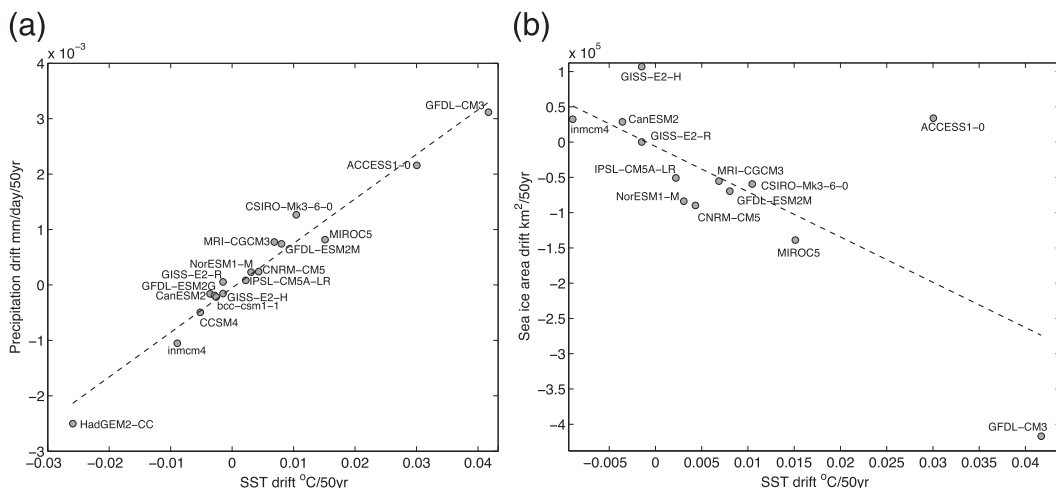


FIG. 6. Globally averaged SST drift vs (a) globally averaged precipitation drift and (b) Antarctic sea ice area drift (all using full linear drift estimates). Dashed line indicates the line of best fit.

tend to be largest at low latitudes, in regions of high precipitation variability. This is also the case for the CMIP5 models. Despite this, the drift estimates at low latitudes are usually not significant. As with SST, statistically significant drift is most likely in the North Atlantic and Pacific Oceans and throughout the Southern Ocean. Moreover, precipitation drift is rarely significant over land areas. This is again consistent with the idea that the precipitation drift is primarily a forced response to drift in SST.

c. Steric sea level

For the CMIP3 models, SG12 showed that subsurface drift in temperature and salinity below about 1–2 km generally dominates over any forced changes at these depths. As a consequence drift may constitute a large component of historical simulation trends in the deep ocean or in depth-integrated properties such as ocean heat content or steric sea level. While data availability

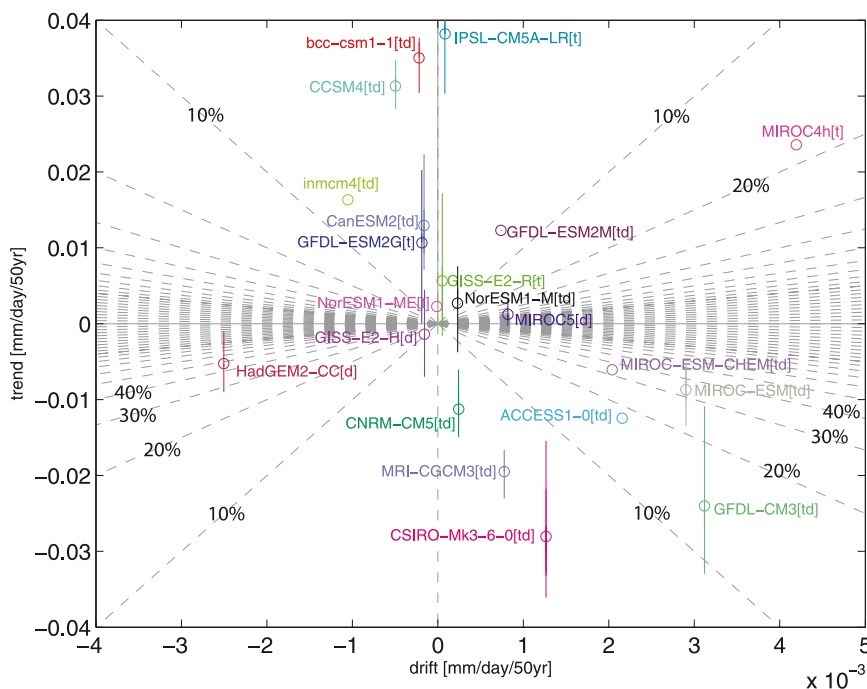


FIG. 7. Full linear drift vs 1945–2005 trend for precipitation. In brackets next to the model names, t denotes a significant multiensemble-mean historical trend, and d denotes a significant full linear drift.

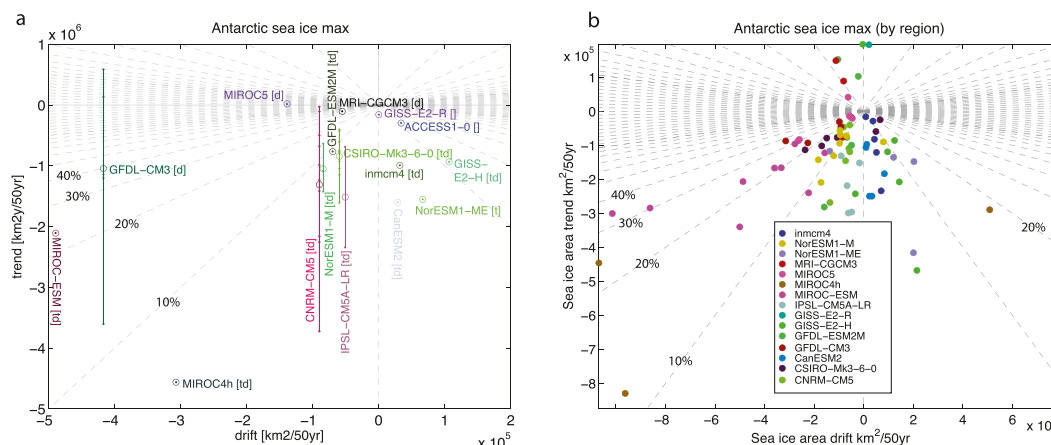


FIG. 8. (a) Full linear drift and 1945–2005 trend for Antarctic sea ice maximum (for the 3-month period of largest sea ice area). Circle indicates multiensemble mean and lines indicate the range across all ensemble members. In brackets next to the model names, t denotes a significant multiensemble-mean trend, and d denotes a significant drift. (b) As in (a), but for 30° longitudinal segments around Antarctica. Region/model points are only shown where both the trend and the drift is statistically significant ($\sim 40\%$ of all regions/model combinations).

constraints have precluded the examination here of three-dimensional temperature and salinity fields for CMIP5, we examine the globally averaged steric sea level that integrates the effect of subsurface temperature and salinity changes.

Figure 1d shows the historical trend and drift information (based on a cubic fit) for steric sea level. All but 7 of the 24 available models (ensemble means) have globally averaged drift that exceeds 20% of the historical trend, and 5 models exceed 60%. We would note that because of the large effect of drift on this variable, it is routinely removed during the calculation of forced trends (e.g., Solomon et al. 2007). The two MIROC-ESM models have the largest overall drift magnitudes that spuriously inflate steric sea level increases by over 1 mm yr^{-1} . Unlike the globally averaged surface metrics described above, the error introduced by drift is for most models larger than the spread in historical trends from different ensemble members. Based on full linear drift estimates, all models except for the IPSL-CM5A-MR have statistically significant drift.

As with the other metrics examined above, there appears to be a substantial overall reduction in the size of drift in steric sea level in CMIP5 as compared to CMIP3. Based on a linear trend over the full preindustrial control period, the mean or median magnitude of the globally averaged drift is almost twice as large for CMIP3 compared to CMIP5, with the multimodel mean drifts significantly different at the 90% level. The median drift magnitude for CMIP5 is approximately $2.5 \text{ cm century}^{-1}$ and for CMIP3 is approximately $4 \text{ cm century}^{-1}$, although the maximum drift magnitude is about $10 \text{ cm century}^{-1}$ for both CMIP3 and CMIP5 models.

The temporal evolution of globally averaged steric sea level is much more linear than for SST, and the variability about the long-term trend is much smaller (Figs. 1c and 2g,h,i). As such, for most models different drift estimation methods result in quite similar drift estimates (e.g., Fig. 2j). This is not always true however. For the MIROC5 model, for example (Figs. 2h,k), drift becomes weaker over the course of the 700-yr control simulation. Consequently, a correction based on a linear fit to the full control time series would not be appropriate, while a quadratic or cubic drift correction would probably be more suitable. For the MRI-CGCM3 model (Figs. 2i,l), it is not obvious what the most appropriate correction method would be (i.e., what is drift and what is internal variability). Even cubic and quadratic methods would result in quite large drift estimate differences over some periods. As steric sea level rise is approximately proportional to the global ocean heat content, the quasi-linear drift in most models could also come about through a nearly constant energy leak. Indeed, significant spurious nontransient energy imbalances have been identified in the various model components for the CMIP3 models (Lucarini and Ragone 2011).

d. Sea ice

Previous studies (e.g., Cai and Gordon 1999) have noted significant drift in sea ice. Here, we examine the Antarctic sea ice area, where historical trends in models and observations are weaker than in the Arctic and as such drift may be a more important source of error. Historical trends for sea ice area around Antarctica for 15 models are presented in Fig. 8a. Most model/ensembles indicate a large-scale reduction in sea ice area

over the 1945–2005 period, consistent with results from CMIP3 (Arzel et al. 2006). Only two ensemble members of the GFDL-CM3 model have nonsignificant increases in area, related to the large multidecadal variability inherent in the different ensemble members. Unlike the models, observations of Antarctic sea ice indicate a small increase over the satellite era (Comiso and Nishio 2008; Parkinson and Cavalieri 2012). Given the limited length of the observational record (~ 30 yr) and the large natural variability inherent across models with multiple ensemble members (even for the longer 60-yr window used in our analysis), it seems plausible that the observed sea ice reduction may be a consequence of natural variability rather than a forced change.

For many models, drift (calculated as a linear trend over the full control) makes up less than 10% of the historical trends. This ratio is high ($> \sim 40\%$) for the GFDL-CM3 and MRI-CGCM3 models; however, while the drift is statistically significant, the trend over the 60-yr historical period is not. Only the MIROC-ESM model has a drift-to-trend ratio that exceeds 20% when both trend and drift are statistically significant. For the models for which there are multiple historical ensemble members, the drift magnitudes are again smaller than the spread in historical trends across the ensemble members. For example, the GFDL-CM3 has a relatively large linear drift [about $-4 \times 10^5 \text{ km}^2 (50 \text{ yr})^{-1}$]; however, the trends in the various historical ensemble members range over approximately $40 \times 10^5 \text{ km}^2 (50 \text{ yr})^{-1}$ [from about -3.6×10^6 to about $7 \times 10^5 \text{ km}^2 (50 \text{ yr})^{-1}$]. That is, the differences in trend related to internal decadal variability are larger than the error introduced by drift (at least for the subset of models examined here). As with globally averaged SST, some models show high levels of low-frequency variability in the control simulations. These models are particularly sensitive to the drift estimation method. For example, the ensemble-mean GFDL-CM3 cubic drift estimate for the 1945–2005 period is almost half that calculated using a full linear drift, while for MIROC-ESM, which has weak low-frequency variability, the drift estimates are very similar.

The trends around Antarctica show regions of both decreasing and increasing ice area. Figure 8b examines trends and associated drifts for twelve 30° longitudinal segments around the poles (based on ensemble means where available). Only regions for which both the historical trend and the drift are statistically significant are shown: approximately 40% of regional segments across the models. Of these, only four models have drift-to-trend ratios that exceed 30% for a limited number of regions (between 1 and 3 out of 12) around Antarctica. In summary, while there is statistically significant

drift in many regions, it only makes up a relatively small component of the forced trend in those regions where the historical trends are discernible.

e. Biogeochemistry

A number of ESMs that incorporate interactive ocean biogeochemistry have been included in CMIP5 (Taylor et al. 2012). The ocean biogeochemistry also has its own long time-scale processes. At the time of retrieval, only a small subset of models have archived biogeochemical properties. To illustrate the possible importance of drift, we examine two biogeochemical variables: depth-integrated dissolved inorganic carbon (IntDIC) and dissolved oxygen (O_2).

As anthropogenic CO_2 levels increase in the atmosphere, a large proportion of that CO_2 is absorbed by the ocean (Sabine et al. 1999) forming various species of dissolved inorganic carbon (DIC). DIC is transported into the ocean interior via both physical processes (related to vertical mixing, subduction, and advection by ocean circulation) and biological processes (export production in the upper ocean and subsequent remineralization of sinking organic material in the deeper ocean). As the ocean plays such a large role in the global carbon cycle, small changes in the uptake of CO_2 by the ocean could substantially affect the rate of future warming. Although still a matter of contention, there is some evidence that the rate of uptake by the ocean may have decreased over recent years (Le Quere et al. 2007, 2008; Law et al. 2008). To investigate these matters using climate models, it will be necessary to not only identify trends but also changes in trends over time.

Figure 9 shows globally averaged trend and drift for depth-integrated DIC. For all six models shown (Fig. 9d), both the globally averaged historical trends and the drift are highly significant. Despite this, the drift usually makes up less than 20% of the historical trend magnitude. The largest globally averaged drift-to-trend ratio is for the IPSL-CM5A-LR model at about 23%. It is interesting to note that the latter model does not have significant global drift for any of the physical properties discussed above, which suggests that the biogeochemical drift can occur independently of changes in the physical environment and that the equilibrium time scales of the physical and biogeochemical model components can be very different. Figures 9a–c show the trend, drift, and spatially resolved drift-to-trend ratio for this model for illustrative purposes (the associated drift and drift-to-trend ratio are different for other models). The trend in integrated DIC shows large uptake in well-ventilated regions of the North Atlantic and Southern Ocean. This is similar to the pattern of observed accumulation of anthropogenic CO_2 [e.g., Sabine et al. (2004)]; although

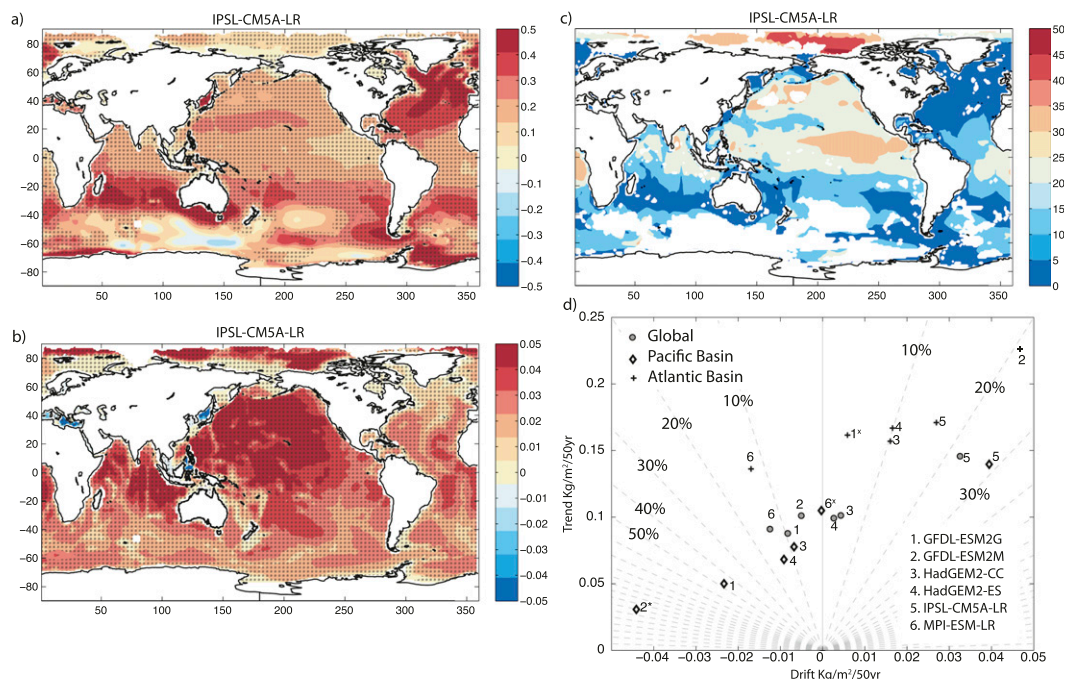


FIG. 9. (a) The 1945–2005 trend in IntDIC, (b) associated drift calculated using linear trend over full control period [mottling in (a) and (b) signifies trends are significant at 90% level], (c) drift-to-trend ratio (%) in regions where the historical trend is significant, and (d) scatterplot of global average (circles), Pacific basin average (diamonds), and Atlantic basin (crosses) drift vs trend for six models.

this does not include the effect of possible changes in natural DIC]; although uptake rates in the North Atlantic in particular appear to be underestimated in the model. It is clear that regionally the importance of drift can be larger than for the global average. This is most evident where the historical trends are relatively weak. In these regions, drift can make up a substantial fraction of the historical trend. For two of the six models, the Pacific basin-averaged drift makes up about 50% or more of the forced trend (Fig. 9d; although in the case of GFDL-ESM2M the historical trend is not significant).

There has been considerable interest in recent changes to subsurface oxygen concentrations and the spatial extent of low oxygen regions in the tropical oceans (Stramma et al. 2008; Keeling et al. 2010; Helm et al. 2011), and we briefly examine this metric here for two models. Figures 10a and 10f show the evolution of globally averaged dissolved oxygen by depth, relative to the first year of the historical simulation in two models. For both these models there is a strong reduction in subsurface oxygen concentration over the course of the historical period. Examination of the control simulation evolution, where both linear (Figs. 10b,g) and cubic (Figs. 10d,i) smoothing has been applied, clearly shows that the subsurface historical simulation changes are spurious and are not driven by changes in external

forcing. Indeed, historical changes in other models (not shown) do not in general exhibit a similar subsurface reduction. Figures 10c and 10h and Figs. 10e and 10j show the drift-corrected evolution based on full linear and cubic methods, respectively. Correcting for drift in both models substantively changes the signal. While the HadGEM2 model is relatively insensitive to the drift correction method, the IPSL model shows very different corrected forced responses depending on the correction method.

4. Discussion

Model drift is still a problem for many climate models when computational resources are limited. Climate models may take thousands of model years to reach equilibrium, beyond practical integration times at many institutes, or they may have chronic problems related to energy or moisture not being fully conserved (e.g., Liepert and Previdi 2012; Lucarini and Ragone 2011). Liepert and Previdi (2012), for example, found that imbalances in the moisture balance in some CMIP3 models caused spurious latent heating or cooling of the atmosphere leading to net energy imbalances with magnitudes of a few watts per square meter that changes over time. Drift exists not only during spinup simulations but

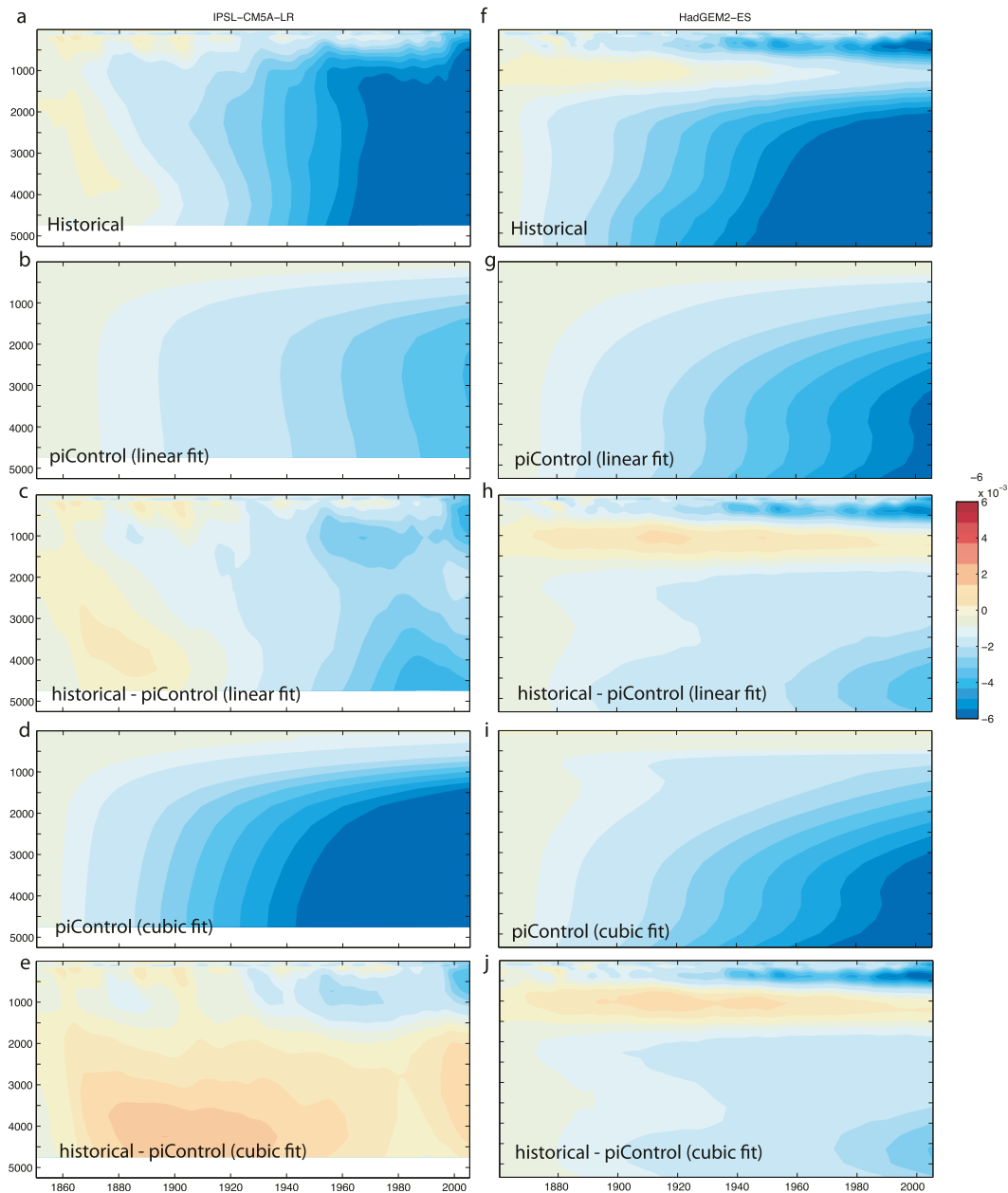


FIG. 10. Evolution of globally averaged O_2 at different depth levels over time for the IPSL-CM5A-LR model (a) historical simulation, (b) control simulation (with linear filtering over full length of the control simulation), (c) drift-corrected concentrations [see (a),(b)], (d) control simulation (with cubic filtering over full length of the control simulation), and (e) drift-corrected concentrations [see (a)–(d)]. (f)–(j) As in (a)–(e), but for the HadGEM2-ES model.

persists in forced climate change simulations. When trying to isolate forced trends, it is important to know how large drift is to determine when drift can be safely ignored and when it must be corrected for. Here, we have examined the drift in a variety of variables using both global and local metrics. We have also examined the differences in selected drift metrics in moving from CMIP3 to CMIP5. By examining the historical period,

we are looking at a worst-case scenario for the importance of drift. The importance of drift relative to projected trends is likely to be considerably smaller as the external forcing is much larger for representative concentration pathway (RCP) simulations compared to historical simulations. As a result, drift should be carefully considered in detection and attribution studies of past climate change, for example.

Based on SST, precipitation, and steric sea level, there is a clear overall reduction in the magnitude of drift in the newer generation of models. In particular, for globally averaged steric sea level, which integrates the drift throughout the ocean, the average size of drift in CMIP5 is about half of that in CMIP3, although there are still a few outlier models with high drift magnitudes. For some models this improvement may in part relate to longer spinup periods. For example all four of the GFDL climate models were spun up in a coupled configuration for around 2000 yr (S. Griffies 2012, personal communication) compared to around 300 yr for the CMIP3 GFDL models. Similarly the new generation of CCCma models were spun up for around 800 yr (O. Saenko 2012, personal communication) compared to <250 yr for the previous generation models. Unfortunately, information regarding model spinup is often not easily available for many of the models (see Table 1). As such making a quantitative comparison of spinup times between CMIP3 and CMIP5 is not possible. For future assessment it would be useful for model developers to provide more detailed spinup information as part of their model descriptions, preferably collated at a central repository [e.g., at the Program for Climate Model Diagnosis and Intercomparison (PCMDI)] as was done for CMIP3. The reduced drift may also result from improved representation of physical parameterizations (e.g., cloud microphysics) and numerical schemes (e.g., advection or diffusion) and/or higher horizontal and vertical resolution (see references in Table 1 for individual model developments). Knutti and Sedlacek (2013) suggest that greater computational resources have been expended on more complete representation of physical and chemical processes than on spatial resolution. However, despite the inclusion of more climate processes underpinned by improved physical understanding, the spread in the mean climate state across the CMIP5 simulations have not reduced and many of the biases inherent in CMIP3 persist.

For the surface properties examined here, the direction of drift is often spatially coherent over large regions and across ensemble members of a given model. Consequently, globally averaged drift will often be statistically significant. However, while significant the importance of drift tends to be relatively small. In particular, the drift in the globally averaged surface properties considered here was in most instances smaller than the differences in historical simulation trends calculated from different ensemble members of a particular model. That is, errors in the calculation of forced trends resulting from the aliasing of natural variability (at least for the 60-yr time period assessed here) are greater than the errors introduced by drift and as a result the use of multiple

ensemble members to calculate trends provides greater benefit than applying a drift correction. The direction of drift across different models does not appear systematic across models for the variables considered. Therefore, the problem of drift is often negligible when considering multimodel means. For example, despite the large positive drifts in steric sea level (particularly for the MIROC models), the multimodel mean drift is still not statistically different from zero. By averaging across multiple models, drift (like internal variability) is substantially reduced. For SST the multimodel mean historical and drift-corrected trends are practically identical.

There is a clear relationship between SST drift and both precipitation and Antarctic sea ice area drift. It appears that slow changes in the ocean manifest as drift in the atmosphere and cryosphere. Moreover, the regions where drift is significant tend to be coincident between SST and precipitation, particularly in the North Atlantic and Pacific and over the Southern Ocean. For most models drift contributes less than 10% to historical precipitations trends, although it can be of greater importance in some models and regions. Similarly, the error introduced by drift when examining sea ice area trends is smaller than the uncertainty associated with natural variability.

The problem of drift is most pronounced when considering applications associated with the deep ocean including depth-integrated properties. As pointed out by SG12, drift below 1–2 km in the ocean usually dominates over any forced signal for historical time periods. Here, we examined drift in steric sea level. The relatively small drifts associated with CMIP5 compared to CMIP3 meant that unlike CMIP3 drift was in no cases large enough to change the direction of the historically simulated global sea level rise. However, at a global scale the drift in most models still exceeded 30% of the historical trend, with a number of models exceeding 60%.

Given the recent inclusion of ocean biogeochemical processes into a number of the CMIP5 models, we also examined depth-integrated dissolved inorganic carbon and dissolved oxygen. In the case of DIC, historical trends are large and positive across most of the upper ocean in all models, while drift varies regionally in sign and magnitude. As a consequence, the drift only becomes important in certain regions and models. In the case of dissolved oxygen however, which has a less systematic response to changes in anthropogenic forcing, drift clearly dominates the historical simulation response in the subsurface ocean.

Particularly for properties where drift makes up a large part of the historical change (e.g., steric sea level rise or dissolved oxygen), the assumed temporal evolution of drift can be important. We have demonstrated that

the method of drift correction can have a substantial effect on the outcome. Using a short time period over which to compute drift may mean that estimates still contain a substantial contamination from the internal variability component. We have shown that this is the case even when using 150-yr control simulation linear trends as the proxy for drift. As a result, a more conservative estimate of drift calculated using the full available control might be less prone to contamination from low-frequency variability. However, even when using the full control, the drift estimate can be sensitive to the correction method (i.e., linear, quadratic, or cubic), as we demonstrated for dissolved oxygen.

Based on our results we would offer some recommendations for the treatment of drift.

- Our results suggest that when estimating forced changes, the use of multiple ensemble members from a historical simulation to minimize the influence of natural variability will often have a larger effect than drift correction. Drift is both model and region dependent. As such, the importance of drift should be assessed based on the application. In many circumstances, such as when considering multimodel means of surface properties, drift is negligible. Even where a clear drift signal can be identified and is statistically significant, its importance may be small.
- The temporal evolution of drift appears to be model and variable dependent. For example globally averaged steric sea level evolution is often quite linear and possesses much less low-frequency variability than SST evolution. Where accurate determination of trends is required—for example, when studying changes in trends over time or in the examination of detection and attribution—the sensitivity of the drift estimate method should be tested. Indeed we have shown that the method of drift correction can substantively change the resulting estimate of a forced trend in some cases. While drift would ideally be assessed on a model-by-model and variable-by-variable basis, in general the use of long portions of the control simulation (if not the full control time series) is recommended to guard against contamination from internal variability. We would suggest that drift estimates based on relatively short periods of control simulation (e.g., 100-yr trends have been commonly used) are probably insufficient to get a robust estimate of drift.
- The structure of drift could in principle be identified using ensembles of perturbed control simulations. As greater precision in trend estimates become necessary such simulations should be considered.
- Careful treatment of drift needs to be considered for more than just physical ocean variables. For some

models and in certain regions, drift in atmospheric properties can be important. This has become important as climate models are increasingly being used for the regional assessment of climate change or regional detection and attribution. Indeed, some regional assessments have started using drift size as one of the metrics of model skill (e.g., Irving et al. 2012; Brown et al. 2012). With respect to detection and attribution, drift is important both in the isolation of the climate change signal and in estimating the magnitude of internal variability [which is often estimated from the control simulation; e.g., Santer et al. (2012)]. Biogeochemical properties within the ocean may also be strongly affected by drift, particularly in the subsurface ocean where forced signals are weak.

- A common method for assessing models is to compare their mean state to observations (usually over the end of the twentieth century). Given that many models are initialized from some observed state and will drift away from this state, care must be taken when interpreting model fidelity where drift is present. A model with a short spinup would have greater fidelity than one with a long spinup, given the same rate of drift. The rate of drift may in itself be a useful metric to gauge model fidelity.

Acknowledgments. 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 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. This research was conducted with the support of the PACCSAP, a program funded by AusAID, in collaboration with the Department of Climate Change and Energy Efficiency and delivered by the Bureau of Meteorology and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). This work was supported by the NCI National Facility at the ANU via the provision of computing resources to the ARC Centre of Excellence for Climate System Science.

APPENDIX

Model Expansions

ACCESS1.0	Australian Community Climate and Earth-System Simulator, version 1.0
BCC-CSM1.1	Beijing Climate Center (BCC), Climate System Model, version 1.1

BCC- CSM1.1(m)	BCC Climate System Model, version 1.1 (moderate resolution)	MPI-ESM-P	Max Planck Institute (MPI) Earth System Model, paleo
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCma) Second Generation Earth System Model	MPI-ESM-LR	MPI Earth System Model, low resolution
		MPI-ESM-MR	MPI Earth System Model, medium resolution
CCSM4	Community Climate System Model, version 4	MRI-CGCM3	Meteorological Research Institute (MRI) Coupled Atmosphere–Ocean General Circulation Model, version 3
CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti Climatici (CMCC) Climate Model	NorESM1-M	Norwegian Climate Centre (NCC) Earth System Model, version 1 (intermediate resolution)
CMCC-CMS	CMCC Climate Model with resolved stratosphere	NorESM1-ME	NCC Earth System Model, version 1 (intermediate resolution), with prognostic biogeochemical cycling
CNRM-CM3	Centre National de Recherches Météorologiques (CNRM) Coupled Global Climate Model, version 3	REFERENCES	
CNRM-CM5	CNRM Coupled Global Climate Model, version 5		
CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation (CSIRO) Mark, version 3.6.0	REFERENCES	
EC-EARTH	European Consortium Earth system model		
GFDL-CM3	Geophysical Fluid Dynamics Laboratory (GFDL) Climate Model, version 3	REFERENCES	
GFDL- ESM2G	GFDL Earth System Model with Generalized Ocean Layer Dynamics (GOLD) component (ESM2G)		
GFDL- ESM2M	GFDL Earth System Model with Modular Ocean Model 4 (MOM4) component (ESM2M)	REFERENCES	
GISS-E2-H	Goddard Institute for Space Studies (GISS) Model E2, coupled with the HYCOM ocean model		
GISS-E2-R	GISS Model E2, coupled with the Russell ocean model	REFERENCES	
HadGEM2- CC	Hadley Centre Global Environment Model, version 2, Carbon Cycle		
HadGEM2- ES	Hadley Centre Global Environment Model, version 2, Earth System	REFERENCES	
INM-CM4.0	Institute of Numerical Mathematics Coupled Model, version 4.0		
IPSL-CM5A- LR	L'Institut Pierre-Simon Laplace (IPSL) Coupled Model, version 5, coupled with NEMO, low resolution	REFERENCES	
IPSL-CM5A- MR	IPSL Coupled Model, version 5, coupled with NEMO, medium resolution		
MIROC-ESM	Model for Interdisciplinary Research on Climate (MIROC) Earth System Model	REFERENCES	
MIROC-ESM- CHEM	MIROC Earth System Model, chemistry coupled		
MIROC4h	MIROC, version 4 (high resolution)	REFERENCES	
MIROC5	MIROC, version 5		

- 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, doi:10.1073/pnas.0605064103.
- Arzel, O., T. Fichefet, and H. Goosse, 2006: Sea ice evolution over the 20th and 21st centuries as simulated by current AOGCMs. *Ocean Modell.*, **12**, 401–415, doi:10.1016/j.ocemod.2005.08.002.
- Brown, J. N., and Coauthors, 2012: Implications of CMIP3 model biases and uncertainties for climate projections in the western tropical Pacific. *Climatic Change*, **113**, 147–161, doi:10.1007/s10584-012-0603-5.
- Bureau of Meteorology, 2010: Operational implementation of the ACCESS numerical weather prediction systems. NMOC Operations Bull. 83. [Available online at <http://www.bom.gov.au/nwp/doc/bulletins/apob83.pdf>.]
- 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.
- Chylek, P., J. Li, M. K. Dubey, M. Wang, and G. Lesins, 2011: Observed and model simulated 20th century Arctic temperature variability: Canadian Earth System Model CanESM2. *Atmos. Chem. Phys. Discuss.*, **11**, 22 893–22 907, doi:10.5194/acpd-11-22893-2011.
- Cohn, T. A., and H. F. Lins, 2005: Nature's style: Naturally trendy. *Geophys. Res. Lett.*, **32**, L23402, doi:10.1029/2005GL024476.
- Collins, W. J., and Coauthors, 2011: Development and evaluation of an Earth-system model—HadGEM2. *Geosci. Model Dev. Discuss.*, **4**, 997–1062.
- Comiso, J. C., and F. Nishio, 2008: Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data. *J. Geophys. Res.*, **113**, C02S07, doi:10.1029/2007JC004257.
- 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.
- Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484–3519.

- 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.
- Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, doi:10.1007/s00382-012-1636-1.
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991.
- 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., and Coauthors, 2001: Comparison of results from several AOGCMs for global and regional sea-level change 1900–2100. *Climate Dyn.*, **18**, 225–240, doi:10.1007/s003820100180.
- , 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.
- Hazeleger, W., and Coauthors, 2010: EC-Earth: A seamless Earth-system prediction approach in action. *Bull. Amer. Meteor. Soc.*, **91**, 1357–1363.
- Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J. Climate*, **9**, 2281–2306.
- Helm, K. P., N. L. Bindoff, and J. A. Church, 2011: Observed decreases in oxygen content of the global ocean. *Geophys. Res. Lett.*, **38**, L23602, doi:10.1029/2011GL049513.
- Irving, D. B., and Coauthors, 2012: Evaluating global climate models for the Pacific island region. *Climate Res.*, **49**, 169–187.
- Keeling, R. F., A. Körtzinger, and N. Gruber, 2010: Ocean deoxygenation in a warming world. *Annu. Rev. Mar. Sci.*, **2**, 199–229, doi:10.1146/annurev.marine.010908.163855.
- Knutti, R., and J. Sedlacek, 2013: Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Climate Change*, **3**, 369–373, doi:10.1038/nclimate1716.
- Law, R. M., R. J. Matear, and R. J. Francey, 2008: Comment on “Saturation of the Southern Ocean CO₂ sink due to recent climate change.” *Science*, **319**, 570, doi:10.1126/science.1149077.
- Le Quere, C., and Coauthors, 2007: Saturation of the Southern Ocean CO₂ sink due to recent climate change. *Science*, **316**, 1735–1738, doi:10.1126/science.1136188.
- , and Coauthors, 2008: Response to comments on “Saturation of the Southern Ocean CO₂ sink due to recent climate change.” *Science*, **319**, 570, doi:10.1126/science.1147315.
- 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.
- Lucarini, V., and F. Ragone, 2011: Energetics of climate models: Net energy balance and meridional enthalpy transport. *Rev. Geophys.*, **49**, RG1001, doi:10.1029/2009RG000323.
- 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., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007: The WCRP CMIP3 multimodel dataset. *Bull. Amer. Meteor. Soc.*, **88**, 1383–1394.
- Mitchell, J. M., B. Dzerdzhevskii, H. Flohn, W. Hofmeyr, H. Lamb, H. N. Rao, and C. C. Wallén, 1966: Climatic change. WMO Tech. Note 79, WMO 195, 81 pp.
- Parkinson, C. L., and D. J. Cavalieri, 2012: Antarctic sea ice variability and trends, 1979–2010. *Cryosphere*, **6**, 871–880, doi:10.5194/tc-6-871-2012.
- Peacock, S., and M. Maltrud, 2006: Transit-time distributions in a global ocean model. *J. Phys. Oceanogr.*, **36**, 474–495.
- Raddatz, T., and Coauthors, 2007: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century? *Climate Dyn.*, **29**, 565–574, doi:10.1007/s00382-007-0247-8.
- Rahmstorf, S., 1995: Climate drift in an ocean model coupled to a simple, perfectly matched atmosphere. *Climate Dyn.*, **11**, 447–458.
- Rotstayn, L. D., M. A. Collier, M. R. Dix, Y. Feng, H. B. Gordon, S. P. O’Farrell, I. N. Smith, and J. Syktus, 2010: Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment. *Int. J. Climatol.*, **30**, 1067–1088, doi:10.1002/joc.1952.
- , S. J. Jeffrey, M. A. Collier, S. M. Dravitzki, A. C. Hirst, J. I. Syktus, and K. K. Wong, 2012: Aerosol-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmos. Chem. Phys. Discuss.*, **12**, 5107–5188, doi:10.5194/acpd-12-5107-2012.
- Sabine, C. L., R. M. Key, K. M. Johnson, F. J. Millero, A. Poisson, J. L. Sarmiento, D. W. R. Wallace, and C. D. Winn, 1999: Anthropogenic CO₂ inventory of the Indian Ocean. *Global Biogeochem. Cycles*, **13**, 179–198.
- , and Coauthors, 2004: The oceanic sink for anthropogenic CO₂. *Science*, **305**, 367–371.
- Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.*, **105**, 7337–7356, doi:10.1029/1999JD901105.
- , and Coauthors, 2012: Identifying human influences on atmospheric temperature. *Proc. Natl. Acad. Sci. USA*, **110**, 26–33, doi:10.1073/pnas.1210514109.
- Schmidt, G. A., and Coauthors, 2006: Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *J. Climate*, **19**, 153–192.
- Sen Gupta, A., A. Santoso, A. S. Taschetto, C. C. Ummerhofer, 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.
- , L. C. Muir, J. N. Brown, S. J. Phipps, P. J. Durack, D. Monselesan, and S. E. Wijffels, 2012: Climate drift in the CMIP3 models. *J. Climate*, **25**, 4621–4640.
- Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. L. Miller Jr., Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- Stramma, L., G. C. Johnson, J. Sprintall, and V. Mohrholz, 2008: Expanding oxygen-minimum zones in the tropical oceans. *Science*, **320**, 655–658.
- 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.
- 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.
- Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the

- atmospheric and oceanic general circulations. *Izv. Atmos. Ocean. Phys.*, **46**, 414–431, doi:10.1134/S000143381004002X.
- Wang, Z., 2013: On the response of Southern Hemisphere subpolar gyres to climate change in coupled climate models. *J. Geophys. Res. Oceans*, **118**, 1070–1086, doi:10.1002/jgrc.20111.
- Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335.
- 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.
- Yukimoto, S., and Coauthors, 2011: Meteorological Research Institute-Earth System Model version 1 (MRI-ESM1). MRI Tech. Rep. 64, 71 pp. [Available online at http://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_64/index_en.html.]