Influence of Ocean and Atmosphere Components on Simulated Climate Sensitivities

MICHAEL WINTON, ALISTAIR ADCROFT, STEPHEN M. GRIFFIES, ROBERT W. HALLBERG,
LARRY W. HOROWITZ, AND RONALD J. STOUFFER

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 28 February 2012, in final form 8 June 2012)

ABSTRACT

The influence of alternative ocean and atmosphere subcomponents on climate model simulation of transient sensitivities is examined by comparing three GFDL climate models used for phase 5 of the Coupled Model Intercomparison Project (CMIP5). The base model ESM2M is closely related to GFDL's CMIP3 climate model version 2.1 (CM2.1), and makes use of a depth coordinate ocean component. The second model, ESM2G, is identical to ESM2M but makes use of an isopycnal coordinate ocean model. The authors compare the impact of this "ocean swap" with an "atmosphere swap" that produces the GFDL Climate Model version 3 (CM3) by replacing the AM2 atmospheric component with AM3 while retaining a depth coordinate ocean model. The atmosphere swap is found to have much larger influence on sensitivities of global surface temperature and Northern Hemisphere sea ice cover. The atmosphere swap also introduces a multidecadal response time scale through its indirect influence on heat uptake. Despite significant differences in their interior ocean mean states, the ESM2M and ESM2G simulations of these metrics of climate change are very similar, except for an enhanced high-latitude salinity response accompanied by temporarily advancing sea ice in ESM2G. In the ESM2G historical simulation this behavior results in the establishment of a strong halocline in the subpolar North Atlantic during the early twentieth century and an associated cooling, which are counter to observations in that region. The Atlantic meridional overturning declines comparably in all three models.

1. Introduction

Differences in the simulation of climate sensitivity are an important contributing factor to uncertainty in climate model projections, with the other major factors being uncertainty in the simulation of radiative forcing and in the emissions scenario itself. There is a large variation in century-scale climate projections. For example, the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4; Solomon et al. 2007) cites twenty-first-century warming projections that vary by more than a factor of 2 for a given socioeconomic scenario. Even the responses of atmosphere–ocean global circulation models (AOGCMs) forced with specified greenhouse gas concentrations and aerosol emissions vary by about a factor of 2 for the middle Special Report on Emissions Scenarios (SRES) A1B

scenario. Aerosol effects account for some of this variation but benchmark transient and equilibrium global temperature responses to doubled CO_2 also vary by a factor of 2. Other metrics also have large variation in projections. The AR4 reports twenty-first-century Atlantic meridional overturning circulation (AMOC) declines of 0% to more than 50%. Zhang and Walsh (2006) cite twenty-first-century Northern Hemisphere annual sea ice extent trends for the phase 3 of the Coupled Model Intercomparison Project (CMIP3) climate models under the SRES A1B scenario that range over more than a factor of 4. Even when normalized by the global temperature increase, the decline of Northern Hemisphere ice extent varies by more than a factor of 2 (Eisenman et al. 2011; Winton 2011).

Sources of uncertainty in sensitivity have commonly been diagnosed by evaluating the individual radiative feedbacks that sum to the total feedback—the inverse sensitivity (e.g., Soden and Held 2006). However, radiative feedbacks are emergent properties of the simulation, so this method does not identify specific sources of the differences within the models. For example, Winton

Corresponding author address: Michael Winton, NOAA/GFDL, Princeton University Forrestal Campus, 201 Forrestal Rd., Princeton, NJ 08540.
E-mail: michael.winton@noaa.gov

TABLE 1. GFDL climate model glossary.

Model	Notes	Primary reference
CM2.0	CMIP3 model	Delworth et al. (2006)
CM2.1	CMIP3 model with climatology improved over CM2.0	Delworth et al. (2006)
CM3	CM2.1 with atmospheric physics, chemistry and vertical resolution improvements for CMIP5	Donner et al. (2011)
ESM2M	CM2.1 enhanced with ESM components for CMIP5	Dunne et al. (2012)
ESM2preG	ESM2M with isopycnal ocean replacing depth-coordinate	Rugenstein et al. (2013)
ESM2G	ESM2preG with improved climatology for CMIP5	Dunne et al. (2012)

et al. (2010) found that the magnitudes of radiative feedbacks depend upon ocean heat uptake in the Geophysical Fluid Dynamics Laboratory Climate Model version 2.1 (GFDL CM2.1) and so vary with time (see also Williams et al. 2008). To narrow the uncertainty to a particular component or parameter in a climate model, it is more useful to perform twin experiments where only a single part of the climate model is altered. To the extent that such an alteration influences climate sensitivity, it represents a source of uncertainty and merits further attention. Through this kind of systematic experimentation, the specific model formulations and parameters that need to be constrained in order to reduce uncertainty can be determined.

In this paper, we employ this strategy at a very coarse level by exploring the impact of the atmosphere and ocean component employed on the coupled climate model responses of global temperature, Northern Hemisphere sea ice, and Atlantic overturning to increasing greenhouse gas concentrations. We use a small ensemble of GFDL climate models that generates a large range of responses in the metrics reviewed above. Our approach is related to perturbed physics sensitivity experiments (e.g., Collins et al. 2007) except that our ensemble is much smaller but has also been developed systematically with the intent of producing models with good climatologies and natural variability. The next section introduces the three models to be used. Conceptually, these can be thought of as a trunk model and two branches: an atmosphere swap model and an ocean swap model. The following three sections compare the sensitivities of global mean surface temperature, Northern Hemisphere sea ice cover, and Atlantic meridional overturning circulation, respectively. The sixth section compares the response time scales of the models. The final section summarizes the conclusions.

2. Models and experiments

Starting from CM2.1, the GFDL CMIP3 generation model (Delworth et al. 2006), major efforts at GFDL have produced two new earth systems models (ESMs),

ESM2M and ESM2G using different ocean components (Dunne et al. 2012), and a new AOGCM, CM3, incorporating a new atmosphere component with a focus on chemistry and aerosol–cloud interactions (Donner et al. 2011; Griffies et al. 2011). Table 1 contains a brief description of these three and several other GFDL climate models discussed in this study.

CM3, ESM2M, and ESM2G have been developed to produce high-quality climatologies and realistic variability comparable to their predecessor CM2.1, which had one of the best climatologies in the CMIP3 group (Reichler and Kim 2008; Gleckler et al. 2008).

Dunne et al. (2012) conclude, based on an evaluation of the preindustrial climatologies of ESM2M and ESM2G, that neither model is fundamentally superior to the other. The ability to simulate the carbon cycle and the response to carbon emissions are the primary capabilities that earth system models have in addition to those of traditional atmosphere–ocean global climate model. In this paper we use only concentration forced experiments in order to allow comparison of AOGCM and ESM results. Griffies et al. (2011) describe CM3's climatology and show that it has similar magnitudes of sea surface temperature and salinity errors to CM2.1.

Table 2 summarizes the formulations of the new GFDL climate models. Although there are some differences in the ocean formulation of CM3 and ESM2M (Dunne et al. 2012) we do not expect them to contribute significantly to differences in the sensitivities of the two models. This expectation is because of the similarity of ESM2M and CM2.1 sensitivities, in spite of these ocean differences, and the fact that CM3 and CM2.1 have nearly identical ocean components. ESM2M and ESM2G are virtually identical in all components except for the ocean (Dunne et al. 2012). The same

TABLE 2. Summary of GFDL climate model components.

	CM3	ESM2M	ESM2G
Atmosphere	AM3	AM2	AM2
Ocean	MOM	MOM	GOLD
Land–sea ice	LM3–SIS	LM3–SIS	LM3–SIS

atmospheric and sea ice parameter tunings are used for both models. Therefore ESM2M may be thought of as the trunk in the three-model ensemble where an atmosphere swap, AM3 for AM2, leads to the CM3 branch, and an ocean swap, the Generalized Ocean Layer Dynamics (GOLD) isopycnal model for the depth-based Modular Ocean Model (MOM), leads to the ESM2G branch. This three-model ensemble is well suited to distinguish the influence of the atmosphere and ocean on the sensitivities.

The development of an isopycnal ocean component for the climate model was motivated in part by concern about spurious mixing and poor representations of overflows in depth-coordinate ocean models. Among other sources of spurious mixing (Griffies et al. 2000; Ilicak et al. 2011), depth-coordinate ocean models suffer from an artificially large mixing by dense plumes as they descend the stair-step topography, which might impact its simulation of meridional overturning and the response of the overturning to climate change (Winton et al. 1998). This concern has motivated the development of bottom boundary layer parameterizations for depth coordinate models (Legg et al. 2009; Danabasoglu et al. 2010). ESM2M partially alleviates this bias by making use of the Beckmann–Doescher parameterization (Beckmann and Döscher 1997). Like other isopycnal-coordinate models, all mixing in overflows must be explicitly parameterized in ESM2G. ESM2G uses the stratified shear-mixing parameterization of Jackson et al. (2008) in conjunction with the bottom-stress parameterization of Legg et al. (2006); this combination works well for various overflows (Legg et al. 2009). ESM2G's control climate AMOC is somewhat deeper but also somewhat weaker than ESM2M's, and closer to observational inferences for both metrics (Dunne et al. 2012). These differences have compensating effects on poleward heat transport, resulting in similar simulations of in the two models (Dunne et al. 2012). A comparison of climate model control states with isopycnal and depth coordinate ocean components models has been made using the third climate configuration of the Met Office Unified Model (HadCM3) (e.g., Megann et al. 2010) and several of the CMIP3 climate models have made use of hybrid ocean coordinates (e.g., Sun and Bleck 2006); however, as far as we know, the present study is the first to address the impact of the alternative ocean coordinate on climate sensitivities.

The atmosphere component of CM3, AM3, was developed to incorporate atmospheric chemistry and interactive aerosols. It contains similar horizontal resolution to AM2 but twice the vertical resolution with the refinement mainly devoted to the stratosphere. The cloud scheme was enhanced to predict droplet numbers based on aerosol concentrations. The goal of these enhancements was to simulate aerosol indirect effects; AM2 only simulated the direct effects. However, as will be shown later, the climate sensitivity was substantially increased by the changes and this accounts for most of the difference in the global temperature response between CM3 and ESM2M. A detailed analysis of these differences is beyond the scope of this paper. The primary focus of this paper is on the comparison of ESM2M and ESM2G sensitivities to radiative forcing while making use of the ESM2M–CM3 differences mainly as a point of comparison. Comparison of the carbon cycle responses of ESM2M and ESM2G is likewise left to future work.

We make use of four standard CMIP5 experiments. The first is a control run made with preindustrial atmospheric composition. After a long spinup of the models under this forcing to remove drift, the beginning of the control period for two idealized and one realistic forcing experiment is marked. The two idealized experiments are a 1% yr^{-1} increasing CO_2 experiment (achieving quadrupled CO_2 in year 140) and an abrupt CO_2 quadrupling experiment. These experiments are useful because the forcing is known, allowing the model sensitivity and time scales of response to be determined. The realistic experiment is forced with a concatenation of historical emissions and atmospheric concentrations of radiatively important species and a future scenario [Representative Concentration Pathway 4.5 (RCP4.5)] for these to extend the runs to 2100.

3. Global surface temperature

Our approach throughout will be to use the idealized 1% yr^{-1} CO_2 increase to quadrupling experiment to interpret the historical/projection experiment with future forcing following the RCP4.5 scenario (Clarke et al. 2007). The historical–RCP4.5 forcing scenario has aerosol radiative forcing rising in the twentieth century, reaching a peak in 2000, and then falling through the twenty-first century. The aerosol forcing in 1950 is about the same as in 2050 in this scenario according to calculations from the International Institute for Applied Systems Analysis (IIASA; IIASA 2011). CM3 and the ESMs calculate different aerosol forcing since only the former prognoses aerosol concentrations and produces aerosol indirect effects through cloud interaction. However, assuming the simulated effect scales with the IIASA-calculated radiative forcing, each model will have about the same aerosol forcing in 1950 and 2050, leaving only the sensitivity to greenhouse gases to account for surface climate changes between these dates. Since greenhouse gas forcings can be accurately estimated, this feature of the forcing allows us to make
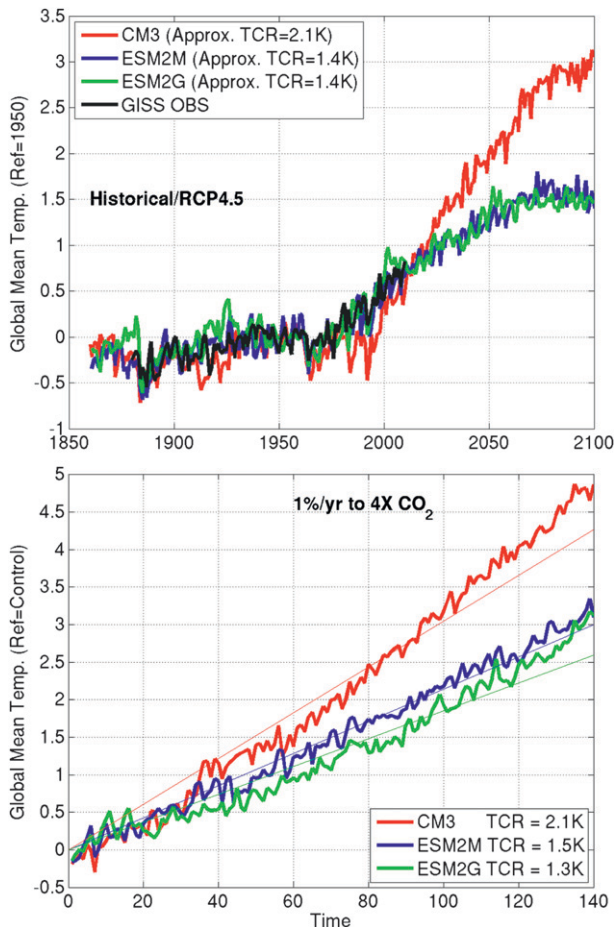
FIG. 1. Global and annual mean surface air temperature anomalies in the (top) historical–scenario and 1% yr^{-1} (bottom) CO_2 increase experiments. Anomalies in the upper panel are from the 1941–60 average and, in the lower panel, from the first century of the control experiment. The fits in the lower panel represent the 140-yr average differences with the control simulation used to calculate the TCRs.

TABLE 3. GFDL model transient climate responses (TCRs) and equilibrium climate sensitivities (ECSs) compared to statistics from a multimodel ensemble compiled by Winton et al. (2010). Transient sensitivities are calculated from 140-yr mean of 1% yr^{-1} CO_2 increase experiment. Values in parentheses use average over years 61–80 instead.

	TCR (K)	ECS (K)
CM3	2.1 (1.9)	4.6
ESM2M	1.5 (1.5)	3.2
ESM2G	1.3 (1.2)	3.2
Multimodel mean	1.8	3.4
Multimodel std dev	0.37	0.78

in model sensitivity—again entirely due to the atmosphere component.

We note that while ESM2M and ESM2G agree very well in their 1950 to 2100 warmings, ESM2G has somewhat less warming along the approach to 1950. Figure 1 (bottom) shows a similar difference in the 1% yr^{-1} experiments where a lag in ESM2G warming relative to ESM2M is established early and then maintained over a large range of forcing magnitudes. The warming simulated by the two models rejoins near 4 times CO_2 when the radiative forcing is 7 W m^{-2}, much larger than that attained in the scenario experiment, nominally 4.5 W m^{-2}. Both ESM2G and CM3 exhibit some nonlinearity in their temperature responses under 1% yr^{-1} CO_2 increases. We can minimize the impact of this nonlinearity by citing TCRs as the difference in the 140-yr averages between the 1% yr^{-1} and control experiments. Since the radiative forcing increase is linear, the 140-yr average gives a response to CO_2 doubling (corresponding to the linear fits in Fig. 1). The TCR values calculated using the 140-yr average, as well as those using the conventional year 61–80 averaging period, are listed in Table 3 along with the mean and standard deviation of 22 climate models compiled by Winton et al. (2010). ESM2M and CM3 straddle the multimodel mean with a difference of about 1.5 standard deviations while ESM2G is about 0.5 standard deviations less sensitive than ESM2M. Cloud feedback is the primary cause of the ESM2M–CM3 transient sensitivity difference (B. Soden 2012, personal communication). CM3 and ESM2G have slightly lower TCR values for the year 61–80 averaging period than for the 140-yr average due to the concave upward shape of their temperature rises (Fig. 1, bottom).

To understand the differences in TCR, it is useful to evaluate the equilibrium climate sensitivities (ECSs) of the models. We use a method that determines the equilibrium response by extrapolating the temperature change–heat uptake relationship to zero heat uptake. This method was proposed by Gregory et al. (2004) as
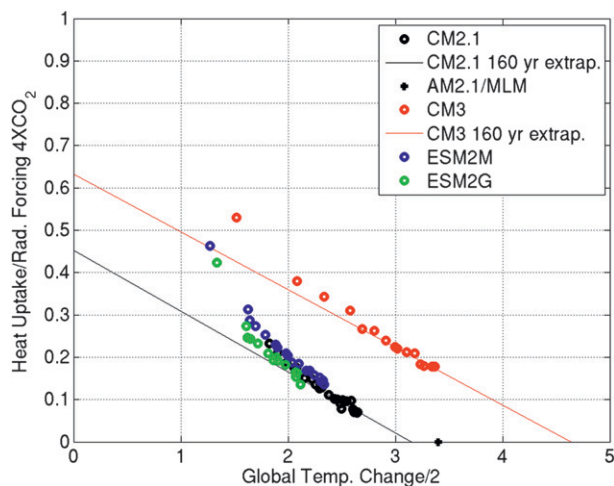
a rough sensitivity estimate from the historical–projection run in addition to that estimated, more conventionally, from the 1% yr^{-1} CO_2 increase runs.

Figure 1 shows the global mean surface temperature for the historical–RCP4.5 and idealized forcing scenarios. The global temperature projections show warming of about 3 K in CM3 and 1.5 K in the ESMs in 2100 relative to 1950. Therefore the ensemble uncertainty in these projections, about a factor of 2, is entirely due to the choice of atmosphere. Transient climate responses (TCRs) can be estimated from the 1950–2050 warming using a small inflation factor (1.09) to account for the slightly less than doubled CO_2 forcing over this interval. This method gives TCRs of 2.1 K for CM3 and 1.4 K for ESM2M and ESM2G. Consequently, we can attribute an intermodel uncertainty of about 50% to differences

FIG. 2. Global surface temperature anomalies (divided by two) plotted against heat uptake anomalies scaled by radiative forcing (assumed 7 W m^{-2} for quadrupled CO_2 in all models). Anomalies are relative to the first century of the control experiment. Twenty-year averages are shown for step CO_2 quadrupling except for CM2.1 where a stabilized section of the 1% CO_2 increase to quadrupling is used. The extrapolations of the last 8 points (160 yr) to 0 heat flux gives an estimate of the ECS for CM2.1 (assume the same for the ESMs) and CM3. The ECS for CM2.1 estimated using an atmosphere–slab-ocean experiment (Stouffer et al. 2006) is also shown.

an alternative to an atmosphere–slab mixed layer experiment. They found that the two methods gave consistent results. Our extrapolation uses ordinary least squares (OLS) on the 20-yr mean perturbation heat uptake and global temperature time series from sections of experiments where the forcing is stabilized at 4 times CO_2 (Fig. 2). The heat uptake is calculated as the perturbation downward heat flux at the top of the atmosphere. The heat uptake and temperature perturbations are calculated relative to the first century of the control experiment. The quadrupled CO_2 forcing, used for scaling the heat uptake, is twice the value for doubling in CM2.1 reported in Table 10.2 of Solomon et al. (2007). Performing the extrapolation on the last 160 years of CM2.1 and CM3, we obtain ECSs of 3.2 K for CM2.1 and 4.6 K for CM3. Noting that the ESM2M and ESM2G series from shorter runs are aligning with the CM2.1 series, we assign an ECS of 3.2K to both ESMs as well. The ECS estimated from this method for CM2.1 is in fairly good agreement with the value estimated using a slab ocean of 3.4 K (Stouffer et al. 2006), consistent with the Gregory et al. (2004) finding. A comparison of the ECSs to those of a multimodel ensemble (Table 3) shows the ESMs to be near the mean while CM3 is at the high end. The difference in ECSs between CM3 and the ESMs is fairly large, about 1.8 standard deviations of

the multimodel ensemble compiled by Winton et al. (2010).

Having the ECSs, we can evaluate the reasons for the differences in TCR. A single-equation model for the role of the ECS in these differences treats the global heat uptake N as the reduction to the radiative forcing R that causes the TCR to be less than the ECS. The degree of equilibration, in terms of these quantities, is written TCR/ECS = $1 - N/R$. However, Winton et al. (2010) found that this model systematically underestimates the impact of a given magnitude of heat uptake on reducing this ratio. This is because the global heat uptake is dominated by the ocean and has large contributions at subpolar latitudes where radiative feedbacks give it a larger influence on surface temperature than an equivalent CO_2 forcing. The ratio of the surface temperature responses to heat uptake and to CO_2 is referred to as the heat uptake efficacy.

Nontrivial efficacy is evident in Fig. 2 as the failure of the fitted lines to intercept the y axis at unity, indicating that ocean heat uptake changes have a larger impact on surface temperature than CO_2. If heat uptake had the same impact on temperature as CO_2, the model states would lie on the line between (0, 1) and (ECS, 0) in the figure. The differences in y intercept between the AM2-based models and CM3 indicate differences in efficacy.

The alternative expression for the degree of equilibration using efficacy is (Winton et al. 2010)

$$\text{TCR/ECS} = 1 - \varepsilon N/R, \qquad (1)$$

where ε is the diagnosed heat uptake efficacy. The heat uptake efficacy is similar to the efficacies used for radiative forcings other than CO_2. Formally, it is derived by treating the temperature response as the sum of two components, one CO_2 forced and the other heat uptake forced, with differing sensitivities (Winton et al. 2010). The appendix reviews the derivation and compares this description of the transient response with the approach used by Raper et al. (2002) and Dufresne and Bony (2008).

Equation (1) allows us to quantify the ocean influence on the TCR and diagnose the source of that influence into heat uptake magnitude and efficacy factors. The values for the terms in (1) for each of the three models are given in Table 4. The TCR and ECS values are taken from Table 1; N is the top of the atmosphere (TOA) heat uptake calculated from the model runs; R is the value for doubling reported in Solomon et al. (2007); and the efficacy is calculated from (1). First addressing the atmosphere swap, we note that CM3 and ESM2M have similar degrees of equilibration. Therefore, CM3's 40% larger TCR is mainly due its larger ECS. In the case of

TABLE 4. Ocean influence metrics: the degree of equilibration (TCR/ECS ratio), heat uptake efficacy (ε), and heat uptake (N) scaled by the radiative forcing (R). Standard values are calculated from 140-yr mean fields of 1% yr^{-1} CO_2 increase experiment. Values in parentheses use averages over years 61–80 instead.

	TCR/ECS	= 1 −	ε	N/R
CM3	0.46 (0.42)		1.4 (1.3)	0.39 (0.43)
ESM2M	0.47 (0.46)		1.5 (1.6)	0.35 (0.33)
ESM2G	0.42 (0.39)		1.8 (1.9)	0.32 (0.32)

the ocean swap, the ECSs are assumed the same so the ocean influence is the sole cause of the TCR differences. Table 4 shows that this difference is not due to the heat uptake, which is smaller in ESM2G than ESM2M. Rather it is due to the substantially larger efficacy of ESM2G's heat uptake.

Winton et al. (2010) found that the TCR had similarly strong positive correlations with ECS and heat uptake, and negative correlation with heat uptake efficacy in a 22-climate model comparison. Efficacy was negatively correlated with heat uptake. The values in Table 4 show that the GFDL model TCRs also correlate positively with ECS and heat uptake and negatively with heat uptake efficacy. Heat uptake efficacy also correlates negatively with heat uptake. Consequently, this small GFDL model ensemble conforms to the intermodel relationships found by Winton et al. (2010) in the larger multimodel ensemble. Using Eq. (1) along with these correlations, intermodel ECS and heat uptake efficacy differences are seen to be drivers of the intermodel TCR variation while heat uptake variation is a damping factor.

Table 4 also lists the degree of equilibration, efficacy, and heat uptake metrics for the TCR determined from narrower year 61–80 averages. The reasons for the nonlinearity of the CM3 and ESM2G temperature changes can be determined from differences in the metrics when averaged over the short and long periods since the differences of these averages are a measure of the nonlinearity. CM3's lower degree of equilibration in years 61–80 is due to a larger heat uptake for the midpoint of the experiment than for the 140-yr average. The ESM2G nonlinearity stems from a different source: the efficacy of the heat uptake is larger at the midpoint than for the experiment average.

The positive heat uptake–TCR relationship evident in Table 4 is particularly striking in the case of the ocean swap where one might expect the ocean model to influence the solution through heat uptake magnitude. Instead ESM2G has smaller warming and smaller heat uptake than ESM2M and so the efficacy difference drives the TCR difference. This behavior is also true of

the warming and heat uptake evaluated separately for each hemisphere (not shown). Winton et al. (2010) show that efficacy stems from the radiative response to ocean heat uptake, which is regionally focused in the subpolar oceans, particularly the North Atlantic and Southern Oceans. In searching for the efficacy difference between the models we should look for the location of heat uptake and its impact on the radiation budget.

The mechanism for the ESM2M–ESM2G efficacy difference is depicted in Fig. 3. ESM2G has a stronger surface freshening response to CO_2 doubling in the subpolar oceans of both hemispheres. Surface freshening stratifies the water column preventing heat from reaching the surface and allowing sea ice to form. Halocline expansion allows an advance of sea ice in some regions in both models but the effect is larger in ESM2G because of a larger salinity response. The sea ice response difference, in turn, causes ESM2G to have reduced shortwave absorption in its subpolar oceans relative to ESM2M (not shown). This radiative response difference accounts for ESM2G's higher heat uptake efficacy. The precise causes for the difference in salinity response are not known. However, Dunne et al. (2012) note that the mixed layers in the preindustrial control experiments are shallower in ESM2G than in ESM2M. A shallower mixed layer will give a larger surface freshening in response to a given perturbation in the surface freshwater budget. Dunne et al. (2012) also document considerable difference in the numerical treatment of the mixed layer in the two ocean models.

The halocline mechanism for the high ESM2G efficacy only works early in the warming simulations when the halocline extent imposes a constraint on the sea ice extent. As the warming proceeds, the sea ice shrinks back from the halocline edge and the ice extent differences between the ESM2M and ESM2G simulations diminish. This behavior is consistent with the finding that the ESM2G efficacy is larger at the midpoint of the 1% yr^{-1} CO_2 increase experiment than for the experiment average (Table 4). The larger efficacy occurs early in the experiment when heat uptake beneath an expanding halocline induces a sea ice response.

4. Northern Hemisphere sea ice cover

Figure 3 shows that, during the CO_2 increase, some regions experience ice advance while others have ice decline in both ESMs, and in both hemispheres. However, the Northern Hemisphere (NH) annual mean extent shows clear differences in the model aggregate responses (Fig. 4). The responses are different both in the historical–RCP4.5 projection and 1% yr^{-1} CO_2 increase runs. Satellite observations are also shown in
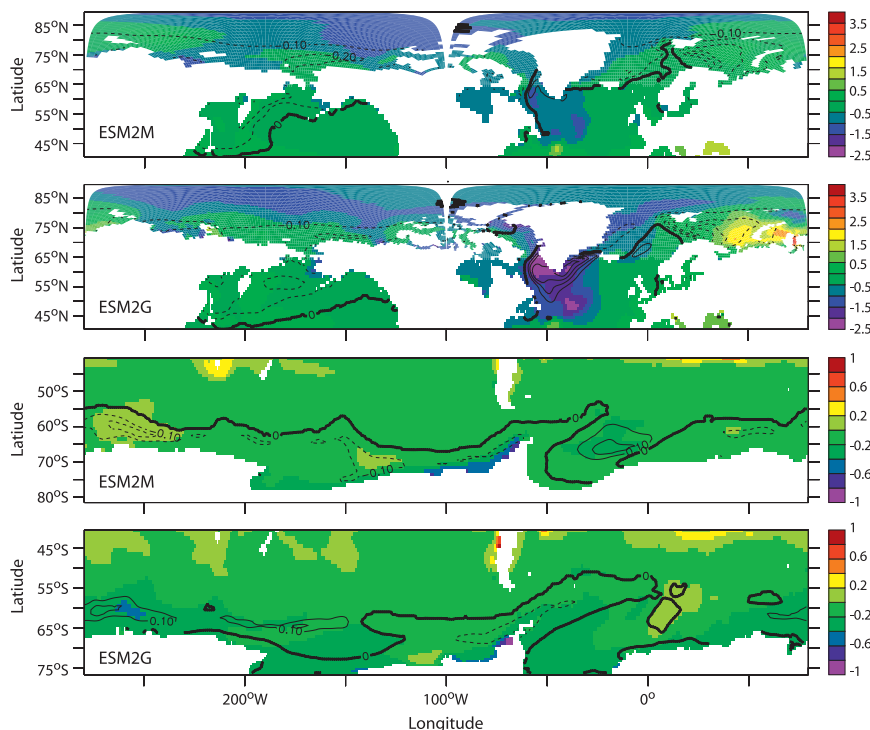
FIG. 3. (top 2 panels) Northern and (bottom 2 panels) Southern Hemisphere high-latitude salinity (shading) and annual mean ice extent (contours) responses under doubled CO_2 forcing (from 140-yr averages of 1% yr^{-1} CO_2 increase to quadrupling) for ESM2M and ESM2G. Note different salinity scales for the two hemispheres. Ice extent change contours are at 0.1 intervals for both hemispheres.

Fig. 4 (Fetterer et al. 2009). ESM2G has a large positive bias in NH extent while the other two models show good agreement with observations. ESM2G's sea ice albedo settings were kept the same as ESM2M's in spite of the bias in order to restrict the differences between the models to the ocean component. After removing this bias, all three models show reasonable agreement with the ongoing decline seen in satellite observations. Furthermore, ESM2M and ESM2G agree in the evolution of the NH extent from 1950 to 2100. As was the case with global temperatures, there is disagreement between the two models prior to 1950 with ESM2M showing a decrease relative to ESM2G. CM3 shows a decline in sea ice cover over the historical–projection run that is about 3 times larger than that in ESM2M and ESM2G. CM3 has ice-free Septembers beginning at the mid-twenty-first century for this scenario while ESM2M has September ice cover to the end of the century (not shown).

The NH ice extent in the 1% yr^{-1} CO_2 increase runs show these model differences even more clearly. ESM2G has no decline 60 years into the run followed by a steep decline that brings its response into better agreement with ESM2M by the time CO_2 quadrupling is reached. The difference in early behavior of the two models is consistent with the early differences in the historical/projection run. CM3 again has about 3 times larger decline than the ESMs, losing almost all of its NH ice cover—summer and winter—by the time of CO_2 quadrupling. This difference is too large to be explained directly by CM3's 40% larger global warming; rather, it is primarily due to its larger Arctic amplification. Recall that all three climate models contain the same sea ice component, indicating that the sea ice formulation does not closely constrain the sea ice cover sensitivity.

Early twentieth-century sea ice observations are not adequate to distinguish the differing behavior of ESM2M and ESM2G over that period. However, hydrographic and surface temperature observations weigh heavily against the behavior of ESM2G. Figure 5 shows salinity (top) and potential temperature (bottom) profiles at the location of ocean weather station Bravo in the Labrador Sea. Model profiles for the early and late twentieth century are compared to a modern observational climatology (Levitus et al. 1994). A surface freshening and cooling occurs over the twentieth century in the Labrador Sea with both ESM2M and ESM2G,
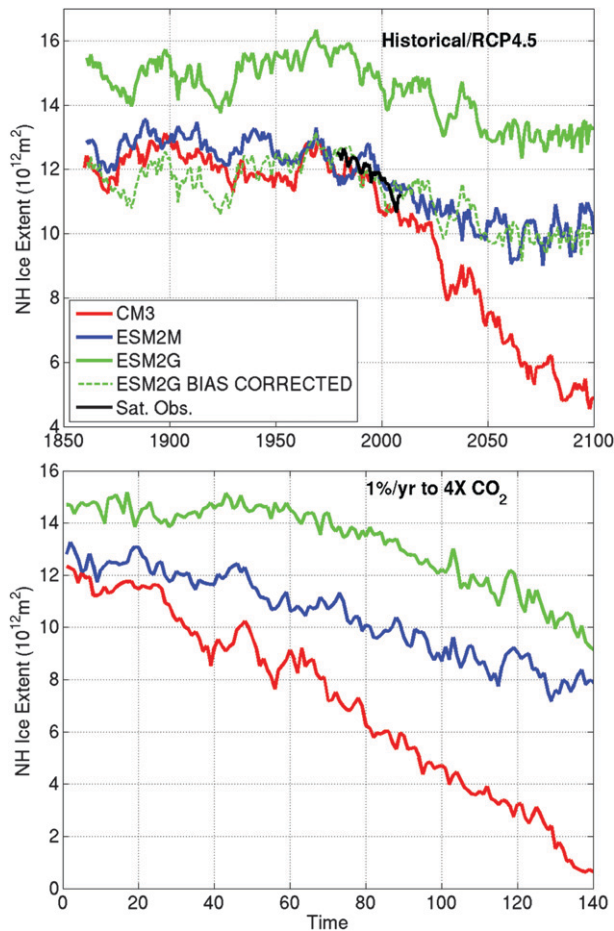
FIG. 4. Annual mean Northern Hemisphere sea ice extent in the (top) historical–scenario and (bottom) 1% yr^{-1} CO$_2$ increase experiments. The National Snow and Ice Data Center (NSIDC) sea ice extent index (Fetterer et al. 2009) is also shown in the upper panel (black).
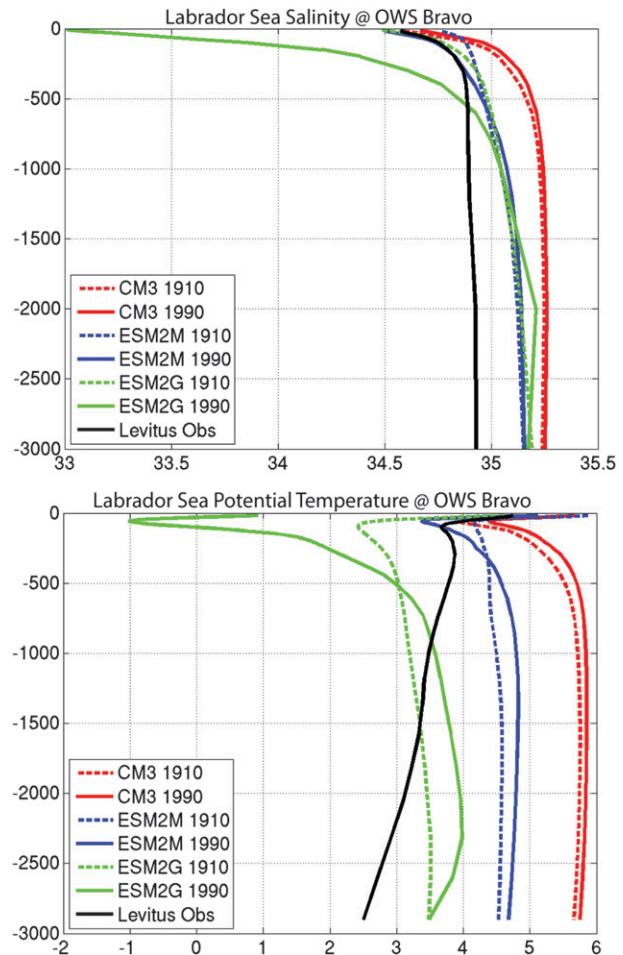

FIG. 5. Simulated and observed (top) salinity and (bottom) potential temperature at the location of ocean weather station Bravo in the Labrador Sea. Simulations are forced with historical forcing.

but the effect is extreme in ESM2G (about 4 times larger than in ESM2M), leading to a large disagreement between its modern profile and the observations. Below 350 m (in ESM2M) or 500 m (in ESM2G), the surface cooling trend is reversed, and all the models show twentieth-century warming trends in the Labrador Sea. It is noteworthy that this location has been monitored since 1949 and shows salinity fluctuations of about 0.2 salinity units at 200-m depth associated with variable mixing anomalies to 2-km depth (Yashayaev et al. 2003). In ESM2G, the decrease in 200-m salinity over the twentieth century is about 0.5 psu—sufficiently large to eliminate deep mixing in the Labrador Sea.

Figure 6 shows the temperature at Nuuk on the southwest coast of Greenland in the models and observations [Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) data are available online at http://data.giss.nasa.gov/gistemp/]. This observation registers the mid-twentieth-century warming followed by several cooling events that have been associated with salinity and temperature anomalies in the adjacent Labrador Sea (Belkin et al. 1998). CM3 and ESM2M have temperature fluctuations of the same magnitude but do not show anomalies as persistent as the observed midcentury warming. ESM2G has an abrupt 4-K cooling that establishes in the 1920s and 1930s and persists to the end of the century associated with the fresh anomaly. This difference in local temperature change is consistent with the global early-twentieth-century differences depicted in Fig. 1. The ESM2G fresh capping is clearly at odds with both temperature and hydrographic observations. A total of six historical runs is available for ESM2G: three concentration forced and three emission forced. All six
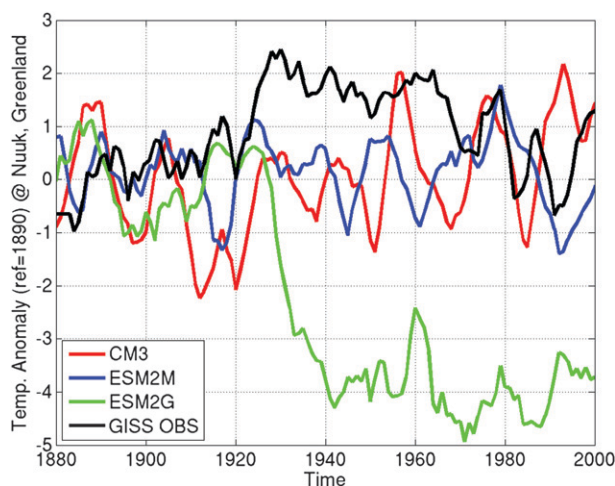
FIG. 6. Simulated and observed air temperature at Nuuk, Greenland. Five-year smoothing has been applied. Simulations are forced with historical forcing.

runs show fresh capping and sea ice advance into the Labrador Sea by the end of the twentieth century. The 1860 control run has large variability in the region but maintains long periods with ice-free conditions throughout the 700-yr run. Therefore the fresh capping behavior seems to be a forced response rather than a result of drift.

5. Atlantic meridional overturning circulation

The Atlantic meridional overturning circulation maximum streamfunction is shown in Fig. 7 for historical–projection and 1% yr^{-1} CO_2 increase runs. While the 1850–2100 evolution of the AMOCs is largely similar among the models, there are differences in the late twentieth and early twenty-first centuries. ESM2G shows a mid-twentieth-century decline relative to ESM2M associated with the fresh capping behavior discussed above. CM3 maintains its preindustrial level of overturning into the early twenty-first century, declining steeply thereafter. This behavior is likely due to the impact of aerosol forcing on the overturning (Delworth and Dixon 2006), which is emphasized in CM3 due to its larger aerosol radiative forcing (Donner et al. 2011). Aerosol forcing increases from preindustrial times up to about 2000, but then reverses abruptly and begins a century-long decline (IIASA 2011). The idealized forcing runs have overturning declines that are quite similar in the three models, with CM3 having a slightly larger decline than the ESMs. Even though ESM2G develops a stronger halocline in the North Atlantic relative to ESM2M (Fig. 3), this fresh capping behavior
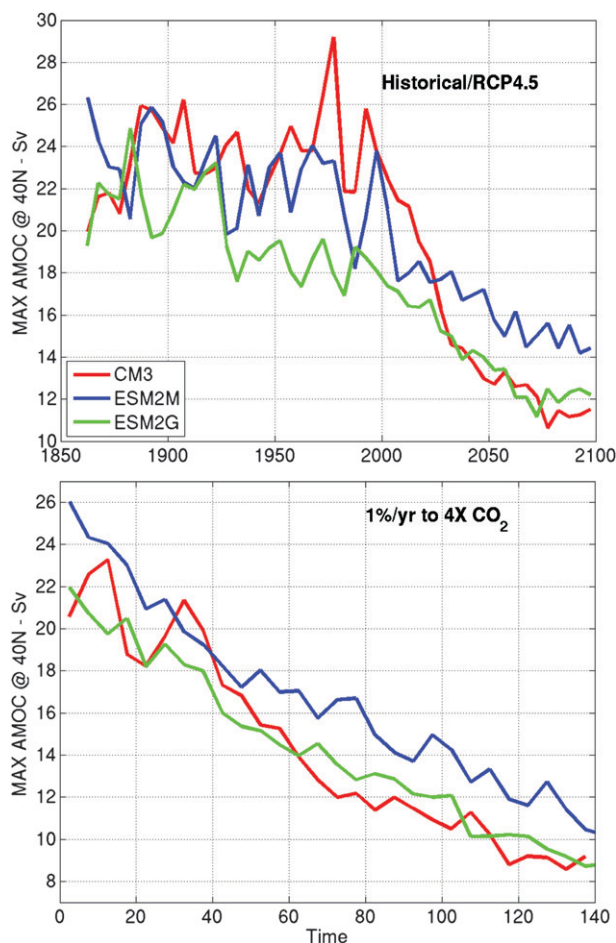

FIG. 7. Maximum Atlantic overturning streamfunction at 40°N in the (top) historical–scenario and (bottom) 1% yr^{-1} CO_2 increase experiments. All values are based on 5-yr means.

does not seem to influence its AMOC decline relative to ESM2M.

To obtain a broader perspective we plot, in Fig. 8, the control overturning and overturning decline near quadrupled CO_2 in in the 1% yr^{-1} CO_2 increase experiment along with values for three additional GFDL models: CM2.0, CM2.1, and ESM2preG (see Table 1 for brief descriptions). ESM2preG is a preliminary version of ESM2G documented in Rugenstein et al. (2013). Aspects of the climatologies and sensitivities of these models are discussed in Delworth et al. (2006), Stouffer et al. (2006), and Rugenstein et al. (2013). Overturning values for the CMIP atmosphere–ocean global circulation models have also been redrafted onto Fig. 8 from Gregory et al. (2005). The models as a group show a positive relationship between control climate overturning magnitude and decline under quadrupled CO_2 forcing. The range of declines varies by more than a factor of 2. All three of the models that are central to
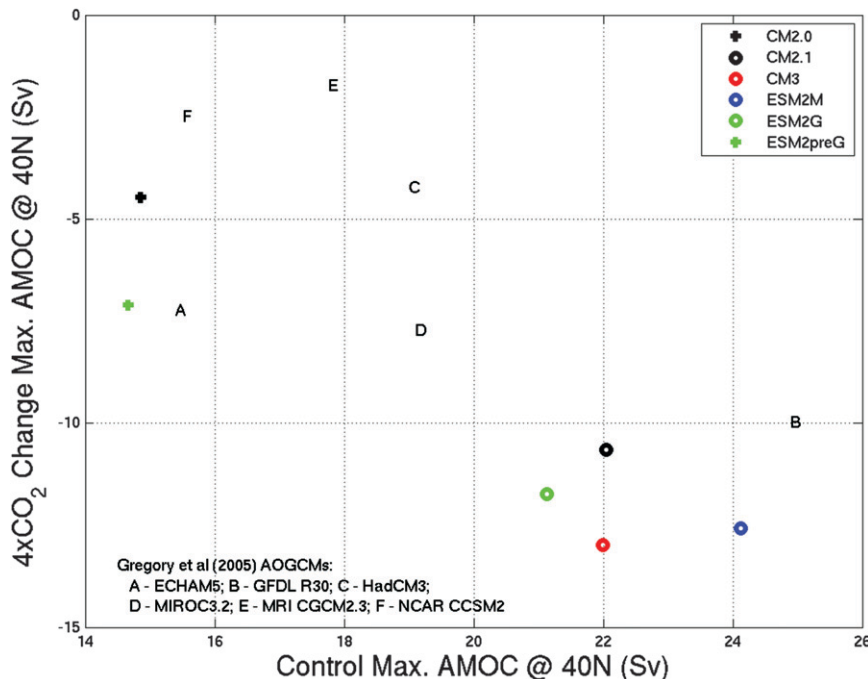
FIG. 8. Decline in maximum Atlantic overturning streamfunction at 40°N in years 120–140 of the 1% yr^{-1} CO_2 increase experiment relative to control plotted against the control value of the overturning. The control values used are first century averages. Values for the CMIP atmosphere–ocean global circulation models (letters) are redrafted from Gregory et al. (2005).

this study are seen to be models with large control overturning that experience large overturning declines. Neither the atmosphere nor the ocean swap has substantial impact relative to the large intermodel spread. However, closely related models based on both depth-coordinate oceans (CM2.0 and CM2.1) and isopycnal-coordinate oceans (ESM2preG and ESM2G) have substantially different overturning sensitivities, consistent with the positive relationship between control AMOC and AMOC decline discussed by Gregory et al. (2005). Rugenstein et al. (2013) found that these differences significantly impact the simulation of warming and sea ice decline at northern high latitudes.

Summarizing the results of Fig. 8, the ESM2M–ESM2G comparison shows that the ocean vertical coordinate does not significantly affect the AMOC strength response while comparison of the response differences of isopycnal coordinate (ESM2preG–ESM2G) and depth coordinate (CM2.0–CM2.1) pairs shows that the range of responses can be obtained with either vertical coordinate. These results suggest that the choice of ocean vertical coordinate is not central to the uncertainty of the overturning response although it does appear to influence the depth of the overturning in the northern North Atlantic and the interior ocean water mass properties (Dunne et al. 2012).

The GFDL models with weak control overturning (CM2.0 and ESM2preG) suffered from an absence of Labrador Sea convection and associated biases in temperature and salinity (Delworth et al. 2006; Rugenstein et al. 2013). Consequently, the models with strong control overturning (CM2.1 and ESM2G) and associated strong responses have been favored for development. However, Delworth et al. (2012) present the climate and climate sensitivity of a new high-resolution model, CM2.5, which has a weak overturning control simulation but less stratification than observed in the Labrador Sea. This model also has a weak overturning response as expected from the Fig. 8 relationship. It is possible that, in the coarse-resolution climate models, a large overturning and overturning response have been selected in the development process to counter a resolution problem in the Labrador Sea. However, it is also possible that the tendency to Labrador Sea biases in the coarse models is due to a problem with their similarly formulated atmosphere models (Delworth et al. 2006). Further experimentation with high-resolution models is needed to resolve this issue.

6. Response time scales

Because of the low heat capacity of the atmosphere, atmospheric temperature anomalies would be short
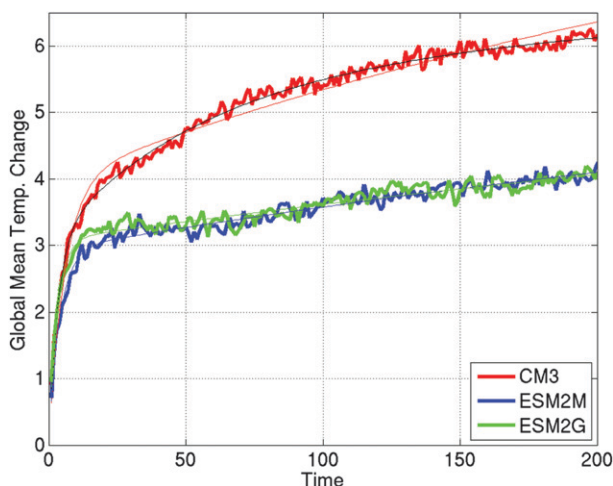
FIG. 9. Global surface temperature response to a step CO_2 quadrupling. A control drift of about 0.1 K century^{-1} has been removed from the CM3 response. A least squares, two-exponential fit is also shown for each model. For CM3 an additional three-exponential fit is shown (black). The fits are constrained to approach the estimated ECS as time goes to infinity. Fit parameters are listed in Table 5.

lived—decaying in a few months—without maintenance from ocean fluxes. Long time scales of response, well beyond those of the atmosphere, are evident in the nonlinear responses of the models to linear radiative forcing increases under a 1% yr^{-1} CO_2 increase in Figs. 1, 4, and 7. A step increase in CO_2 induces responses of a climate model on all of its available time scales, making the experiment a powerful tool for determining the time scales inherent in the simulated climate. The step response function can also be used to predict the model's global temperature response to more complex forcing (Hasselmann et al. 1993; Held et al. 2010; Good et al. 2011).

The model global temperature responses to instant CO_2 quadrupling are shown in Fig. 9. The ESMs respond very similarly while the CM3 response differs in magnitude and shape. The time scales of the responses are determined by a sum of exponentials to these series. The fits are constrained to asymptote to the equilibrium response determined from extrapolation of the temperature change/heat uptake relationship (Fig. 2). The fit

parameters are listed in Table 5. All three model responses have a rapid component with a time scale of a few years accounting for slightly less than half of the response (when two exponentials are used). The time scale of the second exponential is a century or longer for the ESMs than for CM3. A two-box model for this slow temperature response time scale (Held et al. 2010) gives the expectation that the increased equilibrium sensitivity of CM3 would lengthen this time scale. Additionally we note that the two-exponential fit has larger error for CM3 than for the ESMs (Table 5). Three exponentials must be used to achieve a similar accuracy for CM3 as for the ESMs. The three-exponential fit introduces an intermediate time scale of about 60 years between the multiyear and multicentury time scales, accounting for about 30% of the response.

The source of this intermediate time scale can be determined from Fig. 2, which shows the 20-yr mean temperature and heat uptake anomalies for the instant CO_2 quadrupling experiment with the three models. Earlier we used the long time scale behavior of temperature and heat uptake to extrapolate to the equilibrium climate sensitivity. However, the early behavior of the models—the small temperature response–large heat uptake marks in the upper left—show that CM3 also has a distinct behavior on the century time scale. In all three models there is a significant increase in temperature and decrease in heat uptake between the first and second 20-yr periods. Subsequently, however, the ESMs show a tight packing of the marks, indicating little change in either of these quantities. By contrast, CM3 shows significant increases in temperature and decreases in heat uptake over the next 40–60 years. This behavior associates declining heat uptake with the 60-yr time scale temperature increase evident in Fig. 9. This association was also evident in the nonlinear response of CM3's global temperature to the linear 1% yr^{-1} CO_2 increase forcing discussed earlier (Fig. 1). The depression of the TCR calculated near the midpoint of the experiment relative to that calculated with 140-yr average—an indicator of the nonlinearity—is due to larger heat uptake influencing the midpoint measure (Tables 3 and 4). In this experiment, as in the instant CO_2 quadrupling experiment, CM3's heat uptake declines on a multidecadal

TABLE 5. Exponential fits of the form $A(1 - \Sigma w_i e^{-t/\tau_i})$ to the response of global temperature to a step quadrupling of CO_2 shown in Fig. 8. Time scales (τ_i) are in years. The RMSE of the fits is also listed.

Model	A (K)	w_1	τ_1	w_2	τ_2	w_3	τ_3	RMSE (K)
CM3 (three exponentials)	9.2	0.33	3.3	0.28	58	0.38	1242	0.09
CM3	9.2	0.43	5.9	0.57	327	0	—	0.17
ESM2M	6.4	0.46	4.0	0.54	485	0	—	0.10
ESM2G	6.4	0.48	3.0	0.52	566	0	—	0.10

time scale. The mechanism for the introduction of this time scale will be a subject of future work.

7. Conclusions

We have explored the sensitivity of global temperature, Northern Hemisphere sea ice cover, and Atlantic meridional overturning strength in a set of three related CMIP5 generation GFDL climate models. These models roughly correspond to a trunk model (ESM2M) and two branches, an atmosphere swap model (CM3) and an ocean swap model (ESM2G), allowing a comparison of the relative impact of the atmosphere and ocean formulations on the sensitivities using components that have been carefully developed and evaluated. Tables 1 and 2 provide brief descriptions of these models. The ESM2M–ESM2G comparison assesses the impact of the choice of depth (ESM2M) or isopycnal (ESM2G) vertical coordinate and a number of ocean parameterization differences (Dunne et al. 2012) on climate sensitivities. Although the CM3 model was not developed to address sensitivity issues, its replacement of the AM2 atmosphere with AM3 had impacts on the sensitivities comparable to multimodel ranges. Consequently, the atmosphere swap sensitivity changes serve as a good standard of comparison for the ocean swap changes that are our focus.

The difference in transient global warming from the atmosphere swap is much larger than for the ocean swap and is due to increased equilibrium climate sensitivity (ECS). The cause of the increased ECS in CM3 is unknown but is presumably related to differences in moist physics (including convection and aerosol–cloud interactions) between AM2 and AM3. A smaller difference in transient climate response (TCR) between the ESMs is mainly due to a larger heat uptake efficacy in ESM2G stemming from a transient expansion of subpolar halocline and sea ice cover early in forced experiments. This behavior puts ESM2G at odds with twentieth-century observations of the North Atlantic. The larger subpolar surface salinity response in ESM2G may be related to its shallower mixed layers. The relationships of global temperature response and its explanatory metrics among the three models generally agree with the relationships found with a larger set of models by Winton et al. (2010). TCR is correlated with ECS and heat uptake, and anticorrelated with heat uptake efficacy. Heat uptake and heat uptake efficacy are anticorrelated. ECS and heat uptake efficacy differences drive the model TCR differences while the heat uptake differences damp them.

Generally, the result in this paper that the atmosphere formulation plays a larger role than that of the ocean in

transient warming confirms the finding of Collins et al. (2007) using a perturbed physics ensemble of HadCM3. They found that varying ocean mixing parameters such as vertical diffusivity gave a range of TCRs of only a few tenths of a degree while varying atmospheric parameters gave a TCR range of about 1 K. Here we have shown evidence that even fundamental changes to the ocean model formulation have little influence on the sensitivity. Although this result holds for the two ocean models presented in this paper, both of which have been subjected to extensive development to achieve accurate simulation of the climatology, Rugenstein et al. (2013) present a case where changes in ocean mixing parameters and the addition of geothermal heating have large impacts on both the quality of the simulated climatology and response magnitudes.

The ensemble shows that the atmosphere formulation has two sensitivity effects beyond its well-documented role in determining the equilibrium climate sensitivity. The first is the introduction of a new multidecadal ocean heat uptake and global temperature time scale accompanying the atmosphere swap. A second is the large increase in sea ice cover sensitivity leading to a near-complete loss of Northern Hemisphere cover under quadrupled CO_2 in CM3, a response that is roughly triple that of the ESMs. Since CM3's TCR is only 40% larger than for the ESMs, the ice cover loss for each degree of global warming is also substantially larger. All three models share the same sea ice component, indicating that resolving uncertainty in the sea ice formulation will not be sufficient for resolving uncertainty in the ice cover response.

Despite the large difference that the atmosphere swap made to the northern sea ice sensitivity, the AMOC strength sensitivity is fairly similar in all three models aside from transient late twentieth-century differences due to stronger CM3 aerosol effects and ESM2G fresh capping. The three models studied here have large control climate overturning and large overturning responses, consistent with the positive relationship between the two found in the multimodel ensemble of Gregory et al. (2005). Additionally, both depth and isopycnal coordinate models have weak control overturning counterparts that have weak responses. This result suggests that the control overturning represents a major uncertainty for the response that is not strongly affected by the choice of ocean vertical coordinate. Indeed, the similarity of the ESM responses indicates, more generally, that the downslope entrainment problem and other spurious mixing in depth-coordinate models are not major factors in the simulation of the broad, century-scale climate *response* metrics presented in this study.

APPENDIX

Transient Sensitivity Parameters

We briefly review the derivation of (1), which is used in section 3 to understand differences in the global temperature responses of the models. A more complete discussion of the equation is given in Winton et al. (2010). We also compare the parameters of this equation with a more commonly used set introduced by Gregory and Mitchell (1997) and used by Raper et al. (2002) and Dufresne and Bony (2008) for comparison of AOGCM sensitivities.

Both sets of parameters are depicted in Fig. A1, which shows a schematic climate state (T, N), where T is the transient global warming and N is the earth system heat uptake. The radiative forcing R is an important value on the heat uptake (vertical) axis and corresponds to the hypothetical heat uptake that occurs before the climate has responded to the forcing. The transient state evolves toward the equilibrium state $(T_{EQ}, 0)$. The states that we discuss are assumed to represent sufficiently long averages that the ocean mixed layer has adjusted to the forcing. Consequently the relevant heat uptake is occurring at the base of the mixed layer.

First we present the simpler Gregory and Mitchell (1997) description of the climate state that makes use of two feedback parameters: the heat uptake efficiency $\kappa = -N/T$ and the effective feedback parameter $\lambda_{EF} = (N - R)/T$. Combining these definitions, the transient response can be expressed as

$$T = -R/(\kappa + \lambda_{EF}). \tag{A1}$$

In Fig. A1, the heat uptake efficiency corresponds to the negative of the slope of the line from the origin to (T, N) and the effective feedback parameter corresponds to the slope of the line from $(0, R)$ to (T, N).

A difficulty with the Gregory and Mitchell (1997) parameters is their variation over the course of long climate change simulations. This is expected for the heat uptake efficiency, which approaches zero as the climate equilibrates with its radiative forcing. However, many long climate change experiments also show that λ_{EF} increases (becomes less negative) with time (Williams et al. 2008; Winton et al. 2010). Consistent with this, equilibrium sensitivity, T_{EQ}, calculated with atmosphere–slab-ocean model is larger than the effective
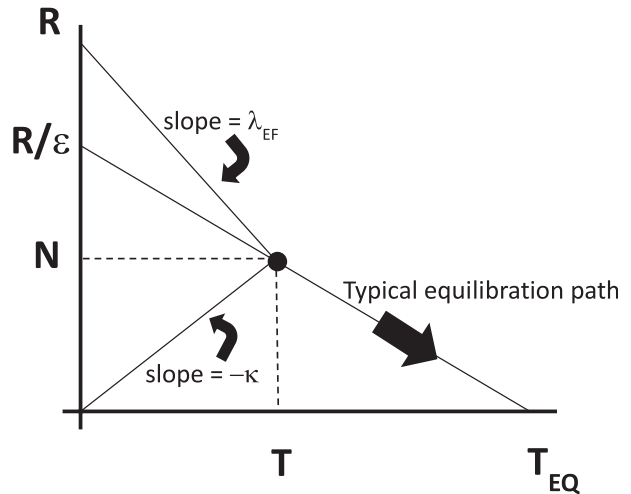

FIG. A1. Schematic climate state: global temperature change, T (x axis); global heat uptake, N (y axis). The parameters of the Gregory and Mitchell (1997) system (R, λ_{EF}, and κ) and the Winton et al. (2010) system (R, T_{EQ}, ε) are indicated on the figure.

sensitivity, $T_{EF} = -R/\lambda_{EF}$, for most AOGCMs (Winton et al. 2010).

Equivalent solutions to this problem were proposed by Williams et al. (2008) and Winton et al. (2010). Williams et al. (2008) proposed using an effective forcing. Winton et al. (2010) proposed applying an efficacy factor to the heat uptake as is commonly done for radiative forcings with global temperature impacts that are different from that of CO_2. Formally this is justified by treating the global temperature response as the sum of a response to the radiative forcing $T_{EQ} = -R/\lambda$ and a response to ocean heat uptake $T - T_{EQ} = N/(\lambda/\varepsilon)$, where ε is the efficacy and λ ($= -R/T_{EQ}$) is the equilibrium feedback parameter [Note: Winton et al. (2010) use the opposite sign convention for λ]. The response to N is therefore a factor of ε larger than the response to R. Summing the two components of the response and using the definition of λ, we obtain (1):

$$T/T_{EQ} = 1 - \varepsilon N/R. \tag{A2}$$

This relationship describes a trajectory from T to T_{EQ} along the line between $(0, R/\varepsilon)$ and $(T_{EQ}, 0)$ in Fig. A1. The Williams et al. (2008) effective forcing is equal to R/ε. Williams et al. (2008) and Winton et al. (2010) show that ε (equivalently R/ε) is relatively constant in long runs with constant forcing. Therefore (A2) replaces (A1) as a description of the transient climate state, substituting the time-invariant parameters ε and T_{EQ} for the time-dependent κ and λ_{EF}. This time invariance comes partly from treating the heat uptake as exogenous rather than incorporating it into a feedback and partly

from allowing different sensitivities to radiative forcing and heat uptake.

REFERENCES

Beckmann, A., and R. Döscher, 1997: A method for improved representation of dense water spreading over topography in geopotential-coordinate models. *J. Phys. Oceanogr.,* **27,** 581–591.

Belkin, I. M., S. Levitus, J. Antonov, and S.-A. Malmberg, 1998: ''Great Salinity Anomalies'' in the North Atlantic. *Prog. Oceanogr.,* **41,** 1–68.

Clarke, L., J. Edmonds, H. Jacoby, H. Pitcher, J. Reilly, and R. Richels, 2007: Scenarios of greenhouse gas emissions and atmospheric concentrations. U.S. Climate Change Science Program and the Subcommittee on Global Change Research Synthesis and Assessment Product 2.1 Subrep. 2.1A, 154 pp.

Collins, M., C. M. Brierley, M. MacVean, B. B. B. Booth, and G. R. Harris, 2007: The sensitivity of the rate of transient climate change to ocean physics perturbations. *J. Climate,* **20,** 2315–2320.

Danabasoglu, G., W. G. Large, and B. P. Briegleb, 2010: Climate impacts of parameterized Nordic Sea overflows. *J. Geophys. Res.,* **115,** C11005, doi:10.1029/2010JC006243.

Delworth, T. L., and K. W. Dixon, 2006: Have anthropogenic aerosols delayed a greenhouse gas–induced weakening of the North Atlantic thermohaline circulation? *Geophys. Res. Lett.,* **33,** L02606, doi:10.1029/2005GL024980.

——, and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate,* **19,** 643–674.

——, and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate,* **25,** 2755–2781.

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.

Dufresne, J.-L., and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *J. Climate,* **21,** 5135–5144.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate,* **25,** 6646–6665.

Eisenman, I., T. Schneider, D. S. Battisti, and C. M. Bitz, 2011: Consistent changes in the sea ice seasonal cycle in response to global warming. *J. Climate,* **24,** 5325–5335.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie, 2009: Sea ice index. National Snow and Ice Data Center, digital media. [Available online at http://nsidc.org/data/seaice_index/.]

Gleckler, P., K. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res.,* **113,** D06104, doi:10.1029/2007JD008972.

Good, P., J. M. Gregory, and J. A. Lowe, 2011: A step-response simple climate model to reconstruct and interpret AOGCM projections. *Geophys. Res. Lett.,* **38,** L01703, doi:10.1029/2010GL045208.

Gregory, J. M., and J. F. B. Mitchell, 1997: The climate response to CO_2 of the Hadley Centre coupled AOGCM with and without flux adjustment. *Geophys. Res. Lett.,* **24,** 1943–1946.

——, and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.,* **31,** L03205, doi:10.1029/2003GL018747.

——, and Coauthors, 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentration. *Geophys. Res. Lett.,* L12703, doi:10.1029/2005GL023209.

Griffies, S. M., R. C. Pacanowski, and R. W. Hallberg, 2000: Spurious diapycnal mixing associated with advection in a z-coordinate ocean model. *Mon. Wea. Rev.,* **128,** 538–564.

——, and Coauthors, 2011: The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate,* **24,** 3520–3544.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1993: On the cold start problem in transient simulations with coupled atmosphere–ocean models. *Climate Dyn.,* **9,** 53–61.

Held, I. M., M. Winton, K. Takahashi, T. L. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate,* **23,** 2418–2427.

IIASA, cited 2011: RCP Database, version 2.0.5. [Available online at http://www.iiasa.ac.at/web-apps/tnt/RcpDb/dsd?Action= htmlpage&page=welcome.]

Ilicak, M., A. Adcroft, S. M. Griffies, and R. W. Hallberg, 2011: Spurious dianeutral mixing and the role of momentum closure. *Ocean Modell.,* **45–46,** 37–58, doi:10.1016/j.ocemod.2011.10.003.

Jackson, L., R. Hallberg, and S. Legg, 2008: A parameterization of shear-driven turbulence for ocean climate models. *J. Phys. Oceanogr.,* **38,** 1033–1053.

Legg, S., R. W. Hallberg, and J. B. Girton, 2006: Comparison of entrainment in overflows simulated by z-coordinate, isopycnal and non-hydrostatic models. *Ocean Modell.,* **11,** 69–97.

——, and Coauthors, 2009: Improving oceanic overflow representation in climate models: The gravity current entrainment climate process team. *Bull. Amer. Meteor. Soc.,* **90,** 657–670.

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

Megann, A. P., A. L. New, A. T. Blaker, and B. Sinha, 2010: The sensitivity of a coupled climate model to its ocean component. *J. Climate,* **23,** 5126–5150.

Raper, S. C., J. M. Gregory, and R. J. Stouffer, 2002: The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *J. Climate,* **15,** 124–130.

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

Rugenstein, M. A. A., M. Winton, R. J. Stouffer, S. M. Griffies, R. W. Hallberg, 2013: Northern high-latitude heat budget decomposition and transient warming. *J. Climate,* in press.

Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *J. Climate,* **19,** 3354–3360.

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.

Stouffer, R. J., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part IV: Idealized climate response. *J. Climate,* **19,** 723–740.

Sun, S., and R. Bleck, 2006: Multi-century simulations with the coupled GISS-HYCOM climate model: Control experiments. *Climate Dyn.,* **26,** 407–428, doi:10.1007/s00382-005-0091-7.

Williams, K. D., W. J. Ingram, and J. M. Gregory, 2008: Time variation of effective climate sensitivity in GCMs. *J. Climate,* **21,** 5076–5090.

Winton, M., 2011: Do climate models underestimate the sensitivity of Northern Hemisphere sea ice cover? *J. Climate,* **24,** 3924–3934.

——, R. W. Hallberg, and A. Gnanadesikan, 1998: Simulation of density-driven frictional downslope flow in *z*-coordinate ocean models. *J. Phys. Oceanogr.,* **28,** 2163–2174.

——, K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate,* **23,** 2333–2344.

Yashayaev, I., J. R. N. Lazier, and R. A. Clarke, 2003: Temperature and salinity in the central Labrador Sea. *Hydrobiological Variability in the ICES Area, 1990–1999,* ICES Marine Symposia Series, Vol. 219, International Council for the Exploration of the Sea, 32–39.

Zhang, X. D., and J. E. Walsh, 2006: Toward a seasonally ice-covered Arctic ocean: Scenarios from the IPCC AR4 model simulations. *J. Climate,* **19,** 1730–1747.