A Moist Benchmark Calculation for Atmospheric General Circulation Models

MYONG-IN LEE

Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, and NASA Goddard Space Flight Center, Greenbelt, Maryland

MAX J. SUAREZ

NASA Goddard Space Flight Center, Greenbelt, Maryland

IN-SIK KANG

School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

ISAAC M. HELD

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

DAEHYUN KIM

School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

(Manuscript received 9 February 2007, in final form 28 March 2008)

ABSTRACT

A benchmark calculation is designed to compare the climate and climate sensitivity of atmospheric general circulation models (AGCMs). The experimental setup basically follows that of the aquaplanet experiment (APE) proposed by Neale and Hoskins, but a simple mixed layer ocean is embedded to enable air–sea coupling and the prediction of surface temperature. In calculations with several AGCMs, this idealization produces very strong zonal-mean flow and exaggerated ITCZ strength, but the model simulations remain sufficiently realistic to justify the use of this framework in isolating key differences between models. Because surface temperatures are free to respond to model differences, the simulation of the cloud distribution, especially in the subtropics, affects many other aspects of the simulations. The analysis of the simulated tropical transients highlights the importance of convection inhibition and air–sea coupling as affected by the depth of the mixed layer. These preliminary comparisons demonstrate that this idealized benchmark provides a discriminating framework for understanding the implications of differing physics parameterization in AGCMs.

1. Introduction

Atmospheric general circulation models (AGCMs) are indispensable tools in weather/climate predictions and climate change studies. Numerous model intercomparisons, however, have shown that climate simulations by current AGCMs still differ among themselves, as well as showing systematic deviations from the observed climate and its variability. For example, the simulation of the Madden–Julian oscillation (MJO) is still poor in most current models (Slingo et al. 1996; Waliser et al. 2003; Lin et al. 2006), and there is wide disagreement in the changes in cloud forcing in climate projections (Houghton et al. 2001). Since the large-scale dynamics is well represented in modern AGCMs, these failures are likely due to deficiencies in the model physical parameterizations. Most of all, large uncertainties still exist in the parameterization of moist processes and in their interaction with radiation and the large-scale flow.

Comparisons of models driven by the same observed climate
sea surface temperatures (SSTs) [i.e., Atmospheric Model Intercomparison Project (AMIP)-type intercomparisons; Gates et al. 1999] have proven very useful, providing a framework to identify common problems and differences in the models. The behavior of the models, however, is often masked by the complexity of the forcing and boundary conditions, including zonal asymmetries due to SSTs, topography, and land–ocean distributions that complicate the interpretation of the model responses. These considerations motivated the proposals by Held and Suarez (1994) and Neale and Hoskins (2000a,b) to study models using much simpler forcings. A limitation of this approach is that results cannot be validated against observations but, if experiments are well designed, this can be compensated by the clarity gained in comparing the model formulations.

One of the idealizations proposed is to run the models under axially symmetric forcings, and this approach has been used in many previous studies, for example, the aquaplanet experiments of Hayashi and Sumi (1986) and Randall et al. (1991). Recently, coordinated research is under way by several modeling groups to assess model simulations in such a simplified aquaplanet experiment (APE), proposed by Neale and Hoskins (2000a, hereafter NH00a). The aquaplanet experiments may be regarded as idealized AMIP simulations in which the full physical parameterizations are retained, but the lower boundary of the model is covered by a fixed, zonally symmetric SST distribution. In this configuration, useful sensitivity experiments can be performed by varying the underlying SST distributions. APE-type simulations serve to explore the interaction between model dynamics and various components of the physics and to assess, in an idealized setting, theories of complex phenomena involving interactions between moist processes and the large-scale flow. As an example, Neale and Hoskins (2000b, hereafter NH00b) investigated the sensitivity of the simulated intertropical convergence zone precipitation to the meridional gradient of the prescribed SST forcing in light of the study by Held and Hou (1980). Other examples include studies of the interactions between convective heating and tropical waves that produce tropical intraseasonal variability (e.g., Hayashi and Sumi 1986; Lau and Peng 1987; Tokioka et al. 1988).

There are, however, several important issues that require careful consideration in this aquaplanet framework with prescribed SST conditions. First, as in the conventional AMIPs, the mean climate is highly constrained by prescribed SSTs, which is a very strong regulator of the modeled convection and diabatic heating distributions. Second, imposed SSTs can also affect the climate’s natural variability. This may be important even at intraseasonal time scales (Wang et al. 2005) and the MJO (Flatau et al. 1997; Waliser et al. 1999; Watherson 2002; Maloney and Sobel 2004). Finally, with prescribed SSTs the surface heat budget is usually unbalanced, allowing the system to simulate states that would be far from equilibrium in a coupled setting. Estimates of climate sensitivity based on sensitivity to prescribed SSTs (e.g., Cess et al. 1990, 1996) are therefore indirect in that one needs to infer sensitivity from changes in the top-of-the-atmosphere energy imbalances.

To avoid those limitations without taking on the complexity of a full ocean GCM, one can consider using a slab ocean model. Examples of aquaplanet simulations with mixed layers are Alexeev (2003) and Langen and Alexeev (2005). In this configuration, the models generate their own surface temperatures and, thus, are freer to express differences in their physical parameterizations. Freeing the SST in this way allows one to study the equilibrium sensitivity of the climate to external perturbations, such as doubling CO$_2$ concentration. A simple slab ocean does, however, require that the surface heat budget be locally balanced, ignoring the effects of ocean advective heat transports. This results in some marked differences in behavior from fixed SST experiments. The coupled slab configuration, for example, does not produce double ITCZ structures that are often observed in AMIP simulations (Williamson and Olson 2003). Some of the effects of ocean transports can be included in this framework by prescribing local imbalance in the surface heat budget, but this does not account for changes in ocean transport with changing climate.

In this study, motivated by the aforementioned benefits of using the ocean–atmosphere coupled system, we extend the APE-type model intercomparison proposed by NH00a to include interactions with a slab ocean—a homogenous, wet, low-heat-capacity surface. We will focus specifically on evaluating the usefulness and limitations of this methodology in comparing the moist parameterizations of AGCMs. Investigations are primarily focused on the simulations of the zonal-mean states, the tropical MJO, and climate sensitivities induced by doubling CO$_2$.

For a preliminary test, we examine four AGCMs that are currently being used at three different institutes: the NASA Global Modeling and Assimilation Office [NASA’s Seasonal-to-Interannual Prediction Project (NSIPP) AGCM and the Goddard Earth Observing System-5 (GEOS-5) AGCM], the Seoul National University (SNU AGCM), and the NOAA Geophysical Fluid Dynamics Laboratory (GFDL AM2 AGCM). It is noted that those four AGCMs were chosen to illustrate the uses of the proposed benchmark simply be-
cause they were readily available and are not necessarily representative of the variety of formulations used in AGCMs. The convection schemes of the four models, for example, are all based on the relaxed Arakawa–Schubert (RAS) parameterization by Moorthi and Suarez (1992). Although this model selection may be less ideal for evaluating the simulation dependence on various closure assumptions in the convection scheme, in fact, the models still differ widely in their behavior, as will be discussed below. This is interesting in its own right since differences in convection schemes are often among the main candidates in explaining differences in model behavior.

Section 2 of the paper describes the experimental design, and section 3 describes the models and the experiments. Results from control runs and CO2 doubled runs are given in section 4, where we discuss the results separately for the basic climate, the tropical transients and subseasonal variability, the sensitivity to the mixed layer depth, and the climate sensitivities to the increased CO2 forcing. Section 5 gives a summary and conclusions.

2. Experimental design

The idealized benchmark proposed in this study is an extension of the APE framework with the atmosphere coupled to a fixed-depth mixed layer ocean. Most of the experimental design follows the detailed specifications in the APE (NH00a; see also the online document at http://www.met.reading.ac.uk/~mike/APE/ape_spec.html). We summarize the APE briefly for completeness. The benchmark idealization is of a rotating planet with a wet saturated surface at constant geopotential 1013.25 hPa and a mean moisture content of about 25 kg m$^{-2}$. Various geophysical constants, such as the earth’s rotation rate, the acceleration of gravity, and gas properties, are also specified as in the APE.

The extension to the APE formulation for this study is the coupling with the mixed layer ocean of fixed heat capacity. The surface temperature ($T_s$) is forced by incoming net heat fluxes ($F$):

$$\frac{dT_s}{dt} = \frac{F}{C_s},$$

where $C_s = 5 \times 10^7$ J m$^{-2}$ K$^{-1}$ is the heat capacity of the mixed layer, chosen to represent about 11.6 m of water. The amplitude of surface temperature variations depends on the specified heat capacity. This should be an important parameter that defines the zonal-mean time-mean states as well as the transient variability. The sensitivity to this parameter will be discussed in section 4c. In Eq. (1), $F$ includes all turbulent and radiative fluxes as well as the latent heat of melting of frozen precipitation. As in the APE, we assume no ice can form in the mixed layer. This eliminates ice–albedo feedback, which in any case would be quite unrealistic in the absence of ocean currents. Frozen precipitation is assumed to melt immediately on reaching the surface. The surface temperature in the polar region is allowed to fall below 0°C (as shown in Fig. 1a). Surface albedo is a global constant of 0.1, and surface roughness is also assumed to be fixed as 3.21 $\times$ $10^{-5}$ m, which is equivalent to a neutral drag coefficient of $10^{-3}$ defined at 10-m height.

3. Model descriptions and experiments

Table 1 gives a brief description of the four AGCMs used in this study. The models use comparable horizontal resolutions (~200–300 km). The vertical resolutions are quite different, ranging from 20 levels in the SNU model to 72 levels in GEOS-5. We note that the two NASA models (NSIPP and GEOS-5) use many of the same physical parameterizations. The other models were developed independently in each institution. While the four models use variations of the same convection parameterization, the Relaxed Arakawa–Schubert parameterization of Moorthi and Suarez (1992), the detailed implementations are different. For example, two models (GFDL and GEOS-5) use a convection trigger function in the RAS that eliminates weakly entraining convective plumes (Tokioka et al. 1988). The impacts of this trigger on tropical transient simulations will be discussed in section 4b.

The time required to reach equilibrium varies depending on the depth assumed for the mixed layer, but is generally less than 20 years, as determined by moni-
We also compared the simulated heat balances at the TOA and the surface from the last 5-yr averages to look for spurious heat sources or sinks in the atmospheric models. The net surface heat fluxes are close to zero in the four models (less than 0.05 W m$^{-2}$), showing that the simple mixed layer is equilibrated. The heat balances at the TOA have also equilibrated. Net heat fluxes at the TOA are about one tenth of a watt per square meter in the NSIPP and GFDL models. However, the net heat fluxes at the TOA in SNU and GEOS-5 are slightly larger than in the other two models, by roughly 0.5 W m$^{-2}$. This implies that the total energy is not exactly conserved in these models owing to spurious sources or sinks of energy in the atmosphere. Because these biases do not change significantly in the climate change experiments, this problem is not believed to affect the findings in this study substantially. We analyzed the model outputs for the last 5 years of each run, after equilibrium had been reached.

4. Results

a. Zonal-mean state

Figure 1a compares the zonal-mean surface temperatures simulated by the models. One of the models (NSIPP) generates a much warmer climate than the others, by more than 10 K at the equator and as much as 20 K near the poles. It is intriguing why the NSIPP model produces a climate so much warmer than the other models. This AGCM has been used as part of a coupled ocean–atmosphere model at the Global Modeling and Assimilation Office (GMAO) for many years and, in that more realistic configuration, it produces a very reasonable mean climatology without any heat flux adjustment for a century-long integration (Waj-
The difference in the global mean surface temperature among the models is fairly well explained by the difference in global-mean planetary albedo. The NISPP model produces a much lower zonal-mean planetary albedo (Fig. 1b), especially in the subtropics, whereas the other models show relatively larger reflection of solar radiation in the subtropics and tend to produce significantly lower surface temperatures.

The simulated meridional gradients of surface temperature are larger than that in the observed climatology. If we use annual mean, rather than equinoctial insolation, the poles are considerably warmer (not shown). We have chosen to retain the equinoctial forcing here to stay close to the APE formulation and because it provides an unambiguous way of retaining the diurnal cycle. Note that the NSIPP model produces higher surface temperature, even in higher latitudes (>60°N/S) where the planetary albedo is actually higher than in the other models. This can be explained as a result of stronger meridional heat transport by transient eddies. We found that the heat flux by transient eddies is strongest in the NSIPP model and weakest in the SNU model, which explains its sharp decrease of surface temperature from midlatitudes toward the poles.

The aquaplanet models also produce an exaggerated intensity of tropical precipitation, with a single ITCZ at the equator (Fig. 1c) and very dry subtropics in all models. Although several observational studies indicate the existence of double ITCZs straddling the equator (e.g., Zhang 2001; Liu and Xie 2002), many current AGCMs and ocean–atmosphere coupled GCMs exaggerate this structure (the “double ITCZs problem”). A number of aquaplanet simulations using GCMs with prescribed SST conditions have shown that models can produce single or double ITCZs depending on the type of convection parameterization used (e.g., Hayashi and Sumi 1986; Hess et al. 1993), on the parameters chosen within the same convection scheme (e.g., Lee et al. 2003), on model resolution, on the meridional gradient of prescribed SST (NH00b), or even on the integration time step (Williamson and Olson 2003). One can argue that slab ocean models with no imposed oceanic heat fluxes, such as we use in our calculations, should not produce multiple ITCZs, regardless of the schemes used. In all cases, there is a net downward radiative flux near the equator at the top the atmosphere. With fixed SSTs, this heating can be passed down through the lower boundary, allowing the atmosphere to produce the equatorial subsidence associated with double ITCZs.

When a slab is used and the surface heat budget is in equilibrium, no net heat flux is allowed at the surface. In this case, the top-of-the-atmosphere flux must be balanced primarily by energy transport away from the equator and, assuming that the circulation is direct, this implies that there is upward motion (an ITCZ) at the equator. Consistent with this expectation, none of these four models generate double ITCZs.

On the other hand, the models exhibit large differences in the magnitude of ITCZ precipitation at the equator. In general, the simulated intensity of the ITCZ precipitation seems to be proportional to the meridional gradient of the zonal-mean surface temperature (Fig. 2). NH00b found a similar result of widening of

<table>
<thead>
<tr>
<th>Model</th>
<th>NASA/NSIPP</th>
<th>GFDL/AM2</th>
<th>SNU AGCM</th>
<th>NASA/GEOS-5</th>
</tr>
</thead>
</table>

[Table 1. Model descriptions.]
the ITCZ precipitation structure in response to a decrease in the gradient of the prescribed SST near the equator.

As might be expected from its warm climate, the subtropical evaporation rate (Fig. 1d) in NSIPP is the largest among the models, implying a large equatorward moisture transport for the maintenance of its strong ITCZ. Overall, the magnitude of subtropical evaporation in the models seems to be inversely proportional to the subtropical planetary albedo, determined by cloudiness in that region. Since the surface albedo is prescribed, the differences in the planetary albedo are caused primarily by differences in cloud distributions.

Figure 3 compares the zonal-mean cloud radiative forcings. In all four models, the shortwave cooling effect of clouds dominates the longwave heating effect, as in observations. Accordingly, the net cloud radiative forcings of the models are strongly negative. The NSIPP model simulates the weakest net cloud radiative forcing, especially in the subtropics. Interestingly, the net forcing in the NSIPP model is almost the same as in observations, albeit due to a cancellation of large differences in the shortwave and longwave components. Nevertheless, it highlights the fact that the other models are maintaining more realistic temperatures with some 20 W m$^{-2}$ more cooling due to clouds than observed.

We compare the latitude–height distributions of the simulated zonal-mean zonal wind from the four control experiments in Fig. 4. The idealized versions of the AGCMs tend to simulate very strong jet streams. This is a common feature in aquaplanet experiments (see, e.g., NH00b), probably due to the use of equinoctial insolation and, in our case, to the absence of oceanic
heat transport. The models simulate the height of the jet streams near the 200 hPa level, close to the observed height, although in the NSIPP model it is much higher, with maxima near 100 hPa.

In Fig. 5, we compare the mass streamfunctions for the zonal-mean meridional circulation. It is interesting to note that the simulated magnitude of the Hadley circulation is not proportional to the strength of ITCZ precipitation among the models (cf. Fig. 1c). In particular, the NSIPP model simulates the weakest Hadley cell, whereas it has the most intense ITCZ precipitation. Since the NSIPP climate is the warmest, this behavior is consistent with the weakening of the tropical circulation with global warming, as discussed by Betts and Ridgway (1989) and more recently by Held and Soden (2006). This point will be revisited in section 4d, where the models are tested under doubling CO₂ concentration.

There are notable differences in the zonal-mean state between NSIPP and GEOS-5, even though they use many of the same physical parameterizations. The two models simulate quite different planetary albedos and cloud radiative forcings, indicating the importance of differences in their cloud formation schemes. One important difference in the models is the use of a scheme in the NSIPP model that destroys low-level stratocumulus cloud over the subtropics. This has been implemented by including an ad hoc sink term for cloud fraction in the presence of convection. (J. T. Bacmeister 2006, personal communication). Inclusion of this parameterization in the NSIPP model results in an improved climate in realistic simulations (both coupled and AMIP-type simulations) by reducing excessive cloudiness over the subtropics. In the current experiments, however, with an interactive, zonally symmetric lower boundary, this parameterization tends to completely eliminate low-level subtropical clouds. This behavior is evidently responsible for the anomalously warm climate produced by this model. Thus, a formulation choice that was merely a fine-tuning of subtropical stratocumulus regimes in a realistic setting takes the model to an extreme climate in the more idealized experiment. This should sound a note of caution that care must be taken when relating these idealized benchmark results to more complete models.

To examine this idea further we conducted an addi-
tional experiment with the NSIPP model by simply disabling the stratocumulus dissipation process. These results are compared with other simulations in Fig. 1 (indicated as NSIPPa). The test run simulates an overall increase in the planetary albedo, particularly over the tropics and subtropical region, lowering of the surface temperature, and resulting in an overall reduction of surface evaporation. However, the simulation still shows a deficit of low-level clouds and warmer surface temperatures compared with the GFDL and GEOS-5 simulations. The ITCZ precipitation also becomes substantially weaker and closer to that in the GEOS-5 run. In the modified run, the increased subtropical shortwave cloud radiative forcing (cf. Fig. 3) may substantially cool down the oceanic surface and suppress the evaporation. This can cause a reduction of the ambient humidity and equatorward moisture transport, and hence weakening of the ITCZ precipitation. This mechanism seems to be one of the more important aspects in explaining the intermodel differences in ITCZ precipitation intensity.

Our results illustrate both a strength and deficiency of the mixed layer formulation. The models can generate very different mean temperatures due to differences in cloud parameterizations, which is itself a potentially useful way of isolating important differences between models. On the other hand, these temperature responses can then dominate many of the other aspects of the circulation, making it more difficult to isolate other effects of alternative parameterizations.

b. Tropical transient variability

In this section, we focus on the model simulations of the tropical MJO and various convectively coupled equatorial waves by considering spectral characteristics in the simulated daily precipitation field. The space–time power spectrum analysis technique (Hayashi 1982) is useful, not only for identifying the dominant temporal scales (frequency) of tropical convective disturbances, but also for identifying the spatial scales (zonal wavenumber) that are associated with them.

The raw power spectrum of tropical convection variables, such as outgoing longwave radiation (OLR) or precipitation, generally shows a very red nature in both wavenumber and frequency. Wheeler and Kiladis (1999) normalized the spectrum of the observed OLR by dividing it by an estimate of the background power (regarded as a red noise) and showed that variance peaks standing above the background spectrum correspond to the normal modes of the linearized shallow-water system identified by Matsuno (1966). This indicates that a significant portion of tropical convective disturbances is organized in waves. We applied the same analysis technique of Wheeler and Kiladis to the simulated daily precipitation from the four control runs. Figure 6 compares the spectral variances, normal-
ized by the estimated red background, of the symmetric component about the equator of precipitation between 15°S and 15°N. We also show the spectrum of the observed Global Precipitation Climatology Project (GPCP) (Huffman et al. 2001) daily precipitation for a comparison with the model simulations.

Observed variances show the dominant signals of the convectively coupled equatorial waves, including the moist Kelvin, equatorial Rossby (ER), and westward inertio–gravity (WIG) waves, whose maximum variances coincide well with the dispersion curves obtained from the shallow-water theory (e.g., Wheeler and Kiladis 1999). The simulated variances are quite different among the models. For the eastward propagating waves, two models (NSIPP and SNU) exhibit very weak variances in the subseasonal time scale (30–60 days). On the other hand, in the spectral band of less than 20-day period, the signatures of the moist Kelvin waves are dominant in these models. The other two models (GFDL and GEOS-5) tend to simulate stronger variances in the longer subseasonal time scales. Simulated phase speeds of the moist Kelvin waves in those two models are quite slow. The variance separation between the high-frequency Kelvin waves and the low-frequency MJO (i.e., the eastward variance of 30–60-day period with zonal wavenumbers 1–3) is not as clear as in the observations; the enhanced eastward variances in the subseasonal time scale seem to be mostly contributed by the retarded Kelvin waves in shallower equivalent depths (as in GEOS-5, for example).

For the westward propagating modes, the models tend to simulate weak variances for the ER waves, and strong variances for the tropical depression (TD)-type disturbances. These disturbances show nondispersive characteristics and their variance is centered on a constant phase speed of about 10 m s⁻¹, drifting with the mean easterlies of the background flow (Lee et al. 2001). Relatively strong zonal-mean easterlies (Fig. 4) in the lower troposphere might enhance the variance of the advective disturbances in the aquaplanet models. Note also that the synoptic-scale (2–5 days) westward variances are prevalent in GFDL and GEOS-5, probably coming from the WIG signal.

Figure 7 shows the variance spectra for the antisymmetric component of the precipitation variability. The observed spectrum shows large variances in the westward mixed Rossby–gravity (MRG), and n = 0 eastward inertio–gravity (EIG) waves. The MJO signal is
also evident in the eastward variances in the observed spectrum. The models commonly fail to reproduce the MRG and \( n = 0 \) EIG waves, as well as the MJO. As in the symmetric variances, the models tend to simulate stronger TD-type signals.

In a preliminary study, we have confirmed that the variance patterns of the space–time power spectrum in this idealized aquaplanet run are quite similar to the patterns obtained from the realistic AMIP-type simulations. The extent to which these aquaplanet configurations can be used to interpret more realistic model intercomparisons deserved further exploration.

As described, the convectively coupled equatorial waves simulated in the NSIPP and SNU models agree best with shallow-water dispersion curves for relatively deep equivalent depths, with fast propagation, whereas the other two models (GFDL and GEOS-5) agree with curves for shallower equivalent depths and exhibit slower propagation.

The difference in phase speeds suggests that the vertical profiles of the diabatic heating should be quite different among the models. Vertically deeper heating tends to excite waves with long wavelength, which propagate faster (Takahashi 1987; Lau and Peng 1987), and this provides a mechanism for slowing down the convectively coupled Kelvin wave. However, Mapes (2000) and Lin et al. (2004) argued that a simple peak-altitude characterization of heating may be misleading, and the vertical heating can be either top-heavy or bottom-heavy at different times. Lin et al. showed that the observed vertical profile of heating at the time of maximum precipitation is very top-heavy owing to the enhanced stratiform precipitation, which contributes more than 60% to the intraseasonal precipitation anomaly. They also showed that simulated vertical heating profiles by many GCMs are substantially different from those observed, with pronounced middle-heavy structures. Based on those results, they suggested that the models systematically underrepresent stratiform-like precipitation by the MJO.

We thus turn to examining how the models separate the precipitation signals into convective and stratiform precipitation. Figure 8 compares Hovmöller diagrams of the simulated precipitation in the equatorial plane (averaged 3°S–3°N) for the total precipitation (the sum of convective and stratiform precipitation) and for the stratiform precipitation only. It is clear that the NSIPP and SNU models simulate relatively fast-propagating moist Kelvin waves. In these models, the stratiform precipitation contributes little to the total precipitation variability, and the total rainfall comes mostly from the parameterized deep convection scheme. On the other hand, the stratiform precipitation appears to be the more dominant type in the GFDL and GEOS-5 mod-
els, which simulate quite slowly propagating equatorial disturbances.

As the deep convection schemes in the four AGCMs are all based on the RAS scheme, the ratio of convective precipitation to the stratiform precipitation must be largely affected by the details of implementation; for example, the use of convection triggers or inhibition functions. Note that the two AGCMs that simulate larger stratiform precipitation set a critical minimum value for the cumulus entrainment rate in the convec-
tion scheme (Tokioka et al. 1988). This modification favors shallower convection with more strongly entraining clouds. With a strong Tokioka constraint, deep convection is suppressed, and the modification works as an inhibition function. In this case, the grid-scale condensation scheme becomes more active.

The suppression of convective rainfall by the Tokioka modification seems to be strongest in GFDL. We further examined the sensitivity of the GFDL model to the degree of suppression, which depends on the minimum allowed entrainment rate in the convective updraft, $\mu_{\min} = \alpha/D$, where $D$ is the planetary boundary layer depth and $\alpha$ is a disposable parameter. We performed experiments at $\alpha = 0.05, 0.025$ (control), 0.01, and 0.0 (no suppression). Figure 9 compares the Hovmöller diagrams of total precipitation at the equator (3°S–3°N) with the different values of $\alpha$. It turns out that increasing the Tokioka inhibition suppresses light rainfall and increases the temporal and spatial rainfall variability; but the zonal-mean rainfall distribution is hardly affected (not shown). Deep convective rainfall is predominant in the no constraint run (Fig. 9a), with significant rainfall intensity greater than 10 mm day$^{-1}$ over most of the equatorial region. These characteristics are similar to those in the unconstrained NSIPP and SNU AGCMs (cf. Fig. 8). As the constraint becomes stronger, the model tends to reduce the convective rainfall (cf. Figs. 9b–d), and the precipitation is more concentrated in propagating bands. Deep convection is strongly suppressed in the strongest constraint case (Fig. 9d), which produces very intense and localized disturbances with most of the rainfall coming from the grid-scale condensation scheme.

As shown in Fig. 9a, the propagation speeds tend to be constant and fast when the deep convection scheme dominates, whereas they become more irregular or disturbances tend to be more stationary when the grid-scale condensation scheme dominates. These characteristics are consistent with the observational findings of Lin et al. (2004) in that shallower vertical heating results in slower propagation speed. To a large degree, our results confirm the findings in many previous modeling studies. The moist convection scheme in GCMs is the key determinant of the nature of MJO simulations and that a threshold for activation of moist convection (i.e., the trigger/inhibition functions) is an important element (Wang and Schlesinger 1999; Lee et al. 2003; Lin et al. 2008).

c. Sensitivity to the ocean mixed layer depth

Three additional mixed layer experiments were conducted for the NSIPP and GFDL models with different mixed layer depths. These values are indicated in Table 2. We compare the zonal-mean states in Figs. 10 and 11 for the NSIPP and GFDL models. Here we show the changes in low latitudes between 30°S and 30°N, where the sensitivity is largest. The two models exhibit different sensitivities in the zonal mean precipitation. While the NSIPP model shows relatively small changes in the ITCZ precipitation (Fig. 10c), the GFDL model tends to increase the equatorial precipitation significantly as the mixed layer depth increases (Fig. 11c). These behaviors are consistent with those of their zonal-mean surface temperature. Although the zonal mean temperatures increase in both models for deeper mixed layers, the meridional temperature gradient between the equator and subtropics becomes much larger in the GFDL model. It is also found that deepening of the mixed layer depth tends to produce stronger zonal-mean circulations in both models (not shown).

Why deeper mixed layers should increase ITCZ precipitation is less clear. In both models, a shallower mixed layer results in greater cloud amount, higher planetary albedo, and a colder climate (Figs. 10 and 11). A smaller heat capacity enables stronger SST variation in time, and this might lead to increased convective activity, especially shallow convection in the subtropics. This will result in an increase of cloudiness, a decrease in surface evaporation over the subtropics, and less moisture available for the equatorward transport to the tropical ITCZ region. This mechanism seems to explain the more meridionally expanded structure of the ITCZ precipitation in the shallower mixed layer experiment, although the change is not monotonic. The time variances of surface temperature and precipitation do increase significantly in the subtropics as the mixed layer depth decreases (not shown), consistent with this line of argument.

Figure 12 compares the space–time power spectra of the symmetric component of precipitation (15°S–15°N) for the four different mixed layer depth cases in the GFDL model. Overall, the eastward variance in subseasonal time scales (periods longer than 10 days) tends to increase as the mixed layer deepens. This result is somewhat contradictory to the findings of Watterson (2002) who found that the amplitude of the simulated MJO-like pattern increases in the shallower mixed layer depth case, in his coupled GCM tests with 10-m and 50-m depths. On the other hand, Maloney and Sobel (2004) showed a nonmonotonic change in variance due to changes in mixed layer depth, with the maximum for a depth of 20 m. Considering that the ITCZ precipitation is systematically increasing with deepening of the mixed layer in the GFDL case, our result is more consistent with the finding of several studies (e.g.,
Slingo et al. 1996; Wang and Schlesinger 1999; Lee et al. 2003) that demonstrated that the tropical subseasonal variability becomes stronger with narrower or stronger ITCZ precipitation at the equator. In this regard, this explanation is consistent with the NSIPP model result, which showed no systematic changes in the precipitation spectra or in the strength of the ITCZ with increased mixed layer depth (not shown).

Our results suggest that the effects of air–sea coupling on tropical variability, as modified by varying the mixed layer depth, are state-dependent and model-dependent. A systematic comparison between mixed

---

**Fig. 9.** Longitude–time cross sections of daily mean (top) total and (bottom) stratiform precipitation rates (mm day$^{-1}$) averaged over the 3°S–3°N latitude belt from the GFDL model sensitivity test to the cumulus entrainment minimum parameter (a) $\alpha = 0.0$, (b) $\alpha = 0.01$, (c) $\alpha = 0.025$ (control), and (d) $\alpha = 0.05$. 
layer coupled runs and fixed SST runs, where the two zonal-mean states are kept close to each other, will be useful to identify the air–sea coupling impacts with no mean state drift. A clean way of making such comparisons is by prescribing the time-averaged zonal-mean SST distribution from the equilibrated mixed layer experiment in the fixed SST run. These are left for future work.

d. Sensitivity to the doubling CO₂

In this section, we consider the effects of doubling the CO₂ concentrations in the four AGCMs. We emphasize that, by construction, sea ice and snow–albedo feedback have been eliminated in this framework. Figure 13 compares changes in the global-mean surface temperatures and percentage change in global-mean water vapor and precipitation for doubled CO₂. The two NASA models show larger increases in temperature compared to the others, but they are similar to each other despite the fact that their control climates are very different (as discussed in section 4a). The percentage change in precipitation in the four models is almost linearly proportional to the magnitude of surface temperature change, corresponding to an increase of 2.5% per degree, consistent with that found in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) coupled GCMs simulations by Held and Soden (2006). Water vapor in the lower troposphere increases at close to 8% K⁻¹, somewhat greater than the 7.5% K⁻¹ obtained by Held and Soden in realistic coupled GCMs. We should, therefore, see the responses in hydrological cycle and circulation in these aquaplanet simulations that Held and Soden describe as resulting

---

**Table 2.** The sensitivity experiments with changing heat capacity ($C_s$) of the mixed layer. Corresponding depths are also indicated in meter unit.

<table>
<thead>
<tr>
<th>Expt</th>
<th>Heat capacity (J kg⁻¹ K⁻¹)</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>$0.5 \times 10^7$</td>
<td>1.1</td>
</tr>
<tr>
<td>D2</td>
<td>$2.5 \times 10^7$</td>
<td>5.8</td>
</tr>
<tr>
<td>D3 (control)</td>
<td>$5.0 \times 10^7$</td>
<td>11.6</td>
</tr>
<tr>
<td>D4</td>
<td>$25.0 \times 10^7$</td>
<td>58.0</td>
</tr>
</tbody>
</table>

---

**Fig. 10.** Zonal-mean distributions of (a) surface temperature, (b) precipitation rate, (c) planetary albedo, and (d) surface evaporation for 30°S–30°N latitudes due to the changes in the ocean mixed layer depth in the NSIPP model. The mixed layer depths are $0.5 \times 10^7$ in D1, $2.5 \times 10^7$ in D2, $5.0 \times 10^7$ in D3 (control), and $25.0 \times 10^7$ J kg⁻¹ K⁻¹ in D4.
from these relations between water vapor, precipitation, and temperature.

Figure 14 compares the latitude–height distributions of the zonal-mean temperature change. All models simulate an overall temperature increase in the troposphere and decrease in the stratosphere. Uptropospheric temperature increases are much larger than surface temperature increases in all models. Changes in the Hadley circulation, vertical stability in the tropics, and lower-tropospheric moisture at doubled CO2 are compared in Fig. 15. The models are in good agreement in simulating a weakening of the Hadley cell, with weaker ascending motions (Fig. 15a) and meridional winds (Fig. 15c) in low latitudes. Equatorial ascending motion decreases in most of the troposphere, except in higher altitudes where changes in upward motion reflect the elevation of the tropopause in the warmer climate. As expected for warmer climates, vertical stability and moisture increase in all the models (Figs. 15b,d). The ITCZ precipitation increases by 0.4–0.8 mm day\(^{-1}\), depending on the model (not shown), implying increased adiabatic cooling in the ITCZ and enhanced low-level moisture transport toward the equator. Since the Hadley circulation weakens in all the models, the enhanced moisture transport is due to the increase in low-level moisture, not stronger trades, and the increased adiabatic cooling results from the increased stability, not an increase in rising motion. Overall, these results are consistent with the robust changes in the coupled GCMs described in Held and Soden (2006).

We compare the changes in the cloud radiative forcings between the control and 2×CO2 experiments (2×CO2 − Control) in Fig. 16. Each value is divided by the surface temperature increase and indicates how much the cloud contributes to amplify or moderate the direct radiative forcing from doubling CO2 (Cess et al. 1990, 1996). The models exhibit relatively good agreement in the projected cloud forcing changes, with negative signs in the longwave radiation and positive signs in the shortwave radiation, as total cloud cover in all the models decreases by doubling CO2. Also shown in the figure are the actual changes in global-mean high, middle, and low cloud amounts for the three models for which we had this diagnostic. Except for an increase in high cloud amounts in GEOS-5, all clouds in all models
decrease with the warming climate. The high cloud increase in GEOS-5 is associated with optically thin clouds, and this model still produces the largest decrease in longwave cloud forcing.

As expected, the models exhibit large cancellations between the negative longwave and positive shortwave changes in cloud radiative forcing. The SNU model is rather exceptional in this, with very small changes both in the longwave and shortwave radiations by cloud. The models are not in agreement in their change in net cloud radiative forcing. In three of the models the shortwave effect dominates. In the GFDL model, the longwave dominates. In summary, responses that depend on changes in water vapor are relatively robust across the models, while changes controlled by the cloud field are more divergent.

5. Summary and concluding remarks

We define a benchmark computation for AGCMs in an idealized framework. In this framework, all physical parameterizations and dynamical assumptions of the models are retained, but the models are run over a flat, homogeneous, saturated ocean surface with a small, but nonzero, heat capacity. Most of the experimental setup follows closely the APE project proposed by NH00a, but a simple mixed layer ocean is added to enable air-sea coupling processes and allow for a crude, but more
FIG. 13. Scatterplots of the percentage change in the global-mean (a) column-integrated moisture and (b) precipitation versus the global-mean change in surface temperature at doubling CO₂ concentration in the four AGCMs. Solid lines indicate the linear fit of the increase rate in column-integrated water vapor (8.0% K⁻¹), and the dashed line in (b) indicates the linear fit of the precipitation increase at a rate of 2.5% K⁻¹.

FIG. 14. Zonal-mean distributions of temperature change (2 × CO₂ − Control). Units are kelvin.
straightforward, analysis of climate sensitivity than with realistic coupled GCMs.

This simple framework has been applied to four different AGCMs. Although the idealized simulations are characterized by very strong jet streams and unrealistically strong ITCZ precipitation, the circulation remains sufficiently realistic, in our view, to justify this study. Comparison of the model results shows that the test differentiates clearly among the different models examined here. For example, the AGCMs exhibit a broad variation in the simulations of the Hadley circulation and ITCZ intensity. Zonal-mean planetary albedo and cloud radiative forcings are also quite different in the four models, primarily owing to differences in simulated cloud distributions. The results suggest that the simulated subtropical cloud distributions are especially important in characterizing different models, determining planetary albedos, global-mean temperatures, and precipitation. The degree to which low-level clouds control the climate in this benchmark was explored in the NSIPP AGCM, which in the control runs was an outlier, producing by far the warmest climate. By changing the cloud parameterization in this model to produce more low-level subtropical cloudiness, its climate is more in agreement with those in the other models.

An analysis of the tropical transients in the control simulations showed that the amplitudes of the subseasonal variances and the eastward propagation speeds are quite different among the models. Two AGCMs (NSIPP and SNU) are characterized by small variances and fast propagations of the convectively coupled equa-

Fig. 15. Vertical profiles of the change at doubling CO₂ concentration (2 × CO₂ − Control) in (a) vertical motion (hPa day⁻¹) and (b) potential temperature (K) averaged over the tropics (5°S–5°N). (c) The change in latitudinal distributions of the zonal-mean meridional wind (m s⁻¹) averaged between 950 and 750 hPa; (d) the change in the vertically integrated moisture (mm) between 950 and 750 hPa. Only the symmetric components about the equator are given in (c) and (d).
Equatorial waves in the subseasonal time scales. The other two AGCMs (GFDL and GEOS-5) are characterized by larger subseasonal variances and much slower propagation of equatorial waves. It was suggested that these differences should be largely due to the differences in the convection inhibition. The shallower heating structure from the grid-scale condensation scheme that results when convection is inhibited seems to explain the slower propagation of the convectively coupled equatorial waves, a result that is consistent with expectations from previous studies (Lau and Peng 1987; Chang and Lim 1988) and also consistent with observations (Lin et al. 2004).

This mixed layer framework introduces an important parameter, the oceanic heat capacity, which can affect the mean climate and variability. The AGCMs tend to simulate stronger zonal-mean circulation and ITCZs as the mixed layer depth increases. In the shallower mixed layer ocean, overall cloud amount is greater, and the model maintains a colder climate with larger planetary albedo. Increase of cloud amount with decreasing mixed layer thickness seems to be caused by increased variance of convective activity due to the increased SST variability. However, the variance increases are selective in time scale. The subseasonal variances are generally increased as the mixed layer depth increases, but these signals seem to be dependent on the zonal-mean state, and also dependent on the model formulation.

The moist physics sensitivity to increased greenhouse forcing was also evaluated. When CO₂ is doubled, the models tend to produce different amplitudes of global mean responses in surface temperature and precipitation. On the other hand, the percentage changes in global-mean precipitation are closely proportional to the global-mean surface temperature change, increasing at a rate of 2.5% per degree of surface temperature increase. This value is quite close to the one (2.2% K⁻¹) found in the IPCC AR4 runs by Held and Soden (2006). In the warmer climate, the models consistently simulate an enhanced water vapor transport toward the equator, a reduction of dry stability, a weakening of the Hadley circulation, and a reduction in cloudiness and tropospheric relative humidity. The net changes in cloud radiative forcing in the four models, on the other hand, disagree in sign and amplitude.

This study has explored the applicability of the benchmark experiment to several climate studies. Although many aspects of the model simulations and their differences have not been fully explained, we feel that the mixed layer aquaplanet framework is a useful complement to the more familiar fixed SST aquaplanet configuration and a simple way of obtaining insights into the behavior of GCMs in climate sensitivity studies.

Acknowledgments. We thank David Straus and three anonymous reviewers whose comments were very constructive in improving the original manuscript. We also thank Julio Bacmeister and In-Sun Song for their helpful discussions and suggestions to this study. This study was supported by NASA Earth Science Enterprise’s Modeling, Analysis, and Prediction (MAP) program. In-Sik Kang and Daehyun Kim have been supported by the Korea Meteorological Administration Research and Development Program under Grant CATER_2006-4206 and the BK21 program.

REFERENCES


Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. Proc. ECMWF Workshop on Convection in Large-Scale Models, Reading, United Kingdom, ECMWF, 297–316.


Wang, W., and M. E. Schlesinger, 1999: The dependence on convection parameterization of the tropical intraseasonal oscillation simulated by the UIUC 11-layer atmospheric GCM. J. Climate, 12, 1424–1457.


