Assimilation of Satellite Cloud Data into the GMAO Finite-Volume Data Assimilation System Using a Parameter Estimation Method. Part I: Motivation and Algorithm Description

PETER M. NORRIS

Goddard Earth Science and Technology Center, University of Maryland, Baltimore County, Baltimore, and Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

ARLINDO M. DA SILVA

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 25 January 2006, in final form 22 June 2006)

ABSTRACT

General circulation models are unable to resolve subgrid-scale moisture variability and associated cloudiness and so must parameterize grid-scale cloud properties. This typically involves various empirical assumptions and a failure to capture the full range (synoptic, geographic, diurnal) of the subgrid-scale variability. A variational parameter estimation technique is employed to adjust empirical model cloud parameters in both space and time, in order to better represent assimilated International Satellite Cloud Climatology Project (ISCCP) cloud fraction and optical depth and Special Sensor Microwave Imager (SSM/I) liquid water path. The value of these adjustments is verified by much improved cloud radiative forcing and persistent improvement in cloud fraction forecasts.

1. Introduction

It is widely recognized that clouds play an essential role in moderating climate and are therefore an important feature to accurately model in GCMs. This is not a simple task, owing to the mismatch in scales between the typical GCM grid box (~100 km) and the smaller scales at which clouds form and evolve and due to the complexity of cloud microphysical processes and their interaction with cloud dynamics and radiative transfer. As a result, clouds continue to represent a major source of uncertainty in GCM studies of future climate. Part of the historical problem has been that, in the face of these complexity and scale mismatch problems, simple empirical cloud parameterizations have been devised and then just tuned to give reasonable top-of-atmosphere radiative forcing in a globally or zonally averaged sense. Sufficient attention has not generally been given to the validation of the predicted cloud properties. In the NWP community even less attention has historically been paid to predicted cloud properties, due to the slower time scales associated with cloud-induced radiative heating rates compared with the forecast duration. Nevertheless, clouds do have an important societal impact from day to day, in terms of their effects on diurnal temperature range and sunlight exposure. Furthermore, since NWP and GCM models have become more merged, typically sharing the same physics, advances in cloud parameterization in either climate or weather studies ought to benefit the other. Now, with the availability of a large and fast growing amount of satellite cloud data [e.g., International Satellite Cloud Climatology Project (ISCCP), Moderate Resolution Imaging Spectroradiometer (MODIS)] there is a need for a careful validation of model cloud properties, and the exciting possibility of directly assimilating such cloud data into NWP models. In this paper we describe the motivation for and development of a new cloud assimilation scheme based on parameter estimation techniques. The new scheme is called the Cloud Parameter Estimation System (hereafter CPES).

There are at least three different approaches to the assimilation of satellite cloud observations.

Corresponding author address: Peter M. Norris, Global Modeling and Assimilation Office, NASA GSFC, Code 610.1, Greenbelt, MD 20771.
E-mail: peter.m.norris@nasa.gov

DOI: 10.1175/2006JAS2046.1

© 2007 American Meteorological Society

1) The direct assimilation of cloudy radiances. This requires a forward radiative transfer model that explicitly accounts for the presence of clouds. Cloud liquid water and ice content may be included as control variables. This approach is the most direct extension of the 3D and 4D variational methods used in assimilating clear-sky fluxes to determine temperature and moisture profiles. This approach is still very new, but progress is being made (Janisková et al. 2002; Greenwald et al. 2004; Chevallier et al. 2004). The challenge is not only to have a forward model that accurately accounts for cloud optical properties, but also that passive radiative observations still only partially constrain the cloud properties (especially in multilayer cloud schemes, which are common). Janisková et al. conduct some preliminary experiments to investigate the assimilation of surface and top-of-atmosphere radiation observations using a one-dimensional variational data assimilation (1DVAR) approach in which linearized cloud diagnostic parameterizations and radiative forward models are used to find increments to the control variables (temperature and humidity profiles and surface pressure) that minimize radiative flux differences between the forward model and cloudy radiance observations. Using both simulated observations and actual ARM site observations, the method shows some positive skill in improving the temperature and humidity profiles and cloud cover, although the following two areas of concern are identified: (i) the method has difficulty in triggering clouds in cases where the observations show them but the background is clear—such triggering only succeeds when the background is already close to saturation and is enabled by a positive dependence of the radiative fluxes on humidity—and (ii) it is difficult to adjust the model state to produce the observed radiative fluxes when the temperature and humidity profiles are very different from the actual profiles that generate the synthetic or observed radiances. It is shown that the use of additional vertical profile observations (of temperature, humidity, or cloud condensate) can help to retrieve more accurate profiles, as does the use of both *surface* and *top-of-atmosphere* radiative flux observations. Chevallier et al. (2004) conclude that there is potential benefit in the direct assimilation of cloud-affected satellite radiances at 4.5, 6.3, and 14.3 μm but not in other more directly cloud-affected channels (e.g., 11 μm) owing to the presence of strong nonlinearities in the forward model. Greenwald et al. (2004) work at smaller cloud-resolving scales (2–5 km) and find that satellite radiances in visible and infrared windows do contain potentially useful information about cloud microphysics, particularly at solar wavelengths.
2) A second simpler approach (Macpherson et al. 1996; Lipton and Modica 1999) is to use cloud observations to generate pseudo-RH data consistent with model diagnostic cloud parameterizations (i.e., such that the model would produce the observed cloud fraction using these pseudo-RH values). These pseudo-RH values are then assimilated into the system in addition to RH observations from other sources (e.g., radiosondes or satellite moisture soundings). Alternatively, the cloud observations can be used to correct collocated RH observations, consistent with the model diagnostic cloud parameterizations. In either case, the model cloud fraction parameterization is never modified, but taken as correctly representative of nature, with the cloud observations being used rather to imply information about the moisture field. Macpherson et al. (1996) use cloud information from satellites and surface reports, together with various semiempirical rules for assigning the vertical profile of cloudiness, to produce an "observed" cloudiness on the model grid. This layer cloud fraction is then used to produce pseudo-RH values that are used to nudge the model moisture fields in the moisture analysis. Lipton and Modica (1999) use a similar strategy to assimilate cloud data into the Mesoscale Model version 4 (MM4) model regional simulations.
3) A third approach, which we will take, is motivated by the rather empirical nature of model cloud parameterizations. This empiricism is due to insufficiently resolved cloud dynamic and thermodynamic processes, and the incredible complexity of clouds on a microphysical level. In view of this inherent empiricism, this approach seeks to modify the cloud parameterizations by using cloud-affected observations to estimate empirical parameters in the parameterizations. An early example of this approach is the work of Wu and Smith (1992) who optimally adjusted the dependence of the grid box cloud fraction on relative humidity such that the model produced an outgoing longwave radiation (OLR) in each column close to Earth Radiation Budget Experiment (ERBE) observations. The improvement in model OLR was confined not just to assimilation time but was retained in forecast mode. Another recent example is the work of Chaboureau and Pinty (2006), who use satellite interchannel brightness temperature differences (BTDs) to constrain the ice to snow autoconversion threshold in regional model

simulations over Brazil, leading to an improved diurnal cycle in cirrus coverage and upper-tropospheric humidity. Our approach is similar to Wu and Smith (1992) except that now, with considerably greater availability of high-quality cloud property observations, we assimilate the cloud properties themselves (not the radiances) and then validate our assimilation against top-of-atmosphere (and surface) radiation measurements. The advantage of this approach is that it allows us to isolate problems or needed tuning in different elements of the cloud–radiation system. We also estimate cloud parameters in a spatially and temporally varying manner to better allow for synoptic and geographic forcings on subgrid-scale moisture variability. This variability cannot be resolved by our model but directly influences cloud properties.

In section 2 we describe the NASA Global Modeling and Assimilation Office (GMAO) finite volume Data Assimilation System (fvDAS), otherwise known as the Goddard Earth Observing System 4 (GEOS-4) DAS. This is the control system from which we test extensions to enable cloud data assimilation. In section 3 we discuss the need for the assimilation of cloud data within GEOS-4. We also give a brief overview of our current cloud assimilation system and describe the various experiments that led us to arrive at that system. Section 4 then presents the technical details of the system. Section 5 draws some conclusions as they relate to the motivation, design, and implementation of the system. A follow-up paper (Part II) will present detailed results of seasonal assimilation experiments using the system and analyze its influence on the quality of the GEOS-4 DAS analyses.

2. Description of the GEOS-4 data assimilation system

The GEOS-4 DAS combines a forecast model, comprising a finite-volume dynamical core (Lin 2004) and the Community Climate Model version 3 (CCM3) physics, with the GMAO's physical space statistical analysis system (PSAS). A 6-h forecast provides a background state, which is merged with various sources of data (moisture, temperature, winds, and surface pressure from meteorology stations, rawindsondes, aircraft, TOVS retrievals, etc.) by PSAS to produce an analysis that is then used to initialize the next 6-h forecast. In this study the model resolution is 1.25° in longitude by 1° in latitude with 55 layers in the vertical and a model top at 0.01 hPa.

The GEOS-4 DAS currently uses the CCM3 physics package including diagnostic parameterizations of both cloud fraction and condensate amount (Kiehl et al. 1996). The cloud fraction parameterization is a generalization of Slingo (1987). The core of this scheme is a quadratic variation of cloud fraction, f, with relative humidity, RH, above a critical value, RH_0, approaching complete cloud cover at 100% RH:

$$f = \begin{cases} 0, & \mathrm{RH} \le \mathrm{RH}_0 \\ \left(\dfrac{\mathrm{RH} - \mathrm{RH}_0}{100\% - \mathrm{RH}_0}\right)^2, & \mathrm{RH}_0 < \mathrm{RH} < 100\% \\ 1, & \mathrm{RH} \ge 100\%. \end{cases} \quad (1)$$

The tuned version used in the GEOS-4 sets RH_0 = 87%. For mid–high clouds (<750 hPa) RH_0 is increased with a positive Brunt–Väisälä frequency to account for reduced subgrid-scale variability under stable conditions. For low clouds (≥750 hPa) RH_0 is reduced to 77% over snow-free land to account for the increased cloud condensation nuclei (CCN) concentrations present there. Two other modulations exist for low clouds: a reduction in f under subsidence conditions with complete removal for $\omega \ge 50$ hPa day^{-1} and the addition of stratus associated with significant low-level inversions. A convective cloud fraction is also calculated based on the mass flux from the convection routines. The cloud fraction, f, discussed above is for the nonconvective region. Also, the relative humidity, RH, discussed above is for the nonconvective fraction of the grid cell assuming saturation in the convective fraction.

The in-cloud condensate water content, ρ_c, is also diagnostic: an exponential decay with height above surface with a scale height dependent on the local column integrated water vapor amount—total precipitable water (TPW)—in kg m^{-2}. Namely,

$$\rho_c = \rho_{cs} \exp\{-z/[h_{c0} \ln(1 + \mathrm{TPW})]\}. \quad (2)$$

The standard CCM3 uses globally uniform values ρ_{cs} = 0.21 g m^{-3} and h_{c0} = 700 m. The in-cloud condensate water path, $\Gamma_c \equiv \int \rho_c \, dz$, for a grid box is just the vertical integral of ρ_c within the grid box.

The treatment of cloudiness in the CCM3 radiation code is an important aspect of the overall cloud parameterization system. Both longwave and shortwave treatments of cloudiness are simplistic. For longwave cloudiness, clouds are assumed to be randomly overlapped, and the cloud fraction, f, for each layer is replaced by an effective cloud fraction that is the product of f and a cloud emissivity that depends on the in-cloud conden-

sate water path, Γ_c, and on details of the size and fractional content of ice crystals in the cloud. For shortwave cloudiness a delta-Eddington approach is used with a cloud extinction optical depth of the form

$$\tau_\lambda^{gb} = C_\lambda \Gamma_c f \sqrt{f}, \tag{3}$$

where λ is the wavelength and C_λ is a factor depending on the wavelength, the fractional split of ice and liquid water, and the effective radii of hydrometeors in each of those phases. The superscript *gb* denotes that this optical depth is used by the model for the whole grid box, not just the cloudy portion. Thus in this very empirical fashion the code places all the effects of partial grid box cloudiness and by implication cloud overlap into the term $f\sqrt{f}$ depending on the cloud fraction f. In essence a grid column containing partially cloudy layers has been converted to a column of overcast cloud layers for simple treatment by the shortwave code, and the actual effect of partial cloudiness and cloud overlap in the column has been parameterized by the use of the multiplicative factor $f\sqrt{f}$ within each grid box. This assumption will be discussed further in section 4b(3).

3. Motivation and design of cloud assimilation system

a. Motivation

Figure 1 shows the state of the GEOS-4 DAS cloud cover prior to any cloud assimilation attempts. In the midlatitudes, a pure climate forecast (i.e., no assimilation) is closest to the observations, while the analysis significantly underpredicts midlatitude cloud fraction (by up to 30% points) both in the low and mid–high bands. The GEOS-4 DAS background state (a 6-h forecast from the analysis) lies in between the analysis and the climate. The climate mode result is closest to the observations because some limited tuning of the climate model to cloud observations has previously been made at the National Center for Atmospheric Research (NCAR) and GMAO. However, this tuning becomes inappropriate when the assimilation corrects the moisture and temperature fields in the model. Other details to note are the massive overprediction of low cloud fraction in winter high latitudes by both the climate run and the analysis, and the somewhat more complicated comparison between climate run, analysis, and data in the Tropics.

The overall poor performance of the model both in climate mode and even more so in assimilation mode was the motivation for our efforts to assimilate observational cloud fraction data into the GEOS-4 DAS.

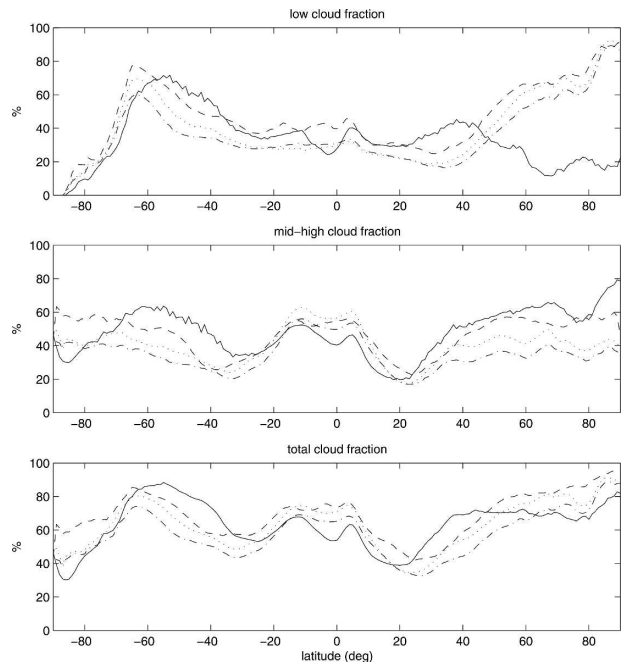

FIG. 1. Zonally averaged (top) low (>750 hPa), (middle) mid–high (30–750 hPa), and (bottom) total cloud fraction for January 2001: ISCCP-DX–derived observations (solid), a GEOS-4 climate (no assimilation) run (dashed), and the analysis (dash–dot) and background (dotted) of a GEOS-4 DAS (standard assimilation) run. Model output is sampled only where observational data are available—this also applies to all subsequent figures.

b. Overview and design of the cloud parameter estimation system

From the beginning we attempted to treat the cloud assimilation problem with a parameter estimation approach, largely because the GEOS-4 cloud parameterizations, which estimate cloud properties based on the modeled moisture field, are rather crude and empirical, while the moisture field itself can be constrained by moisture observations during the assimilation process. Among other things, we regarded the use of globally constant parameters, such as the critical relative humidity for large-scale cloud formation, as unlikely to be correct given that the key goal of the parameterizations was to represent at grid scale the behavior of complex subgrid-scale cloud processes known to have geographical, synoptic, and diurnal dependencies. Our best overall results to date have been produced by a three-stage parameter estimation procedure assimilating in turn 1) ISCCP cloud fraction, 2) SSM/I cloud liquid water path, and 3) ISCCP optical depth. These stages will be described further below and explained in detail in section 4. In the remainder of this section we wish to present, in tutorial fashion, key results from our cloud

assimilation investigations that motivated this final three-stage approach.[1]

The first stage is assimilation of cloud fraction. As detailed in section 4b(1), we replaced the quadratic dependence of the cloud fraction, f, on the relative humidity, RH, namely Eq. (1), with a more general two-parameter S-shaped dependence and then estimated the optimal values of these two parameters as a function of space and time. Figure 2 shows that the scheme produces the desired results; namely, assimilation of ISCCP-DX low and mid–high cloud fractions yields cloud fractions very close to the assimilated observations. Small differences exist due to the influence of the background terms.

However, to validate the usefulness of our approach we compare the resultant changes in model top-of-atmosphere cloud radiative forcing with independent Clouds and the Earth's Radiant Energy System (CERES) observations (Wielicki et al. 1996; Wong et al. 2000). Figure 3 shows the results from a CPES run with single-stage assimilation of ISCCP-DX cloud fraction alone. Clearly the longwave cloud forcing (LWCF) in the midlatitudes is significantly improved by the increased midlatitude cloud fraction in the assimilation run. However, that increased cloud fraction also produces excessive shortwave cloud forcing (SWCF).

This is confirmed by Fig. 4, which shows that the grid-box-averaged cloud[2] optical depth has risen to un-

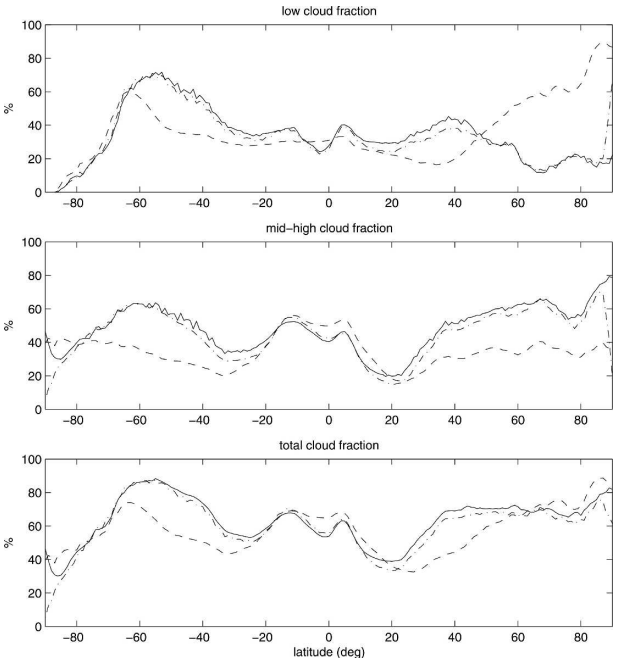

FIG. 2. Zonally averaged (top) low (>750 hPa), (middle) mid–high (30–750 hPa), and (bottom) total cloud fraction for January 2001: ISCCP-DX–derived observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived low and mid–high cloud fraction (dash–dot).

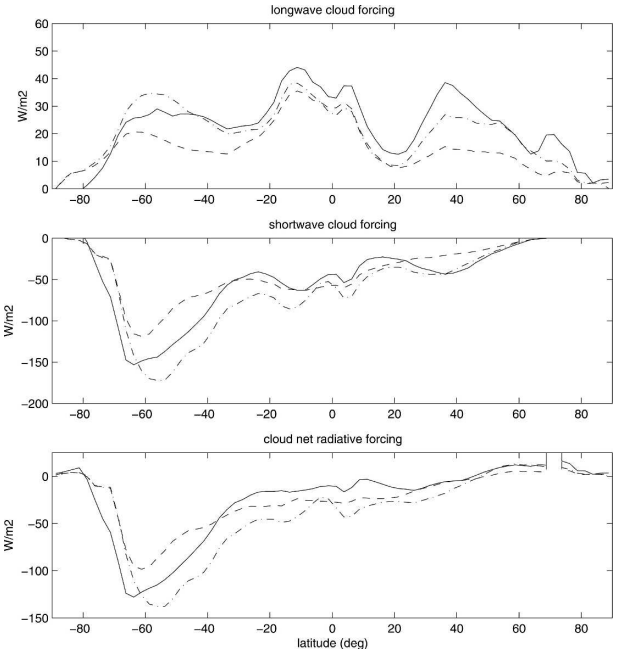

FIG. 3. Zonally averaged (top) longwave, (middle) shortwave, and (bottom) net cloud radiative forcing for January 2001: CERES ES4 observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived cloud fraction *only* (dash–dot).

[1] Note that the use of three separate minimization stages, rather than a single combined minimization, does admit the possibility, for example, that the last stage will provide a worse fit to the observations used in the first and second stages. However, there were practical and theoretical reasons for approaching the problem in three stages. First, given that the model physics was being revised during the course of this investigation, we could not justify developing adjoints for the physics, opting instead for a nongradient-based simplex minimization, which works well for a small number of control variables. Second, parameter identifiability can be a big problem. While the presence of a background term (see section 4) serves the purpose of regularizing this problem, it does so indirectly through the specification of error covariances. While it is true that the last stage could worsen the fit to observations in the first stage, a similar difficulty still exists in the case of a joint estimation subject to multiple observational constraints (e.g., fitting the cloud water observation could in principle impair our ability to fit optical depth data.) We have therefore chosen this safe and minimalistic three-stage approach.

[2] The use of the term "grid-box-averaged cloud" property refers the average over a whole grid box of a cloud property with clear portions contributing zero in the sum. Thus, for example, a half cloudy grid box has a grid-box-averaged cloud water path half that of the in-cloud value. The specific use of the word cloud in "grid-box-averaged cloud" property is to exclude contributions to the property (e.g., optical depth) from noncloud sources (e.g., water vapor or aerosols).

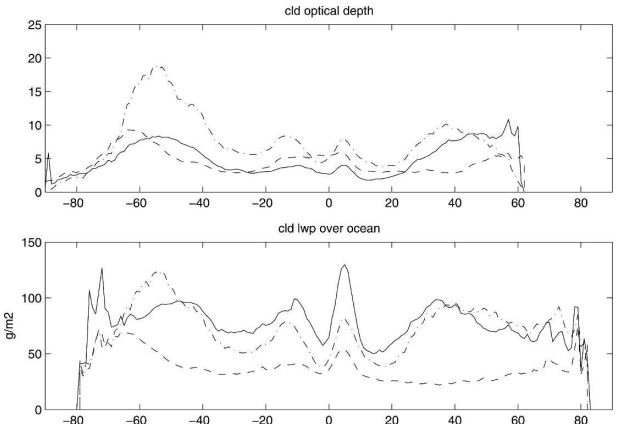

FIG. 4. (top) Zonally averaged grid box cloud optical depth and (bottom) cloud liquid water path over ocean for January 2001: ISCCP-DX–derived optical depth and SSM/I liquid water path observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived cloud fraction *only* (dash–dot).

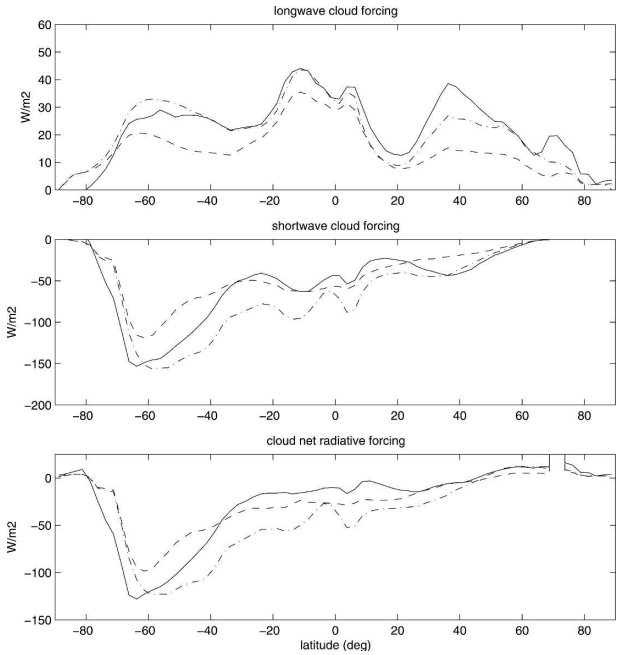

FIG. 5. Zonally averaged (top) longwave, (middle) shortwave, and (bottom) net cloud radiative forcing for January 2001: CERES ES4 observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived cloud fraction and SSM/I liquid water path (dash–dot).

realistic levels as the cloud fraction has increased. Clearly some sort of adjustment of cloud optical properties is needed in concert with the cloud fraction assimilation.

Figure 4 also shows that the cloud fraction assimilation has increased the cloud liquid water path (CLWP) to generally more realistic levels (as judged by SSM/I retrievals) than in the control, which badly underestimated CLWP. Note that CLWP is still underestimated in the Tropics and subtropics and is overestimated in southern high latitudes. Interestingly, Fig. 3 shows that in these same regions longwave cloud forcing is also underestimated and overestimated respectively. This suggests that by also assimilating CLWP, the longwave (LW) cloud forcing might be further improved. This observation motivated (Fig. 5) a new CPES run with an extra SSM/I CLWP assimilation stage after the cloud fraction assimilation stage. As detailed in section 4b(2), this second stage involved estimating an optimal ρ_{cs} in (2) as a function of space and time. With this CLWP assimilation the tropical LW cloud forcing is much improved (Fig. 5). The SWCF remains excessive owing to excessive gridbox-averaged cloud optical depth (not shown).

Evidently, we need to find a way to improve the SWCF without undoing the gains made by cloud fraction and liquid water path assimilations. Examination of the treatment of clouds in the control shortwave code revealed a very empirical (and demonstrably unrealistic) dependence on cloud fraction. By introducing a further parameter to the shortwave code to improve the treatment of partial cloudiness and cloud overlap [see section 4b(3) for the details], it is possible to correct this problem by assimilating optical depth observations (from ISCCP) in a third assimilation stage. Figure 6 shows that this additional stage reduces the excessive tropical and midlatitude SWCF. The resultant zonally averaged SWCF is comparable to that of the control run, although full geographic plots (to be presented in Part II) show that the qualitative spatial distribution of SWCF is much improved in the new system.

While the grid-box-averaged cloud optical depth is well constrained by the third stage (Fig. 7), it is important to remember that this observation is a retrieval with associated bias and statistical errors. Further comparison against MODIS optical depth ought to clarify these errors and is underway.

Nevertheless, the net results of the three-stage assimilation (ISCCP cloud fraction, SSM/I CLWP, and ISCCP optical depth) are vastly improved cloud fraction, CLWP, and optical depth, and *independently* validated longwave cloud forcing, with a more or less neutral impact on shortwave cloud forcing.

Note that our analysis has mainly focused on the midlatitudes where the largest problems existed in the control system. It is interesting, however, to analyze the tropical behavior further. Notice that observed cloudiness has local tropical maxima either side of the equa-

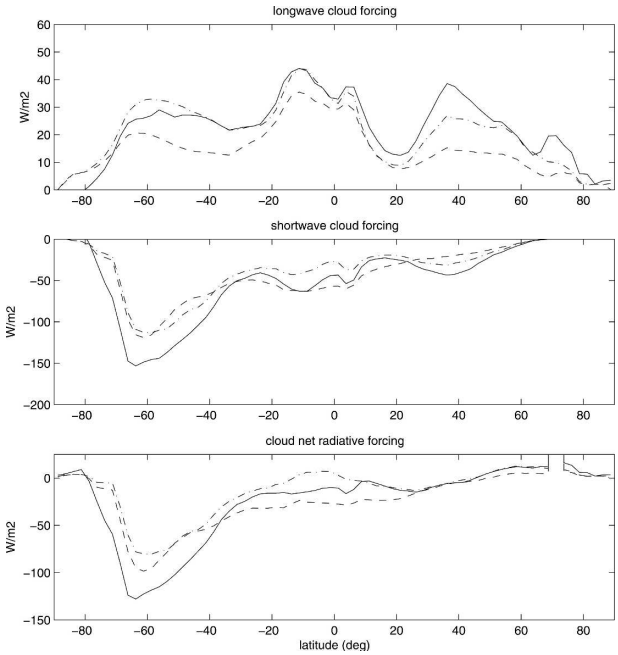

FIG. 6. Zonally averaged (top) longwave, (middle) shortwave, and (bottom) net cloud radiative forcing for January 2001: CERES ES4 observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived cloud fraction and optical depth and SSM/I liquid water path (dash–dot).

tor (Fig. 2) in both the low and mid–high bands. These peaks were underpredicted in the control low band and somewhat overpredicted in the mid–high band. The cloud fraction assimilation corrects these biases, effectively shifting some cloudiness down in the model column. However, the total cloud cover is dominated by the behavior of the mid–high band in both the control and cloud assimilation cases since mid–high cloudiness "obscures" low cloud cover. In terms of formal "cloud cover," this obscuration occurs even if the mid–high cloud is thin. However, the case is very different when we look at the radiative cloud forcings (Fig. 3). The SWCF and even the LWCF are apparently dominated by the low cloud cover increases at the peaks, indicating that much of the mid–high cloud must be thin, so that even the LWCF signal due to the increased low cloud is transmitted. These observations further emphasize the complexity of the problem of cloud data assimilation and the need to assimilate multiple cloud observables to produce desirable results.

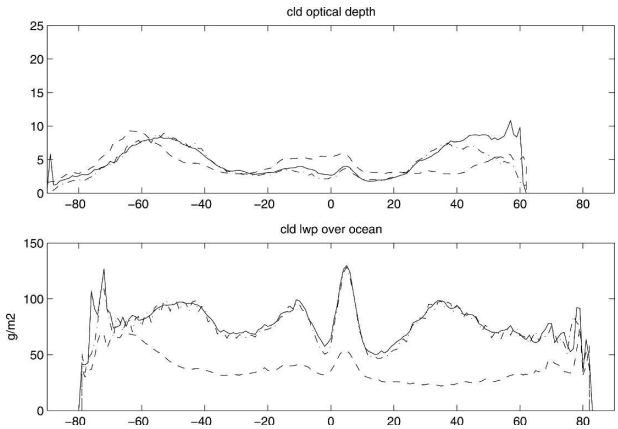

FIG. 7. (top) Zonally averaged gridbox cloud optical depth and (bottom) cloud liquid water path over ocean for January 2001: ISCCP-DX–derived optical depth and SSM/I liquid water path observations (solid), GEOS-4 DAS control run (dashed), and new cloud parameter estimation system with assimilation of ISCCP-DX–derived cloud fraction and optical depth and SSM/I liquid water path (dash–dot).

c. *Forecast skill*

In addition to the expected improvement in cloud fraction in the analysis, we would also like to see if this improvement is maintained in a forecast (i.e., pure simulation mode) if we retain the CPES parameters valid at the initial time of the forecast. To examine this matter, we conducted a series of 5-day forecasts, initialized 3 days apart, throughout the March–May 2001 season. The control run uses the analyses from a control GEOS-4 DAS run (no cloud assimilation) as initial conditions and the standard CCM3 cloud parameterizations in the forecasts. The experiment run uses the analyses from a CPES GEOS-4 DAS run including the cloud parameter restart files created continuously throughout that run. Each experiment run forecast loads the CPES cloud parameters valid at its initialization and uses them in the CPES versions of the model cloud parameterizations during the subsequent 5-day forecast.

There is no significant change in the standard measures of model forecast skill (anomaly correlations for 500-hPa height, sea level pressure, and 200-hPa zonal wind) between the two runs (not shown), but the forecast of cloud fraction is significantly improved with a reduction of up to 25% in the rms error between the two runs (Fig. 8).

This reduction is maintained throughout the 5-day forecasts, showing that the newly parameterized CPES model cloud fraction is a significant improvement over the old scheme, even in a climate sense.

Notice from Fig. 1 (bottom) that in the midlatitudes the zonal monthly mean total cloud fraction error with respect to ISCCP observations decreases with forecast lead, that is, from the analysis (0 h) to the background (6 h) to the climate (∞). Yet we have just seen from

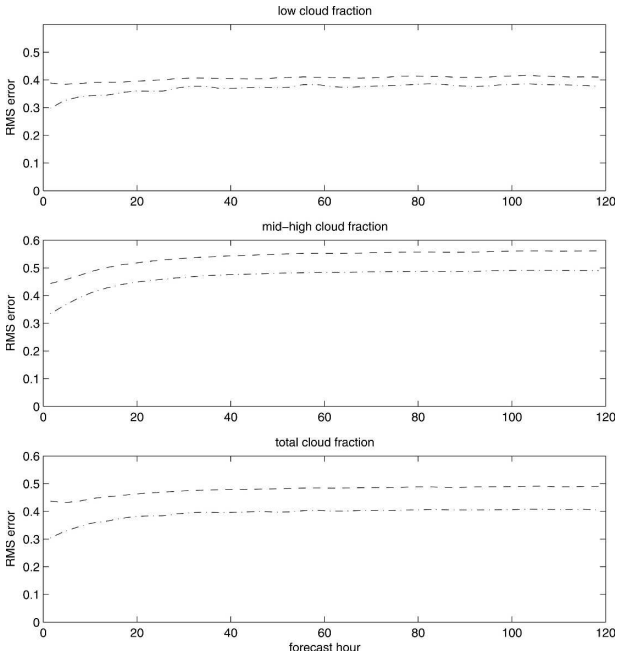

FIG. 8. Global rms (top) low, (middle) mid–high, and (bottom) total cloud fraction error (with respect to ISCCP observations) for a series of 5-day forecasts initialized every 3 days at 0000 UTC throughout March–May 2001: forecast from GEOS-4 DAS control run (dashed), forecast from GEOS-4 DAS CPES base run with cloud fraction using parameters from the CPES cloud parameter restarts of the base run (dash–dot).

Fig. 8 (bottom) that the global rms error tends to slightly increase with time, even in the control case in which the initial condition clouds have not been tuned to match observations. This is not inconsistent for several reasons. First, rms error includes not only the mean error, but the rms of the departures from the mean error. This latter error will increase with time as the meteorological fields in the forecast become increasingly decorrelated with the true atmospheric state. Second, the global rms will have a large tropical component, since the rms has been calculated with area-weighted means. From Fig. 1 we see that the tropical zonal-mean total cloud fraction error increases from analysis to climate, opposite to the midlatitude behavior. But even if we ignored this second reason, the magnitude of the rms total cloud fraction error from Fig. 8 is almost 0.5, which dominates over the corresponding zonal mean error that never exceeds 0.3 even at the latitudes where it is largest (see Fig. 1).

From this analysis, we conclude that the cloud fraction forecast deteriorates with time due to the loss of atmospheric predictability, and yet the forecast that uses CPES-tuned parameters performs better because it is less biased in its production of cloud cover given the meteorological variables.

4. The cloud parameter estimation system in detail

In this section we describe the details of the new cloud parameter estimation system. First we discuss the data that is assimilated. Then we discuss modifications made for cloud assimilation and the parameter estimation process as a whole.

a. Assimilated cloud data

This study will illustrate the assimilation of ISCCP cloud fraction and optical depth and SSM/I cloud liquid water path. While ISCCP is not a real-time source of cloud data, it can be used for reanalysis projects and for a general proof of concept of the assimilation method. Use of MODIS cloud products is currently being investigated.

1) INTERNATIONAL SATELLITE CLOUD CLIMATOLOGY PROJECT CLOUD DATA

ISCCP cloud products are discussed in detail in Rossow et al. (1996). The high-resolution DX data product—namely 3-hourly 25-km sampled cloud data from all available satellites—was chosen to enable automatic use of the cloud assimilation system with the variety of horizontal resolutions used by the GEOS-4 DAS. These studies use the $1.25° \times 1°$ resolution. The assimilation system collects the DX pixels into the longitude/latitude grid boxes of the model. For each grid box it selects pixels from only a single satellite, the one with the best viewing angle, but always preferring geostationary satellites equatorward of 55° and polar satellites poleward of 55°. NOAA-A is always preferred to NOAA-M. This is the same methodology used by ISCCP to produce D1 data from DX data. If no pixels are available for a model grid box within the particular 3-h window, or if the optimal satellite views the grid box from greater than a 72.5° zenith angle, then the grid box cloud observations are marked as missing. Missing data are typically found in parts of the Indian Ocean and Subcontinent and in some polar regions.

Although ISCCP-DX data have a continuous cloud top height, we evaluate a cloud fraction for only two pressure bands (low and mid–high) delineated by 750 hPa as described later. We use only these two bands in our cloud fraction assimilation for two main reasons. First, this band selection is consistent with the CCM3 cloud fraction parameterization present in GEOS-4 (see section 2). Therefore, keeping these two bands required the minimum changes to that cloud parameterization to incorporate cloud fraction assimilation. Second, the use of only two bands makes the assimilation system less sensitive to cloud-top pressure errors that are present in the data.

For each ISSCP DX pixel, a determination of cloudy or clear is provided and, if cloudy, a cloud-top pressure is also available. If clear, the total, mid–high and low cloud fractions for the pixel are set to zero. If cloudy, the total cloud fraction for the pixel is set to one and the band cloud fractions are set as follows: if the cloud-top pressure is less than 750 hPa, the mid–high cloud fraction is set to one and the low cloud fraction to missing, otherwise the mid–high cloud fraction is set to zero and the low cloud fraction to one. When the collection of pixels that falls into a particular model grid box is averaged together, any missing values in the low band are excluded from the low band average. The resultant low cloud fraction is therefore equivalent to that derived under the assumption that the fraction of low cloud in the portion of the grid box obscured by mid–high clouds is the same as the fraction of low cloud in the unobscured portion. Actually, this comparison of the unobscured low cloud fraction with the full grid-column low cloud fraction is equivalent to the assumption of independence between clouds in the low and mid–high bands and, therefore, of random overlap between cloudiness in these two bands. This is an approximate way to remove the bias in low cloud fraction due to obscuration. Note that this strategy is different from that used to produce low cloud fraction in the ISCCP-D2 dataset, where only the fraction of unobscured low clouds is reported.

Each DX pixel also provides estimates of the cloud optical depth. In this study, all mean cloud optical depths are calculated in such a way as to approximately preserve the net shortwave albedo of the ensemble—see section 4b(3) for more details.

2) SSM/I CLOUD LIQUID WATER PATH

We also assimilate the cloud liquid water path retrieved over ocean from the SSM/I instrument onboard the Defense Meteorological Satellite Program (DMSP) series of polar orbiting satellites by Remote Sensing Systems (version 5) using the algorithm of Wentz (1997). The data comes in daily files, one for each satellite, each containing CLWP mapped to a regular grid ($0.25° \times 0.25°$ resolution) complete with data gaps between orbits. Two maps exist per file, one of ascending orbit segments and the other of descending orbit segments. Data on each of the segment maps are overwritten at both the high latitudes where successive orbits cross and at the "seam" or region where the last orbit of the day overlaps the first orbit of the day. Each grid value has a timestamp.

From these data files a single global map at $0.25° \times 0.25°$ resolution is produced containing only data observed within a 6-h window centered on the analysis time and for each "pixel" choosing the satellite and segment map from among the files with the observation time closest to the analysis time. This high-resolution map is then binned and averaged to the GEOS-4 model resolution ($1.25° \times 1°$) with appropriate weighting given to pixels overlapping model grid boundaries since the two grids are offset.

b. *Cloud parameter estimation system*

The cloud parameter estimating system is invoked immediately after PSAS and therefore operates with the analyzed moisture and temperature fields. It uses a minimization procedure to adjust selected parameters of the cloud parameterizations to yield better agreement between the cloud observables (ISCCP cloud fraction and optical depth and SSM/I liquid water path) and their modeled counterparts. This minimization is conducted independently at different geographic locations so that parameters have a geographic dependence, unlike the globally uniform parameters of the non-CPES control assimilation. By examining the variability of the estimated parameters, one can assess the homogeneity assumptions used by the physical parameterizations. Minimization is performed on a reduced resolution longitude–latitude analysis grid—currently, each analysis point is centered on and represents a 3×3 collection of model (longitude–latitude) grid boxes. Each pole is a special analysis point that represents the polar-model grid box and the collection of grid boxes one latitude index equatorward. The tuned parameters at analysis points are interpolated (bilinearly) to the full model grid for use by the model physics during the next 6-h background forecast.

1) CLOUD FRACTION

The cloud fraction parameterization for CPES runs replaces the CCM3 relative humidity dependence (1) with a new form,

$$f = \begin{cases} 0, & r \le 0, \\ (3 - 2s)s^2, & 0 < r < 1, \\ 1, & r \ge 1, \end{cases} \tag{4}$$

where

$$r = \frac{\mathrm{RH} - \mathrm{RH}_0}{\mathrm{RH}_1 - \mathrm{RH}_0} \tag{5}$$

and

$$s = \begin{cases} r^{1+\alpha}, & \alpha \ge 0, \\ 1 - (1 - r)^{1-\alpha}, & \alpha < 0. \end{cases} \tag{6}$$

The function f rises with RH in a smooth S-shaped manner from zero at RH_0 to one at RH_1. This latter relative humidity, at which cloud cover becomes overcast, is currently set at 120% for both low and mid–high

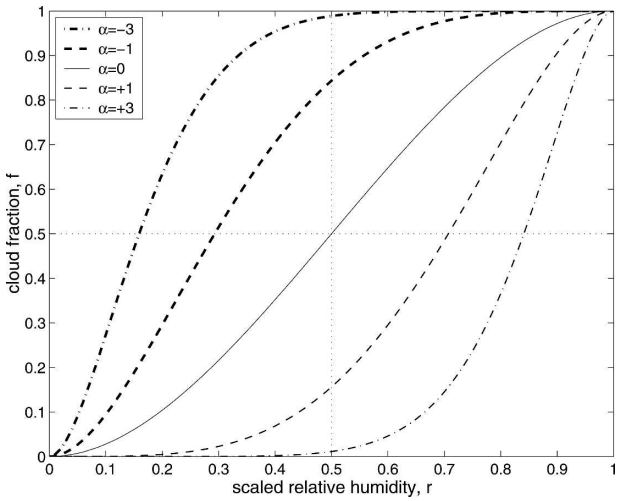

FIG. 9. Variation of cloud fraction with scaled relative humidity: RH scaled to the range [RH_0, RH_1]—for several values of asymmetry parameter α.

clouds. This value was found necessary to account for the observation of nonovercast cloudiness when the analysis indicates a saturated grid box.[3] The parameter α determines the asymmetry of the rise—zero for a symmetric rise, negative for a faster rise and positive for slower rise (see Fig. 9). The CPES uses (RH_0, α) as a two-parameter space in which to tune the cloud fraction.

The introduction of this generalized S-shaped f(RH) admits the possibility of the cloud fraction exceeding the grid box mean relative humidity (i.e., $f > \mathrm{RH}$). If f is taken as the fractional *volume* of cloud in the containing grid box and this volume is (by definition) taken as saturated ($RH_{cloudy} = 1$), then $f > \mathrm{RH}$ necessarily implies a negative relative humidity for the clear portion (i.e., $RH_{clear} < 0$). In physical terms this amounts to the cloudy portion containing more water vapor than is present in the entire grid box, which is clearly impossible. In the range $RH_0 < \mathrm{RH} < RH_1$ f is S-shaped and a physical violation only occurs when this curve rises above the line $f = \mathrm{RH}$ in the range $0 < \mathrm{RH} < 1$. From a technical perspective (see Fig. 9), violations will be greatest for larger negative α and lower RH_0 but will be lessened for $RH_1 > 1$, as in our implementation. From a practical perspective, violations will be forced by the observation of very cloudy or overcast conditions in a pressure band (low or mid–high) of a model grid column that is everywhere subsaturated. This may occur if the analysis is too dry. It may also occur for an observed cloud layer that is thinner in vertical extent than the model grid box in which it falls such that, even for a large areal cloud fraction, the volume cloud fraction is considerably less than unity. This last possibility suggests our practical approach to the removal of these physical violations: f in (4) is regarded as the *areal* cloud fraction. If this fraction is less than RH then the cloud fills the entire vertical extent of the grid box and the volume cloud fraction is also equal to f. If, however, $f >$ RH, then we set the volume cloud fraction equal to RH (the maximum physically allowable value) by having the cloud fill only a fraction $f_{vert} = \mathrm{RH}/f$ of the vertical extent of the grid box. The "vertical cloud fraction," f_{vert}, is used to scale the in-cloud condensate water path, Γ_c, of the grid box and thereby affects both the emissivity and optical depth calculations [see sections 2, 4b(2), and 4b(3)].

The stability, subsidence, and surface-type dependencies of the CPES cloud fraction parameterization remain the same as in CCM3. Currently, the convective cloud fraction is also unchanged and does not participate in the parameter estimation process. Although the convective cloud fraction (unlike the convective rainfall) is generally significantly smaller than the stratiform component in the CCM3, the possibility of adjusting convective cloud fraction parameters remains and will be investigated in the future. Convective parameter estimation will be better addressed once planned changes to the convective cloud parameterization have introduced a tighter coupling between the convective core and anvil/stratiform portions of convective systems. Finally, cloud fraction associated with low-level inversion stratus in the CCM3 code has been removed from the CPES version because this inversion stratus was producing excessive low cloud in winter high-latitudes. The now variable low-cloud relative humidity threshold and asymmetry parameters are able to produce low cloud cover consistent with the ISCCP observations without a specific low-level inversion stratus parameterization.

The parameters RH_0 and α are tuned using nonlinear downhill simplex minimization (Press 1988, section 10.4) under the constraints $0 \le RH_0 \le RH_1$ and $-10 \le \alpha \le 10$. Tuning is conducted for each of the low and mid–high cloud bands separately. The cost function minimized is

$$J(\mathrm{RH}_0, \alpha) = \left(\frac{\mathrm{RH}_0 - \mathrm{RH}_0^0}{\sigma_{\mathrm{RH}_0}}\right)^2 + \left(\frac{\alpha - \alpha^0}{\sigma_\alpha}\right)^2 + \left(\frac{\bar{f}^{\mathrm{mod}}(\mathrm{RH}_0, \alpha) - \bar{f}^{\mathrm{obs}}}{\sigma_{\bar{f}}}\right)^2. \quad (7)$$

[3] In the case of high clouds this may sometimes be physically justified since significant vapor supersaturations can exist in regions of ice nuclei shortage. However, in the main, and certainly for low clouds, the observation of clear areas in saturated grid boxes is most likely due to an imprecise calculation of the actual relative humidity due to temperature and moisture errors.

Here RH_0^0 and α^0 are reference parameter values set to the values of the tuned parameters from the previous CPES tuning cycle. The presence of the first two terms stabilizes the tuning procedure, preventing too rapid a variation in parameters between analysis cycles; $\bar{f}^{\mathrm{mod}}$ and $\bar{f}^{\mathrm{obs}}$ are the means of the modeled and ISCCP-derived grid box cloud fractions for the cloud band in question. These means are taken over the collection of model grid boxes making up an analysis point. Missing values are excluded from the averages and, if all values contributing to either average are missing, the existing parameters at that analysis point are left unchanged. The projection of ISCCP observations to model grid boxes was described in section 4a(1). For modeled band cloud fractions, the random overlap assumption is used, as in CCM3. The errors σ_{RH_0}, σ_α, and $\sigma_{\bar{f}}$ are used to penalize departures of the parameters from their reference values and the modeled cloud fractions from the observations. Currently the errors are globally uniform and set to 0.2, 0.3, and 0.1 respectively. Error values were empirically chosen to produce physically reasonable results, with some sensitivity studies to give reasonably smooth time variability of estimated parameters. More work on estimating optimal and adaptive values for these errors should be conducted.

2) Cloud liquid water path

In the second phase of the CPES, we assimilate SSM/I CLWP. The CPES continues to use the diagnostic CCM3 condensate parameterization (2), but the surface cloud water content, ρ_{cs}, becomes an adjustable parameter with spatial and temporal dependence. The cost function (at each CPES analysis point) is

$$J(\rho_{cs}) = \left(\frac{\rho_{cs} - \rho_{cs}^0}{\sigma_{\rho_{cs}}}\right)^2 + \left(\frac{\overline{\mathrm{CLWP}}^{\mathrm{mod}}(\rho_{cs}) - \overline{\mathrm{CLWP}}^{\mathrm{obs}}}{\sigma_{\overline{\mathrm{CLWP}}}}\right)^2, \quad (8)$$

with globally uniform $\sigma_{\rho cs} = 0.05$ g m^{-3} and $\sigma_{\overline{\mathrm{CLWP}}} = 5$ g m^{-2}, and ρ_{cs} constrained to the range 0–10 g m^{-3}. The reference values ρ_{cs}^0 are set to the ρ_{cs} from the previous CPES tuning cycle. The $\overline{\mathrm{CLWP}}^{\mathrm{obs}}$ and $\overline{\mathrm{CLWP}}^{\mathrm{mod}}$ are analysis point averages of the grid box quantities $\overline{\mathrm{CLWP}}_{gb}^{\mathrm{obs}}$ and $\overline{\mathrm{CLWP}}_{gb}^{\mathrm{mod}}$, these being, respectively, the grid box average of SSM/I pixels, described in section 4a(2), and the model column CLWP given by

$$\overline{\mathrm{CLWP}}_{gb}^{\mathrm{mod}} = \int_0^\infty f f_{\mathrm{vert}} \rho_c (1 - f_{\mathrm{ice}})\, \mathrm{d}z, \quad (9)$$

where f and f_{vert} are the areal and vertical cloud fractions, used to form the grid box average from the *in cloud* condensate ρ_c, and $(1 - f_{\mathrm{ice}})$ selects the liquid water component. Note that, since the parameter ρ_{cs} enters linearly into the model CLWP calculation, a simplex minimization is not actually needed, but we maintain the same general solution method since the procedure may later be generalized to multiple parameters.

The SSM/I CLWP retrievals are available only over ocean, so no observations are available for parameter adjustment over land. In these regions, the CCM3 global value of $\rho_{cs} = 0.21$ g m^{-3} is retained. Upcoming plans to also assimilate MODIS cloud data will allow both for water path data over land and also estimates of the cloud water path in both phases.

3) Cloud optical depth

In section 2 we noted that the entire effect of partial cloudiness and thereby of cloud overlap within the CCM3 shortwave code was parameterized by a single multiplicative term in the grid box optical depth, namely f^Q with $Q = 3/2$. We argue in the appendix that this is overly simplistic and that $Q = 3/2$ is not representative, with values of Q up to 10 being possible for a thick near-overcast single cloud layer in a column. For this reason, we decided to use Q as our adjustable parameter for the third and final CPES phase—optical depth assimilation. The factor $f\sqrt{f}$ in (3)[4] is replaced by f^Q, and Q is estimated using the same sort of minimization procedure as for the cloud fraction parameters or for ρ_{cs}, but in this case using ISCCP-DX–derived cloud optical depth as the observable. This is perhaps the weakest phase of the CPES for the following reasons: 1) there are other empirically specified elements of the model optical depth, namely the effective radii, which could also be chosen for adjustment; 2) adjustment of Q is of limited value in near-overcast or overcast layers where $f \approx 1$; and 3) the partial cloudiness aspect of the shortwave code will, at least in part, become less empirical in a more advanced shortwave code in which partial cloudiness and cloud overlap are more explicitly treated [such as in the new Community Atmosphere Model's shortwave code by Collins (2001)]. However, even with recent improvements (Hogan and Illingworth 2000; Pincus et al. 2005) in the specification of GCM cloud overlap, there will likely continue to remain a level of empiricism [e.g., in the specification of the decorrelation distance in Hogan and Illingworth (2000), or the rank correlation decay scale in Pincus et

[4] Note that in the CPES version of (3) the f in f^Q is the areal cloud fraction. The vertical cloud fraction f_{vert} is multiplied into the integrand in the CPES calculation of the in-cloud condensate path, Γ_c.

al. (2005), under differing synoptic conditions]. We therefore expect there will remain some tunability of this aspect of the shortwave code. Therefore, we have proceeded with optical depth tuning using Q with the realization that we are largely demonstrating a method in this paper and that changes to the adjustment process will come with improved model physics and additional potential observables such as effective radius.

Note that, when we discuss the cloud optical depth of a region, we consistently refer to the visible optical depth of a thin cloud layer covering the entire region that approximates the shortwave radiative effect of the combined cloudy and clear portions of the region. This is to be distinguished from the in-cloud cloud optical depth. This is consistent with the CCM3 shortwave treatment in which the cloudy atmosphere fluxes are calculated based on layer-cloud optical depths that implicitly factor in the model cloud fraction via the f^Q factor. It is also consistent with the pixel-cloud optical depth provided by ISCCP retrievals, which employ radiative models based on a single thin overcast cloud layer in the column.

An observed column cloud optical depth, denoted $\overline{\tau}_{gb}^{\mathrm{obs}}$, is calculated for each model grid box from the collection of contained ISCCP-DX pixels. If any of the ensemble of contained pixels are nighttime pixels, $\overline{\tau}_{gb}^{\mathrm{obs}}$ is set to missing to avoid any bias since the optical depth is unknown for cloudy nighttime pixels. The cloud optical depth of a clear pixel is zero. For cloudy pixels, two ISCCP cloud optical depths are provided—one for a droplet radiative model with an effective radius r_e = 10 μm and one for an ice crystal radiative model with r_e = 30 μm (Rossow et al. 1996). We select the optical depth associated with the ISSCP retrieved phase, which is ice if the cloud top is colder than 260 K (to allow for the existence of supercooled water below 0°C). The calculation of $\overline{\tau}_{gb}^{\mathrm{obs}}$ is not a simple linear average of the cloud optical depths of the pixels but a weighted average that seeks to reproduce the hemispherical shortwave albedo of the ensemble. Specifically, for each pixel, the delta-Eddington approximation is used to calculate the hemispherical shortwave albedo of a single overcast cloud layer assuming conservative scattering and an underlying black surface. This delta-Eddington albedo, $A_{\mathrm{dE}}(\tau, \mu_0, f_s, g)$, is a function of the cloud optical depth τ, the cosine zenith angle μ_0, the asymmetry factor g, and the forward scattered fraction f_s, (see the appendix) and with A_{dE} set to zero for clear pixels. CCM3 formulas are used to evaluate g and f_s as functions of the ISCCP r_e given above for the retrieved phase. We find an iterative solution of $A_{\mathrm{dE}}(\overline{\tau}_{gb}^{\mathrm{obs}}, \overline{\mu_0}, \langle g\rangle, \langle f_s\rangle) = \overline{\mu_0 A_{\mathrm{dE}}}/\overline{\mu_0}$, for a $\overline{\tau}_{gb}^{\mathrm{obs}}$ between the minimum and maximum cloud optical depth of any single pixel, where the overbar is the mean over all ensemble pixels, and $\langle\,\rangle$ is a mean over cloudy pixels only. Finally, the observed cloud optical depth at *analysis* points, $\overline{\tau}_{gb}^{\mathrm{obs}}$ is formed from the ensemble of contributing grid box values, $\overline{\tau}_{gb}^{\mathrm{obs}}$ in a manner analogous to the calculation of $\overline{\tau}_{gb}^{\mathrm{obs}}$ from ISCCP-DX pixel values.

Also needed for the minimization is $\overline{\tau}^{\mathrm{mod}}$, the modeled cloud optical depth at analysis points. The gridcolumn cloud optical depth is evaluated as in the model shortwave code [namely, the vertical integral of (3) with CPES modifications as discussed earlier]. To find the radiatively weighted analysis point mean, using the method described above, the g and f_s values for cloudy grid columns are required. The model does evaluate such values but at each vertical layer. Rather than attempting to devise a method to weight the layer values to calculate an effective column g and f_s, we use the *observed* cloud-top temperature and pressure to evaluate $r_e^{\mathrm{liq}}(T)$, $r_e^{\mathrm{ice}}(P)$ and $f^{\mathrm{ice}}(T)$ as per the CCM3 code (see Kiehl et al. 1996, p. 50) and then evaluate $f(r_e)$ and $g(r_e)$ using the CCM3 routines and $r_e = r_e^{\mathrm{liq}}(1 - f^{\mathrm{ice}}) + r_e^{\mathrm{ice}} f_e^{\mathrm{ice}}$. This use of the model r_e for $\overline{\tau}^{\mathrm{mod}}$ yields a better radiative tuning of the model to the observations since the model shortwave calculation continues to use its own r_e calculation in forecast mode.

Finally, the optical depth cost function is

$$J(Q) = \left(\frac{Q - Q^0}{\sigma_Q}\right)^2 + \left(\frac{\overline{\tau}^{\mathrm{mod}}(Q) - \overline{\tau}^{\mathrm{obs}}}{\sigma_{\overline{\tau}}}\right)^2, \quad (10)$$

with globally uniform σ_Q = 0.5 and $\sigma_{\overline{\tau}}$ = 2, and Q constrained to the range 1–10. The reference value Q^0 is set to the Q from the previous CPES tuning cycle.

c. *Example of spatial and temporal parameter dependence*

Figure 10 shows an example of the mean spatial dependence of the cloud fraction parameters for the month of December 2000. Clearly there is significant spatial variation in the parameters. Numerous interesting features are visible in the figure, such as the regions in the northwest Pacific and Atlantic and in the Southern Ocean in which clouds appear to be forming at very low relative humidities. The northern ocean basin cases may well be due to the occurrence of convective clouds generated by cold air from continents moving out over these ocean regions and destabilizing the warmer and wetter marine boundary layer. If such convective clouds are not being modeled by the convective parameterizations, the CPES will attempt to replicate the observed cloudiness by producing large-scale cloud,

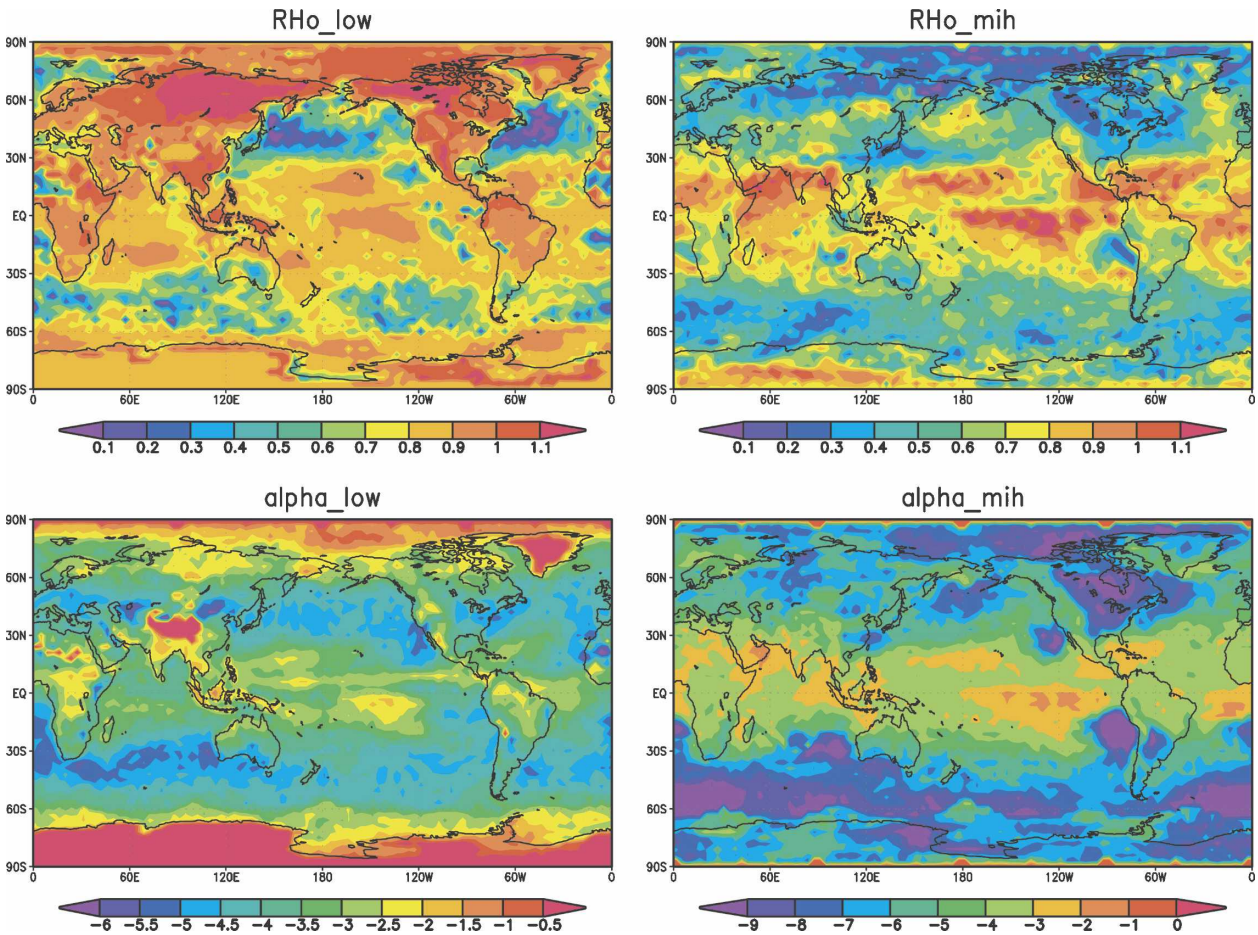

FIG. 10. Mean cloud parameters for December 2000 for a full cloud assimilation run: (top) critical relative humidity, RH_0; (bottom) asymmetry parameter, α; (left) low and (right) mid–high band.

even though the cold outflowing continental air has a low relative humidity. We will analyze such cases further in Part II.

It is important to note that the observed spatial dependence of estimated parameters arises not only from the inhomogeneity of unresolved subgrid-scale variability at the GCM resolution, but also from the fundamental limitations of the cloud fraction parameterization. The optimal use of this information in the future is to aid in the design of an improved cloud fraction parameterization that more fully captures the nature of subgrid-scale moisture and temperature variability and is able to diagnose or prognose that variability in terms of resolved model fields. Such an improved parameterization should show increased homogeneity and stationarity in its estimated parameters, thus showing that it more fully captures the observed physics and spatial variability at the model resolution. Furthermore, because climate runs are an important end user of the parameterizations and because there is no "future data" with which to tune the parameters in future climate mode, it is important to design parameterizations with homogeneous and stationary parameters. Thus, this cloud parameter estimation system can be used as a tool to assess the performance of future candidate cloud parameterizations.

Figure 11 shows the time dependence of the parameter RH_0 in different regions during the first week of an assimilation run (16–23 December 1999). Clearly for mid–high RH_0 there is some spindown in the first few days.

This implies that the initial value of RH_0 used in the run was too large compared to what the mid–high cloud fraction data requires. Specifically, this excessive initial critical relative humidity value produced less mid–high cloud than ISCCP was observing. The assimilation system responds by decreasing the mid–high RH_0, thereby increasing the modeled mid–high cloud fraction. The system clearly reaches quasi equilibrium after a few days. This equilibrium is different for different regions

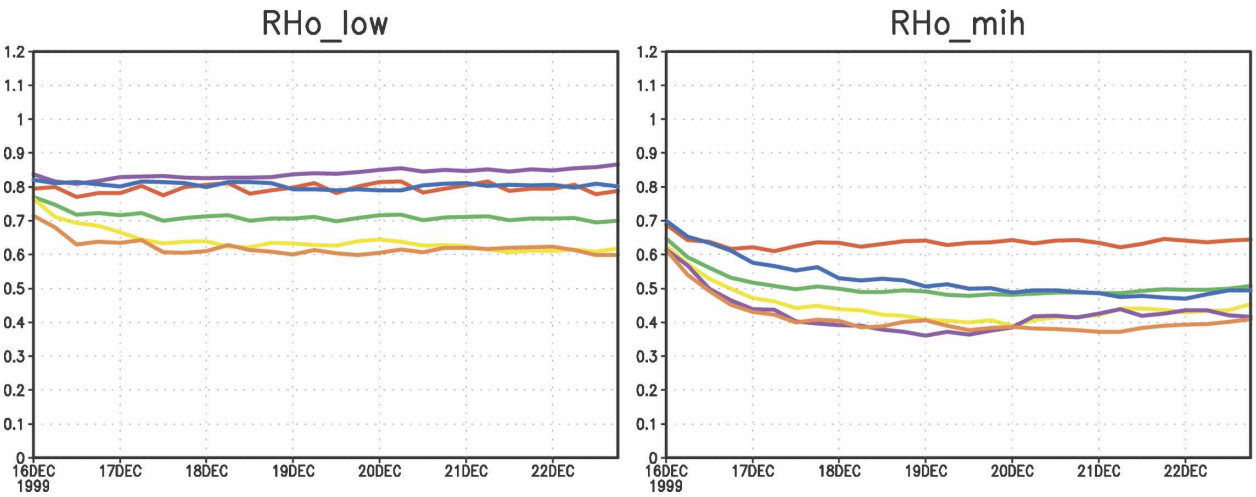

FIG. 11. (left) Low and (right) mid–high RH_0 for the first week of a full cloud assimilation run. Regions are global (green), Tropics (±20°, red), northern midlatitudes (20°–60°, yellow), northern high latitudes (60°–90°, purple), southern midlatitudes (orange), and southern high latitudes (blue).

of the planet, indicating the need for geographically varying parameters. Remaining trends after the spin-down period are forced by synoptic variability.

5. Discussion and conclusions

The development of this new cloud assimilation system was motivated by the poor prediction of zonal cloud cover by the GEOS-4 GCM running in climate mode and the even poorer performance in assimilation mode, for which the introduction of observed moisture drives the model cloud fraction even further from the observations. Given that global climate model cloud physics is highly parameterized and that the CCM3 physics of the GEOS-4 GCM is particularly simple and empirical, we decided to take the route of parameter estimation since, in reality, cloud properties at the GCM scale are strongly dependent on subgrid-scale moisture variability, which varies on synoptic, geographic, and diurnal scales. Thus, we replaced the fixed, globally assigned cloud parameters of the base system by spatially and temporarily varying parameters in the new CPES and estimated these parameters anew at each analysis time so as to improve modeled cloud fraction, liquid water path, and cloud optical depth, yet constraining the estimation to ensure smooth and slow time variation of the parameters.

The first CPES stage estimates a smooth, two-parameter, S-shaped cloud fraction—relative humidity dependence using ISCCP-derived cloud fraction data in low and mid–high pressure bands. As well as vastly improving the model's cloud fraction representation, this stage also greatly improves the model's longwave cloud forcing against independent CERES data. Further improvement in longwave cloud forcing is achieved in a second CPES stage by assimilating SSM/I cloud liquid water path data using adjustment of the model's diagnostic cloud condensate parameterization. A side effect of these two stages, which generally increase both the cloud fraction and cloud water content, is an unrealistic increase in the magnitude of model shortwave cloud forcing. Upon investigation, it was determined that the treatment of partial cloudiness in the CCM3 cloud optical depth code was deficient, but could be adjusted in a third parameter estimation phase, by assimilation of ISCCP cloud optical depth, to restore the model's shortwave cloud forcing to a quality comparable to the base model. Forecast experiments have verified that the improvement in the model cloud fraction representation is maintained in pure simulation mode if the CPES estimated cloud parameters are used during the forecast.

Thus, we have demonstrated, albeit in a system with simplified cloud physics, the value of a parameter estimation approach to cloud data assimilation. This approach is not exclusive of variational state adjustment approaches (e.g., 3D- or 4DVAR) using cloudy radiance assimilation, but complementary to it since we believe that those approaches will benefit from the less biased cloud properties of a model running the CPES. In addition, this parameter estimation approach will continue to be valuable in GCMs with more complicated cloud physics and even at smaller scales in cloud system resolving models (CSRMs), since not only is

there considerable uncertainty in the details of cloud microphysics, but even at CSRM scales unresolved subgrid-scale variability remains a key problem for the parameterizations. This type of parameter estimation method is ideally suited to estimation of a parametric probability density function of subgrid-scale variability as used in several recent statistical cloud parameterizations, (e.g., Tompkins 2002).

Acknowledgments. This work was partially supported by the NASA Interdisciplinary Science Program at Goddard Space Flight Center under WBS-51-291-01-C7. The ISCCP-DX data were obtained from the International Satellite Cloud Climatology Project data archives at the NASA Langley Research Center Atmospheric Sciences Data Center. The SSM/I data were produced by Remote Sensing Systems and sponsored by the NASA Pathfinder Program for early Earth Observing System (EOS) products.

APPENDIX

The Treatment of Partial Grid Box Cloudiness in the CCM3 Shortwave Code

In section 2 we described the simplified treatment of partial cloudiness in the CCM3 shortwave code—namely, that a grid box column containing partially cloudy layers is converted to a column of overcast cloud layers, and the actual effect of partial cloudiness and cloud overlap in the column is parameterized by multiplying the optical depth by $f^{3/2}$ within each layer.

To investigate this simplification, consider the albedo of an isolated cloud layer under the delta-Eddington approximation and assumptions of conservative scattering ($\omega_0 = 1$), zero gaseous scattering or absorption, and a black (nonreflective) surface. The hemispherical albedo of the cloud layer is then

$$A_{\mathrm{dE}}(\tau, \mu_0, f_s, g) = 1 - 2\frac{\left(1 + \frac{3}{2}\mu_0\right) + \left(1 - \frac{3}{2}\mu_0\right)e^{-\tau(1-f_s)/\mu_0}}{4 + 3(1-g)\tau}, \quad \text{(A1)}$$

where τ is the optical depth of the cloud, μ_0 is the cosine of the solar zenith angle, and f_s and g are the forward scattered fraction and asymmetry parameter for cloud droplet scattering. To further simplify this equation, we will consider the case $\mu_0 = 2/3$ in which the second term in the numerator vanishes (as it also does for thick clouds). Then we may write

$$A_{\mathrm{dE}}(\tau, 2/3, f_s, g) = 1 - [1 + \alpha\tau]^{-1} \equiv A(\tau), \quad \text{(A2)}$$

where $\alpha = 3(1 - g)/4$ is taken as a constant and $A(\tau)$ will be used as a shorthand for this expression in what follows. For typical cloud droplets and visible light, $g \approx 0.87$, so $\alpha \approx 0.1 \ll 1$. Hence for thin clouds, $\tau < 1$, $A(\tau) \approx \alpha\tau$.

If the cloud layer occupies an areal fraction f of the grid column, the net albedo of the column is then just $fA(\tau)$. Following the CCM3 approach, which uses an "equivalent overcast optical depth," we seek Q such that $A(\tau f^Q) = fA(\tau)$. Clearly, for thin clouds $Q = 1$. In general,

$$Q^{(1)} = \frac{\log([1 - fA(\tau)]^{-1} - 1) - \log(\alpha\tau)}{\log f}, \quad \text{(A3)}$$

where we are using the superscript $^{-1}$ to denote a single cloud layer in the column. This is plotted in Fig. A1a. Clearly, some very large values of Q are possible for large optical depths and near-overcast conditions.

Now let us consider the slightly more complex case of two randomly overlapped cloud layers. To keep matters simple, we will assume that each layer has an identical cloud fraction f and optical depth τ. Then, looking from above, the column will contain three distinct regions of optical depth 0, τ, and 2τ, and areal fractions $(1 - f)^2$, $2f(1 - f)$, and f^2, respectively. Indeed, for the more general case of N such identical but randomly overlapped cloud layers, there are $N + 1$ distinct regions, denoted $m = 0, \ldots, N$, with optical depths $m\tau$ and fractions

$$\binom{N}{m} f^m (1 - f)^{N-m},$$

respectively. The net column albedo will be

$$A^{(N)} = \sum_{m=1}^{N} \binom{N}{m} f^m (1 - f)^{N-m} A(m\tau).$$

Under the CCM3 approach this must equal the albedo of N overcast layers, each with optical depth τf^Q. Hence we require $A(N\tau f^Q) = A^{(N)}$. Thus, in general,

$$Q^{(N)} = \frac{\log([1 - A^{(N)}]^{-1} - 1) - \log(N\alpha\tau)}{\log f}. \quad \text{(A4)}$$

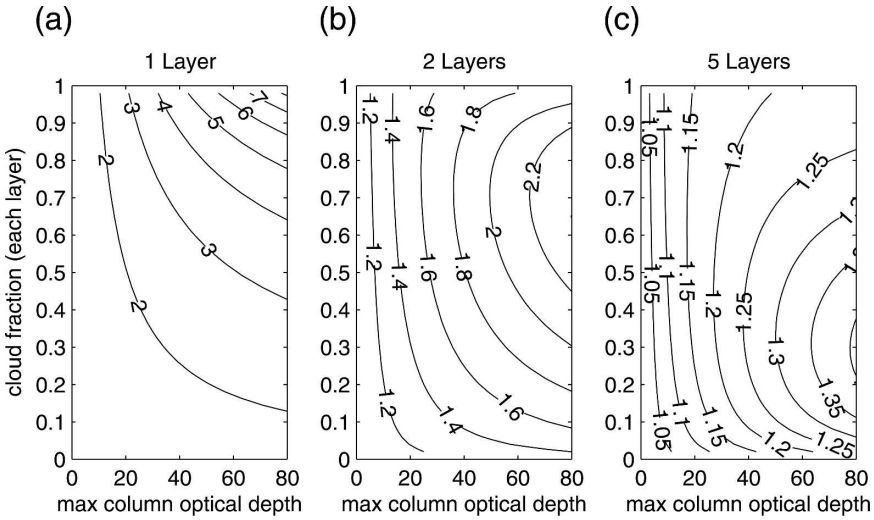

FIG. A1. Plots of the cloud fraction exponent Q against the maximum optical depth in any region of the grid column and against the cloud fraction in each of (a) 1, (b) 2, and (c) 5 randomly overlapped but otherwise identical cloud layers. Specific details are discussed in the text.

Figures A1b and A1c show $Q^{(N)}$ versus $N\tau$ (which is the maximum observed τ anywhere in the column) for $N = 2, 5$.

While we have made several simplifying assumptions in terms of both the radiative transfer and the properties and overlap of the clouds considered, it is clear from Fig. A1 that values of Q anywhere from 1 to 10 are possible and that $Q = 3/2$ is by no means a representative value.

REFERENCES

Chaboureau, J.-P., and J.-P. Pinty, 2006: Validation of a cirrus parameterization with Meteosat Second Generation observations. *Geophys. Res. Lett.,* **33,** L03815, doi:10.1029/2005GL024725.

Chevallier, F., P. Lopez, A. M. Tompkins, M. Janisková, and E. Moreau, 2004: The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Quart. J. Roy. Meteor. Soc.,* **130,** 917–932.

Collins, W. D., 2001: Parameterization of generalized cloud overlap for radiative calculations in general circulation models. *J. Atmos. Sci.,* **58,** 3224–3242.

Greenwald, T. J., T. Vukićević, L. D. Grasso, and T. H. Vonder Haar, 2004: Adjoint sensitivity analysis of an observational operator for visible and infrared cloudy-sky radiance assimilation. *Quart. J. Roy. Meteor. Soc.,* **130,** 685–705.

Hogan, R. J., and A. J. Illingworth, 2000: Deriving cloud overlap statistics from radar. *Quart. J. Roy. Meteor. Soc.,* **126,** 2903–2909.

Janisková, M., J.-F. Mahfouf, and J.-J. Morcrette, 2002: Preliminary studies on the variational assimilation of cloud-radiation observations. *Quart. J. Roy. Meteor. Soc.,* **128,** 2713–2736.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, 1996: Description of the NCAR community climate model (CCM3). NCAR Tech. Note NCAR/TN-420+STR, 152 pp.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.,* **132,** 2293–2307.

Lipton, A. E., and G. D. Modica, 1999: Assimilation of visible-band satellite data for mesoscale forecasting in cloudy conditions. *Mon. Wea. Rev.,* **127,** 265–278.

Macpherson, B., B. J. Wright, W. H. Hand, and A. J. Maycock, 1996: The impact of MOPS moisture data in the U.K. Meteorological Office mesoscale data assimilation scheme. *Mon. Wea. Rev.,* **124,** 1746–1766.

Pincus, R., C. Hannay, S. A. Klein, K.-M. Xu, and R. Hemler, 2005: Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models. *J. Geophys. Res.,* **110,** D15S09, doi:10.1029/2004JD005100.

Press, W. H., 1988: *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 735 pp.

Rossow, W. B., A. W. Walker, D. E. Beuschel, and M. D. Roiter, 1996: International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets. WMO Tech. Doc. 737, 115 pp.

Slingo, J. M., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Meteor. Soc.,* **113,** 899–927.

Tompkins, A. M., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.,* **59,** 1917–1942.

Wentz, F. J., 1997: A well-calibrated ocean algorithm for SSM/I. *J. Geophys. Res.,* **102,** 8703–8718.

Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System experiment. *Bull. Amer. Meteor. Soc.,* **77,** 853–868.

Wong, T., D. F. Young, M. P. Haeffelin, and S. Weckmann, 2000: Validation of the CERES/TRMM ERBE-like monthly mean clear-sky longwave dataset and the effects of the 1998 ENSO event. *J. Climate,* **13,** 4256–4267.

Wu, X., and W. L. Smith, 1992: Assimilation of ERBE data with a nonlinear programming technique to improve cloud-cover diagnosis. *Mon. Wea. Rev.,* **120,** 2009–2024.