



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Critical Evaluation of the ISCCP Simulator Using Ground-Based Remote Sensing Data

Gerald G. Mace, Stephanie Houser, Sally Benson,
Stephen A. Klein, Qilong Min

November 2, 2009

Journal of Climate

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Critical Evaluation of the ISCCP Simulator Using Ground-Based Remote Sensing Data

Gerald G. Mace, Stephanie Houser, Sally Benson

Stephen A. Klein, and Qilong Min

To Be Submitted to Journal of Climate

Abstract:

Given the known shortcomings in representing clouds in Global Climate Models (GCM) comparisons with observations are critical. The International Satellite Cloud Climatology Project (ISCCP) diagnostic products provide global descriptions of cloud top pressure and column optical depth that extends over multiple decades. The necessary limitations of the ISCCP retrieval algorithm require that before comparisons can be made between model output and ISCCP results the model output must be modified to simulate what ISCCP would diagnose under the simulated circumstances. We evaluate one component of the so-called ISCCP simulator in this study by comparing ISCCP and a similar algorithm with various long-term statistics derived from the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) Climate Research Facility ground-based remote sensors. We find that were a model to simulate the cloud radiative profile with the same accuracy as can be derived from the ARM data, then the likelihood of that occurrence being placed in the same cloud top pressure and optical depth bin as ISCCP of the 9 bins that have become standard ranges from 30% to 70% depending on optical depth. While the ISCCP simulator improved the agreement of cloud-top pressure between ground-based remote sensors and satellite observations, we find minor discrepancies due to the parameterization of cloud top pressure in the ISCCP simulator. The primary source of error seems to be related to discrepancies in visible optical depth that are not accounted for in the ISCCP simulator. We show that the optical depth discrepancies are largest when the assumptions necessary for plane parallel radiative transfer optical depths retrievals are violated.

1. Introduction

Clouds play an important role in the earth's climate system through their modification of the earth's radiative energy and hydrologic cycles. Not only do clouds act to modify the energy and water cycles, they are themselves sensitive to changes in the climate state. Among the primary feedback processes in the earth's climate system (water vapor, surface albedo, and lapse rate feedbacks - Soden and Held 2006), uncertainties in the representation of cloud feedbacks in global climate models (GCM) have been consistently identified as the primary source of uncertainty in prediction of anthropogenic climate change (Dufresne and Bony, 2008).

GCMs in the recent IPCC fourth climate assessment (2007) have resolutions that are spatially and temporally much coarser than the spatial and temporal scales important to the evolution of cloud systems. Therefore, the impact of clouds systems (i.e. the radiative and hydrologic forcing) must be represented statistically through parameterizations of the dominant physical processes that result in the forcing (Randall et al. 2003). This task is difficult given the large variety of clouds ranging from deep convection to thin cirrus and the different processes involved. Many studies have shown that shortcomings in the prediction of present day cloud forcing and cloud occurrence represent a major component of the cloud uncertainty associated with cloud feedbacks in future climates (e.g. Dufresne and Bony, 2008; Williams and Tselioudis, 2007). A path forward to improved prediction of cloud feedbacks lies in improved representation of clouds in the present climate state. Comparisons between model output and observations is, therefore, quite important.

While these issues continue to be critical to our present day needs in understanding the climate system and predicting changes to it, the International Satellite Cloud Climatology Project

(ISCCP) was initiated in the early 1980's with similar motivations (Schiffer and Rossow 1983). This level of foresight is clearly a credit to the developers of ISCCP because, more than a quarter century later, ISCCP remains a flagship description of the cloudy atmosphere. By analyzing visible and infrared radiances produced by geostationary and polar orbiting meteorological satellites and applying several assumptions regarding the layering of clouds in the atmosphere, the thermodynamic phases, and their properties, ISCCP produces two key outputs for a cloudy satellite pixel: the column optical depth (τ), and cloud-top pressure (P) of the highest cloudy layer in the column. Hereafter we refer the ISCCP cloud top pressure as P_{CTP} and visible optical depth as τ_{vis} .

It would seem that the long-term global climatology of ISCCP directly addresses the needs of the GCM community. However, before comparing statistics derived from ISCCP and output from GCMs, a bridge must be employed between the model output and the ISCCP diagnostics that accounts for the assumptions made by ISCCP in producing the physical quantities of P_{CTP} and τ_{vis} . In other words, output from the GCMs must be interpreted with a set of equivalent assumptions. This bridge, known as the ISCCP simulator (Klein and Jakob, 1999; Webb et al., 2001) has been developed and is extensively used.

There are two components to the ISCCP simulator. Since a GCM represents clouds within a finite spatial grid that is often much coarser than the satellite measurements, it is necessary to downscale the model data to match the scale of the satellite. This statistical downscaling approach, known as the Subgrid Cloud Overlap Profile Sampler (SCOPS), is described by Klein and Jakob (1999). The other component, and the one we address here, is the representation of CTP and tau from the model output in a manner that is similar to what ISCCP

would produce from radiance measurements. This component of the simulator is known as the ISCCP Clouds and Radiances Using SCOPS (ICARUS).

The ISCCP simulator has become an important tool that has resulted in many model-data comparison studies. Zhang et al. conducted one such study in 2005. They did an extensive cloud climatology comparison with the ISCCP simulator used in ten atmospheric general circulation models. They categorized cloud types using what have become the standard nine ISCCP cloud types and compared them to the ISCCP results and to the Clouds and Earth's Radiant Energy System (CERES) program measurements (Minnis et al. 1995). The results of that study show a large difference among the models in simulation of upper-level clouds. The study also shows that most of the models only simulated less than 30 to 40% of the middle level clouds reported by ISCCP. As for low clouds, none of the models overestimated them while about half of the models underestimated them. Zhang et al. also grouped the nine types into three separate subgroups to better describe the systematic model biases. The first subgroup consisted of the middle and low-level clouds with optically thin ($\tau < 3.6$) and optically intermediate ($3.6 < \tau < 23$) thicknesses. They found that all the models significantly underestimated clouds in this subgroup, with the grand mean frequency of occurrence of all the model results at only about half of the ISCCP (41%) and CERES (43%) measurements. Another subgroup combined all the optically thick ($\tau > 23$) clouds at all three (low, mid and high) height intervals. The majority of the models overestimated this subgroup of clouds with a grand mean of 15.4% of this cloud type in all models with ISSCP and CERES having 6.9% and 8.1% respectively (Zhang et al. 2005).

While the ISCCP simulator has proven to be an indispensable tool in model evaluation, a thorough validation of the ISCCP simulator has yet to be conducted. In an initial examination of the ISCCP simulator, Mace et al. (2006; hereafter referred to as M06) used cloud properties

derived from ground based remote sensors at the Atmospheric Radiation Measurement (ARM) Southern Great Plains site as input to the ICARUS algorithm and then compared the resulting cloud top pressure (hereafter P_{ICARUS} and P_{LBTM}) to P_{ISCCP} and P_{LBTM} and a similar satellite algorithm known as the layer bispectral threshold method (LBTM; Minnis et al. (1995) – hereafter P_{LBTM} and P_{LBTM}). Using data from the year 2000, the P_{ICARUS} - P_{LBTM} statistics compare much better to ISCCP than simply comparing the unaltered P and τ derived from the ground-based ARM data (hereafter P_{ARM} and τ_{ARM}) to P_{ISCCP} and τ_{ISCCP} . However, the statistics of P_{LBTM} tended to be biased lower than P_{ICARUS} and P_{LBTM} (see their Figure 5) and the statistics of τ_{LBTM} tended to be biased high compared to τ_{ICARUS} and τ_{LBTM} (their Figure 6). Because the differences found in M06 have similarities to the differences found between the simulated ISCCP data produced from the models and the actual ISCCP data, a more through examination of the ISCCP simulator is considered here.

2. Data and Technique

Our hypothesis is that if observed cloud property and thermodynamic profiles are provided as input to the ISCCP simulator, then the simulator will produce P_{ISCCP} and τ_{ISCCP} similar to P_{ICARUS} and τ_{ICARUS} . Our goal is not necessarily to evaluate the validity of ISCCP. Our goal is to evaluate the degree to which ICARUS simulates ISCCP when given an observed physical distribution of cloud occurrence and cloud properties.

The ICARUS portion of the ISCCP simulator has two primary components. One parameterizes the IR radiance or brightness temperatures of the clear and cloudy atmospheres that a satellite would observe given the input cloud property and thermodynamic profiles and the other component takes the vertical profile of visible cloud extinction and the simulated IR brightness temperatures and then derives P_{ISCCP} and τ_{ISCCP} using ISCCP-like assumptions (Rossow et

al. 1996). The procedure for this second component is as follows. First, the temperature at cloud-top is derived from the IR brightness temperatures and column visible optical depth assuming, like ISCCP, the presence of a single layer of cloud. Then, the simulated cloud-top pressure is assigned the lowest pressure (highest altitude) in the troposphere for which the temperature of the input sounding matches this derived cloud-top temperature. Finally, the simulated optical depth is identical to the input optical depth in all cases except for optically very thin clouds for which the single-layer cloud retrieval fails. In this case, a nominal value of optical depth is assigned following ISCCP documentation (Rossow et al. 1996). So, except for profiles with . Applying these procedures accounts for the fact that the ISCCP cloud-top pressure differs substantially from the true cloud-top pressure of the highest cloud in the column particularly in multi-layer cloud situations where the highest cloud is transmissive to longwave radiation or in situations of low cloud located underneath a strong temperature inversion. For the first situation, the simulated cloud-top pressure is higher than the true cloud-top pressure and typically results in simulated cloud-top pressures at middle levels of the atmosphere when the true cloud-top pressure is at high levels. For the second situation the simulated cloud-top pressure is lower than the true cloud-top pressure and typically results in simulated cloud-top pressures at middle levels of the atmosphere when the true cloud-top pressure is at low levels.

To verify the radiances that are parameterized by the ISCCP simulator, we calculated clear and cloudy TOA radiances using the Moderate Spectral Resolution Atmospheric Transmittance (MODTRAN) model (Berk et al. 2003). This tests the ability of the simplified radiative transfer in ICARUS to reproduce infrared radiances produced by the MODTRAN model. We then applied the second component of the ICARUS algorithm to these MODTRAN

radiances to calculate the ISCCP simulated cloud top pressure and optical depth (hereafter τ_{sim} and p_{sim}).

For this study, the area of focus is the ARM SGP site in Oklahoma. Ground based observational data are collected continuously for the vertical column directly over the site, and are processed to give a physical description of the atmospheric profile for a 90m vertical grid every 5 minutes. The cloud property profiles are determined using a combination of vertically pointing radar reflectivity, Doppler Velocity, lidar-derived cloud boundaries, and liquid water paths derived from microwave radiometer data. The results have been validated against aircraft in situ data, surface radiometric fluxes, and TOA radiometric fluxes (M06 and Mace et al. 2008). Additionally, the M06 results compared favorably with optical depths derived from Multifilter Rotating Shadowband Radiometer (MFRSR) data using a technique described by Min and Harrison (1996). We also use the Min and Harrison optical depth (hereafter τ_{MH}) as an independent point of comparison in this study.

τ_{sim} and p_{sim} are calculated by passing the cloud property profiles derived from the algorithm suite described in M06 through the ISCCP simulator ICARUS algorithm. The thermodynamic profiles of temperature, pressure, and water vapor are derived from a combination of radiosonde and precipitable water path derived from microwave radiometer measurements. Following the general convention of the modeling community, we use the version of the ISCCP simulator that creates τ_{sim} and p_{sim} from daytime data. As expected, τ_{sim} reported by ICARUS is nearly always identical to τ_{MH} while p_{sim} often varies substantially from p_{MH} depending on the vertical distribution of cloud properties.

There are two significant challenges in testing our hypothesis. First, we assume that the cloud properties input to ICARUS represent a realistic version of the actual cloud properties for

that 5-minute period. Because we use active remote sensing observations and soundings, the vertical locations of the cloud layers and thermodynamics in the vertical column are reasonably certain. The vertical distribution of cloud properties is less certain. However, the algorithms used in M06 and reported on again in M08 were evaluated extensively using radiative closure, aircraft data, and other retrievals and showed minimal bias.

The second challenge in testing our hypothesis is that the ISCCP measurements are derived using radiances collected instantaneously from satellite imagery while the ARM ground-based data are collected from a point on the surface over a period of time. Clearly, situations that have highly variable cloud fields in either space or time are not reasonable candidates for comparison. Therefore, we implement a rather strict set of criteria that are much more stringent than what was used in M06. We define a case to be the union of a 1-hour interval in time during which the ARM data are averaged with a particular point in time (at the center time of the averaging interval) where ISCCP retrievals are conducted within a geographic rectangular domain centered on the SGP site that is 100 km on a side. For a case to be used in the comparison, that case had to have met all of the following criteria:

1. All ISCCP pixels within the averaging domain reported the presence of cloud.
2. The standard deviation of in the 100 km domain must have been less than 100 mb.
3. All ARM 5-minute profiles during the 1-hour averaging period had to have contained cloud at some level.
4. All during the 1-hour averaging period were limited to values between 1 and 100.

We use the reported daylight and from the ISCCP D series data set from 1997 to 2002. These data are reported at 3-hour intervals and sampled every 30km from the native geostationary satellite data providing 5 ISCCP data points within the averaging domain.

We average the ISCCP data within the 100 km domain centered on the ARM SGP central facility to create $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$. $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$ are compared with similar quantities derived from the ground based data that have been averaged during a 1-hour period centered on the ISCCP measurement time (i.e. $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$ and $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$). $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$ are calculated by using the 1-hour averaged cloud property profile as input to ICARUS. The LBTM retrievals are computed from the GOES 8 radiances at the time nearest the center of the averaging interval and reported for the pixel nearest the SGP central facility.

We experimented with various approaches to averaging the data sets and choosing candidate cases. These include bigger and smaller ISCCP and LBTM averaging domains, longer and shorter averaging periods for the ARM data, and more and less stringent criteria to choose cases. While the quantitative details differed somewhat, the qualitative conclusions that we report on in the next section did not change.

3. Results

In Figure 1 we compare various renditions of $\overline{C_{\text{ISCCP}}}$ and in figure 2 the $\overline{C_{\text{LBTM}}}$ quantities are compared. Regression statistics for $\overline{C_{\text{ISCCP}}}$ and $\overline{C_{\text{LBTM}}}$ are listed in Tables 1 and 2. In the comparison of ISCCP with LBTM, a lack of any significant bias suggests that the two satellite algorithms tend to produce reasonably similar results while the scatter in the comparisons likely arises from differences in the algorithms and from comparing the spatially averaged ISCCP to the single pixel LBTM product nearest the SGP central facility. For both ISCCP and LBTM, the improvement relative to observations in the comparison of $\overline{C_{\text{ISCCP}}}$ is evident. $\overline{C_{\text{LBTM}}}$ compared to the satellite products show two clusters of points in the lower and upper troposphere with fewer points in the middle troposphere recorded by the active remote sensors. ICARUS then correctly moves some fraction of those points into the middle troposphere to improve the comparisons

with the satellite products. Interestingly, while the normal deviation is slightly larger, the linear correlation coefficient of τ_{ICARUS} with τ_{ISCCP} and τ_{ICARUS} with τ_{LBTM} is nearly identical to that found comparing τ_{LBTM} with τ_{ISCCP} . This suggests that the alterations of cloud-top pressure performed by ICARUS are performing as well as could be expected. We do identify a small negative bias in both comparisons where τ_{ICARUS} is on average 11 mb less than τ_{LBTM} and τ_{ICARUS} is 16 mb less than τ_{ISCCP} . This negative bias is not found in comparisons with τ_{LBTM} in either case. As expected then, comparing τ_{LBTM} with τ_{ISCCP} reveals this bias more clearly.

We compare the various renditions of τ_{ICARUS} in Figure 2. As in Figure 1, the comparison between LBTM and ISCCP shows minimal bias. The comparison between τ_{ICARUS} and τ_{LBTM} show good agreement with a slight positive bias that seems to be associated with higher optical depth events. Comparing the ground-based techniques to the satellite techniques, however, reveals a bias with the satellite retrievals of τ_{ICARUS} on average 10% lower than the ground-based quantities.

τ_{ICARUS} - τ_{LBTM} histograms are shown in Figure 3 that correspond to the standard 9 ISCCP cloud types. As before, we find that ICARUS seems to bring the ground-based observations into closer agreement with ISCCP. The biggest change in the observations appears in the high and middle cloud categories where ICARUS adjusts downward into the middle troposphere many high clouds. Similarly large changes can be seen in the comparisons with LBTM. However, in comparing the τ_{LBTM} - τ_{ISCCP} statistics with the satellites in other categories we find interesting differences. Specifically, the τ_{LBTM} - τ_{ISCCP} histograms show that the frequency of optically thick clouds are greater than the satellite algorithms while the frequency of optically intermediate clouds are less than the satellite algorithms in approximately equal proportions. A discrepancy that is common between the two satellite algorithms and the surface is that both LBTM and ISCCP diagnose less than half of the optically thick lower tropospheric cloud type (hereafter

referred to as stratus) compared to the surface results while just the opposite is found in the optically intermediate middle troposphere cloud type (hereafter altostratus). The ISCCP and ICARUS results seem to agree in their frequency of optically thick high clouds (hereafter deep clouds). However, LBTM reports substantially fewer of these deep layers but diagnoses proportionally more of the optically thick midlevel clouds referred to commonly as nimbostratus.

Differences in the comparison of cloud type occurrence statistics can arise from errors in the determination of cloud top pressure and in the diagnosis of column optical depth or from simultaneous errors in both of these quantities. For the ground-based results while we are confident in the measurement of the vertical distribution of cloud layers by the cloud radar and micropulse lidar, uncertainties in τ_{total} and $\tau_{\text{total}}^{\text{ground}}$ arise from errors in the microphysical retrieval algorithms that impact the derived vertical profile of optical depth. This source of error needs to be uncoupled from the uncertainty due to the radiative parameterization in the ICARUS algorithm to determine cloud top pressure. To accomplish this we use MODTRAN as described earlier to derive $\tau_{\text{total}}^{\text{MODTRAN}}$ independent of the ICARUS parameterization and we use $\tau_{\text{total}}^{\text{ground}}$ as an independent measure of the ground-based total column optical depth. Figures 3e-h can be compared to get an initial impression of the impact of the uncertainties in the ground-based results. In Figures 3e and 3f where we compare $\tau_{\text{total}}^{\text{MODTRAN}}$ with $\tau_{\text{total}}^{\text{ground}}$ we find, with the exception of the deep and cirrostratus categories, the agreement to be quite close as would be expected from Figure 2. The discrepancy in the cirrostratus and deep categories can be explained as a potential low bias that arises from the MFRSR algorithm in thick clouds with high ice water paths above liquid water layers. The MFRSR optical depth algorithm has no specific information regarding the properties of the ice phase (all condensate in the thick layers is assumed to be water) except to the extent that it reduces the spectral flux at the surface. In

comparing Figures 3e and 3h and Figures 3f and 3g allows us to evaluate  with , and the differences seem small.

Examination of the panels in Figure 3 is instructive. However, one must be cautious not to place too much stock in the quantitative agreement in Figure 3 because there is potential for compensating errors that adjust the counts upward in a particular category that depend on factors unrelated to the agreement between the ground-based and satellite algorithms in that category. To illustrate this point we list in Table 3 the fraction of cases when ISCCP or LBTM diagnose a cloud class that  and  fall within that class. Tables 3c and 3f show the fraction of the number of cases in Table 3a for which ARM and ISCCP or LBTM agree for a particular type without application of the ICARUS algorithm. Tables 3d and 3g, then, illustrates the effect of the ICARUS cloud top pressure corrections. Table 3b illustrates the agreement among the two satellite algorithms. We find the agreement to range from approximately $\frac{1}{2}$ to $\frac{3}{4}$ of cases in most categories with the exception of the stratus, cumulus and altocumulus classes. The small number of cases of altocumulus and cumulus make the interpretation of these results problematic. It is clear, however, that LBTM and ISCCP diagnose stratus clouds differently. 40% of the time that ISCCP diagnoses stratus, LBTM diagnoses nimbostratus suggesting that under these circumstances the interpretation of cloud top pressure is the issue.

We find that when ISCCP or LBTM diagnoses a high cloud, the ICARUS algorithm has little effect and actually acts to reduce the agreement in the cirrostratus and deep categories. This can be understood by considering that the role of ICARUS is to move the cloud top pressure downward in altitude to higher cloud top pressure values from its physical location to match the pressure of the column radiating temperature. ICARUS would not simulate the cloud top pressure to be at lower pressures than it already is physically determined to be. The decrease in

agreement in the cirrostratus and the deep categories are due to the presence of thin cirrus layers overlying thicker layers where ICARUS adjusts downward the cloud top pressure so that the event is counted in the adjacent cloud top pressure bin. While we also find that ICARUS has little influence on the optically thick stratus and stratocumulus agreement statistics, ICARUS does seem to successfully improve the altostratus and nimbostratus agreement, perhaps due to the upward shift in altitude for low clouds under inversions.

The real question is why the overall percentage agreements in Tables 3c and 3e are so small. One could argue, perhaps, that we should not expect the ground-based ICARUS results to agree any better than the two satellite algorithms agree. However, even with that criterion, we find in most cases that the agreement between ICARUS and the satellite results to be smaller. On the other hand, the agreement with ICARUS applied to ARM observations is about 1-2% greater than is the agreement without applying ICARUS to observations, demonstrating that ICARUS does not degrade the comparison between ARM observations and satellite algorithms. We are reasonably certain that the vertical distribution of cloud occurrence in the ARM data is as correct as it could be given a continuously operating millimeter radar and microwave radiometer and other ancillary data used as input to the algorithms. We have established by comparing with MFRSR and in previous publications that the retrieved ARM radiative property profile is largely unbiased and agrees reasonably with downwelling fluxes at the surface, and at the top of the atmosphere (Mace et al., 2008), and that the column optical depth agrees within reasonable uncertainties with independently derived column optical depth from the MFRSR. We have also established that the ICARUS radiative parameterization is in reasonable agreement with similar quantities calculated from a more complicated radiative model, and that the differences in the combined - with - are much smaller for cloud types other than cirrostratus

and deep than the differences in the any of the ground-based results with either of the satellite results.

To help shed light on this issue and examine more closely the differences between the algorithms, we consider the $\frac{N_{\text{ICARUS}}}{N_{\text{ISCCP}}}$ statistics from ISCCP and LBTM when the ICARUS algorithm diagnoses clouds in each of the nine categories (Figures 4 and 5). We present similar distributions in Figure 6 for $\frac{N_{\text{ICARUS}}}{N_{\text{ISCCP}}}$ in order to understand the uncertainty in $\frac{N_{\text{ICARUS}}}{N_{\text{ISCCP}}}$ due to the ICARUS radiative parameterization. The depiction of the statistics in Figures 4-6 is the converse approach as taken in Table 3 and more closely represents the methodology of a model evaluation where the model output would be converted into an ISCCP equivalent using ICARUS. In other words, if a model were to predict the cloud occurrence and properties to within a similar uncertainty as the ground-based results, Figures 4 and 5 show the level of agreement that could be expected with ISCCP and LBTM.

Beginning with the high clouds and moving from optically thickest to thinnest, we find that when ICARUS simulates a deep cloud, 62% of the time this diagnosis will agree with ISCCP (we refer to this as the hit rate: i.e. $\frac{N_{\text{ICARUS}}}{N_{\text{ISCCP}}}=0.62$). Of the 38% of the time ICARUS incorrectly places a cloud in the deep category (we will refer to this as the miss rate: i.e. $\frac{N_{\text{ISCCP}}}{N_{\text{ICARUS}}}=0.38$), ISCCP will report the majority of those in the cirrostratus category. A similar pattern is found with LBTM except that $\frac{N_{\text{ISCCP}}}{N_{\text{LBTM}}}=0.49$. On the other hand, we find that of the deep cloud cases, $\frac{N_{\text{LBTM}}}{N_{\text{ISCCP}}}=0.1$. For the cirrostratus category, $\frac{N_{\text{LBTM}}}{N_{\text{ISCCP}}}=0.53$ and $\frac{N_{\text{ISCCP}}}{N_{\text{LBTM}}}=0.43$. ISCCP and LBTM diagnose a larger τ in about the same fraction of cases although ISCCP places more events into a larger τ bin while LBTM places more into a smaller τ bin. $\frac{N_{\text{LBTM}}}{N_{\text{ISCCP}}}$ is significantly smaller than $\frac{N_{\text{ISCCP}}}{N_{\text{LBTM}}}$ with both algorithms placing the misses at larger τ than at larger τ . For the high cloud category, we find that the MODTRAN results show that the

ICARUS parameterization places the clouds correctly more than 90% of the time. So, the misses in the high cloud category are due to differences in interpretation of $\tau_{0.5}$.

In the middle $\tau_{0.5}$ classes, we see significant differences in skill from high to low $\tau_{0.5}$ with the two satellite algorithms showing very similar hit rates. For the thicker classes at mid levels, $\tau_{0.5} > 1.5$ with the majority of misses being placed at smaller $\tau_{0.5}$ for the nimbostratus category and at larger $\tau_{0.5}$ for the altostratus. While there are only 10-12 cases for the altocumulus category, the miss rate seems quite high – much higher than for optically thin cirrus. ½ of the misses (80% for LBTM) are being diagnosed to have larger $\tau_{0.5}$ while ISCCP diagnoses many of these events as cirrus. The comparison between MODTRAN and ICARUS shows that more uncertainty exists in the ICARUS parameterization with about 15% of the occurrences being placed in the wrong $\tau_{0.5}$ category for nimbostratus and altostratus while this uncertainty rises to 40% being placed in the cirrus category for the altocumulus class.

It is surprising that the agreement is not better between the $\tau_{0.5} > 1.5$ events and the satellite products in the largest $\tau_{0.5}$ classes. We find that $\tau_{0.5} > 1.5$ have about the same hit rate with the ISCCP and LBTM in the stratus and stratocumulus classes being on the order of 30% for stratus and 50% for stratocumulus. One would expect that $\tau_{0.5} > 1.5$ given that the larger $\tau_{0.5}$ would allow for a more accurate determination of $\tau_{0.5}$ although uncertainties associated with surface inversions have been shown to cause these classes of cloud to be erroneously placed in the mid levels by the satellite algorithms. We note that the version of ICARUS that we are using has been specifically modified to correctly place $\tau_{0.5}$ in inversion situations. In the cumulus category, ISCCP diagnoses a higher $\tau_{0.5}$ about half the time (100% for LBTM) and a smaller $\tau_{0.5}$ about 60% of the time (40% of the time for LBTM). The MODTRAN results show

that the accuracy of ICARUS to correctly diagnose τ_{top} decreases from 90% in the stratus category to 56% in the cumulus category.

4. Summary and Discussion

The ISCCP simulator has gained wide usage across the community although it has not been extensively validated. This freely available fortran code is designed to take as input cloud properties and thermodynamics profiles simulated by models and produce τ_{top} and τ_{mid} that would be diagnosed by ISCCP given an atmospheric column with similar cloud properties. Because ISCCP is limited to two pieces of information (the upwelling IR radiance and the visible reflectance) regarding the cloudy atmosphere, simulators of ISCCP are necessary to compare the expansive multi-decade ISCCP data set with model simulations of the recent climate. Such comparisons are critical to improving the representation of clouds in GCMs.

Overall, we find that the ICARUS portion of the ISCCP simulator improves agreement between cloud-top pressures from ISCCP and LBTM with simulated characterizations of the column physical and radiative properties like what would be produced by an atmospheric model (Figure 1). This improvement is accomplished by adjusting some portion of high-topped and low-topped clouds into the middle-troposphere by considering what a satellite radiometer would derive from shortwave and thermal IR radiances.

A careful qualification of this statement needs to be made, however. While comparison of the τ_{top} - τ_{mid} histograms in Figure 3 show patterns that would seem to support several of the findings from previous model-ISCCP intercomparison studies, a more careful analysis revealed that these similarities are as much due to compensating errors as to the ability of ICARUS to

simulate ISCCP or LBTM. While there are uncertainties in the ground-based characterizations of the column radiative properties, these uncertainties appear, based on comparison with independent results, to not be the primary source of this discrepancy.

Overall we find that were a model to predict the actual occurrence of clouds with the same accuracy as a cloud radar and then the model made reasonable diagnostic interpretations of the column radiative properties, agreement with satellite derived $\overline{P} - \overline{\tau}$ after applying ICARUS would be successful in approximately $\frac{1}{2}$ to $\frac{2}{3}$ of cases depending on the cloud type. Here, success is defined by ICARUS placing the simulated column into the same $\overline{P} - \overline{\tau}$ bin of the nine bins typically used for such comparisons. Note, that applying ICARUS to ARM observations results in only slightly better agreement overall ($\sim 2\%$) between ground-based and satellite $\overline{P} - \overline{\tau}$ than not applying ICARUS to ARM observations.

In addition to uncertainties in the derived column radiative properties, uncertainties in the comparison of the ICARUS-modified ground-based results and the satellite results could arise from errors in the parameterization of $\overline{\tau}$, and from an omission in ICARUS to simulate the total visible optical depth that a satellite algorithm would produce. To test the former, we bypassed the ICARUS radiative parameterization of $\overline{\tau}$ using the MODTRAN radiative model. We found that ICARUS tends to accurately parameterize estimates of $\overline{\tau}$ in high clouds and in stratus more than 90% of the time. The uncertainty in the parameterized $\overline{\tau}$ increases as the optical depth of the condensate decreases for middle and low clouds where the errors are on the order of 15% for the nimbostratus, altostratus, and stratocumulus cloud classes. The errors seemed to be larger for altocumulus and cumulus although the number of events in these optically thin categories is small due to our method for selecting candidate cases. Different comparison strategies between

ground-based and satellites are necessary for altocumulus and cumulus cloud types as well as broken cloud fields.

In most of the cloud type classes, errors in $\bar{\tau}$ appear to be the dominant source of discrepancy between the ground-based and satellite-derived results. The general pattern we find is that the higher optical depth classes are often diagnosed by the satellite algorithms to have lower optical depth than derived by either of the ground-based sources and that the optically thin cloud classes are often diagnosed by the satellite algorithms to have higher optical depths than derived from the ground-based sources. Such discrepancies are not new. Min and Harrison (1996) as well as Barker et al. (1998) find that ISCCP and LBTM optical depths are biased low in comparison to optical depths derived from ground-based data. While uncertainties in scattering phase function for ice clouds is one source of error, optical depth retrievals are also sensitive to assumptions regarding bidirectional reflectance and surface albedo in thin clouds. For thicker clouds, the simple fact that the reflected radiances asymptotes beyond optical depths of 10-20 results in increasing uncertainties that are not well documented as shown by Min and Harrison (1996).

The pattern of error we find in this study is consistent with errors that would be expected from satellite τ_{eff} retrieval algorithms that use plane parallel radiative transfer theory when those algorithms are applied to heterogeneous cloud fields with certain properties. These biases occur because solar photons tend to migrate horizontally through multiple scattering and emerge from the cloud field in optically thinner regions causing optically thick regions of the cloud field to appear less reflective (have lower τ_{eff}) and thin regions of the cloud field to appear more reflective (have larger τ_{eff}). If the scale of the plane parallel satellite retrieval is larger than a characteristic radiative smoothing scale, then the effects of horizontal photon transport on the τ_{eff} retrieval tend

to be small. On the other hand, if the scale of the satellite retrieval is smaller than a characteristic smoothing scale, then errors in retrieved τ can occur. Heidinger and Stephens (2002) show, consistent with the earlier findings of Marchak et al., (1995) and Davis et al., (1997), that high optical depth clouds would tend to be interpreted to have smaller τ than reality and that lower optical depth cloud fields would tend to be diagnosed to have higher than actual τ . The exact sign and magnitude of this error will depend on the properties of the cloud field and the spatial size of the satellite pixel retrieval.

Davis et al. (1997) present a convenient measure of this smoothing scale that is valid when the diffusion approximation holds or when $\frac{L}{\lambda} \ll 1$ where g is the asymmetry parameter. Davis et al. (1997) show that this smoothing scale can be expressed,

$\frac{L}{\lambda} \ll 1$ where L is the geometric thickness of the cloud and τ is extinction. So when the scale of the plane parallel satellite pixel retrieval is smaller than $\frac{L}{\tau}$, plane parallel theory may not hold depending on the variability in the cloud field. This theory has been well

developed for stratocumulus clouds and found to be approximately on the order of the physical depth of the cloud field or several hundred meters. When the cloud field is more vertically

extended and optically diffuse, $\frac{L}{\tau}$ can actually become larger than the approximate scale of the LBTM and ISCCP pixels, and when $\frac{L}{\tau}$ increases much beyond 1 km the satellite algorithms should become more sensitive to the heterogeneity of the cloud field. While it is beyond the

scope of this study to exhaustively evaluate the effects of 3D radiative transfer on the satellite retrieved τ , we do examine in Figure 7 comparisons of ground-based and LBTM $\bar{\tau}$ for $\frac{L}{\tau} < 1$ and $\frac{L}{\tau} > 3$ for the pixel-level retrieval nearest the SGP site. Based on our simple theoretical

discussion we expect that any optical depth bias would be smaller for $\frac{L}{\tau} < 1$ than for $\frac{L}{\tau} > 3$ and that is indeed what we find for both satellite-based retrievals when compared to the ground-based

retrievals. While the results in Figure 7 are by no means conclusive, these results are strongly suggestive that biases in the satellite-derived τ not accounted for in the ISCCP simulator may exist in the ISCCP and LBTM products.

The convolution of uncertainties in \bar{P} and $\bar{\tau}$ contrive to cause significant uncertainty in comparing statistics derived from model output with similar statistics derived from satellite radiances. Based on these and earlier findings we recommend that a systematic study of potential errors in visible optical depth be undertaken for ISCCP, LBTM, and the ground-based techniques so that corrections can be made as appropriate and/or the ISCCP simulator can be modified to account for any potential biases in $\bar{\tau}$ that do exist. Finally, we conclude that comparisons of optical depth made between ISCCP and similar algorithms with GCM results whether or not they have been modified to simulate ISCCP with the ISCCP simulator should be viewed with some caution until additional validation of satellite cloud optical depth retrievals can be provided.

References

- Barker, et al., 1998: Optical Depth of Overcast Cloud across Canada: Estimates Based on Surface Pyranometer and Satellite Measurements. *Journal of Climate*, **11**, 2908-2994.
- Berk, A. et al., 1989: MODTRAN: A Moderate Resolution Model for LOWTRAN 7, GL-TR-89-0122, Geophysics Directorate, Phillips Laboratory, Hanscom AFB, MA 01731.
- Dufresne, J.-L. and S. Bony, 2008: An Assessment of the Primary Sources of Spread of Global Warming Estimates from Coupled Ocean-Atmosphere Models. *Journal of Climate*, **21**, 5135-5144.
- Davis, A., A. Marshak, R. Cahalan, and W. Wiscombe, 1997: The Landsat scale break in stratocumulus as a three-dimensional radiative transfer effect: implications for cloud remote sensing. *Journal of Atmospheric Sciences*, **54**, 241-260.
- Heidinger, A. K., and G. L. Stephens, 2002: Molecular line absorption in a scattering atmosphere. Part III: Pathlength characteristics and effects of spatially heterogeneous clouds. *J. Atmos. Sciences*, **59**, 1641-1653.
- Klein, S. A. and C. Jakob, 1999: Validation and Sensitivities of Frontal Clouds Simulated by the ECMWF Model. *Monthly Weather Review*, **127**, 2514-2531.
- Mace et al., 2006: Cloud Radiative Forcing at the Atmospheric Radiation Measurement Program Climate Research Facility: 1. Technique, Validation, and Comparison to Satellite-Derived Diagnostic Quantities. *Journal of Geophysical Research*, 111.
- Mace, G. G., and S. Benson, 2008: The Vertical Structure of Cloud Occurrence and Radiative Forcing at the SGP ARM Site as Revealed by 8 Years of Continuous Data. *Journal of Climate*, **21**, 2591-2610.
- Marshak, A., A. Davis, W. J. Wiscombe, and R. F. Cahalan, 1995: Radiative smoothing in fractal clouds. *J. Geophys. Res.*, **100**, 26247-26261.
- Min, Q., and L. C. Harrison, 1996: Cloud Properties Derived from Surface MFRSR Measurements and Comparison with GOES Results at the ARM SGP Site. *Geophysical Research Letters*, **23**, 1641-1644.
- Minnis, P. et al., 1995: Cloud Properties Derived From GOES-7 for the Spring 1994 ARM Intensive Observing Period Using Version 1.0.0 of the ARM Satellite Data Analysis Program. *NASA Reference Publication*, 1366, 59pp.
- Randall et al., 2003: Breaking the Cloud Parameterization Deadlock. *Bulletin of the American Meteorological Society*, 1547-1564.
- Rossow, W. B., A. W. Walker, D. E. Beusichel, and M. D. Roiter, 1996: International Satellite Cloud Climatology Project (ISCCP) Documentation of New Cloud Datasets. WMO/TD-No. 737, World Meteorological Organization, 115 pp.
- Schiffer, R. A., and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The First Project of the World Climate Research Programme. *Bulletin of the American Meteorological Society*, **64**, No 7.

- Soden, and I.M. Held, 2006: An Assessment of Climate Feedbacks in Coupled Ocean Atmosphere models. *Journal of Climate*, **19**, 3354-3360.
- Webb M. et al., 2001: Combining ERBE and ISCCP Data to Assess Clouds in the Hadley Centre, ECMWF and LMD Atmospheric Climate Models. *Climate Dynamics*, **17**, 605-922.
- Williams, K. D., and G. Tselioudis, 2007: GCM Intercomparison of Global Cloud Regimes: Present-Day Evaluation and Climate Change Response. *Journal of Climate Dynamics*, **29**, 231-250.
- Zhang et al., 2005: Comparing Clouds and Their Seasonal Variations in 10 Atmospheric General Circulation Models with Satellite Measurements. *Journal of Geophysical Research*, **110**.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

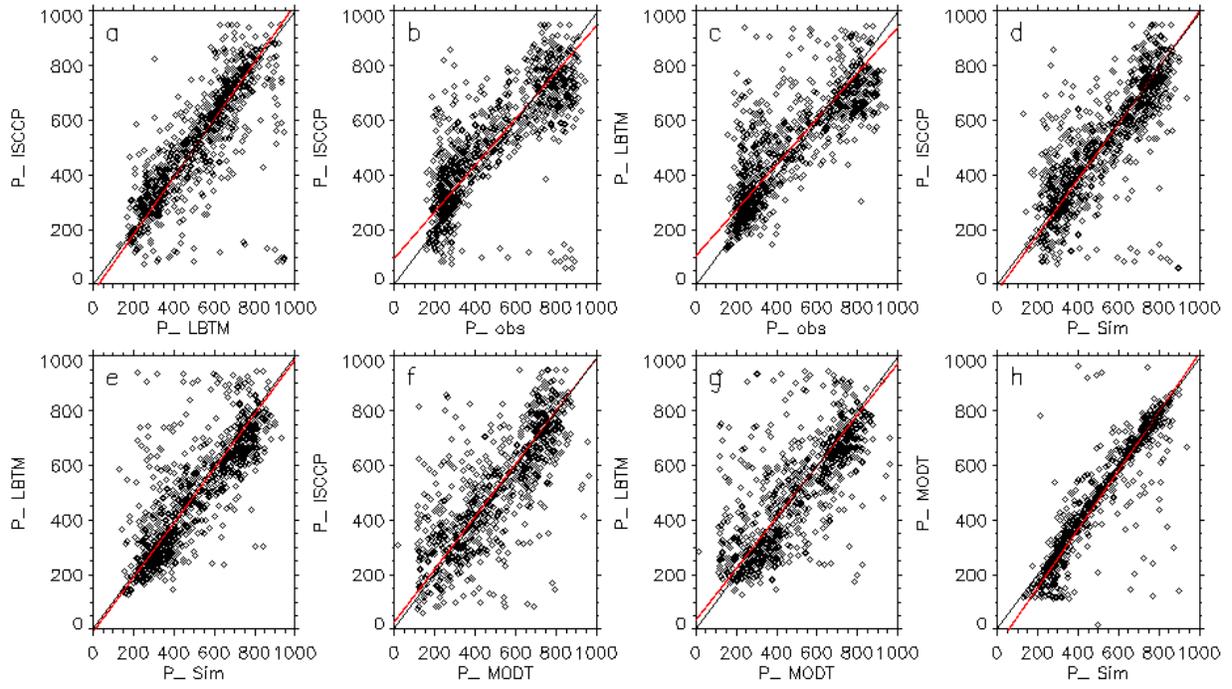


Figure 1. Comparisons of cloud top pressure (mb) between (a) and , (b) and , (c) and , (d) and , (e) and , (f) and , (g) and , (h) and .

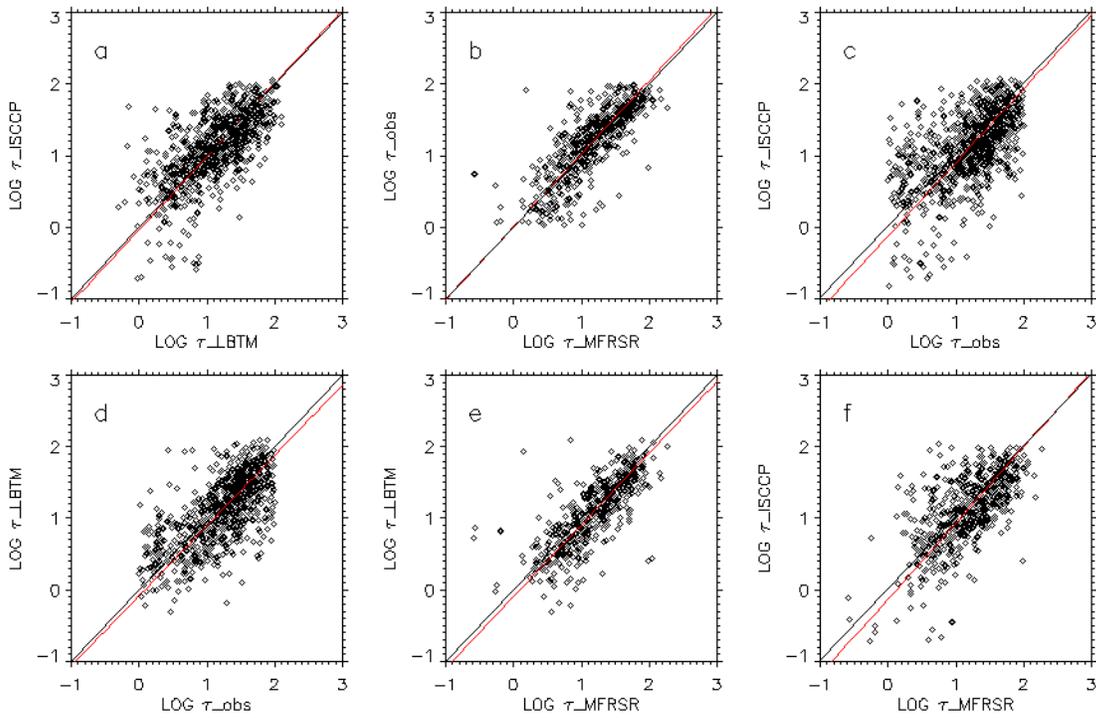


Figure 2. Comparisons of the base 10 logarithm of total optical depth with τ_{JISCCP} capped at 100. The red line in each plot is a linear regression and the black line is 1:1. (a) τ_{JISCCP} - τ_{LBTM} , (b) τ_{obs} - τ_{MFRSR} , (c) τ_{JISCCP} - τ_{obs} , (d) τ_{LBTM} - τ_{obs} , (e) τ_{LBTM} - τ_{MFRSR} , (f) τ_{JISCCP} - τ_{MFRSR}

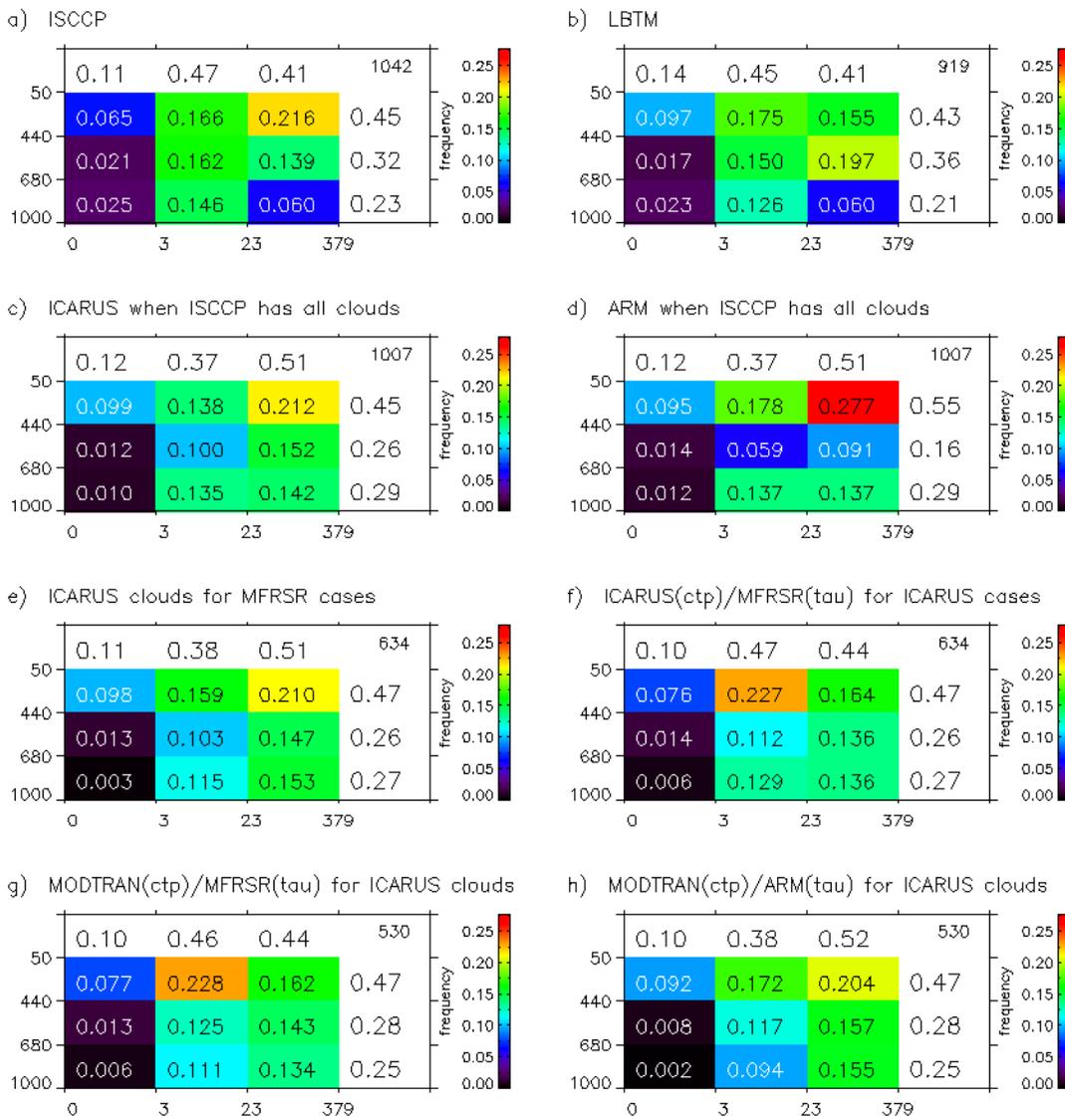


Figure 3.  histograms for the 9 ISCCP cloud type classes with the numerical fraction of the total number of cases (listed in the upper right corner of each plot). Coverage is between 1997-2002 and the events meet the criteria listed in Section 2. The fractions in the right-most column is a summation of the fractional occurrence in each optical depth class. The fractions across the top are summations of the fractions in each cloud top pressure class. (a) ISCCP (b) LBTM, (c) ICARUS applied to ARM events, (d) ARM events before application of ICARUS, (e) As in (c) except for the events with MFRSR measurements, (f) cloud top pressure from ICARUS and optical depth from MFRSR, (g) Cloud top pressure from MODTRAN and optical depth from MFRSR, (h) cloud top pressure from MODTRAN and optical depth from ARM.

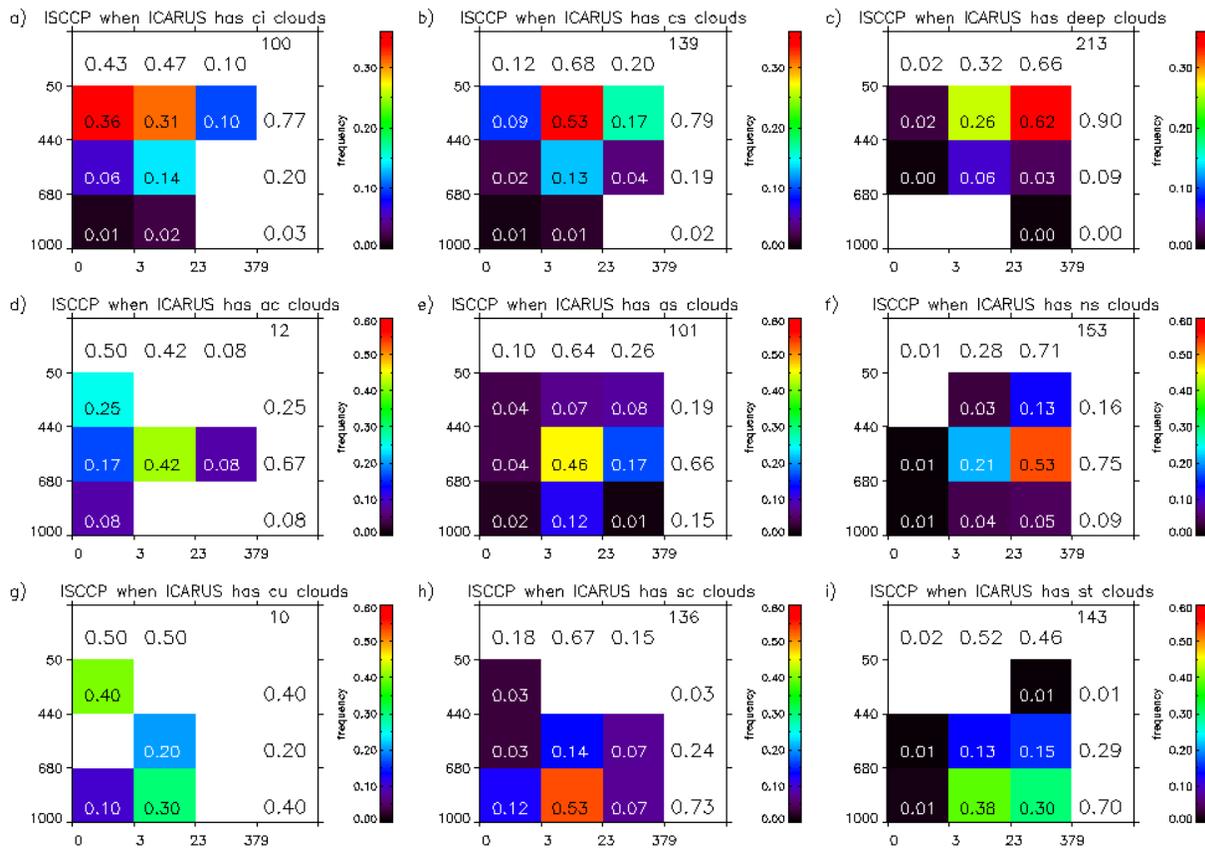


Figure 4. The distribution of $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed in each of the 9 cloud class bins. Each histogram is as described in Figure 3. a) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as cirrus ($\text{ISCCP}_{\text{bin}} < 440$ hPa and $\text{ICARUS}_{\text{bin}} < 3.6$), b) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as cirrostratus ($\text{ISCCP}_{\text{bin}} < 440$ hPa and $3.6 > \text{ICARUS}_{\text{bin}} > 23$), c) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as deep clouds ($\text{ISCCP}_{\text{bin}} < 440$ hPa and $\text{ICARUS}_{\text{bin}} > 23$), d) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as altocumulus (680 hPa $< \text{ISCCP}_{\text{bin}} < 440$ hPa and $\text{ICARUS}_{\text{bin}} < 3.6$), e) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as altostratus (680 hPa $< \text{ISCCP}_{\text{bin}} < 440$ hPa and $3.6 > \text{ICARUS}_{\text{bin}} > 23$), (f) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as nimbostratus (680 hPa $< \text{ISCCP}_{\text{bin}} < 440$ hPa and $\text{ICARUS}_{\text{bin}} > 23$), g) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as cumulus ($\text{ISCCP}_{\text{bin}} > 680$ hPa and $\text{ICARUS}_{\text{bin}} < 3.6$), (h) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as stratocumulus ($\text{ISCCP}_{\text{bin}} > 680$ hPa and $3.6 > \text{ICARUS}_{\text{bin}} > 23$), (i) $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ when $\text{ISCCP}_{\text{bin}}$ - $\text{ICARUS}_{\text{bin}}$ is diagnosed as stratus ($\text{ISCCP}_{\text{bin}} > 680$ hPa and $\text{ICARUS}_{\text{bin}} > 23$).

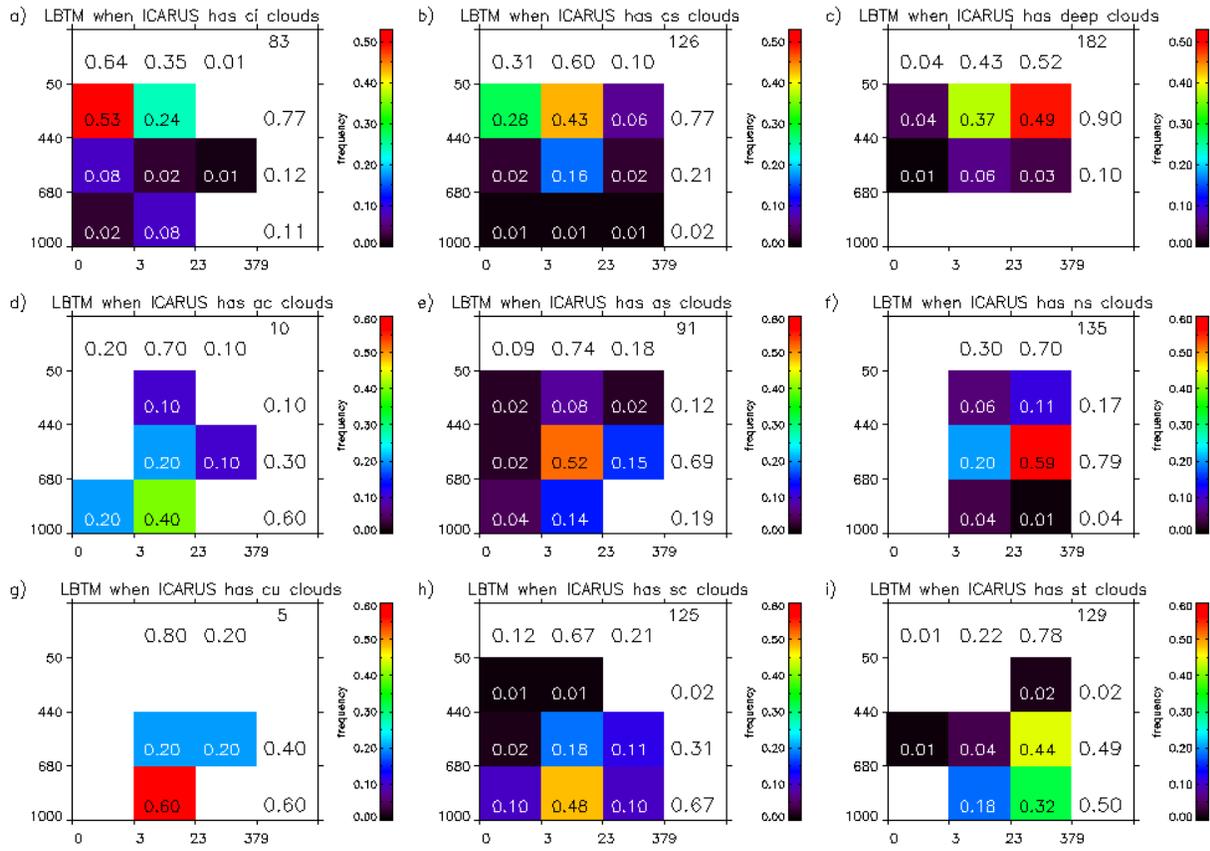


Figure 5. As in Figure 4 except for LBTM.

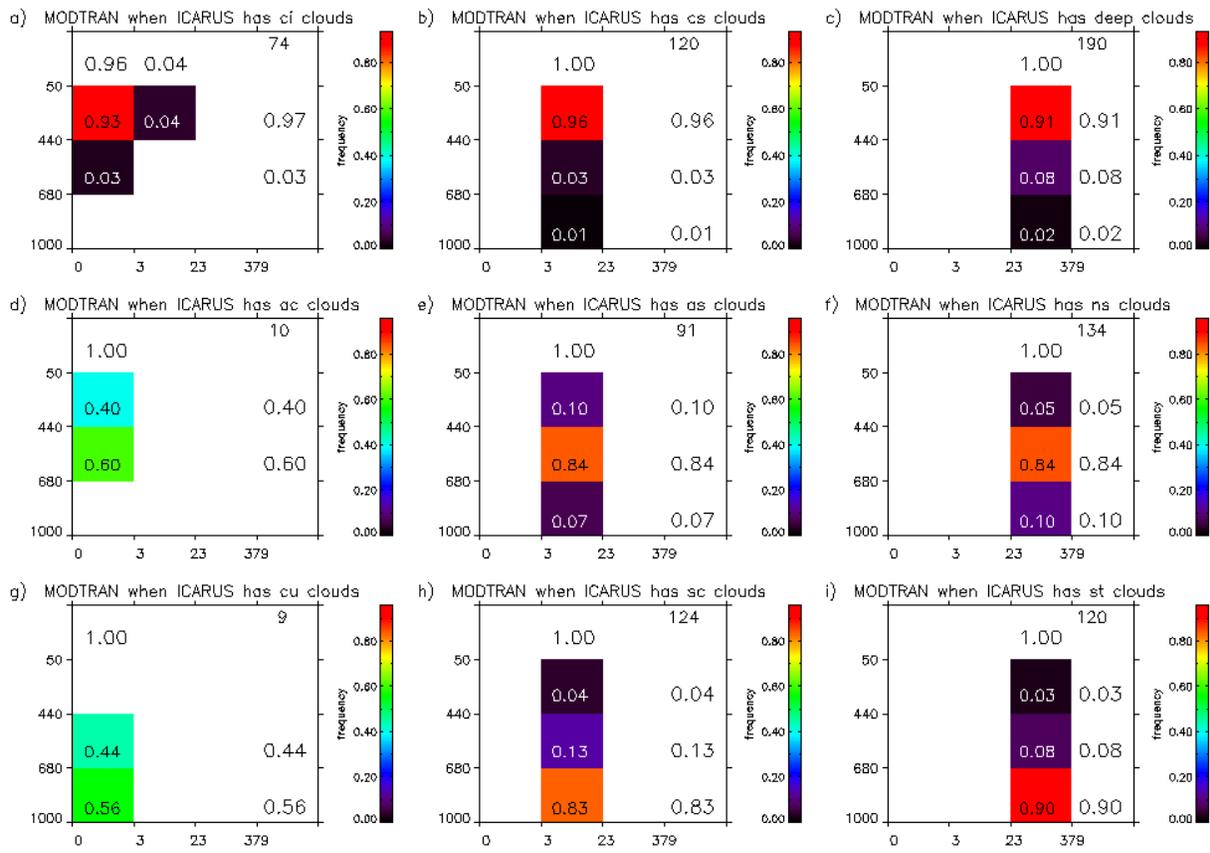


Figure 6. As in Figure 3 except for .

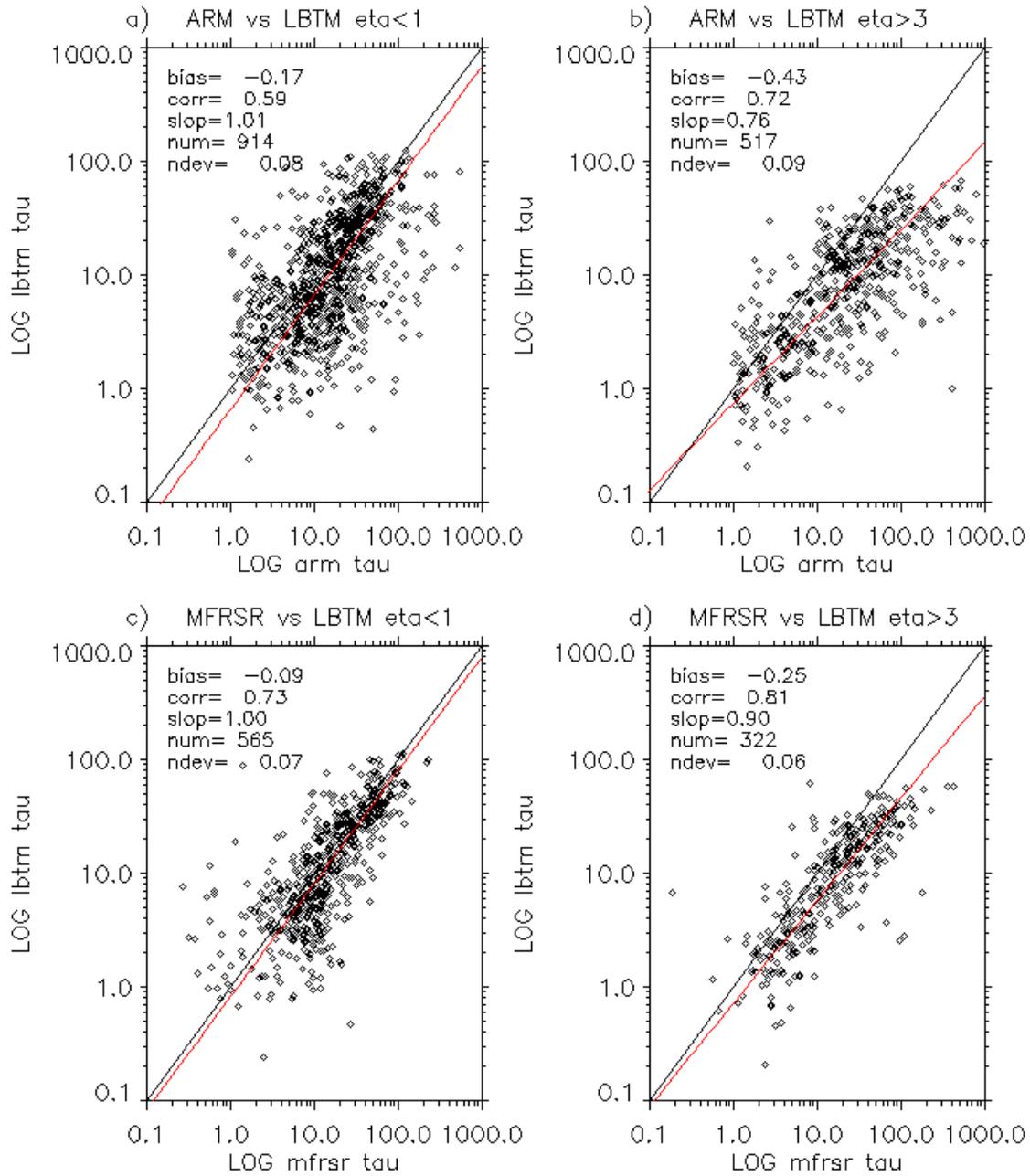


Figure 7. Comparison of τ_{aerosol} for (a) LBTM with ARM ground-based for $\eta < 1$, (b) LBTM with ARM ground-based for $\eta > 3$, (c) LBTM with MFRSR ground-based for $\eta < 1$, (d) LBTM with MFRSR ground-based for $\eta > 3$.

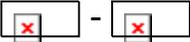
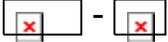
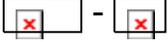
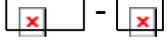
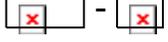
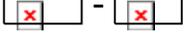
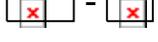
Comparison	Num	Bias	Linear Correlation	Linear Slope	Normal Deviation
 - 	1000	-0.57	0.80	1.06	19.06
 - 	1042	21.94	0.78	0.88	50.17
 - 	919	28.6	0.81	0.84	46.74
 - 	1042	-16.90	0.80	1.05	26.57
 - 	919	-11.01	0.81	1.00	25.48
 - 	900	-1.27	0.75	1.00	17.29
 - 	809	4.31	0.74	0.94	24.18
 - 	900	-16.96	0.89	1.06	20.01

Table 1. Statistics of the cloud top pressure comparisons seen in Figure 1. All quantities are shown in mb except for number of events.

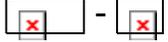
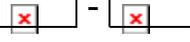
Comparison	Num	Bias	Linear Correlation	Linear Slope	Normal Deviation
 - 	789	-0.01	0.67	1.02	0.07
 - 	555	0.03	0.79	1.02	0.08
 - 	891	-0.09	0.59	1.03	0.10
 - 	789	-0.10	0.68	0.98	0.06
 - 	492	-0.09	0.75	1.00	0.04
 - 	555	-0.05	0.64	1.06	0.06

Table 2. Statistics of the optical depth comparisons seen in Figure 2.

Table 3. Evaluation of the agreement statistics when ISCCP (a-d) and LBTM (e-g) diagnose a particular cloud type. a) number of ISCCP cases, b) the fraction of the ISCCP cases where $\square_x - \square_x$ are in the same class as ISCCP. c) as in b) except $\square_x - \square_x$. d) as in b) except $\square_x - \square_x$. The percentages in parentheses in table 3d show $\square_x - \square_x$. e) as in a) except LBTM. f) as in c) except LBTM and g) as in d) except LBTM.

a. ISCCP # of Cases	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	68	173	225
$680 < \square_x < 440$	22	169	145
$\square_x > 680$	26	152	62

b. LBTM % Agree	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	63	53	57
$680 < \square_x < 440$	16	47	73
$\square_x > 680$	24	49	36

c. ARM % Agree	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	53	47	79
$680 < \square_x < 440$	14	18	37
$\square_x > 680$	8	47	66

d. ICARUS % Agree	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	53 (57)	43 (47)	69 (65)
$680 < \square_x < 440$	9 (0)	27 (26)	57 (52)
$\square_x > 680$	4 (4)	47 (44)	69 (72)

e. LBTM # of Cases	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	89	161	116
$680 < \square_x < 440$	16	138	181
$\square_x > 680$	21	116	55

f. ARM % Agree	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	50	37	87
$680 < \square_x < 440$	0	21	30
$\square_x > 680$	10	53	75

g. ICARUS % Agree	$\square_x < 3.6$	$3.6 < \square_x < 23$	$\square_x > 23$
$\square_x < 440$	50	34	77
$680 < \square_x < 440$	0	34	45
$\square_x > 680$	0	51	75

