Estimation of Liquid Cloud Properties that Conserve Total-Scene Reflectance Using Satellite Measurements

MICHAEL J. FOSTER
Space Science and Engineering Center, Department of Atmospheric and Oceanic Sciences, University of Wisconsin—Madison, Madison, Wisconsin

RALF BENNARTZ
Department of Atmospheric and Oceanic Sciences, University of Wisconsin—Madison, Madison, Wisconsin

ANDREW HEIDINGER
NOAA/NESDIS/Center for Satellite Applications and Research, and University of Wisconsin—Madison, Madison, Wisconsin

(Manuscript received 27 December 2009, in final form 23 July 2010)

ABSTRACT

A new method of deriving statistical moments related to the distribution of liquid water path over partially cloudy scenes is tested using a satellite cloud climatology. The method improves the ability to reconstruct total-scene visible reflectance when compared with an approach that relies on valid liquid water path retrievals, and thus it maintains physical consistency with the primary satellite observations when deriving cloud climatologies. A global application of the new method finds a mean bias of $0.008 \pm 0.017$ when reconstructing total-scene reflectance from liquid water path distributions, as compared with a bias of $0.05 \pm 0.047$ when using a conventional approach. Application of the method to a multidecadal cloud climatology suggests that this may provide a means of identifying data artifacts that could affect long-term cloud property trends. The conservation of reflectance plus the ease of applicability to various satellite datasets makes this method a valuable tool for model validation and comparison of satellite climatologies. Gaussian and gamma functions are used to approximate the distribution of horizontal subgrid-scale liquid water path for $1^\circ \times 1^\circ$ scenes, and while both functions perform well for the majority of atmospheric conditions, it is found that the Gaussian distribution generates a negative bias for cases in which visible reflectance is very high and that neither function is able to represent liquid water path well in the few cases in which the observed distribution is bi- or multimodal.

1. Introduction

Variability in climate sensitivity among general circulation models (GCMs) continues to be dominated by differences in cloud feedback processes, and the largest source of uncertainty in these feedbacks is the shortwave impact caused by changes in boundary layer clouds (Randall et al. 2007). The relatively poor simulation of these clouds may in part be attributed to unresolved subgrid-scale processes. In particular, cloud shortwave albedo is sensitive to in-cloud liquid water variability (Cahalan et al. 1994; Pincus and Klein 2000). The error caused by treating in-cloud microphysical properties as homogenous, often called plane-parallel albedo bias (Cahalan et al. 1994), transfers to atmospheric heating rates, precipitation, and cloud formation and dissipation processes thus contributing to the aforementioned uncertainty.

There have been several studies that attempt to quantify the magnitude of plane-parallel albedo bias generated in model simulations (Barker and Raisanen 2005; Barker et al. 1999; Cahalan et al. 1994; Oreopoulos et al. 2004; Raisanen et al. 2004). The results of these studies vary, in part because they focus on specific cloud scenes generated by cloud-resolving models (CRMs) or large-eddy simulation (LES) models. Other studies have examined plane-parallel albedo bias from observed as opposed to model-generated distributions of cloud properties (Barker et al. 1996; Oreopoulos and Davies 1998a,b; Pincus et al. 1999; Raisanen et al. 2003; Oreopoulos et al. 2007). The most common approach to quantifying this
bias has been to apply an independent pixel approximation (IPA), which simply breaks the spatial domain down into individual vertical columns, each with unique in-cloud properties, and averages the results. The primary disadvantage of this method is that there is no horizontal photon transport between the columns, though for studies such as this one, where the clouds being examined are primarily single-layer water clouds, the 3D effects of cloud field configuration are likely minimal. Another issue with the IPA method is that it is not fast enough to be incorporated into an extended GCM simulation, though there have been some recent studies that address this issue with promising results (Oreopoulos and Barker 1999; Shonk and Hogan 2008; Wood et al. 2005).

The various parameterizations of cloud heterogeneity effects rely on the treatment of cloud water as a probability distribution function (PDF) rather than a mean value. The shape of this PDF has, among other functions, been determined by a lognormal distribution (Cahalan et al. 1994; Oreopoulos and Davies 1998a), a Gaussian function (Bechtold et al. 1992; Bechtold et al. 1995), a gamma function (Oreopoulos and Barker 1999), and a beta function (Tompkins 2002). One challenge that these parameterizations face is to determine statistical moments of these PDFs, such as variance, that are physically defensible. This is one advantage of generating cloud water distributions from observational data, as statistical information can easily be derived over a spatial or temporal domain. Barker et al. (1996) derived gamma distributions of optical depth from Landsat imagery, while Heidinger (2003) derived gamma distributions of optical depth using data from the National Oceanic and Atmospheric Administration’s (NOAA) Advanced Very High Resolution Radiometer (AVHRR). In fact the approach proposed here is similar to that of Heidinger (2003) though it has fewer dependencies and takes full advantage of high-level pixel information.

In this study a method is presented that derives statistical information about the horizontal distribution of cloud liquid water path from satellite data while conserving the visible cloud reflectance. There are a number of advantages to this approach; the first being that the primary quantity observed by the satellites is explicitly conserved in the process of generating climatologies thus minimizing screening biases. It will be shown that it is largely insensitive to pixel-level cloud masking. Processed versions of satellite cloud climatologies generate cloud masks that rely on algorithms that sometimes filter pixels that are difficult to classify as either clear or cloudy (Rossow and Garder 1993; Stowe et al. 1999; Frey et al. 2008). This filtering affects the distribution of liquid water path. Our method avoids this issue by generating the liquid water path PDF from all available visible reflectance measurements. This also has the benefit of providing validation data for radiative transfer models (Oreopoulos and Barker 1999) in the form of both cloud and total-scene reflectance.

A second advantage is that this method may be applied to any satellite climatology that provides visible reflectance and liquid water path. In this paper the method is applied to the NOAA AVHRR Pathfinder Atmospheres Extended (PATMOS-x) climatology as well as Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data. A corollary to this is that the method is fast enough to generate a global dataset of liquid water path PDFs that spans several decades for use in GCM validation. In fact we have applied this method to almost 30 years of PATMOS-x data over an area in the North Pacific, and there are plans to implement this method globally in the next processed version of the PATMOS-x climatology.

Last, the availability of statistical information allows us to test the appropriateness of different functions for different cloud types and atmospheric conditions. In this study we examine both a Gaussian and gamma PDF for liquid water clouds over ocean, but the method is scalable to other distribution functions as well as cloud and surface types.

In section 2 we describe the satellite data used and the methods of retrieving liquid water path statistical information. In section 3, we look at the resulting PDF retrievals for MODIS and PATMOS-x and evaluate them using observations. In section 4 the retrieval method is applied globally for one month of PATMOS-x data and then over a small area in the North Pacific for the entire length of the PATMOS-x climatology and the results are examined. The implications and conclusions from this study are discussed in section 5.

2. Data and methodology

a. Data

1) PATMOS-x

The PATMOS-x project derives atmospheric and surface climate records from the AVHRR instrument, which has been flown on the NOAA meteorological satellite series since June 1981. The AVHRR instrument is a 6-channel cross-track scanning radiometer that covers much of the solar and terrestrial infrared spectral range. The visible reflectance channel (channel 1) spans from 0.58 to 0.68 μm and has a spatial resolution (at nadir) of 1.09 km. The near-IR channel (channel 3B) spans from 3.55 to 3.93 μm, and the infrared channel (channel 4) spans from 10.30 to 11.30 μm. Cloud optical depth and near-cloud-top effective radius are retrieved
using a process based on that developed by Nakajima and King (1990) and uses visible and near-IR radiances. Cloud-top temperature is retrieved with the infrared channel. Other variables used in this study include cloud mask and cloud phase. In this study we use pixel-level (version 5) data.

2) MODIS

The MODIS instrument is a 36-channel cross-track scanning radiometer located on the Aqua and Terra satellites, which are part of the National Aeronautics and Space Administration (NASA) Earth Observing System mission. The spatial resolution varies with the spectral channels but the visible reflectance channel (channel 1) spans from 0.62 to 0.67 μm and has a spatial resolution of 250 m. The MODIS cloud product uses visible and near-infrared radiance measurements to retrieve cloud optical depth and near-cloud-top effective radius for cloudy pixels in daytime (King et al. 1997). The cloud mask algorithm (Ackerman et al. 2002) is used to determine which pixels are cloudy and in turn cloud fraction. Similar to the PATMOS-x data, other variables of interest include cloud-top temperature and cloud phase. For this study pixel-level (collection 5) data from the Aqua satellite are used.

3) COVERAGE AREA AND DATA FORMATTING

The time and spatial domain of the comparison portion of this study is chosen to have collocated and coincident PATMOS-x and MODIS measurements. The time examined is approximately 2200 UTC 6 August 2006. The spatial domain spans 12°S to 50°N and 152° to 114°W, an area located over the northern Pacific Ocean. This region is separated into 1° × 1° boxes and sampled at 2-km intervals, resulting in 2500 measurements per box or “scene” for both MODIS and PATMOS-x. We are interested in warm liquid water clouds, so ice clouds are filtered out, as are times when sun glint significantly affects retrievals. After removing ice clouds and the effects of sun glint, we are left with approximately 600 1° × 1° scenes. Once the comparison portion of this study is complete the liquid water path retrieval methods are applied globally. Daytime PATMOS-x satellite data are examined for June 2007.

b. Methodology

1) MASK METHOD

We explore two methods for deriving PDFs of liquid water path from cloudy scenes, and from these PDFs statistical information (i.e., mean and variance) about the subgrid variability of cloud liquid water path are calculated. The first method uses pixels classified as cloudy by the cloud mask algorithm that have valid retrievals of cloud optical depth and near-cloud-top effective radius. We shall refer to this as the mask method. Liquid water path \( W \) is derived from cloud optical depth \( \tau \) and near-cloud-top effective radius \( r_{\text{eff}} \) using the following equation:

\[
W = (5/9)\rho_l r_{\text{eff}}.
\]

Here \( \rho_l \) is the density of liquid water. The assumption is made, based on observations of boundary layer clouds, that liquid water content maintains a roughly linear relationship with height above a cloud base (Wood and Hartmann 2006). One issue with this method is that only a subset of the available data is used. This is illustrated in Table 1, where the first row shows that the MODIS cloud mask algorithm classifies 66% of the pixels as cloudy, while in only 42% of the pixels is liquid water path successfully calculated. This holds true for the PATMOS-x data as well, though the difference is smaller, with 59% of the pixels being classified as cloudy compared to 56% successful liquid water path calculations. Figure 1 shows a pair of sample images generated from MODIS data, where the first image displays visible reflectance and the second is the same image with all successful cloud optical depth retrievals overlaid. The light gray and white areas found at the edges of larger clouds and in small broken clouds in the second image indicate it is primarily for thin or broken clouds that cloud optical depth retrieval fails or is not attempted. Retrievals not attempted are due to a clear sky restoral (CSR) algorithm introduced in MODIS.

### Table 1. Total number and breakdown of successful retrievals (%) for MODIS and AVHRR. The percent values are with respect to the number of visible reflectance retrievals. The fifth column relates what percentage of satellite retrievals are classified as either probably or definitely cloudy. Retrievals when sun glint is a factor or when ice clouds are determined to be present are removed. Normally optical depth retrievals overlaid. The light gray and white areas found at the edges of larger clouds and in small broken clouds in the second image indicate it is primarily for thin or broken clouds that cloud optical depth retrieval fails or is not attempted. Retrievals not attempted are due to a clear sky restoral (CSR) algorithm introduced in MODIS.

<table>
<thead>
<tr>
<th>Satellite instrument</th>
<th>( R_{\text{vis}} ) retrievals</th>
<th>Clear-sky ( \tau ) retrievals</th>
<th>Cloudy ( \tau ) retrievals</th>
<th>Mask determined to be cloudy</th>
<th>( r_{\text{eff}} ) retrievals</th>
<th>Liquid water path retrievals</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODIS</td>
<td>1 842 064</td>
<td>0</td>
<td>50</td>
<td>66</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>AVHRR</td>
<td>1 396 787</td>
<td>0</td>
<td>56</td>
<td>59</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>AVHRRw/clear</td>
<td>1 396 787</td>
<td>22</td>
<td>56</td>
<td>59</td>
<td>7</td>
<td>78</td>
</tr>
</tbody>
</table>
Collection 005. The purpose of the CSR is to identify pixels considered to be poor retrieval candidates—dust, smoke, sun glint, and partly cloudy pixels—and to “restore” them to a clear-sky designation with a missing value for optical depth (King et al. 2006). These failed/not attempted retrievals have significant effects on the distribution of liquid water path over a given scene, generally causing a low bias since the missing retrievals are often optically thin cloud.

2) FIT METHOD

The method introduced here avoids this failed retrieval issue by defining a relationship between retrieved cloud optical depth \( \tau_{\text{ret}} \) and the corresponding visible reflectance measurements \( R_{\text{vis}} \) for each \( 1^\circ \times 1^\circ \) box:

\[
R_{\text{vis}} = \frac{a \tau_{\text{ret}}}{1 + b \tau_{\text{ret}}}. \tag{2}
\]

This equation is a general relation between cloud albedo and optical depth (e.g., Petty 2004). Its application to directional reflectance observed at multiple viewing angles instead of an albedo term follows similar relationships derived in previous work (Koren et al. 2008, see their appendix B). Specifically, for a relatively small spatial domain such as that used here the coefficients \( a \) and \( b \) account for directionally dependent factors such as solar zenith angle, single-scattering albedo, and asymmetry parameters. The equation is expanded to include a geometrical series expansion to account for the multiple scattering effects between the surface and cloud:

\[
R_{\text{vis}} = R_{\text{vis}} + \frac{c(1 - R_{\text{vis}})^2}{1 - cR_{\text{vis}}}. \tag{3}
\]

Solving for coefficients \( a, b, \) and \( c \) and rearranging Eq. (3) to solve for cloud optical depth in terms of visible reflectance, we can calculate a value for cloud optical depth \( \tau_{\text{fit}} \) for every observed visible reflectance measurement. The second step in this fit method is to convert the fit-generated cloud optical depth into liquid water path. Since we do not have valid droplet effective radius retrievals for each value of \( \tau_{\text{fit}} \), we cannot use Eq. (1) to calculate liquid water path unless we assume a constant droplet effective radius. Instead we apply a method from Bennartz (2007) that uses grid-averaged values of droplet number concentration \( N \) (cm\(^{-3}\)) within clouds and condensation rate \( c_w \) (kg m\(^{-2}\)). The advantage to this method is that measurements of cloud microphysical properties within stratuscumulus clouds indicate that assuming \( N \) to be constant over a relatively small spatial domain is a reasonable approximation (Bennartz 2007). That is, we are assuming that \( N \) varies less over a \( 1^\circ \times 1^\circ \) box than droplet effective radius. Another thing to note is that Eq. (3) is potentially sensitive to variations in surface albedo, which is one reason why we chose to focus on an area over the northern Pacific Ocean where the surface albedo is relatively homogenous. Condensation rate is calculated as 80% of the value that would be produced assuming air parcels rise adiabatically and condense any water in excess of parcel saturation (Duykerke et al. 1995; Pawlowska and Brenguier 2000; Bennartz 2007). The \( N \) is calculated using an equation derived in Bennartz (2007), and then that equation is rearranged to solve for liquid water path:

\[
W = \frac{2^{-5/2} \tau_{\text{fit}}^4}{kN} \left( \frac{3}{\pi Q} \right)^{-3} \left( \frac{3}{4\pi \rho_L} \right)^{-2} c_w^{1/2}. \tag{4}\]
Here $Q$ is the extinction efficiency and is set equal to 2, which is a reasonable approximation in the visible spectrum (Petty 2004), while $k$ is the ratio between the liquid droplet volume mean radius to the third power and the effective radius to the third power. Past studies have found $k$ to vary between 0.5 and 0.9 (Lu and Seinfeld 2006), while in this study we are using a constant value of $k = 0.8$. The cloud optical depth used here is that generated from the visible reflectance fit function in Eq. (3), rather than from the retrievals. This allows us to calculate liquid water path statistics using the full set of visible reflectance measurements, as opposed to a subset of data with the mask method. Figure 2 shows a selection of three PATMOS-x scenes on which this method is performed. This method can be applied to the estimation of cloud optical properties (i.e., optical depth and effective radius), but we have chosen to focus on physical properties such as liquid water path and visible reflectance because these directly link bulk physics to the radiation budget. We should note that this method does not provide a benefit for scenes without failed optical depth retrievals, which is often the case for completely overcast or clear conditions, and requires a minimum number of valid cloud optical depth and effective radius retrievals in order to generate the coefficients for a reasonably good fit function. In this study we set a minimum of 50 valid retrievals for each $1^\circ \times 1^\circ$ box.

3) CLOUD FRACTION

Generating a distribution of liquid water path using all visible reflectance measurements creates liquid water path values for the total scene, including clear-sky and cloudy areas, yet a distinction between clear-sky and cloud is necessary for radiative transfer calculations. As mentioned earlier the mask method uses cloud mask to determine cloud fraction. The fit method requires a
different approach whereby a distribution of clear-sky visible reflectance is removed from the total-sky distribution leaving only the cloudy portion of the scene. To accomplish this we must make an assumption about the boundary between sub- and supersaturated portions of the total-sky distribution, and then integrate over the supersaturated portion to diagnose cloud fraction. This is not unlike many statistical cloud schemes. The median visible reflectance for the clear-sky portion of each $\frac{1}{83}$ box, as determined by the cloud mask, is used as the center of a symmetric distribution where all visible reflectance values left of (less than) the median are considered to be clear sky. Figure 3 shows an idealized illustration of this process. This PDF of clear sky is converted to optical depth via Eq. (3) and subtracted from the total-scene fit distribution of optical depth. The remaining cloudy-sky distribution is converted to liquid water path via Eq. (4). Statistical moments (mean and variance) are calculated from the resulting cloudy-sky distribution of liquid water path.

4) PROBABILITY DISTRIBUTION FUNCTIONS

The statistical moments are used to approximate the shape of the horizontal distribution function of liquid water path PDF($W$) for each $1^\circ \times 1^\circ$ box. Two different functions are used for this purpose: a Gaussian function and a gamma function. The gamma function is of the form

$$PDF(W) = \frac{1}{\Gamma(\nu)} \left( \frac{W}{\bar{W}} \right)^{\nu-1} \exp\left(-\frac{W}{\bar{W}}\right),$$

where $\nu$ is the width parameter. The standard deviation of $W$ is

$$\sigma_W = \bar{W}/\sqrt{\nu},$$

and PDF($W$) is the probability of occurrence within a grid box of a vertically integrated liquid water path of value $W$ (Heidinger 2003).

The use of the Gaussian function follows work performed by Considine et al. (1997) and Wood and Hartmann (2006), who concluded that the relationship between variations in cloud thickness and cloud fraction could be approximated by a Gaussian PDF form (see Fig. 4). The horizontal axis in Fig. 4 represents the physical cloud thickness, so the area where the thickness is negative is clear sky. The positive area represents the cloudy portion of the sky. If we normalize the distribution to unity the integration of the positive portion of the distribution results in the cloud fraction. Furthermore, if we assume an adiabatic cloud layer we can relate the positive cloud thickness to liquid water path and generate a PDF of $W$ (Considine et al. 1997; Wood and Taylor 2001; Wood and Hartmann 2006). The input for this function is cloud fraction, mean liquid water path, and condensation rate. The Gaussian distribution is generated as a function of physical cloud thickness $H$ where the relationship between thickness and liquid water path is defined as

$$W = 0.5c_n H^2.$$  

To convert from $H$ space to $W$ space an integral coordinate conversion has to be performed so that the norm of the PDF remains constant:

$$PDF(W) = PDF(H) \frac{dH}{dW} = \frac{1}{\sqrt{2\pi c_n \bar{W}}} PDF(H).$$

When this conversion is performed the PDF of $W$ is generated for the cloudy portion of PDF($H$) and is no longer strictly of a Gaussian form.
3. Satellite results

Once a PDF of liquid water path has been generated and statistical information has been derived, a method of evaluating the accuracy of the statistical moments is needed. The metric used here is the ability of the derived mean and variance of liquid water path to approximate the observed PDF and in turn reconstruct the total-scene visible reflectance. Several aspects of the fit method are evaluated, including the applicability to different satellite datasets, the advantages and disadvantages of using a Gaussian or gamma function to simulate the subgrid horizontal distribution of liquid water path, and the applicability to a wide array of scenes with differing atmospheric conditions. Figure 5 illustrates the basic suitability of the Gaussian and gamma functions to approximate the observed horizontal distribution of liquid water path. A reduced $\chi^2$ test (Taylor 1982) is performed for $1^\circ \times 1^\circ$ scenes, each with a minimum of 50 degrees of freedom. A value of $\chi^2 < 1$ indicates the function provides a reasonable fit to the observed distribution, and both the Gaussian and gamma functions achieve this for over 90% of the scenes.

a. MODIS versus PATMOS-x

We next calculate a total-scene reflectance from PDF($W$) for comparison against the observed total-scene reflectance. The mean cloudy portion of the scene reflectance is

$$\overline{R_{\text{cld}}} = \int_0^\infty R_{\text{vis}} \cdot \text{PDF}(W) \, dW.$$  

(8)

To calculate this PDF($W$) is converted to a PDF of optical depth using Eq. (4), which is in turn converted to visible reflectance using the inverse of Eqs. (2) and (3). The reflectance for the clear portion of the scene $R_{\text{clr}}$ is the mean reflectance for clear pixels as determined by the cloud mask. The calculated total-scene reflectance $R_{\text{tot}}$ is found using the cloud fraction CF:

$$R_{\text{tot}} = \text{CF}(\overline{R_{\text{cld}}}) + (1 - \text{CF})(R_{\text{clr}}).$$  

(9)

Figure 6 lists the basic steps to calculate PDF($W$) and then total-scene reflectance for the fit and mask methods, highlighting the differences between the two.

Figure 7 shows a scatterplot of observed versus calculated total-scene visible reflectance of $1^\circ \times 1^\circ$ boxes that span $12^\circ S$–$50^\circ N$ and $152^\circ$–$114^\circ W$ taken from PATMOS-x data on 6 Aug 2006. Values of $\chi^2 < 1$ indicate the function is doing a satisfactory job of modeling the observed horizontal distribution of liquid water path.

The top row of Fig. 7 shows the mask method of calculating PDF($W$) often overestimates total-scene reflectance for PATMOS-x and MODIS. The bias is greater in the MODIS data, which is because MODIS filters more pixels than PATMOS-x. As observed in section 2b(1), many of the pixels filtered by cloud-masking algorithms are thin clouds with low liquid water paths, creating a high bias in the distribution of liquid water path. The gamma distribution reflectance bias is $0.023 \pm 0.039$ for PATMOS-x and $0.045 \pm 0.023$ for MODIS, while for the Gaussian distribution the bias is $0.015 \pm 0.038$ and $0.030 \pm 0.024$, respectively.

The bottom row in Fig. 7 uses the fit method to calculate PDF($W$), which reduces the bias. The gamma distribution reflectance bias is $-0.002 \pm 0.008$ for PATMOS-x and $-0.004 \pm 0.007$ for MODIS, while for the Gaussian distribution the bias is $-0.007 \pm 0.011$ and $-0.008 \pm 0.012$, respectively.

Figure 8 displays these biases bin-averaged by visible reflectance; for PATMOS-x the bias generally increases with visible reflectance when using the mask method, while for the fit method there is little to no bias until the visible reflectance reaches 0.35, at which time negative biases appear when using the Gaussian function to approximate PDF($W$). For MODIS the mask method produces a positive bias until the visible reflectance reaches 0.35, at which time the gamma distribution maintains the positive bias while the Gaussian distribution bias reduces to zero and even becomes slightly negative. For the fit method both distributions behave in a similar manner:
little to no bias until the reflectance reaches 0.35, at which time the Gaussian distribution produces a negative bias.

b. Gamma versus Gaussian fit

One distinction between the Gaussian and gamma distributions is that the Gaussian function generates a negative mean bias for large reflectance (greater than 0.35), while the gamma function does not. Another difference is that the Gaussian distribution has a greater variance than that of the gamma distribution for large reflectance. These differences suggest a potential dependence on cloud fraction and cloud physical thickness as it relates to liquid water path. To examine this we use the fit method to generate two-dimensional histograms of liquid water path and cloud fraction. Figure 9 shows the reflectance bias and relative frequency of occurrence for each liquid water path-cloud fraction pairing for PATMOS-x and MODIS. For PATMOS-x the mean reflectance bias is small (less than 0.01) for mean liquid water paths less than 40 g m\(^{-2}\), regardless of cloud fraction or whether the Gaussian or gamma function is used to approximate PDF\(W\). This accounts for 84% of the 1° x 1° scenes being examined, suggesting that both functions perform well for a majority of atmospheric conditions, at least for the area being examined here. For mean liquid water path greater than 40 g m\(^{-2}\) there is more variation in the performance of the functions. Most notably the Gaussian function generates a negative bias for high mean liquid water path (>80 g m\(^{-2}\)) and cloud fraction less than 0.8. Figure 10 shows some examples of the observed liquid water path PDFs along with gamma and Gaussian approximations, and examination of the bottom right panels show why the Gaussian function creates the negative bias. When the PDF shifts far enough to the right the Gaussian function, which is generated using cloud fraction and mean liquid water path, does not have the required degrees of freedom to shift along with it. This is because, as described in section 2b(4), the Gaussian distribution of \(W\) is converted from a Gaussian distribution of \(H\), whose \(H = 5\) as the separation point between clear and cloudy sky. When the conversion from \(H\) space to \(W\) space takes place that separation point becomes \(W = 0\), meaning regardless of cloud fraction the left tail of the Gaussian is constrained to \(W = 0\). The gamma function has the necessary degrees of freedom to begin with an offset based on a minimum value of \(W\). The bottom-right-hand panel in Fig. 10 shows a good example of this. MODIS, like PATMOS-x, has a small mean bias (less than 0.01) when
mean liquid water path is low, though for MODIS the threshold is 60 g m\(^{-2}\) instead of 40 g m\(^{-2}\), which accounts for 88% of the 1° × 1° scenes examined. For mean liquid water path greater than 60 g m\(^{-2}\) the Gaussian function generates the same negative bias for MODIS as it does for PATMOS-x.

Among the 600+ scenes examined, most can be well represented with a unimodal distribution such as the Gaussian and gamma functions provide. However, there are a handful of scenes for both PATMOS-x and MODIS for which the observed liquid water path more closely resembles a bi- or multimodal distribution, and the relative frequency of occurrence of these scenes appears to be correlated with mean liquid water path. The right-hand panel in the middle row of Fig. 10 is a good example of such a distribution. In these cases the ability of either the Gaussian or gamma functions to approximate the observed liquid water path PDF is diminished.

4. Extending the fit method temporally and spatially

Having established some of the advantages of deriving liquid water path statistical moments using the fit method versus the mask method, it is useful to examine more closely some possible applications. To this end the fit method is extended spatially by applying it to one month of global data, and then temporally by applying it to over 20 years of data over the North Pacific. PATMOS-x data are used for both applications.

**Fig. 7.** Scatterplot of observed vs calculated visible-scene reflectance for 1° × 1° boxes: (left) PATMOS-x and (right) MODIS; (top) mask-derived \(R_{vis}\) and (bottom) fit-derived \(R_{vis}\).

**Fig. 8.** Bin-averaged fit visible reflectance plotted as a relative difference from that observed: (left) PATMOS-x and (right) MODIS.
Using the gamma distribution and PATMOS-x data described in section 3 the liquid water path retrieval method is examined globally. For daytime hours, the fit and mask methods are applied to those areas that contain liquid water cloud, as ice clouds are not included in this study. The results are then averaged for June 2007. Figure 11 shows the difference between observed and derived reflectance for 1° × 1° scenes globally. Figure 11a uses the mask method to derive liquid water path statistics and shows the reconstructed reflectance bias. Similar to the scenes examined in section 3, the mask method produces a global overestimate of reflectance that has little dependence on location. The mean global reflectance bias is 0.050 ± 0.047. Figure 11b displays the same total-scene reflectance bias but with liquid water path information derived using the fit method, which produces a reflectance bias of −0.008 ± 0.017. The noticeable improvement can be attributed to the fit method deriving values of liquid water path from all available reflectance measurements over the spatial domain (excluding areas with ice clouds and sun glint, which are not used for reflectance calculation either). Since the total-scene reflectance is the mean value of all available visible reflectance measurements, the performance of the fit method is dependent on only three factors: 1) the goodness of the fit function [Eq. (3)], 2) the approximation of the Gaussian and gamma function to the observed liquid water path PDF, and 3) the distinction between clear and cloudy portions of the scene examined [Eq. (9)]. Meanwhile, as mentioned earlier, the mask method uses only those reflectance measurements from

![Fig. 9. The 2D histograms of visible reflectance bias relative to observations as a function of liquid water path and cloud fraction. Numbers in each bin represent the relative frequency of occurrence for that liquid water path–cloud fraction pairing. The data for this figure are approximately 600 1° × 1° scenes over the northern Pacific Ocean. (top) A gamma function is used to approximate in-cloud liquid water path distribution; (bottom) a Gaussian function is used. (left) PATMOS-x data and (right) MODIS data.](image-url)
which successful retrievals of cloud optical depth and droplet effective radius are derived, meaning in addition to the above factors the mask method is also dependent on how well a subset of the total available reflectance measurements represents the whole. Since the failed cloud optical depth and droplet effective radius retrievals often occur around the optically thin cloud edges, a positive bias is generated.

b. Multidecadal results

Figure 12 shows the mean liquid water path derived using both the fit and mask methods from 1988 through June of 2009. The area examined is in the North Pacific spanning from 15° to 35°N latitude and 140° to 120°W longitude. Similar to the previous analysis the spatial domain is broken into $1° \times 1°$ boxes. The daily mean liquid water path is then defined as the mean of the cloudy and clear portions of sky for all the boxes. A polynomial fit function is used to interpolate each data point to 1400 local time, which helps remove bias associated with satellite drift. The slope of each line is calculated with a simple linear fit and includes one standard deviation. Calculating $t$ values for the slope of January and July shows a significant negative trend in January and no significant trend in July for both methods. However, the slopes of the fit and mask methods differ, indicating the method used affects the trend found. For example, a shift in cloud regimes from a small number of large clouds to greater numbers of smaller clouds would affect the amount of optically thin cloud edges, which would produce different results for the fit and mask methods. In this context the fit method is a potentially valuable way to identify data artifacts affecting long-term trends in cloud properties.

![Figure 10](image-url)
5. Discussion and conclusions

A method of deriving statistical information related to the distribution of liquid water path over partially cloudy scenes is tested using satellite data. It is found that using the fit method improves our ability to reconstruct total-scene visible reflectance when compared to an approach that relies only on valid liquid water path retrievals. This disagreement may in part be explained by differences in the cloud masking algorithms, which filter pixels that are difficult to classify as either clear sky or cloud. Furthermore, when using these distributions to reconstruct total-scene reflectance, both climatologies have positive mean biases when compared to that observed. When the fit method is applied to the same data the reconstructed reflectance from both climatologies more closely match that observed, generating mean biases of less than 0.01. This allows conservation of observed visible reflectance when generating distributions of liquid water path, which in turn has the advantage of not only being a tool to compare satellite datasets but also to generate validation data for use in climate models.

Gaussian and gamma functions are used to approximate the distribution of horizontal subgrid-scale liquid water path for partially cloudy scenes, the advantage being they can be generated using little information while being accurate for a majority of atmospheric conditions. The Gaussian function requires only cloud fraction and mean liquid water path, while the gamma function uses the mean and standard deviation of liquid water path. All of these values can be derived from satellite data given a sufficient period of time or spatial domain.

When using the fit method, both functions reconstruct total-scene reflectance with mean biases of less than 0.01 over all cases; however, the variance in the Gaussian results is greater than that of the gamma function for reflectance greater than 0.35. In addition, for scene reflectance greater than 0.35, the Gaussian function produces a negative bias: $-0.022 \pm 0.02$ for MODIS and $-0.023 \pm 0.025$ for PATMOS-x. Finally, a handful of partially cloudy scenes are found to have bi- or multimodal distributions of liquid water path, and the occurrence of these scenes increases with mean liquid water path. For these scenes the observed liquid water path PDFs are not reproduced well by either the Gaussian or gamma function.

The fit and mask methods of deriving liquid water path PDFs are applied globally to the PATMOS-x climatology for June 2007. The gamma function is chosen to approximate the distribution of liquid water path for each partially cloudy scene. It is found that using the mask method there is a mean reflectance bias of $0.050 \pm 0.047$ globally with no discernible dependence on location, while using the fit method produces a mean bias of $-0.008 \pm 0.017$. The global improvement is similar to that found over the northern Pacific and does not appear to be location dependent.

These methods are also applied to a small area in the North Pacific for the past 30 years. Small differences in
the trends of mean liquid water path are found for the two methods, suggesting the fit method may provide a means of identifying data artifacts that could affect long-term trends in cloud properties.

It is not surprising that the fit method produces total-scene reflectance closer to that observed than does the mask method, as the fit method incorporates all visible reflectance measurements over a spatial domain. A cloud climatology that includes liquid water path distributions that conserve cloud and total-scene reflectance is a valuable tool for use in model validation, and a climatology that extends back nearly 30 years, as does PATMOS-x, may be used as validation data for climate models. The simplicity of the fit method and its input requirements makes it easy to apply to various satellite climatologies, and could potentially be used as a comparison tool for these climatologies. Future work includes modifying the fit method to incorporate ice clouds.

Acknowledgments. The views, opinions, and findings contained in this article are those of the authors and should not be construed as an official NOAA or U.S. government position, policy, or decision.

REFERENCES
—, B. A. Wielicki, and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous