Impact of Rain on Retrieved Warm Cloud Properties Using Visible and Near-Infrared Reflectances Using Markov Chain Monte Carlo Techniques

Zhuocan Xu, Gerald G. Mace, and Derek J. Posselt

Abstract—Estimates of cloud droplet effective radius \( (r_e) \) and optical thickness \( (\tau) \) can be derived using reflected sunlight in a visible non-absorbing channel combined with reflectances from a near IR channel that is absorbing [e.g., the bispectral method (BSM)]. Discrepancies between BSM-estimated \( r_e \) and collocated \textit{in situ} measurements are commonly attributed to a violation of the assumptions used in the BSM algorithm such as plane parallel geometry, and a single mode droplet size distribution (DSD). This research uses Markov chain Monte Carlo (MCMC) experiments to examine the impact of precipitation on BSM-retrieved \( r_e \) near optical cloud top by comparing the retrievals and associated uncertainties obtained from two types of experiments assuming a unimodal or bimodal drop size distribution. Where rain is present, BSM-retrieved \( r_e \) overestimates the true cloud mode \( r_e \). Moreover, there is no longer a unique measure of \( r_e \) within the precipitating liquid-phased clouds, resulting in a substantial increase in retrieval uncertainties. This leads to a corresponding loss of information on the total number concentration and liquid water content \( (LWC) \) near the cloud top. It is found that \( r_e \) biases are not strongly correlated with properties exclusively pertaining to rain, such as rain water content \( (RWC) \) or precipitation rates, but tend to be a function of the ratio between rain and cloud water content \( (CWC) \) and the cloud total number concentration. These results highlight the need for additional independent information such as from an active or passive microwave sensor that can identify the presence of precipitation and constrain additional aspects of bimodal droplet distributions.

Index Terms—Bayesian methods, cloud microphysics, cloud remote sensing, error analysis, Moderate Resolution Imaging Spectroradiometer (MODIS).

I. INTRODUCTION

MUCH of what has been determined observationally about aerosol indirect effects in warm clouds [1] and assessments of cloud parametrizations in global climate models [2] have been inferred from cloud properties derived from data collected by satellite spectrometers in the visible spectrum [3]. In these algorithms, cloud optical thickness \( (\tau) \) and effective radius \( (r_e) \) are simultaneously estimated by combining solar reflectance measurements in the visible and near-infrared regions of the electromagnetic spectrum [4]. This approach is often referred to as the bispectral method (BSM), which rests on the principles that the reflectances at visible wavelengths are primarily a function of optical thickness and not effective radius, and the near-IR reflectance can depend upon both cloud-effective radius and optical thickness. With assumptions about the nature of the properties of droplet size distribution (DSD) in the cloud column, \( \tau \) and \( r_e \) can be used to further derive other variables of fundamental importance such as liquid water path and cloud number concentration \( (N_c) \). Assumptions inherent in satellite BSM retrievals include plane-parallel transport of photons as well as vertical and horizontal homogeneity of clouds within a pixel, which can be easily violated within natural clouds and therefore cause biases.

Discrepancies between \( r_e \) estimated with the BSM and collocated \textit{in situ} measurements have been noted in a body of literature that seeks to validate satellite retrievals against airborne measurements. For stratuscumulus clouds observed at solar zenith angles less than 50°, the Moderate Resolution Imaging Spectroradiometer (MODIS, [3]) retrievals have been found to overestimate cloud top \( r_e \) by: 15%–20% \((\sim 2.08 \mu m) [5]\), up to 1.75 \( \mu m \) for cloud droplets with size between 6 and 11 \( \mu m \) [6], and 13% on average [7], in comparison with the collocated aircraft data collected over the Eastern Pacific Ocean. Cloud scenes investigated in these studies are relatively homogeneous, and the potential impacts of subpixel variability and 3-D effect were examined and deemed to be limited. King et al. [7] noted that the largest positive biases occur where drizzle-size drops are detected by the two-dimensional stereo probe (2DS). Witte et al. [8] argue that these differences between satellite and \textit{in situ} \( r_e \) result from the fact that the aircraft probes (i.e., CAS + CIP or CDP + 2DC/CIP) undermeasure droplets with diameters between 40 and 80 \( \mu m \) [9], and assert that MODIS products show little biases for \( r_e \) that range from 5 to 16 \( \mu m \) within stratuscumulus. More recently, higher magnitude biases have been recorded over the Southern Ocean, which is unlikely explainable solely by accounting for the uncertainties and errors of \textit{in situ}
probes. Focusing on broken and fast-moving clouds under low sun conditions, Ahn et al. [10] reported an average of 13-μm overestimation for non-drizzling clouds and 10-μm underestimation for heavily drizzling clouds in the MODIS \(r_e\). Over the approximately same region, Kang et al. [11] conducted an intercomparison in the overcast stratocumulus and found a 3–4 μm negative offset in satellite products when significant drizzle is present near the cloud top.

Biases emerging in the in situ validation studies are often an intertwined outcome of various error sources. Numerous studies have demonstrated that the impact of 3-D radiative effects, subpixel variability, and view geometry can be significant (see [12], [13], [14], [15], [16], [17]; among many others). There have been far fewer investigations of errors due to the formation and presence of liquid precipitation in the atmospheric cloudy column. As hydrometeor droplets collide and coalesce within the turbulent air, DSDs tend to broaden, which could ultimately result in a distinct mode associated with drizzle or rain coexisting with the cloud droplet mode [18]. Microphysical profiles can also depart from adiabatic cloud droplet growth approximations in precipitating scenes.

In the MODIS operational algorithms, DSDs are assumed to follow unimodal gamma distributions, and the effective variance that characterizes the DSD width is fixed at 0.1 [19], in which case the potential existence of rain is omitted. Evidence suggests a tendency of \(r_e\) being overestimated, with the true DSDs being narrower than the prescribed width (see [5], [20]), while the effective variance is deemed to be secondary to \(r_e\) in determining reflectance measurements (see [4], [21]). Studies of bimodal DSDs were conducted primarily with numerical sensitivity tests and large eddy simulations. Nakajima et al. [22] suggested that discrepancies between an assumed unimodal \(r_e\) and the true cloud mode \(r_e\) in bimodal distributions grow with the volume ratio between drizzle and cloud mode. Similar comparisons of retrieved \(r_e\) under different DSD assumptions were carried out by Platnick and Valero [20], who perturbed the mean radius and water content of the drizzle mode, and found that the differences of \(r_e\) can rise up to 8 μm. In large eddy simulation studies (see [13], [23]), the impact of bimodality was found to be insignificant for cases with very light rain (<1.2 mm/day) and remains unclear for cumulus with local rain rates greater than 24 mm/day, as reflectances over large regions with drizzle were too low for retrievals in the case examined.

Previous studies exhibit substantial variability in terms of the \(r_e\) bias magnitude. Most studies were targeted at clouds with light drizzle representative of stratocumulus while few examined cloud scenes with stronger precipitation. Continuing along this line of inquiry, we explore how the presence of rain affects the physical constraints of bispectral reflectances posed on \(r_e\) in a retrieval context. The vertical profiles of distinct cloud and rain microphysics can be derived from 0.55 to 2.1-μm passive reflectances with the Markov chain Monte Carlo (MCMC) approach developed in [24], which also allows an explicit assessment of observational information carried by the solar reflectances as well as the biases and uncertainties of retrieved \(r_e\) that arise due to the bimodality of DSDs.

II. METHODOLOGY

A. Retrieval Algorithm

Following [24], DSDs are assumed to be bimodal and composed of cloud and rain drop populations, which are individually characterized by the modified gamma distributions:

\[
N(D) = N_{x,i} \left( \frac{D}{D_{x,i}} \right)^{\alpha} \exp \left( -\frac{D}{D_{x,i}} \right) \tag{1}
\]

where \(i\) is an index that indicates either rain (\(p\)) or cloud (\(c\)) mode. \(N_x\) and \(D_x\) denote characteristic numbers associated with the total number concentration and droplet size, respectively. \(\alpha\) is a shape parameter that governs the width of the DSDs. If both cloud and rain modes are assumed to be present, a total of six DSD parameters, \(N_{x,c}, \alpha_c, D_x,c, N_{x,p}, \alpha_p,\) and \(D_{x,p}\), are retrieved at each cloudy level.

If the precipitation amount is assumed to be constant, only vertical profiles of cloud DSD parameters are retrieved. Liquid water content (LWC), total number concentration (\(N_r\)), and effective radius (\(r_e\)) can be subsequently derived from the DSD parameters with the assumption that the liquid-phased drops are approximately spherical in the absence of heavy precipitation

\[
\begin{align*}
\text{LWC}_i &= \frac{\pi}{6} N_{x,i} D_{x,i}^4 \Gamma(\alpha_i + 4) \tag{2} \\
N_{x,i} &= N_{x,c} D_{x,c} \Gamma(\alpha_i + 1) \tag{3} \\
r_{e,i} &= \frac{1}{2} D_{x,i} \Gamma(\alpha_i + 3). \tag{4}
\end{align*}
\]

As a Bayesian methodology, MCMC determines the posterior probability distributions of state variables \(P(x | y)\) from a conjunction of probability density functions (PDFs), which quantitatively characterize the information contained in measurements \(P(y)\), the prior knowledge of state variables \(P(x)\), and the likelihood \(P(y | x)\) that the estimated state can represent the “true” state consistent with observations

\[
P(x | y) = \frac{P(y | x) P(x)}{P(y)}. \tag{5}
\]

One or more sets of forward models are often employed in order to map from the space of state variables to the measurement space, and we describe those that we use in our experiments below. The prior applied in our MCMC algorithm (see Table I) is cast in cloud space, the statistics of which are determined primarily from shallow cumulus data collected during the RICO field campaign [25]. The median and interquartile ranges (IQR) of rain rates derived from the in situ DSD measurements are 9.2 and 48.6 mm/day, respectively.

Twelve Markov chains with varying first guesses are constructed to effectively explore the solution space. Proceeding from one iteration to the next within each chain, a new set of state variables (\(\hat{x}\)) for consideration is randomly drawn from the “proposal” distribution \(q(\hat{x}, x)\), which is assumed to take the form of a Gaussian probability distribution with its mean located at the current state (\(x\)). Forward simulations are then conducted using the proposed set (\(\hat{x}\)) as inputs. The key probabilistic procedure to decide if \(\hat{x}\) is accepted as a new
TABLE I

|                | LWC_L [gm⁻³] | Re_L [microns] | Nt_L [cm⁻³] | LWC_S [gm⁻³] | Re_S [microns] | Nt_S [cm⁻³] |
|----------------|---------------|----------------|--------------|---------------|----------------|--------------|
| **Median**     | 0.038         | 122.4          | 0.021        | 0.31          | 11.2           | 221.6        |
| **IQR**        | 0.24          | 221.4          | 0.27         | 0.63          | 8.2            | 575.2        |

member of the posterior samples via the computation of the acceptance ratio, which is formulated as

\[ \rho = \frac{P(\hat{x})P(\hat{y} | \hat{x})q(\hat{x}, x_i)}{P(x_i)P(y | x_i)q(x_i, \hat{x})} \]  

(6)

\( q(\hat{x}, x_i) \) indicates the probability of transitioning from \( x_i \) to \( \hat{x} \), and \( q(x_i, \hat{x}) \) is similar but reverses the direction of the transition. If the distribution is symmetrical, \( q(x_i, \hat{x}) \) and \( q(\hat{x}, x_i) \) are equal. In our case where the Gaussian PDFs are assumed, the acceptance ratio is simplified to a function of the prior and likelihood functions.

The new set will be accepted instantly if it generates forward model output that is in an improved agreement with measurements (\( \rho > 1 \)). Otherwise, the accept/reject decision will depend on the comparison of \( \rho \) with a random number from a uniform distribution. \( \hat{x} \) will be accepted to the posterior sampling space only if \( \rho \) is larger than the random number. The procedure is intended to reject the candidate set if the forward simulated measurables diverge from observations, and allow candidate sets with \( \rho \) close to 1 still to be considered. As such, the high-probability regions in the solution space are favored, while the sampling processes do not terminate upon finding a local probability maxima (i.e., mode). If there exist multiple modes in the posterior probability space, MCMC algorithms are capable of returning the solution as such providing reasonable prior, well-functioning forward models, and informative observations. In addition, there is no restriction on the form of the statistical distribution for the PDFs involved in (5) and (6). The flexibility intrinsic to MCMC differentiates it from simple optimization techniques, enabling the method to yield a complete solution space that is representative of the variability within natural clouds to the extent that the computational demands of the methodology can be accommodated. We refer to [24] regarding further details of the MCMC algorithms. Upgrades from the established methodology are primarily associated with the choice of forward models and subsequent implementation of radiative transfer simulations.

B. Radiative Transfer Simulations

For the purpose of obtaining optimal computational accuracy, the forward simulator used in [24] is replaced with a 16-stream discrete-ordinate numerical model (DISORT, [26]). Radiative transfer simulations are conducted under the assumption that low-level warm-phase clouds consist of three vertical levels. The number of levels is determined based on the profile generated from the RICO data collected on January 7, 2005 (refer to [24] for further details), intending to represent a measure of the vertical variability observed in maritime shallow cumulus without requiring the algorithm to carry the additional burden of resolving the details of cloud vertical structure. Perturbations of DSD parameters in the MCMC retrievals are performed on a 350-m resolution vertical coordinate, which is converted to a coordinate measured in optical depth in the radiative computations with DISORT. Thermodynamic profiles consistent with a standard tropical atmosphere are adopted, and the underlying surface is assumed to be Lambertian with an albedo of 0.0644 at the visible wavelength. Unpolarized and unnormalized phase function Legendre moments, as well as scattering and extinction efficiency factors, are first computed based on Mie theory and saved as a function of size parameter with 2100 terms in the lookup tables for the operating wavelengths at 0.55 and 2.1 μm [27]. Intending to generate the coefficients in the Legendre polynomial expansions of the composite phase function for a cloud volume, the previously stored phase function Legendre moments are integrated over the unimodal or bimodal DSDs proposed in each iteration of the MCMC runs, and then normalized by the integral of scattering cross sections over sizes. The ultimate phase function moments are a weighted sum of cloud and Rayleigh scattering with aerosols omitted. In terms of gaseous absorption, ozone and water vapor are of radiative importance, respectively, at 0.55 and 2.1 channels [28]. They are taken into account with the SBDART LOWTRAN band models with a 20 cm⁻¹ resolution, which are coarsened from the detailed line-by-line computations [29]. Since we are focusing on the microphysical aspects of the BSM, the solar zenith angle is fixed at 45°, the sensor view angle at 18°, and the azimuth angle at 0°.

C. Experimental Design

Imaging radiometers measure the radiances emerging from the tops of clouds, which are a weighted sum of the radiometric return from drops of various sizes within the entire cloudy column. Without additional ancillary observations, the contribution from cloud and rain drops cannot be differentiated. Duplicating the challenges a BSM algorithm would encounter in reality, two types of MCMC experiments, each with contrasting assumptions, are performed for the same thermodynamics profile with an identical set of VIS and near-IR reflectances. One assumes the co-occurrence of cloud and rain, and the other assumes that rain is absent. In the former, the cloud and rain mode DSD parameters are perturbed and retrieved, while in the latter only the cloud mode is perturbed with rain water content (RWC) remaining constant at an extremely small value (i.e., 10⁻⁶ gm⁻³). The latter is referred...
to as the “cloud-only” experiment. How the presence of rain exerts an influence on $r_e$ retrieval accuracies and uncertainties can be evaluated by comparing the MCMC outcomes with the 

A BSM lookup table is produced with $r_e$ at the cloud-top level and total optical thickness (see Fig. 1 and Table II) generated from the solution space of a cloud-only MCMC run with input reflectances at p12_24. The penetration depth of photons at water-absorbing bands is known to vary with the wavelength selected as well as the size of droplets within the upper layer of a cloudy profile [30]. Our retrievals show the constraints provided by 2.1-$\mu$m reflectance on cloud $r_e$ are largely restricted to within five optical depths from cloud top, which will be discussed in Section V. As such, the investigations are carried out using $r_e$ at cloud top only so as to avoid the effects of rain being confounded by the lack of information on effective radius deeper in the clouds.

In essence, the BSM lookup table (see Fig. 1) provides the result that a BSM algorithm would report given the visible and near-IR reflectances, regardless of whether a precipitation droplet mode is present or absent, since the BSM algorithm assumes a unimodal cloud DSD. An MCMC BSM algorithm that assumes a single mode cloud DSD should return effective radii and optical depths very close to the BSM values in Fig. 1.

In the bimodal DSD case, the MCMC algorithm converges on cloud and precipitation droplet characteristics that would, together, produce the visible and near IR reflectances assumed to within the uncertainties of the reflectance measurements. We then compare the cloud properties of the bimodal case with the results of the lookup table in Fig. 1. The degree to which the cloud properties of the bimodal case differ from the single-mode lookup table in Fig. 1 represents the bias that a single mode BSM algorithm would suffer in the presence of precipitation. One might assume that the effective radius that a single-mode BSM algorithm would infer in the presence of both cloud and rain might be some weighted combination of the reflectance contribution from both modes. If we were able to know that weighting function, then it might be possible to separate the single mode BSM-derived microphysical state into separate cloud and precipitation contributions. The bimodal MCMC experiment essentially provides this weighting function, and also returns a quantitative estimate of uncertainty. What we will show, however, is that the uncertainty in the bimodal retrieval precludes knowledge of this weighting under the limited information provided by bispectral reflectance pairs.

A suitably low uncertainty of 2% is assumed for reflectance measurements, provided that the MODIS instrument uncertainty is approximately 1%–2% [31]. We note that potentially significant sources of forward model uncertainties (e.g., 3-D radiative effects) are not the focus of this study and are not considered. Our objective is to evaluate the effect of precipitation on the BSM retrievals. Reflectances of 0.55 and 2.1 $\mu$m at eight points marked in Fig. 1 are used to perform the two types of MCMC runs as designed, which allows for the dependence of rain effects on cloud microphysical and optical properties to be examined.

III. BIASES AND UNCERTAINTIES OF CLOUD EFFECTIVE RADIUS

We initiate this section with a brief overview of retrieval results. The inferred PDFs of rain $r_e$ [see Fig. 2(c) and (e)] and other rain variables (not shown) to a large part resemble the prior [see Fig. 2(a)], and negligible differences exist among the points, which imply that little knowledge in regard to precipitation is gained beyond what is known a priori from the in situ data with the addition of reflectance measurements. As opposed to the rain mode, the PDFs of the cloud mode are noticeably more confined than the prior [see Fig. 2(b)], suggesting a reduction of uncertainties due to the observational information. The median of cloud $r_e$ is about 11 $\mu$m in the prior (see Table I). The introduction of observations is driving the posterior PDFs of $r_e$ away from the prior median at p8_36 and p20_16. The low $R_{21}$ at p20_16 suggests absorption by large droplets is significant, and therefore the algorithm returns 17.5 $\mu$m as the modal $r_e$. The relatively high $R_{21}$ at p8_36 suggests the opposite would likely be the case, which is evidenced by the posterior PDF maximum probability at 7.6 $\mu$m. Also notable is the cloud $r_e$ retrieval uncertainty is more diminished from the prior at p8_36 [see Fig. 2(d)] relative to p20_16 [see Fig. 2(f)]. The uncertainties of cloud $r_e$ are enlarged in the bimodal experiments in comparison with those derived in the cloud-only experiments, the magnitude of which appears dependent on the input reflectances.

The medians of the posterior PDFs generated within the rain-free and rain-present MCMC runs are compared against the BSM $r_e$ at each chosen point. Fig. 3(a) shows their relative differences scaled by the corresponding BSM estimates, where the rain-free experiments are indicated by black bars and rain-occuring experiments are grouped into four different colors based on the BSM $r_e$. The median fractional biases are almost negligible when rain is absent and only a cloud droplet mode is used in the MCMC inversion algorithm. In contrast, the BSM $r_e$ is found to be consistently greater than the cloud mode effective radius in the experiments assuming the presence of both precipitation and cloud droplet modes. This can be interpreted as implying that for a certain reflectance pair, a cloud
with precipitation will, in reality, have a smaller effective radius than what a BSM algorithm would infer. In a cloud where a distinct rain mode emerges, bimodal distributions are necessitated to faithfully characterize the DSDs. A single $r_e$ that is derived from an actual bimodal DSD cannot represent the cloud mode nor the rain mode in a physically meaningful manner. Regarding the magnitudes of the biases, the strength of absorption governed by the cloud drop sizes appears to be the first-order influential factor. The biases are smaller than 8% ($\sim 0.9$ $\mu m$) when the BSM $r_e$ is 12 $\mu m$ or less (points p12_24, p8_16, and p8_36) and rises to 27% ($\sim 7.7$ $\mu m$) at p28_16 with $r_e$ around 28 $\mu m$. Overall, the largest biases occur when the reflectances in the near-IR are smaller. With similar 2.1-$\mu m$ reflectances (indicated by the same color in Fig. 3), the significance of precipitation droplet mode is modulated by the brightness of clouds. For instance, the input measurements of $R_{21}$ at p28_16 and p28_64 are nearly identical, while the biases resulting from the bimodality of DSDs are notably less pronounced at p28_64. Since the optical thickness of the cloud mode constitutes over 90% of the total in our retrievals, without a sufficient proportion of small-sized cloud droplets, the relatively high $R_{055}$ at p28_64 can hardly be obtained.

Signatures contained in the reflected solar energy at VIS and near-IR wavelengths are a total measure of the scattering and absorption that occur over a spectrum of drop sizes. The presence of rain runs afoul of the unimodal assumption in the BSM algorithm, the degree of which is linked with the dominance of cloud mode in the DSDs. It is found that $r_e$ biases are not strongly correlated with properties exclusively pertaining to rain, such as RWC or precipitation rates in our retrievals, but tend to be commensurate with the ratio between rain and cloud water content (CWC) [see Fig. 4(a)]. At p20_36 where RWC/CWC is about 0.2, the BSM $r_e$ is 2.2 $\mu m$ larger than the cloud mode $r_e$ in the rain-present experiment. The discrepancies grow to approximately 7.5 $\mu m$ with RWC/CWC exceeding 0.5 at p28_16. The magnitudes and their evident dependence on the comparative role of cloud and rain are largely in line with the numerical analysis in [22], where the derived unimodal $r_e$ increases from 22 to 25 $\mu m$, as the volume
Fig. 3. (a) Median fractional bias of cloud effective radius at the level of cloud top. Median fractional bias is defined as the difference between the BSR lookup table estimate and the median of MCMC retrievals, which is then divided by the BSM estimate (i.e., (BSM–MCMC median)/BSM). The bars in color indicate the biases within the rain-occurring experiments, and the black bar next to the one in color on the left represents the biases for the same point but for the cloud-only experiment. (b) IQRs scaled by BSM estimate in percentage for \( r_e \) at cloud top. Yellow squares indicate the rain-present experiments. Black squares indicate the cloud-only experiments. Bars mark the increase of retrieval uncertainties resulting from rain effects. Legend provides the cloud effective radius and total optical thickness from BSM.

Fig. 4. (a) Median fractional bias of cloud effective radius versus the ratio between the MCMC-retrieved median of RWC and CWC. (b) Median fractional bias of cloud effective radius versus the retrieved median of cloud mode total number concentration at cloud top in the rain-present experiments. See legend for the colors indicating the cloud \( r_e \) estimated from BSM.

ratio between the rain and cloud mode is perturbed from 0.2 to 0.5 given cloud \( r_e \) at 20 \( \mu m \). A similar sensitivity of \( r_e \) biases to the ratio of water mass is also observed in [11], which validated satellite \( r_e \) retrievals with \textit{in situ} data collected over the Southern Ocean.

Meanwhile, we note that cloud total number concentration independently may be able to imply the relative robustness of the cloud mode within the bimodal DSDs, given that the \( N_t \) of the rain mode tends to be several orders of magnitude smaller than that of the cloud mode for the maritime shallow cumulus considered here. \( r_e \) biases reduce from 27.5% to 10% with cloud \( N_t \) increasing from 20 to 60 \( cm^{-3} \), and seemingly reaches an asymptotic minimum as \( N_t \) exceeds \( \approx 180 \text{ cm}^{-3} \). While the number of cases is limited, the tendency seems robust. Further studies are necessitated to determine under what circumstances cloud total number concentration may also be an effective indicator for \( r_e \) biases due to the existence of rain.

When the BSM derives the same \( r_e \) (e.g., p20_36, p20_16, p20_10), retrievals tend to be more accurate for optically thicker clouds [see Fig. 3(b)], as also noted by King and Vaughan [31]. The reflectance at 2.1 \( \mu m \) affords information primarily in regards to \( r_e \) via the droplet size controlling how strong the water absorption is, while neither \( R_{21} \) nor \( \partial R_{21}/\partial r_e \) is independent of \( \tau \). Along with clouds being more reflective, the sensitivity of \( R_{21} \) to droplet size grows, and isolines of \( r_e \) and \( \tau \) are increasingly orthogonal in Fig. 1, which implies a reduction in covariance (i.e., \( \partial R_{21}/\partial \tau \) and \( \partial R_{055}/\partial r_e \)). Retrieval uncertainties vary from 18% to 32% in cloud-only experiments. Constraints on \( r_e \) are weaker at p20_10 and p8_16 where the covariance is quite significant. The enlarged uncertainties due to the consideration of the rain mode, as indicated by the length of bars in Fig. 3(b), are shown to scale with decreasing \( R_{21} \), but are practically invariant with \( \tau \). A >20% increase in IQR is observed at p28_64 and p28_16, where the optical thicknesses are very different.
Fig. 5. (a) Joint PDFs of cloud effective radius ($R_{\text{e, top}}$) and single scattering albedo (SSA$_{\text{21, top}}$) at cloud-top level in the cloud-only experiment at p28_16. Joint PDFs of cloud effective radius ($R_{\text{e, top}}$) and similarity parameters ($S_{\text{top}}$) at cloud-top level in (b) cloud-only and (c) rain-present experiments at p28_16. (d) and (e) same as (b) and (c), but for experiments at p8_36.

### TABLE III

| DKL | Cloud top | Middle layer | Cloud bottom |
|-----|-----------|--------------|--------------|
| Cloud LWC | 1.849 | 1.052 | 1.069 |
| Rain LWC | 0.05 | 0.02 | 0.014 |
| Cloud $r_e$ | 0.629 | 0.008 | 0.006 |
| Rain $r_e$ | 0.19 | 0.12 | 0.003 |
| Cloud $N_e$ | 1.538 | 1.069 | 1.23 |
| Rain $N_e$ | 0.12 | 0.058 | 0.13 |

### IV. SIMILARITY PARAMETERS

The region deep within an optically thick cloud layer where diffuse radiation dominates is referred to as the diffusion domain. Considering the asymptotic expression of the reflection function at the water-absorbing wavelength (e.g., 2.1 $\mu$m) within this domain, all the asymptotic constants involved show a strong dependence on single scatter albedo ($\omega$) together with a weaker relation with asymmetry parameter ($g$, [4], [32]). Accordingly, a similarity parameter ($s$) was proposed taking into account both factors simultaneously, which is formulated as follows:

$$s = \left( \frac{1 - \omega}{1 - \omega g} \right)^{\frac{1}{2}}. \quad (7)$$

Consistent with the range of $\omega$, values of $s$ vary between 0 and 1. $s$ increases with stronger absorption (smaller $\omega$) and forward scattering (greater $g$), occurring with the prevalence of larger drops. In comparison with the nearly linear correlation between cloud effective radius and single scattering albedo at 2.1 $\mu$m [see Fig. 5(a)], the dependence of ($\partial R_{\text{21, top}}/\partial r_e$) on $r_e$ can be more truthfully represented via the curve characterized by similarity parameters [see Fig. 5(b)]. The measured reflectances are modulated by optical thickness, the albedo of the underlying surface, and sun/viewing geometry, while the sensitivity of $s$ to $r_e$ is at the core regarding the reliability of the BSM $r_e$.

When there are solely cloud droplets present, a robust relationship between the similarity parameters and cloud effective radii is manifest [see Fig. 5(b) and (d)], which corroborates that 2.1-$\mu$m reflectance is appreciably informative regarding $r_e$ near cloud top with the absence of a rain mode. With bimodal DSDs involved in the computation of the scattering phase function for the cloud volume [see Fig. 5(c) and (e)], this unique relationship becomes invalidated. Provided $s$ is 0.4, $r_e$ is estimated to be between 10 and 14 $\mu$m when clouds are precipitation-free [see Fig. 5(b)]. Given the same reflectance measurements and $s$, the PDF of derived $r_e$ extends to as low as 4 $\mu$m in Fig. 5(c) with the mode of $r_e$ moving below 10 $\mu$m, which implies the cloud $r_e$ in reality tends to be smaller than the prediction based on BSM if bimodal DSDs emerge with the occurrence of rain. As displayed, a BSM algorithm is liable to overestimate the cloud $r_e$ under such circumstances and retrievals are more uncertain with the enhanced possibility of multiple solutions. At pE where a considerable number of cloud droplets dominate, the coexistence of cloud and rain induces a marginal impact on the spread of $s$. Contrastingly at p28_16 where RWC/CWC is over 50%, the deviation from the well-established relation is substantial. To what degree the
existence of rain interferes with the physical reliance of cloud 
\( r_e \) on similarity parameters ultimately maps to the biases and 
retrieval uncertainties of various magnitudes shown in Fig. 3.

V. CALCULATION OF INFORMATION CONTENT

Shannon entropy \( (H) \) is commonly accepted as a quantita-
tive means to assess the information added by remote sensing 
observations beyond what is known \textit{a priori} within optimal 
estimation retrievals, which are premised on Gaussian error 
statistics [33], [34]. However, negative \( H \) may be yielded when 
the addition of observations in fact improves the knowledge of 
inferred variables, if the prior and posterior PDFs are non-
Gaussian and the mapping from observation to solution space 
is non-linear in retrievals [35]. Given many PDFs in our 
algorithms are indeed not in the form of Gaussian statistics, 
an alternative known as relative entropy (i.e., Kullback–Leibler 
divergence, [36]) is adopted, which is defined as

\[
D_{\text{KL}} = \int_{-\infty}^{\infty} p(x) \log \left( \frac{p(x)}{q(x)} \right) dx. 
\]

\( x \) indicates the variable to be assessed. \( p(x) \) and \( q(x) \) represent its posterior and prior PDFs, respectively. By its definition, 
\( D_{\text{KL}} \) is an information metric that quantifies the change relative 
to the prior as a consequence of measurements, which can 
be associated with the movement of the mode in posterior 
PDFs, the reduction in retrieval uncertainties, or the combined 
influence of both.

As exemplified with p20_16 (see Table III), the information 
from bispectral solar reflectances is concentrated on the cloud 
mode and little knowledge is delivered in terms of rain. For 
LWC and \( N_t \), \( D_{\text{KL}} \) of the cloud mode is more than a factor of 
10 greater than that of the rain mode. As such, retrieved rain 
microphysics to a great extent are determined by the prior. 
The \( D_{\text{KL}} \) value of \( r_e \) at cloud top is smaller than those of \( N_t \) 
and LWC, which is mainly due to the fact that the standard 
deviation of cloud \( r_e \) is less than 10 while \( N_t \) spans a few 
magnitudes in the prior adopted for retrievals. Consequently, 
less improvement resulting from observations is not surprising, 
given \( r_e \) is much better known \textit{a priori}. The occurrence of 
water absorption is heavily weighted toward cloud top and 
\( R_{21} \) measurements are mostly informative regarding cloud 
\( r_e \) within approximately five optical depths. Cloud \( N_t \) and 
LWC, dependent on \( r_e \) and \( \tau \) simultaneously, are therefore 
constrained by \( R_{21} \) and \( R_{055} \) combined. Their reductions of 
\( D_{\text{KL}} \) are less significant from the top to the middle cloud level, 
as scattering taking place deeper within clouds continues to 
contribute to the reflectances.

The information regarding the microphysical variables at 
cloud top gleaned from the bispectral reflectance observations 
declines with the magnitude of the absorption due to drop size 
(see Fig. 6). The amount of change in \( D_{\text{KL}} \) resulting from 
bimodality is limited below 30% within experiments where the 
BSM \( r_e \) is 8 or 12 \( \mu \)m, while the loss of information 
can be dramatic with respect to cloud effective radius, with 
the maximum reaching about 80%. LWC and cloud \( N_t \) are 
computed from the retrieved DSD parameters following (2) 
and (3), the retrieval of which is ultimately determined by 
the physical constraints of bispectral reflectances on cloud \( r_e \) 
and \( \tau \). As such, the errors in \( r_e \) propagate into derived LWC 
and \( N_t \) inevitably. In Fig. 6, Cloud \( N_t \) appears to be more 
subject to the presence of rain relative to LWC, which might 
be associated with its higher susceptibility to \( r_e \), given \( N_t \) 
and LWC are, respectively, proportional to \( \tau^{(1/2)}r_e^{-(5/2)} \) and \( r_e \tau \) 
[37], and rain impact is insignificant on \( \tau \).

VI. SUMMARY AND DISCUSSION

An approach built on the MCMC techniques allowed us 
to probe how bimodal DSDs present with precipitating 
liquid-phased clouds physically impose effects on the accu-
racies and uncertainties of \( r_e \) derived with the bispectral 
reflectance methods [4]. Vertical profiles of unimodal or 
bimodal DSDs are, respectively, inverted from an identical set 
of passive reflectances at 0.55 and 2.1 \( \mu \)m, under the assump-
tions of a single cloud mode or the co-occurrence of cloud 
and rain. The retrieved \( r_e \) in the MCMC experiments assuming 
unimodal DSDs largely replicate the estimate derived from a 
BSM algorithm. The impact of precipitation can be ascertained 
by comparing it with a corresponding experiment where rain 
is included in the inversion. As the observational information 
on \( r_e \) is mostly restricted near the optical top of the cloud, 
only \( r_e \) at the cloud-top level in the profile is used in the 
evaluation of biases and uncertainties. Generating the prior 
from the RICO field campaign enables the rain intensities 
under investigation to be extended to what is observed in 
shallow cumulus clouds, provided that the rain mode is largely 
driven by the climatological prior due to lack of constraints 
from reflectance measurements. Rain rates retrieved in the
upper third of the cloud vary from 1 to 34 mm/day with approximately a factor of 5 uncertainty in the examined cases, which is consistent with the natural variability presented in the RICO dataset.

BSM $r_e$ tends to overestimate the cloud mode $r_e$ in the bimodal DSDs, as also remarked in [22]. The biases can be as little as 0.3 $\mu$m ($\sim$4%) or as substantial as 7.5 $\mu$m ($\sim$28%), with a first-order dependence on the amount of absorption that occurred, which can be moderated given brighter clouds. Considered from the perspective of microphysics, the rain effects on the accuracies of the BSM derived $r_e$ are seemingly amplified with the rain mass composing an increasingly significant part of the total water mass, instead of being commensurate with the amount of rain water alone. Also, the median fractional biases of $r_e$ appear to be negatively correlated with cloud mode $N_e$. Both the high ratio between cloud and RWC and $N_e$ being large imply that the cloud mode dominates the DSDs. As such, little biases are incurred provided that the unimodal assumption entailed in the BSMs is not severely violated.

A key finding of this study is that the presence of rain can largely dampen the sensitivity of similarity parameters to $r_e$, the validity of which permits an accurate retrieval of $r_e$ fundamentally. It is found that the cloud mode $r_e$ has a tendency to be lower than the estimate according to the BSM and the reliance of similarity parameters on $r_e$ is considerably randomized, when cloud and rain coexist. The ambiguities caused by the bimodality become remarkable, as 2.1-$\mu$m reflectance drops to 0.18 where the BSM $r_e$ is about 20 $\mu$m in our simulations. With the underlying physics obscured by the existence of rain, the retrieval uncertainties can be enlarged by a factor of 2 and the information content with regards to $r_e$ can be reduced by nearly 80%. The bottom-line result is that the presence of rain that tends to cause high biases in $r_e$ cannot be ultimately distinguished from actual changes in the cloud mode. In essence, the biased $r_e$ cannot be separated into weighted contributions from the cloud and rain modes without additional independent information.

Reflected sunlight measurements alone are not expected to be capable of separating the cloud and precipitation signals. A readily effective manner to detect and inform about precipitation is to incorporate information that is dependent on a DSD weighting that is different from the reflectance measurements such as coincident active or passive microwave observations. Radar reflectivity, for instance, is approximately proportional to the sixth moment of the DSDs and renders radar reflectivity as being highly sensitive to the largest-size drops in warm-phase clouds. Radar Doppler spectra are also demonstrated to possess information on identifying cloud and drizzle modes [38], [39]. Some previous studies have synergistically combined radar reflectivity, shortwave radiance with lidar [40], [41], or microwave brightness temperature observations [24], [42] to infer concurrent cloud and drizzle, while the emphasis was not on investigating the improvement of the BSM retrievals in a more integrative observing system. In the near future, we plan to further explore how the addition of active radar observations may reduce the biases and uncertainties of cloud $r_e$ estimated from the BSM in the precipitating scenario. Simply knowing whether precipitation is present or not may allow for a significant advance in establishing the validity of BSM retrievals.

REFERENCES

[1] I. L. McCoy et al., “The hemispheric contrast in cloud microphysical properties constrains aerosol forcing,” Proc. Nat. Acad. Sci. USA, vol. 117, no. 32, pp. 18998–19006, Aug. 2020, doi: 10.1073/pnas.1922502117.
[2] R. Pincus, S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, “Reconciling simulated and observed views of MODIS, ISCCP, and the limits of instrument simulators,” J. Climate, vol. 25, no. 13, pp. 4699–4720, Jul. 2012, doi: 10.1175/JCLI-D-11-00267.1.
[3] S. Platnick et al., “The MODIS cloud products: Algorithms and examples from Terra,” IEEE Trans. Geosci. Remote Sens., vol. 41, no. 2, pp. 459–473, Feb. 2003, doi: 10.1109/TGRS.2002.808301.
[4] T. Nakajima and M. D. King, “Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory,” J. Atmos. Sci., vol. 47, no. 15, pp. 1878–1893, 1990, doi: 10.1175/1520-0469(1990)047<1878:DOTOPT>2.0.CO;2.
[5] D. Painemal and P. Zuidema, “Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements,” J. Geophys. Res., Atmos., vol. 116, no. D24, Dec. 2011, Art. no. D24206, doi: 10.1029/2011JD016155.
[6] Q. Min et al., “Comparison of MODIS cloud microphysical properties with in-situ measurements over the Southeast Pacific,” Atmos. Chem. Phys., vol. 12, no. 23, pp. 11261–11273, Dec. 2012.
[7] N. J. King, K. N. Bower, J. Crosier, and I. Crawford, “Evaluating MODIS cloud retrievals with in situ observations from VOCALS-REX,” Atmos. Chem. Phys., vol. 13, no. 1, pp. 191–209, Jan. 2013, doi: 10.5194/acp-13-191-2013.
[8] M. K. Witte et al., “MODIS retrievals of cloud effective radius in marine stratocumulus exhibit no significant bias,” Geophys. Res. Lett., vol. 44, no. 19, pp. 1–9, Oct. 2017, doi: 10.1002/2017GL079325.
[9] S. Glienke, A. Kostinski, J. Fuglai, R. A. Shaw, S. Bormann, and J. Stith, “Cloud droplets to drizzle: Contribution of transition drops to microphysical and optical properties of marine stratocumulus clouds,” Geophys. Res. Lett., vol. 44, no. 15, pp. 8002–8010, Aug. 2017, doi: 10.1002/2017GL074430.
[10] E. Ahn, Y. Huang, S. T. Siems, and M. J. Manton, “A comparison of cloud microphysical properties derived from MODIS and CALIPSO with in situ measurements over the wintertime southern ocean,” J. Geophys. Res., Atmos., vol. 123, no. 19, p. 11, Oct. 2018, doi: 10.1002/2018JD028535.
[11] L. Kang, R. Marchand, and W. Smith, “Evaluation of MODIS and Himawari-8 low clouds retrievals over the southern ocean with in situ measurements from the Socrates campaign,” Earth Space Sci., vol. 8, no. 3, Mar. 2021, Art. no. e2020EA001397, doi: 10.1029/2020EA001397.
[12] A. Marshak, S. Platnick, F. Várnai, G. Wen, and R. F. Cahalan, “Impact of three-dimensional radiative effects on satellite retrievals of cloud droplet sizes,” J. Geophys. Res., vol. 111, no. D9, 2006, Art. no. D09207, doi: 10.1029/2005JD006686.
[13] Z. Zhang, A. S. Ackerman, G. Feingold, S. Platnick, R. Pincus, and H. Xue, “Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations,” J. Geophys. Res., Atmos., vol. 117, no. D19, Oct. 2012, Art. no. D19208, doi: 10.1029/2012JD017655.
[14] Z. Zhang et al., “A framework based on 2-D Taylor expansion for quantifying the impacts of subpixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the spectral method,” J. Geophys. Res., Atmos., vol. 123, no. 7, pp. 7007–7025, Jun. 2016, doi: 10.1002/2016JD024837.
[15] R. F. Cahalan, W. Ridgway, W. J. Wiscombe, and S. Gollmer, “Independent pixel and Monte Carlo estimates of stratocumulus albedo,” J. Atmos. Sci., vol. 51, pp. 3776–3790, 1994, doi: 10.1175/1520-0469(1994)051<3776:IPAMCE>2.0.CO;2.
[16] L. Di Girolamo, L. Liang, and S. Platnick, “A global view of one-dimensional solar radiative transfer through oceanic water clouds,” Geophys. Res. Lett., vol. 37, no. 18, Sep. 2010, Art. no. L18809, doi: 10.1029/2010GL044094.
[17] L. Liang, L. Di Girolamo, and W. Sun, “Bias in MODIS cloud drop effective radius for oceanic water clouds as deduced from optical thickness variability across scattering angles,” J. Geophys. Res., Atmos., vol. 120, no. 15, pp. 7661–7681, Aug. 2015, doi: 10.1002/2015JD023256.
