

Earth and Space Science

RESEARCH ARTICLE

10.1029/2019EA000900

Key Points:

- A bilateral filter scheme that was developed to compress cloud radar noise is applied to CloudSat hydrometeor detection
- This method decreases background noise but retains true signals, thus increases the accuracy of cloud mask, especially for cirrus
- This method would lead to an improvement of the retrievals of cirrus microphysical properties

Correspondence to:

J. Ge, gejm@lzu.edu.cn

Citation:

Hu, X., Ge, J., Li, Y., Marchand, R., Huang, J., & Fu, Q. (2020). Improved hydrometeor detection method: An application to CloudSat. *Earth and Space Science*, 7, e2019EA000900. https://doi.org/10.1029/2019EA000900

Received 19 SEP 2019 Accepted 28 JAN 2020 Accepted article online 03 FEB 2020

© 2020 The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Improved Hydrometeor Detection Method: An Application to CloudSat

Xiaoyu Hu¹, Jinming Ge¹, Yanrong Li¹, Roger Marchand², Jianping Huang¹, and Qiang Fu²

¹Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou, China, ²Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA

Abstract Clouds play an important role in the climate system and are a principal source of uncertainty in climate projections. CloudSat has provided an unprecedented opportunity to study the vertical structure of clouds, and its observations are being widely used in scientific studies. However, some clouds are not detected or are only weakly detected by CloudSat. In most studies, the weakest detections, specifically those detected by the so-called along-track integration scheme, are typically ignored due to the high rate of false detections, namely, a significant probability that a detected cloud is actually a region of increased measurement noise, rather than a true cloud signal. False detections have been reduced in the latest version (called R05 for release 5) of the CloudSat cloud mask product but at a cost of a significant loss in the true weak signals (i.e., a higher false omission rate). In this study, the CloudSat hydrometeor detection algorithm used in R05 is modified by adding a bilateral filter scheme to improve the detection of weak signals. By comparing with the CALIPSO lidar vertical feature mask, it is shown that the new scheme largely reduces the false detection rate compared to the R04 version, while retaining a large fraction of the true weak signals that have been lost in the R05 version. Implementing this scheme in future CloudSat data processing is expected to lead to a better detection of thin clouds.

Plain Language Summary Cloud is important to the Earth-atmosphere system. Different types of cloud may produce opposing effects, cooling or warming the system. But the simulation of clouds in weather forecast and climate prediction remains challenging, due to the numerous nonlinear processes that govern cloud formation and evolution. Continued accurate observation of clouds remains crucial to improve our knowledge of the processes that control clouds. Data from CloudSat, which carries a 94 GHz cloud radar, are being widely used in climate research. A first step in using the radar data is to separate signals that correspond to energy reflected from real cloud and precipitation from noise in the measurement. This is difficult when the signals from clouds are of similar magnitude to the noise. In this study, we report on an improved algorithm to separate real signals from noise by using a bilateral filter. This improved method can increase the accuracy of CloudSat cloud detection of weak signals and will hopefully lead to a better understanding of thin cloud in the climate system.

1. Introduction

Clouds are a critical component of the Earth-atmosphere system. They modulate the energy budget by reflecting solar radiation and trapping terrestrial radiation through the interactions between cloud particles and radiation and are thus critical to understanding climate (Fu et al., 2000; Fu et al., 2000; Huang et al., 2014; Rosenfeld et al., 2019; Su et al., 2008). The formation of clouds and phase transitions from liquid-to-ice heat the atmosphere while evaporation of clouds and precipitation play the opposite role, also making clouds important as regards both atmospheric dynamics and hydrological cycle (Fu et al., 2002; Huang et al., 2015; Marchand, 2012; Worden et al., 2007). However, the mechanisms that govern the evolution of clouds are associated with microphysical processes and dynamical interactions that span spatial scales from the microscopic through the largest atmospheric waves and circulations and make the representation of cloud processes difficult in general circulation models (Fu & Hollars, 2004; Hourdin et al., 2017; McCoy et al., 2016; Randall et al., 2003).

In the context of increasing amounts of atmospheric carbon dioxide and global warming, cloud properties such as amount, height, location, phase, and particle size are expected to change in a variety of ways, with evidence for some of these changes now starting to be found in satellite observations (Gupta et al., 2018;

Liu et al., 2018; Marchand, 2013; Norris et al., 2016). These changes are associated with a series of feedbacks that will amplify or dampen global warming, which leads the largest source of uncertainty in current predictions of climate sensitivity (Fu, 2013; Huang et al., 2006; Stevens & Bony, 2013; Tan et al., 2016). For example, a warmer surface promotes strong convection that may increase high-cloud amount, which would reduce terrestrial radiation emitted to space and generate a positive feedback (Dessler & Sherwood, 2004; Hanisco et al., 2007; Huang et al., 2017). However, global warming may lead to increases in the occurrence of clouds composed of liquid water rather than ice (and possibly clouds with more liquid water), which makes clouds more reflective to solar radiation and thus induces a negative feedback (Ceppi et al., 2016; McCoy et al., 2014). Narrowing the range in estimates of cloud feedback requires accurate long-term observations to understand and constrain physical mechanisms responsible for those changes (Bony et al., 2006; Kollias et al., 2005). In particular, satellite observations over decades are necessary to directly detect expected changes and separate them from natural variability (Geer et al., 2017; Winker et al., 2017).

With the aim of improving knowledge of cloud properties, especially their vertical structure, CloudSat carrying a Cloud Profiling Radar was launched on 28 April 2006, providing active profiling of clouds from space (Stephens et al., 2002). The raw return power measured by the radar is a combination of energy reflected from hydrometeors (cloud and precipitation) and background noise. To be useful for science studies, the hydrometeor signals first need to be distinguished and extracted from the noise, that is why a cloud detection algorithm is applied. For CloudSat, a "cloud mask" is provided in the 2B Geometric Profile (GEOPROF) cloud mask product (Marchand et al., 2008), hereafter M08, and is one of the primary data sets used for cloud research. Thousands of papers have now been published based on the CloudSat observations (Stephens et al., 2018).

The GEOPROF cloud mask product provides a value between 0 and 40 for each radar range bin, with increasing values indicating a reduced probability of a false detection). However, in most studies, range bins with cloud mask values between 6 and 10 are typically ignored because of its high false detection rate, 44% compared with CALIPSO for GEOPROF R04 (see Table 1 in M08). These low-confidence cloud detections are identified using an along-track integration scheme, which is designed to detect horizontally extensive clouds whose signal-to-noise ratio (SNR) is near or below 1 dB at the nominal 1.4×1.7 km radar resolution [M08]. Namely, the noise is equal or larger than the true cloud signal and the cloud can only be detected by further reducing the noise through averaging. Many of these weak detections are cirrus clouds and liquid clouds with small droplets. For the remainder of this paper, we use the terms "weak signals" or "weak detections" to mean clouds detected by the along-track integration scheme (cloud mask value 6–10) and "strong signals" to mean detections which do not rely on this integration schemed (cloud mask value greater than 10).

The false detection rate for weak signals has been reduced in the newest version of the CloudSat products, release 5 (R05), by improving estimates of the radar noise variance and optimizing filter parameters. However, as is shown below, this decrease in false detections comes at the cost of a large reduction in the identification of true signals, meaning the false detections were reduced but so was the amount of correctly identified clouds (details in section 3). Independently, Ge et al. (2017), hereafter G17, proposed an improved hydrometeor detection algorithm by adopting a bilateral filter, which is initially used in images process (Tomasi & Manduchi, 1998) to improve the weak signal detection. It has demonstrated good performance when applied to ground-based Ka-band cloud radar data collected at Semi-Arid Climate and Environment Observatory of Lanzhou University (Ge et al., 2018, 2019; Huang et al., 2008; Zhu et al., 2017), which can reduce radar noise while preserving cloud edges. In this study, we modified and applied this bilateral filter scheme to CloudSat hydrometeor detection algorithm in order to decrease the false detection rate of weak signals but preserve the detected real signals in the cloud mask. This method is given in section 2. Section 3 demonstrates several case comparisons and multiorbit evaluations. Section 4 presents a summary of this study and discussion.

2. Data and Methods

2.1. Data Sets

The input to the hydrometeor detection algorithm is the measured raw radar return power (P_{raw}) in the level 1B Cloud Profiling Radar data product. Here, the resulting cloud mask using the improved hydrometeor

detection from G17 (described below as the G17 version) is compared with the operational CloudSat R04 and R05 version cloud masks, which are stored in level 2B GEOPROF data products. Following M08, we use the CALIPSO version 4 lidar vertical feature mask (VFM) (Getzewich et al., 2018; Vaughan et al., 2004) as a reference to evaluate hydrometeor detection algorithm, because the lidar is more sensitive to thin cloud than the radar. For most of their operational life (and for all the data examined here), CALIPSO and CloudSat were both parts of the A-Train satellite constellation. The VFM product identifies where there is backscatter from cloud (value of 2), aerosol particle (value of 3), or stratospheric feature (value of 4) and where the signal does not penetrate the whole cloud layer (value of 7). We also use CALIPSO version 4 level 2, 5 km cloud profile product (L2_05km_CPro) (Vaughan et al., 2009; Young et al., 2018) to identify cirrus cloud to help make the evaluation.

2.2. Hydrometeor Detection Algorithm

The CloudSat hydrometeor detection algorithm is described in M08 and is based on an earlier method developed for ground-based radar (Clothiaux et al., 1995, 2000) but with two significant changes, a power probability weighting scheme, and an along-track integration scheme. This algorithm was optimized for improving estimates of the radar noise variance and filter parameters in R05. In this study, we modified the algorithm by adding a bilateral filter to the CloudSat along-track integration scheme as shown in the schematic diagram of Figure 1. The first several steps (estimating noise level, generating initial mask, and applying spatial filter) are the same as R05 version algorithm, until one reaches the point in the R05 algorithm when along-track averaging is used.

The algorithm in Figure 1 works as follows. The mean noise power (P_n) and its standard deviation (σ_{p_n}) are first calculated using a moving average applied to the measured raw radar return power (P_{raw}) from the stratosphere at each along-track sample, with a box that is 10 range bins in the vertical and up to 2 along-track bins wide. An initial mask is then created by comparing P_{raw} with the estimated noise. Specifically, range bins where $P_{raw} > P_n + \sigma_{p_n}$ are assumed to be cloudy (contain a hydrometeor) and are set to a cloud mask value of 20, 30, or 40 if P_{raw} are greater than $P_n + \sigma_{p_n}$, $P_n + 2\sigma_{p_n}$, or $P_n + 3\sigma_{p_n}$, respectively. As a first guess, range bins where $P_{raw} \le P_n + \sigma_{p_n}$ are likewise taken to have no cloud and given a value of 0 in the initial mask.

A range bin in initial mask has a 16% probability to be falsely masked as cloudy bin based on the Gaussian distribution assumption of the radar noise. Note that the noise is randomly distributed while hydrometeors are highly correlated in time and space. One can take advantage of this correlation to further reduce the occurrence of false detections. A spatial filter (with five bins in the vertical and seven bins in the horizontal) is then used to remove false detections in the initial mask. A cloudy range bin in the initial mask is kept, if it is surrounded by more cloudy range bins than N_{thresh} (20 was used following M08). Otherwise, the cloudy range bin is considered to be false detection and the mask value is set to 0 after all range bins have been evaluated. The confidence value in the initial mask is also taken consideration, by the so-called power probability weighting scheme. A range bin initially masked as 40 will have a smaller probability being a false detection (less than 16%) compared to a range bin with an initial value of 20; thus, a higher threshold (more empty cloud neighbors) are required before a range bin will be declared empty (set to 0), if it was initially identified as cloudy at level 30 or 40 (details are given in M08). Likewise, a range bin that was initially empty can be turned on (with a mask value set to 20), if it is surrounded by many range bins that are cloudy. This spatial filter is applied several times (typically three).

The two steps above can detect most of the cloud-filled range bins, except some very weak signals (cloud mask between 6 and 10), that are close to the sensitivity limit of the radar. An along-track integration scheme is then used to detect these very weak clouds in R04 and R05, by averaging the $P_{\rm raw}$ horizontally, that is averaging over time at a given altitude above mean sea level. In R04 and R05, the average includes all range bins regardless of whether or not they have sufficient power to have been identified as potentially cloudy in the initial mask. In brief, the raw radar power is averaged (effectively reduced in temporal resolution by summing neighboring columns) and a new (coarse) resolution mask is generated using the same procedure described in the preceding paragraph. The new (coarse) resolution mask is then merged with the previous (full resolution) mask, by giving range bins in the previous mask which were empty (value 0) but now have a detection in the coarse mask, a value between 6 and 10, depending on averaging steps. However, the along-track averaging of all range bins can cause range bins with little return power, which is having power due



Figure 1. Schematic flow diagram for cloud mask. P_n and σ_{p_n} are the mean and standard deviation of noise, respectively. The red box indicates where the bilateral filter is introduced.

only to background noise, to be incorrectly identified as cloud, because neighboring potentially cloudy range bins can contribute enough so that the averaged power (potentially cloudy + not cloudy in initial mask) becomes significantly larger than the noise. Not only has the background noise been raised but also cloud edges been blurred by this unconditional averaging, since both signal and noise are mixed together around cloud edges. The merging process is therefore restricted to adding range bins where the averaging includes no clouds in the full resolution mask.

Here we apply a bilateral filter before averaging the power, which was added to the cloud detection algorithm in G17 to preserve the cloud edges while reducing the radar noise. The approach basically follows the same idea with along-track averaging, except that averaging in the along-track direction is conditioned upon the radar range bins being initially identified as cloudy or not cloudy. For radar range bins which do have strong signal to initially be identified as cloudy $(P_{raw} > P_n + \sigma_{p_n})$, only neighboring range bins with $P_{raw} > P_n + \sigma_{p_n}$ are averaged and vice versa; if a volume is initially thought to be not cloudy $(P_{raw} \le P_n + \sigma_{p_n})$, only neighboring range bins with $P_{raw} \le P_n + \sigma_{p_n}$ are averaged. The advantage of this conditional averaging is that it decreases the variance of the noise (σ_{p_n}) without blurring cloud edges (i.e., averaging power from range bins with strong scattering into those with no real detections) and yet still allowing range bins which are initially identified as not cloudy $(P_{raw} \le P_n + \sigma_{p_n})$ to potentially be identified as cloudy (after averaging). This is because noise level (σ_{p_n}) goes down with averaging while the power level of weak signal does not change much due to highly correlated cloud range bins in time and space. It also promotes the efficiency of CloudSat cloud mask processes by removing the need for the merging coarse resolution mask to consider only pixels which are nonadjacent to cloud in the full resolution mask.

3. Results

3.1. Case Study

Case comparisons between CloudSat and CALIPSO were shown in M08 when describing the R04 version algorithm. We use the same colocation method as M08 to re-examine one of these cases, featuring thin cirrus and consisting largely of weak signals (Figure 2 in the present paper and Figure 8 in M08), and also examine another three cases featuring weak signals, strong signals, and stratocumulus (Figures 3–5, respectively).

Figure 2 highlights the strengths and weakness of the along-track integration scheme. Clouds are difficult to identify in the measured power (Figure 2a). Considering the low SNR, it is not surprising that three versions



Figure 2. Example of thin cirrus on 3 October 2006 (same as Figure 8 in M08). CloudSat radar return power (a). CloudSat cloud mask from R04 (b), R05 (c), and G17 (d). CALIPSO VFM (e). Difference mask between VFM and R04 cloud mask (f), R05 cloud mask (g), and G17 cloud mask (h). The black line in (b)–(d) is the surface (ground clutter) and is not considered a detection.

of cloud mask (Figures 2b–2d) all fail to detect much of the cirrus layer identified in the VFM (Figure 2e). Note that our focus in this paper is primarily on the weak detections, that is, cloud mask between 6 and 10, detected by the along-track averaging scheme, colored blue in Figures 2b–2d. The bias in the binary cloud masks (CALIPSO-CloudSat) is shown in Figures 2f–2h for R04, R05, and the new mask with the bilateral filter from G17, respectively. Range bins where CALIPSO identified cloud but CloudSat did not appear red; we refer to these as false negatives. Blue bins, on the other hand, represent false positives, that is, range bins where the CloudSat masks indicate the presence of hydrometeors but CALIPSO does not. Canary yellow bins represent those range bins where cloud (or precipitation) is detected by both sensors, that is, true positives, while true negatives, detected by neither, are colored white.

Among the three radar cloud masks, the R04 (Figures 2b and 2f) detects the most weak signals (masks valued 6–10, blue bins) but with a 19.6% false detection rate of weak signals (FDRW, fraction of weak false positives to total weak signals). R05 (Figures 2c and 2g) has a much cleaner background than R04, with a 0.4% FDRW, but has lost many of the true detections (55.1%) as well. G17 (Figures 2d and 2h) has little noise in the background (0.1% FDRW), with even more true signals (9.1%) being detected than R04. Further, the weak detections connect better with (are more adjacent to) the strong detections (cloud mask values >10) than in the R04 and R05 masks. We examined the minimum SNR of true weak signals. It is found that the SNRs for the three versions are basically the same (about -0.59 dB). Namely, the bilateral filter does not extend the minimum SNR of weak true signal. Rather, it effectively reduces more false detections compared with R04, thus allowing a lower threshold to be used in along-track averaging scheme and consequently detects more true signals than R05.

Figure 3 shows an example with two thinner cirrus clouds than in Figure 2. Much of the cirrus in Figure 3 are close to or below the sensitivity limit of the radar and can barely be noticed in the raw radar return power (Figure 3a). We note that the blocky appearance of the cloud in the lidar data (Figure 3e) occurs because, even for the lidar, much of the cloud volume is being detected via along-track averaging (Vaughan et al.,



1**0**C

Earth and Space Science



Figure 3. As in Figure 2 but for two thinner clouds on 4 July 2016.



Figure 4. As in Figure 2 but for a case with strong signals on 25 March 2011.





Figure 5. As in Figure 2 but for a case with stratocumulus on 2 February 2011.

2009). For the cloud around 20°S, we see that R04 detects the most cloud among the three versions and R05 detects the least, with again, many more false detections in R04 than in R05. The G17 version decreases the false detections compared to R04 (FDRW from 42.2% to 8.8%, with R05 being 24.3%), while retaining much of the true cloud (a 4.8% decrease of true weak detections for G17, while 65.3% for R05, compared to R04).

From the two examples above (Figures 2 and 3), it can be seen that our improved method does well in detecting weak clouds and decreasing random background noise. But one might be concerned that focusing on the detection of weak signals might cause some false detection around the edges of thick clouds. Figure 4 shows a case of strong signals to address this question. Very different from previous cases, there are many orange bins in Figures 4b–4d, which indicates that the cloud mask confidence is 40 and there is little probability of a false detection (<1%) for these range bins. Note that lidar cannot penetrate optically thick clouds, denoted by the red regions shown in Figure 4e. As in M08, we take any cloud detected by CloudSat when the lidar is attenuated to be a true positive (in this figure and section 3.2).

There are two clouds in this case, north and south of 72°N. For the northern (leftmost) one, the three versions of cloud mask (Figures 4b–4d) are all very similar as there are essentially no weak detections associated with this cloud. The small number of false positives and negatives around the cloud boundary (Figures 4f–4h) may be caused by the mismatches in time and resolution between CloudSat and CALIPSO, though R04 has slightly more false detections here. The southern (rightmost) cloud features a convective cloud with an extended anvil. G17 are able to detect more (11.5%) optically thin ice cloud at the top of the anvil than R04 (from the perspective of CloudSat, actually still far away from the top detected by CALIPSO, Figure 4e), while almost all of which is lost in R05 (95.4%). In summary, this case shows that G17 is able to preserve sharp boundaries due to strong scatters while also detecting neighboring weak signals when such are present.

From the above three cases (Figures 2–4), it can be seen that the G17 version shows good ability to detect weak signals. However, many of these weak signals are adjacent to strong signals. In general, it is a natural for weak signals to surround strong signals, because cloud boundaries are often marked by regions where cloud droplets begin to dissipate due to entrainment and have lower reflectivity and are therefore hard to detect (Chernykh et al., 2001; Pinsky & Khain, 2019). In the absence of precipitation or large ice particles,





(a) Coverage of true weak signal from R04

Figure 6. Coverage of true weak signal from R04 version (a), R05 version (b), and G17 version (c). Percentage in the upperright corner of each panel is the global mean (treating each CloudSat measurement equally. The data are not equal-area weighted). The resolution is 4.0° latitude by 6.0° longitude.

cloud composed of only small liquid droplets typically have low reflectivity and are composed entirely of weak radar signals (in spite of being optically thick). Figure 5 shows an example of such a stratocumulus cloud, with a cloud top height around 4 km. CALIPSO (Figure 5e) detects the slowly decreasing cloud top height from north to south, while all three versions of CloudSat mask (Figures 5b-5d) miss much of this cloud layer, especially from 28°N to 26°N, only show some discontinuous signals, away from strong signals. We note that this cloud is high enough above the surface that ground clutter is not a factor. The aggressive R04 mask (Figures 5b and 5f) detects the more cloud between 28°N and 26°N than the other mask and also has a high FDRW of 42.1%. R05 (Figures 5c and 5g) and G17 (Figures 5d and 5h) both have clearer background than R04 (FDRW is 9.5% and 22.3%, respectively) but with about 83.5% and 37.9% loss of true weak signals compared to R04. As discussed in section 2.2, the bilateral filter is designed to preserve the edge between strong signals $(P_{\text{raw}} > P_n + \sigma_{p_n})$ and potential noise $(P_{\text{raw}} \le P_n + \sigma_{p_n})$; thus, it improves the detection of weak signals near strong signals but does little to improve detection of isolated and weak signals.

3.2. Multiorbit Evaluation

To further demonstrate the performance of our improved method, Figure 6 shows the true weak signal coverage for each version, which is the ratio of the number of radar profiles containing a true weak signals to the



(a) Zonal vertical frequency of true weak signal from R04

Figure 7. Zonal vertical frequency of true weak signal from R04 version (a), R05 version (b), and G17 version (c). Solid and dash lines are mean cloud top and base height of single-layer cirrus cloud. The vertical resolution is 1 km, and the horizontal resolution is 2.0° latitude.

total number of radar profiles, using 5,315 orbits data (orbit number from 3,607 to 8,921) for the whole year of 2007. All of them show similar patterns to high-cloud coverage (Mace et al., 2009) indicative of the large degree to which detection of weak signals from cirrus dominates this statistic. Although the patterns are similar, the global coverage of R05 (Figure 6b, 3.6%) is considerably less than the other two (R04 is 15.6% and G17 is 15.4%). This is because, as shown in section 3.1, R05 loses a large amount of true weak signals compared with R04, while G17 maintains these true weak signals.

Figure 7 shows zonal vertical frequency of true weak signal detection for the three versions, which is the number of range bins containing true weak signal divided by the numbers of total range bins. Superimposed is the mean cloud top (solid line) and base height (dash line) of single-layer cirrus identified through the criterion that cloud optical depth does not exceed 3 and cloud base temperature is lower than -40 °C from CALIPSO version 4 level 2, 5 km cloud profile product (L2_05km_CPro) (Gasparini et al., 2018). Note that polar stratospheric clouds detected by CALIPSO, which are optically very thin (Kato et al., 2010; Kohma & Sato, 2011; Vaughan et al., 2019), cause the cirrus top and base lines to rise poleward of 60°S. Overall, G17 (Figure 7c) maintains a similar pattern as R04 (Figure 7a). The main difference here between G17 and R04 is the somewhat greater rate of weak detections near the surface (which as Figure 5



Figure 8. Relative changes of true (a, b) and false (c, d) weak signals. Positive values indicate increase, while negative values denote decrease. Percentage in the upper-right corner of each panel is the global mean, not equal-area weighted. The resolution is 4.0° latitude by 6.0° longitude.



Figure 9. Weak signal composition for the three versions. The FDRWs are marked on the top of each column.

demonstrates comes with a large increase in FDRW and is likely due to non or weakly precipitating stratocumulus). As expected from Figure 6, R05 (Figure 7b) loses many true signals compared with R04 and G17.

The global patterns of true weak signals in Figures 6 and 7 show a large difference between R05 and G17. In Figure 8, we examine the relative changes of true and false weak signals computed as (R05-R04)/R04 and (G17-R04)/R04. As expected, R05-R04 (Figure 8a) shows large negative values, indicating a significant loss of true cloud signals in R05 that is uniformly distributed over the globe. The relative change for G17-R04 (Figure 8b), on the other hand, has a clear latitudinal dependence, where there is an increase of the true weak signals in G17 in the Intertropical Convergence Zone and middle to high latitude of both hemispheres poleward of 30°, while there is a decrease of true weak signals in subtropics. Thus, as one might expect from our case studies, G17 is doing somewhat better where there is a mixture of strong and weak signals, while it loses some true isolated weak signals compared with R04 version, in particular associated with nonprecipitating stratocumulus. As for the false weak signals, false weak signals are mainly caused by random noise and are largely independent of latitude and longitude for both R05 and G17 (Figures 8c and 8d).

The global results are summarized in Figure 9. The black columns represent the numbers of true signals, and gray columns are the false signals. The FDRWs are marked on top of each column. One can see that the total number of weak signals (the whole column) is dramatically decreased from R04 to R05 version from about 141M range bins dropping to 20M (a reduction of 85.9%) while a smaller reduction of 31.4% (~141M to 97M) from R04 to G17 version. Ideally, this reduction should be mainly due to the removal of false signals, leaving the number of true signals as untouched as possible. However, the number of true signals (the black column) in R05 is considerably decreased compared with R04 (i.e., reduced from ~70M to about 14M, a reduction of 80.0%), which can also be seen in Figures 6–8. In G17, the number of true signals is 67M which is comparable to 70M in R04 (a small change of about 4.6%) and the FDRW is only 1.2% larger than that of R05. Thus, G17 can detect more than four times of the true detections than current R05 version and yet is still able to reduce the FDRW from about 50.1% to 30.5%.

Figure 10a shows the zonally averaged FDR of total signals (strong + weak) for R04 version (Figure 10a1), which is defined as the number of all false positives divided by the sum of false positives and true positives, and the differences of (R05–R04, Figure 10a2) and (G17–R04, Figure 10a3). The FDR values of R04 are



Figure 10. Zonally averaged false detections rate (FDR, column a), false omission rate (FOR, column b) and accuracy (ACC, column c) of all detections of R04 (row 1) and differences of R05–R04 (row 2), and differences of G17–R04 (row 3). Positive values in the difference panels indicate that R05 and G17 version have a higher value. Percentage in the upper-right corner of each panel is the global mean, not equal-area weighted. The vertical resolution is 1 km, and the horizontal resolution is 2.0° latitude.

relatively high in the subtropics where cloud occurrence is low (i.e., less true positives) (Hagihara et al., 2010; Mace et al., 2007), though the false positives are randomly (spatial uniformly) caused by along-track averaging scheme. In Figures 10a2 and 10a3, it is explicit that both FDRs for R05 and G17 versions are decreased globally, and unsurprisingly, R05 has a stronger decline (-5.3%) than G17 (-3.7%), especially in the subtropics where few clouds occur.

We further show the false omission rate (FOR), which is defined as the number of false negatives divided by the sum of false negatives and true negatives in Figure 10b1 to describe the proportion of incorrectly marked clear range bins. A high-FOR band appears coincidently with the cirrus cloud layer (except poleward of 60°S, where optically very thin PSC occurs). It reaches the maximum at tropical troposphere, where there are a large number of cirrus clouds (Dessler & Yang, 2003; Fu et al., 2007). A high-positive band (mostly more than 1.5%) is clearly around cirrus cloud base in Figure 10b2, indicating loss of true positives in R05. Contrary to this positive region, our G17 version shows slight negative values, meaning more true positives around cloud base than R04. Neither R05 nor G17 shows any significant improvement in the upper level of cirrus cloud, where ice crystals are smaller (Garrett, 2003). The top of cirrus clouds detected by CALIPSO is always missed by all the three versions of CloudSat mask (e.g., Figures 3 and 4), so the improvement of G17, mostly at the lower level of cirrus cloud form CALIPSO's perspective, is mainly at the cloud top boundary detected by CloudSat.

The accuracy (ACC) of CloudSat cloud mask, which is defined as the sum of true positives and true negatives divided by the total sample range bins, can provide an overall evaluation in a comprehensive perspective by considering both how much false signals are marked (FDR) and how much true signals are lost (FOR). Increased (decreased) FDR or FOR will lead to a drop (rise) of ACC. The zonally averaged ACC (Figure 10c1) shows similar distribution as FOR (Figure 10b1), due to the uniformly spatial distribution of FDR (Figure 10a1). Figure 10c2 shows that the ACC is raised almost globally for R05 version, except at high altitudes where decreased FDR is offset by the increase of FOR (i.e., the algorithm lost more true positives than false positives). Unlike Figure 10c2, there is no negative value in Figure 10c3, indicating that the accuracy of cloud mask in G17 version is significantly increased (0.4%, while 0.1% in R05 version) in a balanced way (both FDR and FOR are reduced simultaneously) by applying the bilateral filter scheme.

4. Summary and Discussion

We have added a new bilateral filter scheme to the operational CloudSat hydrometeor detection method to improve the along-track integration scheme. The bilateral filter scheme produces a cleaner background than R04 version (fewer false positives) and more true signals than R05 version (fewer false negatives). The signals that R05 loses but the bilateral filter scheme detects are mostly associated with cirrus clouds (that are detected by CALIPSO). Overall, the bilateral filter improves the global cloud distribution pattern of weak signals, reaches a better balance between false detection rates and false omission rates, and raises the accuracy of cloud mask. For the quantitative evaluation in this paper, the CALIPSO VFM is assumed to be perfectly correct; the small temporal and spatial offsets between the radar and lidar are assumed to have only a little impact on the evaluation. In general, the temporal and spatial mismatch (as well as difference in instruments fields of view) likely causes some overestimate in both false and failed detection rates. Nonetheless, while the VFM is certainly not perfect, it is much more sensitive to cirrus clouds and is thus a reasonable benchmark for checking radar algorithms.

While it remains a topic for future research, we expect that this improvement in weak signal detection will enhance the retrieval of cirrus microphysical properties and the bilateral filter will hopefully be included in the next major revision of the CloudSat operational product (R06) after a systematic test by CloudSat science team. At present the CloudSat microphysical retrieval for cirrus (2C-ICE) (Deng et al., 2013, 2015) ignores weak detections (cloud mask values ≤ 10), because the false detection rate in R04 was judged to be too high (causing more error than benefit), with the results that a larger fraction of cirrus retrievals (than need to be) are based on extrapolations from lidar backscatter and a priori data (climatological constraints). Parameterization based on CloudSat reflectivity (Deng et al., 2010) is used to retrieve cirrus in the lidar-only region (where CALIPSO detects cloud while the CloudSat mask value has value ≤ 10). Our new method will allow weak detections to be incorporated into the 2C-ICE retrieval, instead of relying on

Acknowledgments

This work was supported by the National Science Foundation of China (41430425, 41875028, and 41922032), and the National Key R & D Program of China (2016YFC0401003). We also would like to thank the CloudSat and CALIPSO science teams for providing excellent and accessible data that made this study possible. The data were obtained from the NASA CloudSat data processing center (http://cloudsat.cira. colostate.edu/) and the NASA Langley Research Center Atmospheric Science Data Center (http://eosweb.larc.nasa. gov/). parameterization, and the benefit will outweigh the error for better understanding clouds impact on our climate system.

References

- Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J. L., et al. (2006). How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, 19(15), 3445–3482. https://doi.org/10.1175/jcli3819.1
- Ceppi, P., Hartmann, D. L., & Webb, M. J. (2016). Mechanisms of the negative shortwave cloud feedback in middle to high latitudes. *Journal of Climate*, 29(1), 139–157. https://doi.org/10.1175/jcli-d-15-0327.1
- Chernykh, I. V., Alduchov, O. A., & Eskridge, R. E. (2001). Trends in low and high cloud boundaries and errors in height determination of cloud boundaries. *Bulletin of the American Meteorological Society*, 82(9), 1941–1947. https://doi.org/10.1175/1520-0477(2001)082<1941: Tilahc>2.3.Co;2
- Clothiaux, E. E., Ackerman, T. P., Mace, G. G., Moran, K. P., Marchand, R. T., Miller, M. A., & Martner, B. E. (2000). Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *Journal of Applied Meteorology*, 39(5), 645–665. https://doi.org/10.1175/1520-0450(2000)039<0645:odocha>2.0.co;2
- Clothiaux, E. E., Miller, M. A., Albrecht, B. A., Ackerman, T. P., Verlinde, J., Babb, D. M., et al. (1995). An evaluation of a 94-GHz radar for remote sensing of cloud properties. *Journal of Atmospheric and Oceanic Technology*, 12(2), 201–229. https://doi.org/10.1175/1520-0426(1995)012<0201:aeoagr>2.0.co;2
- Deng, M., Mace, G. G., Wang, Z., & Berry, E. (2015). CloudSat 2C-ICE product update with a new Ze parameterization in lidar-only region. Journal of Geophysical Research: Atmospheres, 120, 12198–12208. https://doi.org/10.1002/2015jd023600
- Deng, M., Mace, G. G., Wang, Z., & Lawson, R. P. (2013). Evaluation of several A-Train ice cloud retrieval products with in situ measurements collected during the SPARTICUS campaign. *Journal of Applied Meteorology and Climatology*, 52(4), 1014–1030. https://doi. org/10.1175/jamc-d-12-054.1
- Deng, M., Mace, G. G., Wang, Z., & Okamoto, H. (2010). Tropical composition, cloud and climate coupling experiment validation for cirrus cloud profiling retrieval using CloudSat radar and CALIPSO lidar. *Journal of Geophysical Research*, 115, D00J15. https://doi.org/10.1029/ 2009jd013104
- Dessler, A. E., & Sherwood, S. C. (2004). Effect of convection on the summertime extratropical lower stratosphere. Journal of Geophysical Research, 109, D23301. https://doi.org/10.1029/2004jd005209
- Dessler, A. E., & Yang, P. (2003). The distribution of tropical thin cirrus clouds inferred fromTerraMODIS data. Journal of Climate, 16(8), 1241–1247. https://doi.org/10.1175/1520-0442(2003)16<1241:Tdottc>2.0.Co;2
- Fu, Q. (2013). Bottom up in the tropics. Nature Climate Change, 3(11), 957-958. https://doi.org/10.1038/nclimate2039
- Fu, Q., Baker, M., & Hartmann, D. L. (2002). Tropical cirrus and water vapor: An effective Earth infrared iris feedback? Atmospheric Chemistry and Physics, 2, 31–37. https://doi.org/10.5194/acp-2-31-2002
- Fu, Q., Carlin, B., & Mace, G. (2000). Cirrus horizontal inhomogeneity and OLR bias. Geophysical Research Letters, 27(20), 3341–3344. https://doi.org/10.1029/2000gl011944
- Fu, Q., Cribb, M. C., & Barker, H. W. (2000). Cloud geometry effects on atmospheric solar absorption. *Journal of the Atmospheric Sciences*, 57(8), 1156–1168. https://doi.org/10.1175/1520-0469(2000)057<1156:CGEOAS>2.0.CO;2
- Fu, Q., & Hollars, S. (2004). Testing mixed-phase cloud water vapor parameterizations with SHEBA/FIRE-ACE observations. Journal of the Atmospheric Sciences, 61(16), 2083–2091. https://doi.org/10.1175/1520-0469(2004)061<2083:Tmcwvp>2.0.Co;2
- Fu, Q., Hu, Y., & Yang, Q. (2007). Identifying the top of the tropical tropopause layer from vertical mass flux analysis and CALIPSO lidar cloud observations. *Geophysical Research Letters*, 34, L14813. https://doi.org/10.1029/2007gl030099
- Garrett, T. J. (2003). Small, highly reflective ice crystals in low-latitude cirrus. Geophysical Research Letters, 30(21), 2132. https://doi.org/ 10.1029/2003gl018153
- Gasparini, B., Meyer, A., Neubauer, D., Münch, S., & Lohmann, U. (2018). Cirrus Cloud Properties as seen by the CALIPSO satellite and ECHAM-HAM global climate model. *Journal of Climate*, *31*(5), 1983–2003. https://doi.org/10.1175/jcli-d-16-0608.1
- Ge, J., Wang, Z., Liu, Y., Su, J., Wang, C., & Dong, Z. (2019). Linkages between mid-latitude cirrus cloud properties and large-scale meteorology at the SACOL site. *Climate Dynamics*, 53(7-8), 5035–5046. https://doi.org/10.1007/s00382-019-04843-9
- Ge, J., Zheng, C., Xie, H., Xin, Y., Huang, J., & Fu, Q. (2018). Midlatitude cirrus clouds at the SACOL site: Macrophysical properties and large-scale atmospheric states. *Journal of Geophysical Research: Atmospheres*, 123, 2256–2271. https://doi.org/10.1002/ 2017jd027724
- Ge, J., Zhu, Z., Zheng, C., Xie, H., Zhou, T., Huang, J., & Fu, Q. (2017). An improved hydrometeor detection method for millimeterwavelength cloud radar. *Atmospheric Chemistry and Physics*, *17*(14), 9035–9047. https://doi.org/10.5194/acp-17-9035-2017
- Geer, A. J., Baordo, F., Bormann, N., Chambon, P., English, S. J., Kazumori, M., et al. (2017). The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quarterly Journal of the Royal Meteorological Society*, *143*(709), 3189–3206. https://doi.org/10.1002/qj.3172
- Getzewich, B. J., Vaughan, M. A., Hunt, W. H., Avery, M. A., Powell, K. A., Tackett, J. L., et al. (2018). CALIPSO lidar calibration at 532 nm: Version 4 daytime algorithm. *Atmospheric Measurement Techniques*, 11(11), 6309–6326. https://doi.org/10.5194/amt-11-6309-2018
- Gupta, A. K., Rajeev, K., Sijikumar, S., & Nair, A. K. M. (2018). Enhanced daytime occurrence of clouds in the tropical upper troposphere over land and ocean. Atmospheric Research, 201, 133–143. https://doi.org/10.1016/j.atmosres.2017.10.018
- Hagihara, Y., Okamoto, H., & Yoshida, R. (2010). Development of a combined CloudSat-CALIPSO cloud mask to show global cloud distribution. *Journal of Geophysical Research*, 115, D00H33. https://doi.org/10.1029/2009jd012344
- Hanisco, T. F., Moyer, E. J., Weinstock, E. M., St. Clair, J. M., Sayres, D. S., Smith, J. B., et al. (2007). Observations of deep convective influence on stratospheric water vapor and its isotopic composition. *Geophysical Research Letters*, 34, L04814. https://doi.org/10.1029/ 2006gl027899
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J. C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. Bulletin of the American Meteorological Society, 98(3), 589–602. https://doi.org/10.1175/bams-d-15-00135.1
- Huang, J., Ji, M., Higuchi, K., & Shabbar, A. (2006). Temporal structures of the North Atlantic Oscillation and its impact on the regional climate variability. Advances in Atmospheric Sciences, 23(1), 23–32. https://doi.org/10.1007/s00376-006-0003-8
- Huang, J., Ji, M., Xie, Y., Wang, S., He, Y., & Ran, J. (2015). Global semi-arid climate change over last 60 years. *Climate Dynamics*, 46(3-4), 1131–1150. https://doi.org/10.1007/s00382-015-2636-8

Huang, J., Wang, T., Wang, W., Li, Z., & Yan, H. (2014). Climate effects of dust aerosols over East Asian arid and semiarid regions. Journal of Geophysical Research: Atmospheres, 119, 11398–11416. https://doi.org/10.1002/2014jd021796

Huang, J., Yu, H., Dai, A., Wei, Y., & Kang, L. (2017). Drylands face potential threat under 2 °C global warming target. Nature Climate Change, 7(6), 417–422. https://doi.org/10.1038/nclimate3275

Huang, J., Zhang, W., Zuo, J., Bi, J., Shi, J., Wang, X., et al. (2008). An overview of the semi-arid climate and environment research observatory over the Loess Plateau. Advances in Atmospheric Sciences, 25(6), 906–921. https://doi.org/10.1007/s00376-008-0906-7

Kato, S., Sun-Mack, S., Miller, W. F., Rose, F. G., Chen, Y., Minnis, P., & Wielicki, B. A. (2010). Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles. *Journal of Geophysical Research*, 115, D00H28. https://doi.org/10.1029/2009jd012277

Kohma, M., & Sato, K. (2011). The effects of atmospheric waves on the amounts of polar stratospheric clouds. Atmospheric Chemistry and Physics, 11(22), 11535–11552. https://doi.org/10.5194/acp-11-11535-2011

Kollias, P., Clothiaux, E. E., Albrecht, B. A., Miller, M. A., Moran, K. P., & Johnson, K. L. (2005). The atmospheric radiation measurement program cloud profiling radars: An evaluation of signal processing and sampling strategies. *Journal of Atmospheric and Oceanic Technology*, 22(7), 930–948. https://doi.org/10.1175/jtech1749.1

Liu, Y., Chen, H., Wang, H., & Qiu, Y. (2018). The impact of the NAO on the delayed break-up date of lake ice over the southern Tibetan Plateau. Journal of Climate, 31(22), 9073–9086. https://doi.org/10.1175/jcli-d-18-0197.1

Mace, G. G., Marchand, R., Zhang, Q. Q., & Stephens, G. (2007). Global hydrometeor occurrence as observed by CloudSat: Initial observations from summer 2006. Geophysical Research Letters, 34, L09808. https://doi.org/10.1029/2006GL029017

Mace, G. G., Zhang, Q., Vaughan, M., Marchand, R., Stephens, G., Trepte, C., & Winker, D. (2009). A description of hydrometeor layer occurrence statistics derived from the first year of merged Cloudsat and CALIPSO data. *Journal of Geophysical Research*, 114, D00A26. https://doi.org/10.1029/2007jd009755

Marchand, R. (2012). Spatial correlation of hydrometeor occurrence, reflectivity, and rain rate from CloudSat. Journal of Geophysical Research, 117, D06202. https://doi.org/10.1029/2011jd016678

Marchand, R. (2013). Trends in ISCCP, MISR, and MODIS cloud-top-height and optical-depth histograms. Journal of Geophysical Research: Atmospheres, 118, 1941–1949. https://doi.org/10.1002/jgrd.50207

Marchand, R., Mace, G. G., Ackerman, T., & Stephens, G. (2008). Hydrometeor detection using Cloudsat—An Earth-orbiting 94-GHz cloud radar. Journal of Atmospheric and Oceanic Technology, 25(4), 519–533. https://doi.org/10.1175/2007jtecha1006.1

McCoy, D. T., Hartmann, D. L., & Grosvenor, D. P. (2014). Observed Southern Ocean cloud properties and shortwave reflection. Part II: Phase changes and low cloud feedback. *Journal of Climate*, 27(23), 8858–8868. https://doi.org/10.1175/jcli-d-14-00288.1

McCoy, D. T., Tan, I., Hartmann, D. L., Zelinka, M. D., & Storelvmo, T. (2016). On the relationships among cloud cover, mixed-phase partitioning, and planetary albedo in GCMs. *Journal of Advances in Modeling Earth Systems*, 8(2), 650–668. https://doi.org/10.1002/ 2015ms000589

Norris, J. R., Allen, R. J., Evan, A. T., Zelinka, M. D., O'Dell, C. W., & Klein, S. A. (2016). Evidence for climate change in the satellite cloud record. *Nature*, 536(7614), 72–75. https://doi.org/10.1038/nature18273

Pinsky, M., & Khain, A. (2019). Theoretical analysis of the entrainment-mixing process at cloud boundaries. Part II: Motion of cloud interface. Journal of the Atmospheric Sciences, 76(8), 2599–2616. https://doi.org/10.1175/jas-d-18-0314.1

Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. Bulletin of the American Meteorological Society, 84(11), 1547–1564. https://doi.org/10.1175/bams-84-11-1547

Rosenfeld, D., Zhu, Y., Wang, M., Zheng, Y., Goren, T., & Yu, S. (2019). Aerosol-driven droplet concentrations dominate coverage and water of oceanic low-level clouds. *Science*, *363*(6427). https://doi.org/10.1126/science.aav0566

Stephens, G., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., et al. (2018). CloudSat and CALIPSO within the A-Train: Ten years of actively observing the Earth system. Bulletin of the American Meteorological Society, 99(3), 569–581. https://doi.org/10.1175/bams-d-16-0324.1

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al., & the CloudSat Science Team (2002). The cloudsat mission and the A-Train—A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, 83(12), 1771–1790. https://doi.org/10.1175/bams-83-12-1771

Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, *340*(6136), 1053–1054. https://doi.org/10.1126/ science.1237554

Su, J., Huang, J., Fu, Q., Minnis, P., Ge, J., & Bi, J. (2008). Estimation of Asian dust aerosol effect on cloud radiation forcing using Fu-Liou radiative model and CERES measurements. *Atmospheric Chemistry and Physics*, 8(10), 2763–2771. https://doi.org/10.5194/acp-8-2763-2008

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016). Observational constraints on mixed-phase clouds imply higher climate sensitivity. Science, 352(6282), 224–227. https://doi.org/10.1126/science.aad5300

Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images, IEEE International Conference on Computer Vision, Bombay, India, doi:https://doi.org/10.1109/ICCV.1998.710815.

Vaughan, M., Garnier, A., Josset, D., Avery, M., Lee, K. P., Liu, Z., et al. (2019). CALIPSO lidar calibration at 1064 nm: Version 4 algorithm. Atmospheric Measurement Techniques, 12(1), 51–82. https://doi.org/10.5194/amt-12-51-2019

Vaughan, M. A., Powell, K. A., Winker, D. M., Hostetler, C. A., Kuehn, R. E., Hunt, W. H., et al. (2009). Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements. *Journal of Atmospheric and Oceanic Technology*, 26(10), 2034–2050. https://doi. org/10.1175/2009jtecha1228.1

Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., et al. (2004). Fully automated analysis of space-based lidar data: An overview of the CALIPSO retrieval algorithms and data products. *Laser Radar Techniques for Atmospheric Sensing*, 5575, 16. https://doi.org/10.1117/12.572024

Winker, D., Chepfer, H., Noel, V., & Cai, X. (2017). Observational constraints on cloud feedbacks: The role of active satellite sensors. Surv Geophys, 38(6), 1483–1508. https://doi.org/10.1007/s10712-017-9452-0

Worden, J., Noone, D., & Bowman, K. (2007). Importance of rain evaporation and continental convection in the tropical water cycle. *Nature*, 445(7127), 528–532. https://doi.org/10.1038/nature05508

 Young, S. A., Vaughan, M. A., Garnier, A., Tackett, J. L., Lambeth, J. D., & Powell, K. A. (2018). Extinction and optical depth retrievals for CALIPSO's version 4 data release. *Atmospheric Measurement Techniques*, 11(10), 5701–5727. https://doi.org/10.5194/amt-11-5701-2018
Zhang, C., Ge, L. Huang, L. & Fu, O. (2017). Cloud macrophysical properties from KAZR at the SACOL. *Chinese Science Bulletin*.

Zhu, Z., Zheng, C., Ge, J., Huang, J., & Fu, Q. (2017). Cloud macrophysical properties from KAZR at the SACOL. *Chinese Science Bulletin*, 62(8), 824–835. https://doi.org/10.1360/n972016-00857