Author's Response for

Shallow Cumuli Cover and Its Uncertainties from Ground-based Lidar-Radar Data and Sky Images
Erin A. Riley, Jessica M. Kleiss, Laura D. Riihimaki, Charles N. Long, Larry K. Berg and Evgueni Kassianov

This document contains a point-by-point response to the reviews, a list of all relevant changes made in the manuscript, and a marked-up manuscript version.

Table of Contents

Response to Reviewer 1

Thank you for the careful review of our paper and your thoughtful suggestions. We hope that you will find our responses and the corresponding revisions for the original manuscript satisfactory. Please find below your comments/suggestions (bold) and our responses with manuscript changes indicated in italic.

Positive points
o Figures are legible and have appropriate font size.
o Sentences are clear and properly structured.
o I think that the authors description of the many sensor and algorithm upgrades will add value to the literature since these are often difficult to track down and are poorly documented.
o A great novelty of this article includes comparing temporal estimates of cloud cover – from radar, ceilometer and MPL - to domain estimates of cloud cover – from TSI trying to assess the impact of field of view.

We are delighted that the Reviewer highlighted these positive points.

Major Comments:

o 1) Confusing uses of the terms "cloud fraction" and "cloud cover" and other derivatives - The term "cloud fraction" is most commonly used to represent the amount of clouds present at different levels in the atmosphere and is most often presented as a profile. Here, where the authors effectively refer to the projected area of clouds at the surface, I think the term "cloud cover" would be more appropriate. - Here, it would also seem appropriate to introduce two district cloud cover concepts: 1)"Domain cloud cover" which would be the number of cloudy pixels in each TSI image relative to the total numberof pixels in each TSI image. This definition would be closest to what a large-scale model would simulate/report and is what we would ideally like to measure (i.e., "truth") and 2) "Temporal cloud cover" which is the number of cloudy pixels in time series relative to the total number of observations collected over a defined time period (e.g., in ceilometer or radar time-series over 30-min or in a TSI pixel over 15-min).

We agree with the Reviewer that clarifications are needed for different estimates of cloud cover. Both "cloud fraction" and "fractional sky cover" are widely accepted terms for describing cloud cover (e.g., Qian et al., 2012 and references therein), while the term "domain" frequently defines a volume with specified grid spacing and number of vertical levels for model simulations (e.g., Berg et al., 2013). Therefore, we prefer to keep these two conventional terms in our paper. To address valuable comment from the Reviewer regarding their "temporal" and "areal" representations, we have modified the abstract slightly (lines 15-18), clarified introduction of these terms in the Introduction (Sect. 1, lines 46-51), and added a reminder of terminology to the Results and Discussion (Sect. 4.2, lines 270-272).

--Abstract (lines 15-18): *Enhanced observations at this site combine the advantages of the ceilometer, micropulse lidar (MPL) and cloud radar in merged data products. Data collected by these three instruments are used to calculate narrow-FOV cloud fraction (CF) as a temporal fraction of cloudy returns within a given period. Sky images provided by TSI are used to calculate the wide-FOV fractional sky cover (FSC) as a fraction of cloudy pixels within a given image.*

--Section 1, (lines 46-51):

There are two conventional measurement-based estimates of cloud cover: (1) cloud fraction (CF) obtained from zenith-pointing narrow-FOV observations and defined as the fraction of time when a cloud is detected within a specified period, and (2) fractional sky coverage (FSC) obtained from wide-FOV observations and defined as the fraction of cloudy pixels in a sky image. Note that FSC is similar to that estimated by a cloudy-sky observer (e.g., Henderson-Sellers and McGuffie, 1990; Kassianov et al., 2005; Long et al. 2006).

--Section 4.2 (lines 270-272): *Recall that the CF obtained from lidar-radar observations with narrow FOV represents a transect of a cloudy sky along wind direction, while FSC acquired from wide-FOV TSI data defines an area of cloudy sky. Both the CF and the FSC are widespread measurement-based estimates of cloud cover.*

o 2) The first goal stated by the authors is very similar to work undertaken by Kennedy et al. 2014. The authors first goal reads: "1) Have significant changes in the observations of ShCu cover occurred at the SGP site due to instrumental and algorithmic upgrades?". While I can appreciate that the current work tackles an extended dataset, to further add value, I wish it also went into more details about what are the exact algorithm changes or sensor upgrades responsible for the observed differences. For instance:

- Given what Kennedy et al. 2014 stated about the radar: "sudden change in CF occurs around the time the radar was upgraded, suggesting that this decrease is tied to hardware sensitivity or scanning strategy changes." The current study could quantify the additional amount of cloud detected solely by the radar sensitivity increase following the change from the MMCR to the KAZR.

- Given what Kennedy et al 2014 reported about the MPL: "Addition of the MPL increases the 14-year average CF by 9 %, mainly through an increase in optically thin high clouds year-round, and mid-level clouds during the summer months." In the current study, what can the authors say about the relative importance of such a sensitivity boost relative to the number of MPL false classification for ShCu. What exact changes were implemented following 2011 to improve the MPL cloud mask and can you recommend any further algorithm modification which could help mitigate the number of false ShCu detections?

The Reviewer is right that the algorithm changes and/or sensor upgrades are interesting topics, but their detailed discussion is beyond the scope of our current work.

However, we have expanded and clarified several important issues associated with these important topics.

- The radar-related changes and upgrades: Appendix A summarizes the instrumental and algorithmic changes of the radar. The primary limitation of the radar is the difficulty of observing optically thin clouds with low liquid water content (LWC), such as continental ShCu. The radar upgrades have primarily decreased the dwell time resulting in an increased number of pulses averaged. However, these upgrades very likely have not improved detection of the optically thin clouds. We have added a clarification regarding the sensitivity of the radar to cloud optical depth. (Section A.2, lines 463-465; lines 477-479).

- The MPL-related changes: The original cloud mask involved a signal-to-noise threshold based on cloud droplet scattering *(Campbell et al. 1998)*, whereas the updated cloud mask incorporated information on different scattering properties of aerosol and clouds (e.g., Wang and Sussen 2001). The corresponding clarifications have been added (Section 4.1; lines 254-258) and (Section A.3, Lines 481-483).

The specific changes are documented below.

--Section A.2 (lines 463-465):
Many continental cumuli are optically thin (liquid water content less than 50 g/m2) and have small droplets. As a result, the cloud radars (both MMCR and KAZR) are not able to "see" the majority of these clouds (e.g., Chandra et al. 2013; Lamer and Kollias, 2015).

--Section A.2 (lines 477-479)
Cloud radars (both the MMCR and KAZR) are known to miss detection of ShCu with small droplets compared to the ceilometer. For example, Chandra et al (2013) have found that the MMCR misses the majority of continental ShCu observed at the SGP site.

--Section 4.1 (lines 254-258): *Though a number of differences exist, the incorporation of MPL data (below 3 km) in the original cloud top height retrieval would increase number of detected cloud tops compared to those retrieved from the radar data alone for the initial period (2000-2010). Reliance only on the radar data for cloud top detection in the updated algorithm would result in fewer cloud top height detections and therefore a lower CF (see Sect. A.4 for more details).*

--Section A.3 (lines 481-483):
The original cloud mask involved a signal-to-noise threshold based on cloud droplet scattering (Campbell et al. 1998), whereas the updated cloud mask incorporated information on different scattering properties of aerosol and clouds (Wang and Sussen 2001). … It is expected that the updated cloud mask would improve cloud returns.

O 3) I think the second goal of the authors should take the forefront as it tackles something that remains poorly documented in literature. As it reads in the manuscript, the authors second goal is: "2) what is the impact of FOV configurations on hourly and sub-hourly observations of ShCu cover?". Taking it a step further I would be curious to know:
- Can narrow-field of view sensors be used to estimate a cloud cover representative of a domain?
- if so under what circumstances (e.g., strong horizontal wind, high cloud cover, ect.)?

In contrast to simulation-based studies, we use observations to quantify the impact of instrumental FOVs on cloud cover estimates. However, we should clarify that areas "seen" by ground-based instruments with narrow- and wide-FOV at the cloud base height are much smaller than areas (~ 30 km) considered in simulation-based studies (e.g., Oue et al. 2016) For example, TSI with 100-deg FOV "sees" a moderate area (~ 2.4 km diameter) at typical cloud based height (1km).

A clarification of differences between our approach and previous studies is included in section 4.2 (Lines 278-284):

In particular, the previous model studies (Astin et al., 2001; Berg and Stull, 2002) have demonstrated that the cloud cover obtained from the transect measurements mimics the area-averaged cloud cover for non-organized (e.g. random) cloud fields well if the sample size is relatively large (or numerous individual clouds are sampled). Recently Oue et al. (2016) have showed that 10 or more ceilometers equally spaced across a 25 km width in the cross-wind direction are required to estimate the simulated cloud cover in the small (30 km) domain. Certainly, the number of ceilometers, their locations and averaging time required for an accurate estimation of the cloud cover depend on the spatial arrangement of clouds and wind speed.

With this clarification in mind, we respond to the Reviewer's questions by explicitly stating that we interpret the term "domain" in the question to be an area "seen" by TSI with the 100-deg FOV at the cloud base height.

1) **Can narrow-field of view sensors be used to estimate a cloud cover representative of a domain?**
Our results (Section 4.1) suggest that a narrow-FOV data from ceilometer as compared with those from lidar and radar can be used to estimate wide-FOV FSC reasonably well, on average, in terms of bias (Table 3), slope and intercept of linear regression (Table B2), and correlation coefficient (Table B3). For example, the correlation coefficient values for sub-hourly and hourly time scales are 0.83 and 0.89, respectively (Table B3). It should be emphasized that about 30-34% of the corresponding CF-FSC comparisons still have difference greater than 0.1 (section 4.2, lines 301-303). The level of agreement between the narrow-FOV CF and wide-FOV FSC depends on several factors described below (the second question).

2) **if so under what circumstances (e.g., strong horizontal wind, high cloud cover, ect.)?**
Several factors, such as wind speed and spatial arrangement of clouds, determine whether or not the narrow-FOV CF and wide-FOV FSC are comparable, on average. The spatial arrangement of clouds (e.g., organized versus non-organized spatial distribution) defines representation of a 1D transect along a wind direction of a 2D cloud field for a given area of interest, while the wind speed determines the number of sampled clouds (or sample size). The "quick-look" tool (Section 4.3) can be used for characterization of the spatial arrangement of clouds while available information on the wind speed can be used for estimation of the sample size, and therefore for assessment of the wind speed impact on the CF-FSC comparison.

To address these points in the manuscript, two new paragraphs have been added. The first new paragraph discusses possible applications of the "quick-look" tool to assess the impact of cloud field organization on agreement between narrow- and wide-FOV observations (Section 4.3, lines 365-375). The second new paragraph describes the impact of wind speed on the agreement of narrow- and wide-FOV observations (Section 4.2, lines 306-3017).   A new Figure 6 has been added to illustrate the concepts discussed in the second new paragraph.

The new paragraph discussing the impact of cloud field organization on the agreement of narrow- and wide-FOV observations (Section 4.3, lines 365-375):

"There are two main expected applications of the introduced "quick-look" tool. The first potential application is a classification of spatial organization of cloud fields using, for example, cross-wind cloud field variability (e.g. peaks and valleys in Fig. 7b) and within-lane variance of cloud amount (e.g. vertical bars in Fig. 7b). Numerous images generated by the "quick-look" tool (e.g., Figure 8b) for the extended period (2000-2017) can be considered as a valuable training dataset for machine learning with focus on automated detection of desired features of the cloud fields (e.g., "cloud streets") and unwanted contaminations of TSI images (e.g., Figure 9). Second potential application is a visual inspection of the generated images for a given period of interest (e.g., a short-term field campaign) to check for the impact of instrumental detection differences and cloud field organization on the observed cloud amount. Visual inspection may be feasible given a limited number (about 40) of ShCu events annually during the warm season. For example, a spread of the lane CFs (gray region in Fig. 8c) gives an idea about the cross-wind cloud field variability within a given FOV, and thus aids in understanding the difference between cloud amounts obtained from the narrow- and wide-FOV observations."

The new paragraph discussing the impact of wind speed and the sample size on the agreement of narrow- and wide-FOV observations (Section 4.2, lines 306-317).

"The effective spatial area sampled by either narrow or wide FOV instruments is a function of both sampling duration and wind speed. High wind speed in comparison with low wind speed (1) increases sample size for a given period and (2) tends to organize horizontal arrangement of clouds (e.g., Weckworth et al. 1999, Atkinson and Zhang 1996). These two factors associated with sample size and spatial arrangement of clouds should be considered when differences between cloud cover obtained from narrow- and wide-FOV observations as function of wind speed are considered (Fig. 6). In particular, Figure 6 illustrates that both CF-FSC and "CF-like"-FSC differences are reduced noticeably as the wind speed increases from 1 m/s to 3 m/s, and continue to reduce slightly as the wind speed grows up to 11 m/s. The CF-FSC and "CF-like"-FSC differences obtained at a higher wind speed (above 11 m/s) should be considered with caution due to limited number of the corresponding cases with high wind speed (e.g., fewer than 100 cases for 60-min time average). The increased sampling area associated with increased wind speed does not necessarily result in an improved agreement between the narrow- and wide-FOV observations for both hourly and sub-hourly observations due to the impact of wind speed on cloud organization."

o 4) Did the authors consider the effect of horizontal wind speed in the comparison of domain cloud cover and temporal cloud cover? I would expect that higher horizontal wind speeds advect clouds more rapidly such that, under higher wind speeds, shorter time periods of narrow-field of view measurements would be required to capture the CF observed by the wider field of view TSI.

To address your important questions, a new paragraph regarding the wind speed and the sample size (Sect. 4.2, lines 306-317) together with a new plot (Figure 6) have been added. This paragraph has been included in our reply above.

o 5) The spatial analysis of TSI cloud mask is very interesting but I had to read the article twice to understand where is fits in with the other cloud fraction definitions. What would make this clearer for me would be to state that the TSI "lane by lane" cloud fraction estimates are effectively temporal cloud fraction estimate (and not FOV or domain cloud fraction estimates) and that each lane can be interpreted as a time series observed by a narrow field of view sensor. I would also perhaps bring information about this lane by lane methodology and information about the radar wind profiler to section 4.3 where you describe your tool.

Thank you for the valuable suggestions. We have moved the description of the lane-by-lane methodology and the radar wind profiler (previous section 3.5) to section 4.3, added clarification to section 4.3, and changed the labels on Figure 7 from FSC to CF.

The clarification (Section 4.3; lines 328-330)
Each pixel in the averaged image can be interpreted as a 15-min CF measurement from a narrow-FOV sensor. The variability of CF in the cross-wind direction can indicate the possible influence of cloud field organization on cloud cover estimates provided by narrow-FOV observations.

Also, the vertical axis label for Figure 7b has been changed from FSC to CF for consistency with the above changes.

o 6) Clarifications regarding the impact of insects on ShCu top detections - Multiple studies have reported that the presence of insect hinders the radars ability to accurately detect cloud top. I think more information is needed here about how insect contamination is handled in ARSCL both pre and post 2011 where the authors hint that the MPL stopped being applied in the boundary layer. This could offer an alternative explanation to the changes in radar-lidar CF post 2011 where

the increase in radar detected cloud top could be due both to the KAZR being more sensitive than the MMCR and to the KAZR insect filtering having changed such that more insect returns are misclassified as cloud tops. If both effects are in play, then I would like to see their relative importance quantified.

We agree that insects can contaminate accurate determination of cloud boundaries by radar. However, accurate cloud top height retrievals by radar is not required in our analysis because a simple threshold is used to determine the presence of ShCu. Moreover, cloud base height estimation involves lidar observations (both ceilometer and MPL) which are not impacted by the presence of insects,

Text has been added to clarify this point: (Section 3.1, lines 141-144)
Insect contamination may contribute to significant uncertainty of the radar-based retrievals of cloud boundaries. Therefore, our analysis employs a semi-quantitative threshold approach when using the cloud top heights. This approach is less sensitive to the insect contamination.

o 7) The idea of compensating bias introduced on Page 8 "introduction of compensating errors using the cloud top height criteria in the updated merged lidar-radar product." needs clarification.
- If I understand correctly the hypothesis is that in the 2000-2010 period the MPL was overly sensitive to aerosols leading to a CF overestimation while the MMCR was underly sensitive to cloud leading to a CF underestimation hence the compensating bias

Thank you for pointing this out. For the later sub-period (2010-2017), the merged cloud radar-lidar product relies on the "shallow" (< 3 km) radar data instead of the combined MPL-radar observations for determining the cloud top. The radar misses a substantial fraction (about 30%) of ShCu, therefore the cloud top height (below 3 km) is very likely to be missed. Meanwhile, the merged lidars (ceilometer and MPL) data are used to detect the cloud base height and exhibit higher CF than that from the ceilometer alone. A compensating error could potentially arise from the over-detection of clouds in the merged lidar data with the under-detection of cloud from the radar observations. The RMSD for the CF including cloud top heights for the later sub-period (2010-2017) is higher than those for the CF obtained from ceilometer alone (even for near-zero bias). This indicates that the instrument detection differences in the merged lidar-radar product contribute mostly to the CF uncertainty.

o There are gaps in the literature review
- Beyond the few studies cited on Page 2 line 30, others before have attempted to assess the representativeness/reconcile multiple cloud fraction measurements (e.g., Dr. Mariko Oue work with scanning cloud radar or Dr. Steve Schwartz work with photography or Dr. Wei Wu work with ISCC).
- Some references are missing for the bibliography (e.g., Tatarevic and Kollias, 2015)
- Some references are to meeting abstracts rather than to the published journal articles (e.g., Lamer et al. 2017 abstract work has since been published in GMD)
- Some references are miscited (e.g. Chandra et al., 2013, Zhang and Klein, 2010, 2013 and Lamer and Kollias, 2015 do not show any model-observation comparison).

We thank the Reviewer for the careful editing and helpful suggestions. We have made the following adjustments to the manuscript.

Introduction (Lines 69-71)
- We have added Oue et al. (2016) reference along with two citations in the text.

Moreover, sampling of LES-generated cloud fields by a virtual instrument can be a helpful way to reconcile debated differences between the retrieved and predicted values of cloud cover (Oue et al., 2016).

Section 4.2 (lines 278-284)

In particular, the previous model studies (Astin et al., 2001; Berg and Stull, 2002) have demonstrated that the cloud cover obtained from the transect measurements mimics the area-averaged cloud cover for non-organized (e.g. random) cloud fields well if the sample size is relatively large (or numerous individual clouds are sampled). Recently Oue et al. (2016) have showed that 10 or more ceilometers equally spaced across a 25 km width in the cross-wind direction are required to estimate the simulated cloud cover in the small (30 km) domain. Certainly, the number of ceilometers, their locations and averaging time required for an accurate estimation of the cloud cover depend on the spatial arrangement of clouds and wind speed.

- We have added Tatarevic and Kollias (2015) to the bibliography,

Conclusions (line 408)
We have replaced the Lamer et al. (2017) reference with Lamer et al. (2018) GMD article, and adjusted the bibliography accordingly.

Introduction (lines 52-53)
- We have removed the citations to Chandra et al., 2013, Zhang and Klein, 2010, 2013 and Lamer and Kollias, 2015 for model-observation comparison, and inserted Zhang et al., 2017 and Endo et al., 2019.

o 9) Although I understand that there are many ways to organize a methods section and that we all prefer to receive information in different sequences, for me, the layout of the data and methods section was confusing.
- I would rather the authors merge the data and methods sections which go hand in hand and preface such a section stating what quantities they are after 1) Identification of cumulus cases, which requires cloud top height and cloud cover estimates 2) Temporal cloud fraction, which will be obtained from ceilometer, ceilometer+MPL, and ceilometer+MPL+radar cloud base height time series 3) Domain cloud fraction, which will be obtained from TSI using different angular domains 4) Providing context using horizontal wind direction and TSI lane-by-lane decomposition

We thank the Reviewer for this helpful comment. Indeed, during development of this manuscript we have attempted many possible organizational approaches to communicating the data and methods used to produce this dataset. One challenge with the approach that the Reviewer suggests is the interdependent nature of these data. For example, the natural first quantity of the paper, as you list above, is the identification of ShCu events. However, this method requires TSI, ceilometer, and lidar/radar data (items #2 and 3 above). Thus we settled on the current organizational structure of the data and methods sections.

While the current layout may be somewhat cumbersome to read as prose, we believe this structure will be most conducive for readers to access the data and perform their own analyses in extension of this work.

o 10) Clarification are needed when it comes to ShCu case selection
Thank you for your careful attention to detail in the review of this paper. We address each point independently below.

- Page 4, Line 3: Which observations are used in the ARM Shallow Cumulus data product to identify ShCu cases?

The reference to the new Shallow Cumulus data product has been added for clarification of the observations used in our study. (Section 2, lines 108-110)
The newly released ARM Shallow Cumulus data product (Data Reference: Shi et al., 2000) identifies times of ShCu from lidar / radar cloud boundary heights and includes FSC from TSI observations (Lim et al., 2018).

- Table 1 somewhat helps in understanding which portion of the cloud field is of interest. However, I do not see a reference to Table 1 in Sect. 3.1. where I would expect it.

We respectfully note that Table 1 is referenced in the third sentence in Sect. 3.1. To facilitate readability, we have added another reference to it later in this paragraph. (Line 146)

- Figure 2 shows cloud fractions ranging from 0 to 1, do you consider overcast conditions to be ShCu?

The FSC obtained from 100-deg FOV TSI data was used to determine clear-sky or overcast conditions. Only partly-cloudy conditions (0.05<FSC<0.95) were considered in our study. This is stated in Section 3.4 (line 185), and the captions of Figure 2, Table B1, Table 2, Table 3, and Figure B1.

- Page 4, Line 8: Why did you chose to: "additionally extending the start and end-times by 1 hour each.". This could include periods presenting deeper clouds or cloud aloft.

The text describing Shallow Cumulus event detection was revised to improve clarity. (Section 2, lines 134-138)
The extension allows for more accurate determination of the start- and end- times of the event on the finer time scale of the TSI FSC (15 min). Quality control procedures (Sect. 3.4) are used to censor multi-layer clouds and clear sky conditions on the 15-min and hourly observations of cloud cover.

- Table 1. Why is a minimum cloud base height threshold applied to most CF estimates? If insects and clutter have been properly filtered from the radar data, I cannot see why this could be necessary.

This helps additionally screen out smoke from agricultural burning of biomass, which happens several days each warm season. It would eliminate fog too, which has never been an issue.

- Table 1 What is the value of estimating a CFtot if the cases discussed are purely single layer ShCu? Shouldn't all cloud observed have tops and bases below 3 km?

The ShCu data product includes ShCu cases with overlying cirrus. The TSI cannot differentiate low clouds from cirrus, which can be sporadic with short duration. The selected cases were required to have at least 2 hours without cirrus.

o 11) Pertinent information missing in the various definitions of cloud fraction

- Page 3 Line 15: "Appendix A contains pertinent information for their application". I believe all pertinent text should be in the main text and the appendix should be reserved for details. I am especially wanting to know how the ARSCL reports cloud top height when both the radar and MPL are used (in the 2000-2010 period) since this is very relevant to the sensitivity versus insect detection compensating effect.
Appendix A provides detailed information about data streams and data pre-processing. This is valuable background knowledge for researchers.

The issue of cloud top height observations in the earlier (2000-2010) and later (2011-2017) periods is indeed interesting and relevant and discussed in full in the Conclusions. In particular, the corresponding discussion under the first research question "*1) Have significant changes in the observations of ShCu cover occurred at the SGP site due to instrumental and algorithmic upgrades?*" is included. (lines 383-394)

Minor Comments:

1) In the future, when submitting articles for review, please number the lines continuously rather than restarting the numbering process on each page.

We agree that continuous line numbering is preferable - The revised manuscript uses continuous line numbering.

2) The abstract is very "number focused" and could benefit from including more "conclusions", for instance, the abstract does not provide information about which sensor upgrade had the largest impact on the cloud cover estimates or about the fact that cloud field organization (e.g., cloud streets) parallel to the horizontal wind direction can create large biases between narrow and wide field of view cloud cover estimates.

We have addressed the first point by adding a statement to the abstract regarding the algorithmic upgrade that had the largest impact on the cloud cover estimate (lines 26-27), and we have addressed the second point regarding the potential bias in observations due to cloud field organization (lines 32-33). The specific edited text is below.

Abstract (lines 26-27): *The strong period dependence of CF obtained from the combined ceilometer-MPL-radar data is likely due to increased reliance on the radar for cloud top height returns.*

Abstract (lines 32-33): *The influence of cloud field organization, such cloud streets parallel to the wind direction, on narrow- and wide-FOV cloud cover estimates can be visually assessed.*

3) Page 1 Line 22: What is meant by "mean cloud cover"

The term "mean cloud cover" has been replaced with the term "multi-year mean cloud cover" (Abstract, line 23)

4) Page 2 Line 2: Missing some "the"

Thank you - this has been changed. (Section 1, line 40)

5) Page 2 in a few places. I would suggest using the word "variability" instead of the word "changes" when referring to the cloud field

Thank you - this has been changed in the Introduction, lines 43 and 45.

6) The acronym for the Ka-band ARM Zenith Radar should be entirely capitalized (i.e., "KAZR" not "KaZR")

Thank you - all instances have been changed.

7) Page 2 Line 18: What do you mean by "consistent"?

The word "consistent" has been removed to avoid confusion. (Page 2, line 56)

8) Page 2, Line 18: Zhang and Klein 2013 used 13 years of ARSCL data, the sentence as you have it constructed is somewhat misleading as it suggests that they used 20 years of data. It would be more appropriate to state that previous studies have used ARSCL (cite here) and this data record is now reaching 20 years in length.

The text has been changed to indicate that the ARSCL data have been available since Nov. 1996 to avoid confusion. The revised sentence reads:

(Introduction, lines 56-57) *A merged lidar-radar data product is available from Nov. 1996 to the present at the SGP site and has served as a basis for developing ShCu climatology at the SGP site …*

9) Page 2 Line 23: Following my suggestion above "Areal cloud cover" would become domain cloud cover.

We've elected to maintain our conventions. Please see the discussion in the first question of this review. Thank you for your suggestion.

10) Page 5 Line 1: Given that the radar can be affected by insects, I would avoid using the word "reliably".

Thank you for this suggestion. The wording of these sentences was changed in response to major comment number 6.

11) Page 4, Line 4: Shouldn't "ShCu cloud coverage" read "ShCu cloud periods"?

Yes, thank you. This has been changed to *times of ShCu*. (Section 2, line 109)

12) Figure 2 Panels a and b are missing a legend

Note that this is now Figure 1. The caption has been updated. Note that the 1D histograms were changed to line plots in response to a separate Reviewer comment.

Figure 1 (caption): *merged ceilometer-MPL (blue) and merged lidar-radar (red)*

13) Figure 2 c and d and all figures of this style are missing colorbar labels

Thank you. The captions for figures 1, has been updated to include specific mention of the colorbar.

Figure 1 (caption): *joint histograms (counts)*
Figure 2 (caption): *Color scale represents counts in increments of 10.*

14) Page 4, line 31: "This method has the advantages of low missing data due to multiple instruments used and limits the vertical extent of clouds." Please rephrase. Using a cloud top detection criteria does not "limits the vertical extent of clouds".

Thank you for this suggested clarification. The relevant passage has been changed:
(Section 3.1, lines 148-149) *This method has the advantages of low missing data due to use of multiple instruments and incorporates information about cloud top height consistent with the definition of shallow convection.*

15) Page 9 line 4: Add "altitude" after "1.5 km"
We respectfully note that the sentence already reads "*1.5 km cloud base*", which employs the correct terminology. (Section 4.2, line 272)

16) Page 2, Line 29 "In addition, long-term averages of CF obtained from merged ceilometer-MPL data tend to be larger than FSC (Boers et al., 2010; Qian et al., 2012; Wu et al., 2014; Kennedy et al., 2014), indicating a potential consequence of instrument-dependent cloud detection differences." Could this difference not also be attributable to FOV differences? If so, please add this caveat.

Long-term averages are monthly and yearly long, and have sufficient sample size that reduce the impacts of limited FOV. A potential long-term impact of FOV configuration on cloud cover estimates could be associated with cloud sides visible at moderate zenith (viewing) angles. However, one would then expect that wide-FOV FSC would exceed narrow-FOV CF. The opposite is demonstrated by our analysis. This suggests that the obtained difference between narrow-FOV CF and wide-FOV FSC is mainly due to instrument-dependent cloud detection differences. (revised manuscript Page 3, line 67)

17) Page 1 Line 19 "We demonstrate that CF obtained from ceilometer data alone and FSC obtained from sky images provide the most similar and consistent cloud cover estimates: bias and root-mean-square difference (RMSD) are within 0.04 and 0.12, respectively."
According to your analysis of the impact of the "Field Of View (FOV)" performed by comparing the two TSI FOV, the averaging period of the narrow field of view sensor can affect the RMSD between cloud cover obtained by the narrow and wide FOV. Am I correct to understand that this result also applies to the comparison between the ceilometer or any other "beam" observation (e.g., radar, MPL) and the TSI? If so, I think the statement above should include information about the averaging time period used for the ceilometer in this comparison.

Thank you for the careful attention to detail. Yes, the "CF-like" aims to characterize the effect of FOV on cloud amount observations, since it is derived from the same instrument (the TSI) as the wide-FOV FSC. Indeed, the characterization of the RMSD between "CF-like" and FSC is 0.1 for 60-minute averages, and 0.15 for fine temporal scale (30-minute CF-like, 15-minute FSC). We accept your suggestion, and edit the text to read as follows:

(Abstract, lines 21-23) *We demonstrate that CF obtained from ceilometer data alone and FSC obtained from sky images provide the most similar and consistent cloud cover estimates: hourly bias and root-mean-square difference (RMSD) are within 0.04 and 0.12, respectively.*

18) Page 8 line 24: "Though a number of differences exist, the incorporation of MPL data below 3 km in the initial cloud top height retrieval algorithm between 2000-2010 but not the updated algorithm likely has a large impact (see Sect. A.4 for more details)."
I think it would help the reader to explain if an overestimation or an underestimation is expected and why?

This sentence has been changed to read:

(Section 4.1, lines 253-257) *Though a number of differences exist, the incorporation of MPL data below 3 km in the original cloud top height retrieval would increase the number of detected cloud tops compared to those retrieved from the radar data alone for the initial period (2000-2010). Reliance only on the radar data for cloud top detection in the updated algorithm would result in fewer cloud top height detections and therefore a lower CF (see Sect. A.4 for more details).*

Response to Reviewer 2

Thank you for the careful review of our paper and your thoughtful suggestions. We hope that you will find our responses and the corresponding revisions for the original manuscript satisfactory. Please find below your comments/suggestions (bold) and our responses with manuscript changes indicated in italic.

Major suggestions

1. End of Page 5 and the beginning of page 6 you have mentioned the differences in the field of view of the instruments. This is good. However, the effective field of view of the radar and lidar is essentially FOV plus dwell-time times the wind speed (advection). I suggest you add a few sentences to describe this effect. When you are generating the 15-min or 30-min statistics, then the differences from the two methods will largely governed by the wind speed.

We believe your comment is related to narrow-FOV FSC ("CF-like") introduced in Section 3.3. We've added the following three sentences (section 3.3, lines 174-78):

The narrow-FOV (3x3 pixel region) "CF-like" observation has 4.1° (original) and 2.3°(improved) angular resolution. For comparison, the narrow-FOV ARSCL cloud products, including the lidar-radar CF, have much finer angular resolution (about 0.2°). The corresponding spatial resolution of the lidar-radar CF can be estimated by multiplying wind speed at cloud base height by lidar-radar dwell time (about 10 sec). For example, its spatial resolution is about 100 m for 10 m/sec wind speed.

2. In a similar vein, it will be good if you can make a figure of the CF from radar-lidar and FSC as a function of the wind speed. You already have the wind speed from the radar wind profiler, and the other two are shown in Figure 3. A figure like this will tell us how much high or low wind contributes to the differences in the two methods. Thanks.

As you suggested, we've added a new Figure 6 in the revised manuscript to illustrate relative contributions of low and high wind speed to the difference between cloud cover obtained from the narrow- and wide-FOV observations. Also, we've added the following paragraph (section 4.2, lines 305-316).

The effective spatial area sampled by either narrow or wide FOV instruments is a function of both sampling duration and wind speed. High wind speed in comparison with low wind speed (1) increases sample size for a given period and (2) tends to organize horizontal arrangement of clouds (e.g., Weckworth et al. 1999, Atkinson and Zhang 1996). These two factors associated with sample size and spatial arrangement of clouds should be considered when differences between cloud cover obtained from narrow- and wide-FOV observations as function of wind speed are considered (Fig. 6). In particular, Figure 6 illustrates that both CF-FSC and "CF-like"-FSC differences are reduced noticeably as the wind speed increases from 1 m/s to 3 m/s, and

continue to reduce slightly as the wind speed grows up to 11 m/s. The CF-FSC and "CF-like"-FSC differences obtained at a higher wind speed (above 11 m/s) should be considered with caution due to limited number of the corresponding cases with high wind speed (e.g., fewer than 100 cases for 60-min time average). The increased sampling area associated with increased wind speed does not necessarily result in an improved agreement between the narrow- and wide-FOV observations for both hourly and sub-hourly observations due to the impact of wind speed on cloud organization.

3. Lastly, the lane approach is very novel. If incorporated properly, it will tell us how the clouds are organized within a cloud field and move in relation to each other. Such an analysis is outside the scope of this article. However, I suggest you add a paragraph on the potential scientific usage of the statistics derived by this approach. Thanks.

Thank you for your valuable suggestion. We've added a new paragraph (section 4.3, lines 464-374) in the revised manuscript to highlight expected scientific applications of our "quick-look" tool.

There are two main expected applications of the introduced "quick-look" tool. The first potential application is a classification of spatial organization of cloud fields using, for example, cross-wind cloud field variability (e.g. peaks and valleys in Fig. 7b) and within-lane variance of cloud amount (e.g. vertical bars in Fig. 7b). Numerous images generated by the "quick-look" tool (e.g., Figure 8b) for the extended period (2000-2017) can be considered as a valuable training dataset for machine learning with focus on automated detection of desired features of the cloud fields (e.g., "cloud streets") and unwanted contaminations of TSI images (e.g., Figure 9). Second potential application is a visual inspection of the generated images for a given period of interest (e.g., a short-term field campaign) to check for the impact of instrumental detection differences and cloud field organization on the observed cloud amount. Visual inspection may be feasible given a limited number (about 40) of ShCu events annually during the warm season. For example, a spread of the lane CFs (gray region in Fig. 8c) gives an idea about the cross-wind cloud field variability within a given FOV, and thus aids in understanding the difference between cloud amounts obtained from the narrow- and wide-FOV observations.

Minor Suggestions

Page 2 line 5: I think you mean "partitioning" and not "proportioning". Thanks.
- Corrected, thank you.

Page 2 line 17: Might be better to refer to the ARM monograph in the AMS
- Citation to the ARM Program website (https://www.arm.gov/) was added.

Page 2, line 22: Remove "for example, a recent report" and just say "Zhang et al. (2017) suggested .."
- Corrected, thank you.

Page 3, Line 2 and Line 16: Also, at other locations. Please either use radar-lidar or lidar-radar for consistency.
- Changed to lidar-radar throughout.

Page 3, line 30: "height is used here"
- Corrected, thank you.

Page 7 line 5: Figure B1 not 1B
- Figure 1a and 1b are now Figure 7a and 7b in the revised manuscript.
- All figure names have been changed as a result of inserting new Figure 6.

Figure 1 caption: I suggest using "vertical bars" rather than "error bars" to avoid confusion.
- Suggestion accepted, thank you.

Figure 2: As the brown and blue bars are on top of each other, maybe it will be better to show them as line plots. It will be good to know how far apart they are for low CF values. Also, I see light brown bars in (a) and (b), and a dashed red line in (d). Both of these have not been explained in the caption.
- Figure 2 is now Figure 1 in the revised manuscript.
- Thanks for the suggestion. This figure has been updated to use a line plot instead of bars. The figure caption has been edited.

Figure 3-6: It will be good if you can bin the shades in bins of 10% and use only one color for each bin. It is difficult to identify the actual values in the current versions. There are also dashed magenta lines in some of the panels.
- These are count histograms. We've changed the colorbar to be in increments of 10 counts, which seems to improve the legibility.
- We have clarified that the units are "counts" in the legends of figures 1 and 2.

List of all relevant changes made in the manuscript

- The Abstract has been clarified to emphasize the qualitative (as well as quantitative) findings from this study.
- The text has been clarified to differentiate between the wide field of view (FOV) observations from the total sky imager (TSI) and the narrow-FOV observations from active remote sensing (lidar / radar).
- The description of the process to determine times of shallow cumulus was clarified.
- The possible contamination by insects on radar data in the context of this study was included
- An expanded discussion is included on the impact of insects on the radar retrievals of cloud top height, and benefits of the merged lidar/radar product for cloud top height estimates.
- The influence of wind at cloud base height on the effective field of view of narrow remote sensors was included.
- The methodology to perform the novel spatial analysis of wide-FOV sky images is moved from the methods section to the Results and Discussion section, immediately before the implications of this analysis are presented.
- A new discussion and new figure 6 are added to address the influence of wind at cloud base on the effective field of view of both narrow- and wide-FOV observations.
- A discussion is added to indicate possible applications and future uses of the spatial analysis "quick look" tool.
- Additional technical details on the algorithm changes and instrumental upgrades for the radar and lidar are included in the appendices.

- Some stylistic changes to the text have been applied throughout, including consistent use of the acronym KAZR, and "lidar/radar" to indicate active remote sensing.
- 8 additional references were added describing previous work that is relevant to this study.

- Figures have been re-ordered to present the instrumental and time-averaging differences prior to introduction of the spatial analysis "quick look" tool.
- Figure color bars and line types have been adjusted for increased clarity.

Shallow Cumuli Cover and Its Uncertainties from Ground-based Lidar-Radar Data and Sky Images

Erin A. Riley[1], Jessica M. Kleiss[1,*], Laura D. Riihimaki[2,3], Charles N. Long[2,3], Larry K. Berg[4] and Evgueni Kassianov[4,*]

[1] Environmental Studies, Lewis and Clark College, Portland, OR 97219, USA
[2] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA
[3] Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, CO, 80305, USA
[4] Pacific Northwest National Laboratory, Richland, WA 99354, USA

Correspondence to: Jessica M. Kleiss, jkleiss@lclark.edu, Evgueni Kassianov, Evgueni.Kassianov@pnnl.gov

Abstract

Cloud cover estimates of single-layer shallow cumuli obtained from narrow field-of-view (FOV) lidar-radar and wide-FOV Total Sky Imager (TSI) data are compared over an extended period (2000-2017 summers) at the established United States Atmospheric Radiation Measurement mid-continental Southern Great Plains site. We quantify the impacts of two factors on hourly and sub-hourly cloud cover estimates: 1) instrument-dependent cloud detection and data merging criteria, and 2) FOV configuration. ~~Popular E~~enhanced observations at this site combine the advantages of the ceilometer, micropulse lidar (MPL) and cloud radar in merged data products. Data collected by these three instruments are used to calculate narrow-FOV cloud fraction (CF) as a temporal fraction of cloudy returns within a given period. Sky images provided by TSI are used to calculate the wide-FOV fractional sky cover (FSC) as a fraction of cloudy pixels within a given image. ~~, and are used to calculate temporal cloud fractions (CF). Sky images provide the spatial fractional sky cover (FSC) within the visible sky dome.~~ To assess the impact of the first factor on CF obtained from the merged data products, we consider two additional sub-periods (2000-2010 and 2011-2017 summers) that mark significant instrumentation and algorithmic advances in the cloud detection and data merging. We demonstrate that CF obtained from ceilometer data alone and FSC obtained from sky images provide the most similar and consistent cloud cover estimates: hourly bias and root-mean-square difference (RMSD) are within 0.04 and 0.12, respectively. Whereas CF from merged MPL-ceilometer data provides the largest estimates of the ~~mean~~ multi-year mean cloud cover: about 0.12 (35%) and 0.08 (24%) greater than FSC for the first and second sub-periods, respectively. CF from merged ceilometer-MPL-radar data has the strongest sub-period dependence with a bias of 0.08 (24%) compared to FSC for the first sub-period and shows no bias for the second sub-period. The strong period dependence of CF obtained from the combined ceilometer-MPL-radar data is likely due to increased reliance on the radar for cloud top height returns. To quantify the FOV impact, a narrow-FOV FSC is derived from the TSI images. We demonstrate that FOV configuration does not modify the bias, but impacts the RMSD (0.1 hourly, 0.15 sub-hourly). In particular, the FOV impact is significant for sub-hourly observations, where 41% of narrow- and wide-FOV FSC differ by more than 0.1. A new "quick-

look" tool is introduced to visualize impacts of these two factors through integration of CF and FSC data with novel TSI-based images of the spatial variability in cloud cover. The influence of cloud field organization, such cloud streets parallel to the wind direction, on narrow- and wide- FOV cloud cover estimates can be visually assessed.

35

1 Introduction

Shallow cumuli (ShCu) have a number of important roles in the Earth's climate system due to their complex interactions
40 with radiation, the atmosphere and the surface (e.g., Arakawa, 2004; Vial et al., 2017; Park and Kwon, 2018). For example, the amount of surface moisture can influence the cloud properties by altering the humidity of the boundary layer and by modifying the partitioningroportioning of the sensible and latent heat fluxes (Zhang and Klein 2013; Berg et al., 2013), while the presence of ShCu provides a negative feedback by shading the surface and reducing its solar heating (Berg et al., 2011; Xiao et al., 2018). The ShCu exhibit strong spatial and temporal changes and these changes have variability that has been a
45 topic of increasing interest in recent years for both observational (Berg and Kassianov, 2008; Chandra et al., 2013) and model (Yun et al., 2017; Angevine et al., 2018) studies. To improve our understanding of these changescloud cover variability and their its impact on the intricate cloud-atmosphere-surface interactions, both long-term and detailed diurnal observations of the ShCu properties are highly desirable. There are two conventional measurement-based estimates of cloud cover: (1) cloud fraction (CF) obtained from zenith-pointing narrow-FOV observations and defined as the fraction of time
50 when a cloud is detected within a specified period, and (2) fractional sky coverage (FSC) obtained from wide-FOV observations and defined as the fraction of cloudy pixels in a sky image. Note that FSC is similar to that estimated by a cloudy-sky observer (e.g., Henderson-Sellers and McGuffie, 1990; Kassianov et al., 2005; Long et al. 2006).ShCu cover is a fundamental property that is estimated from ground based observations in two distinct ways: 1) Cloud fraction (CF) is defined for zenith pointing observations as the fraction of time cloud is detected within a defined time period. Alternately,
55 the fractional sky coverage (FSC) is the fraction of the visible sky covered with clouds. 2) Fractional sky coverage (FSC) is the fraction of cloudy pixels in a hemispherical sky image, most similar to historic manual observations (Henderson-Sellers and McGuffie, 1990; Kassianov et al., 2005; Long et al. 2006).

Long-term measurement-based statistics of ShCu have been employed for model-observation comparison (Chandra et al., 2013; Zhang and Klein, 2010, 2013;Lamer and Kollias, 2015Zhang et al., 2017; Endo et al., 2018). The observational
60 studies take advantage of the synergistic use of ground-based observations from the ceilometer, micropulse lidar (MPL) and millimeter cloud radar at the Atmospheric Radiation Measurement (ARM) mid-continental Southern Great Plains (SGP) site to obtain vertically-resolved hydrometeor layers with high temporal resolution (https://www.arm.gov/). A consistent merged radar-lidar-radar data product is available for 20 years from Nov. 1996 to the present at the SGP site and has served as a

2

basis for developing ShCu climatology at the SGP site (Berg and Kassianov 2008; Zhang and Klein 2013) and observational
testbeds for model evaluation (Zhang et al., 2017).

Efforts in improving both ground-based observations and modeling of continental ShCu at the SGP site are currently underway (e.g., Gustafson et al., 2017; Zhang et al., 2017). ~~For example, a recent report (~~Zhang et al., (2017) ~~reveals~~ suggested that, on average, the modeled areal cloud cover tends to underestimate observed CF substantially (up to 0.1 or 65% for clouds with vertical extent greater than 0.3 km). Although several potential factors for the obtained model-data discrepancy have been considered, including different model setups and selection of events with different vertical cloud extents, the role of observational uncertainties has not been directly addressed. For example, FSC obtained from the wide-field-of-view (FOV) Total Sky Imager (TSI) data as the fraction of cloudy pixels in a hemispherical image provides a better agreement with model outputs than the narrow-FOV CF from merged ceilometer-MPL data (Gustafson et al., 2018), indicating a potential consequence of the FOV configuration. In addition, long-term averages of CF obtained from merged ceilometer-MPL data tend to be larger than FSC acquired from collocated TSI observations (Boers et al., 2010; Qian et al., 2012; Wu et al., 2014; Kennedy et al., 2014), indicating a potential consequence of instrument-dependent cloud detection differences. Moreover, sampling of LES-generated cloud fields by a virtual instrument can be a helpful way to reconcile debated differences between the retrieved and predicted values of cloud cover (Oue et al., 2016).

It is expected that the impacts of instrument-dependent cloud detection and FOV configurations are entangled, and our study aims to assess their relative importance on cloud cover estimates of single-layer continental ShCu observed at the SGP site. In 2010, upgrades of three instruments (TSI, MPL and radar) took place, and data merging improvements were implemented on the ceilometer-MPL and ceilometer-MPL-radar (hereafter lidar-radar) merged datasets. Our study uses the natural separation provided by these upgrades and improvements to assess the relative impact of instrument-dependent cloud detection on the ShCu cover derived from three commonly used CF calculations, and FSC. Our study also uses wide-FOV and narrow-FOV TSI-based observations to assess the relative impact of the FOV configuration on FSC. Two main questions guide our investigation: 1) Have significant changes in the observations of ShCu cover occurred at the SGP site due to instrumental and algorithmic upgrades? And, 2) what is the impact of FOV configurations on hourly and sub-hourly observations of ShCu cover? These questions are addressed using statistical approaches as well as a new tool that integrates the narrow- and wide-FOV observations to visualize the two considered impacts for a given day of interest at user-specified temporal and spatial scales.

2 Data

Eighteen years (2000-2017) summertime (May-Sept) observations of FSC, cloud base height (CBH), cloud top height, and wind speed and direction were collected at the ARM SGP Central Facility. The data described below are available at https://www.arm.gov/data, and references to detailed instrument descriptions and records are in Table A1. Here is a short summary of data used in our analysis, and Appendix A contains pertinent information for their application.

1) **Merged lidar and ~~radar~~ lidar-radar data products**. A time-series of CBH and cloud top height values are obtained from the combined 34.86 GHz millimeter cloud radar (MMCR), MPL, and ceilometer data in the Active Remotely-Sensed Cloud Locations (ARSCL) value added product (Data reference: Johnson and Giangrande, 1996). The merged dataset improves upon the MMCR vertical profiling of hydrometeors by employing an MPL cloud mask to identify CBH and aide in mitigating the significant impact of insect returns (Clothiaux et al., 2000). The screened high temporal (10 s) and vertical (45 m) radar reflectivity best estimate is used to identify the vertical distribution of hydrometeor layers above the SGP. The CBH best estimate is obtained from merged ceilometer-MPL data. A cloud top height estimate is derived from merging the radar reflectivity best estimate, CBH best estimate, and MPL cloud mask. In 2011 the MMCR was replaced and renamed to Ka-band Zenith Radar (K~~Aa~~ZR) (Kollias et al. 2016) and ARSCL was subsequently updated and renamed to KAZRARSCL (Data reference: Johnson et al., 2011). A summary of the pertinent instrumental and algorithmic differences to the results herein for the periods before and after 2011 are provided in Appendix A.

2) **Ceilometer.** A time series of CBH values is obtained from the ceilometer (Data Reference: Ermold and Morris, 1997). The ceilometer records up to three cloud bases at heights up to a maximum of 7.7 km and 16 s resolution. Only the lowest cloud-base ~~height~~ is used here.

3) **Total Sky Imager cloud mask**. In addition to the sky image, the TSI processing produces a cloud mask (Data reference: Holben et al., 1994) every 30 seconds that classifies each image pixel as clear, opaque, or thin cloud, and identifies a 25° FOV circumsolar region and an inner zenith region (Long et al., 2006). The percent error of in FSC retrieval is estimated to be 10% for 95% of observations (Long et al., 2001).

4) **ShCu Event Selection**. The newly released ARM Shallow Cumulus data product (Data Reference: Shi et al., 2000) identifies times of ShCu from lidar / radar cloud boundary heights and includes FSC from TSI observations (Lim et al., 2018~~identifies ShCu cloud coverage using cloud boundary heights and cloud cover observations (Lim et al., 2018~~). The dataset provides hourly flags for the presence of ShCu, and a flag indicating the additional presence of other cloud types, such as commonly overlying cirrus that can either be intermittent or persistent throughout the day. From the Shallow Cumulus data product, we initially select days with single-layer (no upper-level) ShCu with at least 2-h duration. Once selected, all hours in this day with ShCu are included, additionally extending the start and end-times by 1 hour each. The extension allows for more accurate determination of the start- and end- times of the event on the finer time scale of the TSI FSC (15 min). Quality control procedures (Sect. 3.4) are used to censor multi-layer clouds and clear sky conditions on the 15-min and hourly observations of cloud cover.~~to more accurately capture the start and end of the event on a 15-min time scale, as defined by the total sky imager FSC.~~ The initial selection provided 614 candidate days with a total of 2393 hours.

5) **915 MHz Radar Wind Profiler**. Wind data are obtained at CBH from the low power setting of the 915 MHz Radar Wind Profiler which has a range-gate spacing of 60 m and hourly reporting intervals (Data reference: Muradyan and Coulter, 1997). In the event of missing data at CBH, the first available reading above 500 m is used.

4

130   In the event of missing data, 10m surface wind data are used as a tertiary means to impute wind data (Data reference: ARM, 1994). The wind data is used qualitatively in this analysis.

3 Methods

We calculate five estimates of the cloud cover from active and passive observations. Three estimates define narrow-FOV cloud fractions (CFs) retrieved from the lidar-radar observations (Sect. 3.1) to assess the impact of the instrument-dependent 135 cloud detection and data merging on cloud cover estimates (Sect.4.1). Two estimates define fractional sky cover (FSC) with 100°-FOV (Sect. 3.2) and narrow-FOV (3x3 pixels within 100°FOV) (Sect.3.3) obtained from TSI observations to estimate the impact of the FOV configuration on CF and FSC comparisons (Sect. 4.2). The resulting five ShCu cloud cover estimates are then combined, and data selection and quality control procedures are applied to the dataset (Sect. 3.4). The cloud cover estimates are compared for the entire period (2000-2017) and two sub-periods (2000-2010, and 2011-2017) separated by 140 instrumental and algorithmic upgrades (Sect.4). Finally, a heuristic tool is developed ~~(Sects. 3.5 and 4.3)~~ to visualize impacts of the cloud detection, data merging and FOV configuration on the cloud cover estimates (Sect. 4.3).

3.1. Narrow-FOV Cloud Fraction (CF)

We consider CF calculated directly from the ARSCL (or KAZRARSCL) data products as a fraction of 10 s (or 4 s) cloudy returns over a given time (e.g., Xi et al., 2010). Thus, the CF defines the frequency of cloud occurrence from the combined 145 narrow-FOV lidar-radar measurements. ShCu returns are defined using thresholds on 10 s (or 4 s) CBH and cloud top height information (Table 1) in a binary fashion. The three estimates of ShCu cover from lidar-radar observations are 1) CF from the merged lidar-radar products uses CBH information from the merged ceilometer-MPL and cloud top height from the merged MPL-radar. This method has the advantages of low missing data due to use of multiple instruments ~~used~~ and incorporates information about cloud top height consistent with the definition of shallow convection. ~~limits the vertical extent~~ 150 ~~of clouds.~~ The MPL has the potential to attenuate in thick clouds, necessitating radar data for cloud top height retrieval. ~~The radar is the only instrument that can reliably return cloud top height.~~ Insect contamination may contribute to significant uncertainty of the radar-based retrievals of cloud boundaries. Therefore, our analysis employs a semi-quantitative threshold approach when using the cloud top heights. This approach is less sensitive to the insect contamination. 2) CF from merged ceilometer-MPL uses CBH only and also has the advantage of low missing data as either ceilometer or MPL are used to 155 determine CBH. 3) CF from the ceilometer alone is the most common method that only uses CBH information, but has a disadvantage of missing data and limited vertical range for detecting high level cloud (Table 1). Additionally, total cloud fraction CF is calculated from merged ceilometer-MPL for any cloudy return detected within the narrow-FOV vertical column to screen multiple cloud layers (Sect. 3.4), this merged product has an extended range to 10 km from the MPL.

        Narrow-FOV observations of CF, CBH, and wind are computed for two averaging periods: 1) a 30-min period 160 centered on the 15-min averaging time for FSC, meaning the CF time period begins 7.5 minutes before the FSC averaging

5

time-bin and ends 7.5 afterwards; and 2) a 60 min period. The extended observation time (30-min instead of 15-min) for the CF measurement is aimed to compensate for the additional sky area observed by the wide-FOV TSI during a 15-min period. Hourly CF observations and FSC averages are computed over identical averaging periods for consistency with previous studies.

3.2 Wide-FOV Fractional Sky Cover (FSC)

FSCs obtained from TSI observations within the 100° and 160° FOV are available from routine TSI processing. The 100° FOV is used in this analysis because it was previously established to best correspond to nadir observations due to a reduced impact of cloud sides on estimated cloud cover (Kassianov et al. 2005). Prior to 2005, the TSI used 20° FOV rather than 100°; therefore, we calculate 100° FOV FSC for an additional period (2000-2005) directly from the TSI cloud mask as the fraction of thin and opaque pixels over total number of pixels. In our calculations, we remove a 25° FOV ellipse surrounding the sun location for consistency with the routine TSI processing. Our product compares well with the available TSI product, only 2% of the calculated FSCs exceed the processed FSCs by more than 0.05. The total (opaque + thin) FSC is calculated for each cloud mask, and then averaged in non-overlapping 15-min and 60-min intervals. Note that 15-min is consistent with the expected decorrelation time of a ShCu field (Kassianov et al. 2005). The TSI does not contain CBH information therefore selection of single-layer ShCu time periods requires the joint use of active sensors.

3.3 Narrow-FOV FSC ("CF-like")

A narrow-FOV FSC is calculated from the TSI cloud mask with the sampling, sensitivity and processing of the wide-FOV FSC. These calculations are aimed to mimic the narrow-FOV lidar-radar measurements ("CF-like"). The narrow-FOV represents a 3x3 pixel region located at a zenith angle of 20° and azimuthal angle of 315° from each TSI cloud mask. There are three main reasons for selection of this region. This region (1) is close to zenith (the zenith region is obstructed by the camera box), (2) is located outside of the sun circle region during the study hours (local daylight 9:00-18:00), and (3) corresponds to the best agreement with the wide-FOV FSC observations (Wagner & Kleiss, 2016). The TSI image resolution has been increased from 352 x 288 pixels (original resolution) to 480 x 640 pixels (improved resolution) in August 2011. The narrow-FOV (3x3 pixel region) "CF-like" observation has 4.1° (original) and 2.3°(improved) angular resolution. For comparison, the narrow-FOV ARSCL cloud products, including the lidar-radar CF, have much finer angular resolution (about 0.2°). The corresponding spatial resolution of the lidar-radar CF can be estimated by multiplying wind speed at cloud base height by lidar-radar dwell time (about 10 sec). For example, its spatial resolution is about 100 m for 10 m/sec wind speed. ~~The corresponding narrow FOV (3x3 pixel region) is 4.1° (original resolution) and 2.3°(improved resolution). For reference, the ARSCL product has a 0.2° FOV.~~

6

3.4 Data Selection and Quality Control Procedures

Data selection and quality control procedures are applied to the averaged and merged dataset. The scatterplots for 30-min CF and 15-min FSC (Fig. B1) illustrate application of the quality control criteria while resulting data completeness is provided in Table B1. These screening criteria are executed on the times with ShCu identified in Sect. 2 and includes the following steps.

1) Periods with clear (averaged FSC < 0.05) or overcast (averaged FSC>0.95) conditions are removed from this analysis for all variables.

2) Periods with multiple cloud layers are screened using merged ceilometer-MPL data where the CF from all clouds is required to be within 0.1 of the ShCu fraction (Berg et al., 2011) (Fig. B1a)

3) Individual 10 s values from the merged ceilometer-MPL product that were flagged as suspect are excluded from the analysis. Incomplete data contributes to errors in the CF calculation (Fig.B1b), and so only periods with 100% available data are used in the analysis.

4) Periods with questionable TSI retrievals are censored using a threshold on allowed thin clouds. The thin cloud classification of the TSI represents an uncertain pixel, which is "not really clear or cloudy" and typically appears as a thin border surrounding the bright and easily visible ShCu clouds. However, camera degradation and normal automated white balance adjustments of the camera can lead to excessive thin cloud as well as erroneous opaque regions. Manual inspection of these time periods results in a QC criterion that discards intervals with thin cloud fraction greater than 0.3 (Fig. B1c).

5) Additional requirement for comparisons using the ceilometer dataset. Individual observations in which the quality-control flag was marked as good are retained and those with faults are censored from the dataset. Averaging periods with missed or censored ceilometer data are discarded. The ten-year (2000-2010) period has a substantial (~40%) fraction of missed ceilometer data (Table B1).

Initial selection criteria for the partially-cloudy events with single-layer ShCu (criteria 1 and 2) reduced the dataset from 614 to 609 days. Quality control (steps 3 and 4) on the 15-min averaged data further reduced it to 569 days, and from 2393 to 2048 hours (Table B1). It is important to note that this dataset is not strictly selective against cases that transitioned from another cloud type (i.e. stratus to cumulus), nor is it selective against cases that are driven by large-scale weather systems.

3.5 Spatial analysis of TSI cloud mask

TSI cloud mask images are analyzed for FSC differences along the cross-wind direction to visually assess the spatial inhomogeneity of the cloud field and its potential impact on disagreement between narrow- and wide-FOV observations. Prior to our analysis, the 25° FOV sun circle is removed, the cloud mask is cropped to the 100° FOV and undistorted from hemispherical to rectilinear coordinates (Chow et al., 2011). Mean composite images are generated by summing all images within 15-min and dividing by the number of images (Fig. 1). The composite image is then divided into 21 equally spaced

7

4 Results and Discussion

4.1. Cloud detection and data merging

Instrument-dependent values of the estimated CF (Table 1, Sect. 3.1) and the narrow- and wide-FOV FSCs (Sect. 3.2, 3.3) are generated for the whole 18-yr period (2000-2017) and two sub-periods (2000-2010 and 2011-2017). The comparison of the cloud cover estimates from the sub-periods illustrates the joint and individual impacts of instrumental upgrades (Sect. A.2 - A.3) and data merging schemes (Sect. A.4) on the long-term statistics of cloud cover. The comparison includes both 15-min (e.g., Lareau et al., 2018) and 60-min (e.g., Zhang et al., 2017) averaging windows to assess the impact of the shorter averaging time on the CF-FSC agreement.

We begin our investigation of differences in cloud cover estimates by comparing mean values (Tables 2 and 3). The 60-min average FSCs obtained for the two subperiods are comparable (0.34 vs. 0.33), and their distributions are very similar (comparison not shown) indicating weak changes in the mean ShCu cloud cover for the two sub-periods. The ceilometer-based CF supports this interpretation as mean values show good agreement with FSC, particularly for the later sub-period (2011-2017) where the ceilometer data have good completeness. However, the mean CF calculated from the merged instruments are larger than FSC by 0.12 for the ceilometer-MPL, and 0.08 for the merged lidar-radar in the early sub-period (2000-2010). The larger values of the CF from the merged data in comparison with the ceilometer and FSC in the early sub-period are consistent with previous studies which focus on all cloud types and sky conditions (Boers et al., 2010; Qian et al., 2012; Wu et al., 2014; Kennedy et al., 2014). However, the FSC-CF correlation obtained for the single-layer ShCu is greatly improved over those from the previous studies with all cloud types. For example, Wu et al. (2014) reported a correlation coefficient of 0.54 comparing hourly CF and FSC when ignoring clear and overcast conditions. Here, the correlation coefficients for hourly data are much higher (0.8-0.84) (Table B3). For the later sub-period (2011-2017), the mean CF from the merged instruments decreases by 0.05 for the combined ceilometer-MPL and 0.09 for the combined lidar-radar. As a result, the combined lidar-radar CF shows an improved agreement with the FSC, whereas the CF from ceilometer-MPL is still larger than the FSC by 0.08 (Table 2).

The changes in mean values of the merged instrument CFs between the two subperiods are examined in the 1D and 2D histograms in Fig. 12. The figure shows strong correlation between merged ceilometer-MPL and merged lidar-radar CFs in the early sub-period, indicating that the cloud top height threshold employed has little influence on cloud fraction. In contrast, in the later sub-period the RMSD between the CFs is twice as large and the correlation is only moderate, indicating

a higher influence of the cloud top height criteria. The 1D histograms (Fig. 12a,b) show that both variables shifted to have more observations of low CF in the second subperiod, although this is most pronounced for the lidar-radar. These results suggest that changes in merging the ceilometer-MPL cloud base height detection impacted CF moderately, whereas changes in cloud top height detection in the merged lidar-radar data created significant differences in CF between the two periods.

The 2D histogram comparisons of merged lidar-radar CF and FSC (Fig. 23) show the elimination of bias across the entire range of FSC in the later sub-period; however, this may be due to the introduction of compensating errors using the cloud top height criteria in the updated merged lidar-radar product. Though the mean bias is reduced by 0.08 (Table 2) in the later period, the scatter about the 1:1 line is little affected as evidenced by only a 0.04 reduction in the RMSD in the later sub-period; additionally, 51% of CF values differ from FSC by more than 0.1 in the early period compared to 41% in the later sub-period. The modest improvement in agreement indicates the continued presence of cloud detection differences. The 2D histograms for the merged-lidar data CF vs FSC (Fig. 34) show better agreement across the entire range in the later sub-period, consistent with an improvement in ShCu cloud detection. The compensating error hypothesis is supported further by the comparisons of ceilometer CF to FSC (Fig. 45), which return the highest correlation coefficients (Table B3) and the lowest RMSD (Tables 2,3).

The new cloud radar (deployed in 2010) is expected to have improved ShCu detection in the lower atmosphere over the previous instrument. Therefore, it is not expected that the changes in radar detection are responsible for the *reduction* of merged lidar-radar CF in the later sub-period. We should note that the small droplet size of continental cumuli has been reported to impede their detection by both the new (Lamer and Kollias, 2015) and retired radar (Chandra et al., 2013). For example, Lamer and Kollias (2015) report that 37% of the hydrometeor detections by the ceilometer were missed by the radar at the SGP site. Most certainly, in either period, the radar would not contribute to an overestimation of CF. Instead it is expected that changes in how MPL and radar data are merged to retrieve cloud top height is likely the source of these significant differences seen between sub-periods (Sect. A.4). Though a number of differences exist, the incorporation of MPL data (below 3 km) in the ~~initial~~ original cloud top height retrieval would increase the number of detected cloud tops compared to those retrieved from the radar data alone for the initial period (2000-2010). Reliance only on the radar data for cloud top detection in the updated algorithm would result in fewer cloud top height detections and therefore a lower CF ~~algorithm between 2000-2010 but not the updated algorithm likely has a large impact~~ (see Sect. A.4 for more details).

Since ceilometer CF does not demonstrate dependence on sub-periods it is unlikely that the decrease in merged ceilometer-MPL CF between the two periods is due to the ceilometer. The merged ceilometer-MPL CF is always the same or greater than CF from the ceilometer alone, and the bias decreases between the two sub-periods from 0.19 to 0.11 (comparing times with complete ceilometer observations). This decrease in bias also corresponds to a decrease in the RMSD from 0.25 to 0.16, and an increase in Pearson's correlation from 0.86 to 0.91. The improvement in agreement might be attributable to improvements in the MPL mask that reduce false cloudy returns from boundary layer aerosols, resulting in a lower MPL CF (Sect. A3). It is also consistent with an increased reliance on the ceilometer for boundary layer retrievals in the cloud base best estimate retrieval algorithm (Sect. A.4). Specifically, the updates of the merging algorithm would likely result in more

frequent "clear" returns. The TSI and ceilometer data in comparison with the merged ceilometer-MPL and/or ceilometer-MPL-radar data may be preferable for consistent estimates of ShCu cloud cover at the SGP site from 2000 to present.

4.2. Impact of FOV on cloud cover measurements

Recall that the CF obtained from lidar-radar observations with narrow FOV represents a transect of a cloudy sky along the wind direction, while FSC acquired from wide-FOV TSI data defines an area of cloudy sky. Both the CF and the FSC are widespread measurement-based estimates of cloud cover. Fractional sky cover (FSC) from a 100° image spans about a 3 km width of the sky for a 1.5 km cloud base, whereas the active remote sensing instruments (radar, MPL, and ceilometer) observe only a "soda straw" approximately 10-m wide – essentially a point in the sky. Cloud fraction obtained from these narrow field of view instruments is thus essentially a 1-D transect through clouds that pass overhead. Previous studies have reported sampling uncertainties in transect measurements for modeled random cloud fields (Astin et al., 2001; Berg and Stull, 2002; Kassianov et al., 2005). The term "random" refers to the random arrangement of clouds on a horizontal plane within a given domain. In particular, the previous model studies (Astin et al., 2001; Berg and Stull, 2002) have demonstrated that the cloud cover obtained from the transect measurements mimics the area-averaged cloud cover for non-organized (e.g. random) cloud fields well if the sample size is relatively large (or numerous individual clouds are sampled). Recently Oue et al. (2016) have showed that 10 or more ceilometers equally spaced across a 25 km width in the cross-wind direction are required to estimate the simulated cloud cover in the small (30 km) domain. Certainly, the number of ceilometers, their locations and averaging time required for an accurate estimation of the cloud cover depend on the spatial arrangement of clouds and wind speed. Conversely, poor agreement is expected for organized cloud fields, such as so-called "cloud streets," where individual clouds are arranged in rows along the mean wind direction within a given period of interest. To estimate the relative impact of the FOV configuration on the estimated cloud cover, we use the "CF-like" observations (Sect. 3.3) following an approach previously introduced by Wagner and Kleiss (2016). Recall that the narrow-FOV "CF-like" and the wide-FOV observations are both from the TSI and thus have identical sensitivities for ShCu detection. Thus, the differences between "CF-like" and FSC illustrate the FOV impact on the estimated cloud cover.

Comparison of "CF-like" and FSC shows symmetric variability around the 1:1 line (Fig. 56), and no indication of bias in average values (Tables 2 & 3). For the 1 h (30 min) observations, 23% (41%) of "CF-like" values differ from FSC by more than 0.1, and 5% (14%) differ by more than 0.2. Reducing the averaging time from 1 hr to 30 min increases the RMSD by 0.05 (Table 2) and decreases the correlation from 0.92 to 0.85 (Table B3). Similar trends in the RMSD, and correlation coefficients are seen for all the CF measurements as well. This result is consistent with a 15-min decorrelation time for the temporal FSC fluctuations (Kassianov et al., 2005). The "CF-like"-FSC comparison allows us to estimate that narrow-FOV observations are responsible for a ±0.1 uncertainty of for hourly observations of CF (Table 2).

Using the passive "CF-like" observations, we also emphasize the relative impact of the instrument-dependent cloud detection and data merging for actively sensed CF measurements. The variability about the best fits in Figs. 2-43 5 (comparing FSC to CF) is greater than is expected from FOV impacts alone (Fig. 56) for both hourly and sub-hourly

10

observations. For hourly measurements, 23% of the "CF-like" values differ from FSCs by more than 0.1, this is the expected impact of FOV alone. For the earlier sub-period (2000-2010), the percentages of the CF values differing by more than 0.1 from the wide-FOV FSC are 62%, 51%, and 34% for the merged ceilometer-MPL, lidar-radar, and ceilometer retrievals, respectively. For the later sub-period (2011-2017), which has the least bias, the corresponding percentages are 53%, 41%, and 30%. The large percentages of observations with significant uncertainty in both time periods highlight the importance of distinguishing between FOV effects from instrumental detection differences when using data for sub-hourly applications.

The effective spatial area sampled by either narrow or wide FOV instruments is a function of both sampling duration and wind speed. High wind speed in comparison with low wind speed (1) increases sample size for a given period and (2) tends to organize horizontal arrangement of clouds (e.g., Weckworth et al. 1999, Atkinson and Zhang 1996). These two factors associated with sample size and spatial arrangement of clouds should be considered when differences between cloud cover obtained from narrow- and wide-FOV observations as function of wind speed are considered (Fig. 6). In particular, Figure 6 illustrates that both CF-FSC and "CF-like"-FSC differences are reduced noticeably as the wind speed increases from 1 m/s to 3 m/s, and continue to reduce slightly as the wind speed grows up to 11 m/s. The CF-FSC and "CF-like"-FSC differences obtained at a higher wind speed (above 11 m/s) should be considered with caution due to limited number of the corresponding cases with high wind speed (e.g., fewer than 100 cases for 60-min time average). The increased sampling area associated with increased wind speed does not necessarily result in an improved agreement between the narrow- and wide-FOV observations for both hourly and sub-hourly observations due to the impact of wind speed on cloud organization.

4.3 Heuristic tool to evaluate individual cloud cover estimates.

We developed a tool to help understand the impact of different sources on individual CF measurements. These sources include data quality control, detection differences, and spatial cloud organization during the time period of interest (FOV-impact). Many factors contribute to cloud field organization including cloud cover, cloud size (Zhu et al., 1992), wind speed, cloud growth and decay rates, and the presence of cloud streets and clusters. The "quick-look" tool (Sect. 3.5) is used to distinguish the detection-based uncertainty from the FOV-impact during different sky conditions.

TSI cloud mask images are analyzed for FSC differences along the cross-wind direction to visually assess the spatial inhomogeneity of the cloud field and its potential impact on disagreement between narrow- and wide-FOV observations. Prior to our analysis, the 25° FOV sun circle is removed, the cloud mask is cropped to the 100° FOV and undistorted from hemispherical to rectilinear coordinates (Chow et al., 2011). Mean composite images are generated by summing all images within 15-min and dividing by the number of images (Fig. 7a1). Each pixel in the averaged image can be interpreted as a 15-min CF measurement from a narrow-FOV sensor. The variability of CF in the cross-wind direction can indicate the possible influence of cloud field organization on cloud cover estimates by narrow-FOV observations. The composite image is then divided into 21 equally spaced "lanes" parallel to the wind direction at CBH (lane FOV is 5°). The mean and interquartile FSCCF from each lane is then computed for this 15-min average, as well as the minimum, maximum,

11

25th and 75th percentile "lane" FSC values, e.g. the 1st, 5th, 16th, and 21st ranked lane FSC. Figure 71b shows that the cross-wind variability represented by the lane FSCCF means exceeds substantially the within-lane variability (errorvertical bars). The low variability within each lane indicates small changes of FSC within a lane for this period.

Figure 87 shows an example of "quick-look" results generated for one day (June 14, 2017). The narrow-FOV (5°) lane-FSC (Fig. 71) is used to estimate the FOV impact for each 15-min interval. The variability in lane-FSC for each 15-min interval is displayed as one row in the heatmap (Fig. 87a), and provides the spatial distribution of the narrow-FOV FSC perpendicular to the wind direction. The source of between lane-FSC variability can be inferred from the 15-min mean composite images themselves (Fig. 87b). For perfect data, the thumbnail images are a representation of the mean cloud field within the 100° FOV image. For typical windspeeds (4-15 ms^{-1}) it is common to observe streaks, as is seen in this example day, generated by the motion of clouds (see also Fig. 98a, b). For low wind speeds (less than 4 ms^{-1}), a mottled field of clouds emerges with pixel CF dependent upon the rate of cloud formation and decay (Fig. 98c). Less variability in cross-lane FSC is observed for large cloud cover (FSC > 0.7) especially for typical winds (Fig. 98d).

The lane-FSC is compared to actual 30-min CF measurements in Fig. 87c. The intention of this plot is to compare the narrow-FOV CF to the range of possible narrow-FOV FSC observed over different sky regions. By quantifying the uncertainty in cloud cover estimates due to the spatial variations in the mean cloud field, we are better equipped to separate FOV impacts from detection differences. For example, a noticeable cloud street appears from about 12:30-13:15 local time, characterized by a wider spread in the lane-FSC min-max and interquartile range (IQR). The mean FSC is approximately 0.5 during this time and the CF varies between 0.25-0.3, and is nearly equal to the lane-FSC minimum. The corresponding thumbnails in Fig. 87b appear non-uniform, suggesting that the CF measurement is impacted by the narrow FOV configuration. In contrast from 13:15 to 14:15 a more uniform mean cloud field is depicted, and the lane-FSC IQR is within 0.1 of the FSC. During this period, ceilometer CF is also within the lane-IQR. Surprisingly, CF from the merged lidar-radar divergesradar diverges from the ceilometer value for the first time in the day, and exceeds the lane FSC-maximum significantly. We interpret the larger merged lidar-radar CF to be an *over-estimation* of cloud cover associated with detection differences or data merging issues.

Figure 87 is also useful for quality checks as the "quick-looks" are not censored based on the criteria of Sect. 3.4. For example, on this day we can see little evidence of upper-level clouds by the close agreement of the merged ceilometer-MPL total CF (blue) and ShCu CF (red-dashed). The thumbnails illustrate image artifacts such as incomplete removal of sun glare (17:00-19:00), which would cause FSC to be overestimated. In this same period, the images otherwise show relatively uniform cloud cover, therefore CF$_{ceil}$ might be preferred over FSC. Sun glare is common in the TSI, and a method has been proposed to identify and correct its effects on the FSC (Long, 2010). The wind-direction, indicated by the red arrow, is provided to verify cloud motion with observed wind direction. This is particularly helpful if missing wind data at CBH has been imputed from other sources, as it can be incorrect. Sun glare and incorrect wind direction are significant limitations of this method's extensibility to a statistical analysis.

There are two main expected applications of the introduced "quick-look" tool. The first potential application is a classification of spatial organization of cloud fields using, for example, cross-wind cloud field variability (e.g. peaks and valleys in Fig. 7b) and within-lane variance of cloud amount (e.g. vertical bars in Fig. 7b). Numerous images generated by the "quick-look" tool (e.g., Figure 8b) for the extended period (2000-2017) can be considered as a valuable training dataset for machine learning with focus on automated detection of desired features of the cloud fields (e.g., "cloud streets") and unwanted contaminations of TSI images (e.g., Figure 9). Second potential application is a visual inspection of the generated images for a given period of interest (e.g., a short-term field campaign) to check for the impact of instrumental detection differences and cloud field organization on the observed cloud amount. Visual inspection may be feasible given a limited number (about 40) of ShCu events annually during the warm season. For example, a spread of the lane CFs (gray region in Fig. 8c) gives an idea about the cross-wind cloud field variability within a given FOV, and thus aids in understanding the difference between cloud amounts obtained from the narrow- and wide-FOV observations.

5 Conclusions

We compare single-layer ShCu cover estimates obtained from the narrow-FOV lidar-radar and the wide-FOV TSI data collected at the continental SGP site. The data represent an extended period (2000-2017 summers) and two sub-periods (2000-2010 and 2011-2017 summers), which mark the instrumentation upgrades and corresponding improvements in data merging. Our comparison includes the bias between the long-term mean values of the cloud cover estimates, the corresponding RMSDs, and correlation coefficients. The main conclusions are organized along the two guiding questions highlighted in the introduction:

1) *Have significant changes in the observations of ShCu cover occurred at the SGP site due to instrumental and algorithmic upgrades?* We demonstrate that the best agreement (bias and RMSD are within 0.04 and 0.12, respectively) occurs between the CF obtained from the ceilometer alone and FSC obtained from TSI data for the entire period and two sub-periods. In contrast, the CF calculated from combined ceilometer-MPL data has the largest disagreement with FSC in the first sub-period (positive bias of 0.12, RMSD of 0.2), with improvement in the second sub-period (positive bias of 0.08, RMSD of 0.17). This improvement is likely associated with updates to the MPL cloud mask, and improved merging strategies that rely more on the ceilometer. When incorporating a threshold on the cloud top height, the CF obtained from the combined ceilometer-MPL-radar data has better agreement with FSC in the first sub-period (positive bias of 0.08, RMSD of 0.18), and good agreement in the second sub-period (no bias, RMSD of 0.14). The strong period dependence of CF obtained from the combined ceilometer-MPL-radar data, is likely due to increased reliance on the radar for cloud top height returns (MPL is not used below 3km in the second sub-period). However, the radar has known difficulties in detecting cumuli; therefore, the improved agreement is likely due to partial compensating errors in cloud detection.

2) *What is the impact of FOV configurations on hourly and sub-hourly observations of ShCu cover?* TSI-based narrow- and wide-FOV FSC comparisons have no bias and RMSD of 0.1 for hourly observations for all considered time

periods. The obtained small RMSD shows that, on average, the narrow-FOV introduces an uncertainty in cloud cover of ±0.1
for 1-h CF measurements. The "penalty" for decreasing the averaging time to 30-min (15-min FSC) is to increase the
average uncertainty in CF to ±0.15. The majority (77%) of the comparisons are within 0.1 for hourly observations, whereas
only 59% of the 30-min (narrow-FOV) versus 15-min (wide-FOV) comparisons are within 0.1. This considerable impact of
the FOV configuration on sub-hourly observations of ShCu, confounded with instrumental detection differences motivates
the introduction of a new "quick-look" tool.

The "quick-look" tool uses a spatially resolved analysis of the cross-wind variability in cloud cover in the TSI 100°
FOV to provide an expected range of cloud cover that would be detected by a set of 21-narrow-FOV instruments with the
same detection properties as the TSI. We demonstrate the utility of the "quick-look" tool to identify the impacts of the FOV
configuration on the cloud cover estimates and to identify periods with organized cloud fields. In addition, the "quick-look"
tool can identify times with potential issues associated with data merging and cloud detection, as well as periods with TSI
data quality concerns. The developed data-centric tool may supplement outputs from the existing and future narrow-FOV
lidar-radar simulators (Bouniol et al., 2010; Tatarevic and Kollias, 2015; Lamer et al., 20187).

The presented dataset provides a unique opportunity to (1) contrast the long-term cloud cover estimates with
different instrumental sensitivities and data merging strategies and (2) describe the across-wind variability of cloud cover at
user-specified temporal and spatial scales. These data provide an observational foundation for a better interpretation of
quandaries such as model-to-data discrepancy of cloud cover (Zhang et al., 2017). Finally, the developed dataset is expected
to facilitate evaluation of the modeled shallow cumuli over small domains (~4 km x 4 km) at the SGP site. These studies
with strong model and observational components include the recent large eddy simulation (LES) ARM Symbiotic
Simulation and Observation (LASSO) project (Gustafson et al., 2015) and the Holistic Interactions of Shallow Clouds,
Aerosols, and Land-Ecosystems (HI-SCALE) field campaign (Fast et al., 2018).

Data Availability

The analysis dataset is derived from observations collected at the Atmospheric Radiation Measurement (ARM) Southern
Great Plains (SGP) site Central Facility in northern Oklahoma, USA and are freely available at https://www.arm.gov/data.
Specific data streams in Sect 2 can be requested through the interactive search tool, or accessed through the DOI in the data
reference.

The data from the figures, and novel quick look tool have been made available for all 614 days with detected
shallow cumuli on the ARM archive (https://www.archive.arm.gov/) as principal investigator data (PI reference name is
Kleiss): TSI composite images merged cloud fraction product for shallow cumulus cases (tsiQLtable). Searchable by DOI:
10.5439/1523254.

14

Appendix A: Description of Millimeter Cloud Radar (MMCR), Micropulse Lidar (MPL), and merging algorithm 2000-2017

A.1 Active **R**emotely **S**ensed **C**loud **L**ocations **product Overview**

MMCR, MPL, and ceilometer data are combined in the Active Remotely-Sensed Cloud Locations (ARSCL) value added product (Clothiaux et al., 2000). The product improves upon the MMCR vertical profiling of hydrometeors by employing an MPL cloud mask to identify CBH and aide in mitigating the significant impact of insect returns. The screened high temporal (10 s) and vertical (45 m) radar reflectivity best estimate (Kollias et al. 2016) is further distilled to generate a user-friendly data stream of the vertical distribution of hydrometeors above the SGP site (sgparsclbnd1clothC1.c1). Within this data stream is a field for best estimate of CBH ("CloudBaseBestEstimate") from merging MPL and ceilometer data, and two fields containing up to 10 hydrometeor layer determinations of cloud bottom ("CloudLayerBottomHeightMplCloth") and top ("CloudLayerTopHeightMplCloth") heights. These two fields are derived from the reflectivity best estimate, cloud base best estimate, and MPL cloud mask. In 2011, the ARSCL was replaced with the KAZRARSCL with corresponding data stream name "sgparsclkazrbnd1kolliasC1.c1". Although the reporting interval of the MPL and ceilometer are 30 s and 16 s respectively, the ARSCL reporting interval is 10 s corresponding to the MMCR, and 4 s for KAZRARSCL corresponding to the KAaZR.

Each instrument (MMCR, MPL and ceilometer) has challenges for detecting ShCu, therefore the data merging process can significantly impact the resulting cloud fraction. Radar reflectivity is weak for shallow cumuli with low liquid water path (below 50gm^{-2}) due to the small cloud droplet radii (3-7 μm); for this reason, approximately half of the clouds identified by the ceilometer are not seen by the radar (Chandra et al. 2013). Consequently, the lidar data (ceilometer and MPL) are used for CBH information.

The MPL has a greater sensitivity to cloud droplets than the ceilometer and MMCR, resulting in a larger CF (Donovan and van Lammeren, 2001; Kennedy et al. 2014). Some weaknesses of the MPL include 1) misidentification of aerosol layers as clouds, due to the increased sensitivity to aerosols of shorter wavelength of the laser, 2) difficulty retrieving cloud bases in the "blind spot" below 500 m (Welton and Campbell, 2002; Sivaraman and Comstock, 2011), 3) quick attenuation of the backscatter signal within clouds resulting in inaccurate cloud top height, 4) Longer integration time (30 s) may complicate returns in partially cloudy conditions. It was previously found that when MPL is merged with the ceilometer, the combined lidar CF is larger by 0.03-0.06 than CF from the MPL alone (Kennedy et al. 2014). The ceilometer has better visibility and accuracy in the lower atmosphere than the MPL, but its range is limited to 7.7 km. Below 3km, the ceilometer is used to define CBH if both MPL and ceilometer detect clouds, otherwise if just one instrument detects cloud then a CBH is returned from that instrument. Therefore, there is a preference to report a 10s period as cloudy if the lidars disagree.

The lidars are limited in determining cloud top height owing to the attenuation of the backscatter signal. Cloud top height from the MPL is typically assessed by setting a threshold on the attenuation of the return; for this reason, MPL cloud

15

top heights are usually considered effective top heights. Multiple cloud layer detection is better performed by the radar and the algorithmic updates in KAZRARSCL processing reflect this strength (Sect. A4).

A.2 Radar Updates (2011)

485 Many continental cumuli are optically thin (liquid water content less than 50 g/m^2) and have small droplets. As a result, the cloud radars (both MMCR and KAZR) are not able to "see" the majority of these clouds (e.g., Chandra et al., 2013; Lamer and Kollias, 2015). The ARM radars have undergone a number of upgrades over the years as was recently reviewed by Kollias et al. (2016) (see their Table 17-2). The millimeter cloud radar (MMCR, 1997-2010) was replaced by Ka-Band Zenith Radar (KaZR KAZR, 2011-present) with expected improved sensitivity to cloud properties. The KaZR KAZR differs

490 from the MMCR particularly due to the change in radar pulse compression type from Barker to non-linear frequency modulation modes, which improve the radar's sensitivity. Importantly, the cirrus mode (pulse length of 8000 ns) was dropped from the pulse sequence and a more sensitive moderate mode (4000 ns) was added. The MMCR and KaZR KAZR both employ a general mode (300 ns). The shorter pulse widths serve to increase the radar's visibility in the lower atmosphere. This is a consequence of an expanded range in the lower atmosphere that is gained when the radar's

495 transmission time is decreased. Another key feature of the upgrade was a transition from sequential pulse sequences employed in the MMCR to dual acquisition pulses in the KaZR KAZR. The sequential pulse sequence required a sampling period of 36 s in the earliest period prior to 2004, then 12-14s in 2004-2010 to cycle through the different modes. The KaZR KAZR only employs two modes (general and moderate), which can be transmitted in tandem, this reduces the sampling period to 4 s. The improvement in sampling period is reflected in the shortened reporting interval in the KAZRARSCL data

500 product. Cloud radars (both the MMCR and KAZR) are known to miss detection of ShCu with small droplets compared to the ceilometer. For example, Chandra et al (2013) have found that the MMCR misses the majority of continental ShCu observed at the SGP site.

A.3 MPL Updates (2010)

The MPL cloud mask calculates cloud boundaries predominately from the attenuated backscatter signal. The original cloud
505 mask involved a signal-to-noise threshold based on cloud droplet scattering (Campbell et al. 1998), whereas the updated cloud mask incorporated information on different scattering properties of aerosol and clouds (Wang and Sussen 2001). In 2004 the MPL non-polarized implementation was replaced with a polarized system. The MPL hardware was upgraded in 2010 to a fast switching mode that allowed for switching between linear and circular polarization channels. The cloud mask algorithm was also updated at this time to a methodology developed by Wang and Sassen (2001) which has been used to

510 process the entire MPL record at the ARM facilities, though this cloud boundary processing has only been used in ARSCL and KAZRARSCL since 2010. It is expected that the updated cloud mask would improve cloud returns. The updated mask accounts for the different scattering properties of aerosol and cloud droplets, and is expected to result in improved cloud returns. Interested readers are directed to the technical documentation for the MPL mask (Sivaraman and Comstock, 2011).

16

A.4 Cloud Detection Algorithm Developments (2011)

515 Many changes occurred between the summers of 2010 and 2011: 1) ARSCL was changed to use the MPL cloud mask of Wang and Sassen (2001) in 2010, 2) the new ~~KaZR~~ KAZR was deployed, and the MMCR was decommissioned, and 3) KAZRARSCL replaced ARSCL in 2011 including updates to how retrievals of cloud layer boundaries (bases and tops) are handled in the MPL cloud mask. It is beyond the scope of this paper to identify all the possible changes that could result in differences in the ARSCL- and KAZARSCL-based cloud detections. However, it is important to note that both the MPL and

520 MMCR are used to determine the cloud top heights within the boundary layer in ARSCL, whereas in KAZRARSCL only the ~~KaZR~~ KAZR is used to determine cloud top heights below 3 km due to well-known MPL difficulties detecting low clouds (Berg and Kassianov, 2008; K. Johnson, private communication, 2018). In KAZRARSCL the MPL mask is still used for removing clutter in the radar returns.

Updates to the cloud base best estimate algorithm are also implemented. In particular, when the ceilometer observed

525 clear skies and the MPL had no return, the ARSCL data is flagged as "possibly clear" whereas the KAZRARSCL data are flagged as clear. The ceilometer is used to judge clear or cloudy below 500 m in KAZRARSCL. Otherwise in both versions the ceilometer's cloud base is preferred if both instruments return cloud below 3 km, and if either detects cloud then a cloud base height is returned (with the exception noted above).

Appendix B. Supporting Figures and Tables

530 Scatter plots of sequential application of quality control procedures (Sect 3.4) for 30-min CF and 15-min FSC are shown in Fig. B1, while data completeness for each summer in this study is provided in Table B1. The regression coefficients and correlation coefficients for ~~F~~figures 3-6 are provided in Tables B2 and B3, respectively.

Author Contributions

All authors contributed to the experimental design, analysis plan, and editing of the manuscript. ER conducted the statistical

535 data analysis. JK implemented the "quick-look" tool. LR and LB advised on the active remote sensing instrumentation. CL advised on the TSI quality control. EK provided project oversight and direction. EK and ER lead the composition of the manuscript.

Competing Interests

The authors declare that they have no conflict of interest.

References

Angevine, W. M., Olson, J., Kenyon, J., Gustafson, W. I., Endo, S., Suselj, K. and Turner, D. D.: Shallow Cumulus in WRF Parameterizations Evaluated against LASSO Large-Eddy Simulations, Monthly Weather Review, 146(12), 4303–4322, doi:10.1175/MWR-D-18-0115.1, 2018.

Arakawa, A.: The Cumulus Parameterization Problem: Past, Present, and Future, J. Climate, 17(13), 2493–2525, doi:10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2, 2004.

ARM: ARM Best Estimate Data Pproducts (ARMBE) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility. [online] Available from: https://www.arm.gov/capabilities/vaps/armbe (Accessed 25 February 2019), 1994.

Astin, I., Girolamo, L. D. and Poll, H. M. van de: Bayesian confidence intervals for true fractional coverage from finite transect measurements: Implications for cloud studies from space, Journal of Geophysical Research: Atmospheres, 106(D15), 17303–17310, doi:10.1029/2001JD900168, 2001.

Atkinson, B. W. and Zhang, J. W.: Mesoscale shallow convection in the atmosphere, Reviews of Geophysics, 34(4), 403–431, doi:10.1029/96RG02623, 1996.

Berg, L. K. and Kassianov, E. I.: Temporal Variability of Fair-Weather Cumulus Statistics at the ACRF SGP Site, J. Climate, 21(13), 3344–3358, doi:10.1175/2007JCLI2266.1, 2008.

Berg, L. K. and Stull, R. B.: Accuracy of Point and Line Measures of Boundary Layer Cloud Amount, J. Appl. Meteor., 41(6), 640–650, doi:10.1175/1520-0450(2002)041<0640:AOPALM>2.0.CO;2, 2002.

Berg, L. K., Kassianov, E. I., Long, C. N. and Mills, D. L.: Surface summertime radiative forcing by shallow cumuli at the Atmospheric Radiation Measurement Southern Great Plains site, J. Geophys. Res., 116(D1), D01202, doi:10.1029/2010JD014593, 2011.

Berg, L. K., Gustafson, W. I., Kassianov, E. I. and Deng, L.: Evaluation of a Modified Scheme for Shallow Convection: Implementation of CuP and Case Studies, Mon. Wea. Rev., 141(1), 134–147, doi:10.1175/MWR-D-12-00136.1, 2013.

Boers, R., de Haij, M. J., Wauben, W. M. F., Baltink, H. K., van Ulft, L. H., Savenije, M. and Long, C. N.: Optimized fractional cloudiness determination from five ground-based remote sensing techniques, Journal of Geophysical Research, 115(D24), doi:10.1029/2010JD014661, 2010.

575 Bouniol, D., Protat, A., Delanoë, J., Pelon, J., Piriou, J.-M., Bouyssel, F., Tompkins, A. M., Wilson, D. R., Morille, Y., Haeffelin, M., O'Connor, E. J., Hogan, R. J., Illingworth, A. J., Donovan, D. P. and Baltink, H.-K.: Using Continuous Ground-Based Radar and Lidar Measurements for Evaluating the Representation of Clouds in Four Operational Models, J. Appl. Meteor. Climatol., 49(9), 1971–1991, doi:10.1175/2010JAMC2333.1, 2010.

Campbell, J. R., Hlavka, D. L., Spinhirne, J. D., Turner, D. D. and Flynn, C. J.: Operational Cloud Boundary Detection and
580 Analysis from Micropulse Lidar Data, in Proc. Eighth ARM Science Team Meeting, pp. 119–122, U.S. Department of Energy, Tucson, AZ, USA., 1998.

Chandra, A. S., Kollias, P. and Albrecht, B. A.: Multiyear Summertime Observations of Daytime Fair-Weather Cumuli at the ARM Southern Great Plains Facility, J. Climate, 26(24), 10031–10050, doi:10.1175/JCLI-D-12-00223.1, 2013.

Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J. and Washom, B.: Intra-hour forecasting with a
585 total sky imager at the UC San Diego solar energy testbed, Solar Energy, 85(11), 2881–2893, doi:10.1016/j.solener.2011.08.025, 2011.

Clothiaux, E. E., Ackerman, T. P., Mace, G. G., Moran, K. P., Marchand, R. T., Miller, M. A. and Martner, B. E.: Objective Determination of Cloud Heights and Radar Reflectivities Using a Combination of Active Remote Sensors at the ARM CART Sites, J. Appl. Meteor., 39(5), 645–665, doi:10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2, 2000.

590 Coulter, R.: Micropulse Lidar (MPL) Handbook DOE/SC-ARM/TR-019, ARM Climate Research Facility [online] Available from: https://www.arm.gov/publications/tech_reports/handbooks/mpl_handbook.pdf (Accessed 8 February 2019), 2012.

Donovan, D. P. and Lammeren, A. C. A. P. van: Cloud effective particle size and water content profile retrievals using combined lidar and radar observations: 1. Theory and examples, Journal of Geophysical Research: Atmospheres, 106(D21), 27425–27448, doi:10.1029/2001JD900243, 2001.

595 Ermold, B. and Morris, V. R.: Ceilometer (CEIL) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility. [online] Available from: http://dx.doi.org/ 10.5439/1181954 (Accessed 25 February 2019), 1997.

Fast, J. D., Berg, L. K., Burleyson, C., Fan, J., Feng, Z., Hagos, S., Huang, M., Guenther, A., Laskin, A. and Ovchinnikov, M.: Holistic Interactions of Shallow Clouds, Aerosols, and Land-Ecosystems (HI-SCALE) Field Campaign Report
600 DOE/SC-ARM-17-014, ARM Climate Research Facility [online] Available from: https://www.osti.gov/servlets/purl/1354732, 2017.

Endo, S., Zhang, D., Vogelmann, A. M., Kollias, P., Lamer, K., Oue, M., Xiao, H., Gustafson, W. I. and Romps, D. M.: Reconciling differences between large-eddy simulations and Doppler-lidar observations of continental shallow cumulus cloud-base vertical velocity, Geophysical Research Letters, 0(ja), doi:10.1029/2019GL084893, 2019.

19

605 Gustafson, W. I. and Vogelmann, A.: LES ARM Symbiotic Simulation and Observation (LASSO) Implementation Strategy: DOE/SC-ARM-15-039, ARM Climate Research Facility [online] Available from: https://www.arm.gov/publications/programdocs/doe-sc-arm-15-039.pdf, 2015.

Gustafson, W., Cheng, X., Krishna, B., Toto, T., Vogelmann, A., Endo, S., Li, Z. and Xiao, H.: Description of the LASSO Alpha 2 Release: DOE/SC-ARM-TR-199., [online] Available from:
610 https://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-199.pdf, 2018.

Gustafson, W. I., Vogelmann, A., Cheng, X., Endo, S., Krishna, B., Li, Z., Toto, T. and Xiao, H.: Recommendations for Implementation of the LASSO Workflow. DOE/SC-ARM-17-031., [online] Available from: https://www.osti.gov/servlets/purl/1406259/, 2017.

Henderson-Sellers, A. and McGuffie, K.: Are cloud amounts estimated from satellite sensor and conventional surface-based
615 observations related?, International Journal of Remote Sensing, 11(3), 543–550, doi:10.1080/01431169008955038, 1990.

Johnson, K. and Giangrande, S. E.: Active Remotely-Sensed Cloud Locations (ARSCLBND1CLOTH) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility, Lamont, OK. [online] Available from: http://dx.doi.org/10.5439/1027283 (Accessed 25 February 2019), 1996.

620 Johnson, K., Giangrande, S. and Toto, T.: Active Remote Sensing of CLouds (ARSCL) product using Ka-band ARM Zenith Radars (ARSCLKAZRBND1KOLLIAS) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) User Facility, Oak Ridge, Tennessee, USA. [online] Available from: http://dx.doi.org/10.5439/1350630 (Accessed 1 May 2018), 2011.

Kassianov, E., Long, C. N. and Ovtchinnikov, M.: Cloud Sky Cover versus Cloud Fraction: Whole-Sky Simulations and
625 Observations, J. Appl. Meteor., 44(1), 86–98, doi:10.1175/JAM-2184.1, 2005.

Kennedy, A. D., Dong, X. and Xi, B.: Cloud fraction at the ARM SGP site, Theor Appl Climatol, 115(1–2), 91–105, doi:10.1007/s00704-013-0853-9, 2014.

Kollias, P., Clothiaux, E. E., Ackerman, T. P., Albrecht, B. A., Widener, K. B., Moran, K. P., Luke, E. P., Johnson, K. L., Bharadwaj, N., Mead, J. B., Miller, M. A., Verlinde, J., Marchand, R. T. and Mace, G. G.: Development and
630 Applications of ARM Millimeter-Wavelength Cloud Radars, Meteorological Monographs, 57, 17.1-17.19, doi:10.1175/AMSMONOGRAPHS-D-15-0037.1, 2016.

Lamer, K. and Kollias, P.: Observations of fair-weather cumuli over land: Dynamical factors controlling cloud size and cover, Geophys. Res. Lett., 42(20), 2015GL064534, doi:10.1002/2015GL064534, 2015.

Lamer, K., Fridlind, A. M., Ackerman, A. S., Kollias, P., Clothiaux, E. E. and Kelley, M.: (GO)^2-SIM: a GCM-oriented
635 ground-observation forward-simulator framework for objective evaluation of cloud and precipitation phase, Geoscientific Model Development, 11(10), 4195–4214, doi:https://doi.org/10.5194/gmd-11-4195-2018, 2018. Lamer, K., Fridlind, A. M., Ackerman, A. S., Kollias, P. and Clothiaux, E. E.: A Ground-Based Doppler Radar and Micropulse Lidar Forward Simulator for GCM Evaluation of Arctic Mixed-Phase Clouds: Moving Forward Towards an Apples-to-

20

apples Comparison of Hydrometeor Phase, AGU Fall Meeting Abstracts, 51 [online] Available from: http://adsabs.harvard.edu/abs/2017AGUFM.A51E2118L (Accessed 8 February 2019), 2017.

Lareau, N. P., Zhang, Y. and Klein, S. A.: Observed Boundary Layer Controls on Shallow Cumulus at the ARM Southern Great Plains Site, J. Atmos. Sci., doi:10.1175/JAS-D-17-0244.1, 2018.

Lim, K.-S., Riihimaki, L. D., Shi, Y., Flynn, D., Kleiss, J. M., Berg, L. K., Gustafson Jr., W. I., Zhang, Y. and Johnson, K. L.: Long-term cloud type retrieval using a combination of active remote sensors and a total sky imager at the ARM SGP site, J. Atmospheric Ocean. Technol, submitted, 2018.

Long, C. N.: Correcting for Circumsolar and Near-Horizon Errors in Sky Cover Retrievals from Sky Images., Open Atmos. Sci. J, 4, 45–52, 2010.

Long, C. N., Slater, D. W. and Tooman, T. P.: Total sky imager model 880 status and testing results, Pacific Northwest National Laboratory. [online] Available from: https://arm.gov/publications/tech_reports/arm-tr-006.pdf (Accessed 28 May 2015), 2001.

Long, C. N., Sabburg, J. M. and Calbó, J.: Retrieving cloud characteristics from ground-based daytime color all-sky images., Journal of Atmospheric & Oceanic Technology, 23(5), 633–652, 2006.

Morris, V. R.: Total Sky Imager (TSISKYCOVER) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility., Oak Ridge, Tennessee, USA. [online] Available from: http://dx.doi.org/10.5439/1025308 (Accessed 1 May 2018), 2000.

Morris, V. R.: Ceilometer Instrument Handbook DOE/SC-ARM-TR-020, DOE ARM Climate Research Facility [online] Available from: https://www.arm.gov/publications/tech_reports/handbooks/ceil_handbook.pdf (Accessed 8 February 2019), 2016.

Muradyan, P. and Coulter, R.: Radar Wind Profiler (915RWPWINDCON) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility., Oak Ridge, Tennessee, USA. [online] Available from: http://dx.doi.org/10.5439/1025135 (Accessed 1 May 2018), 1997.

Oue, M., Kollias, P., North, K. W., Tatarevic, A., Endo, S., Vogelmann, A. M. and Gustafson, W. I.: Estimation of cloud fraction profile in shallow convection using a scanning cloud radar, Geophys. Res. Lett., 43(20), 2016GL070776, doi:10.1002/2016GL070776, 2016.

Park, R.-S. and Kwon, Y. C.: The Implications for Radiative Cloud Forcing via the Link Between Shallow Convection and Planetary Boundary Layer Mixing, Journal of Geophysical Research: Atmospheres, 123(23), 13,203-13,218, doi:10.1029/2018JD028678, 2018.

Qian, Y., Long, C. N., Wang, H., Comstock, J. M., McFarlane, S. A. and Xie, S.: Evaluation of cloud fraction and its radiative effect simulated by IPCC AR4 global models against ARM surface observations, Atmos. Chem. Phys., 12(4), 1785–1810, doi:10.5194/acp-12-1785-2012, 2012.

Tatarevic, A. and Kollias, P.: User's Guide CR-SIM SOFTWARE v 2.0, McGill University Clouds Research Group. [online] Available from: http://radarscience.weebly.com/radar-simulators.html, 2015.

Shi, Y., Flynn, D. and Riihimaki, L.: Fair-Weather Shallow Cumulus Identification (SHCUSUMMARY) at Southern Great Plains (SGP) Central Facility, Lamont, OK (C1)., Atmospheric Radiation Measurement (ARM) user facility., Oak Ridge, Tennessee, USA. [online] Available from: http://dx.doi.org/10.5439/1392569 (Accessed 1 May 2018), 2000.

Sivaraman, C. and Comstock, J.: Micropulse Lidar Cloud Mask Value-Added Product Technical Report DOE/SC-ARM/TR-098, ARM Climate Research Facility, doi:10.2172/1019283, 2011.

Vial, J., Bony, S., Stevens, B. and Vogel, R.: Mechanisms and Model Diversity of Trade-Wind Shallow Cumulus Cloud Feedbacks: A Review, Surv Geophys, 38(6), 1331–1353, doi:10.1007/s10712-017-9418-2, 2017.

Wagner, T. J. and Kleiss, J. M.: Error characteristics of ceilometer-based observations of cloud amount, Journal of Atmospheric & Oceanic Technology, 2016.

Wang, Z. and Sassen, K.: Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors, J. Appl. Meteor., 40(10), 1665–1682, doi:10.1175/1520-0450(2001)040<1665:CTAMPR>2.0.CO;2, 2001.

Weckwerth, T. M., Horst, T. W. and Wilson, J. W.: An Observational Study of the Evolution of Horizontal Convective Rolls, Monthly Weather Review, 127(9), 2160–2179, doi:10.1175/1520-0493(1999)127<2160:AOSOTE>2.0.CO;2, 1999.

Welton, E. J. and Campbell, J. R.: Micropulse Lidar Signals: Uncertainty Analysis, J. Atmos. Oceanic Technol., 19(12), 2089–2094, doi:10.1175/1520-0426(2002)019<2089:MLSUA>2.0.CO;2, 2002.

Widener, K. B., Bharadwaj, N. and Johnson, K. L.: Ka-Band ARM Zenith Radar Handbook DOE/SC-ARM/TR-106, ARM Climate Research Facility [online] Available from: https://www.arm.gov/publications/tech_reports/handbooks/kazr_handbook.pdf (Accessed 8 February 2019), 2012.

Wu, W., Liu, Y., Jensen, M. P., Toto, T., Foster, M. J. and Long, C. N.: A comparison of multiscale variations of decade-long cloud fractions from six different platforms over the Southern Great Plains in the United States, Journal of Geophysical Research: Atmospheres, 119(6), 3438–3459, doi:10.1002/2013JD019813, 2014.

Xi, B., Dong, X., Minnis, P. and Khaiyer, M. M.: A 10 year climatology of cloud fraction and vertical distribution derived from both surface and GOES observations over the DOE ARM SPG site, J. Geophys. Res., 115(D12), D12124, doi:10.1029/2009JD012800, 2010.

Xiao, H., Berg, L. K. and Huang, M.: The Impact of Surface Heterogeneities and Land-Atmosphere Interactions on Shallow Clouds Over ARM SGP Site, Journal of Advances in Modeling Earth Systems, 10, doi:10.1029/2018MS001286, 2018.

Yun, Y., Fan, J., Xiao, H., Zhang, G. J., Ghan, S. J., Xu, K.-M., Ma, P.-L. and Gustafson, W. I.: Assessing the Resolution Adaptability of the Zhang-McFarlane Cumulus Parameterization With Spatial and Temporal Averaging, Journal of Advances in Modeling Earth Systems, 9(7), 2753–2770, doi:10.1002/2017MS001035, 2017.

Zhang, Y. and Klein, S. A.: Mechanisms Affecting the Transition from Shallow to Deep Convection over Land: Inferences from Observations of the Diurnal Cycle Collected at the ARM Southern Great Plains Site, J. Atmos. Sci., 67(9), 2943–
2959, doi:10.1175/2010JAS3366.1, 2010.

Zhang, Y. and Klein, S. A.: Factors Controlling the Vertical Extent of Fair-Weather Shallow Cumulus Clouds over Land: Investigation of Diurnal-Cycle Observations Collected at the ARM Southern Great Plains Site, J. Atmos. Sci., 70(4), 1297–1315, doi:10.1175/JAS-D-12-0131.1, 2013.

Zhang, Y., Klein, S. A., Fan, J., Chandra, A. S., Kollias, P., Xie, S. and Tang, S.: Large-Eddy Simulation of Shallow Cumulus over Land: A Composite Case Based on ARM Long-Term Observations at Its Southern Great Plains Site, J. Atmos. Sci., 74(10), 3229–3251, doi:10.1175/JAS-D-16-0317.1, 2017.

Zhu, T., Lee, J., Weger, R. C. and Welch, R. M.: Clustering, randomness, and regularity in cloud fields: 2. Cumulus cloud fields, Journal of Geophysical Research: Atmospheres, 97(D18), 20537–20558, doi:10.1029/92JD02022, 1992.

715

Figures
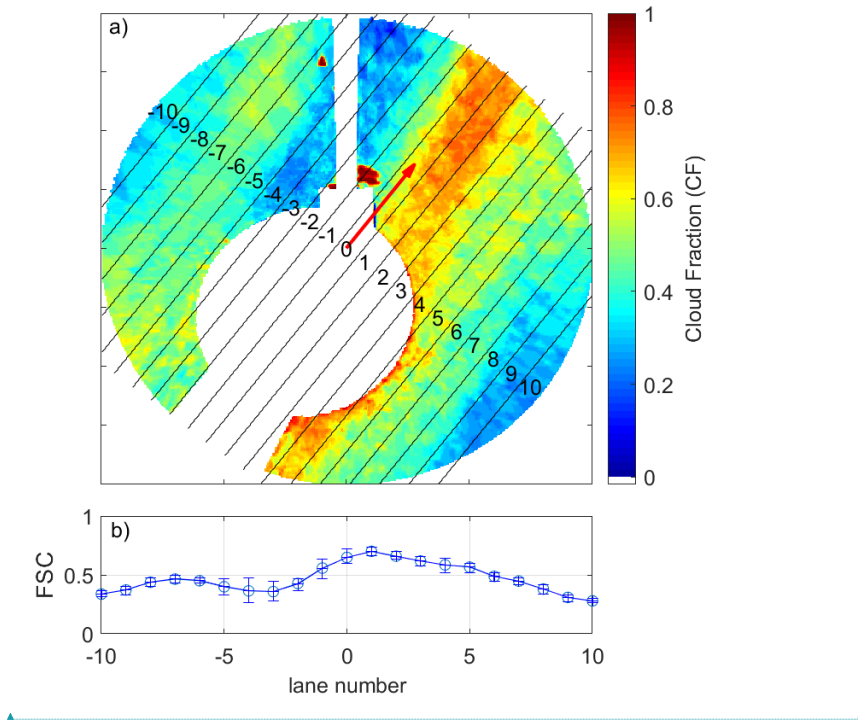
720   ~~Figure 1. a) A composite 15-min average cloud cover image (Sect. 3.5) created from 100° FOV TSI cloud masks captured every 30 s with sun circle region removed and projected to a rectilinear grid. Color bar indicates the fraction of time each pixel observes a cloud (thin or opaque). North is top of figure; streaks result from cloud motion across the image field. Image is divided into 21 lanes oriented parallel to the wind direction (red arrow points down-wind) obtained from a radar wind profiler at CBH. b) The mean and interquartile (error bars) of FSC observed within each of the 21 lanes. Note, the moderate (~20°) misalignment of the cloud streaks and the wind vector, illustrating the challenges of wind-based analyses.~~

725
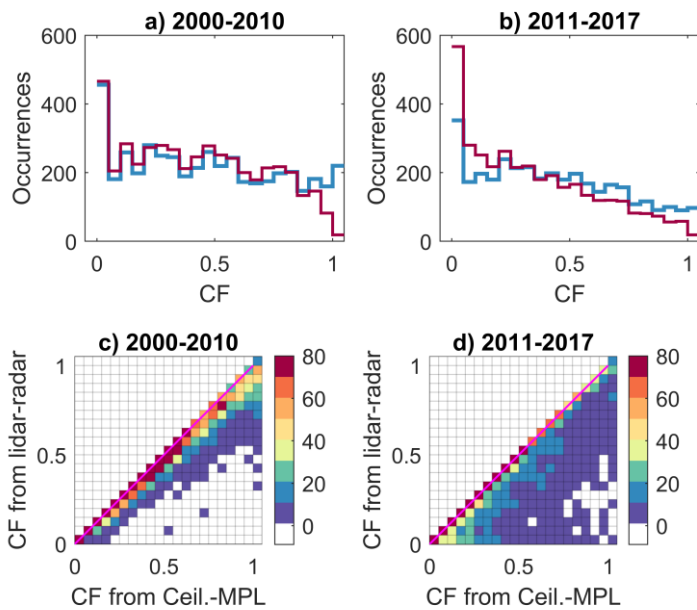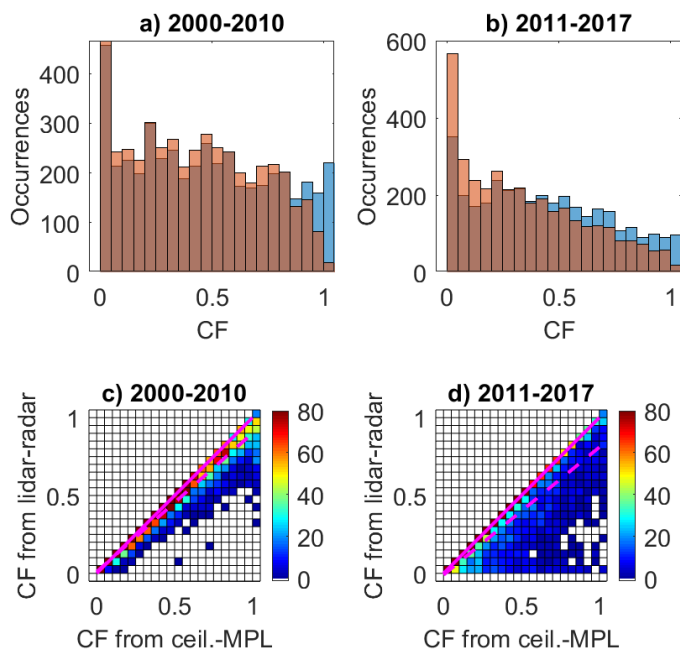
Figure 12. Relationship between CF from merged ceilometer-MPL (clouds identified by CBH) and from merged lidar-radar (clouds identified by both CBH and cloud top height). (a,b) histograms of CF from merged ceilometer-MPL (blue) and merged lidar-radar (~~brown~~red). (c,d) joint histograms (counts) for the two methods, magenta line is 1:1. (a,c) present data obtained from 2000-2010 (N = 4628), where the radar and MPL data are used jointly to determine cloud top height (Table 1). (b,d) present data obtained from 2011-2017, (N = 3564), where the cloud tope height is determined from radar data alone below 3 km. Each observation represents a 30-min average. RMSD (Pearson's correlation coefficient) in (c) is 0.08 (0.98) and 0.17 (0.87) in (d).

730
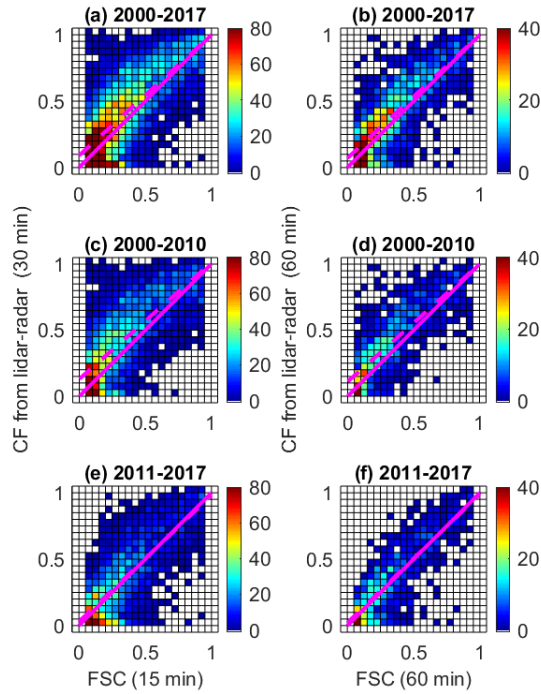
Figure 23. Two-dimensional (2D) occurrence distributions of paired-in-time CF and FSC for the three periods of interest: 2000-2017(a,b), 2000-2010 (c,d), and 2011-2017(e,f). Solid magenta line is 1:1 and dashed is least squares fit, with regression coefficients given in Appendix B (Tables B2,B3). Left (a,c,e) and right (b,d,f) columns define fine (15-min FSC and 30-min CF) and coarser (60-min FSC and CF) temporal scales respectively. Color scale represents counts in increments of 10. Each included data point excludes clear and overcast conditions using criteria: 0.05 < FSC <0.95. CF from merged lidar-radar method uses information about cloud base height (ceilometer-MPL) and cloud top height (radar-MPL).
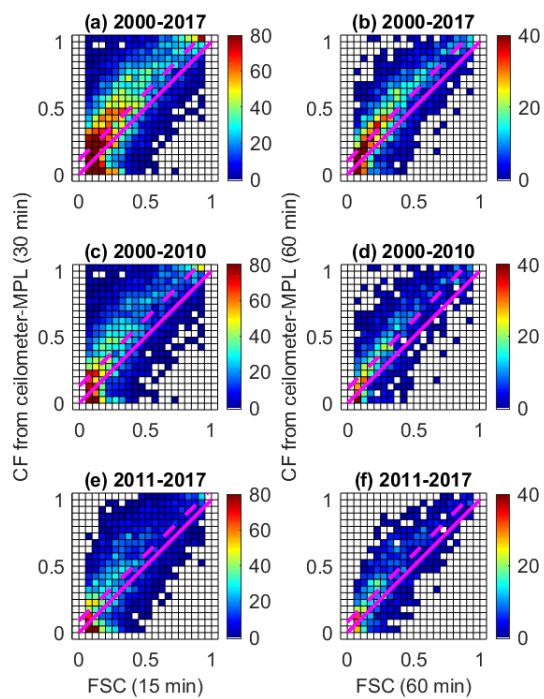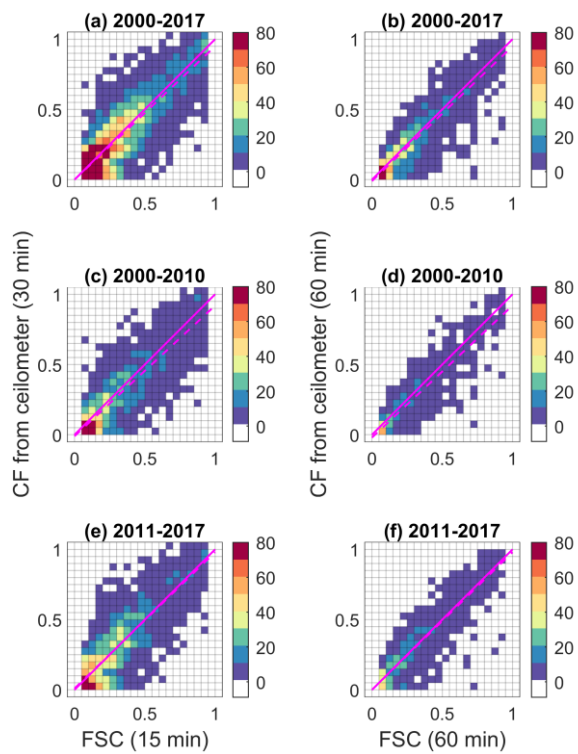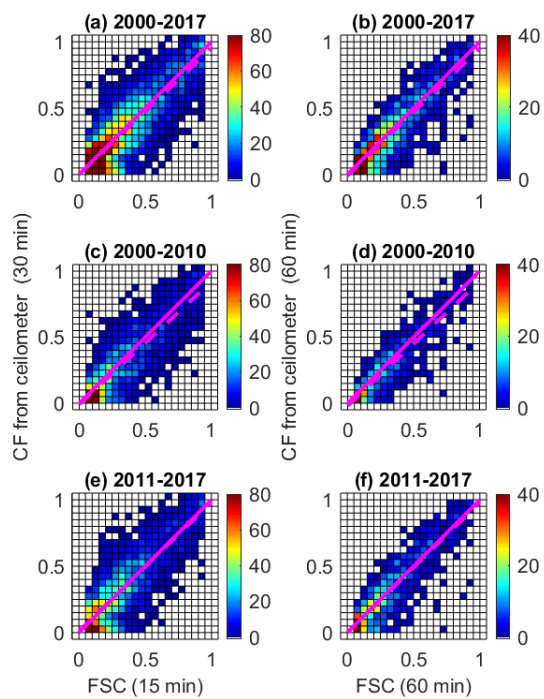
28
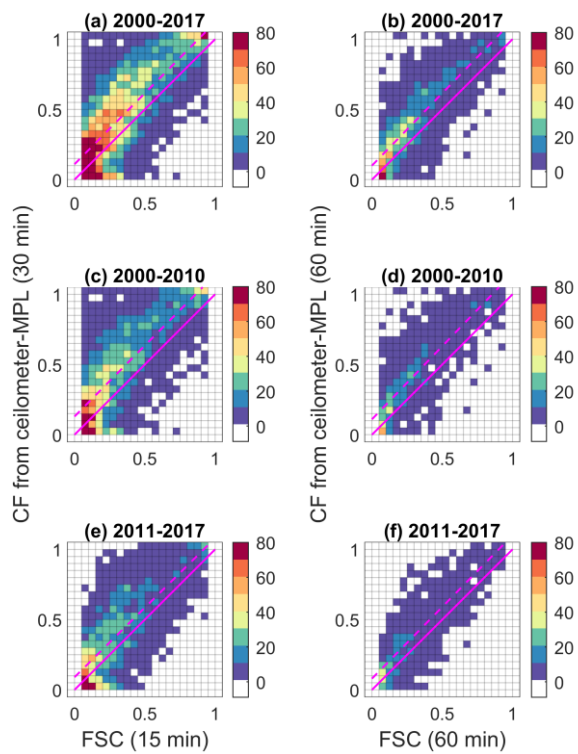
Figure 34. The same as Fig. 23, except CF from merged ceilometer-MPL method only uses information about cloud base height.

745

31

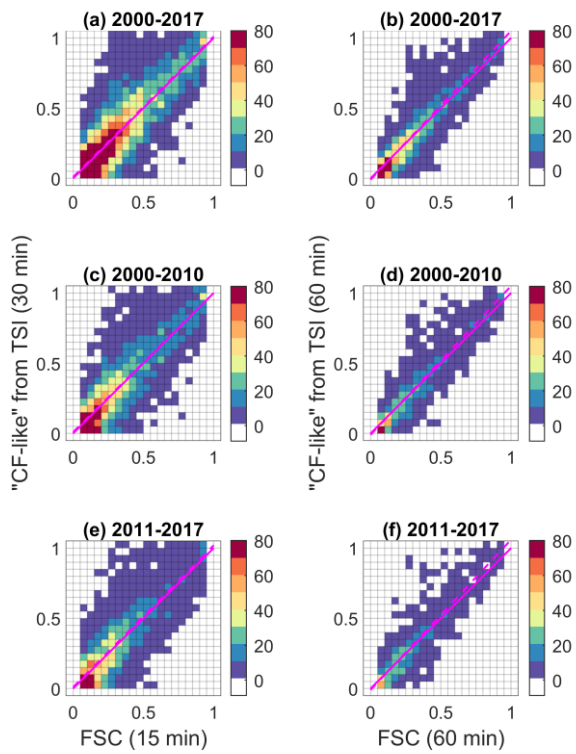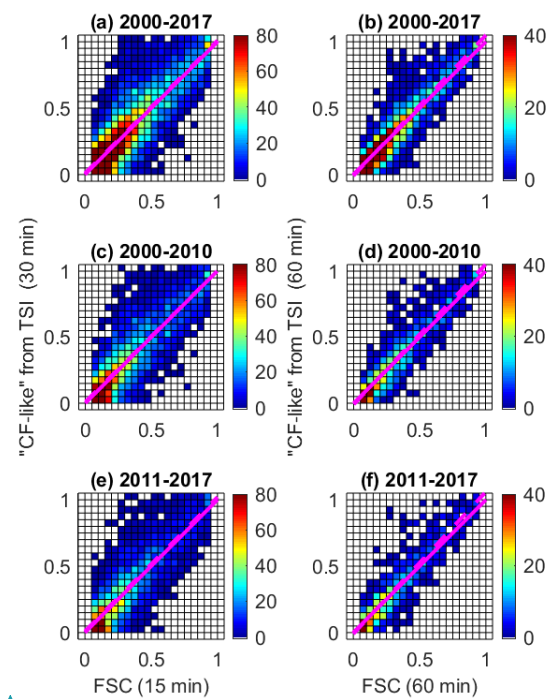750   **Figure 45. The same as Fig. 23, except CF only uses cloud base height information from the ceilometer .**

755

Figure 56. The same as Fig. 23, except for comparison between "CF-like" and FSC, where both are derived from the TSI instrument.
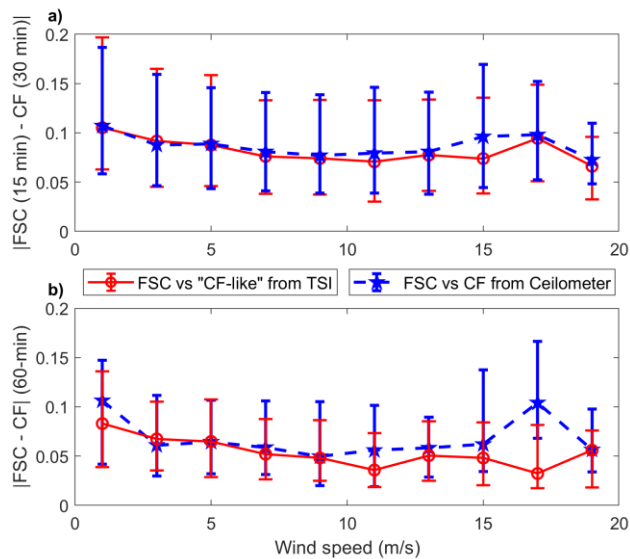
760

Figure 6. Absolute difference between FSC and CF as a function of wind speed at CBH (2 m/s bins). Markers indicate median, and bars indicate 25th and 75th quartiles. Red: absolute difference between FSC and "CF-like" from the TSI. Blue: absolute difference between FSC and CF from the ceilometer. a) 15-min FSC and 30-min CF and b) 60 min FSC and CF. Missing wind speed at CBH is replaced by RWP wind speeds at 500 m, missing RWP data is replaced using surface wind speeds.
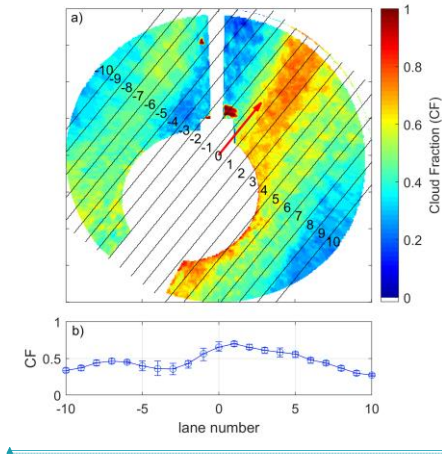
765

Formatted: Font: (Default) +Body (Times New Roman), 11 pt, Font color: Text 1

Figure 7~~1~~. a) A composite 15-min average cloud cover image (Sect. 4.3~~3.5~~) created from 100° FOV TSI cloud masks captured every 30 s with sun circle region removed and projected to a rectilinear grid. Color bar indicates the fraction of time each pixel observes a cloud (thin or opaque). North is top of figure; streaks result from cloud motion across the image field. Image is divided into 21 lanes oriented parallel to the wind direction (red arrow points down-wind) obtained from a radar wind profiler at CBH. b) The mean and interquartile (vertical~~(error~~ bars) of ~~FSC~~CF observed within each of the 21 lanes. Note, the moderate (~20°) misalignment of the cloud streaks and the wind vector, illustrating the challenges of wind-based analyses.
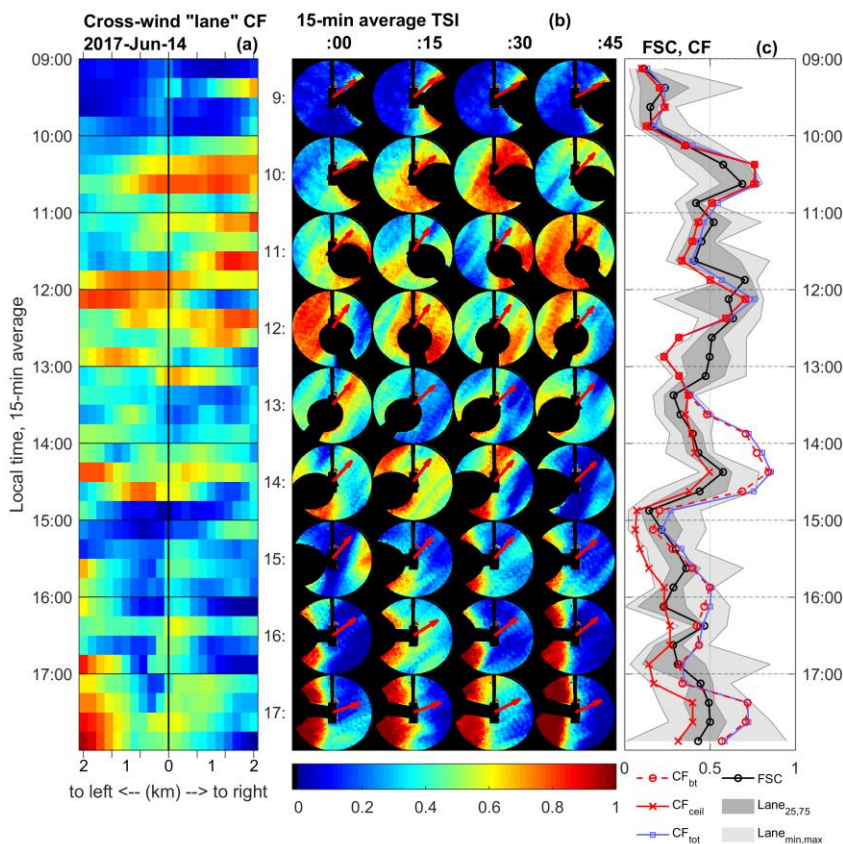
770

Formatted: Centered

775

Figure 87. Example of results from "quick look" tool for comparing FSC and CF for 14 Jun, 2017. (a) Heat map of 15-min average FSC CF indicated by the color bar. Cross-wind distance is calculated from the average CBH of 1.5 km, with zero representing the centre lane of the composite image (the row from 13:00-13:15 (LST) corresponds to the FSC values in Fig. 74b. (b) 15-min
780 **composite images within the 100° FOV, projected onto rectilinear grid. Red arrows indicate wind direction at CBH. (c) Comparison of cloud amounts obtained from TSI and lidar-radar data: 30-min averages of CF_{bt} (CF "cloud bases and tops") from merged lidar-radar for ShCu only (red dashed with circles), CF_{ceil} (CF "from ceilometer", red line with "x" symbol), and CF_{tot} (CF "total") from merged ceilometer-MPL for any cloud detection (blue line with squares). 15-min average FSC from TSI data (black with circles), 15-min minimum and maximum lane-averaged CF FSC (light grey shading) and interquartile lane FSC CF**
785 **(dark grey shading). Markers are placed at time bin centre.**
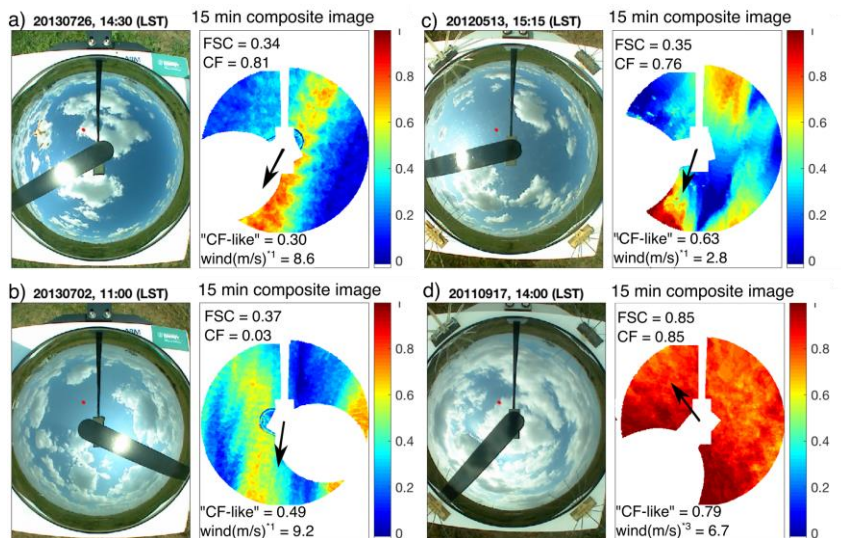
38

Figure 98. Illustrations of spatial patterns of ShCu manifest in 15-min averaged composite images (100° FOV, rectilinear coordinates) and a representative all sky image from the averaging period. Color bar is 15-min mean FSC for each pixel in the composite image, black arrow indicates wind direction, red dot in sky image indicates location of the 3x3 pixel for "CF-like". 15-min FSC, 30-min CF from merged lidar-radar method, "CF-like", and wind speed are provided as text. a) A bit of bird feces contaminates the mirror on the left-hand side of the image, fortunately not within the 100-degree FOV resulting in no error of FSC. a-c) *1: wind source is radar wind profiler at CBH, d) *3 is 10 m surface wind speed.
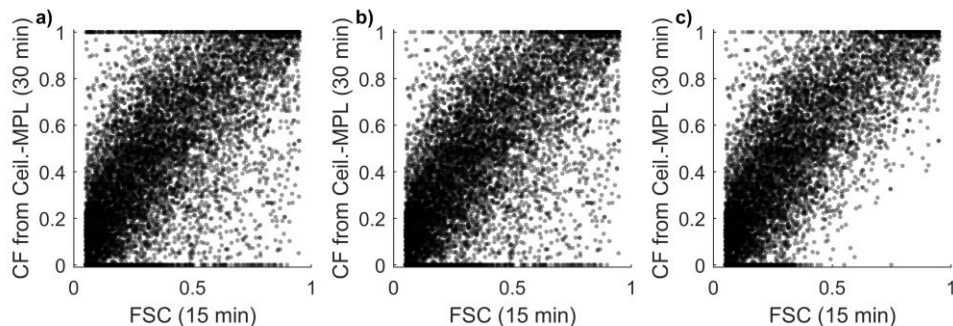
Figure B1. Comparison of 15-min FSC and 30-min CF from the merged ceilometer-MPL method after data screening steps: (a) Single-layer cases: CF (all clouds) – CF (ShCu) < 0.1 and 0.05<FSC < 0.95 (N = 9571), (b) data as in (a), with additional

39

requirement to discard periods with missing CBH returns (N = 9128), (c) data as in (b), with periods when 15-min FSC from thin cloud exceeds 0.3 are removed (N = 8192). Each point represents 15-min period.

40

Tables

Table 1. Variants of cloud fraction (CF) products used in this work: contributing instruments, thresholds on cloud base height (CBH) and cloud top height (CTH).

Instruments used for CF	Abbreviation	CBH thresholds	CTH threshold
Ceil. + MPL[†] + radar	CF_{bt} (bases & tops)	0.3 km < CBH < 3 km	CTH < 4 km & CTH > CBH
Ceilometer + MPL	CF_b (bases only)	0.3 km < CBH < 3 km	-
Ceilometer + MPL	CF_{tot} (all clouds)	CBH > 0	-
Ceilometer	CF_{ceil} (ceil. only)	0.3 m < CBH < 3 km	-

[†]In 2011-2017 period, the MPL is not used to determine cloud top height below 3 km (Sect. A4).

805 Table 2. Total number of observations (N) and summary statistics for cloud amount estimated by five methods for three periods of interest (second-fourth rows) at coarser temporal scale: 60-min FSC, "CF-like" and CF. These statistics include the mean (second column), and root mean-square difference (RMSD) (third column). Values more than 0.05 different from FSC are bold.

Year	N	N_{ceil}	Mean cloud cover					RMSD: FSC vs CF			
			FSC	CF_b	CF_{bt}	CF_{ceil}	CF-like	CF_b	CF_{bt}	CF_{ceil}	CF-like
All	2250	1665	0.33	**0.44**	**0.38**	0.31	0.34	0.19	0.16	0.11	0.1
2000-2010	1259	708	0.34	**0.46**	**0.42**	0.3*	0.34	0.2	0.18	0.11	0.1
2011-2017	991	957	0.33	**0.41**	0.33	0.32	0.33	0.17	0.14	0.12	0.1

Each 60 min period excludes clear and overcast conditions using criteria: 0.05 < FSC <0.95. *Note that the ceilometer data in 2000-2010 had a large percentage of missing during quality control (see Table B1), and that the mean is calculated from a subset of
810 times reflected by N_{ceil}. CF abbreviations defined in Table 1.

Table 3. The same as Table 2, except for fine temporal scale: 15-min FSC and 30-min "CF-like" and CF.

Year	N	N_{ceil}	Mean cloud cover					RMSD: FSC vs CF			
			FSC	CF_b	CF_{bt}	CF_{ceil}	CF-like	CF_b	CF_{bt}	CF_{ceil}	CF-like
All	8192	6070	0.34	**0.45**	**0.39**	0.32	0.35	0.22	0.2	0.14	0.15
2000-2010	4628	2586	0.34	**0.47**	**0.43**	0.3*	0.35	0.24	0.21	0.14	0.15
2011-2017	3564	3484	0.33	**0.43**	0.35	0.33	0.34	0.21	0.18	0.15	0.14

Each 15 min period excludes clear and overcast conditions using criteria: 0.05 < FSC <0.95. *This value is calculated from a much smaller subset of times, N_{ceil}. CF abbreviations defined in Table 1.

815

Table A1. References to instrumentation descriptions and service records at the SGP site:

Instrument (abbreviation)	Reference
Millimeter cloud radar (MMCR, 2000-2010)	https://www.arm.gov/capabilities/instruments/mmcr
Ka-band zenith pointing radar (KAaZR, 2011-present)	https://www.arm.gov/capabilities/instruments/kazr
Micropulse lidar (MPL)	https://www.arm.gov/capabilities/instruments/mpl
Ceilometer (Ceil)	https://www.arm.gov/capabilities/instruments/ceil
915 MHz Radar Wind Profiler (RWP)	https://www.arm.gov/capabilities/instruments/rwp
Total Sky Imager (TSI)	https://www.arm.gov/capabilities/instruments/tsi

820 **Table B1. Summary of single-layer ShCu events by year and percent data completeness after applying quality control steps. Entries correspond to scatterplots in Figure B1. In particular these periods comply to the requirement that 0.05 <FSC < 0.95.**

Year 20xx	'00*	'01	'02	'03	'04	'05†	'06	'07	'08	'09	'10¥	'11	'12	'13	'14	'15	'16	'17	All
Days[a]	12	29	43	28	44	22	36	37	56	43	6	21	33	47	45	30	42	35	609
Hours[a]	52	93	205	124	191	88	135	150	227	155	20	86	98	179	182	112	149	148	2393
QC pval (%)[b]	100	70	100	91	93	98	97	96	98	86	14	100	100	100	100	99	99	100	95
QC thin (%)[c]	91	98	68	88	100	41	100	98	91	86	97	90	96	99	95	79	91	98	90
QC "ceil ok" (%)	91	2	0	0	1	1	99	97	91	86	95	90	94	99	95	76	84	96	65
Final Days[c]	12	20	36	28	41	12	36	36	56	38	3	21	33	47	45	29	42	34	569
Final Hrs[c]	47	63	139	98	176	35	131	141	203	121	2	78	95	178	173	88	135	146	2048

[a] Data in Fig. B1a. [b] Percent of observations that pass a 100% Quality Control (QC) percent valid test for ARSCL / KAZRARSCL; data shown in Fig. B1b. [c] Data shown in Fig. B1c.*TSI data begins July 2. †TSI missing most data in June and July. ¥In 2010, 19 days with shallow cumuli are identified, of which 6 contained events longer than 2 hours that also satisfied single
825 layer conditions: CF(all clouds) – CF (ShCu) < 0.1). Only 14% of the observations satisfied the 100% instrument validity criterion. The values in this table are generated from the 15-min dataset. These values do not correspond to the 1-hr values (i.e. the 1-hr values are not generated from the 15-min values, and are subject to passing the same quality control thresholds, but on the hourly timescale).

830

Table B2. Regression coefficients for Figures 3-6.

Year	*15- min: $CF = \beta_0 + \beta_1 * FSC$*								*1 hr: $CF = \beta_0 + \beta_1 * FSC$*							
	CF-like		**CF$_{ceil}$**		**CF$_b$**		**CF$_{bt}$**		**CF-like**		**CF$_{ceil}$**		**CF$_b$**		**CF$_{bt}$**	
	(β_0	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**	**(β_0**	**β_1)**
All	0.01	0.99	0.01	0.92	0.12	1	0.09	0.9	-0.01	1.06	-0.01	0.96	0.1	1.03	0.07	0.93
2000-2010	0.01	0.99	0	0.91	0.13	1.01	0.13	0.87	-0.01	1.06	-0.02	0.94	0.11	1.04	0.11	0.89
2011-2017	0.01	0.99	0.02	0.93	0.11	0.98	0.04	0.92	-0.01	1.07	0	0.98	0.08	1.02	0.02	0.97

CF abbreviations defined in Table 1.

Table B3. Pearson's Correlation coefficient for FSC vs CF for fine and hourly time scale.

Year	30 min CF, 15-min FSC				hourly CF and FSC			
	CF$_b$	CF$_{bt}$	CF$_{ceil}$	CF-like	CF$_b$	CF$_{bt}$	CF$_{ceil}$	CF-like
All	0.77	0.74	0.83	0.85	0.84	0.81	0.89	0.92
2000-2010	0.78	0.74	0.86*	0.85	0.84	0.8	0.9*	0.92
2011-2017	0.77	0.77	0.82	0.84	0.84	0.83	0.88	0.92

835 ***This value is calculated from a much smaller subset of times (see N$_{ceil}$, Tables 2,3). CF abbreviations defined in Table 1.**