Identifying Insects, Clouds, and Precipitation using Vertically Pointing Polarimetric Radar Doppler Velocity Spectra

Christopher R. Williams[1], Karen L. Johnson[2], Scott E. Giangrande[2], Joseph C. Hardin[3], Ruşen Öktem[4,5], and David M. Romps[4,5]

[1] Ann and H.J. Smead Aerospace Engineering Sciences Department, University of Colorado Boulder, CO, 80309, United States
[2] Brookhaven National Laboratory, Upton, NY, 11973, United States
[3] Pacific Northwest National Laboratory, Richland, WA, 99354, United States
[4] Department of Earth and Planetary Science, University of California, Berkeley, CA, 94720, United States
10 [5] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, CA, 94720, United States

Correspondence to: Christopher R. Williams (Christopher.Williams@colorado.edu)

Abstract. This study presents a method to identify and distinguish insects, clouds, and precipitation in 35 GHz (Ka-band) vertically pointing polarimetric radar Doppler velocity power spectra and then produce masks indicating the occurrence of hydrometeors (i.e., clouds or precipitation) and insects at each range gate. The polarimetric radar used in this study transmits a linear polarized 15 wave and receives signals in collinear (CoPol) and cross-linear (XPol) polarized channels. The insect-hydrometeor discrimination method uses CoPol and XPol spectral information in two separate algorithms with their spectral results merged and then filtered into single value products at each range gate. The first algorithm discriminates between insects and clouds in the CoPol Doppler velocity power spectra based on the spectra texture, or spectra roughness, which varies due to the scattering characteristics of insects versus cloud particles. The second algorithm distinguishes insects from raindrops and ice particles by exploiting the larger 20 Doppler velocity spectra linear depolarization ratio (LDR) produced by asymmetric insects. Since XPol power return is always less than CoPol power return for the same target (i.e., insect or hydrometeor), fewer insects and hydrometeors are detected in the LDR algorithm than the CoPol algorithm, which drives this need for a CoPol based algorithm. After performing both CoPol and LDR detection algorithms, regions of insect and hydrometeor scattering from both algorithms are combined in the Doppler velocity spectra domain and then filtered to produce a binary hydrometeor mask indicating the occurrence of cloud, raindrops, or ice 25 particles at each range gate. Comparison with a collocated ceilometer indicates that hydrometeor mask column bottoms are within +/-100 meters of simultaneous ceilometer cloud base heights. Forty-seven (47) summer-time days were processed with the insect-hydrometeor discrimination method using U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) program Ka-band zenith pointing radar observations in northern Oklahoma (USA). All datasets and images are available on public repositories.

30

1. Introduction

The vertical structure of non-precipitating clouds plays an important role in the Earth's radiation balance. These clouds absorb longwave radiation emitted from the surface and reflect shortwave solar radiation back into space (Cess et al., 1990). The proportion of these two processes determines whether these clouds act as a net radiation sink or source in the Earth's radiation budget (Ramanathan et al., 1989). Vertically pointing cloud radars have been used for decades to quantify the extent to which non-precipitating clouds can be used as inputs to Earth radiation budget studies to understand cloud dynamics and cloud life cycles (Moran et al., 1998; Ackerman and Stokes, 2003).

In addition to measuring cloud properties, cloud radars are sensitive enough to detect individual insects within the radar volume (for overviews, see Riley, 1989, Westbrook et al., 2014; Nansen and Elliot, 2016). The field of radar entomology exploits this sensitivity by pointing polarimetric radar beams a few degrees off vertical and rotating the beam 360 degrees in azimuth to estimate insect population and migration direction (Drake et al., 2020). The field of radar meteorology has used polarimetric scanning radar observations to track insect flying direction and altitude outside of clouds (Mueller and Larkin, 1985) and to estimate gust-front motions ahead of convective cells because insects and small pieces of vegetation act as radar reflectors trapped within the strong boundary layer outflow (Klingle el al., 1987). Insects are considered clutter in vertically pointing cloud radar observations and the polarimetric signal from asymmetric insects is a good indicator of insect scattering (Lohmeier et al., 1997; Khandwalla et al., 2001 and 2002; Martner and Moran, 2001). Due to their large power fluctuations in the velocity spectra, insects have also been detected in cloud radar Doppler velocity power spectra (Luke et al., 2008).

Identifying and removing radar scattering from insects and other sources of "atmospheric plankton" (Lhermitte, 1966) has been a known problem in developing operational cloud products (Kollias et al., 2016). The U. S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) program merges observations from multiple sensors (including radars, lidars, and ceilometers) to produce an estimate of hydrometeors (i.e., cloud particles, raindrops, ice particles) in the vertical column, called the Active Remote Sensing of CLouds (ARSCL) product (Clothiaux et al., 2000). ARSCL is a high temporal (~4 s) and vertical (~30 m) resolution operational product that is being used to validate cloud resolving model and climate model simulations including shallow cloud simulations produced from the Large Eddy Simulation (LES) ARM Symbiotic Simulation and Observation (LASSO) workflow (Gustafson et al., 2017).

One of the most difficult aspects of the operational ARSCL processing is identifying and removing insect clutter from the Ka-band ARM Zenith pointing Radar (KAZR) observations. Figure 1 shows one hour of ARSCL processed reflectivity from 31-July-2018 (ARM user facility, 2014). ARSCL classifies each radar volume as either hydrometeor-only, clutter-only, or some combination of hydrometeor-plus-clutter. Figure 1a (top panel) shows ARSCL reflectivity for radar volumes classified as either hydrometeor-only or hydrometeor-plus-clutter. Thus, these radar volumes contain hydrometeors with some radar volumes also containing clutter. To explore which radar volumes are clutter-free and can be used to validate cloud model simulations, Fig. 1b (bottom panel) shows the hydrometeor-only radar volumes. The black symbols represent ceilometer-derived cloud base, which is also used in the ARSCL operational algorithm. One interpretation of Fig. 1b, is that drizzle is falling below cloud base at about 1.5 km. As will be shown in Section 3, these fall streaks below cloud base are actually insect clutter. Due to ARSCL misclassifying insect clutter as hydrometeor scattering as exemplified in Fig. 1, this study was initiated to develop a KAZR hydrometeor mask identifying locations of clouds and precipitation in the vertical column to help validate shallow cloud simulations produced by the ARM LASSO workflow (Gustafson et al., 2017) and to provide another input to the ARSCL operational processing chain to help identify and remove insect clutter from future ARSCL products.

40
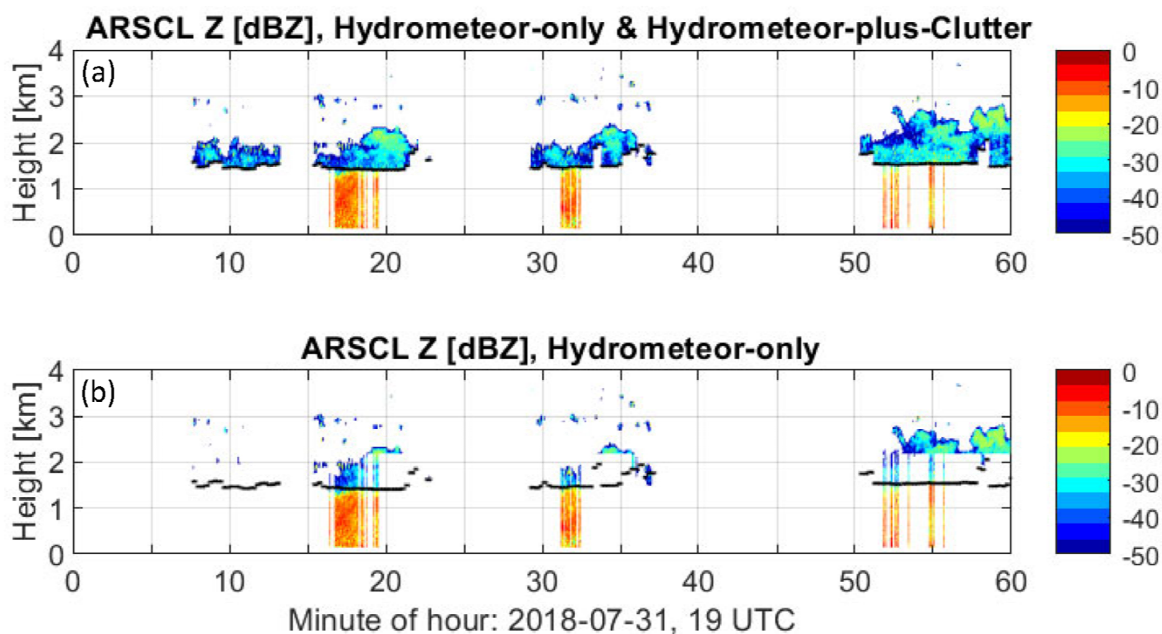
Figure 1: Active Remote Sensing of CLouds product (ARSCL) for hour 19 UTC (hour 14 local) on 31-July-2018. (a) ARSCL reflectivity for radar volumes ARSCL classified as either hydrometeor-only or hydrometeor-plus-clutter. (b) ARSCL reflectivity for radar volumes ARSCL classified as hydrometeor-only. Black symbols in both panels are ceilometer derived cloud base stored in ARSCL product. Note the hydrometeor misclassification below ceilometer cloud base in (b) motivates the need for improved insect clutter detection.

The method to identify insects and hydrometeors presented herein builds on prior work using polarimetric diversity and Doppler velocity power spectra variability (e.g., Martner and Moran, 2001; Luke et al., 2008). The observations used in this study and the signatures of insect and hydrometeor scattering are discussed in Sections 2 and 3. Section 4 presents the main concept behind the algorithms developed in this study, specifically, that by examining small regions of Doppler velocity power spectra over a limited height range (typically a few velocity bins in width and three range gates in height), it is possible to assess whether the power in that small spectra-height region resulted from insect or hydrometeor scattering. Once all small spectra-height regions have been classified by texture and polarimetric algorithms, low-pass filters are used to identify larger spectra-height regions of hydrometeor scattering. If the hydrometeor regions are large enough to be detected as a primary hydrometeor peak in operational Doppler velocity spectra processing routines, then the hydrometeor mask is set to affirmative for that range gate. As a bonus output product, insect activity is estimated by counting the number of Doppler velocity bins with insect scattering at each range gate. Three hydrometeor mask products are produced with different time-height filtering to allow the end-user the ability to select outputs for their particular application. The KAZR hydrometeor masks are compared with lidar and camera observations in Section 5 with conclusions discussed in Section 6. The online Supplemental Material Section contains images of insect and hydrometeor classifications for forty-seven (47) summer-time days in northern Oklahoma, U.S.A., identified as LASSO cloud simulation events (LASSO, 2020).

2. Observations

The observations used in this study were collected by the U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) program at their Southern Great Plains (SGP) central facility located in northern Oklahoma. Vertically pointing Ka-band radar co-polarized (CoPol) and cross-polarized (XPol) Doppler velocity power spectra are processed to identify insects, clouds,
5   and precipitation in the vertical column. Verification of those classifications are based on observations from co-located lidar, ceilometer, Total Sky Imager (TSI), and cloud boundaries contained in the Clouds Optically Gridded by Stereo (COGS) product (Romps and Öktem, 2018).

2.1 Ka-band ARM Zenith Pointing Radar (KAZR)

10 The DOE ARM program deploys atmospheric observing systems to characterize the radiative properties of clouds in the atmosphere (Mather and Voyles, 2013). One of ARM's hallmark instruments is the Ka-band (35 GHz) ARM zenith pointing cloud radar (KAZR), which transmits linear polarized waves that are detected simultaneously with collinear polarized (CoPol) and cross-linear polarized (XPol) receivers. The received signals are processed to yield co-polarized $S_{signal}^{CoPol}(v_i, h_j)$ [Watts] and cross-polarized $S_{signal}^{XPol}(v_i, h_j)$ [Watts] Doppler velocity power spectra at each velocity bin v_i and range gate h_j. The linear depolarization
15 ratio spectra profile $S_{dB}^{LDR}(v_i, h_j)$ [dB] is the ratio of polarized signal magnitudes defined as

$$S_{dB}^{LDR}(v_i, h_j) = 10log\left[\frac{S_{signal}^{Xpol}(v_i, h_j)}{S_{signal}^{CoPol}(v_i, h_j)}\right], \text{ or as} \tag{1a}$$

$$S_{dB}^{LDR}(v_i, h_j) = S_{signal,dB}^{XPol}(v_i, h_j) - S_{signal,dB}^{CoPol}(v_i, h_j) \tag{1b}$$

where $S_{signal,dB}^{XPol}(v_i, h_j)$ and $S_{signal,dB}^{CoPol}(v_i, h_j)$ are expressed in decibel units [dB] using $X_{dB} = 10log[X]$. The linear depolarization ratio LDR [dB] is the integration of Xpol and CoPol signals over the spectrum and is defined as

$$LDR(h_j) = 10log\left(\frac{\sum_{v_{min}}^{v_{max}}\left[\frac{S_{signal}^{Xpol}(v_i, h_j)}{S_{signal}^{CoPol}(v_i, h_j)}\right]\Delta v}{\sum_{v_{min}}^{v_{max}}\Delta v}\right), \tag{2}$$

where v_{min} to v_{max} define the velocity range of valid $S_{signal}^{CoPol}(v_i, h_j)$ and $S_{signal}^{XPol}(v_i, h_j)$ observations.

At SGP, KAZR operates in the general 'GE' and medium 'MD' sensitivity modes to sense clouds at different altitudes with operating parameters during 2018 and 2019 shown in Table 1 (ARM user facility, 2011a, 2011b, 2011c, 2011d). Insects are detected in all KAZR operating modes. Figures and algorithm descriptions use the MD mode with retrievals for the GE mode
25 available in the ARM data archive. Since the MD mode uses a long coded transmitted pulse, the first resolved range gate is 570 m above the radar. The KAZR 0.2° antenna beamwidth produces a 3.6 m diameter horizontal disk field-of-view at 1 km range. With a 30 m range resolution, the radar volume at 1 km is approximately 300 m³. To save computer hard disk space, the KAZR CoPol and XPol Doppler velocity power spectra are saved only at range gates with significant power above a noise threshold.

2.2 Validation Observations

30 Four different observational datasets are used to validate the derived KAZR insect and hydrometeor classifications. Specifically, qualitative observations included the backscattered attenuated power from vertically pointing Doppler lidar (ARM user facility, 2010a) and photos from the TSI (ARM user facility, 2000). These observations provide context for cloud bottom and cloud type (neither instrument observe cloud top). Quantitative comparisons come from ceilometer cloud base estimates from a Vaisala 25K ceilometer (ARM user facility, 2010b) and from cloud bottom and top estimates from the COGS product (ARM user facility,
35 2017). The ceilometer is an automated system estimating cloud base with 10-m vertical and 16-s temporal resolution. The COGS

cloud boundaries are derived from three pairs of stereo cameras positioned around the SGP Central Facility and represent cloud boundaries over a cubic domain 6 km to a side (Romps and Öktem, 2018). Due to camera visual occlusion during precipitation, COGS cloud boundaries are only estimated for cases of shallow cumulus clouds, which allow the three cameras to view the vertical extent of each cloud. Likewise, estimates from COGS are only available during daylight hours.

Table 1. Operating parameters for KAZR deployed at ARM Southern Great Plains (SGP) during 2018 and 2019. Operating modes included General Purpose (GE) and Medium Sensitivity (MD) modes. Tabulated parameters include: pulse repetition frequency (PRF) [Hz], inter-pulse period (IPP) [μsec], number of points in FFT (N_{FFT}), number of averaged spectra (also known as number of incoherent integrations) (N_{incoh}), Nyquist velocity ($V_{Nyquist}$) [m s⁻¹], velocity resolution (Δv) [m s⁻¹], range to first range gate [m], range resolution [m], time-on target (which is calculated using $IPP\ N_{FFT}\ N_{incoh}$) [s], and time between samples [s].

Parameter	2018 and 2019	
Sensitivity Mode	general 'GE'	medium 'MD'
Pulse Repetition Frequency (PRF) [Hz]	2777	2777
Inter-Pulse Period (IPP) [μsec]	360	360
Number of points in FFT (N_{FFT})	256	256
Number of incoherent integrations N_{incoh}	18	20
$V_{Nyquist}$ [m s⁻¹]	5.97	5.97
Δv [cm s⁻¹]	4.67	4.67
Range to first range gate R_1 [m]	100	570
Range resolution ΔR [m]	30	30
Time-on Target $t_{target} = IPP\ N_{FFT}\ N_{incoh}$ [s]	1.8	1.8
Time between samples t_{sample} [s]	3.7	3.7

3. Insect, Cloud Droplet, and Precipitation Spectral Characteristics

This section discusses the scattering characteristics of insects, clouds, and precipitation as observed in KAZR CoPol and XPol Doppler velocity power spectra. The first subsection discusses characteristics when it is not raining and the radar is observing insects and shallow cumulus clouds. The second subsection describes the characteristics when insects and raindrops occur simultaneously in the radar volume.

3.1 Insects and Shallow Cumulus Clouds

Figure 2 shows an hour of KAZR observations when both insects and shallow cumulus clouds are observed over the radar during 1900 UTC (1400 Local Time) on 31-July-2018. From top to the bottom, Fig. 2 shows KAZR (a) CoPol reflectivity [dBZ], (b) mean Doppler velocity [m s⁻¹], (c) Doppler velocity spectrum width [m s⁻¹], (d) linear depolarization ratio (LDR) [dB], and (e) KAZR CoPol reflectivity at time-height locations (also called 'pixels' in this study) that do not have a LDR measurement. The black symbols in each panel indicate ceilometer-derived cloud base height, which is near 1.5 km for this hour. Below cloud base, reflectivity (Fig. 2a) and spectrum width (Fig. 2c) have a coherent pattern, but vertical motion (Fig. 2b) appears more variable. If drizzle or rain were below cloud base, then all three quantities would be coherent with downward motions increasing as reflectivity and spectrum width increase (Williams and Gage, 2009). Thus, it is not raining below cloud base. Above the ceilometer derived

cloud base height, there are CoPol reflectivity observations (Fig. 2a), but not as many LDR estimates (Fig. 2d). For example, near minute 20, there is an enhancement of CoPol reflectivity above the ceilometer cloud base and extending above 2 km, yet, there are very few LDR observations in this time-height region. Since LDR requires both CoPol and XPol reflectivity observations, the lack of LDR above cloud base indicates that the XPol channel is not detecting cloud particles. The abrupt omission of LDR observations

5   above the ceilometer cloud base height appears suspicious as it produces a nearly horizontal feature in Fig. 2d. This CoPol vs. XPol sensitivity is illustrated in the bottom panel which shows CoPol reflectivity for all pixels that do not also have a LDR observation. The continuous time-height CoPol reflectivity observations above 1.5 km are cloud features that are easily discernible by eye. Return signals from individual insects appear as speckles up to 4 km in all panels.

The CoPol and XPol Doppler velocity power spectra produced by individual insects and by cloud droplet distributions

10 have different characteristics as illustrated in Fig. 3, which shows CoPol (Fig. 3a) and XPol (Fig. 3b) Doppler velocity power spectral density profiles at 19:19:02 UTC on 31-July-2018. The vertical axis extends from 0 to 3 km in height and the horizontal axis extends +/- 6 m s^{-1} radial velocities. The Nyquist velocity is 5.95 m s^{-1} and downward motions have positive values consistent with positive raindrop diameters having positive fall speeds due to gravity. Due to the long coded transmitted pulse, the first observations occur at 0.57 km range. The colors represent the return signal power expressed in dB with warmer colors indicating

15 larger return signal power. The mean noise power is approximately -100 dB.

Figure 3c shows CoPol Doppler velocity power spectra at 1 and 2 km heights (black and red lines, respectively). The power spectrum at 1 km has more variability between velocity bins compared to the spectrum at 2 km. This variability is because the radar is detecting individual insects within the 300 m^3 field-of-view with each insect moving at its own radial velocity. If an insect is the only insect moving at a particular velocity, the spectrum will have an isolated peak (e.g., near -1.7 m s^{-1} radial velocity

20 in Fig. 3c). If multiple insects are moving at similar speeds, the spectrum will be broader, yet, will still have variability. For example, between -1 and +3 m s^{-1} radial velocities, the 1 km height spectrum (black line) is both elevated in magnitude and has more bin-to-bin variability than the spectrum from 2 km (red line). Also, the backscattered power from insects is primarily confined to one range gate with some power leaking into neighboring range gates due to radar signal processing limitations, which produce point enhancements in the spectra profile. Shown in sequential spectra profiles in the Supplemental Material, point enhancements

25 often appear in only one spectra profile and not in neighboring profiles separated 4 seconds apart. This indicates that individual insects appear in the 3.6 m diameter by 30 m field-of-view for less than 4 seconds.

In contrast to individual insects, clouds and precipitation are distributed targets filling the radar volume with hundreds or thousands of hydrometeors of different sizes with different radial velocities. Since the number of particles in the hydrometeor size distribution varies gradually over neighboring particle sizes and the hydrometeor spectrum is extended in the velocity dimension

30 due to antenna broadening effects, the return power spectrum has a gradual change over neighboring velocity bins. Thus, the power spectrum from a distribution of many hydrometeors is smoother than the return from a few individual insects. The smoother power spectra at 2 km height shown in Fig. 3c are consistent with a distribution of small cloud droplets moving at different velocities within the radar volume. In addition to smooth power spectra across the velocity dimension, power spectra from cloud droplets are also more continuous in range due to the vertical extent of clouds as can be seen with a continuity of clouds with height in Fig. 3a.

35
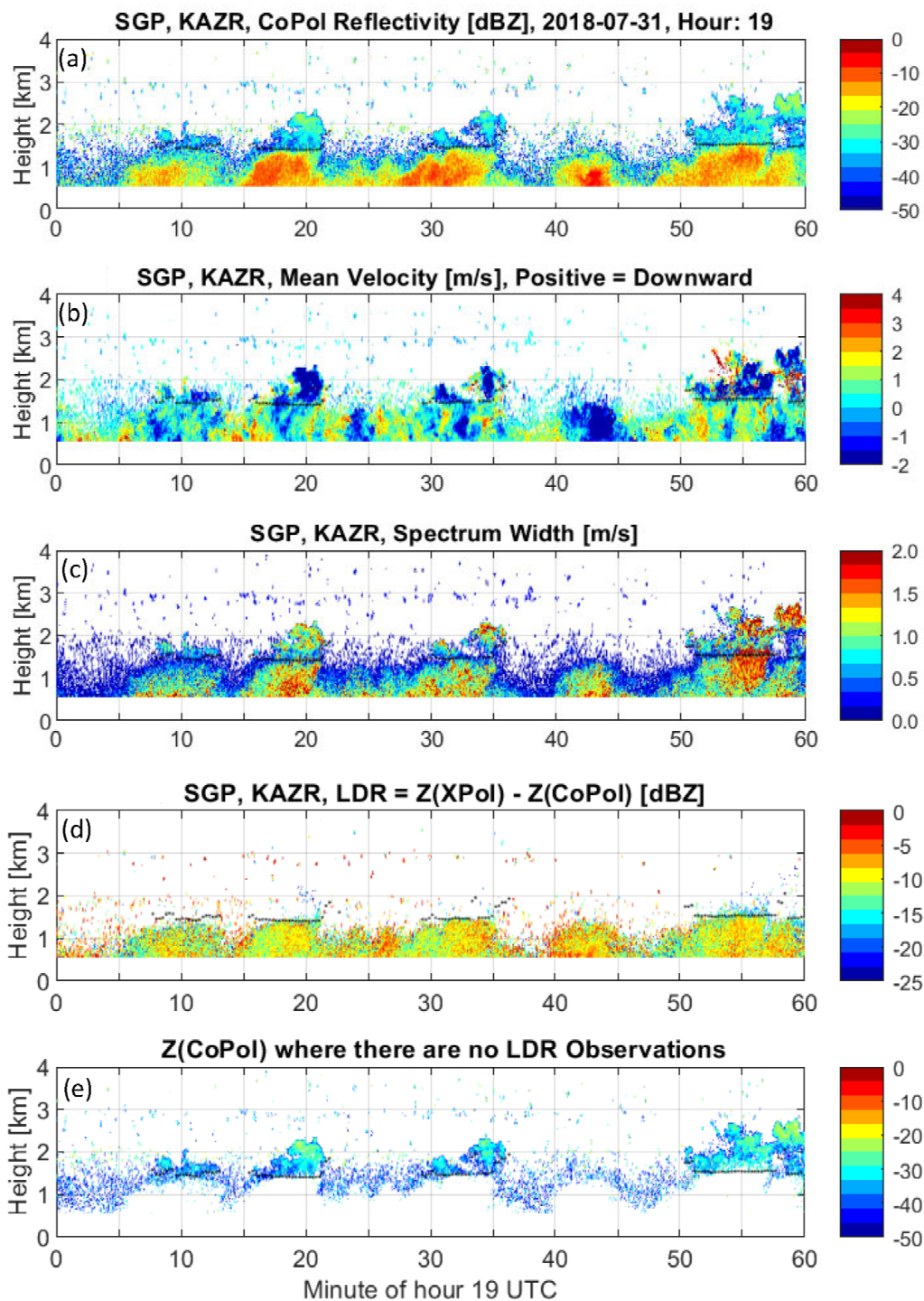
Figure 2: Moments calculated from raw spectra for hour 19 UTC on 31-July-2018. (a) CoPol reflectivity [dBZ]. (b) Mean radial velocity [m s⁻¹], positive values are downward motion. (c) Spectrum width [m s⁻¹]. (d) Linear depolarization ration (LDR) [dB]. (e) CoPol reflectivity [dBZ] at pixels that do not have an LDR measurement. Black symbols in both panels are ceilometer derived cloud base.
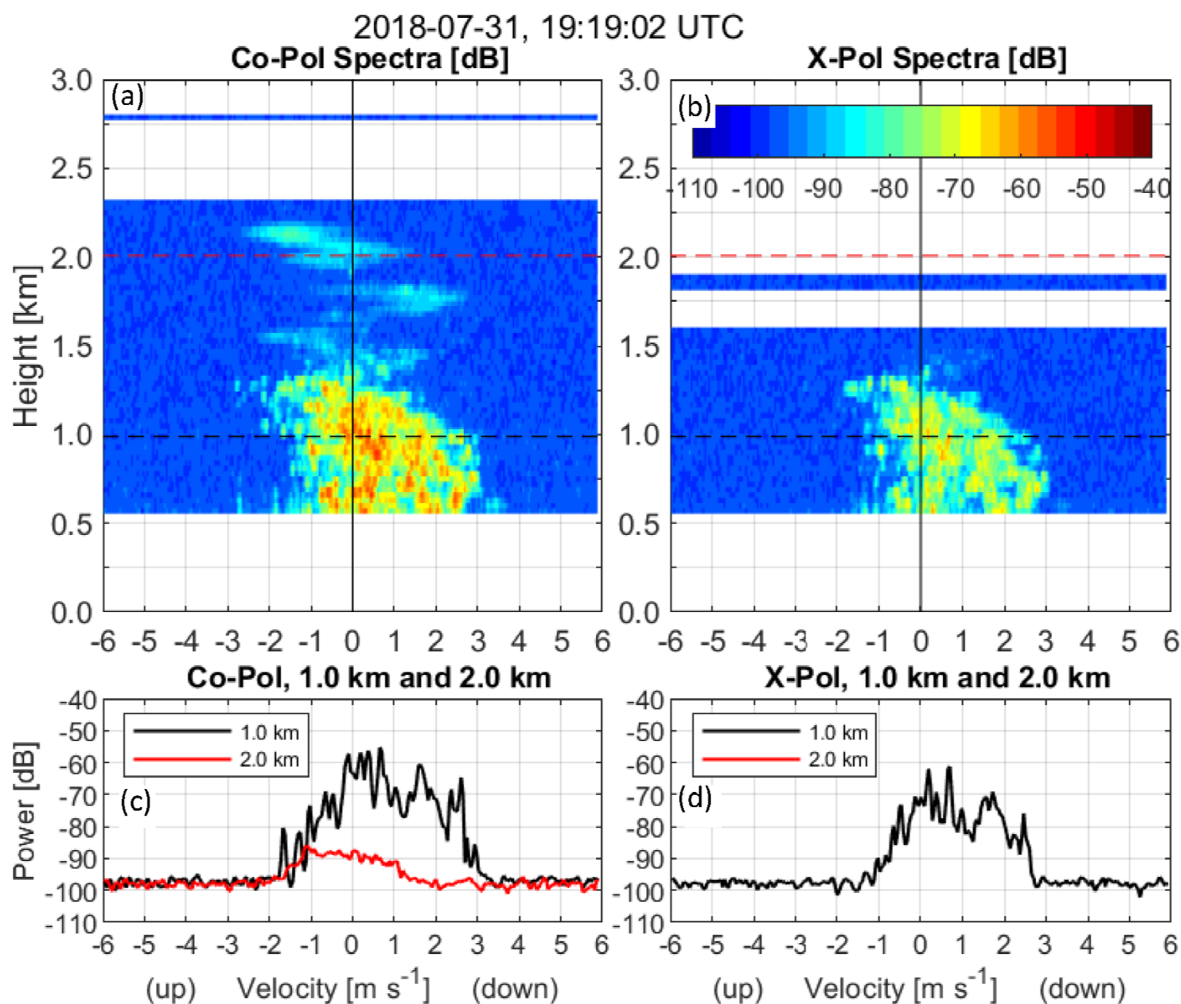
Figure 3: Spectra from profile at 19:19:02 UTC on 31-July-2018. (a) CoPol Doppler velocity power spectra [dB] as a function of range and radial velocity. (b) XPol Doppler velocity power spectra [dB] as a function of range and radial velocity. (c) CoPol Doppler velocity power spectra at 1.0 km (black line) and 2.0 km (red line). (d) XPol Doppler velocity power spectra at 1.0 km (black line).

5 *3.2 Insects and Precipitation*

Figure 4 shows time-height cross-sections of KAZR CoPol reflectivity (Fig. 4a) and LDR observations (Fig. 4b) when insects, clouds, and precipitation are observed in the same hour. Observations were collected during 0400 UTC (2300 Local) on 04-April-2019. From minutes 0-to-20, the approximate 1.5 km ceilometer cloud base height (black symbols) is above the insect layer that has LDR values between approximately -5 and -10 dB (see Fig. 4b), while the CoPol reflectivity is continuous in time and height

10 above the ceilometer cloud base height (see Fig. 4a and 4c). At the beginning of the hour, the CoPol reflectivity (Fig. 4a) time-height structure indicates a precipitating cloud system between 3 and 5 km that evolves in time with precipitation reaching the lowest resolved height of 0.57 km after minute 20. The LDR shows a similar time-height structure (with reduced vertical depth) with LDR values ranging between -25 to -20 dB. The LDR enhancement near 2.4 km and after minute 20 is due to a mixture of liquid and frozen particles near the melting layer (Baldini and Gorgucci, 2006). Below the melting layer, the LDR has values near

15 -25 dB that is due to scattering from rain drops. Above the melting layer, scattering from asymmetrical ice particles leads to LDR

values near -20 dB (Baldini and Gorgucci, 2006). In contrast to the shallow cumulus cloud droplet observations in Figs. 2 and 3, KAZR has enough sensitivity to detect XPol signal returns from large spherical raindrops and ice particles.
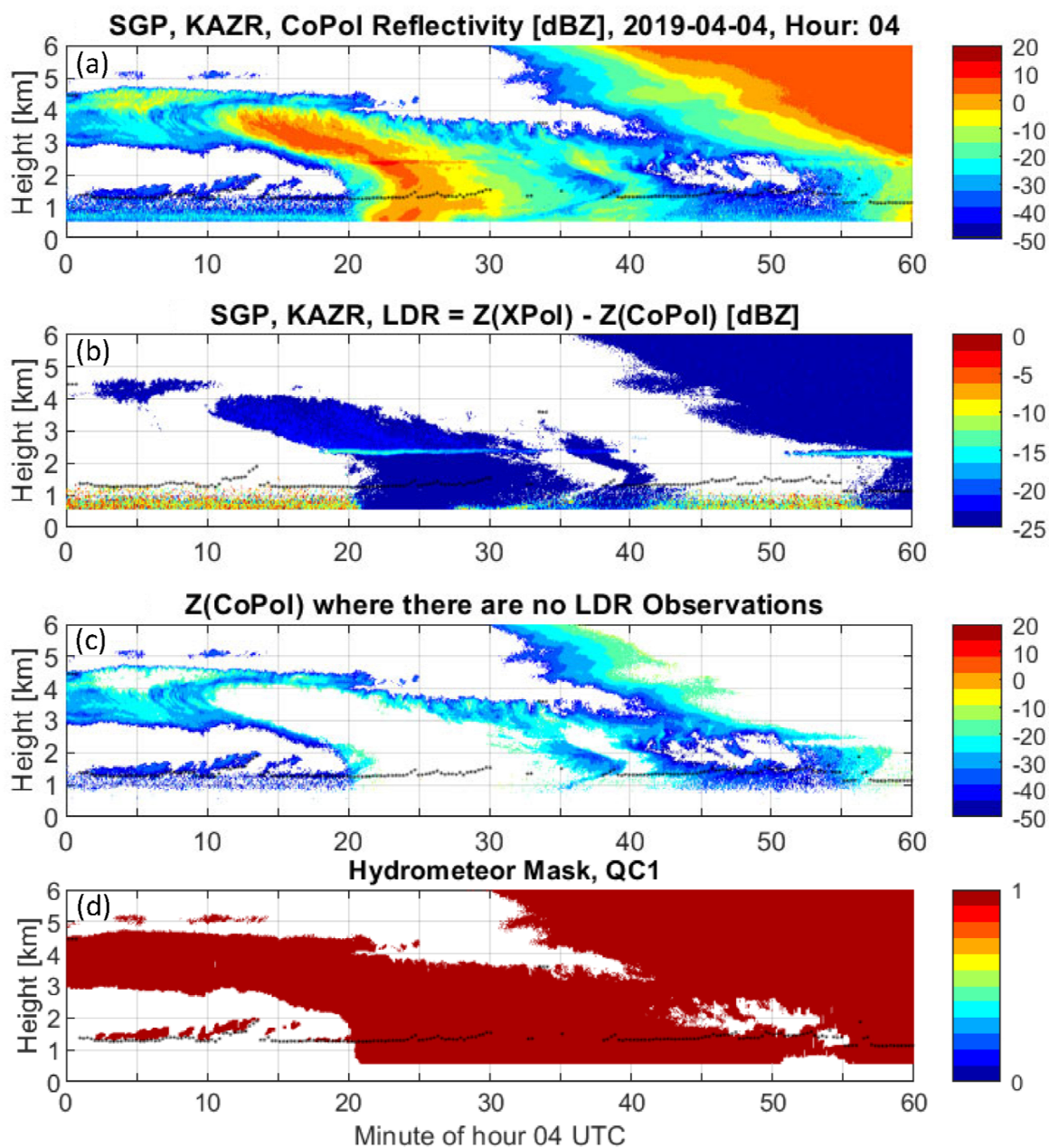
Figure 4: Moments calculated from raw spectra and retrieved hydrometeor QC1 mask for hour 4 UTC on 4-April-2019. (a) CoPol reflectivity [dBZ]. (b) Linear depolarization ration (LDR) [dB]. (c) CoPol reflectivity [dBZ] at pixels that do not have an LDR measurement. (d) Retrieved hydrometeor QC1 mask. Black symbols in both panels are ceilometer derived cloud base.

Figure 4c shows the CoPol reflectivity at time-height pixels that do not have a LDR measurement. As with the shallow cloud observations (see Fig. 2e), there are more CoPol observations than LDR observations. By eye, regions of insects, clouds, and precipitation can be identified in CoPol reflectivities in Fig. 4c. The bottom panel (Fig. 3d) shows the QC1 hydrometeor mask produced by the insect-hydrometeor detection algorithm described later in this manuscript. The mask identifies the shallow clouds

5 near 1.5 km from 0-to-10 minutes and the precipitating anvil at the beginning of the hour between 3-to-4 km that descends to the lowest range gate just after minute 20. Surprisingly, the hydrometeor mask is affirmative below 1.5 km starting at about minute 21 and continues until the end of the hour except for a shallow gap between minutes 50-to-55. Spectra will be investigated in Section 4 and will confirm that hydrometeors are present in this time-height interval and the affirmative mask is correct. The events shown in Figs. 2-4 highlight three important attributes of CoPol and LDR measurements:

10 • LDR measurements detect some, but not all, insect and precipitation observations.

• KAZR LDR measurements do not have the sensitivity to detect shallow non-precipitating liquid clouds.

• Doppler velocity spectral features need to be examined to discriminate between insects, clouds, and precipitation.

These LDR attributes indicate that CoPol spectra must be used in addition to LDR measurements to discriminate between insects, clouds, and precipitation.

15 **4. Anatomy of Insect-Hydrometeor Detection Algorithms**

As described previously, the radar returned signal results from scattering from insects (including "atmospheric plankton") and/or hydrometeors (aka, cloud droplets or precipitation sized particles). The insect-hydrometeor detection algorithms described in this section aim to classify each region of the CoPol and LDR Doppler velocity spectra as either insect or hydrometeor scattering. Next, the two CoPol and XPol regional spectral classifications are combined and then filtered to produce masks indicating the occurrence

20 of insect or hydrometeor scattering at every range gate.

The detection algorithms start with the observed CoPol and XPol spectra profiles $S_{obs}^{CoPol/XPol}(v_i, h_j)$ [Watts]. These are a combination of signal power $S_{signal}^{CoPol/XPol}(v_i, h_j)$ [Watts] and random noise power $n(v_i, h_j)$ [Watts]

$$S_{obs}^{CoPol}(v_i, h_j) = S_{signal}^{CoPol}(v_i, h_j) + n(v_i, h_j) \tag{3a}$$

and

25 $$S_{obs}^{XPol}(v_i, h_j) = S_{signal}^{XPol}(v_i, h_j) + n(v_i, h_j) . \tag{3b}$$

The signal powers are a combination of insect signal power $S_{insect}^{CoPol/XPol}(v_i, h_j)$ [Watts] and hydrometeor signal power $S_{hydro}^{CoPol/XPol}(v_i, h_j)$ [Watts] for both polarizations. This can be expressed as

$$S_{signal}^{CoPol}(v_i, h_j) = S_{insect}^{CoPol}(v_i, h_j) + S_{hydro}^{CoPol}(v_i, h_j) \text{ and} \tag{4a}$$

$$S_{signal}^{XPol}(v_i, h_j) = S_{insect}^{XPol}(v_i, h_j) + S_{hydro}^{XPol}(v_i, h_j) . \tag{4b}$$

30 The goal of the insect-hydrometeor detection algorithms is to identify the insect and hydrometeor scattering contributions that produce the observed signal power spectra profile. In the discussed algorithms, it is assumed that the power in any (v_i, h_i) location is due to either insect or hydrometeor scattering, and not both.

The observed KAZR CoPol and XPol spectra profiles (Fig. 3) are the inputs to the insect-hydrometeor algorithms, with the processing steps for both algorithms outlined in Fig. 5. The methodology consists of two parallel algorithms. The *CoPol Texture*

35 *Algorithm* classifies insects and hydrometeors based on the CoPol spectra texture, with the understanding that scattering from insects produces more spectrum variability than cloud droplet or raindrop distributions. The *LDR algorithm* classifies insects and hydrometeors based on the understanding that asymmetric insects produce larger LDR magnitudes than spherical raindrops (when

viewed from the bottom), and that the KAZR XPol channel is not sensitive enough to detect non-precipitating liquid cloud droplets. Both algorithms produce insect-hydrometeor membership classes for every region of the spectra profile. The membership classes are combined and then reduced to binary insect and hydrometeor masks that have affirmative values for insect or hydrometeor scattering at each range gate. After processing individual spectra profiles, two time-height continuity quality control filters are

5 applied to the hydrometeor masks to remove outliers. This is based on the understanding that clouds and precipitation are persistent over 10's of seconds and 10's of meters. Details of each algorithm module are described in the following sections.
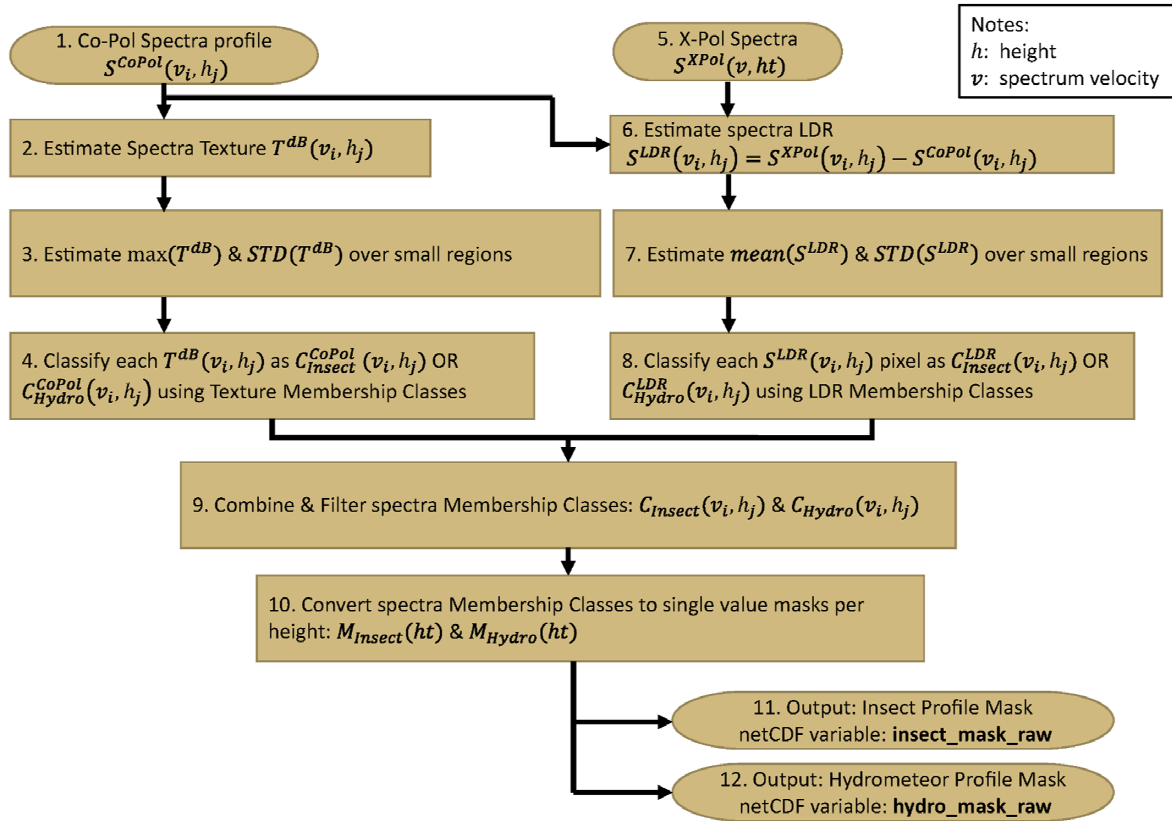
Figure 5: Retrieval logical flow diagram

4.1 CoPol Texture Algorithm

10 This section describes the *CoPol Texture Algorithm* by applying the processing steps (Boxes 1-4 of Fig. 5) to the observed spectra profile shown in Fig. 3a. An objective noise threshold $n_{HS}(h_j)$ is estimated from the CoPol spectra at each height (Hildebrand and Sekhon 1974; Carter et al. 1995). The CoPol spectra with magnitudes greater than $n_{HS}(h_j)$ are defined as signal power (see equation 3). The CoPol signal power for the boundary layer spectra in Fig. 3a is shown in Fig. 6a. As discussed before, insect scattering produces delta functions in the power spectra that are broadened in the velocity domain because of hardware limitations

15 (e.g., antenna beamwidth) and signal processing techniques (e.g., FFT processing). A texture parameter $T^{dB}(v_i, h_j)$ [dB] (Box 2 of Fig. 5) captures delta function variability in the CoPol power spectra, and is defined as

$$T^{dB}(v_i, h_j) = max \begin{bmatrix} \left| S^{CoPol}_{signal,dB}(v_i) - S^{CoPol}_{signal,dB}(v_{i-1}) \right| \\ \left| S^{CoPol}_{signal,dB}(v_i) - S^{CoPol}_{signal,dB}(v_{i+1}) \right| \end{bmatrix} \qquad . \tag{5}$$

To capture both positive and negative changes equally, $T^{dB}(v_i, h_j)$ uses the absolute magnitude, then selects the largest difference between the neighbors (i.e., v_{i-1} or v_{i+1}). Figure 6b shows the texture $T^{dB}(v_i, h_j)$ for the CoPol power spectra shown in Fig. 6a. Note that the small magnitude texture values in the upper heights are due to cloud droplet scattering and larger magnitude texture values in the lower heights are caused by insect scattering. Several features make texture $T^{dB}(v_i, h_j)$ well suited for identifying

5   insect produced delta function variability. First, signal power is expressed in decibel units to remove signal magnitude dependencies that occur between cloud droplet (order of 10 dB) and raindrop (order of 50 dB) observations. Second, a narrow KAZR antenna beamwidth allows the difference between nearest neighbors (i.e., v_i and v_{i+1}) to quantify delta functions. (Note that for radars with broader beamwidths, the insect peak would broaden, and power differences using further neighbors may be necessary in order to identify delta functions, e.g., between v_i and $v_{i\pm2}$.)
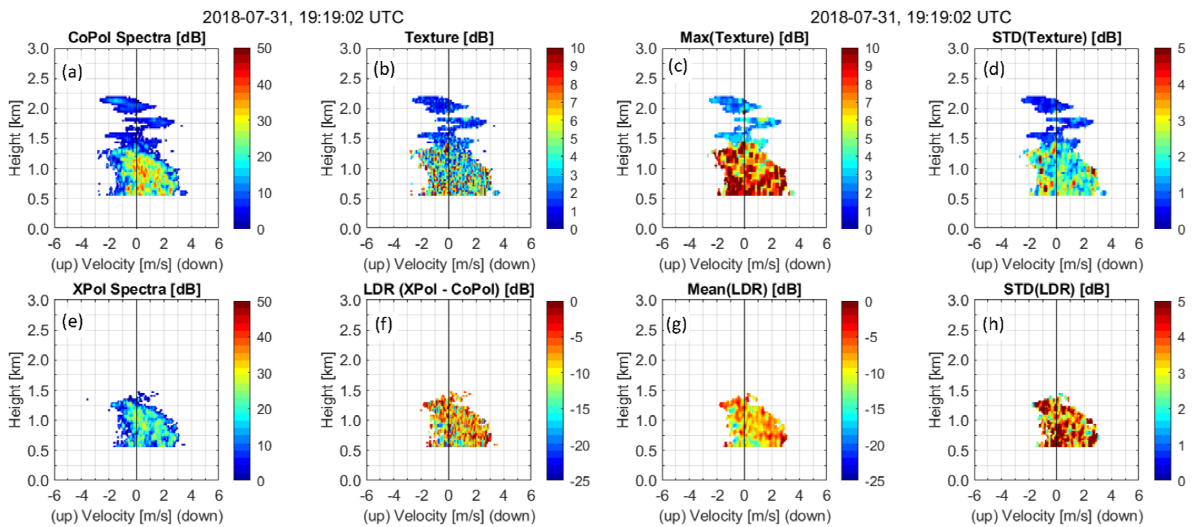
Figure 6: Spectra profile measurements and calculations from profile at 19:19:02 UTC on 31-July-2018. (a) CoPol spectra. (b) CoPol Texture. (c) Max(Texture). (d) STD(Texture). (e) XPol spectra. (f) LDR spectra. (g) mean(LDR). (h) STD(LDR). All measurements and calculations are in units of dB.

With a goal of identifying regions of insect and hydrometeor scattering, a small velocity-height window is moved

15   throughout the $T^{dB}(v_i, h_j)$ domain and texture statistics are calculated for each small region. For this KAZR dataset, a velocity-height window of five (5) velocity bins (total width of 0.186 m s⁻¹) and three (3) range gates (total depth of 90 m) was large enough to capture regional texture variability. For each small region, maximum texture $T_{max}^{dB} = \max[T^{dB}(v_{i\pm2}, h_{j\pm1})]$ and standard deviation $T_{STD}^{dB} = STD[T^{dB}(v_{i\pm2}, h_{j\pm1})]$ are estimated and assigned to the location (v_i, h_j). Figures 6c and 6d show the regional maximum and standard deviation for the texture shown in Fig. 6a. Note that both quantities are larger at lower altitudes where

20   insect scattering dominates compared to higher altitudes that are dominated by cloud droplet scattering. Interestingly, enhancements in both max texture and STD texture are visible near 1.8 and 2 km indicating that insect scattering is occurring within cloud scattering regions.

With an objective of separating insect and cloud scattering regions based on CoPol texture statistics, Fig. 7 shows 1-dimensional (1D) and 2-dimensional (2D) probability distribution functions (PDFs) of $T_{STD}^{dB} = STD[T^{dB}]$ and $T_{max}^{dB} = \max[T^{dB}]$

25   for all spectral regions that do not have a LDR measurement. To increase the number of samples in the PDFs, Fig. 7 uses spectra texture statistics for all profiles in hour 19 of 31-July-2018. The color coding in the 2D plot represents the percent occurrence relative to the cell with maximum number of observations. The 1D PDFs produced from the observations are shown with black

curves in Figs. 7a and 7c using 953,136 samples, each representing a small spectral region, distributed into two populations. The peak with smaller $STD[T^{dB}]$ and smaller $\max[T^{dB}]$ is due to cloud particle scattering. The peak with larger texture attributes is caused by insect scattering. The contour lines in Fig. 7b represent 90%, 75%, 63% and 50% occurrence of 2D Gaussian functions estimated for both populations. The red lines in Figs. 7a and 7c are 1D Gaussian function fits to the observations. Better fits were

5   obtained using Generalized Gaussian functions that accounted for skewness in the observed distributions. However, these better fits did not yield better classifications, as better classifications depend on the samples between the two peaks and not on the outer tails of the distributions that determined the distribution skewness.
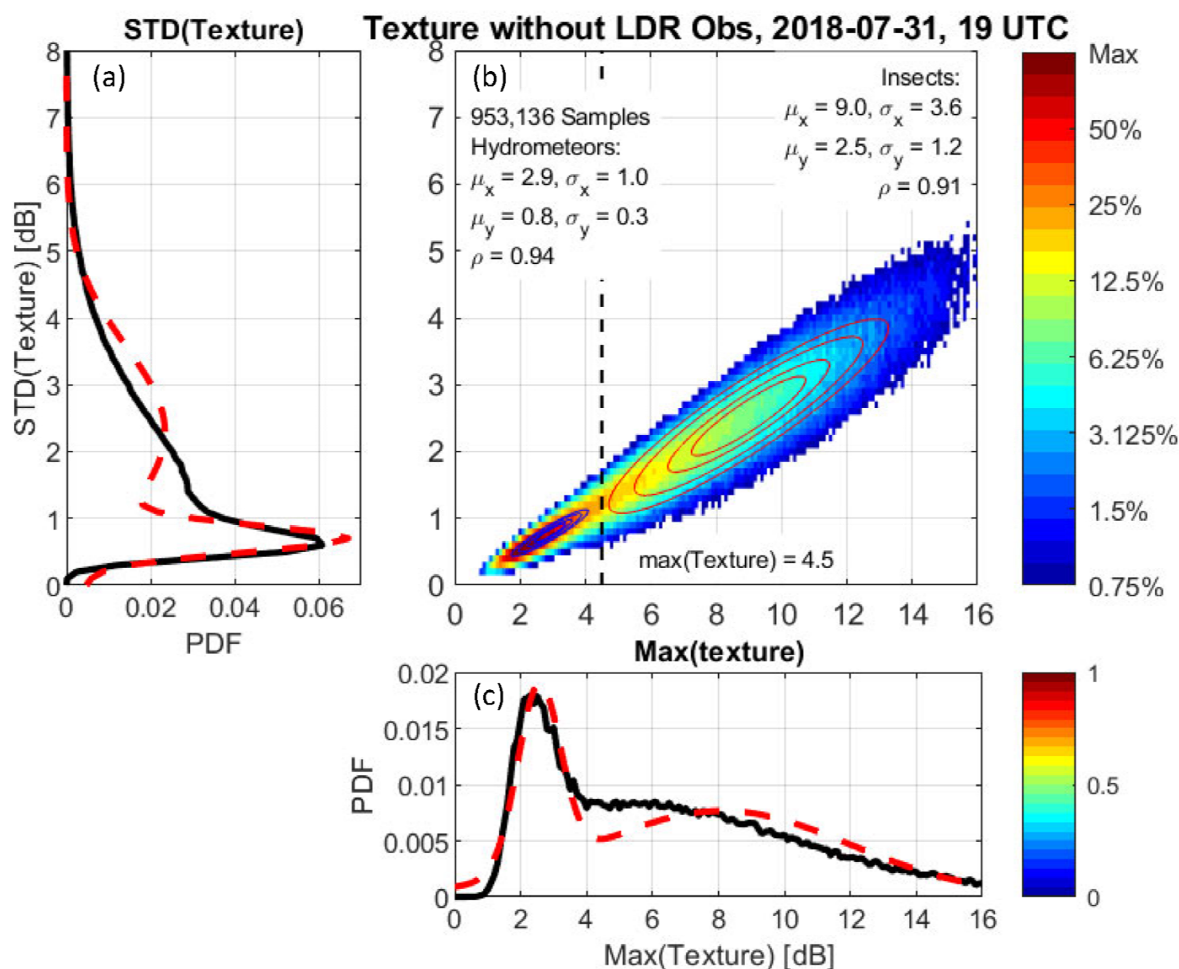
Figure 7: 1D and 2D distributions of texture statistics from hour 19 UTC on 31-July-2018. There are 953,136 spectra regions used in
10   **these distributions. (a) 1D PDF of STD(Texture). Black line is observations and red dashed line is fit to two Gaussian distributions. (b) Colors are observed 2D distribution of STD(Texture) vs. Max(Texture). Colors represent drop from pixel with most occurrences. Blue and red contours are 2D Gaussian fits to hydrometeors (blue) and insects (red). Contours represent 90%, 75%, 63%, and 50% occurrence. Gaussian fit parameters are displayed in panel. Threshold between hydrometeor and insect indicated by dashed black line at max(texture) = 4.5. (c) 1D PDF of Max(Texture). Black line is observations and red line is fit to two Gaussian distributions.**

15   The insect and hydrometeor 2D Gaussian functional fits shown in Fig. 7b have correlation coefficients greater than 0.9 and suggest that simple thresholds in 1D space may be used to define binary insect and hydrometeor texture membership classes. Using the $\max[T^{dB}]$ PDFs shown in Fig. 7c, a threshold of $\max[T^{dB}]_{threshold} = 4.5$ is defined to separate the two texture

membership classes with measured $\max[T^{dB}]$ values greater than $\max[T^{dB}]_{threshold}$ classified as due to insect scattering. Applying the CoPol texture thresholds to the example profile from 19:19:02 UTC, Fig. 8a shows the insect (blue shading) and hydrometeor (red shading) texture membership classes. Also in Fig. 8 are the LDR insect-hydrometeor classes; the combined classes; and the profile mask; all of which are discussed in the next section.
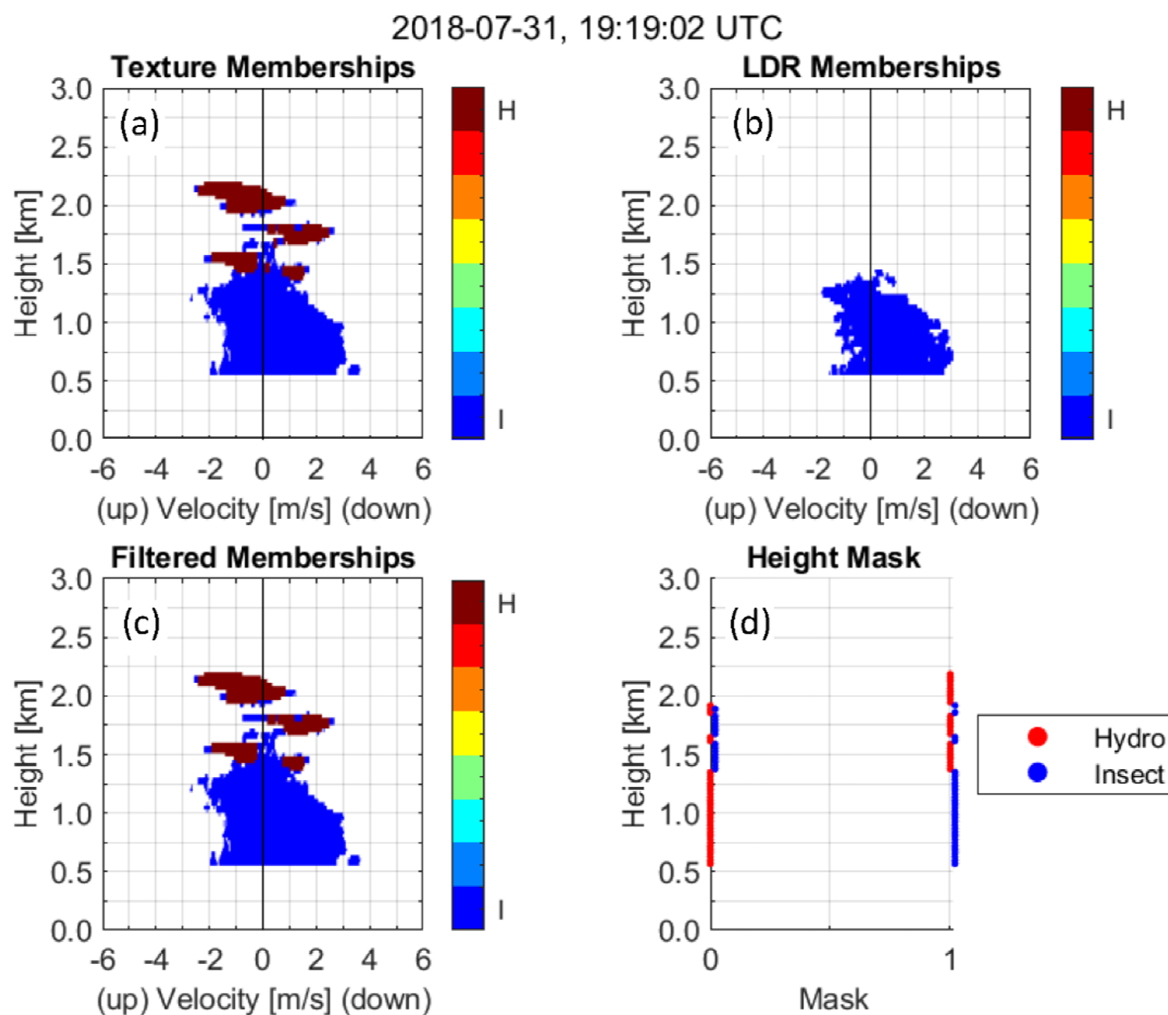
Figure 8: Spectral memberships and binary mask for profile at 19:19:02 UTC on 31-July-2018. Red color indicates hydrometeor membership and blue color represents insect membership (a) Texture algorithm spectral membership. (b) LDR algorithm spectral membership. (c) Filtered spectral membership. (d) Binary hydrometeor and insect mask.

4.2 LDR Algorithm Branch

This section describes the processing steps of the *LDR Algorithm* (Boxes 5-8 of Fig. 5) using the CoPol and XPol spectra profiles at 04:48:17 UTC from the precipitation event on 4-April-2019 shown in Fig. 4. The top row of Fig. 9 (Fig. 9a-d) shows CoPol observations and CoPol texture statistics used in the CoPol texture algorithm. Figures 9e and 9f show XPol and LDR spectra profiles. To estimate regional scattering properties, the same 5x3 velocity-height window used in the texture algorithm is used to calculate regional LDR statistics throughout the $S_{dB}^{LDR}(v_i, h_j)$ spectra profile (Box 7 of Fig. 5). Figures 9g and 9h show the

$mean[S_{dB}^{LDR}(v_i, h_j)]$ and $STD[S_{dB}^{LDR}(v_i, h_j)]$ estimates and suggest that insects are present below 1 km with near zero vertical velocity and falling hydrometeors are present above 3 km. The insects are deduced by $mean[S_{dB}^{LDR}(v_i, h_j)]$ between -10 and -5 dB and the falling hydrometeors by $mean[S_{dB}^{LDR}(v_i, h_j)]$ less than -20 dB. These inferences are supported by the CoPol texture statistics (Figs. 9c and 9d) with insects having large $max[T^{dB}(v_{i\pm2}, h_{j\pm1})]$ near zero vertical velocities below 1 km and smaller

5 values elsewhere. As with the warm shallow cumulus cloud event shown in Fig. 6, there are more CoPol observations (Fig. 6a-d) than LDR measurements (Fig. 6e-h) below 1.5 km.
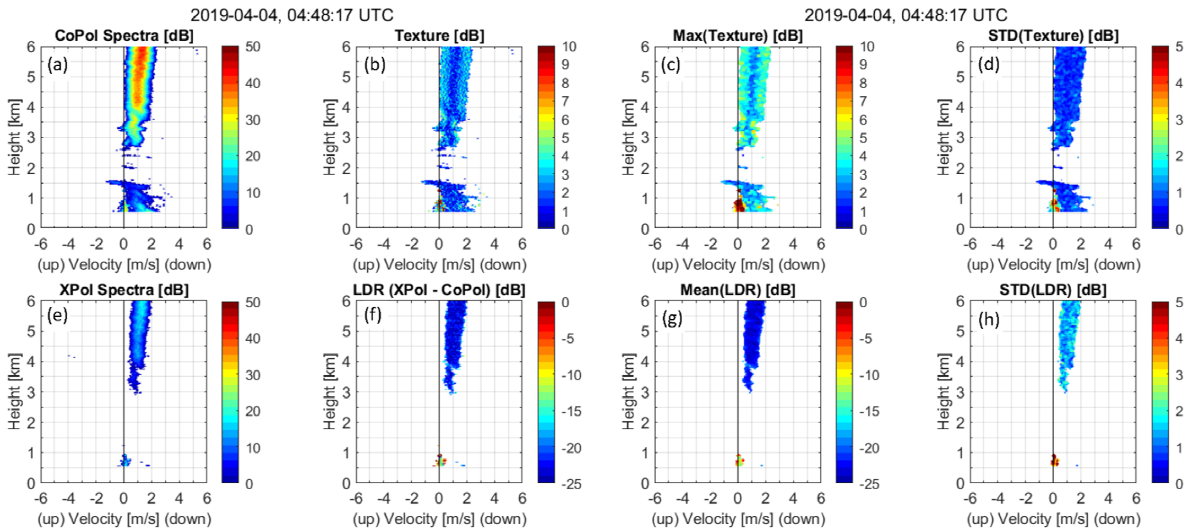
Figure 9: Same as Fig. 6 except for profile at 04:48:17 UTC on 4-April-2019.

With an objective of separating insect and hydrometeor scattering regions based on LDR statistics, Fig. 10 shows 2D and

10 1D PDFs of the LDR statistics estimated for all observations below 1.5 km (to avoid ice particle scattering) for hour 04 on 04-April-2019. Figure 10 contains over 1 million LDR statistic samples each calculated over a separate 5x3 spectral region. The distribution near $mean[S_{dB}^{LDR}(v_i, h_j)] = -8$ dB is due to insect scattering and the distribution near $mean[S_{dB}^{LDR}(v_i, h_j)] = -20$ dB is due to hydrometeor scattering. A threshold of $mean[S_{dB}^{LDR}]_{threshold} = -15$ dB clearly separates the two distributions and is indicated with a dashed line in Fig. 10b.
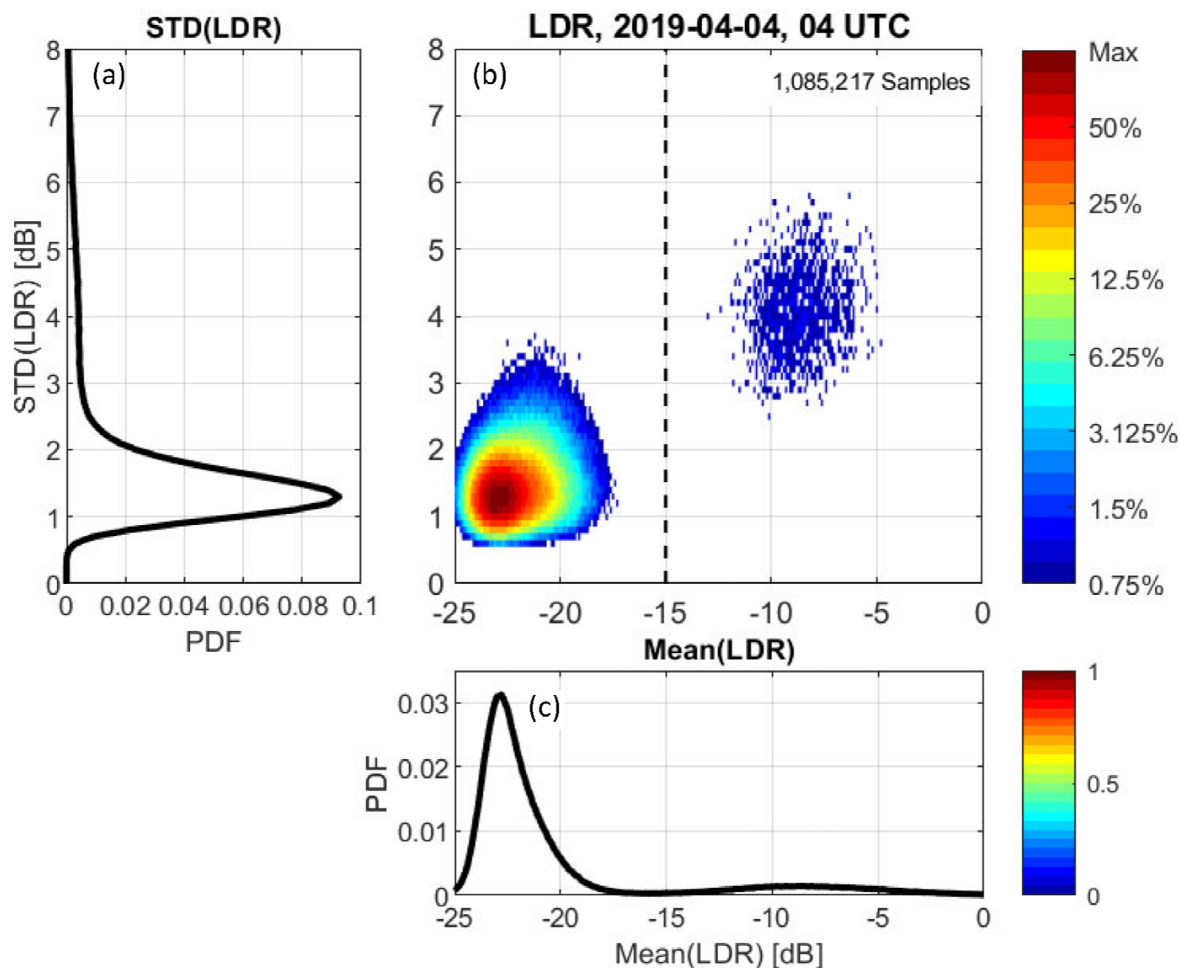
Figure 10: Similar to Fig. 7 except for hour 4 UTC on 4-April-2019 and for STD(LDR) and mean(LDR) statistics. There are 1,085,217 samples collected below 1.5 km height.

Figures 11a and 11b show the CoPol texture and LDR membership classes for this spectra profile. Blue shading indicates insect scattering and red shading indicates hydrometeor scattering. Note that the texture algorithm identifies both insect and hydrometeor scattering below 1.5 km while the LDR algorithm only detects a few insects at these lower range gates. Both algorithms identify hydrometeor scattering above about 3 km.
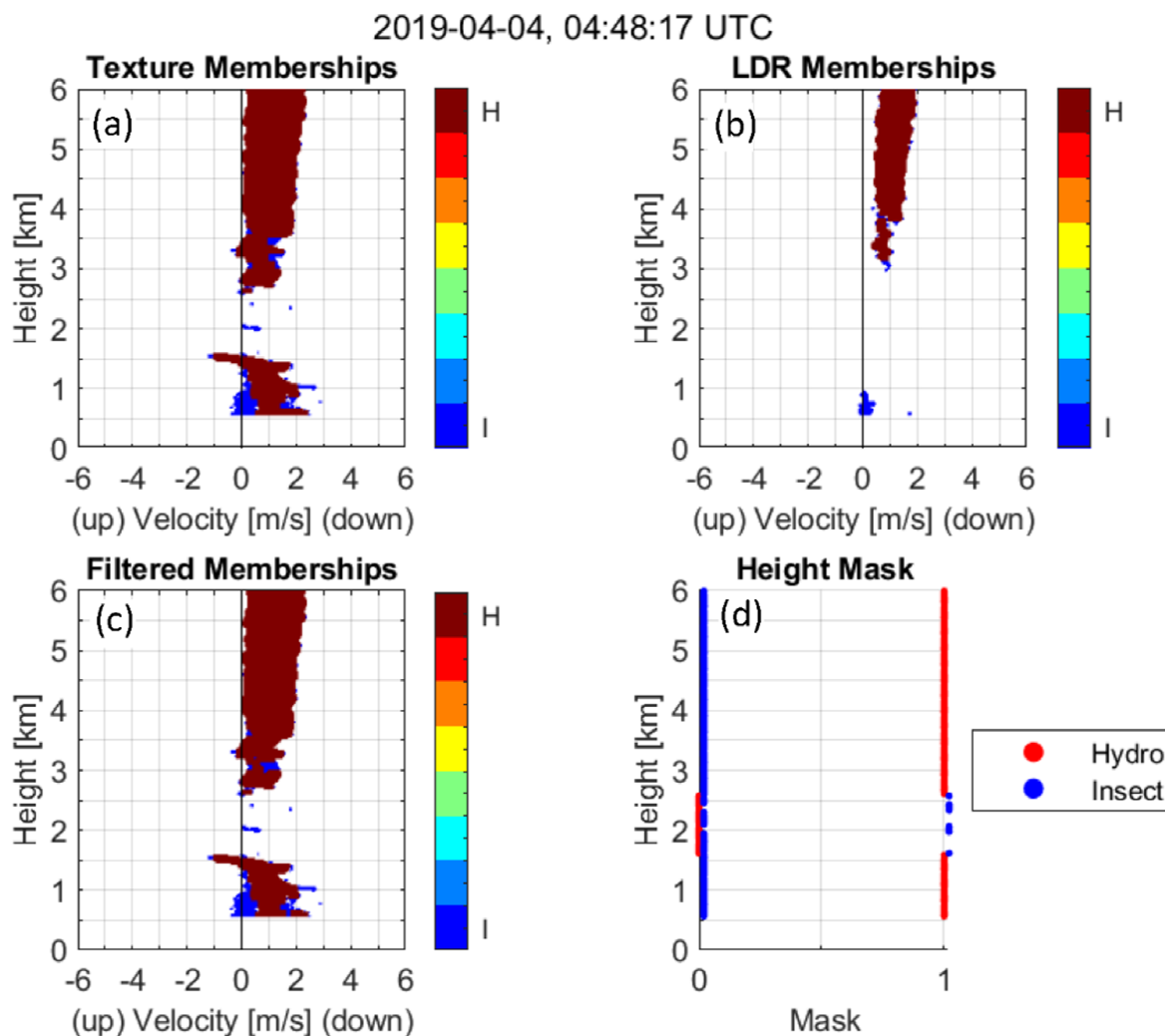
Figure 11: Same as Fig. 8 except for profile at 04:48:17 UTC on 4-April-2019.

4.3 Combining Co-Pol Texture and LDR Algorithm Classifications

After performing the CoPol texture and LDR algorithms, the binary insect and hydrometeor spectral classifications from both
5   algorithms are combined and then filtered (e.g., see Figs. 8a and 8b, and Figs. 11a and 11b). Initially, the combined spectral
classification is the texture classification because the LDR classification will always have fewer valid observations than the CoPol
observations. To incorporate the LDR classification, the combined classification is changed only if the LDR algorithm produced a
hydrometeor class when the texture classification was set to insect class. This logic places more emphasis on identifying
hydrometeors than insects.

10   One of the physical attributes of hydrometeor scattering is that the Doppler velocity spectra span multiple continuous
velocity bins and over several range gates. Accordingly, isolated hydrometeor pixels in the combined spectral classification are
removed by requiring at least 7 continuous hydrometeor pixels in the velocity dimension. All hydrometeor pixels not satisfying
this constraint are set to the insect scattering class. The filtered memberships for the two example profiles are shown in Figs. 8c

for the warm liquid cloud event on 31-July-2018 and Fig. 11c for the precipitation event on 4-April-2019. The red and blue shading corresponds to hydrometeor and insect scattering classes, respectively.

The final processing step is to reduce the filtered membership classes into binary masks indicating the presence of insect or hydrometeor scattering at each range gate (Box 10 of Fig. 5). The insect and hydrometeor masks are set to unity if that filtered

5   membership class exists for that range gate h_j. In the case when both insect and hydrometeor scattering are detected at the same range gate, the hydrometeor mask is set to unity and the insect mask is set to zero. This logic places more emphasis on identifying robust hydrometeor masks and defining masks resulting from either insect or hydrometeor scattering at each range gate. Figures 8d and 11d show the binary insect and hydrometeor masks for the two example profiles. Both masks are saved in output data files and have the variable names *insect_mask_raw* and *hydro_mask_raw* (Boxes 11 and 12 of Fig. 5). The suffix *raw* designates that

10   these masks were estimated from individual profiles and without any temporal information from neighboring profiles, which is discussed in Section 4.4.

In addition to the binary insect mask, an insect activity index is generated by counting the number of insect scattering velocity bins at each height. This insect index $I_{insect}(h_j)$ is defined as

$$I_{insect}(h_j) = \sum_{i=1}^{i_{max}} C_{insect}^{filtered}(v_i, h_j) \qquad (6)$$

15   where $C_{insect}^{filtered}(v_i, h_j)$ is the insect spectral classification and has a value of either 0 or 1. This insect index is not an estimate of the insect number concentration because the magnitude of the insect scattering is not being taken into account. The authors hypothesize that the insect index should be related to insect number density, as the breadth of insect velocities should increase as the number of insects increases. The insect index is available in the output data files with the variable name *insect_index_raw*.

4.4   QC filtering of the Cloud Profile Mask

20   Figure 12 shows the time-height cross-section of observed CoPol KAZR reflectivity (Fig. 12a), the raw hydrometeor mask (Fig. 12b), a time-height filtered hydrometeor mask (Fig. 12c), and the insect index (Fig. 12d) for hour 19 on 31-July-2018. This is the same event shown in Figs. 1 and 2, except with the vertical axis limited to 3 km height. The ceilometer cloud base height is shown in each panel with black dots. The blue and red plus symbols are cloud top and base determined from the COGS stereo camera system, which is discussed in more detail in Section 5. The hydrometeor mask in Fig. 12b is the raw mask produced from each

25   spectra profile. These raw hydrometeor masks contain random misclassified pixels of hydrometeors below the ceilometer cloud base height. Most of these false positive hydrometeor mask pixels are removed by sequentially applying two time-height quality control filters.
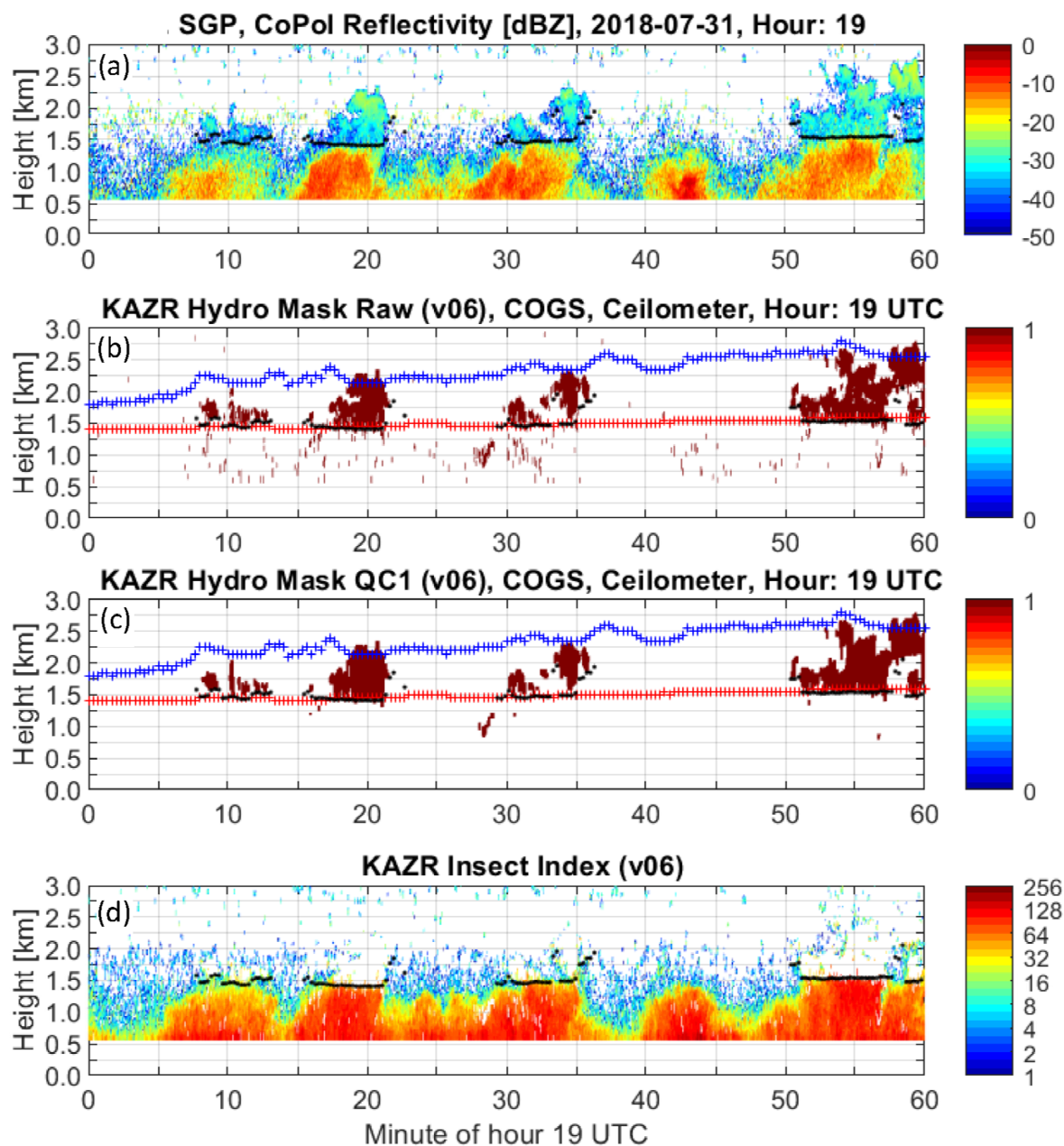
Figure 12: Measurements and retrievals for hour 19 UTC on 31-July-2018. (a) CoPol reflectivity [dB]. (b) Retrieved hydrometeor raw mask (red shading). COGS-derived 6-km x 6-km domain average cloud base (red symbols) and cloud top (blue symbols). (c) Same as panel (b) expect for retrieved hydrometeor QC1 mask. (d) Retrieved insect activity index. Black symbols are ceilometeor-derived cloud base.

The first quality control filter, named QC1 (shown in Fig. 12c), removes temporal outliers by applying a 3-member temporal continuity filter. The QC1 filter also inserts up to three consecutive hydrometeor mask pixels in vertical profiles to fill small gaps in the raw hydrometeor mask. The second quality control filter, named QC2 (not shown, but available in the online Supplemental Material), applies a low-pass filter to the QC1 filtered mask by moving a 3x3 time-height (approximately 12 s by 90

m) continuity constraint throughout the domain to robustly identify hydrometeors that are persistent in time and height. Both the QC1 and QC2 filtered hydrometeor masks are available in the output data files with variable names *hydro_mask_qc1* and *hydro_mask_qc2*. Figure 12d shows the insect index and estimates the number of velocity bins in the spectra that contained insect scattering. The color scale is logarithmic with maximum value 256 representing the number of velocity bins in the spectra. The
5   QC1 hydrometeor mask is plotted for the 4-April-2019 precipitation event in Fig. 4d. There is strong agreement between the ceilometer cloud base height estimates and the hydrometeor mask before minute 20. After this time, the hydrometeor mask identifies raindrops, while the ceilometer is identifying cloud base. COGS measurements are unavailable for comparison purposes during this event because COGS is an optical system requiring daylight.

5.   Comparing Cloud Mask with Independent Measurements

10   Figures 4 and 12 show significant agreement between ceilometer cloud base estimates and retrieved QC1 hydrometeor masks. Figure 12 also shows agreement between COGS cloud base and top estimates with the QC1 hydrometeor mask. In comparing the three products, the KAZR hydrometeor masks and ceilometer cloud base estimates appear as discrete cloud events. Conversely, the COGS estimates appear continuous in time, as if COGS is detecting a persistent cloud layer. This difference is because KAZR and ceilometer are 'soda-straw' observations and COGS is a 6-km x 6-km domain-averaged product produced from three pairs of
15   stereo cameras positioned around the radar and ceilometer (Romps and Öktem, 2018). Figure 12c shows that when the radar and ceilometer both detect clouds, COGS also had a similar cloud base height estimate. The ceilometer and radar cloud bases also showed consistency even at the cloud edges (see near minute 35). Regarding cloud top estimates, COGS estimates are higher than the radar because COGS is a domain average. The online Supplemental Material section contains images of QC1 hydrometeor mask, ceilometer, and COGS retrievals for forty-seven (47) days corresponding to 2018 and 2019 LASSO shallow cloud events
20   (LASSO, 2020). The COGS product is available only for shallow cumuliform clouds and only during daylight hours.

       Figure 13 compares hydrometeor mask QC1 column bottoms with ceilometer cloud bases for the 47 LASSO days. The hydrometeor mask QC1 columns were at least 90 m thick (i.e., 3 consecutive range gates). Using the same format as Figs. 7 and 10, Fig. 13b shows the 2D distribution of height differences with the line graphs showing 1D PDFs. Over 70% of the 12,141 simultaneous profiles had height differences within +/-100 m, which represents +/-3 thirty-meter radar range gates. There is a small
25   skewness to the height difference PDF (Fig. 13a) that is consistent with the ceilometer detecting clouds before the radar detects hydrometeors. Also, during the few precipitation events, the hydrometeor mask bottom was significantly lower than the ceilometer cloud base as the hydrometeor mask detects falling raindrops far below the ceilometer detected cloud base.
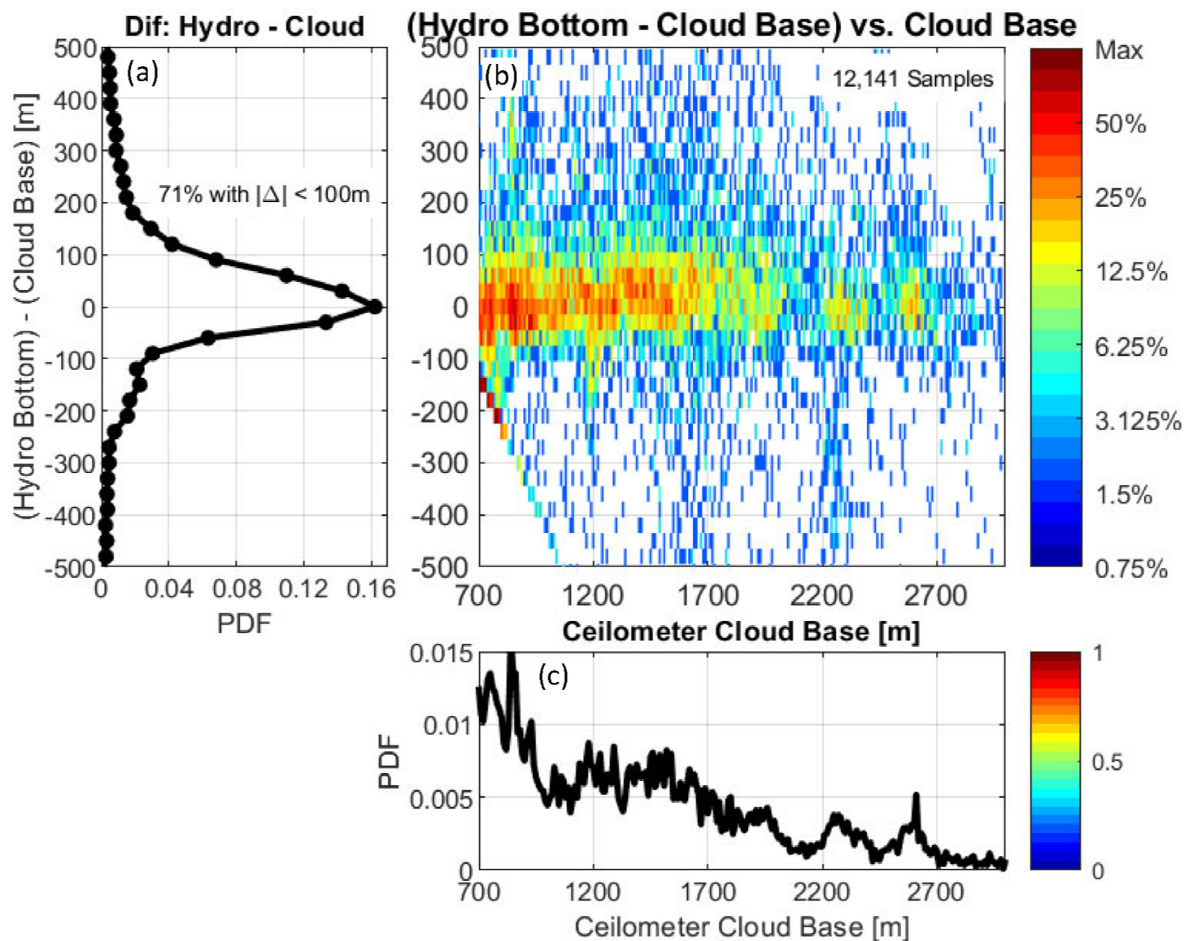
Figure 13: Difference in hydrometeor mask QC1 column bottom and ceilometer cloud base height using 47 days at SGP during 2018 and 2019. There were 12,141 profiles with simultaneous hydrometeor mask QC1 and ceilometer cloud bases. (a) 1D PDF of height difference defined as (Hydrometeor mask column bottom) – (Ceilometer cloud base) [m] with 30-m resolution corresponding to radar range resolution. (b) Colors are 2D distributions of height difference vs. ceilometer cloud base. Colors represent drop from pixel with most occurrences. Artifact at negative height differences and low ceilometer cloud base is due to radar first range gate at 570 m. (c) 1D PDF of ceilometer cloud base [m] with 30-m resolution.

The TSI and Doppler lidar, positioned near KAZR, provide qualitative cloud observations. Figure 14a shows the full spectra profile on 31-July-2018 at 19:19:02 UTC (same as Fig. 3a). The TSI produces a photograph every 30 seconds. Figure 14b shows intermittent clouds passing over KAZR that confirms the broken hydrometeor mask and ceilometer cloud base height estimates shown in Fig. 12. The spectra in Fig. 14c have been screened by the filtered spectra hydrometeor classification to show spectral pixels classified as hydrometeors. While Figs. 14a-c show single profiles, Fig. 14d shows observations for +/- 60 seconds centered on the KAZR profile time. The black horizontal lines are QC2 hydrometeor masks at discrete range gates. The color shading in Fig. 14d is the Doppler lidar attenuated backscattered power [dB]. The decrease in lidar backscattered power with increasing height below 0.5 km is due to decreasing aerosol concentrations and the enhancement near 1.5 km is due to scattering from cloud particles. Thus, the Doppler lidar is detecting cloud base within 30 to 90 meters below the lowest KAZR hydrometeor mask height. The differences between these two sensors can be attributed to their different sensitivities to small cloud particles.
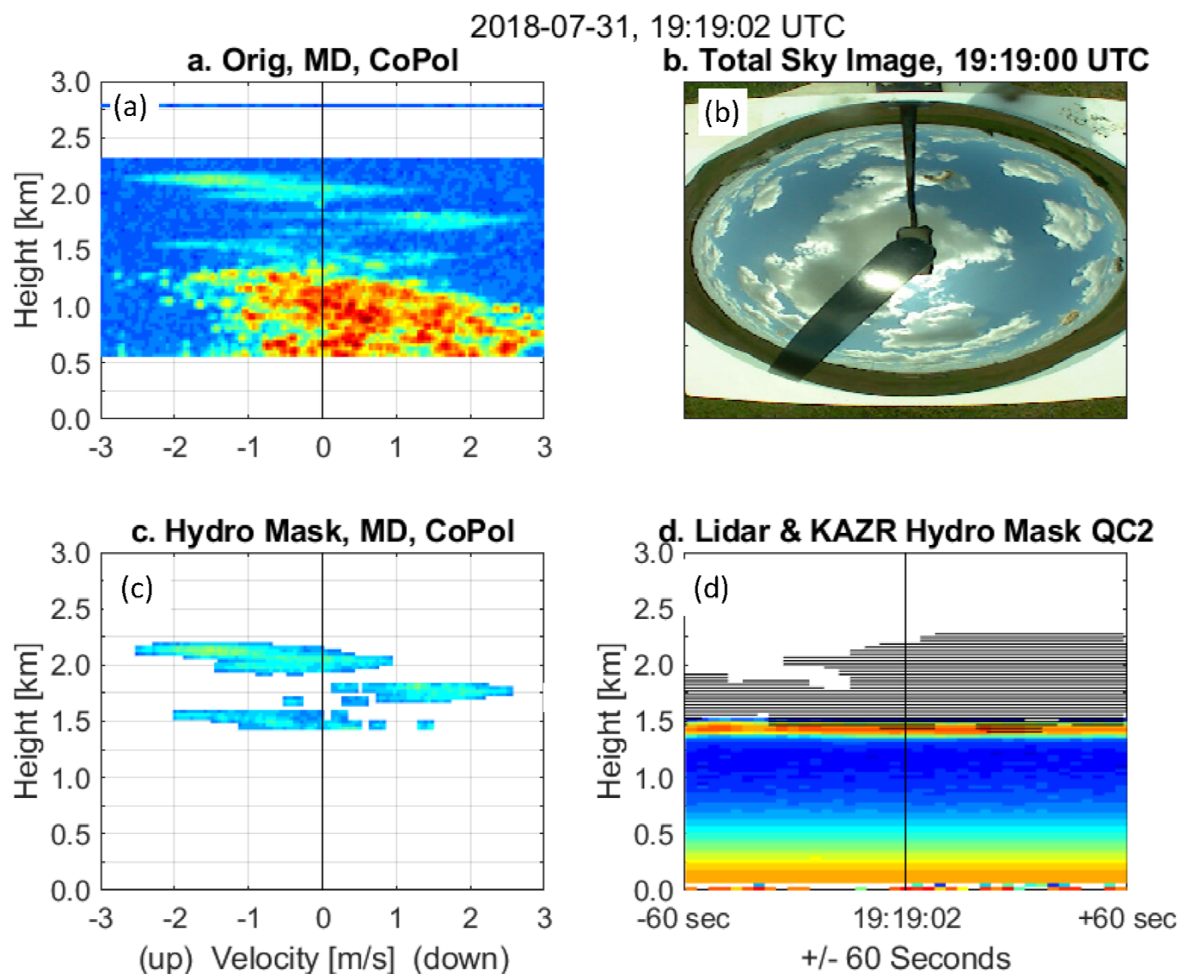
Figure 14: Measurement and retrievals near profile 19:19:02 UTC on 31-July-2018. (a) Doppler velocity power spectra. (b) Total Sky Imager (TSI) photo. (c) Doppler velocity power spectra screened by spectral hydrometeor mask. (d) Doppler lidar attenuated backscattered power for +/- 60 s of profile 19:19:02 UTC. Black lines indicate hydrometeor QC2 mask for +/- 60 s of profile 19:19:02 UTC.

6. Conclusions

In addition to detecting cloud particles, vertically pointing cloud radars are sensitive enough to detect individual insects. If insect contamination is not identified and removed, then radar derived cloud properties will be incorrect and will not help with validating cloud resolving models and / or climate simulations. This study used polarimetric radar observations to develop two insect-hydrometeor detection algorithms. The two algorithms use different radar scattering principles to identify small velocity-height regions in the Doppler velocity power spectra profile as resulting from either insect or hydrometeor scattering. The results of both algorithms are combined and then filtered to produce single value insect and hydrometeor masks at each range gate. One interesting and cautionary note for future studies is that XPol spectra observations observe fewer insects than CoPol observations. This implies that using a polarimetric signal processing method to identify insects will not identify all insect clutter affecting CoPol observations and that insect clutter mitigation methods must use CoPol observations to identify all insect clutter.

One algorithm uses the texture of CoPol Doppler velocity power spectra to identify small velocity-height regions of spectra attributed to insect or hydrometeor scattering. Since insects are individual point targets, their radar power return is confined to narrow intervals of Doppler velocity and range gates. In contrast, cloud particles and raindrops occur in distributions that extend over many velocity bins and several range gates. The CoPol and XPol Doppler velocity power spectra from insect scattering have

5 large variability, or texture, while scattering from cloud particles and raindrops produce smoother, less variable, spectra. The CoPol texture algorithm uses the texture information to identify small regions of insect and hydrometeor scattering.

The other algorithm uses the linear depolarization ratio (LDR) at each point in the Doppler velocity power spectra to identify regions of scattering due to spherical raindrops, asymmetric ice particles, or asymmetric insects. Unlike previous studies, this work uses the LDR at each spectra bin. After identifying small velocity-height regions of insect and hydrometeor scattering in

10 both algorithms, the spectra classifications are combined and then filtered to account for continuity in the Doppler velocity and vertical range dimensions. The filtered spectra classifications are reduced to binary affirmative insect and hydrometeor masks with a single value at each range gate. An insect activity index is estimated at each range gate by counting the number Doppler velocity spectra bins with insect scattering. There appears to be relationships between the insect activity index, radar reflectivity and cloud formation. Future studies will explore these relationships, as well as vertical air motion to explore whether insects are passive

15 tracers or actively propelling themselves through the atmosphere. Often, insects occur at the same height as clouds and during the onset of precipitation. While these are interesting phenomena, the focus of this work is producing robust hydrometeor masks to help identify cloud boundaries, which can be used, for example, to study the evolution of shallow cumulus clouds in the planetary boundary layer (Gustafson et al., 2017). Using over 12,000 simultaneous ceilometer and radar profiles, it was found that over 70% of the hydrometeor mask column bottoms were within +/-100 m of the ceilometer cloud base (i.e., +/-3 thirty-meter radar range

20 gates). The hydrometeor mask column bottom was slightly higher than the ceilometer cloud base. This is to be expected, as the ceilometer detects cloud particles at lower heights than the radar detecting hydrometeors within the cloud.

The online Supplemental Material includes images of observed KAZR reflectivity, retrieved hydrometeor masks, and verification observations from ceilometer, COGS, TSI, and Doppler lidar. The processing described herein was applied to KAZR observations for thirty (30) days in 2018 and seventeen (17) days 2019 deployed at the Southern Great Plains (SGP) facility

25 corresponding to LASSO event days (LASSO, 2020). The insect and hydrometeor masks for these days are available online on the DOE ARM Archive in netCDF format.

Data availability. Original raw KAZR spectra are available on the DOE ARM archive. After release of the manuscript, forty-seven (47) days of insect and hydrometeor mask data files at Southern Great Plains (SGP) Central Facility will be available at the DOE

30 ARM archive as a Principle Investigator (PI) Product (Williams, 2021).

Source code availability. The code used to generate the insect, cloud, precipitation, and hydrometeor masks stored on the DOE ARM archive is available upon request from the lead author. With this source code, users can repeat the analysis presented in this study and develop improved insect-cloud and insect-precipitation detection algorithms for their vertically pointing radar

35 observations.

Supplemental Material. Images of observed KAZR reflectivity, retrieved hydrometeor masks, and verification observations from ceilometer, COGS, TSI, and Doppler lidar are available in the online Supplemental Material.

Author contribution. CRW: insect and hydrometeor mask code development; KLJ: testing and evaluation of products and verification; SEG and JCH: Evaluation of products against collocated observations; RO and DMR: processing of COGS product.

Competing interests. The authors declare that they have no conflict of interest.

5 *Disclaimer*. The authors declare no disclaimers.

Acknowledgements. We recognize and appreciate the work of field technicians tasked with keeping instruments running at U.S. Department of Energy (DOE) field sites. CRW received support from the DOE Atmospheric Radiation Mission (ARM) program under Contract No. 508641. This research was supported by the DOE Office of Biological and Environmental Research as part of the ARM Climate Research Facility, an Office of Science scientific user facility. RO and DMR acknowledge support from the

10 DOE Atmospheric System Research (ASR), an Office of Science, Office of Biological and Environmental Research program; Lawrence Berkeley National Laboratory is operated for the DOE by the University of California under Contract No. DE-AC02-05CH11231. This paper has been authored by employees (KLJ and SEG) of Brookhaven Science Associates, LLC, under contract No. DE-SC0012704 with the U.S. Department of Energy (DOE). The publisher by accepting the paper for publication acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or

15 reproduce the published form of this paper, or allow others to do so, for United States Government purposes.

References

Ackerman, T. P., and G. M. Stokes: The Atmospheric Radiation Measurement Program. *Phys. Today*, **56**, 38–44, doi:10.1063/1.1554135, 2003.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Total Sky Imager (TSISKYCOVER). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by V. Morris and D. Flynn. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1025308, 2000.

Atmospheric Radiation Measurement (ARM) user facility: 2010a, updated hourly. Doppler Lidar (DLFPT). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by R. Newsom and R. Krishnamurthy. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1025185, 2010a.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Ceilometer (CEIL10M). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by V. Morris, D. Zhang and B. Ermold. ARM Data Center. Data set accessed 2020-05-01, 2010b.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Ka ARM Zenith Radar (KAZRSPECCMASKGECOPOL). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by I. Lindenmaier, N. Bharadwaj, D. Nelson, B. Isom, J. Hardin, A. Matthews, T. Wendler and V. Castro. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1025218, 2011a.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Ka ARM Zenith Radar (KAZRSPECCMASKGEXPOL). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by I. Lindenmaier, N. Bharadwaj, D. Nelson, B. Isom, J. Hardin, A. Matthews, T. Wendler and V. Castro. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1025219, 2011b.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Ka ARM Zenith Radar (KAZRSPECCMASKMDCOPOL). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by I. Lindenmaier, N. Bharadwaj, D. Nelson, B. Isom, J. Hardin, A. Matthews, T. Wendler and V. Castro. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1095603, 2011c.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Ka ARM Zenith Radar (KAZRSPECCMASKMDXPOL). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by I. Lindenmaier, N. Bharadwaj, D. Nelson, B. Isom, J. Hardin, A. Matthews, T. Wendler and V. Castro. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1095604, 2011d.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Active Remote Sensing of CLouds (ARSCL) product using Ka-band ARM Zenith Radars (ARSCLKAZR1KOLLIAS). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by K. Johnson and T. Scott. ARM Data Center. Data set accessed 2020-05-01 at http://dx.doi.org/10.5439/1393437, 2014.

Atmospheric Radiation Measurement (ARM) user facility: updated hourly. Clouds Optically Gridded by Stereo (COGS) product (COGSPI). 2018-01-01 to 2019-12-31, Southern Great Plains (SGP) Southern Great Plains Network (N1). Compiled by K. Gaustad. ARM Data Center. Data set accessed 2021-01-16 at http://dx.doi.org/10.5439/1644322, 2017.

Baldini, L., and E. Gorcucci: Identification of the melting layer through dual-polarization radar measurements at vertical incidence. *J. Atmos. Ocean. Technol.*, **23**, 829-839, 2006.

Carter, D. A., Gage, K. S., Ecklund, W. L., Angevine, W. M., Johnston, P. E., Riddle, A. C., Wilson, J. S. and Williams, C. R.: Developments in UHF lower tropospheric wind profiling at NOAA's Aeronomy Laboratory, Radio Science, 30, 977-1001, 1995.

Cess, R. D., and Coauthors: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. J. Geophys. Res., 95, 16 601–16 615, doi:10.1029/JD095iD10p16601, 1990.

Clothiaux, E. E, K. P. Moran, B. E. Marther, T. P. Ackerman, G. G. Mace, T. Uttal, J. H. Mather, K. B. Widener, M. A. Miller, and D. J. Rodriguez: The Atmospheric Radiation Measurement program cloud radars: Operational modes. J . Atmos.

5   Oceanic Technol., 16, 819-827, 1999.

Drake, V. A., S. Hatty, C. Symons, and H. Wang: Insect Monitoring Radar: Maximizing Performance and Utility. *Remote Sens*. 12(4), 596; https://doi.org/10.3390/rs12040596, 2020.

Gustafson, WI, AM Vogelmann, X Cheng, S Endo, B Krishna, Z Li, T Toto, and H Xiao: Recommendations for Implementation of the LASSO Workflow. Ed. by R. Stafford, DOE Atmospheric Radiation Measurement Climate Research Facility.

10   DOE/SC-ARM-17-031, doi:10.2172/1406259, 2017.

Hildebrand, P. H. and Sekhon, R.: Objective determination of the noise level in Doppler spectra, J. Appl. Meteorol., 13, 808-811, https://doi.org/10.1175/1520-0450(1974) 013,0808:ODOTNL.2.0.CO;2, 1974.

Khandwalla, A., S. Sekelsky, L. Li, and M. Bergada: Theory and observations between Ka-band and W-band to explain scattering differences between insects. Proc. 11th ARM Science Team Meeting, Atlanta, GA DOE ARM. [Available

15   online at http://www.arm.gov/publications/proceedings/conf11/extended_abs/khandwalla_a.pdf], 2001.

Khandwalla, A., N. Majurec, S. M. Sekelsky, C. R. Williams, and K. S. Gage: Characterization of radar boundary layer data collected during the 2001 multi-frequency radar IOP. Proc. 12th ARM Science Team Meeting, St. Petersburg, FL, ARM [Available online at http:www.arm.gov/publications/proceedings/conf12/extended_abs/khandwalla-a.pdf.], 2002.

Klingle, D. L., D. R. Smith, and M. M. Wolfson: Gust front characteristics as detected by Doppler radar, Mon. Wea. Rev., 115,

20   905-918, 1987.

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

25   LASSO: Large Eddy Simulation (LES) ARM Symbiotic Simulation and Observation (LASSO) Bundle Browser – Visualization & Access, https://adc.arm.gov/lassobrowser, accessed 1-May-2020.

Lhermitte, R. M.: Probing air motion by Doppler analysis of radar clear air returns. J. Atmos. Sci., 23, 575–591, doi:10.1175/1520-0469(1966)023,0575:PAMBDA.2.0.CO;2, 1966.

Lohmeier S., P., S. M. Sekelsky, J. M. Firda, G. A. Sadowy, and R. E. McIntosh: Classification of particles in stratiform clouds

30   using the 33 and 95 GHz polarimetric Cloud Profiling Radar System, IEEE Trans. Geosci. Remote Sens., 35, 256-270, 1997.

Luke, E.P, P. Kollias, K.L. Johnson, and E.E. Clothiaux: A Technique for the Automatic Detection of Insect Clutter in Cloud Radar Returns. J. Atmos. Ocean. Technol., 25, 1498-1513, doi:10.1175/2007JTECHA953.1, 2008.

Mather, J. H., and J. W. Voyles, J. W.: The ARM Climate Research Facility: A review of structure and capabilities, Bull. Amer.

35   Meteor. Soc., 94, 377-392, doi: 10.1175/BAMS-D-11-00218.1, 2013.

Martner, B. E., and K. P. Moran: Using cloud radar polarization measurements to evaluate stratus cloud and insect echoes. J. Geophys. Res., 106, 4891-4897, https://doi.org/10.1029/2000JD900623, 2001.

Moran, K. P., B. E. Martner, M. J. Post, R. A. Kropfli, D. C. Welsh, and K. B. Widener: An unattended cloud-profiling radar for use in climate research. Bull. Amer. Meteor. Soc., 79, 443–455, doi:10.1175/1520-

40   0477(1998)079,0443:AUCPRF.2.0.CO;2, 1998.

Mueller, E. A., and R. P. Larkin: Insects observed using dual polarization radar, J. Atmos. Oceanic Technol., 2, 49-54, doi: https://doi.org/10.1175/1520-0426(1985)002<0049:IOUDPR>2.0.CO;2, 1985.

Nansen, C., and N. Elliott: Remote sensing and reflectance profiling in Entomology. Ann. Rev. Entomol., Vol. 61:139-158, https://doi.org/10.1146/annurev-ento-010715-023834, 2016.

5   Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. Hartmann: Cloud radiative forcing and climate: Results from the Earth Radiation Budget Experiment. Science, 243, 57–63, doi:10.1126/science.243.4887.57, 1989.

Riley,J . R., Remote sensing in entomology, Ann. Rev. Entomol., 34, 247-271, 1989.

Romps, D. M., and R. Öktem: Observing clouds in 4D with multiview stereophotogrammetry. Bull. Amer. Meteor. Soc., 99,
10   2575-2586, https://doi.org/10.1175/BAMS-D-18-0029.1, 2018.

Sekelsky, S. M., L. Li, J. Calloway, R. E. McIntosh, M. A. Miller, E. E. Clothiaux, S. Haimov, G. C. Mace, and K. Sassen: Comparison of millimeter-wave cloud radar measurements for the fall 1997 cloud IOP, in Proceedings of 8th ARM Science Team Meeting, Dep. of Energy, pp. 671-675, Tucson, Ariz., 1998.

Williams, C. R., and K. S. Gage: Raindrop size distribution variability estimated using ensemble statistics. Ann. Geophys., 27,
15   555-567, 2009, www.ann-geophys.net/27/55/2009/, 2009.

Williams, C. R., Hydrometer mask dataset for 2018 and 2019 to be submitted to DOE ARM Archive after release of this manuscript. 2021.

Westbrook, J. K., R. S. Eyster, and W. W. Wolf: WSR-88D Doppler radar detection of corn earworm moth migration. Int. J. Biometeorol., doi: 10.1007/s00484-013-0676-5, 2014.