Interactive comment on "A kernel-driven BRDF model to inform satellite-derived visible anvil cloud detection" by Benjamin Scarino et al.

Anonymous Referee #1

General comments

In general, quantitative uses of the visible bands on the GOES are under-utilized. This work will greatly help with that issue, to better automatically detector and characterize anvil clouds.

Thank you for taking the time to review and offer feedback for our paper. Our responses are offered below, and any applicable changes are indicated by page and line number.

Specific Comments

Page 1. Line 28. Consider these missing references: Brunner, J. C., S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin,2007: A quantitative analysis of the enhanced-V feature in relation to severe weather. Wea. Forecasting, 22, 853–872. McCann, D. W., 1983: The enhanced-V: A satellite observable severe storm signature. Mon. Wea. Rev., 111, 887–894.

Please reference Ted Fujita's "jumping (stratospheric) cirrus": Fujita, T. T., 1974: Overshooting thunderheads observed from ATS and Learjet. Satellite and Mesometeorology Research Project Rep. 117, Texas Tech University, Lubbock, TX, 29 pp.

Menzel, W. P., and J. F. W.Purdom, 1994: Introducing GOES-I: The first of a newgeneration of Geostationary Operational Environmental Satellites. Bull. Amer. Meteor. Soc., 75, 757–781.

Thank you for these suggestions, the references have been included.

Page 2. Line 5. It should be noted that the legacy GOES (eg, GOES-13/15) were calibrated pre-launch, but not on orbit. Unless, you have applied a X. Wu (NOAA STAR) visible calibration correction (to account for degradation over time). If this is the case, it should be stated. Details of the correction: https://www.star.nesdis.noaa.gov/smcd/spb/fwu/homepage/GOES_Imager_Vis_OpCal.php Is this what the authors are referring to, when they refer to "spectral band normalization"?

We apologize for not making this clear initially, but the "calibrated visible reflectance" we're discussing here and elsewhere in the paper is neither referring to the pre-launch nor to Fred Wu's on-orbit calibration corrections. Rather we are referring to CERES Edition 4 visible imager calibration method coefficients of Doelling et al. 2018, which are determined from the monthly gain trends of GEO and Aqua MODIS spectrally consistent, ray-matched radiance pairs over all-sky tropical ocean, deep convective clouds, and deserts. The spectral consistency, or spectral band normalization, aspect of the calibration is ensured by the application of spectral band adjustment factors (SBAFs) derived by convolving scene-specific hyperspectral data from the SCIAMACHY instrument with the spectral response functions of the reference (MODIS) and target (GEO) sensors, further described by Scarino et al. 2017. Although we did reference these two articles and pointed to the Aqua MODIS standard shortly after the initial mention of *calibrated visible reflectance* and *spectral band normalization*

(originally Page 2, Line 9), further explanation about the calibration origin did not come until Page 4, starting at Line 34. Also, the SBAF was not specifically mentioned.

To address this, we have added mention of the CERES Edition 4 calibration method and spectral band adjustment factors to Page 2, starting at Line 5 where the term *calibrated* is first introduced.

Page 2. Line 29. Note also that the brightness temperature of a water vapor band varies with view angle. For example, a cooling of over 5K for the legacy GOES imager water vapor band at 55 Local (Viewing) Zenith Angle: http://cimss.ssec.wisc.edu/goes/calibration/GOES12_IMGR_LZAvsTEMP.jpg

Thank you for this note. We have added discussion and reference of the strong VZA dependency in the WV band for legacy GOES starting at Page 2, Line 30 and Page 11, Line 38.

Page 4. Line 26. "The VIS retrievals are subsampled to the IR data resolution". Was an averaging approach considered? Or ran on one day to understand if there are any differences between sampling or averaging?

Your question reveals that we did not clearly explain how we handle resampling, subsampling, their distinction, and what data resolutions are utilized at the different steps of the process. For a short explanation, the key points are: 1) There is generally little resultant difference from subsampling vs. averaging in regard the method being described, which is aggregation of anvil reflectance, given the homogeneity of the scene. 2) Resampling to a fixed grid, subsampling, and averaging are all occurring, but the exact process depends on the imager and the variable being output. 3) Retrieval processes do in fact take advantage of the higher VIS resolution. That is, VIS downscaling does not happen until the final output. The full explanation is as follows, which is summarized in the restructured section starting at Page 4, Line 30:

We have explored DCC reflectance measurement sensitivity to subsampling vs. averaging and have found that the difference is insignificant for scenes that are relatively homogeneous across many kilometers, which is the case for the anvil clouds we are targeting in this study. The largest deviations between the two methods come in the vicinity of cloud edges, within a field of broken clouds, near OT, or amongst gravity waves. Cloud edges and scattered clouds are easily avoided with IR- and VIS-based homogeneity filters, threshold checks, and continuity assurances regardless of sampling method. Even smaller-scale VIS variabilities caused by subtle OT and gravity wave features are filtered comparably whether employing subsampling or averaging because they had first been damped through the 6 × 6 array Lanczos-filtering-based resampling process. The point is that the final reflectance aggregation is comparable whether subsampling or averaging because large homogeneities are discovered either way, and subtle homogeneities have been muted through the resampling process.

That said, we still must explain what method was employed. The text originally did not make it clear that the Lanczos resampling process was also performed on GOES data – not just on Himawari-8. We also did not adequately convey that the Himawari-8 data were first subsampled (by skipping every other line and element using McIDAS-X software) to GOES "1-km" resolution before resampling. The final result of the resampling is a fixed-scale grid at 1-km resolution for VIS, and 4-km resolution for IR. As such, hereafter, unless otherwise stated,

the term *pixel* refers specifically to individual data samples of the Lanczos-interpolated fixed grid rather than the actual imager pixel data. Given the two fixed-grid scales, algorithm processes involving VIS can take advantage of 1-km data. Only upon output are the data sampled to 4 km. The method of sampling in this step, which is with respect to the fix-grid, depends on the variable being output. In the case of reflectance, the sixteen 1-km pixels of the corresponding 4-km IR pixel are averaged. In the case of one of our product outputs like visible texture rating, the maximum value is output (although with Fourier smoothing applied – such details that are beyond the scope of this paper but rather are being prepared for an ATBD). The important point is that as far as VIS output is concerned, our method of creating the anvil BRDF of expected reflectance relies only on the resampled reflectance, which prior to output was downscaled to IR resolution by averaging.

It is important to point out that these examples are specific to the satellite data being discussed in this section for the particular purpose of aggregating homogeneous reflectance measurements coincident with the IR anvil mask across legacy GEOs. For a different application we may utilize full-resolution Himawari-8 or GOES-16, in which case VIS would be resampled to a fixed 0.5-km grid, IR to a fixed 2-km grid, algorithm processes would consider 0.5-km VIS data, and final output would be 2-km for all parameters. In this case, sixteen 0.5-km reflectance pixels would be averaged for each 2-km IR box, or, again in the case of visible texture, the maximum value of sixteen would be output.

Page 5. Line 3. It's stated that Meteosat aren't used due to a lack of 1km vis data over full disk, but the 1km is sampled to 4km in your analysis anyway. So, would using 3km vs 4km be that different? Especially considering that the 3km is at the subpoint? Isn't a larger difference with Meteosat is the timing, given those data are scanned south to north, while the GOES scan north to south?

It is correct that that main reason we did not use Meteosat data for the construction of the anvil BRDF (or more precisely, the look lookup table of average anvil reflectance) is because rather than having 1-km VIS imagery across the full field of view as with GOES and Himawari, 1-km MSG VIS imagery is broadly available only for the Northern Hemisphere. In the Southern Hemisphere, 1-km MSG VIS follows the Sun, which is disadvantageous for an angular dependency model. As described above, we resample VIS reflectance to the Lanczos-interpolated 1-km fixed grid, and then average to a 4-km fixed grid for the purposes of harmonized output with IR. Therefore, the GOES and Himawari processes utilize 1-km VIS, whereas comparable MSG processes would have to use 3-km. The question then is whether we expect a difference in anvil identification and reflectance measurements between Lanczos-interpolated 1-km GOES/Himawari VIS data and 3-km MSG VIS data. For IR we would not expect a significant brightness temperature variation for 3 km vs 4 km given the homogeneous clouds we are targeting. For VIS, however, small-scale OT shadowing and texture variances are likely to yield differences in identification and measurement for 1 km vs 3 km. Also, regarding the point about the 3-km MSG resolution being at the sub-satellite point, the resampling process to achieve the fixed grid masks the effect of FOV growth as view moves from nadir to limb. More discussion on this, however, can be found in response to a comment further below.

The opposite full disk scanning pattern compared to GOES and Himawari would not be a significant concern for the purpose discussed in this section, which is construction of the multi-angle LUT for anvil reflectance. There is no time matching involved in this process, as the only dependency being investigated here is that on combined VZA, SZA, and RAA. That is,

we collected half-hourly reflectance measurements independently for each satellite and averaged all results into a single angular-dependent LUT without requirement for time consideration. Convective processes of each satellite region are distinct enough (spatially and over a 1-year period) that inconsistency in scan direction should not introduce systematic bias in reflectance measurements for specific angular configurations.

Page 5. Line 10. Please clarify how this relates to the paragraph: "Note that for some analyses, satellite data are supplemented by modeled atmospheric profiles provided by the Global Modeling and Assimilation Office (GMAO) Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) product."

The intent was to conclude Section 2.1 with mention of the MERRA-2 inclusion. We agree, however, that it did not fit naturally as the final sentence of the MSG discussion paragraph. Therefore, we have moved the sentence to the end of the first paragraph of Section 2.1 (Page 4, Line 27), where we feel its description of being a satellite data supplement fits more intuitively.

Page 6. Line 5. How was the 5x5 determined? Given that a projected field-of-view size grows away from nadir, this mean that less area is considered near the sub-point compared to near the limb. How might this affect your results?

The 5x5 array was chosen for homogeneity assessment as the compromise between the model having higher uncertainty or being excessively constrained. First of all, we chose an odd number like 5 rather than 4 or 6 because doing so simplifies evaluation of a centered pixel relative to its surrounding pixels, which are equidistant on the fixed grid. We initially constructed the LUT based on a 3x3-array homogeneity evaluation, but this allowed for inclusion of questionable anvil pixels given that there was less we could quantitatively interpret in order to filter the results. Our LUT uncertainty (Fig. 5) was significantly higher than what is shown for the 5x5 array case, especially at the higher SZA bins. Aggregating many questionable pixels would not be helpful for a model designed to provide expected anvil reflectance. On the other hand, it is expected that choosing a 7x7 array for evaluation of homogeneity would be too conservative. Spatial coherency and consistency in positive identification from an empirical perspective are the key aspects we aim to capture in our anvil and OT detections products, and thus we believe homogeneity evaluation on a 7x7-pixel scale would exclude too many clouds that an informed person would otherwise identify as anvil.

We don't specifically account for the FOV growth, but it is effectively masked by the resampling process. Of course, the effect does not simply disappear because we resample. Distortion of an input pixel will in some form persist through to the resampled pixel, and the share of satellite pixels being combined in the resampling process is dependent on the viewing angle. Nevertheless, for the purpose of anvil reflectance aggregation, resampling an enlarged satellite pixel from an extreme viewing angle is not detrimental because the FOV is less likely to feature subtle shadowing and fine-scale texture. It is true then that OT detection is negatively impacted by FOV stretching, but that is beyond the scope of this paper. The main importance is that we construct a model that provides the expected reflectance for a given view, regardless of whether extreme views are distorted or are masking features that would otherwise result in exclusion at less extreme angles. In other words, what the imager "sees" at a particular angular configuration is the truth we are aiming to model.

Page 8. Line 6. This is every other 4km line and 4km element, correct? If so, this should be stated here, for clarity.

This is actually in reference to the Lanczos-interpolated fixed VIS grid. The actual numbers are dependent on the satellite input. For example, for a stand-alone GOES-16 VIS anvil mask computation, the resampled 0.5-km VIS reflectance imagery is subset to 1-km by taking every other row and column of the fixed grid. We hope that the edits made in Section 2.1 (explained above) provide better context for this statement. We have also added specific mention of the subsampling being in reference to the fixed grid, and have provided example scale values so that there is no ambiguity (Page 8, Line 21).

Page 8. Line 11. Do you screen out the sun glint times/location?

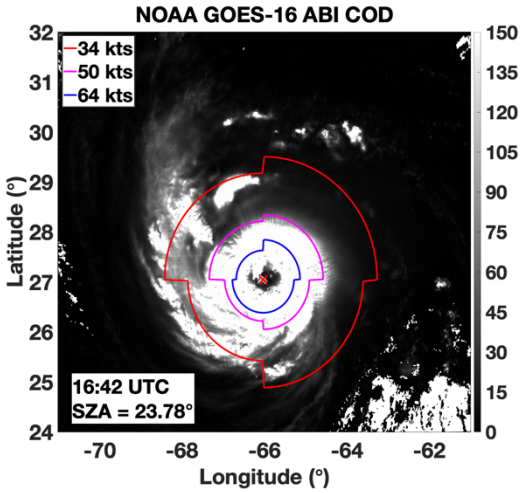
No, we do not employ any special handling of sun glint regions because any anvil reflectance enhancement owed to glint would be predicted by the BRDF model. Thus, observed reflectance should closely match the nominal reflectance, and anvil determination can continue. It is possible the pixels will be saturated – an occurrence that is specifically considered when the anvil mask is being constructed because OT edges can appear very bright in certain scattering configurations. This is the type of saturation being discussed on the line you indicated. Unlike in the BRDF construction, the goal when defining the anvil mask is to be inclusive of anomalous shadowed or saturated pixels that are otherwise surrounded by anvil. So rather than screening out sun glint, the aim is to correctly classify glint-saturated pixels. We have updated the text to include mention of not only bright OT edges, but also sun glint (Page 8, Line 29).

Page 9. Line 19. When it is stated "adapted to GOES-16 imagery", does this still mean using 4km spatial resolutions for both the vis and IR inputs? If so, consider running a case (for comparison) at 2km, to use more of both the vis and IR information. Add a GOES-16 reference: http://nwafiles.nwas.org/jom/articles/2018/2018- JOM4/2018-JOM4.pdf

We were attempting to explain that the COD retrieval methodology used for CERES instruments, which is described by Minnis et al. (2020), has been adapted to the SatCORPS Framework as applied to GOES-16, which is nominally run at 4-km resolution over CONUS. Note that the sensitivity of SatCORPS ice cloud COD to retrieval resolution was investigated by Minnis et al. 2016 (https://www1.ncdc.noaa.gov/pub/data/sds/cdr/CDRs/AVHRR_Cloud_Properties_NASA/Algorithm_Description_01B-30b.pdf**), who found no variation in COD whether retrieving at 1-km, 2-km, or 4-km resolution. So to be clear, "adapted to GOES-16 imagery" specifically concerns adaptation of the CERES COD method to SatCORPS GOES-16. That said, the VIS anvil mask that we aligned with the SatCORPS GOES-16 COD for this section on COD parameterization was from a run done with 0.5-km VIS, which hopefully we have now provided better context for given our previous responses and edits. We have changed this sentence to clarify that CERES COD was adapted to SatCORPS GOES-16, referenced the COD retrieval sensitivity analysis, and provided specific scale values for the datasets involved in this process. Thank you for the GOES-16 reference, we have included that as well.**

Page 9. Line 25. Consider adding a GOES-16 COD reference.

We have performed a comparison with the NOAA GOES-R Series Advanced Baseline Imager Level 2 Cloud Optical Depth (ABI-L2-COD). Using only "good_quality_qf" pixels from ABI-L2-COD, geo-spatially aligned with SatCORPS COD for the same analysis period and regions, the NOAA GOES-16 COD reference is on average ~8% greater than SatCORPS and the parameterized COD, even when seemingly normalized to the same scale. This is a large difference and we feel that further investigation is merited, especially in regard to the nuances of scale normalization and handling of saturated pixels, but such efforts are more suitable for a separate paper that is meant to emphasize product comparisons and validations. The goal of the parameterization, after all, is to quickly estimate SatCORPS COD, which has been accomplished. Exploration of the absolute agreement of SatCORPS COD relative to NOAA ABI-L2-COD is beyond the scope of this work. We did, however, add a GOES-16 COD reference (Page 9, Line 27).



Page 11. Line 17. Even more importantly, water vapor BT varies strongly with few angle, much more than an IR longwave window.

We have added discussion and reference of the strong VZA dependency in the WV band for legacy GOES, with specific mention of its significance in the context of historical consistency and comparison to IR window bands (Page 11, Line 38).

Page 13. Line 6. This might be a good spot to reference: Line, W. E., T. J. Schmit, D. T. Lindsey, and S. J. Goodman, 2016: Use of Geostationary Super Rapid Scan Satellite Imagery by the Storm Prediction Center. Wea. Forecasting, 31, 483–494, https://doi.org/10.1175/WAF-D-15-0135.1.

Thank you for the reference, we have included it (Page 13, Line 30).

Page 14. Line 8. Again, is the every other line/element based on 4km pixels? If so, this should be stated for clarity.

As before, this in reference to the Lanczos-interpolated fixed grid, but this time for the IR resolution. Also as before, the actual scale values are dependent on the application and satellite input. E.g., for a GOES-16 IR anvil mask computation, the resampled 2-km IR BT imagery is subset to 4-km by taking every other row and column of the fixed grid. For the application of anvil reflectance aggregation across GOES-13, GOES-15, and Himawari-8, as this paper describes, the IR anvil mask was developed from the 4-km fixed-grid IR BT, subsampled

to 8-km resolution. We have changed the sentence to now offer both of these examples (Page 14, Line 32).

Figure 1. Define what calibration means. Pre-launch? Adjusted for degradation over time? State that the images have been re-mapped to a common projection.

We now define the reflectance as inter-calibrated with reference to MODIS, with further explanation provided in the text. We also now state that the images were remapped to a common projection.

Figure 2. Severe storms based on NOAA storm reports of hail? Wind?

We changed the caption to start as, "Severe wind- and hail-producing storms (as reported by NOAA)…"

Technical corrections

Page 5. Line 22. Typo: founFu, D.rthermore

Typo corrected.

Interactive comment on "A kernel-driven BRDF model to inform satellite-derived visible anvil cloud detection" by Benjamin Scarino et al.

Referee Comments – Martin Setvák, CHMI

03 July 2020

General Comments

The paper focuses on automatic detection of anvils of deep convective clouds (DCC), based on BDRF model developed by the authors. Various anvil detection techniques or products are being used in (satellite-based) nowcasting systems, thus any new similar method can enhance credit of the satellite data. This gains on importance with recent onset of new generations of GEO satellites, such as Himawari-8/9, GOES-R series, FY-4A series, or GEO-COMPSAT-2A, or the upcoming third generation of Meteosat satellites (MTG). For these reasons I welcome the submitted paper and recommend it for publication.

Thank you for your time and insight in reviewing our manuscript. You will find our responses to your comments below with indicated changes in the text where applicable.

Specific Comments

Page 2, Lines 25 – 29

I would be somewhat more conservative about usefulness of the WV-IR BTD method, namely for the overshooting tops (OT) detection. It depends not only on availability of appropriate WV channel and scanning geometry, but for specific cases namely on presence and total amount of water vapor in the lower stratosphere, above the storms, and its vertical thermal profile. Reading this part as it is written now may impose an impression that this technique is broadly used for OT detections, being reliable – which is far from the reality. However, I do not dispute its use for detection of DCC in general.

We agree with your assessment. Although Ai et al. (2017) did demonstrate capability for WV-IR BTD to detect OT, it was neither reliable nor the main focus of their BTD to noise ratio, which rather was DCC detection. Also, given that our manuscript does not directly concern OT, it is probably better to not mention OT in this discussion. Therefore, we have removed the sentence that started, "Ai et al. (2017)… " originally appearing on Page 2, Line 26. We have also muted our emphasis on WV-IR BTD reliability in general (Page 12, Line 4 and Page 14, Line 8) and introduced the fact that WV bands on legacy GOES have very strong VZA dependency (Page 2, Line 30 and Page 11, Line 38)

Page 4, Lines 1 – 2

As written, *its application in enhancing anvil cloud detection (and thereby OT detection) capability*, it may seem that the method can be directly used for OT detection. Though the authors elaborate this statement later in the paper, perhaps a more accurate wording might help here.

In order to better convey that the BRDF model does not directly enhance OT detection, we have adjusted the sentence to read, "… its application in enhancing anvil cloud detection capability and cloud optical depth (COD) parameterization" (Page 4, Line 3).

Page 4, Lines 16 – 17

Can there be any impact of the location of the satellites – Himawari-8 providing data namely for DCC above the ocean, while GOES satellites depicting namely storms above the continent? I'm not speaking here about different underlying surface, but rather about different types and concentrations of the condensation nuclei above continent and oceans, which may affect the cloud top microphysics and thus also its reflectance (BRDF) …

True there are regional differences in DCC reflectance owed to different microphysics, but they should not have a significant impact on our results. Doelling et al. (2018) measured a 0.8% difference in the count response between the TWP region and the Meteosat region from ~2003-2007, which is the largest regional difference they observed. This corresponds to ~0.9% difference in reflectance and is not enough to have meaningful impact on predicted nominal reflectance from the BRDF model.

Page 14, Line 13 and 24

… *should exhibit spatially uniform cold temperature values* … You discuss here the impact of colder overshooting tops, but how about the enclosed warm areas of storms exhibiting cold-Vs or cold rings? How does the algorithm deal with these?

The warm areas of such features are evaluated in the same way as the rest of the anvil using the 22-km moving window. These portions are not warm enough to negate detection, but they are likely to be assigned lesser IR anvil ratings than the colder portions. As far the anvil BRDF is concerned, enclosed warm areas would likely be excluded from the reflectance LUT aggregations because they would not pass the BT homogeneity test. These rare exclusions, however, should not significantly influence the nominal reflectance predicted by the BRDF model, and therefore the final VIS anvil mask results would not be affected. We have added these details to the text on (Page 15, Line 17).

Other comments

I can hardly discuss the technical details of this work (as I have no personal experience in this area), however from an observational perspective and long-term personal experience with satellite observations of storm tops, the individual steps, their settings and parametrization seem to be reasonable and justified. I hope that the authors plan extension (or verification) of this work also to the GOES-16 and GOES-17 data, and possibly also to Meteosat's SEVIRI and future FCI data.

Thank you again for your comments. Extension of the BRDF model of expected anvil reflectance is planned to follow the launch of MTG.

A kernel-driven BRDF model to inform satellite-derived visible anvil cloud detection

Benjamin Scarino[1], Kristopher Bedka[2], Rajendra Bhatt[1], Konstantin Khlopenkov[1], David R. Doelling[2], and William L. Smith Jr.[2]

[1]Science Systems and Applications, Inc., One Enterprise Pkwy Ste 200, Hampton, VA 23666 USA
[2]NASA Langley Research Center, 21 Langley Blvd MS 420, Hampton, VA 23681-2199 USA

Correspondence to: Benjamin R. Scarino (benjamin.r.scarino@nasa.gov)

Abstract. Satellites routinely observe deep convective clouds across the world. The cirrus outflow from deep convection, commonly referred to as anvil cloud, has a ubiquitous appearance in visible and infrared (IR) wavelength imagery. Anvil clouds appear as broad areas of highly reflective and cold pixels relative to the darker and warmer clear sky background, often with embedded textured and colder pixels that indicate updrafts and gravity waves. These characteristics would suggest that creating automated anvil cloud detection products useful for weather forecasting and research should be straightforward, yet in practice such product development can be challenging. Some anvil detection methods have used reflectance or temperature thresholding, but anvil reflectance varies significantly throughout a day as a function of combined solar illumination and satellite viewing geometry, and anvil cloud top temperature varies as a function of convective equilibrium level and tropopause height. This paper highlights a technique for facilitating anvil cloud detection based on visible observations that relies on comparative analysis with expected cloud reflectance for a given set of angles, thereby addressing limitations of previous methods. A one-year database of anvil-identified pixels, as determined from IR observations, from several geostationary satellites was used to construct a bi-directional reflectance distribution function (BRDF) model to quantify typical anvil reflectance across almost all expected viewing, solar, and azimuth angle configurations, in addition to the reflectance uncertainty for each angular bin. Application of the BRDF model for cloud optical depth retrieval in deep convection is described as well.

1 Introduction

Satellite imagery offers a valuable perspective for tracking deep convection that is advantageous in being both spatially broad and contiguous, and also persistent in time. Deep convective clouds (DCC) are comprised of one or more updraft regions, some of which penetrate into the lower stratosphere and are referred to as overshooting tops (OTs), and cirrus outflow emanating from the updrafts that are referred to as anvil cloud (Fujita 1974). Deep convection appears as spatially coherent cold and highly reflective regions in infrared (IR) and visible imagery (McCann 1983; Kirk-Davidoff et al. 1999; Shindell 2001; Brunner et al. 2007; Setvák et al. 2010; Homeyer and Kumjian 2015; Bedka et al. 2015; Bedka et al. 2016). Automated satellite-observed DCC detection based on recognition of these patterns is important for a variety of reasons. Deep convective clouds generate hazardous weather conditions such as heavy rainfall, lightning, aviation turbulence and icing, damaging wind, hail, and tornados (Bedka et al. 2010; Bedka and Khlopenkov 2016; Yost et al. 2018). Forecasters can benefit from satellite-based guidance that can identify these hazardous weather conditions (Gravelle et al. 2016a; Gravelle et al. 2016b). Deep convective clouds are also a common Earth target used for vicarious satellite instrument calibration (Doelling et al. 2013; Doelling et al. 2016; Bhatt et al. 2017a; Bhatt et al. 2017b; Doelling et al. 2018). Furthermore, researchers studying upper troposphere and lower stratosphere composition benefit from knowing where DCC and OT occurred for use in trajectory models (Herman et al. 2016; Smith et al. 2017; Vernier et al. 2018).

The human eye is rather adept at identifying patterns indicative of deep convection – easily being able to locate coherent and circular or elliptical regions of bright, cold, and persistent clouds. Replicating a human-like recognition approach can be problematic, however, because solar illumination and viewing geometry variations affect the appearance of deep convection. That is, a basic pattern recognition algorithm may miss or falsely detect DCC depending on time of day, the angle at which the clouds are viewed, and tropopause temperature. For example, Figs. 1a and 1b show the calibrated visible (VIS) reflectance, based on Clouds and the Earth's Radiant Energy System (CERES) Edition 4 calibration coefficients, of the same mesoscale convective system (MCS) over the Texas panhandle as viewed from GOES-West (GOES-15) and GOES-East (GOES-13), respectively, on 25 May 2015 (Menzel and Purdom 1994; Doelling et al. 2018). Despite spectral consistency ensured by application of DCC-based spectral band adjustment factors and reference to the same calibration standard of Aqua-MODIS, the GOES-West view shows significantly higher reflectance values due to being in the forward-scatter position at 12:30 UTC (Scarino et al. 2017; Doelling et al. 2018). Later, at 23:45 UTC when GOES-East is in the forward-scatter position, the GOES-East view of the mature MCS now appears to be the brighter of the two images (Fig. 1c and 1d). It is easy to see, therefore, how a DCC mask based on a simple reflectance threshold would mischaracterize this storm depending on whether it was viewed from the east or west. The appearance of a storm can vary significantly owing only to illumination and/or viewing conditions, and therefore a justifiable need exists to carefully quantify the expected anisotropic reflectance before visual-based judgements of DCC can be reliably made.

The increased interest in the development of automated, geostationary- (GEO) satellite-sourced means of severe weather and climate analysis in recent years necessitates continual advancement in skillful identification of DCC. Bedka et al. (2010), relied on fixed temperature thresholds based on longwave IR window (~11 μm) observations, which often resulted in seasonal and regional biases. Bedka and Khlopenkov (2016) employed better IR pattern recognition and the addition of a VIS-channel-based (~0.65 μm) algorithm that emulates processes used by humans to identify DCC top anvils and their embedded OTs, which yielded significantly improved detection consistency and quality. The product is valuable for purposes of operational forecasting of severe weather, especially in regions without adequate contiguous weather radar coverage, whether due to isolation, terrain, or influence of national borders. Continued quantification of algorithm accuracy and validation relative to other remotely sensed datasets are key to product development.

Other methods have been introduced that attempt to objectively recognize DCC features. A common technique is the well-documented multispectral band water vapor (WV) minus IR (WV−IR) brightness temperature (BT) difference (BTD) method (Schmetz et al. 1997; Setvák et al. 2007; Martin et al. 2008; Young et al. 2012; Aumann and Ruzmaikin 2013; Ai et al. 2017). Although this approach can be effective, it relies on the presence of a WV channel, which can have significant spectral variation across the global constellation of geostationary imagers or may be absent entirely on some imagers. Furthermore, WV BT is known to vary significantly with viewing angle, by as much as 5 K at 55º relative to nadir view, for legacy GOES imagers, which impacts both the spatial and historical consistency of measurements (CIMSS 2016). For broad applicability, many methods often rely on fixed single-channel BT thresholds, often in the range of 195 K to 225 K, with the threshold dependent on the product application, method, and satellite (Bedka et al. 2010; Young et al. 2012; Doelling et al. 2013; Bhatt et al. 2017a; Bhatt et al. 2017b; Doelling et al. 2018). One drawback of relying on such thresholds is the zonal dependence of cloud-top IR BT. That is, severe storms at mid-to-high latitudes will have warmer cloud tops than comparably severe storms at low/tropical latitudes. Figures 2a and 2b illustrate this point, showing the IR BT of two groups of severe-weather-producing storm systems on 31 August 2018 at 20:47 UTC – one group at higher latitudes of the Contiguous United States (CONUS) over Minnesota, and the other group at lower CONUS latitudes over the Gulf of Mexico and southeastern states. The coldest IR BT value found in Fig. 2a is near 207 K, whereas Fig. 2b shows IR BT approaching 196 K in many areas. As such, a 205-K IR BT threshold (see DCC-

based vicarious calibration methodologies of Doelling et al. and Bhatt et al.), for example, would classify abundant DCC in the southern latitudes, but no DCC at the higher latitudes, despite reports of 2-3-inch (~5.0-7.5-cm) diameter hail in Minnesota. If one normalizes for tropopause height, by computing the IR minus tropopause (IR−Trop) BTD, the apparent intensity of the storms in the north (Fig. 2c), especially the pronounced cell near the western central Minnesota border that produced the severe hail, is more comparable to that of the southern storms (Fig. 2d). Note that the aforementioned vicarious calibration approaches that employ a fixed 205-K threshold are limited to tropical latitudes, and thus suffer no consequence for lack of tropopause normalization (Doelling et al. 2013; Bhatt et al. 2017a; Bhatt et al. 2017b, Doelling et al. 2018). For broader study areas, however, the Fig. 2 example demonstrates the importance of consideration, and potentially normalizing, for one's analysis environment when developing a globally applicable approach to DCC characterization.

Although the OT/anvil detection technique developed by Bedka and Khlopenkov (2016) works reasonably well, the VIS component of the two-channel methodology suffered from a lack of information on expected anvil cloud reflectance for given viewing, solar, and azimuthal angle combinations. Visible-wavelength satellite observations have been routinely used to develop models for bidirectional reflectance distribution function (BRDF) over non-Lambertian land surfaces, which is useful for angle-dependent pattern recognition through provision of expected reflectance values to which measurements are compared. With the proper atmospheric correction (Hu et al. 1999; Lucht et al. 2000; Radkevich 2018), the high-resolution and varied angular sampling retrieval capabilities of instruments like the Moderate Resolution Imaging Spectroradiometer (MODIS), the Multiangle Imaging Spectrometer (MISR), the polarization and directionality of Earth reflectances (POLDER) radiometer, and the Advanced Very High Resolution Spectrometer (AVHRR), have yielded reliable operational land surface BRDF and albedo products for nearly two decades, with significant benefit to the research community (Lucht et al. 2000; Schaaf et al. 2002; Jin et al. 2003; Huang et al. 2013; Vasilkov et al. 2017). These high-resolution retrieval capabilities also extend to modern, advanced GEO satellite imagers, such as that from Himawari-8, thereby granting the benefit of high temporal frequency to multi-angular pattern-analysis-based BRDF modeling, i.e., a kernel-driven concept (Matsuoka et al. 2016). A similar kernel-driven BRDF model technique based on DCC may therefore help normalize VIS-imagery-based anvil cloud identification across almost all illumination and viewing conditions.

Although DCC anvils are, in certain conditions, of the most Lambertian Earth reflectance targets, a BRDF correction is necessary to properly characterize cloud-top surface scattering as a function of illumination and viewing geometry, especially for larger angles (Bhatt et al., 2017b). Unlike the case of land surface BRDF retrieval, which requires atmospheric correction, DCC tops reside near the tropopause, above which absorption effects are relatively small and thus albedo distribution is assumed to be effectively constant month to month (Hu et al. 2004). Owing to these characteristics, a DCC-based VIS BRDF model was developed from CERES and Visible Infrared Scanner (VIRS) observations for the purpose of post-launch calibration of satellite sensors. The main premise of this vicarious calibration approach is that the distribution of DCC albedo remains stable in time, and any temporal shift observed in the DCC reflectance distribution can be attributed to satellite sensor degradation (Hu et al., 2004, Doelling et al., 2013). Bhatt et al. (2017b) expanded this DCC calibration technique to shortwave infrared channels by constructing channel-specific seasonal BRDFs from Suomi National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (SNPP) Visible Infrared Imaging Radiometer Suite (VIIRS) observations and applying the result to the corresponding MODIS bands. The DCC technique allows for characterization of sensor gain stability early in an instrument's lifetime – forgoing the two-year time period necessary for traditional deseasonalization methods, and thereby enabling more-timely calibration stability analyses for any imager with a similar sun-synchronous orbit (Bhatt et al. 2017a; Bhatt et al. 2017b). These studies demonstrate that a DCC-sourced BRDF can accurately predict expected cloud reflectance for a given range of viewing zenith angle (VZA), solar zenith angle (SZA), and relative azimuth angle (RAA) conditions, thereby allowing

for accurate monitoring of satellite imager stability. Expanding such a technique for cloud reflectance prediction to GEO satellites can aid inter-consistency studies and benefit anvil cloud detection.

This article proposes a new kernel-driven BRDF model, which finds its application in enhancing anvil cloud detection capability and cloud optical depth (COD) parameterization. We will describe the satellite-derived data used to formulate the model, as well as explain the development and uncertainty of an anvil reflectance prediction look-up-table (LUT), which shapes the BRDF model. We will show that the kernel-driven approach provides reasonable estimates of expected reflectance for widely varying solar illumination and viewing conditions, thereby promoting consistent identification of anvil cloud regardless of time of day or satellite view. Furthermore, a BRDF of expected reflectance for any viewing condition allows for a quick parameterization of COD based only on the difference between observed reflectance and the model prediction. The parameterization is developed based on multispectral retrieval employed within the NASA Langley Research Center SATellite ClOud and Radiation Property Retrieval System (SatCORPS) framework in support of the CERES project, in which GEO cloud retrieval relies on the CERES Edition 4 algorithm for global cloud detection (Trepte et al. 2019; Minnis et al. 2020). The timeliness and consistency of the anvil detection scheme and related COD parameterization owed to the BRDF model are beneficial to operational forecasting and nowcasting efforts, e.g., convection avoidance or interception for flight routing purposes or airborne science campaigns, which rely on accurate, real-time information.

2 Data and Methods

2.1 Geostationary satellite imagery

A twelve-month database of GOES-13, GOES-15, and Himawari-8 VIS and IR satellite imagery from December 2016 through November 2017 was used to develop the BRDF model. This time period was chosen such that a full seasonal cycle is characterized without influence of orbital shifts, which did occur for GOES-13 when it was replaced by GOES-16 in December of 2017 (Schmit et al. 2018). Half-hourly observations are acquired from sunrise to sunset between 65° N and 65° S, and from 130° W to 30° W for GOES-13, 175° E to 90° W for GOES-15, and 88° E to 178° W for Himawari-8, combining all partial hemisphere, hemisphere, and full disk scanning patterns. It is assumed that using observations across four seasons, high latitudes, diverse regions, and to solar terminator will yield a reasonably full range of possible VZA, SZA, and RAA combinations for observed anvil clouds, with significant statistical population – serving as a strong empirical foundation to shape the BRDF model. All geostationary imagery was acquired from the University of Wisconsin-Madison Space Science and Engineering Center (SSEC). Note that for some analyses, satellite data are supplemented by modeled atmospheric profiles provided by the Global Modeling and Assimilation Office (GMAO) Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) product.

The 0.5-km VIS and 2-km IR Himawari-8 nadir spatial resolution scales were subsampled upon acquisition by skipping every other line and element in order to better match those for GOES-13 and GOES-15. All Himawari-8 and GOES data were then resampled to 1-km and 4-km fixed scales for VIS and IR, respectively. The resampling function is based on Lanczos filtering with the parameter a=3 extended to the two-dimensional case and is applied over an array of 6 × 6 pixels, which is padded with replicated edge-pixel values near image boundaries (Duchon 1979). This interpolation is based on the sinc filter, which is known to be an optimal reconstruction filter for band-limited signals such as digital satellite imagery (Bedka and Khlopenkov 2016). The resampling process typically preserves the BT signal of anvil clouds, which ideally are relatively homogenous across many kilometers. Resampling over the pixel array also acts to dampen any potential small-scale VIS and BT variability in convective

anvils that may otherwise influence construction of the model. Note that hereafter, unless otherwise stated, the term *pixel* refers specifically to individual data samples of the Lanczos-interpolated fixed grid rather than the imager pixel measurements.

The VIS processes take advantage of the higher-resolution (1 km in this case) fixed-scale, but upon output the reflectance measurements are averaged to the IR data resolution. Therefore, in this case the final output resolution is 4 km for all parameters and all satellites. Note that for VIS-related products, such as texture rating, a maximum pixel value is selected for output rather than the average. It is important to state that the scales discussed here are relevant to the combination of satellite data being used for this specific purpose of developing a BRDF model of expected anvil reflectance. In other words, applications that do not involve aggregation with legacy imagers can utilize the full 0.5-km VIS and 2-km IR resolution of third-generation GEO satellites like Himawari-8, GOES-16, and GOES-17. In such cases, the resampling process results in 0.5-km and 2-km fixed grids, respectively, for VIS and IR, VIS processes take advantage of the 0.5-km resolution, and final VIS reflectance or cloud product output are, respectively, averaged or subset to the 2-km IR resolution.

Relative consistency of reflectance observations between the three instruments is ensured by application of CERES Edition 4 VIS imager calibration coefficients for each GEO, which are determined from the monthly gain trends of GEO and Aqua MODIS spectrally-consistent, ray-matched radiance pairs over all-sky tropical ocean, DCC, and invariant desert scenes, based on the best-practices of the Global Space-based Inter-Calibration System (GSICS), and with uncertainty less than 1% (Goldberg et al. 2011; Scarino et al. 2017; Doelling et al. 2018). Furthermore, although relative consistency of BT values is not necessary to develop the BRDF model, IR calibration is based on hourly adjustments to GSICS-referenced VIRS ray-matched pairs (Scarino et al. 2017).

The Meteosat Second Generation (MSG) satellites are not included in this analysis because 1-km VIS imagery is not collected across the entire 65 °N to 65 °S domain throughout a day. Much of the Northern Hemisphere is observed at 1 km but, over the Southern Hemisphere, a moving window of 1-km data is collected, which follows the Sun and captures data during well-illuminated periods of the day. Visible data are only available at 3-km resolution across the full disk view, which would be inconsistent with the GOES and Himawari-based analyses. Given that MSG data is incorporated into the GSICS inter-calibration analysis, we expect that methods developed from GOES and Himawari will perform consistently when applied to MSG data, which is a claim supported by analyses not shown in this paper.

2.2 Multi-angle lookup table for anvil cloud reflectance

A three-dimensional lookup table (LUT) of anvil cloud reflectance was built from pixels classified as anvil using IR observations that satisfy a set of conditions (Bedka and Khlopenkov 2016; Khlopenkov and Bedka 2018). The specifics of the IR-based anvil classification process are described in Appendix A. The three dimensions of the LUT are VZA, SZA, and RAA, yielding the mean anvil reflectance and standard deviation derived from one year of satellite observations. There are 18 bins along each dimension, with 5° bin increments from 2.5° to 87.5° (±2.5°) for both VZA and SZA, and 10° bin increments from 5° to 175° (±5°) for RAA, where 0° RAA is the backscattering angle. Figure 3 illustrates the average anvil reflectance for overhead Sun, for a solar zenith angle near 45°, and approaching early sunset. Although the allowable VZA limit for actual observations is ~88.4°, it should be noted that the maximum observed VZA for this model is ~77.6°, and sampling per VZA bin can suffer with increasing viewing angle in the poleward direction, where deep convection is less likely to be found. Furthermore, because VIS-based observations can become highly shadowed and variable at large SZAs, we caution the use of this method where SZA is greater than 82° despite an allowable LUT limit of up to 90°. Even below this limit, increased sun angle can cause significant shadowing from OT, thereby excluding potential samples (see next paragraph). Therefore, it is not surprising that sample size is sparser, and uncertainty is higher, at the most extreme angular bins, especially when compounded by both high SZA and VZA.

This pattern can be observed in Figs. 4 and 5, which illustrate the bin sampling and reflectance standard deviation (σ), respectively.

Anvil reflectance pixels must satisfy three homogeneity criteria before being included in the LUT. First, each pixel of a 5 × 5 window/array centered on the pixel of interest must initially be classified as an anvil using the IR-based method (see Appendix A). As such, each pixel of the 5 × 5 window must have a rating greater than 0, and the rating of the pixel being considered for the LUT (i.e., the center pixel of the 5 × 5 window) must be greater than 75. These requirements help ensure that each included pixel is in an area of reasonably contiguous anvil detection as determined by the IR method, and is therefore likely to be separated from anvil cloud edges where the quality of anvil detection is more suspect due to semitransparency effects. This 5 × 5 array technique must inherently exclude pixels within two spaces of the image edge, but this is necessary given that the continuity of the anvil cloud beyond the image boundary is unknown, and therefore we are unable to reliably judge those edge pixels.

The IR anvil rating threshold of 75 was chosen for the LUT based on empirical judgement of the relationship between the detection rating and remarkably obvious, spatially coherent cold clouds, which leads to higher certainty in the model. This rating limit of 75 corresponds to an anvil false detection occurrence of less than 1% based on comparison with anvil inferences from CloudSat profiles (see Sect. 2.5). Compared to having no anvil threshold criterion, setting the anvil rating cutoff to 75 reduces the overall dataset size by 6.5% while slightly lowering LUT bin standard deviations, but the magnitude of the LUT does not meaningfully change. Nominally and in this study, both the IR and VIS anvil masks are defined by an anvil detection rating of 15 or more, which is again a determination made based on empirical assessment of satellite imagery from varied regions and seasons. Having different thresholds for defining a mask and developing the LUT of expected reflectance is acceptable because it is critically important that the LUT observations are exceptionally consistent and predictable. As such, with this first criterion we establish a solid foundation for the BRDF model.

The second and third homogeneity criteria require that the standard deviation of the 5 × 5 VIS reflectance array is less than 3% of the 5 × 5 reflectance average, and that the standard deviation of the 5 × 5 IR BT is less than 1 K. These requirements build upon the first criterion by independently quantifying a standard for homogeneity of the surrounding pixels. The thresholds of 3% and 1 K are well-suited for filtering pixels near anvil edges as a secondary check to the anvil continuity test above. More importantly, perhaps, is the ability of these standard deviation checks to exclude anomalous anvil pixels, particularly those that may be associated with OT or gravity waves. These features and other irregularities in the anvil generate localized temperature variability and shadowing effects that would not satisfy the filter thresholds. In the VIS case, comparing the standard deviation to the 5 × 5 average rather than the center pixel reflectance helps mitigate the rare occurrence that an abnormally bright center pixel surrounded by dark anvil pixels will satisfy the homogeneity filter, given that the standard deviation in such a case is likely to be less than 3% of the bright pixel but not less than 3% of the array average.

2.3 Kernel-driven BRDF

Despite the fact that the three-dimensional LUT approach described in the previous section is relatively simple to construct and computationally fast for estimating anvil reflectance across the spectrum of viewing and illumination angles, it suffers from certain drawbacks. The anisotropy of an Earth target is expected to vary continuously with viewing and solar geometry. The finite discretization of angular bins in the LUT, however, can generate sharp discontinuities in the anvil reflectance between neighboring angular bins. The non-uniformity in the sample sizes between bins also impairs the smoother transition of reflectance across the bins resulting in the discontinuous patterns that are seen in the reflectance contour lines of Fig. 3, especially at higher SZA. In addition, the LUT approach is unable to define an anvil reflectance for bins without convection.

Here we describe the construction of a kernel-based BRDF model for characterization of anvil top-of-atmosphere reflectance at continuously varying SZA, VZA, and RAA. This approach not only mitigates the discretization discontinuities that are an effect of the LUT approach, but also fills in the missing intermediate bins. The BRDF is described by a linear superposition of a set of weighting functions, e.g., geometric and optical, that characterize its shape, which is the defining concept of a kernel-driven model (Roujean et al. 1992; Wanner et al. 1995). Despite needing to be flexible enough for application to a variety of inhomogeneous scene types, kernel-driven BRDF models are able to adequately provide description of the anisotropic reflectance of natural surfaces (Hu et al. 1997; Wanner et al. 1997; Hu et al. 1999; Breon and Maignan 2017). The BRDF model is based on the work of Roujean et al. (1992) that describes a bidirectional reflectance R as a linear sum of three kernels in the following form:

$$R(\theta,\psi,\varphi) = K_0 + K_1 f_1(\theta,\psi,\varphi) + K_2 f_2(\theta,\psi,\varphi), \quad (1)$$

where θ, ψ, and φ are the SZA, VZA, and RAA, respectively. Expressions f_1 and f_2 are the model kernels defined as analytical functions of θ, ψ, and φ that represent the geometric and volume scattering components, respectively. These functions are given in the forms:

$$f_1(\theta,\psi,\varphi) = \frac{1}{2\pi}[(\pi-\varphi)\cos\varphi + \sin\varphi]\tan\theta\tan\psi - \frac{1}{\pi}(\tan\theta + \tan\psi + \chi) \quad (2)$$

and

$$f_2(\theta,\psi,\varphi) = \frac{4}{3\pi(\cos\theta+\cos\psi)}\left[\left(\frac{\pi}{2}-\xi\right)\cos\xi + \sin\xi\right] - \frac{1}{3}, \quad (3)$$

where

$$\chi = \sqrt{\tan\theta^2 + \tan\psi^2 - 2\tan\theta\tan\psi\cos\varphi} \quad (4)$$

and

$$\xi = \cos^{-1}(\cos\theta\cos\psi + \sin\theta\sin\psi\cos\varphi). \quad (5)$$

Terms K_0, K_1, and K_2 are scene-specific kernel coefficients that are determined using the least-squares solution of the linear BRDF function for a given set of observations. That is, K_0, K_1, and K_2 are derived from the solution to Eq. (1) for the available empirical data and at the certain angular configuration, thereby providing the best fit for the analytical functions and the observed reflectance of each bin. Furthermore, the coefficients are linearly interpolated across adjacent three-dimensional bins in order to yield an even smoother transition of predicted R across continuous angular variation. Finally, the analytical expressions of f_1 and f_2 are defined such that these terms vanish at nadir viewing and overhead sun conditions. Thus, K_0 is the isotropic component representing the overhead sun reflectance at nadir view. An illustration of this model is seen in Fig. 6, using input data shown in Fig. 3. Figures 7 shows the BRDF difference relative to the LUT, and Fig. 8 shows the model uncertainty.

It is noteworthy that the Roujean et al. BRDF model was originally derived for characterizing surface reflectance. However, previous studies have shown that it is also applicable for anisotropic correction of the TOA reflectance over pseudo-invariant ground sites, provided that the atmospheric effect above the sites is repeatable and predictable over time (Angal et al. 2010; Bhatt et al. 2017b). For anvil measurements from GEO imagers, the fact that atmospheric absorption is minimal above anvils and that the GEO imagers have consistent imaging schedules with repeating angular combinations supports the argument for using the kernel-based approach for modeling TOA anvil reflectance (Hu et al. 2004). In this study, K_0, K_1, and K_2 are computed for each angular bin utilizing the satellite-observed reflectance acquired within the bin supplemented by additional measurements from the neighboring bins that are ±5° apart in SZA and VZA, and ±25° apart in RAA from the center bin. Using the extended set of input data for computing the kernel coefficients ensures that the transition of modeled reflectance is uniform across the bins. In addition, this method also allows for modeling of the bidirectional reflectance for the non-filled bins based on the measurements from surrounding bins. The 1-σ uncertainty of the modeled reflectance for a given angular bin is defined by the standard error of

the regression that is computed for the least-squares fit between the analytical functions and the observed reflectance values during the determination of K_0, K_1, and K_2.

2.4 Visible anvil mask overview

As was noted in the Sect. 1, many satellite-based methods have been developed to identify anvil clouds. These methods typically rely on IR and/or WV absorption-band BT and are perhaps augmented by ancillary information such as tropopause temperature from weather prediction models or reanalysis. Bedka and Khlopenkov (2016) demonstrated a method that incorporates spatial analysis for identification of anvil cloud pixels, but this approach is designed to capture anvil regions near OT and not the entire anvil cloud. A COD threshold-based method for recognizing DCC or anvil features, like that described by Hong et al. (2007), can help to address assumptions of full anvil extents, but because the method is pixel-based and does not incorporate spatial analysis it can be adversely impacted by shadowing due to texture or OT. As such, there existed a need for a reliable VIS-based anvil mask as an important prerequisite to the Bedka and Khlopenkov (2016) VIS texture and OT detection algorithm, in that a search for texture and OT should only occur within the accurate full extents of an anvil cloud.

Anvil reflectance prediction using the method described in Sects. 2.2 and 2.3 coupled with spatial analysis offers an opportunity to address previous limitations and enable efficient texture and OT detection. The VIS texture detection process is based on Fourier analysis of spatial frequencies in the VIS imagery. Spatial frequencies consistent with texture are present not only within convection, but also amongst scattered clouds, and especially near cloud edges. The anvil mask is used to limit the Fourier analysis to only the actual anvil clouds, thereby eliminating false detections associated with other cloud types while also reducing processing time. Classification of the VIS anvil mask is similar to that for the IR mask (Appendix A), except with scoring based on VIS input with reference to the BRDF model.

As described by Bedka and Khlopenkov (2016), the VIS anvil mask is determined by a scoring system that uses an accumulation of histogram-derived information from a nearby ensemble of pixels relative to the assessed pixel. The resampled input VIS reflectance imagery is processed in subsets described by a 50-km-diameter circular window evaluated at every other column and every other row with respect to the Lanczos-interpolated fixed grid. For example, in the case of full-resolution processing for GOES-16, the 0.5-km fixed-grid reflectance is subsampled to 1-km resolution for the purpose of VIS anvil mask determination. The peak of the reflectance-based histogram within each subset is evaluated in a way such that the smooth, uniform signature associated with an anvil cloud is detected, which should ideally exhibit a tall and narrow distribution. An initial VIS anvil rating is constructed based on three main considerations: 1) the width and height of the histogram peak, excluding a possible peak at low-reflectance bins that correspond to clear-sky areas, 2) the difference between the observed reflectance and some nominal reflectance predicted by the BRDF model, and 3) the existence of saturated pixels corresponding to bright OT edges or sun glint from clouds. The second consideration listed above exemplifies the main purpose of this article – marking the major distinction between the VIS anvil mask derivation described here and the original method of Bedka and Khlopenkov (2016), who relied on an empirically-derived function of $\cos(\theta)$ to formulate a nominal anvil reflectance that lacked consideration for bidirectional effects. The nominal reflectance relative to the histogram maximum prominence determines whether a VIS anvil rating of either 8 or 16 is assigned for the assessed subset pixel and surrounding pixels. As such, anvil rating for a given pixel will accumulate as the window moves through the image, until finally the accumulated mask rating is resampled to the original non-subset resolution (Bedka and Khlopenkov 2016).

The result of the initial considerations is a preliminary mask of pixels corresponding to evenly bright areas, although possibly with some saturated pixels. The area for inclusion in the mask is then expanded by 6-10 km (larger expansion is used for higher SZA) to include any regions that potentially contain cloud shadows such as those around OT cores, which may have been missed

by the histogram analysis. A shadowed pixel is then included in the mask if it is sufficiently surrounded by pixels defined in the preliminary mask. Finally, the expanded mask is multiplied by a scaled difference between the tropopause temperature and the pixel BT. This last step ensures that the resulting anvil mask corresponds to sufficiently cold areas that match the actual extents of an anvil cloud, and furthermore indicates that there remains an IR component to the VIS mask despite its designation.

2.5 Comparison with CloudSat anvil cloud detection

Anvil detections from GOES-16 based on the VIS mask, the IR mask, the WV−IR BTD method, and a tropopause-normalized IR temperature threshold method (i.e., IR−Trop BTD) were validated against independent determinations of anvil clouds from CloudSat using ~800 CloudSat granules from January, April, July, and October of 2018. The CloudSat definition of anvil is based on the method of Young et al. (2013) and relies on the 2B-CLDCLASS product. Following their technique, an anvil cloud is determined when high/cirriform clouds are connected to within 33 product profiles of a vertical DCC cluster of sufficient depth, provided that the region below the high clouds is cloud free or only partially filled by single-layer, low-level clouds (Young et al. 2013). Receiver operating characteristics (ROC) curves, which report true positive rate against false positive rate, are then determined relating the rate of agreement with CloudSat anvil indications (probability of anvil detection) within 5 minutes using each of the four methods listed above to the rate of false alarms (positive indications that disagree with CloudSat). Note that because of the nature of CloudSat measurements (two-dimensional vertical profiles along the satellite track), these comparisons can only be considered validations in a relative sense, rather than an absolute sense. That is, because the CloudSat profile must encounter a DCC cluster within 33 profiles of high/cirriform clouds in order for the clouds to be assessed as anvils, it is possible that true anvil clouds remain unidentified by CloudSat simply because the DCC cluster associated with them was not along the scan path. Therefore, false alarm rates may be incorrectly inflated in this validation approach. Also, given the ~13:30 local equator crossing time of CloudSat, these validation results are only representative of low-SZA conditions. Nevertheless, being that all four methods are assessed with these same limitations, we believe their relative comparison remains fair.

2.6 Cloud optical depth parameterization

A simple parameterization for anvil COD based on the difference between observed (Obs) VIS reflectance and BRDF-model-predicted anvil reflectance was developed. This feature is important for nowcasting, i.e., flight routing for airborne science campaigns, because it allows for rapid estimation of COD based on readily available input, whereas multi-band cloud retrieval algorithms, such as those of CERES/SatCORPS or NOAA (Minnis et al. 2011; GOES-R 2018; Minnis et al. 2020), require ancillary datasets, additional preliminary computations, and longer processing time (e.g., cloud masking is required before COD retrieval). That is, this parameterization can approximate imager multi-band-retrieved COD in a matter of seconds rather than minutes, which is significant for real-time weather applications. The parameterization is developed based on the SatCORPS COD product, and thus will introduce additional error on top of the uncertainty of those retrievals. Therefore, the approximation should not be used as a replacement for the true retrievals but rather should be employed as a general estimate of COD when timeliness is the chief concern.

The approximation is defined by an exponential fit of COD as a function of Obs minus BRDF (Obs−BRDF) reflectance, calculated as a function of SZA, with twenty-eight 3° SZA bins from 0°-3° to 81°-84°. Figure 9 shows the 0°-3°, 45°-48°, and 78°-81° fits in order to highlight how the shape of the functional relationship of COD to Obs−BRDF reflectance changes with SZA. The fits are based on calibrated VIS reflectance (Scarino et al. 2017; Doelling et al. 2018) and 4-km ice-cloud COD using the Minnis et al. (2020) CERES MODIS Edition 4 methodology, which has been adapted to the SatCORPS GEO framework for

GOES-16 imagery over the CONUS in July 2018 (Schmit et al. 2018). The VIS and COD pixels are analyzed only where they are coincident with the VIS anvil mask, which was derived from resampled 0.5-km GOES-16 VIS imagery and output at 2-km resolution. Note that ice-cloud COD derived in the SatCORPS framework is effectively insensitive to whether retrieval resolution is 1 km, 2 km, or 4 km (Minnis et al. 2016). Rather than being fit to the entirety of the dataset that satisfies each SZA bin, the two-term exponential model is simply guided by the maximum density of data found along the curve, as indicated by black circles in Fig. 9. Fitting as such prevents influence from outliers and bad retrievals, and therefore better models the most common functional relationship between COD and Obs−BRDF reflectance.

The COD parameterization consistency was evaluated relative to its SatCORPS reference by comparing with SatCORPS GOES-16 COD that is independent from the training dataset, derived from daytime imagery of Hurricane Florence on 11 September 2018. This date was chosen because Florence maintained Category 4 intensity throughout the day and thus sustained a large area of persistent anvil cloud across the full spectrum of SZA, which is unlike land-based convection that typically exist for a few hours in the late afternoon. With land-based convection, it is difficult to distinguish whether variations in COD are due to increasing/decreasing storm intensity, or if they are being caused by a parameterization that is dependent on SZA. It is important to note that although an intense hurricane should have relatively consistent COD throughout a day when averaged across the storm anvil, variations in reflectance associated with spiral band development or eyewall replacement do occur, which can appear in our data.

3 Results and discussion

3.1 Features of the kernel-driven model

As was described in Sect. 2.3, the most significant result of the kernel-driven BRDF is the mitigation of discretization discontinuities between adjacent LUT angular bins and completion of bidirectional reflectance for non-filled bins based on measurements from surrounding bins. These effects are apparent, as the patterns seen in Fig. 6 are smoother and more coherent than those of Fig. 3, with gaps filled, thereby creating a more complete three-dimensional model. Keep in mind that the kernel-driven approach fills gaps in LUT based on an interpolation scheme that draws not only from neighboring VZA and RAA bins, but also from adjacent SZA increments. This is the reason why Fig. 6c, for example, appears exceptionally more continuous than Fig. 3c despite an apparent lack of valid VZA and RAA bins with which to interpolate from to complete the model as shown. The kernel-driven model allows for filling of angular bins to exactly one bin beyond the valid coverage of the LUT.

Figure 7 reveals, in a qualitative sense, the amount of smoothing accomplished by the kernel model, with the difference pattern highlighting the chaotic nature of the observation-based LUT as owed to sampling inadequacies. The kernel-driven BRDF is largely consistent with the LUT, excepting smoothing differences, at low- and mid-SZA positions (Figs. 7a and 7b, respectively) as indicated by pale purple (0-1% difference) to dark blue (2-3% difference) shading. The largest divergences of the model from the LUT are at high-SZA positions and where sampling is low. That is, the largest differences shown in Fig. 7b align with areas of low sampling or high uncertainty in Figs. 4b and 5b, respectively. The large differences in Fig. 7c are a result of the greater amount of interpolation required given the discrete nature of the LUT in this volume, which has a much higher associated 1-σ uncertainty compared to that of the lower SZA conditions (compare Fig. 5c to Figs. 5a and 5b), and also overall higher uncertainty than that of the model (Fig. 8). Note that the 1-σ uncertainty of Fig. 5 is based on the actual samples and average of each LUT bin, whereas in Fig. 8, recall that 1-σ uncertainty comes from the standard error of the regression computed for the least-squares fit between the analytical functions and the observed reflectance values, which benefits from the incorporation of measurements from surrounding bins and thus has lower uncertainty for each bin compared to that of the same bin from the LUT,

assuming the LUT bin is filled. Despite the uncertainties in high-SZA conditions, the kernel-driven BRDF model exemplifies significant improvements over the simple LUT in terms of anvil characterization because continuous, smooth transitions across bin thresholds lends to a more realistic pattern of anvil reflectance.

3.2 Anvil mask comparisons

A robust anvil mask based on VIS interpretation should perform similarly regardless of perceived changes in cloud brightness owed to viewing or illumination conditions, which is the motivation behind the BRDF model. This premise was introduced earlier with Fig. 1, where the same cloud structures were viewed simultaneously with either GOES-East or GOES-West in the forward-scatter position. Figure 10 revisits that imagery, but now with the VIS anvil mask (Sect. 2.4) determined from either GOES-East or GOES-West indicated with a red line. In either case (12:30 UTC or 23:45 UTC), the general shapes of the masks are comparable despite rather large apparent differences in the calibrated reflectance values, especially in the 12:30 UTC example (Figs. 10a and 10b). Fine-scale differences in the shapes of the red line are present but overall the masks are in agreement.

Another way to qualitatively evaluate anvil mask performance is through comparison with other IR-based methods described above. Examples of these methods are provided in Fig. 11, which shows developing thunderstorms near Kansas and Missouri observed by GOES-16. Figure 11a overlays the IR anvil mask (Appendix A), which had been used to construct the VIS anvil reflectance database for the LUT, and the VIS anvil mask described above. The VIS mask identifies the bright DCC whereas the IR mask extends beyond the boundaries of the VIS mask into pixels that are still cold but not as bright. Even in the northeast portion of the image where the explanation for disagreement is perhaps more questionable, close examination reveals that the VIS mask is outlining a narrow region of bright cloud whereas the IR mask identifies less cohesive shapes. Both approaches define the anvil, and the preferable approach depends on application and tolerances of those that may use the data. Commercial aircraft, for instance, may choose to avoid any indication of anvil cloud out of an abundance of caution. In contrast, airborne research field campaign mission planners, such as those during the High Ice Water Content – Radar missions (Bedka et al. 2019), were interested in only the coldest and most reflective clouds.

The VIS and IR masks aim for self-consistency and reliability but with broad applicability, independent of changes in time, location, or satellite source. The drawbacks of other anvil identification methods were mentioned in Sect. 1, some of which are highlighted and expanded upon in Figs. 11b-11d. For instance, Fig. 11b showcases an example of why DCC or anvil characterization based solely on simple IR BT thresholds is problematic. A 205-K threshold, which is suitable for the tropics, performs poorly for this CONUS scene. In this case, a 225-K threshold is generally more consistent with the VIS and IR mask results.

Rather than using a simple threshold, one can normalize IR BT by the MERRA-2 tropopause temperature reanalysis (i.e., "TROPT" from the "inst1_2d_asm_Nx" dataset), which is interpolated in time and space to the satellite pixels and accounts for latitudinal dependencies in cloud top height/temperature. Figure 11c shows the difference of GOES-16 IR BT and the MERRA-2 tropopause temperature (IR−Trop BTD). In terms of anvil detection, it is unclear what IR−Trop BTD threshold would perform best. A BTD < 10 K provides results consistent with the VIS and IR masks. Less-positive BTD values will gradually restrict the mask to only smaller areas in and around OT regions, whereas greater positive values can easily yield false detections.

The WV−IR BTD method for anvil detection is also demonstrated (Fig. 11d). Aside from the similar question of which BTD best defines anvil clouds, there is a more practical limitation of this technique to consider. Because the WV spectral response functions vary across imagers, WV−IR BTD uniformity across multiple sensors is difficult to achieve. Perhaps more importantly, especially from the perspective of historical consistency, is the fact that certain legacy GOES suffer from strong VZA

dependency in the WV band – significantly more-so compared to that of IR window bands (CIMSS 2016). Furthermore, the signal-to-noise ratio is low for cold WV BT, and the fairly broad WV channels limit the WV−IR BTD detection of anvil clouds to the most extreme cases (Ai et al. 2017). Also, having a WV channel is not guaranteed for common sensors, whether historical instruments such as AVHRR or for newer instruments such as VIIRS. Therefore, although the WV−IR BTD approach is relatively accurate, it is not a consistent option. As such, given that the VIS and IR mask techniques rely only on two standard channels, the methods are more portable and dependably applicable to past, current, and future platforms.

Figure 12 offers a more quantitative comparison of the relative performance of each anvil identification technique with respect to the CloudSat reference (see Sect. 2.5). Probability of anvil detection as defined by CloudSat, or agreement of each identification method with CloudSat, is related to the false alarm rate, or occurrences when a method identifies an anvil but CloudSat does not. Naturally, as the definition of what qualifies as an anvil becomes more restrictive by each method's standards (e.g., VIS mask anvil defined by 1 or greater vs. 20 or greater, or WV−IR BTD anvil defined by -8 K or greater vs. 0 K or greater), the probability of detection decreases. Similarly, less restrictive definitions expectedly result in greater false alarm rates. Therefore, in an ideal case, the perfect anvil definition would maximize probability of detection while minimizing false alarm rate and would thereby be situated as near as possible to the top left corner of the ROC graph. In the context of this relative comparison, the VIS mask offers comparable or slightly better probability of anvil detection and lower false alarm rate than any of the other three methods. That is, overall the four methods are quantifiably similar in effectiveness. The VIS mask, however, when applicable in low-SZA conditions, has perhaps a slight advantage in anvil designation. For high SZA or VZA conditions, in which shadowing can be prevalent, IR-centric methods are preferred.

3.3 Hurricane Florence cloud optical depth parameterization

Morning (10:52 UTC top row), mid-day (16:42 UTC center row), and evening (21:47 UTC bottom row) GOES-16 observed reflectance (left column), COD derived from Obs−BRDF anvil reflectance (center column), and COD from SatCORPS (right column) values for Hurricane Florence are shown in Fig. 13. These morning, mid-day, and evening views of Florence have average SZAs of about 79.1°, 23.8°, and 79.8°, respectively. The SZA values are computed from the mean SZA within the red contour, which signifies the 34-kt wind radii (one radius per quadrant for a total of four 34-kt radii). The NOAA National Hurricane Center provides wind quadrant radii for maximum sustained wind values of 34, 50, and 64 kts. The coordinates defining each quadrant extend in the NE, SE, SW, and NW directions, radiating outward from the center of the storm to a distance where the indicated windspeed is expected to be possible. This study uses these radii as a basis for COD evaluation, with each one drawn on the Fig. 13 imagery in red (34 kts), magenta (50 kts), and blue (64 kts).

In the morning, the high SZA creates shadows on Florence cloud tops due to OT and gravity waves, seen most prominently near the eyewall but also within the outflow shield (Fig. 13a). Compared to mid-day (Fig. 13d), the overall observed reflectance is lower at around 0.7 to 0.8 on average, rather than ≥ 0.9 when the Sun is higher overhead. According to the BRDF model for this morning angular combination (Fig. 14), the predicted anvil reflectance is between 0.71 and 0.81 – gradually increasing from the southwest to the northeast part of the image. The COD derived by the SZA-dependent functional relationship (Fig. 9), namely Fig. 13a minus Fig. 14, is shown in Fig. 13 b. The areas where observed reflectance is much less than predicted reflectance is where low COD values are expected as signified by the darkest shades. The veracity of these dark shades within the outflow shield is questionable, however, because it is unlikely that COD is fluctuating so rapidly across these short distances where deep convective clouds are located. In other words, shadowing and texture generates severely under-estimated COD within three-dimensional cloud-top structures. This pattern, however, is consistent with SatCORPS results (Fig. 13c), which is our current baseline for comparison as the parameterization is dependent on the SatCORPS reference. Bedka et al. (2019) show that

reflectance smoothing prior to COD computation dampens shadowing effects at high SZA, which results in a more spatially consistent product. That smoothing approach, however, was purposefully excluded for this study. Observed reflectance that is only slightly less than predicted values signifies a greater COD, shown in gray shades. Cloud optical depth grows exponentially as the observed reflectance matches and surpasses the predicted anvil reflectance. Note that despite the potential for the exponential function to predict excessive COD values given rather modest changes in either reflectance value, COD is capped at a maximum value of 150 to be consistent with SatCORPS output.

At high SZA, the exponential growth of COD with Obs−BRDF is rather extreme and therefore uncertainty is high. Based on Fig. 9c, when Obs−BRDF is close to 0, a combined 0.05 error in observed or predicted reflectance could amount to the difference between 70 and 150 COD, or an 80-COD unit variance. At such early and late times in the day (as Figs. 13g, 13h, and 13i behave similarly to 13a, 13b, and 13c) the function is steep and therefore highly sensitive to reflectance uncertainty, not to mention the increased standard error of the fit itself due to variable SatCORPS retrievals at these high-SZA conditions. On the other hand, at midday (Figs. 13d, 13e, and 13f), the exponential function is less steep, with a shape somewhere between that of Figs. 9a and 9b. Here the BRDF only varies between 0.91 and 0.94 across the entire image (even less across the wind radii). A combined 0.05 error close to where Obs−BRDF is 0 in this case results in about 10-COD variance, and with lower standard error of the fit compared to the previous case. Exactly how impactful 80-COD error would be in near-sunset applications, or 10-COD error during midday operations, is dependent on the product application. For a simple and immediate means of estimating the broad-scale COD conditions, however, this method performs well relative the more computationally intensive, although likely overall more accurate, SatCORPS multi-band retrieval method.

The Obs−BRDF parameterized COD is compared to SatCORPS COD within the 34-, 50-, and 64-kt radii of Hurricane Florence using the daytime (SZA < 82°) 5-minute imagery from GOES-16 on 11 September 2018. The COD results as a function of UTC hour for each radii set are shown in Fig. 15, with Obs−BRDF COD in red and SatCORPS COD in blue. The SZA is also displayed above the x-axis of each plot. Overall, the Obs−BRDF COD agrees rather well with SatCORPS throughout the day, which is reassuring given that the parameterization was developed using independent SatCORPS COD data. The mean COD differences between Obs−BRDF and SatCORPS (Obs−BRDF minus SatCORPS) for the 34-, 50-, and 64-kt radii are 1.9%, 0.5%, and, 0.9%, respectively. The agreement with SatCORPS is encouraging, especially near 80°-82° SZA where COD differences are around 11 at worst and 0 at best, because it validates a consistency in approach that is independent of viewing and illumination conditions, at least to the extent that SatCORPS is similarly independent, which is the purpose of the BRDF model.

4. Summary

Operational forecasting of severe and aviation weather, especially in regions without adequate contiguous weather radar coverage, can benefit significantly from rapid and highly detailed imaging offered by geostationary satellite measurements (Line et al. 2016). Consistent imagery-based identification of severe weather indicators, such as deep convective updrafts, anvil clouds, and OT, can be difficult to achieve, however. This article highlights a kernel-driven BRDF model for informed prediction of anvil reflectance, which helps anvil cloud detection efforts and ultimately improves the two-channel, passive satellite imager OT detection algorithm. A satellite VIS-based detection algorithm that incorporates predicted anvil reflectance for known angular conditions is able to more consistently identify DCC anvils on the scale afforded by satellite imagery, regardless of viewing and solar conditions, which is beneficial to a variety of stakeholders.

The kernel-driven BRDF model, which is described by a linear superposition of a set of geometric and optical weighting functions, is employed to characterize the anvil top-of-atmosphere reflectance at continuously varying SZA, VZA, and RAA.

This approach effectively mitigates discretization discontinuities and fills missing intermediate bins in the LUT, thereby creating a reliable model for predicted anvil reflectance. Despite lingering uncertainties at high-SZA positions, the kernel-driven BRDF model improvement over the LUT is significant because continuous, smooth transitions across discrete angular bins results in a more natural pattern of predicted reflectance, which benefits anvil characterization efforts.

The VIS mask is a more conservative approach in anvil detection than the initial IR mask. The masks may disagree on what strictly defines the limit of an anvil, but either offers advantages depending on the application. These techniques are also independent of geographic considerations, unlike methods based on static IR thresholds, and are not as susceptible to false positives that arise due to ill-defined BTD allowances such as is the case for IR−Trop-based anvil detection. The WV−IR BTD method is a well-documented technique for anvil classification, but consistent identification across the entire constellation of geostationary satellites is not guaranteed. Therefore, a two-channel approach based on widely available VIS and IR imagery grants broader applicability and dependability for well-calibrated imagers. Furthermore, in a relative comparison study with a CloudSat anvil reference, the VIS anvil mask offered better skill in anvil identification for low-SZA conditions than any of the IR-centric methods.

By subtracting BRDF-model-predicted anvil reflectance from observed VIS reflectance we are able to develop a simple parameterization for anvil COD. An SZA-dependent exponential fit of SatCORPS-derived COD as a function of Obs−BRDF reflectance defines the parameterization, which produces an approximation of SatCORPS COD but with significantly less computational demand. The exponential growth of COD with Obs−BRDF is rather extreme at high SZA, and thus COD is sensitive to small changes in observed or predicted reflectance. Regardless, for a simple and immediate means of estimating the broad-scale COD conditions that is at least comparable to SatCORPS to within ~1% on average, this parameterization works well.

Acknowledgments

This work was supported by the NASA Applied Sciences Disasters Program under award number 18-DISASTER18-0008. We thank Douglas Spangenberg (SSAI at NASA LaRC) for assisting with the data processing. We also thanks Patrick Minnis for his valuable insight in regard to the SatCORPS cloud products. The angular-dependent kernel coefficients are available upon request to those looking to compute specific bidirectional reflectance values.

Appendix A: Calculating the IR anvil mask

The steps below describe how spatial infrared temperature patterns are quantified to identify convective anvil clouds in the form of an "anvil mask." The anvil mask is a rating that indicates a confidence in anvil detection, with values above 10-15 roughly corresponding to human perception of anvil cloud extents and values above 100 indicating a high level of confidence. Quantification of anvil confidence is a crucial step in development of the kernel-driven BRDF.

Pixel-level IR BT data are first subtracted from the local tropopause temperature in order to obtain the brightness temperature difference (BTD) relative to the tropopause. This BTD is processed in circular subsets of 22-km diameter extracted at every other column and every other row with respect to the Lanczos-interpolated fixed grid. For example, in the case of full-resolution processing for GOES-16, the 2-km fixed-grid IR BT is subsampled to 4-km resolution for the purpose of anvil mask determination. For the application of anvil reflectance aggregation across GOES-13, GOES-15, and Himawari-8, as described in this paper, the IR anvil mask was developed from the 4-km fixed-grid IR BT, subsampled to 8-km resolution. The local distribution of BTD within each subset is analyzed by constructing a histogram H having N=32 bins and covering the range from

-35 K (i.e., warmer than the tropopause), which is a low tropopause-relative bound for anvil clouds, to 13 K colder than the tropopause, which only occurs in updraft regions. Pixels colder than the 13-K threshold are accumulated in the last histogram bin. Figure A1 shows examples of BTD histograms calculated for a typical anvil cloud within a convective system (red columns) and for a region outside the anvil cloud (blue columns).

The BTD within anvils should exhibit spatially uniform cold temperature values, which in most cases will result in a sharply peaked histogram. As such, it follows that anvil rating should be made proportional to the peak height H_i, which indicates the number of counts in the i-th bin, and therefore is also proportional to that bin's index i given that higher bin indices correspond to colder regions, e.g., Fig. A1. The following formula, refined through extensive testing, describes the dependence of anvil rating r_{anvil} on index i:

$$r_{anvil} = \frac{0.35}{D^2} \cdot H_i \cdot i \cdot (2N + 8 - i). \quad \text{(A1)}$$

Here, D is the diameter of the histogram window in pixels and the term in parentheses acts to gradually flatten the $r_{anvil}(i)$ dependence at higher levels of confidence in anvil detection as BTD reaches zero and becomes strongly positive. Based on the example shown in red in Fig. A1, using an 11-pixel diameter D at 2-km pixel resolution, the peak in the red histogram at bin 20 has a height of 23, which yields an IR anvil rating of 69. In most cases the formula above describes uniformly cold anvil clouds reasonably well. If non-uniform regions around OT cores are causing the histogram peak to split over several bins, the major peak can be counted together with neighboring bins to make the total contribution equivalent to a single strong peak, thereby lending stability in resultant r_{anvil} across the whole anvil. Note that in the case of enclosed warm areas within an anvil (e.g., an enhanced-v or cold ring), the IR anvil rating is likely to be lower in the warmer portions but should not be negated completely (McCann 1983; Brunner et al. 2007; Bedka et al. 2013). Such areas may be excluded during construction of the anvil reflectance prediction look-up-table, but the influence on the nominal reflectance predicted by the BRDF model should be minimal, and thus resultant VIS anvil ratings should not be affected.

Finally, the obtained anvil mask has to be expanded in order to include pixels along the anvil boundary, where there is only partial anvil coverage in the subsetting window. This expansion is implemented by raising the rating for all anvil pixels inside the 22-km circular window that have BTD larger than 7.5 K below the histogram's peak. Their anvil rating is increased to reach the level of the peak bin. After this spatial expansion, the IR anvil mask presents a reasonable match relative to the actual anvil cloud region, with the anvil extents filled with the nearly uniform field of r_{anvil}.

Although this method attempts to identify uniform cloud areas near the tropopause, it does not guarantee that a broad area of extremely cold cloud is indeed an anvil cloud. For example, a large area of cold jet stream cirrus in a winter storm may be assigned a significant anvil rating if a local histogram happens to have sufficient criteria. This leaves some room for improving the IR anvil rating, for instance by 1) incorporating a difference relative to the regional background in order to help define convective environments (i.e., cold cloud vs. much warmer clear sky background), or 2) using model-derived atmospheric instability indices, such as convective available potential energy, to restrict detections to regions where deep convection is assumed to be possible. Nevertheless, practical experience with developing the mask, the graphical examples shown in this paper, and comparisons with CloudSat indicate that the IR anvil mask performs reasonably well and is suitable for constructing the BRDF model.

References

Ai, Y., Li, J., Shi, W., Schmit, T. J., Cao, C., and Li, W.: Deep convective cloud characterizations from both broadband imager and hyperspectral infrared sounder measurements, J. Geophys. Res., 122, 1700–1712, doi:10.1002/2016JD025408, 2017.

Angal, A., Xiong, X., Choi, T., Chander, G., Wu, A.: Using the Sonoran and Libyan desert test sites to monitor the temporal stability of reflective solar bands for Landsat 7 ETM+ and Terra MODIS sensors, J. Appl. Remote Sens. 4(1), 043525, 2010.

Bedka, K. M., Brunner, J., and Feltz, W.: Objective overshooting top and enhanced-V signature detection for the GOES-R Advanced Baseline Imager: Algorithm theoretical basis document, Available online at: http://clouds.larc.nasa.gov/site/people/data/kbedka/GOES-R_ABI_ATBD_OvershootingTop_Enhanced-V_100perc.doc, 2010.

Aumann, H. H. and Ruzmaikin, A.: Frequency of deep convective clouds in the tropical zone from 10 years of AIRS data, Atmos. Chem. Phys., 13, 10795–10806, https://doi.org/10.5194/acp-13-10795-2013, 2013.

Bedka, K, Brunner, J., Dworak, R., Feltz, W., Otkin, J., and Greenwald, T.: Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients, J. Appl. Meteor. And Climatol., 49, 181–22, doi: 10.1175/2009JAMC2286.1, 2010.

Bedka, K. M., Wang, C., Rogers, R., Cerey, L., Feltz, W., and Kanak, J.: Examining deep convective cloud evolution using total lightning, WSR-88D, and GOES-14 super rapid scan datasets, Wea. Forecasting, 30, 571–590, doi:10.1175/WAF-D-14-00062.1, 2015.

Bedka, K. M., and Khlopenkov, K.: A probabilistic pattern recognition method for detection of overshooting cloud tops using satellite imager data, J. Appl. Meteor. And Climatol., 55, 1983–2005, doi: 10.1175/JAMC-D-15-0249.1, 2016.

Bedka, K. M., Yost, C., Nguyen, L., Strapp, W., Ratvasky, T., Khlopenkov, K., Scarino, B., Bhatt, R., Spangenberg, D., Palikonda, R.: Analysis and automated detection of ice crystal icing conditions using geostationary satellite datasets and in situ ice water content measurements, Proceedings of the 2019 SAE International Conference on Icing of Aircraft, Engines, and Structures, 2019.

Bhatt, R., Doelling, D. R., Angal, A., Xiong, X., Scarino, B., Gopalan, A., Haney, C., and Wu, A.: Characterizing response versus scan-angle for MODIS reflective solar bands using deep convective clouds, J. Appl. Remote Sens., 11, 016014, 2017.

Bhatt, R., Doelling, D. R., Scarino, B., Haney, C., and Gopalan, A.: Development of seasonal BRDF models to extend the use of convective clouds as invariant targets for satellite SWIR-band calibration, Remote Sens., 9, 1061; doi:10.3390/rs9101061, 2017.

Breon, F-.M. and Maignan, F.: A BRDF–BPDF database for the analysis of Earth target reflectances, Earth Syst. Sci. Data, 9, 31–45, doi:10.5194/essd-9-31-2017, 2017.

Brunner, J. C., Ackerman, S. A., Bachmeier, A. S., and Rabin, R. M.: A quantitative analysis of the enhanced-V feature in relation to severe weather, Wea. Forecasting, 22, 853–872, 2007.

Doelling, D. R., Haney, C., Bhatt, R., Scarino, B., and Gopalan, A.: Geostationary visible imager calibration for the CERES SYN1deg Edition 4 product, Remote Sens., 10, 288, doi:10.3390/rs10020288, 2018.

Doelling, D. R., Haney, C., Morstad, D., Scarino, B. R., Bhatt, R., and Gopalan, A.: The characterization of deep convective clouds as an invariant calibration target and as a visible calibration technique, IEEE Trans. Geosci. Remote Sens., 51, 1147–1159, 2013.

Doelling, D. R., Haney, C., Scarino, B. R., Gopalan, A., and Bhatt, R.: Improvements to the geostationary visible imager ray-matching calibration algorithm for CERES Edition 4, JTECH, 33, 2679–2698, 2016.

Duchon, C. E.: Lanczos filtering in one and two dimensions, J. Appl. Meteor., 18, 1016–1022, doi:https://doi.org/10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2, 1979.

Feng, Z., Dong, X., Xi, B., Schumacher, C., Minnis, P., and Khaiyer, M.: Top-of-atmosphere radiation budget of convective core/stratiform rain and anvil clouds from deep convective systems, JGR, 116, D23202, doi:10.1029/ 2011JD016451, 2011.

Fujita, T. T.: Overshooting thunderheads observed from ATS and Learjet, Satellite and Mesometeorology Research Project Rep. 117, Texas Tech University, Lubbock, TX, 29 pp, 1974.

Goldberg, M., Ohrin, G., Butler, J., Cao, C., Doelling, D. R., Gaertner, V., Hewinson, T., Iacovazzi, B., Kim, D., Kurino, T., Lafeuille, J., Minnis, P., Renaut, D., Schmetz, J., Tobin, D., Wang, L., Weng, F., Wu, X., Yu, F., Zhang, P., and Zhu, T.: The global space-based inter-calibration system (GSICS), Bull. Am. Meteorol. Soc., 92, 467–475, 2011.

Gravelle, C. M., Mecikalski, J. R., Line, W. E., Bedka, K. M., Petersen, R. A., Sieglaff, J. M. , Stano, G. T., and Goodman, S. J.: Demonstration of a GOES-R satellite convective toolkit to ''bridge the gap'' between severe weather watches and warnings: An example from the 20 May 2013 Moore, Oklahoma, tornado outbreak, Bull. Amer. Meteor. Soc., 97, 69–84, doi:10.1175/BAMS-D-14-00054.1, 2016.

Gravelle, C. M., Runk, K. J., Crandall, K. L., and Snyder, D. W.: Forecaster evaluations of high temporal satellite imagery for the GOES-R era at the NWS operations proving ground, Wea. Forecasting, 31, 1157–1177, doi:10.1175/WAF-D-15-0133.1, 2016.

GOES-R Algorithm Working Group and GOES-R Program Office: NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 2 Cloud Optical Depth (COD), ABI-L2-COD, NOAA National Centers for Environmental Information. doi:10.7289/V58G8J02, last access 8 July 2020, 2018.

Herman, R. L., Ray, E. A., Rosenlof, K. H., Bedka, K. M., Schwartz, M. J., Read, W. G., Troy, R. F., Chin, K., Christensen, L. E., Fu, D., Stachnik, R. A., Bui, T. P., and Dean-Day, J. M.: Enhanced stratospheric water vapor over the summertime continental United States and the role of overshooting convection, Atmos. Chem. Phys., 17, 6113-6124, doi: 10.5194/acp-17-6113-2017, 2017.

Homeyer, C. and Kumjian, M. R.: Microphysical characteristics of overshooting convection from polarimetric radar observations, J. Atmos. Sci., 72, 870–891, doi:10.1175/JAS-D-13-0388.1, 2015.

Huang, X., Jiao, Z., Dong, Y., Zhang, H., and Li, X.: Analysis of BRDF and albedo retrieved by kernel-driven models using field measurements, IEEE J-STARS, 6, 1, 149–161, 2013.

Hu, B., Lucht, W., Li, X., and Strahler, A. H.: Validation of kernel-driven models for global modeling of bidirectional reflectance, Remote Sens. Environ., 62, 201–214, 1997.

Hu, B., Lucht, W., and Strahler, A. H.: The interrelationship of atmospheric correction of reflectances and surface BRDF retrieval: A sensitivity study, IEEE Trans. Geosci. Remote Sens., 37, 724–738, 1999.

Hu, Y., Wielicki, B., Yang, P., Stackhouse, P., Lin, B., and Young, D.: Application of deep convective cloud albedo observations to satellite-based study of terrestrial atmosphere: Monitoring stability of space-borne measurements and assessing absorption anomaly, IEEE Trans. Geosci. Remote Sens., 42, 2594–2599, 2004.

Hong, G., Yang, P., Huang, H. L., Baum, B. A., Hu, Y. X., and Platnick, S.: The sensitivity of ice cloud optical and microphysical passive satellite retrievals to cloud geometrical thickness, IEEE T. Geosci. Remote, 45, 1315–1323, 2007.

Jin, Y., Schaaf, C. B., Gao, F., Li, X., and Strahler, A. H.: Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 1. Algorithm performance, J. Geophys. Res., 108, D5, 4158, doi:10.1029/2002JD002803, 2003.

Kirk-Davidoff, D. B., Hintsa, E. J., Anderson, J. G., and Keith, D. W.: The effect of climate change on ozone depletion through changes in stratospheric water vapour. Nature, 402, 399–401, doi:10.1038/46521, 1999.

Khlopenkov K. and Bedka K.: Development of pattern recognition algorithms to detect intense convective storms from multispectral satellite imagery, IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium; IEEE Xplore: November 2018; doi: 10.1109/IGARSS.2018.8518596, 2018.

Line, W. E., Schmit, T. J., Lindsey, D. T., and Goodman, S. J.: Use of geostationary super rapid scan satellite imagery by the storm prediction center, Wea. Forecasting, 31, 483–494, https://doi.org/10.1175/WAF-D-15-0135.1, 2016.

Lucht, W., Schaaf, C. B., and Strahler, A. H.: An Algorithm for the retrieval of albedo from space using semiempirical BRDF models, IEEE Trans. Geosci. Remote Sens., 38, 977–998, 2000.

Martin, D.W., Kohrs, R. A., Mosher, F. R., Medaglia, C. M., and Adamo, C.: Over-ocean validation of the global convective diagnostic, J. Appl. Meteor. Climatol., 47,525–543, https://doi.org/10.1175/2007JAMC1525.1, 2008.

Matsuoka, M., Takagi, M., Akatsuka, S., Honda, R., Nonomura, A., Moriya, D., and Yoshioka, H.: Bidirectional reflectance modeling of the geostationary sensor Himawari-8/AHI using a kernel-driven BRDF model, ISPRS Annals Proc., III-7, XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic, 2016.

McCann, D. W.: The enhanced-V: A satellite observable severe storm signature, Mon. Wea. Rev., 111, 887–894, 1983.

Menzel, W. P., and Purdom, J. F. W.: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. Bull. Amer. Meteor. Soc., 75, 757–781, 1994.

Minnis, P., Bedka, K., Trepte, Q., Yost, C. R., Bedka, S. T., Scarino, B., Khlopenkov, K., and Khaiyer, M. M.: A consistent long-term cloud and clear-sky radiation property dataset from the Advanced Very High Resolution Radiometer (AVHRR), Climate Algorithm Theoretical Basis Document (C-ATBD), CDRP-ATBD-0826 AVHRR Cloud Properties - NASA, NOAA CDR Program, 159 pp., DOI:10.7289/V5HT2M8T, 2016.

Minnis, P., Sun-Mack, S., Chen, Y., Chang, F.-L., Yost, C. R., Smith Jr., W. L., Heck, P. W., Arduini, R. F., Bedka, S. T., Yi, Y., Hong, G., Jin, Z., Painemal, D., Palikonda, R., Scarino, B., Spangenberg, D. A., Smith, R. A., Trepte, Q. Z., Yang, P., and Xie, Y.: CERES MODIS cloud product retrievals for Edition 4, Part I: Algorithm changes, IEEE Trans. Geosci. Remote Sens., early on-line, 2020.

Minnis, P., Sun-Mack, S., Young, D. F., Heck, P. W., Garber, D. P., Chen, Y., Spangenberg, D. A., Arduini, R. F., Trepte, Q. Z., Smith Jr., W. L., Ayer, J. K., Gibson, S. C., Miller, W. F., Chakrapani, V., Takano, Y., Liou, K.-N., Xie, Y., and Yang, P.: CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part I: Algorithms, IEEE Trans. Geosci. Remote Sens., 49, 4374-4400, 2011.

Radkevich, A.: A method of retrieving BRDF from surface-reflected radiance using decoupling of atmospheric radiative transfer and surface reflection, Remote Sens., 10, 591; doi:10.3390/rs10040591, 2018.

Roujean, J.-L., Leroy, M., and Deschamps, P. Y.: A bi-directional reflectance model of the Earth's surface for the correction of remote sensing data, J. Geophys. Res., D-97, 20455–20468, 1992.

Scarino, B., Doelling, D. R., Haney, C., Bedka, K., Minnis, P., and Gopalan, A.: Utilizing the precessing orbit of TRMM to produce hourly corrections of geostationary infrared imager data with the VIRS sensor, SPIE Proc., 10403, Infrared Remote Sensing and Instrumentation XXV; 104030H (2017) https://doi.org/10.1117/12.2273883, 2017.

Scarino, B. R., Doelling, D. R., Minnis, P., Gopalan, A., Chee, T., Bhatt, R., Lukashin, C., and Haney, C. O.: A web-based tool for calculating spectral band difference adjustment factors derived from SCIAMACHY hyperspectral data, IEEE Trans. Geosci. Remote Sens., 54, 5, 2529–2542, 2016.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sens. Env., 83, 135–148, 2002.

Schmetz, J., Tjemkes, S. A., Gube, M., and van de Berg, L.: Monitoring deep convection and convective overshooting with METEOSAT, Adv. Space Res., 19, 433–441, 1997.

Schmit, T. J., Lindstrom, S. S., Gerth, J. J., and Gunshor, M. M.: Applications of the 16 spectral bands on the Advanced Baseline Imager (ABI), J. Operational Meteor., 6, 4, 33–46, doi: https://doi.org/10.15191/nwajom.2018.0604, 2018.

Setvák, M., Lindsey, D.T., Novák, P., Wang, P.K., Radová, M., Kerkmann, J., Grasso, L., Su, S.-H., Rabin, R.M., Šťástka, J., Charvát, Z., Kyznarová, H.: Cold-ring-shaped cloud top features atop convective storms, Atmos. Res. (ISSN: 0169-8095) 97, 80–96, http://dx.doi.org/10.1016/ j.atmosres.2010.03.009, 2010.

Setvák, M., Rabin, R. M., and Wang, P. K.: Contribution of the MODIS instrument to observations of deep convective storms and stratospheric moisture detection in GOES and MSG imagery, Atmos. Res., 83, 505–518, 2007.

Shindell, D. T.: Climate and ozone response to increased stratospheric water vapour, Geophys. Res. Lett., 28, 1551– 1554, doi:10.1029/1999GL011197, 2001.

Smith, J. B., Wilmouth, D. M., Bedka, K. M., Bowman, K. P., Homeyer, C. R., Dykema, J. A., Sargent, M. R., Clapp, C. E., Leroy, S. S., Sayres, D. S., Dean-Day, J. M., Bui, T. P., and Anderson, J. G.: A case study of convectively sourced water vapor observed in the overworld stratosphere over the United States, J. Geophys. Res., 122, 9529–9554, doi: doi.org/10.1002/2017JD026831, 2017.

Trepte, Q. Z., Minnis, P., Sun-Mack, S., Yost, C. R., Chen, Y., Jin, Z., Hong, G., Chang, F.-L., Smith, W. L., Bedka, K. M., and Chee, T. L.: Global cloud detection for CERES edition 4 using Terra and Aqua MODIS data, IEEE Trans. Geosci. Remote Sens., 57, 9410–9449, doi: 10.1109/TGRS.2019.2926620, 2019.

Vasilkov, A., Qin, W., Krotkov, N., Lamsal, L., Spurr, R., Haffner, D., Joiner, J., Yang, E.-S., and Marchenko, S.: Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: a new approach based on geometry-dependent Lambertian equivalent reflectivity applied to OMI algorithms, Atmos. Meas. Tech., 10, 333–349, doi:10.5194/amt-10-333-2017, 2017.

Vernier, J., Fairlie, T. D., Deshler, T., Venkat Ratnam, M., Gadhavi, H., Kumar, B. S., Natarajan, M., Pandit, A. K., Akhil Raj, S. T., Hemanth Kumar, A., Jayaraman, A., Singh, A. K., Rastogi, N., Sinha, P. R., Kumar, S., Tiwari, S., Wegner, T., Baker, N., Vignelles, D., Stenchikov, G., Shevchenko, I., Smith, J., Bedka, K., Kesarkar, A., Singh, V., Bhate. J., Ravikiran, V., Durga Rao, M., Ravindrababu, S., Patel, A., Vernier, H., Wienhold, F. G., Liu, H., Knepp, T. N., Thomason, L., Crawford, J., Ziemba, L., Moore, J., Crumeyrolle, S., Williamson, M., Berthet, G., Jégou, F., and Renard, J.: BATAL: The Balloon measurement campaigns of the Asian tropopause aerosol layer, Bull. Amer. Meteor. Soc., 99, 955–973, https://doi.org/10.1175/BAMS-D-17-0014.1, 2018.

View angle considerations for GOES I–M imagers: http://cimss.ssec.wisc.edu/goes/calibration/GOES12_IMGR_LZAvsTEMP.jpg, last access 6 July 2020, 2016.

Wanner, W., Strahler, A. H., Hu, B., Lewis, P., Muller, J.-P., Li, X., Barker Schaaf, C. L., and Barnsley, M. J.: Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm, J. Geophys. Res., 102, D14, 17143–17161, 1997.

Wanner, W., Li, X., and Strahler, A. H.: On the derivation of kernels for kernel-driven models of bidirectional reflectance, J. Geophys. Res., I00, 21077–21090, 1995.

Yost, C. R., Bedka, K. M., Minnis, P., Nguyen, L., Strapp, J. W., Palikonda, R., Khlopenkov, K., Spangenberg, D., Smith Jr., W. L., Protat, A., and Delanoe, J.: A prototype method for diagnosing high ice water content probability using satellite imager data, Atmos. Meas. Tech., 11, 1615–1637, doi: 10.5194/amt-11-1615-2018, 2018.

Young, A. H., Bates, J. J., and Curry, J. A.: Complementary use of passive and active remote sensing for detection of penetrating convection from CloudSat, CALIPSO, and Aqua MODIS, J. Geophys. Res., 117, D13205, doi:10.1029/ 2011JD016749, 2012.

Young, A. H., Bates, J. J., and Curry, J. A.: Application of cloud vertical structure from CloudSat to investigate MODIS-derived cloud properties of cirriform, anvil, and deep convective clouds, J. Geophys. Res., 118, 4689–4699, 2013.

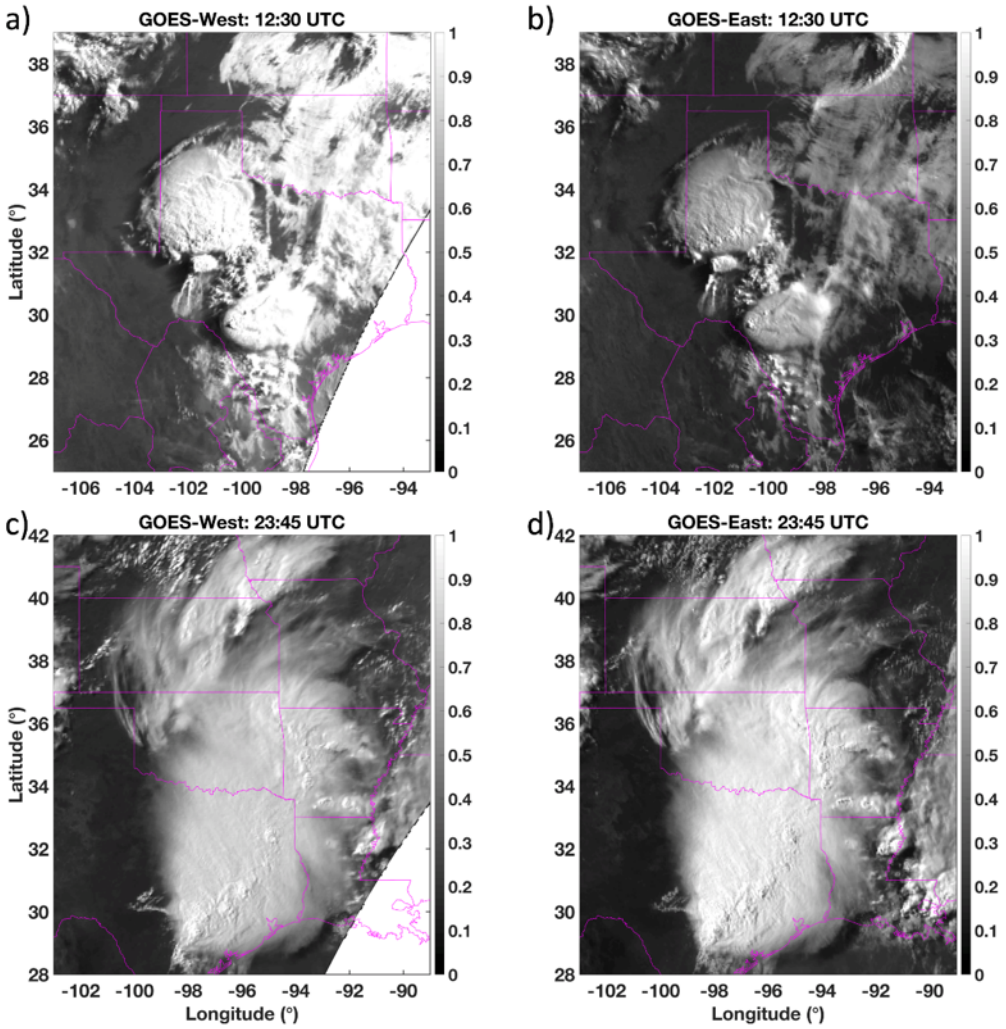


Figure 1: Inter-calibrated, MODIS-referenced VIS reflectance of a 25 May 2015 MCS over Texas and Oklahoma as viewed by a) GOES-West (GOES-15) at 12:30 UTC, b) GOES-East (GOES-13) at 12:30 UTC, c) GOES-West at 23:45 UTC, and d) GOES-East at 23:45 UTC, remapped to a common projection. The apparent brightness of the MCS changes as the solar illumination and viewing conditions vary.

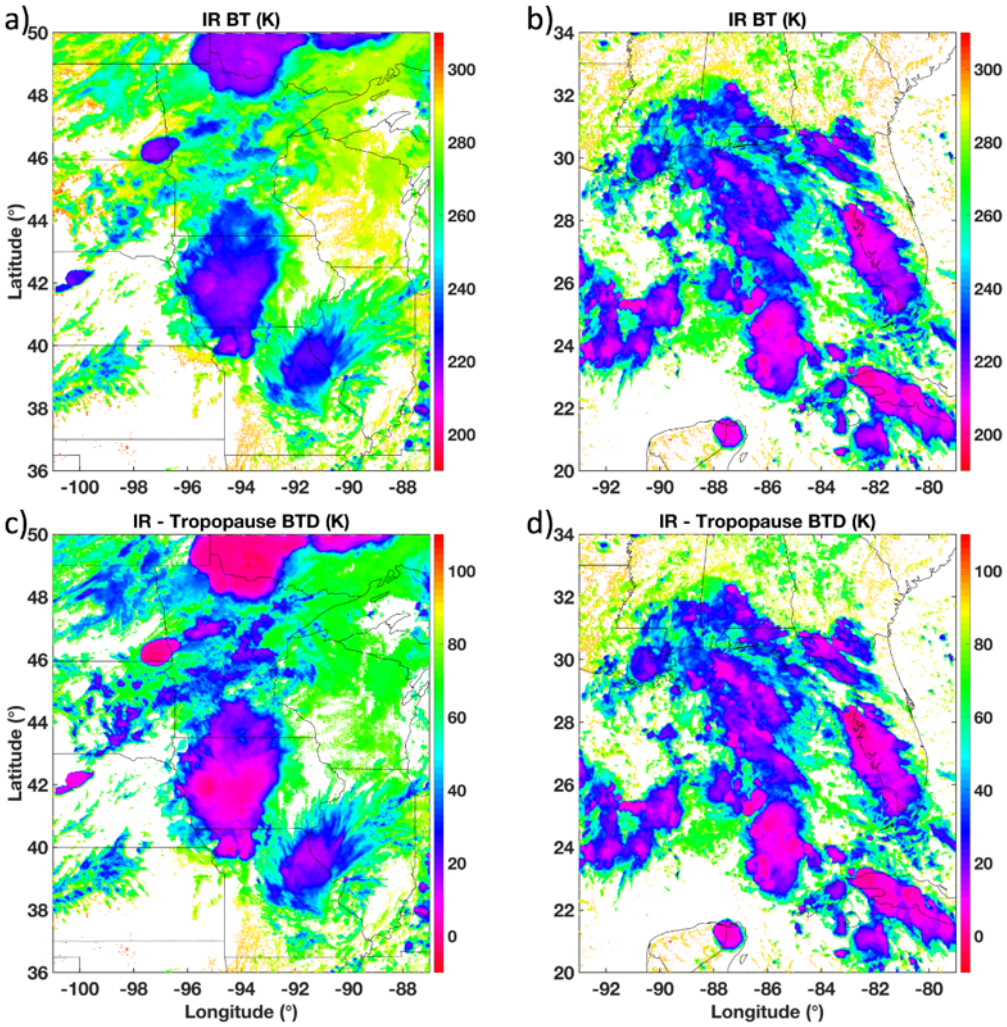


Figure 2: Severe wind- and hail-producing storms (as reported by NOAA) on 31 August 2018 at 20:47 UTC visualized by GOES-16 a) IR BT at northern CONUS latitudes, b) IR BT at southern CONUS latitudes, c) IR−Trop BTD at northern CONUS latitudes, and d) IR−Trop BTD at southern CONUS latitudes. Without tropopause normalization, the relative intensity of the northern storms appears less than that of the southern storms despite both being significant producers of severe weather.

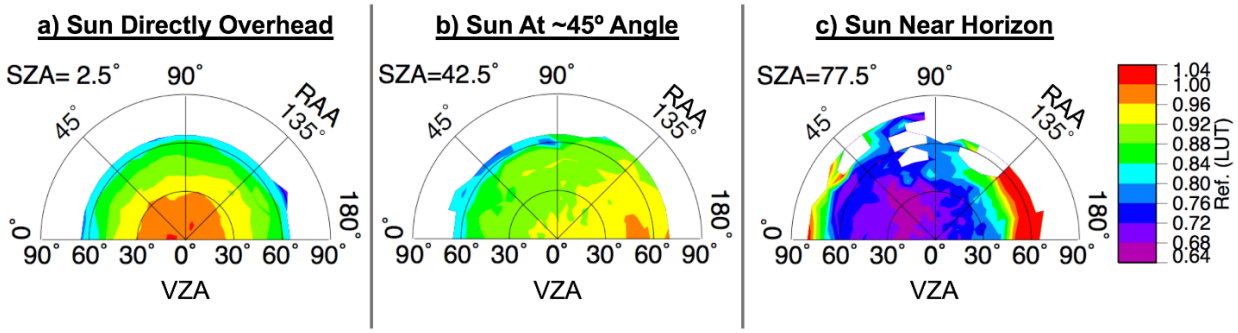


Figure 3: Illustration of LUT for average anvil reflectance as a function of VZA, SZA, and RAA, based on 2016 December through 2017 November GOES-13, GOES-15, and Himawari-8 retrievals. Polar plots are shown for the a) 2.5º, b) 42.5º, and c) 77.5º SZA bins (±2.5°). The radial coordinates of each plot indicate the change in VZA, demarcated into 18 bins with 5° bin increments from 2.5° to 87.5° (±2.5°). The polar coordinates of each plot indicate the change in RAA, demarcated into 18 bins with 10° bin increments from 5° to 175° (±5°), where 0° RAA is the backscattering angle. Gaps at certain bin indices indicate a lack of anvil sampling for that angular configuration.

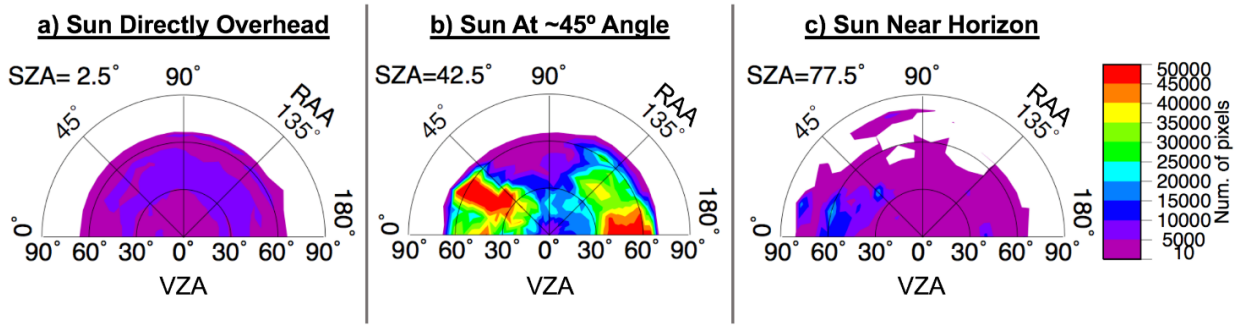


Figure 4: Sampling distribution for the LUT shown in Fig. 3.

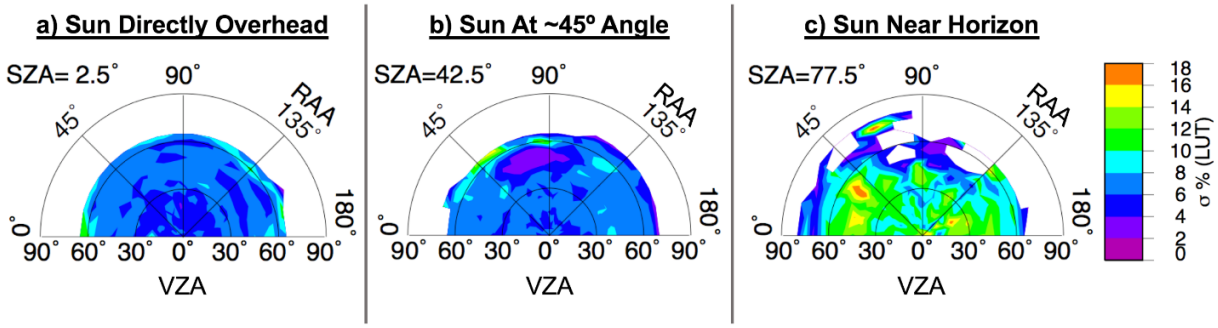


Figure 5: Uncertainty of the LUT shown in Fig. 3 given as the standard deviation percentage of the average bin reflectance.

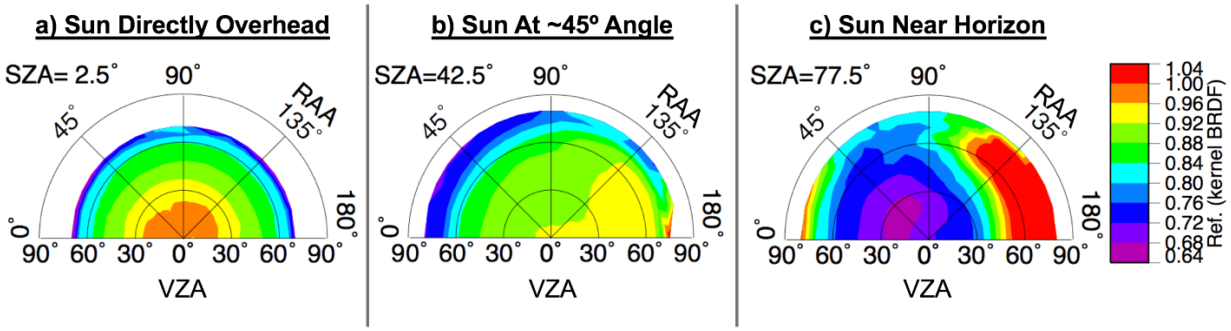


Figure 6: A semi-empirical kernel-based BRDF model of anvil top-of-atmosphere reflectance at continuously varying SZA, VZA, and RAA. The coordinate system is the same as that described in Fig. 3.

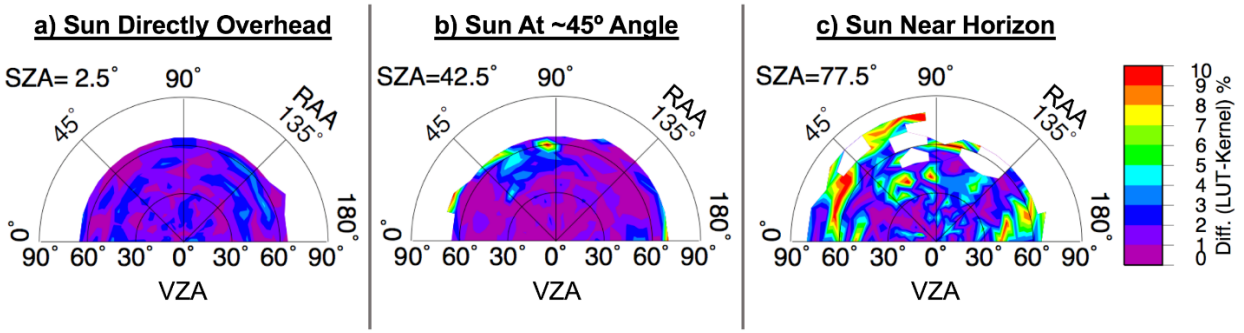


Figure 7: Percent difference in predicted anvil reflectance between the LUT shown in Fig. 3 and the kernel-based BRDF model shown in Fig. 6.

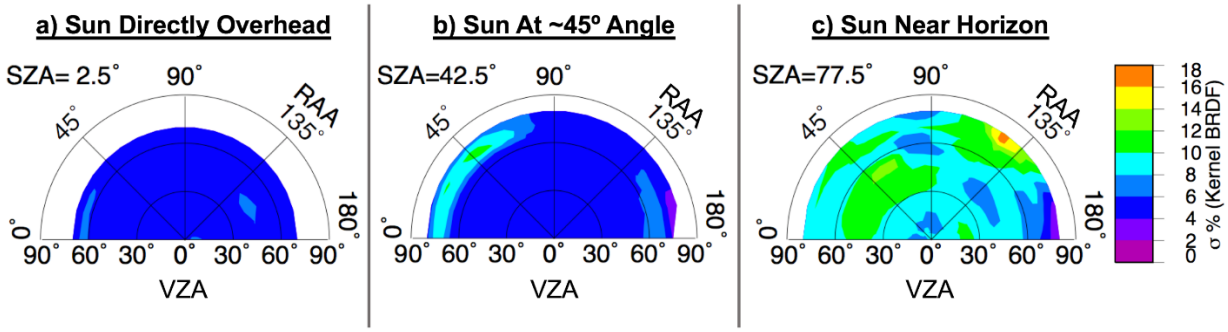


Figure 8: Uncertainty of the kernel-based BRDF model shown in Fig. 6 based on the standard error of the regression computed for the least-squares fit between modeled and observed reflectance values.

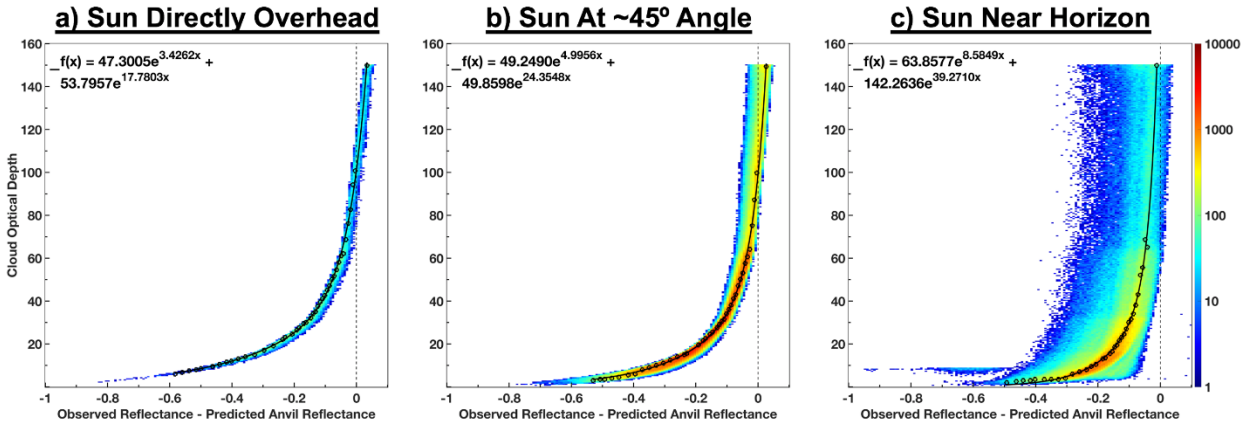


Figure 9: Cloud optical depth as a function of SatCORPS Obs−BRDF based on July 2018 CONUS retrievals for SZA ranges of a) 0º to 3º, b) 45º to 48º, and c) 78º to 81º.

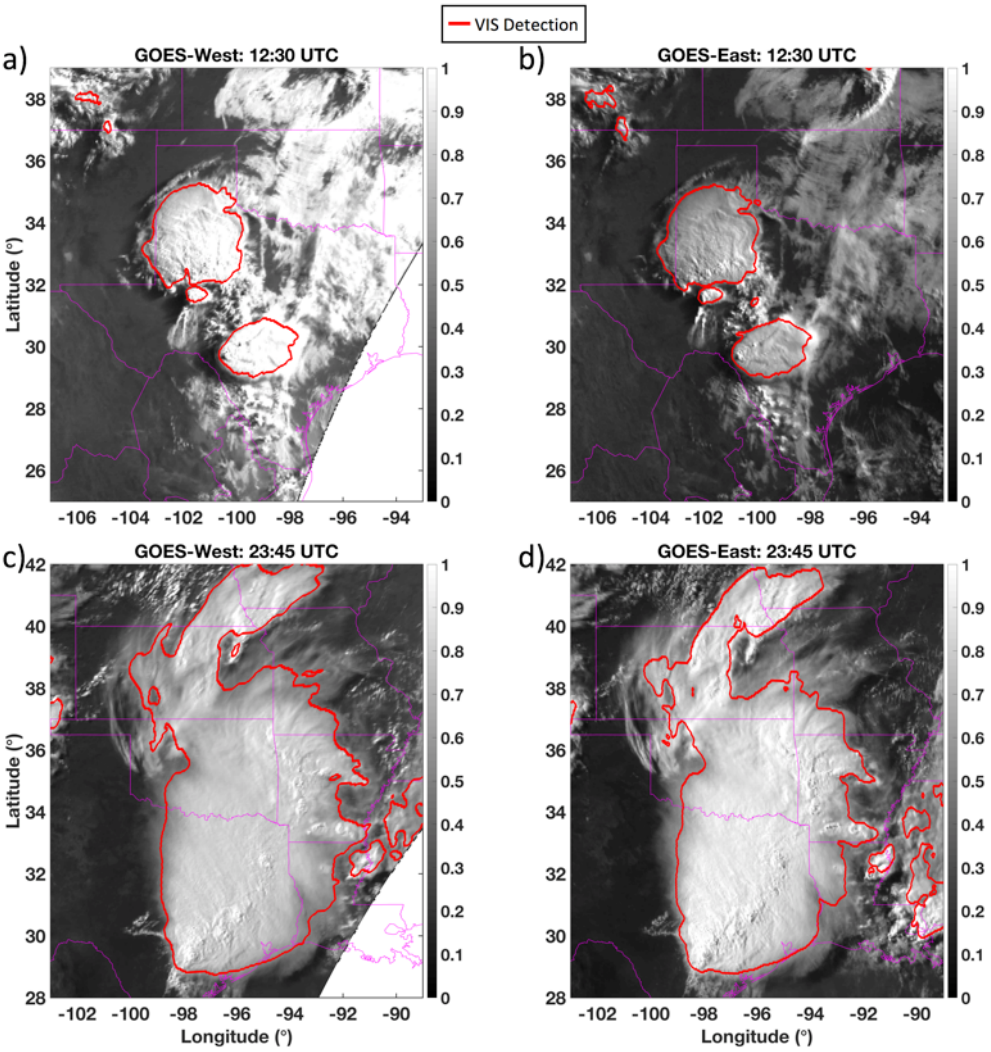


Figure 10: Same as Fig. 1 except with the GOES-West or GOES-East VIS anvil mask outlined in red. The masks from each satellite are similar at the corresponding times despite extreme viewing and illumination differences.

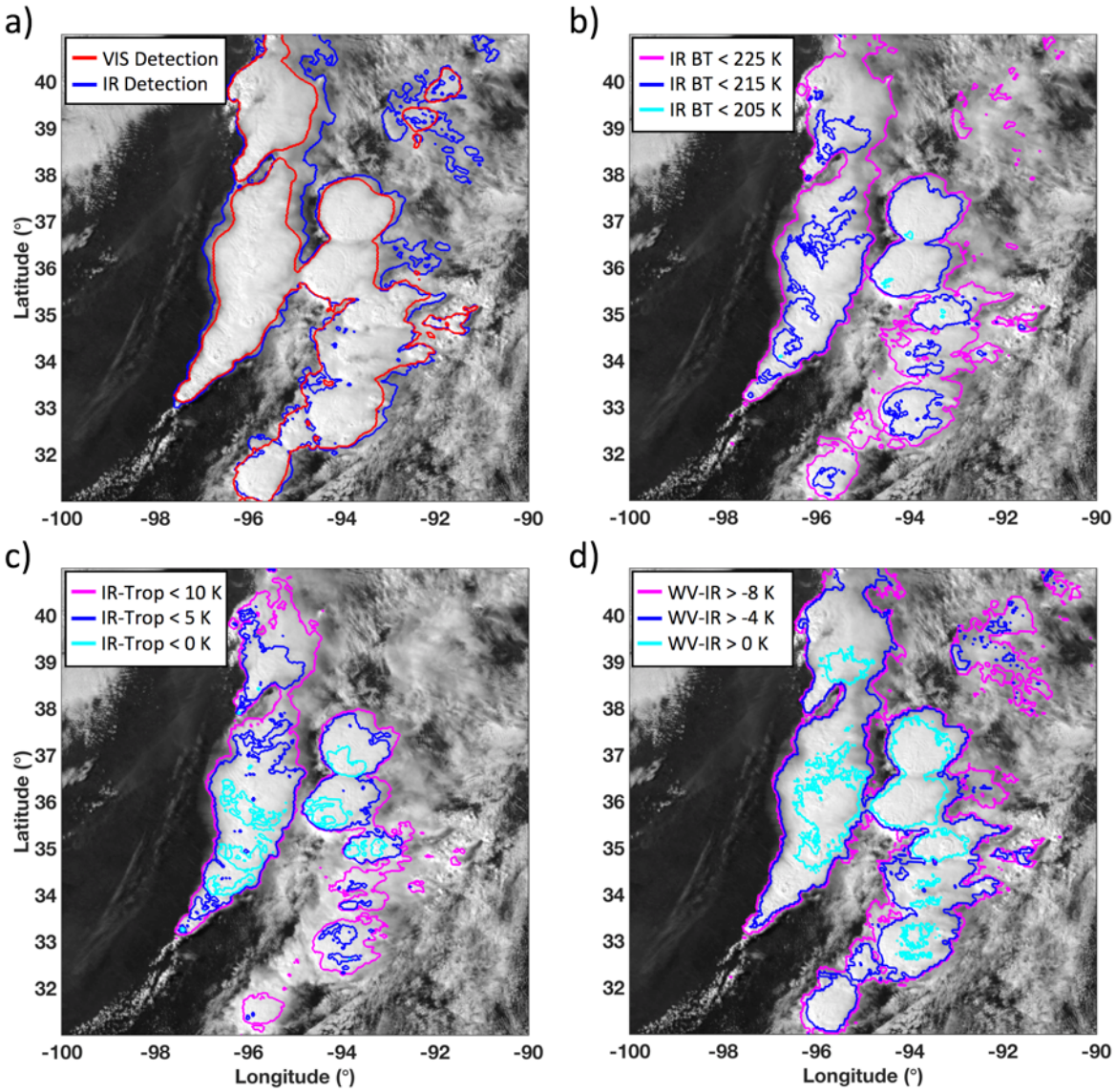


Figure 11: A cluster of developing thunderstorms near Kansas and Missouri on 13 April 2018 viewed from GOES-16 at 20:45 UTC overlaid with a) the VIS and IR anvil masks, b) contours for IR BT less than 225, 215, and 205 K, c) contours for IR−Trop BTD less than 10, 5 and 0 K, and d) contours for WV−IR BTD greater than -8, -4, and 0 K.

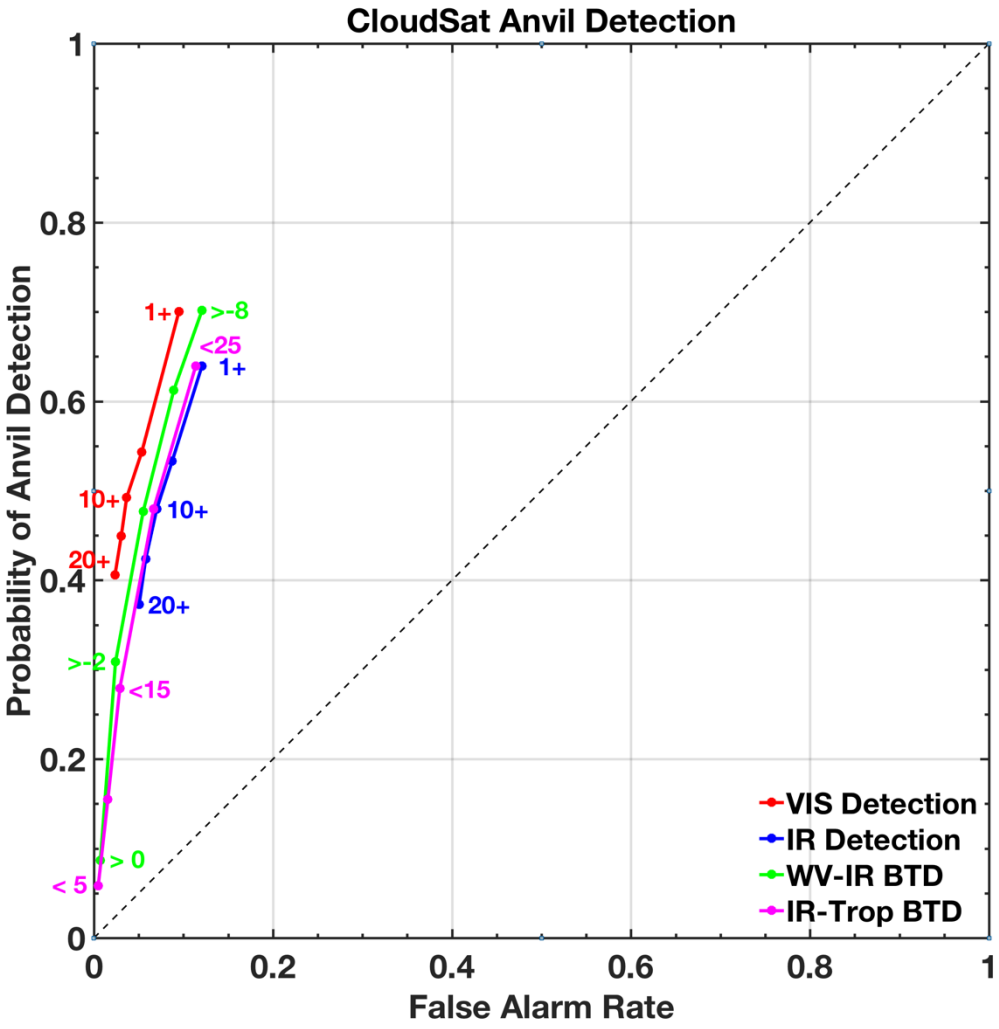


Figure 12: Receiver operating characteristic curves highlighting the rate of CloudSat anvil detection based on the VIS mask, the IR mask, the WV−IR BTD test, and the IR−Trop BTD test relative to the rate of false alarms. The VIS and IR masks are evaluated from ratings of 1 or greater up to ratings of 20 or greater, the WV−IR BTD test is evaluated from differences of -8 K or greater to 0 K or greater, and the IR−Trop BTD test is evaluated from differences of 25 K or less to 5 K or less.

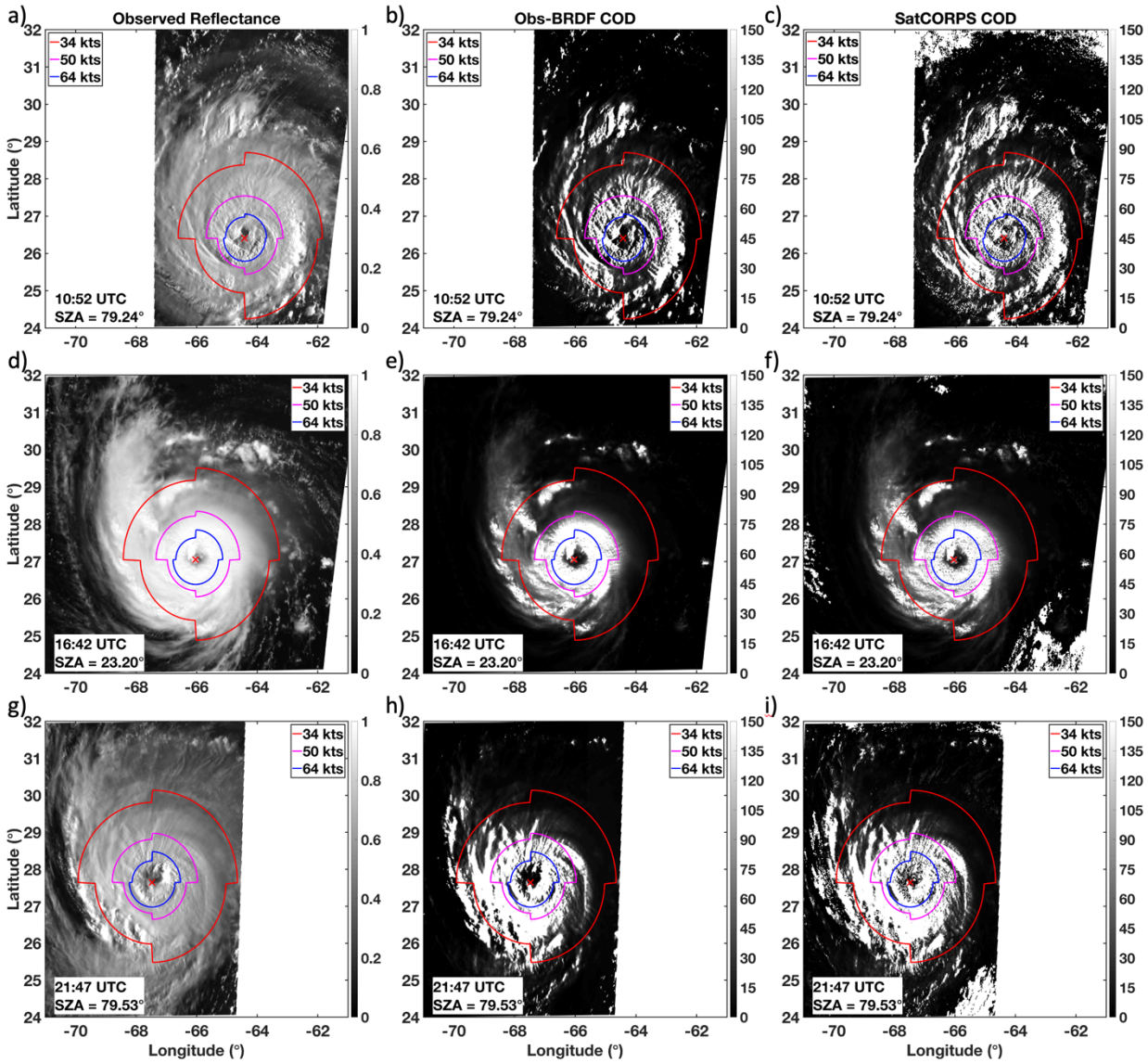


Figure 13: Observed VIS reflectance (left column), Obs−BRDF COD (center column), and SatCORPS COD (right column) imagery of Hurricane Florence from GOES-16 on 11 September 2018 at 10:52 (top row), 16:42 (center row), and 21:47 UTC (bottom row). Wind radii contours provided by the NOAA National Hurricane Center are indicated at 34, 50, and 64 kts, and the storm center is marked with a red 'X.' The average SZA within the 34-kts radii is displayed at the bottom of each panel. White areas beyond the edge of the image are either unprocessed parts of the domain or are regions with SZA greater than 82º.

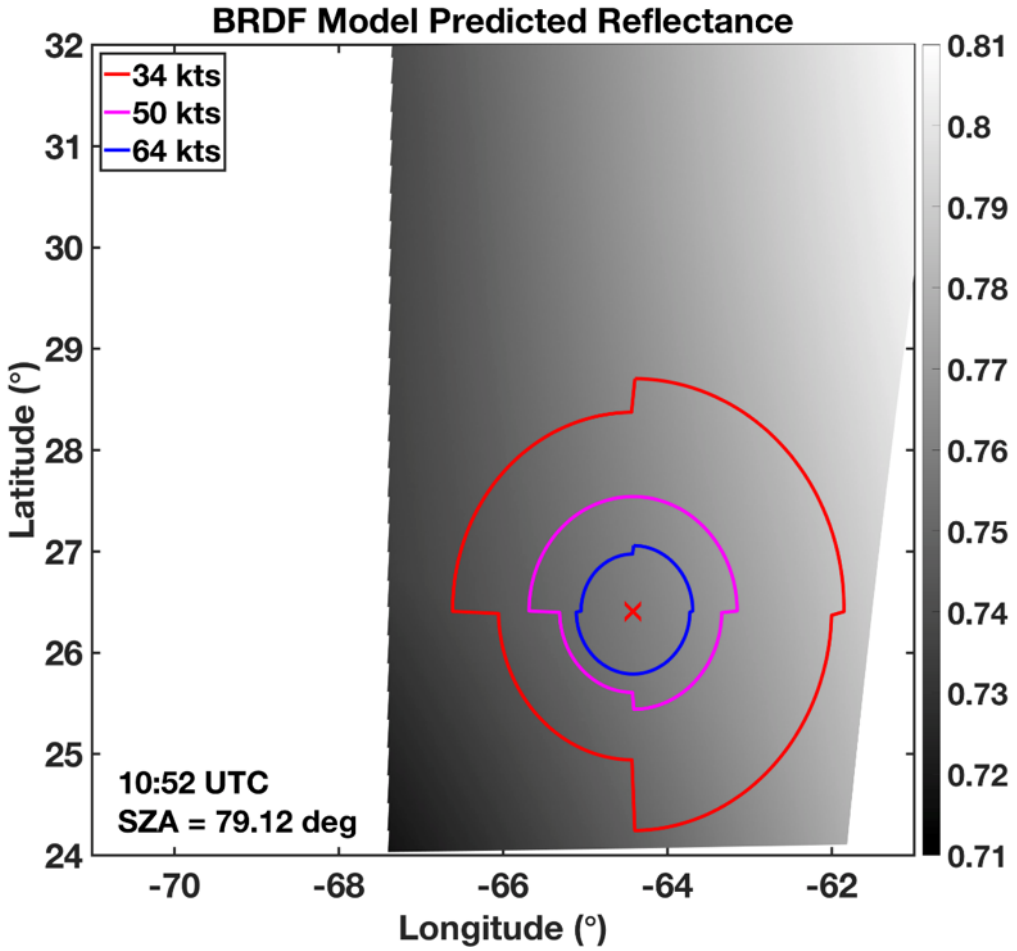


Figure 14: Predicted anvil reflectance based on the kernel-driven BRDF and GOES-16 viewing/illumination geometry at 10:52 UTC on 11 September 2018 over Hurricane Florence. White areas beyond on the edge of the image are either unprocessed parts of the domain or are regions with SZA greater than 82º.

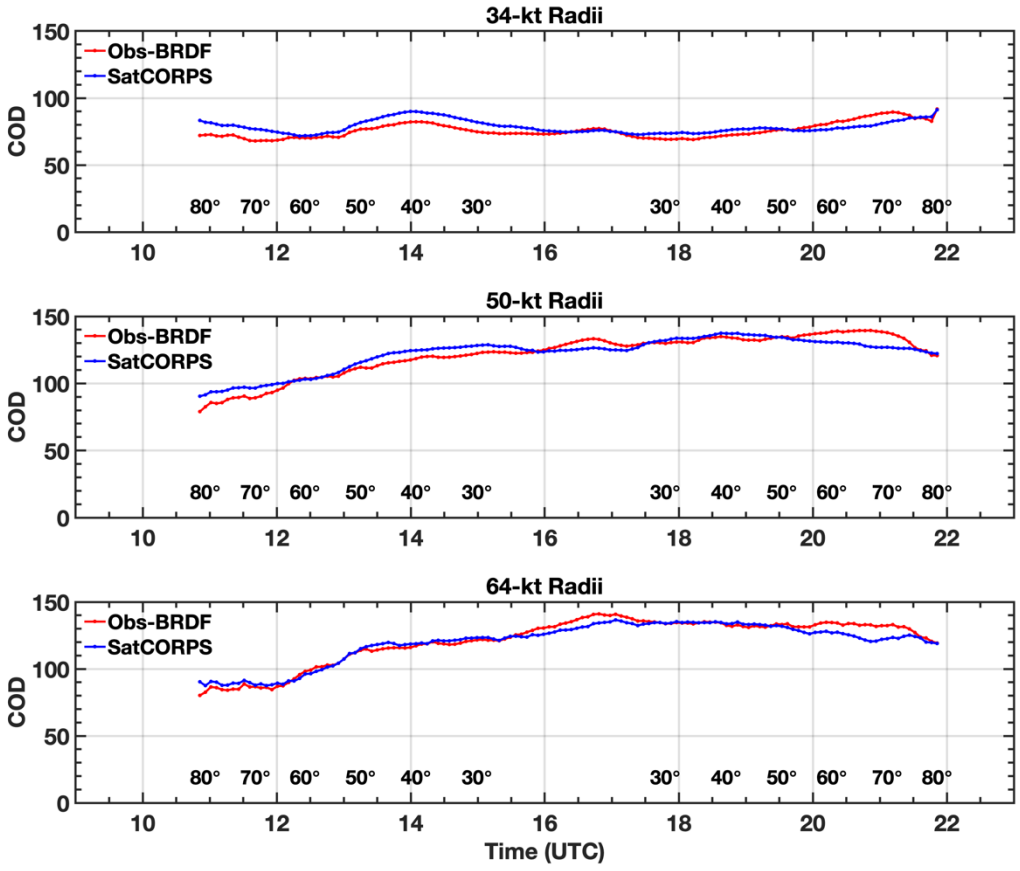


Figure 15: Cloud optical depth as a function of time (UTC hour) and SZA (listed in degrees) from SatCORPS (blue) and as determined from the SZA-dependent Obs−BRDF function (red) as shown in Fig. 9, derived over Hurricane Florence on 11 September 2018 within the 34-, 50-, and 64-kt radii as shown in Figs. 13 and 14.

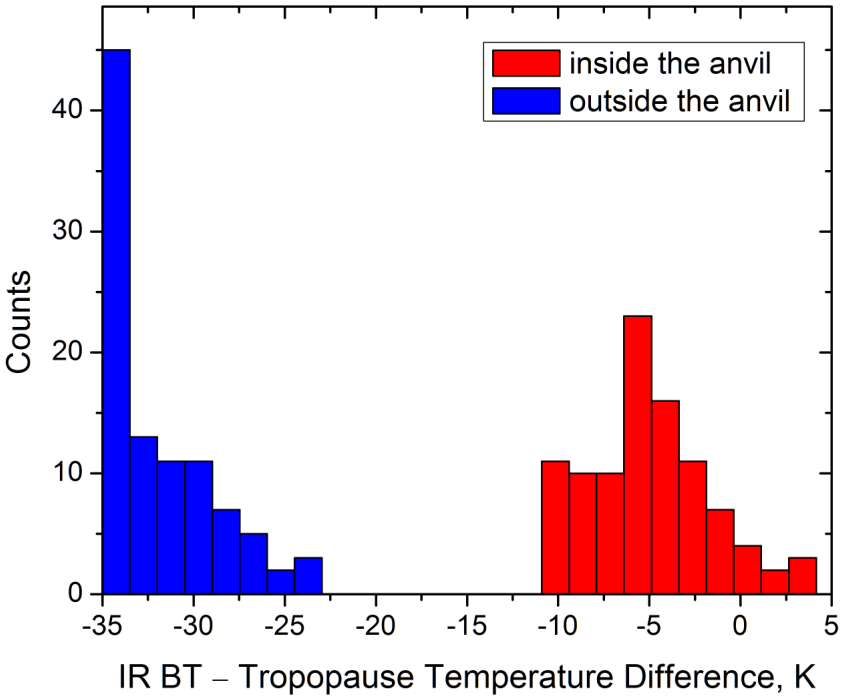


Figure A1: BTD histograms calculated over two regions observed by GOES-16 on 5 May 2019 at 23:00 UTC: inside a typical anvil cloud (red) and outside the anvil cloud (blue). On the red histogram, the peak is identified at bin 20 equaling 23 counts.