ASSESSING GOES-DERIVED SATELLITE PRODUCTS UTILITY IN AUGMENTING NEXRAD RADAR COVERAGE DURING SNOW SQUALL EVENTS IN CENTRAL PENNSYLVANIA

by

D’andre Jarobi Tillman

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Science
Atmospheric and Oceanic Sciences

at the
UNIVERSITY OF WISCONSIN-MADISON

2026

Abstract

Overland convective snow squalls are shallow (~2-3 km) and spatially discrete bursts of snow which most frequently happen downwind of the Great Lakes—sometimes in tandem with lake-effect snow. Snow squalls result from cold-season static instability within the planetary boundary layer which, in Pennsylvania, most frequently occurs from November to April. Snow squalls pose significant and sometimes deadly hazards to unsuspecting motorists. To compound forecaster difficulty, “nowcasting” (~0-3 hours of lead time) using ground-based radar is hindered by: a) beam overshooting from squalls being far away from the nearest radar, and b) spatial deficiencies in national radar coverage. This study proposes the use of satellite data to ameliorate these radar deficiencies and improve nowcasting capabilities. Data is ingested and colocated from both the State College radar (within 125 km) and the Geostationary Operational Environmental Satellite (GOES-16), the latter being used to generate quantitative cloud products using the Extended Clouds from the Advanced Very High Resolution Radiometer Organizations algorithm (CLAVR-x). A radar-based snow squall climatology with 51 unique snow squall cases from 2018-2022 in Central Pennsylvania were used as the basis for this study (Schneider et al., 2024), with 4 distinct snow squall types identified. CLAVR-x variables were found to be adept in identifying key characteristics of all squall types except parallel, “streamer” band types. Discrete, surface-based squalls had the highest (3-4 km) cloud top heights and greatest quantity of ice-phase cloud pixels, indicating skill in the identification of cold convective cloud processes like glaciation. Quasi-linear, synoptically forced squalls had the greatest skill in empirical fits with cloud water path and radar reflectivity, diminished skill for discrete, surface-based squalls, and no skill for streamer-band squalls. This study lays the groundwork for more sophisticated statistical analysis and nowcasting applications with machine learning.

Acknowledgements

It may sound cliche, but I could truly write a whole chapter on all the people I'd like to thank in my life. If I could be indulged just for a moment, there are some very special people whose help deserves to be verbalized.

I'll start with my family. My stepmom, dad, my mom, and my grandparents for all the long hours that they've toiled and labored through to get me through school. Thinking back on everything, it seems almost statistically impossible that someone like me could get to a place like this, and yet it is because of them that I was ever granted this opportunity in the first place. No acknowledgement could ever reveal the depth of my gratitude! I'd also like to thank my brother, for always bringing a good laugh when it was needed most, and standing by me through everything. Of course, I must acknowledge my beautiful dog Candy. I love you all so much!

My research advisors, Mark Kulie and Mike Foster. Again, no blots of black ink on a page can represent just how much you guys have meant to me, how invaluable your thoughts and inputs and ideas have been on this work week after week. Every weekly meeting has always felt so invigorating and motivating and I can't thank you guys enough. I have to give a special shoutout to Mark, who gave me a chance three years ago, and who helped me through a rocky transition to get here. I truly am indebted to you all. I'd also like to thank Karl Schneider and everyone at Penn State who have helped me put this work together, and who contributed their time and data.

In AOS I have to first thank my academic advisor and GAC committee member Tristan L'Ecuyer. Your insight and work has been so helpful in putting this thing together and I'm grateful for your time. My second GAC member Angela Rowe, who I've been honored to help TA and work with. When I think of scientists who push me to be better like no one else, you're on the top of my list, and I can’t thank you enough for devoting the time you do to helping me out and giving me opportunities. Of course thanks to the rest of the AOS faculty as well! Truly an absolutely standout department in our field. In these times where it feels harder and harder to do science everyday, it’s more important than ever to have a group of passionate, hardworking, and dedicated people like the ones in AOS, and I’m so grateful to be a part of it and to be given an opportunity to make my tiny mark in our field.

To my friends, both in AOS and out of it, thank you. There were times when I doubted myself, when I felt like I didn’t belong, like I took a wrong turn and by dumb luck ended up in my office. Every single one of those times you were all there to reassure me, to help me, to let me take a load off and recharge. Moving to a new state and new town halfway across the country was scary, but you all made the gamble worth it. Thank you so, so much.

Table of Contents

Abstract......i
Acknowledgements......ii
Table of Contents......iv
List of Figures......v
List of Tables......vii
Chapter 1: Background......1
1.1. Introduction......1
1.2. Background......2
1.3. Observing Strategies......8
1.4. Motivation......11
Chapter 2: Methodology......13
2.1. Snow Squall Climatology......13
2.2. KCCX Radar Data......19
2.3. CLAVR-x......22
2.4. Colocation......27
2.5. Analytical Methods......30
Chapter 3: Results......33
3.1. RQ1: Spatial Match-up Results......33
3.2. RQ2: Empirical Relationship......40
3.3. Data Filtering......45
Chapter 4: Application: GREMLIN as an Example......79
Chapter 5: Discussion and Conclusion......84
5.1. Key Takeaways......85
5.2. Future Work......86
References......88

List of Figures

Figure 1. Heatmap showing the number and location of snow squall warnings from January 2018 to January 2026. Source: Iowa Environmental Mesonet....4

Figure 2. Map of NEXRAD radar coverage above 4 kt (1.2 km), 6 kt (1.8 km), and 10 kft (approx. 3 km) across the contiguous United States....9

Figure 4. Stacked histogram showing the number of timestamps within each snow squall type. Orange colors indicate the number of timestamps which overlap with timestamps associated with NWS snow squall warnings within the NWS State College warning area.... 18

Figure 5. Histogram showing the frequency of different VCPs within the dataset.... 22

Figure 6. Histogram showing the CLAVR-x - radar colocated timestamp difference, in minutes.... 28

Figure 7. Scatterplot showing the total pixel occurrence count (y-axis) vs. distance from the KCCX radar (x-axis). Colored in red is this pixel occurrence for radar pixels, and in blue for CWP pixels.... 34

Figure 8. Plot centered on Milwaukee, Wisconsin demonstrating the "spatial overlap" index. Yellow pixels indicate radar and satellite have identified a pixel, green is satellite only, and purple is radar only. 36

Figure 9. Box and whisker plot showing the success ratio distribution as a function of snow squall type within 0-125 km from the radar.... 38

Figure 10. Violin plot showing the distribution of success ratio values based on the VCP used.... 39

Figure 11. Box and whisker plot showing the success ratio distribution as a function of snow squall type within 0-75 km from the radar.... 40

Figure 12. 2D Histograms of logged CWP (y-axis) vs. reflectivity (x-axis), broken down by snow squall type. Warmer colors indicate a higher density of points.... 42

Figure 13. As in Figure 12, but instead broken down by each unique snow squall case....43

Figure 14. As in Figure 11, but now showing R2 scores for each snow squall type.... 44

Figure 15. Multi-panel plot showing CAPPI reflectivity for various VCPs. The R2 value shown is with respect to CWP (no gaussian filtering).... 47

Figure 16. Boxplot showing R2 score vs. VCP at three distance thresholds....48

Figure 17. Histogram of VCP, broken down and colored by snow squall type.... 49

Figure 18. Four panel plot showing CWP (logged) with varying levels of Gaussian filtering applied....50

Figure 19. Histograms of CWP across the entire dataset, with progressively higher sigma Gaussian filters applied....51

Figure 20. As in Figure 12, but with the Gaussian filtered CWP....52

Figure 21. Box-and-whisker plots showing R2 score variation based on snow squall type for all points a) 0-75 km and b) 0-125 km away from the radar....53

Figure 22. Histogram showing success ratio change for all timestamps.... 54

Figure 23. Plot showing how R2 scores vary with increasing Gaussian filtering threshold values. The solid lines are the median for each snow squall type, and the dotted line is the maximum value....55

Figure 24. Histogram showing the number of features associated with the most common phase type within the feature....57

Figure 25. Cloud phase plots from two different times on the same day. (a) is at 19:10 UTC, and (b) is at 21:20 UTC....59

Figure 26. Box-and-whisker plots showing cloud phase proportions per timestamp by snow squall type. 60

Figure 27. Box and whisker plot showing the R2 values as a function of scene cloud phase proportion... 61

Figure 28. Box-and-whisker plot showing cloud top height per pixel based on snow squall type....62

Figure 29. Plot showing a) radar reflectivity PPI at the 0.5 degree elevation angle, b) CLAVR-x cloud top height, and c) a psuedo-range height indicator plot from several PPIs, with an estimated cloud top height drawn with a dotted line. All images are valid at 17:11 UTC on March 28th, 2022.................................. 64

Figure 30. Histogram showing the distribution of cloud top heights..65

Figure 31. Box-and-whiskers plot showing cloud top height as a function of snow squall type, filtered by ice content.. 66

Figure 32. 2D Histogram showing the distribution of points based on mean tobac feature height and phase mode..67

Figure 33. CWP versus Cloud Top Height by snow squall type. Black contours show all pixels which were identified as SLW..68

Figure 34. As in Figure 33, but contoured everywhere that ice clouds are identified................................ 69

Figure 35. Boxplot of effective radius versus snow squall type for each DCOMP mode. Outliers past the 5th and 95th percentile whiskers are omitted...71

Figure 36. Cloud Effective Radius plots for each DCOMP mode centered over central Pennsylvania...... 71

Figure 37. As in Figure 36, but for a different time..72

Figure 38. 2D histograms as in figure blue, but showing CWP against cloud effective radius for each DCOMP mode...72

Figure 39. As in Figure 38, but broken down by snow squall type and contoured at points which reflectivity > 25 dBZ...73

Figure 40. As in Figure 39, but for DCOMP mode 2.. 74

Figure 41. 4-panel plot as in figure blue, except points which had an effective radius DCOMP1 > 50 μm is contoured..75

Figure 42. Box-and-whiskers plot showing R2 values for each snow squall type, broken down by DCOMP1 effective radius filtering thresholds...76

Figure 43. As in Figure 42, but with a gaussian filter applied to the underlying CWP data....................... 77

Figure 44. 2D histograms of CWP vs. CAPPI reflectivity, with a >50 μm effective radius DCOMP1 filter applied...78

Figure 45. GREMLIN synthetic reflectivity centered over Central Pennsylvania..................................... 80

Figure 46. Radar reflectivity from GREMLIN (left) and KCCX (right) centred over Central Pennsylvania. 81

Figure 47. Density plot comparing CAPPI reflectivity to GREMLIN synthetic reflectivity.......................82

Figure 48. Radar reflectivity centered on Central Pennsylvania...83

Figure 49. As in Figure 47, but with CWP-derived synthetic reflectivity.. 84

List of Tables

Table 1. Characteristics of the CAPPI grid created from the radar data..21
Table 2. List of CLAVR-x variables used, a brief description for each, and the subsequent units.... 27
Table 3. The number of colocated timestamps and snow squall events attributable to each snow squall type..29
Table 4. Specifications of the TOBAC feature identification and tracking algorithm................. 31
Table 5. Table showing the breakdown of SLW vs. ice pixels... 58

Chapter 1: Background

1.1. Introduction

Snow squalls are shallow, discrete, and intense bursts of snowfall which are convective in nature, not unlike the typical diurnally-driven summer time thunderstorm. Often concurrent but not predicated by lake-effect snow (LES), these intense bursts of precipitation rapidly reduce visibility, and can produce significant snowfall accumulations over relatively short durations (Pettegrew et al. 2009; Milrad et al. 2011; Milrad et al. 2014) . Snow squalls are strongly modulated by diurnal processes, with the majority occurring during the day as a result of boundary layer destabilization from solar heating (Banacos et al., 2014). The daytime tendency of these phenomena is pernicious for motorists, as the primary danger of snow squalls is often not the amount of snow, but rather the thin layer of road ice formed via flash freeze that significantly reduces friction on roadways. It is in this manner that numerous traffic accidents can be attributed to snow squalls (Devoir 2004). These motor vehicle accidents occur with such frequency and lethality that recent research has focused on social science and outreach to better warn and inform the public, including the creation of a virtual reality simulation (Martin et al., 2026). Snow squalls are shallow—on the order of 2-3 km (Brown et al., 2007)—which complicates real-time surveillance of these systems by National Weather Service (NWS) radars since Weather Surveillance Radars (WSR-88Ds) will fail to adequately sample any squall greater than 75-100+ km in range from the radar site. For regions like Southeastern Pennsylvania, snow squalls undetected by radar (and therefore unwarned) led to the death of six and causing injuries to over two dozen on March 28th, 2022 (Medina 2022). There are also equity issues; those already vulnerable in poor, rural, and “radar-dearth” locations are the most at risk for experiencing unwarnable snow squalls. With new radar sites being expensive to build and maintain, another option for identifying and quantifying the intensity of these snow squalls in conjunction with radar is imperative. This study therefore utilizes satellite-based measurements of cloud properties to augment the spatially limited radar network—an area in which shallow convective snow like snow squalls have yet to be meaningfully investigated—and assesses their utility in describing the distinct macrophysical characteristics across different snow squall morphologies. The desired outcome of this analysis would be increased “nowcasting” capabilities (forecasting at very small lead times, typically within a 0-3 hours) independent of radar.

1.2. Background

1.2.1. Lake-Effect and Snow Squalls: Shared History and Terminology

Although LES and snow squalls are contemporarily treated as separate phenomena (and are thus warned separately by the NWS), they were often not separated in the past. It was common for lake-effect to be referred to as snow squalls, flurries, or bursts (Wiggins 1950), and some of the earliest literature on mesoscale phenomena did classify brief, intense bursts of snow associated with pressure rises and gusty winds as “squalls” of the snow variety (Loisel 1909), perhaps leading to the adoption of the term by the Weather Bureau (i.e., precursor to the NWS) for LES. Both LES and snow squalls can be considered as subsets of convective snow, and it would be appropriate to refer to either as such (with the distinction being one originating from lake-air interaction and the other land-air). However, another challenge posed is that most convective snow originating over land does not meet official NWS snow squall warning criteria, so referring to all such activities as “snow squalls” may not necessarily be appropriate either. Schneider et al. (2024) suggests referring to all snow squalls (including those which have intensities less than the NWS warning thresholds) originating in the boundary layer over land as “overland convective snow”, or simply “convective snow”. For the purposes of this study, “overland convective snow” and “snow squalls” will be used interchangeably, with the intention that they refer specifically to any convective activity originating in the boundary layer with snow/graupel as the primary precipitation type, often accompanied by gusty winds, blowing snow, and pressure rises. Seldom is lightning observed; Pettegrew et al. (2009) documents such an event, though Milrad et al. (2011) notes in their study of a snow squall in Southern Canada that research into thundersnow events outside of mid-latitude winter storms is limited.

Although lake-effect and snow squalls often occur together downwind of the Great Lakes in the northeast United States, this is far from the only condition in which they occur. Despite being biased by NWS-warning metrics (including non-meteorological socioeconomic factors like population and interstates) Figure 1 shows that snow squalls occur across the country, far from the Great Lakes. Focus for background literature will first be placed on studies in the northeast (where shallow convective snow has been studied most heavily), then broadened to research elsewhere, with the caveat that ultimately this study will focus on snow squalls occurring in Central Pennsylvania—primarily because the climatology of snow squalls used focuses on Central Pennsylvania (Schneider et al. 2024). Future research would be well equipped to extend this work to other regions in the northeast United States and across the country.

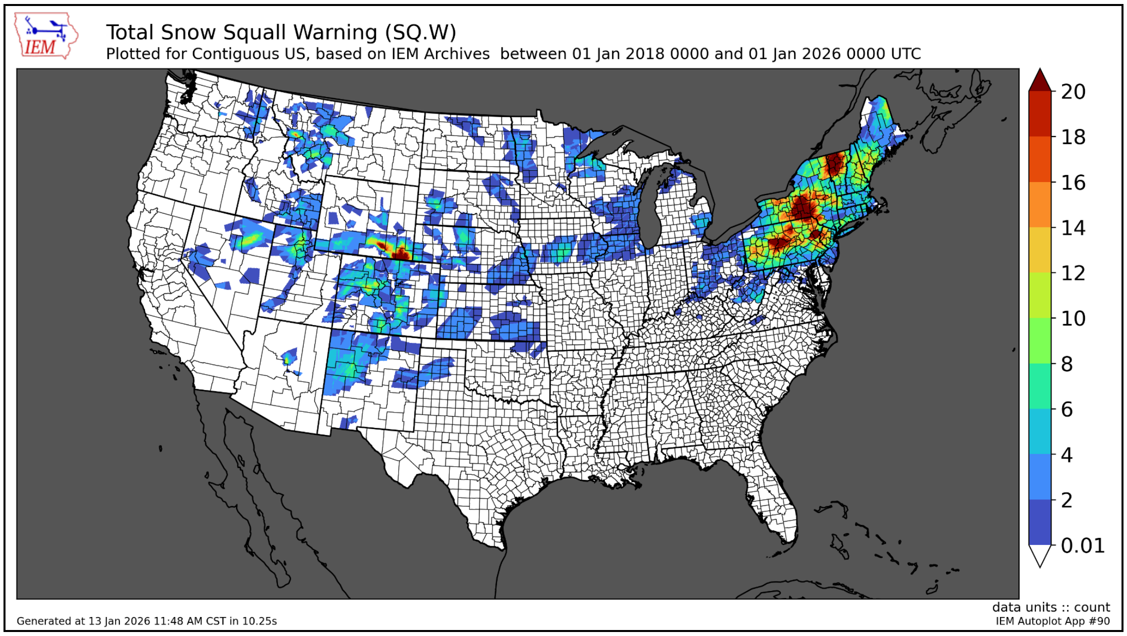

Figure 1. Heatmap showing the number and location of snow squall warnings from January 2018 to January 2026. Source: Iowa Environmental Mesonet.

1.2.2. Shallow Convective Snow Over Lake/Ocean

Lake-effect clouds are formed much like snow squalls or overland convective snow clouds are; by relatively cool air aloft flowing over warmer near-ocean surface air, creating instability and initiating convection, often as a result of a cold-frontal passage as a part of a mid-latitude cyclone. They are characterized by upper-level subsidence (manifested as a temperature inversion and relatively higher dewpoint depression) and well-mixed planetary boundary layer (PBL) within the depth of the cloud layer and below the temperature inversion. There have been several field campaigns attempting to quantify the variability of these clouds with in-situ observation.

Over the Laurentian Great Lakes, studies in the 1980s and 1990s like Lake Ontario Winter Storms project (LOWS; Reinking et al., 1993), and project Lakesnow (Braham 1990) sought to quantify, with higher spatiotemporal resolution than previously afforded, variability in convective snow at macrophysical and microphysical scales. Another field campaign during this time, the Lake-Induced Convection Experiment or "Lake-ICE" (Kristovich et al. 2000), lead to Barthold and Kristovich (2011) exploring how the size and distribution of ice particles within lake-effect clouds varied from their earliest stages upwind to maturity downwind (relative to the west-east coasts of Lake Michigan). They found that the size of cloud particles increased downwind as expected, but that the location of the largest cloud particles vertically was below the boundary layer top and above cloud base. The authors also hypothesized the potential source of cloud and ice condensation nuclei for the lake clouds as blowing snow upwind, but highlighted uncertainties in this estimate, demonstrating the lingering microphysical ambiguity in cold, shallow convective snow with implications for precipitation processes.

These field campaigns laid the foundation for more sophisticated study: the Ontario Winter Lake-Effect Systems Field Campaign (OWLeS, Kristovich et al., 2017). Several studies followed from OWLeS, of which a few detailed how microphysical processes within lake-effect can vary depending on mesoscale fronts and topography (Campbell and Steenburgh 2017). From the breadth of these works, a microphysical baseline for over-lake CS (albeit within specific geographic regions over limited observation periods) is obtained, laying the investigative groundwork for other types of CS.

1.2.3. Snow Squall Case Studies and Composites

Although more limited than studies on lake and ocean-effect, there are numerous studies which have analyzed snow squalls (overland CS) specifically. One of the first studies to make such a distinction was Lundstedt (1993). Snow squall events were determined using surface station data in Concord, New Hampshire. The study used a composite of snow squall events from 1989 to 1992 (about 30 events) to form a forecasting parameter called WINDEX (winter instability index) which forecasters could use to analyze whether the large-scale synoptic environment was favorable for snow squall development. Lundstedt found that, similar to lake effect, the passage of a cold front and subsequent 500 mb trough passage were critical. However, the instability generated over land was contingent on moisture being advected by an antecedent lake (as opposed to moisture flux from the surface directly in lake-effect), or from remnant moisture (i.e., a low pressure system associated with the trough passage). This study laid the foundation for snow squalls to be studied separate from lake-effect, and was the first to show the potential utility of a forecasting parameter for snow squalls.

After Lundstedt's initial study, interest in studying snow squalls grew into the 21st century. DeVoir (2004) vocalized the need for the NWS to develop a new way to warn of snow squalls, which at the time they classified as “high impact, sub-advisory” events, meaning that while they were impactful on both life and property, the NWS had no formal advisories and/or warnings they could issue. Pettegrew et al. (2009) studied a particularly severe snow squall which moved through southeast Iowa and central Illinois in February 2003. They note that previous studies on convective snowfall was focused on that which occurred in large-scale synoptic storms, opting to focus on mesoscale snow banding resulting from frontogenesis and/or conditional slantwise instability as opposed to processes linked to the boundary layer convection (Kocin et al., 1985; Schneider 1990). They observe that the meteorological characteristics of the snow squall were not dissimilar from warm-season convection, with lightning even being observed. Importantly, they note that despite the intense impacts observed at the surface (~50 kt winds, downed trees and utility poles), storm cloud tops “never exceeded 3.7 km”, with maximum reflectivity values averaging out between 40-45 dBz.

Snow squall events have been documented in many geographic regions around the United States and Canada (e.g., the front range of the Rocky Mountains in Canada and the United States, western Quebec/eastern Ontario, the United States Great Plains region). All events exhibited the sudden, intense onset of snow, with lower quantitative precipitation (QPF) values than in synoptic-scale snows or longer-duration LES, but accompanied by intense winds, blowing snow, and very rarely lightning (Milrad et al. 2011, 2014; Schumacher et al. 2010; Rosenow 2018; Bender and French 2020). These studies collectively highlighted the relative dearth of resources available for meteorological agencies to warn of snow squalls and radar coverage deficiencies.

The environmental conditions associated with snow squalls which ultimately control their morphological characteristics have also been documented. For instance, Rosenow et al. (2018) studied a snow squall which occurred in the San Luis Valley in Colorado, an event that was almost completely undetected by the operational radar network (2018). One of the more notable aspects of this squall was the relatively high (~600 J/kg) amount of surface–based convective available potential energy (CAPE) present; a value much higher than seen in other literature for similar snow squalls. Additionally, Bender and French (2020) studied the environmental constraints which control the morphology of snow squalls in western South Dakota. They found distinct differences in both wind shear and CAPE between so-called “banded” and “cellular” snow squalls, with the former characterized by a moderately sheared profile (20-40 m/s) and low (approx. 0-20 J/kg) CAPE. Although these studies occurred far from the great lakes and were largely absent of 500 mb troughing, low-level instability was key in their formation. Still, the parameter spaces in which these squalls occupied deviated from prior studies in the northeast, especially in their shear profiles, especially in Bender and French (2020). Despite their differences, these studies collectively support the need to develop better operational forecasting/nowcasting tools for snow squalls.

The first operationally-relevant snow squall tool was developed by Banacos et al. (2014), whereby a “snow squall parameter” was defined in likeness to the WINDEX parameter of Lundstedt 1993. This parameter was created by using a climatology of snow squalls created from surface station data in the northeast United States (Burlington and Montpelier, Vermont, and Massena New York) from 10 winter seasons from 2001-02 through 2010-11. They retread the work of Lundstedt (1993), with updated data and modeling capabilities, to establish their so-called snow squall parameter, of which they demonstrate its effectiveness in identifying favorable environments for snow squall formation, though it did not account for different snow squall types as Bender and French (2020) would do. His work would, in part, motivate the NWS to formalize the “snow squall warning” as a forecaster warning tool analogous to the severe thunderstorm warning (Hawkins 2018).

1.3. Observing Strategies

1.3.1. Radar-Based

The NWS operates and maintains a network of WSR-88D radars nationwide. Figure 2 shows the extent of this network and the effective range of each radar. Readily apparent is the limited spatial coverage broadly across the United States, especially in rural locations and smaller cities. While radar has long been used to monitor precipitation at a high spatiotemporal cadence, including convective snow, its range limitations have long been known (Peace and Sykes 1966), with some trying to account for this issue by introducing certain volume coverage patterns (VCPs) which incorporate very low, or even negative, elevation angles (Brown 2007). Such elevation angles, while proven useful in some circumstances, are not used operationally and so surveillance of shallow convective snow from WSR-88D radars in the United States remains limited. At MQT, the NEXRAD is typically kept in clear air mode for LES events to allow for high radar sensitivity—see Pettersen et al. (2020). However, that scanning strategy also limits temporal coverage. Kulie et al. (2021) also highlights NEXRAD coverage issues for LES events at KMQT (Marquette).

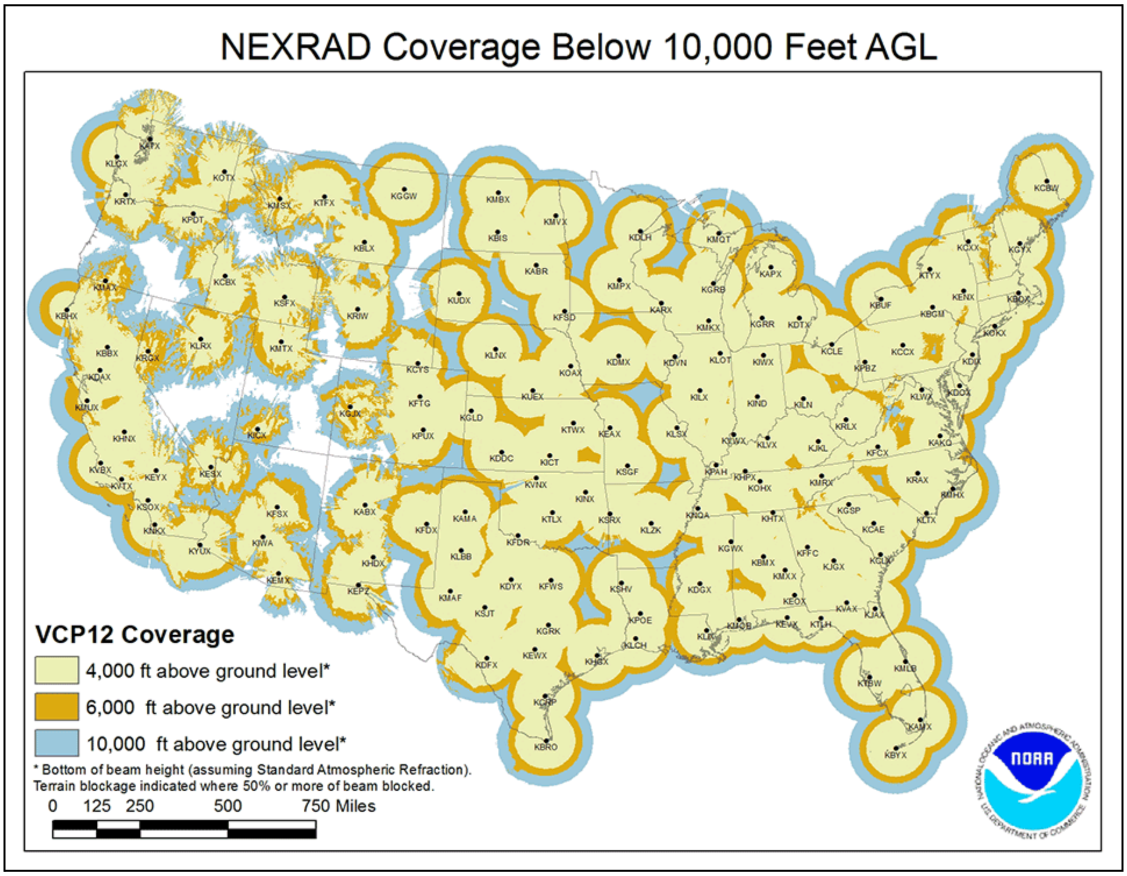

Figure 2. Map of NEXRAD radar coverage above 4 kt (1.2 km), 6 kt (1.8 km), and 10 kft (approx. 3 km) across the contiguous United States

1.3.2. Satellite-Based

The remote sensing of precipitation from space has a long and storied history, especially when considering snowfall estimation. Both passive and active sensing techniques have been employed towards this goal. For example, Kulie et al. (2016) determined, from the cloud profiling radar on the *CloudSat* (Stephens et al. 2002) satellite, that shallow snowfall constitutes a sizable portion (~36%) of snowfall mode across the globe.

There has also been an extensive history of colocating ground-based radar with satellite observations. Of particular relevance to this study is the Geostationary Operational Environmental Satellite (GOES) Radar Estimation via Machine Learning to Inform NWP (GREMLIN) project, which as the name suggests, utilizes the GOES products to create a "synthetic reflectivity" product (Hilburn et al., 2021). It does so by training a machine learning model off the following GOES Advanced Baseline Imager (ABI) channels:

1. Channel 7 (3.9-mm, shortwave infrared window),
2. Channel 9 [6.9-mm, midlevel water vapor (442 hPa)], and
3. Channel 13 (10.3-mm, clean longwave infrared window).

These channels, along with the Multi-Radar Multi-Sensor (MRMS) product, are ingested into the ML model, with the end goal of predicting reflectivity based on GOES ABI output alone. An important observation is the use of only IR channels, which was done for "maximum portability" between new generation and legacy observations, and unified day/night observations. GREMLIN does not ingest any of the visible channels (in particular, the highest resolution 0.5 km resolution 0.47 μm "red" channel). Hilburn et al (2021) say that, while visible and near-IR channels would provide additional information content, the use of such channels was outside the scope of their study.

GREMLIN was made specifically for deep convection, for use during the warm season (boreal summer); therefore, its skill decreases significantly during boreal winter (Lee and Hilburn 2024). By extension, its skill should also be greatly decreased for snow squalls, though this has not been directly studied. A comprehensive comparison between the results of this study and GREMLIN is outside the scope of this paper, but a case study comparing an individual snow squall event can be found in Chapter 4. In addition, while contiguous day/night coverage is advantageous, its use of only-IR channels comes at the cost of resolution (2 km, then downscaled to 3 km for NWP data assimilation purposes). Despite these limitations, initial testbed activities with warm season convection found GREMLIN aided NWS forecasters when radar coverage was poor (Lyza et al, 2026), demonstrating the utility of a satellite-based product augmenting radar coverage, and by extension, motivates the creation of an analogous product for snow squalls.

1.4. Motivation

Although previous studies have examined snow squalls both from the perspective of case studies and climatology through the use of surface station data and radar (among other in-situ, active, and passive observations), none have quantitatively used satellite products to examine their macrophysical cloud characteristics, and by extension none have done this across multiple snow squall events in different thermodynamic environments. The closest product to such an analysis is GREMLIN, and while it was found to be useful for warm season convection, its coarse spatial resolution and diminished skill in boreal winter make it impractical for use with snow squalls (Chapter 4 quantitatively shows this discrepancy). Given the limited utility of radar to detect snow squalls past 75 km, and the threat they pose to both life and property, a high-resolution satellite-based method to 1) detect where a snow squall is, and 2) provide a probabilistic assessment of its intensity without radar would prove invaluable.

To achieve such a goal, there must first be a thorough investigation into how exactly these satellite-derived cloud products relate to the radar data, and whether such a comparison even holds significance in the first place. Therefore, we ask the following two research questions:

1. How well do satellite-derived cloud products match up spatially with radar products during snow squall events?
2. Is there a relationship between snow squall cloud properties observed from satellite-derived cloud products and radar reflectivity?

The first question will primarily concern itself with matching up, both in space and time, the cloud features observed from satellite products with cloud features identified in radar. This task is critical to establish initially, as it is important to determine, both in space and time, how CLAVR-x colocates with radar pixels (which is presumably indicative of precipitation). Once the rate and distribution of these collocations is elucidated, then the second research question is asked: can a meaningful empirical relationship be formed from the macrophysical cloud properties observed from satellite-derived cloud products, and those same properties observed from radar? If such a relationship proves untenable, then why?

A sufficiently comprehensive answer to both of these questions will result in: 1) an improved understanding of snow squall evolution in Central Pennsylvania, 2) methods for probabilistic estimation of snow squall location and intensity from satellite only, 3) a large and comprehensive radar-satellite colocated snow squall dataset, which will serve as critical ground-level work for 4) a refined, comprehensive training dataset for future machine learning techniques from which outcome 2 can be accomplished.

Chapter 2: Methodology

2.1. Snow Squall Climatology

2.1.1. Overview

In order to adequately study how satellite-derived cloud products can augment radar during snow squall events, there must first exist a large enough database of snow squall occurrences to analyze. Though many have tried to make composites and/or climatologies of snow squalls (Lundstedt 1993; Banacos et al. 2014; Colby et al. 2022), these studies utilized reanalysis or surface station data as primary factors in subsetting snow squall dates from non-snow squall dates. Although those two datasets are (and will continue to be) critical in the identification of snow squalls, an issue arises; if only using surface station data or reanalysis, one runs the risk of only capturing environments and/or events characteristic of only the most intense snow squalls, or only characteristic of squalls which impact the surface station(s) in question. While this still holds research value, one would be missing the large breadth of convective snow which falls outside those filter criteria. Schneider et al (2024) give the example of a quasi-linear, synoptically forced snow squalls versus discrete, cellular squalls. The former is much more likely to be identified by surface station data due to its (relative to cellular snow squalls) larger coverage and intensity, while the latter is much less likely due to being much smaller and less intense. Additionally, aside from Colby et al. (2022), most studies which made snow squall climatologies are over a decade old, and would require reproducing their methodology to create an updated dataset for the last five or ten years.

To create an updated climatology, and to ameliorate concerns of bias toward more intense, synoptically forced squalls, Schneider et al. (2024) created a radar-based snow squall climatology, using the KCCX radar site just outside of State College, PA, though they note their methodology can be modified and applied to any radar across the country. While the full methodology is described within their paper, the key features will be summarized here.

The snow squall climatology was created by analyzing KCCX data during the cool season months (November-April) during which (the authors note) snow squalls are observed in Central Pennsylvania. They use Rapid Refresh (RAP; Benjamin et al. 2016) model data to rule out any thermal vertical profiles unsupportive of snow. The only exception being if surface station data from the nearest Automated Surface Observing System (ASOS) site reported snow anytime within the previous hour timestamp being analyzed. Once a set of favorable snow squall vertical thermal profiles are obtained (at an hourly cadence according to the temporal frequency of RAP), radar data from KCCX closest to the hourly RAP time is retrieved. The radar data (taken from the lowest elevation scan, typically 0.5°) is then regridded from its native polarimetric form to a 1 km Cartesian grid within 100 km of KCCX. They employ an algorithm on the regridded radar field to check that:

1) The cells are discrete (at least 50 km^2 in size) and no stratiform precipitation is detected (< 67% of the domain has dBZ $\geq$ 5),
2) the discrete cells have a maximum reflectivity of at least 20 dBZ, and
3) there is a sufficiently large reflectivity gradient along the discrete cell edges ($\nabla Z \geq 5$ dBZ km^{-1}, confirming that they are indeed convective in nature).

Schneider et al. (2024) cite the 20 dBZ figure as coming from multiple studies of snow squalls who used radar analysis (Pettegrew et al. 2009; Rosenow et al. 2018) and experimentation. They also cite 20 dBZ as generally being at or just above the NWS snow squall warning criterion for visibility ($\leq$ 0.4 km), with 20 dBZ generally corresponding to visibilities of ~0.5-1.5 km.

From this process, a list of potential snow squall times is ascertained, which are then filtered by the vertical thermodynamic profile described earlier, and then filtered once again by human intervention. The final snow squall climatology is then produced, with an hourly temporal frequency.

2.1.2. Snow Squall Types

Given the ability to detect the presence of snow squalls from radar reflectivity, it follows that the algorithm can then ascertain some measure of the squall's morphological characteristics (cell length, aspect ratio, etc.). As a result, squalls were categorized and typed by these characteristics. These types were classified according to the 20 dBZ reflectivity contours as described in the previous section. The dominant category was chosen if multiple occurred simultaneously within a given hour. The snow squall typing scheme as described in Schneider et al. (2024) is as follows :

1. **S1**: Single, laterally extensive frontal band oriented approximately perpendicular to the mean wind.
2. **S2:** Single cells with small areal extent (closer to the 50 km^2 threshold).
3. **S3:** Multicells/mixed-type cells. Cells which cannot be strictly classified as S1 or S2. Further broken down into cells approximately perpendicular (S3-A) or parallel (S3-B) to the mean wind.
4. **S4:** Bands oriented approximately parallel to the mean wind, similar in appearance to lake-effect snow.

Where the "mean wind" is defined as the layer-averaged wind between the surface and cloud top, with the cloud top being the point at which the environmental potential temperature (θ_e) becomes larger than 2°C of the surface value. Each of these snow squall types were able to be linked back to the synoptic scale environmental characteristics, the details of which are described in Schneider et al (2024) and are summarized here. Vertical profiles of meteorologically relevant variables (temperature, dewpoint, wind speed, etc.) were taken and composited from the RAP model at the nearest grid point to State College. These profiles were then analyzed and compared against each other to determine to what extent (if any) they varied across snow squall type. They found three variables were most important (i.e., had the most statistically significant differences among the snow squall types): the 500-mb trough position with respect to State College, the amount of surface-based CAPE, and the depth of the cloud layer. These results are summarized in Figure 3. It is from this dataset that we derive the snow squall events and timestamps used in this study, the characteristics of which are discussed in the next section.

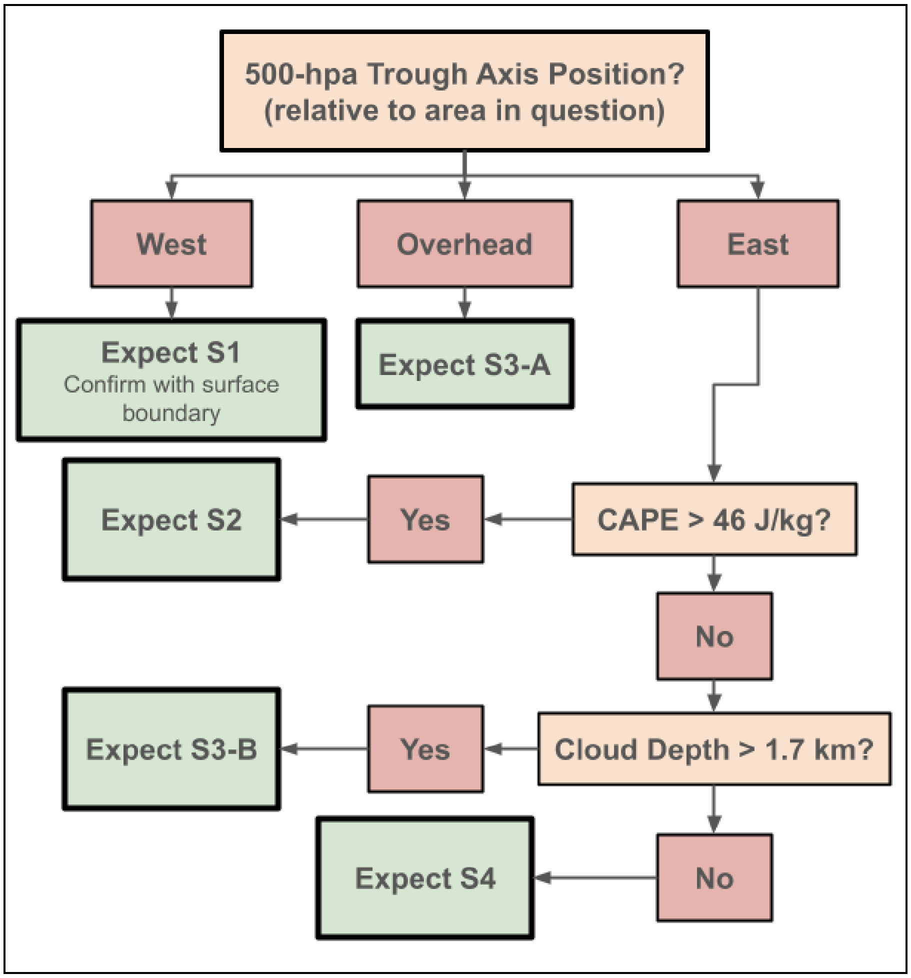

Figure 3. Flow chart adapted from Schneider et al (2024) showing which snow squall type to expect depending on the environmental conditions.

2.1.3. Overlap with NWS Snow Squall Warnings

Figure 4 shows the number of snow squall timestamps within the dataset, and the number of those timestamps which overlapped with a NWS snow squall warning. NWS snow squall warning data from 2017-2022 for KCCX was retrieved from an online archive maintained by the Iowa Environmental Mesonet (IEM). To determine whether a climatology timestamp "overlapped" with an NWS warning, each timestamp was converted such that its minutes and seconds information were omitted; in other words, the times were specific only to the hour. It was then calculated where the NWS warned-timestamps and snow squall climatology timestamps overlapped.

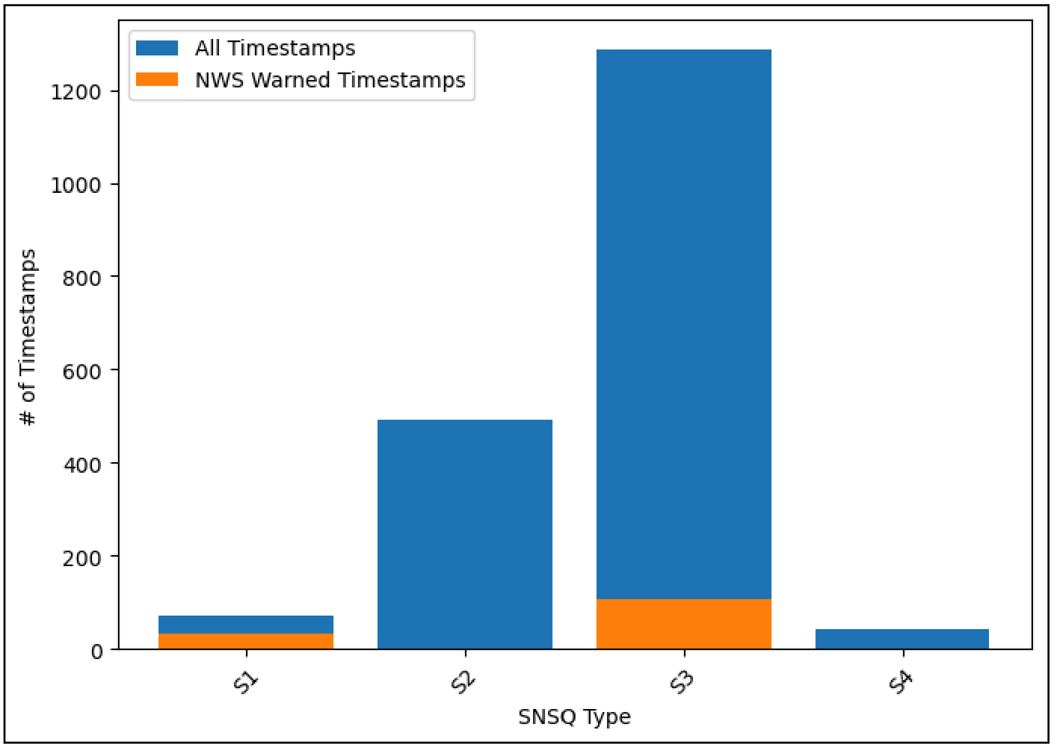

Figure 4. Stacked histogram showing the number of timestamps within each snow squall type. Orange colors indicate the number of timestamps which overlap with timestamps associated with NWS snow squall warnings within the NWS State College warning area.

This methodology produced only a 7.44% overlap rate between warned and non-warned squalls. That statistic can also be interpreted as only 7% of climatology timestamps falling within an hour in which a snow squall warning was issued. It is clear to see how even this somewhat small figure could be inflated. One timestamp will often hold several individual squalls, of which only a few may be warned (i.e., warning polygons). It is also possible that, despite the climatology and NWS warned timestamp taking place within the same hour, that they do not occur at the same time (e.g., the warning may be issued at 3:05 pm, but the timestamp being assessed is at 3:55 pm). These statistics were meant to give more general, broad information about how many timestamps contain warning-level squalls.

Perhaps a more telling statistic, only 8 of the 51 individual snow squall cases had a snow squall warning issued at all, a rate of 15.67%. It is difficult at this point to say whether these are being under or over warned since we don't truly know how many warning-level squalls happen within a year, but at the very least it insinuates that the dataset created is representative of both snow squall warning-level convective snow, and sub-warning level snow.

2.2. KCCX Radar Data

Radar reflectivity data was taken from the local WSR-88D radar KCCX. Radar reflectivity is defined by the sixth moment of the particle size distribution, given by the following expression:

$$Z = \int n(D) D^6 dD \tag{1}$$

where the radar reflectivity factor Z is equal to the integrated particle size distribution (PSD) (in terms of diameter instead of radius here). Radar reflectivity is measured in dBZ, expressed as

$$dBZ = 10 \log_{10} \left[\frac{Z}{1\ mm^6\ m^{-3}} \right] \tag{2}$$

For the purposes of this study, dual-polarization variables and doppler velocity were not considered. However, previous literature (summarized in Kumjian et al, 2022) has found utility in distinguishing snow crystal habit using differential reflectivity, and correlation coefficient has historically been used as a proxy for the melting layer, and also has implications for particle morphology, especially in conjunction with other dual-pol variables. Additionally, vertical gradients in one or more of these dual-pol variables can also elucidate other physical cloud processes. A more robust study in the future would include these as to include a more comprehensive microphysical picture of snow squalls from radar.

Reflectivity data was processed using the Py-ART python package (Helmus et al., 2016), and interpolated from multiple elevation angles to a constant altitude plan position indicator (CAPPI) grid of 0.5 km spacing and 1 km AGL height, with further specifications found in Table 1. The exact methodology of this scheme can be found within the Py-ART documentation, but the "map.grid_from_radars" function was applied on the level II radar files using a nearest-neighbor interpolation scheme. Although recent literature such as Brook et al. (2022) acknowledge the potential limitations of the nearest-neighbor interpolation scheme, they still note relative accuracy of nearest-neighbor when compared to other interpolation methods. Future work may look to refine and/or change this scheme. 1 km AGL was chosen as the constant height to interpolate the data to, as this was elevated enough from the surface to negate ground clutter and noise, while not being so high as to deviate from the core of the precipitation, which has been found in previous literature (Lundstedt 1993; Pettegrew et al. 2009; Banacos et al. 2014) and qualitatively through this work to be close (~0.5-1 km) to the surface.

Grid Extent	(125 km x 125 km)
Grid Resolution	0.5 km
Altitude of Interpolated Grid Points	1 km
Interpolation Scheme	Nearest-Neighbor

Table 1. Characteristics of the CAPPI grid created from the radar data

There are a few caveats to this approach: first, interpolating polar coordinate data to a Cartesian grid has to account for impacts of varying VCPs. WSR-88D operators will change the frequency, and duration of azimuthal and zenith angle scans underneath changing weather conditions and phenomena, typically at the behest of NWS forecasters. This creates a wide and varied range of PPI elevation angles for any given radar scene, which will therefore affect the quality of the resultant CAPPI. To understand and quantify the effect that this might have on the resultant data, the VCP was retrieved for each timestamp and is plotted in Figure 5. A general overview of the VCPs and their coverage patterns is laid out in NOAA (2022), but generally this dataset consists of equal amounts of “clear-air” (31, 32, 35) and “precipitation mode” (212, 215) VCPs.

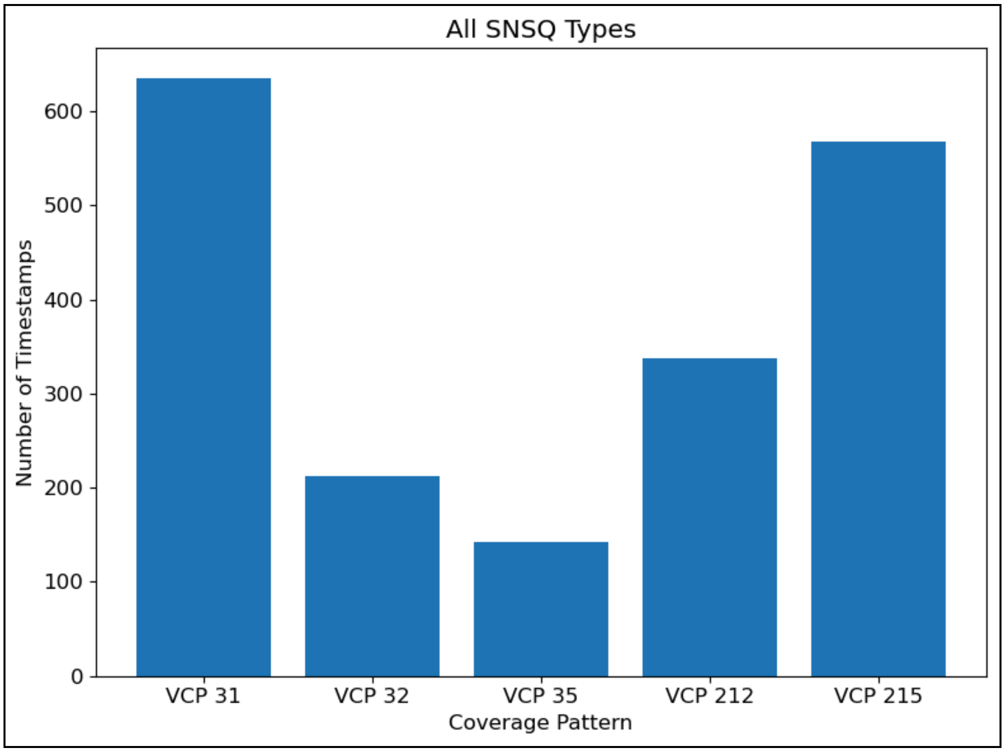

Figure 5. Histogram showing the frequency of different VCPs within the dataset.

2.3. CLAVR-x

2.3.1. Overview

The Extended Clouds from AVHRR system (CLAVR-x) is a satellite processing algorithm which retrieves microphysical cloud properties from satellite imagers, though this study only uses geostationary data from GOES-16. The processing framework for CLAVR-x was developed as a replacement for the original CLAVR-1 clear/cloudy sky classification framework described in Stowe et al. (1999). Initially developed as a joint project between the National Oceanic and Atmospheric Administration (NOAA) and the University of Wisconsin-Madison Cooperative Institute for Meteorological Satellite Studies (CIMSS), the algorithm has been developed to support several satellite imagers, including the GOES Advanced Baseline Imager (ABI).

CLAVR-x produces several atmospheric cloud and surface products, but most products are either produced from or derived by three main algorithms: the Enterprise Cloud Mask (ECM; Heidinger et al., 2012), the Enterprise Algorithm working group Height Algorithm (ACHA; Heidinger and Pavolonis, 2009), and the Daytime Cloud Optical and Microphysical Properties algorithm (DCOMP; Walther and Heidinger, 2012). These algorithms are operated underneath several modes to reflect the different combinations of channels available from supported imagers. Each DCOMP mode requires at least two channels; one from the visible spectrum, and one from the infrared (IR). For GOES-16, every mode (for the version of CLAVR-x used in this study) utilizes channel 2 [0.64 μm, “red” visible band] for the visible. For the IR, the following channels in the near-IR are used (with respect to each DCOMP mode):

1. DCOMP mode 1: Channel 3 [0.86-μm, “veggie” band],
2. DCOMP mode 2: Channel 4 [1.46-μm, “cirrus” band],
3. DCOMP mode 3: Channel 5 [1.6-μm, “snow/ice” band].

Note the difference in channels when compared to GREMLIN. While GREMLIN focuses on the far-IR and microwave, CLAVR-x uses the visible and near-IR. Because DCOMP uses channels within the visible, the algorithm can only generate data during the daytime (defined as solar zenith angle at any given pixel being less than or equal to 65°). The difference between these DCOMP modes and their subsequent output is reflected in several of the produced CLAVR-x cloud products discussed in the following subsections.

2.3.2. CLAVR-x Products

While CLAVR-x generates several products, only a few are particularly relevant to shallow convective snow. Recall from equation 1 how radar reflectivity varies with the integrated sixth moment of the PSD. The PSD describes fundamental characteristics of a cloud; along with quantifying the number and distribution of cloud particles, it can describe the total water content and mass of a cloud. Therefore, it is an often sought-after variable within both passive and active remote sensing. Two critical fields retrieved by CLAVR-x, which in part rely on the PSD, are cloud optical depth and effective radius.

These products will be described broadly in the following subsections, with the caveat that a more detailed technical and theoretical description can be found within Walther and Straka (2020), and Pavolonis and Calvert (2020).

Cloud Optical Depth

Cloud optical depth is a quantity which describes the extinction of radiation at a particular wavelength by a mass of cloud within some given volume, and can be mathematically described with the following expression:

$$\tau = \int n(r) Q(r, \lambda) \pi r^2 dr \tag{3}$$

where τ, the optical depth of a column of the atmosphere of unitless dimension, is equal to the integrated particle size distribution ($n(r)$) and scattering coefficient ($Q(r, \lambda)$) across all particle size radii "r". The scattering coefficient is pre-calculated using Mie code within the CLAVR-x algorithm. The particle size distribution is approximated by a typical gamma distribution, expressed as:

$$n(r) = r^{\frac{1-3r_b}{r_b}} \exp\left[\frac{-r}{r_e r_b}\right] \tag{4}$$

where r_b is a dispersion constant shaping the particle size distribution (PSD), and r_e is the effective radius, to be described in more detail in the following section. The dispersion constant r_b (described as the mean effective variance, a measure for the distribution width) is assigned a value of 0.1. The particle size distribution within clouds and precipitation is a notoriously difficult variable to constrain, especially for snow/ice (Wood et al. 2013). This uncertainty compounds downstream and has implications for the remote sensing of snowfall from spaceborne observation (Liu 2008; Kulie et al. 2010; von Lerber et al. 2018). Therefore, the PSD will be a major source of uncertainty going forward.

Cloud Effective Radius

The effective radius of a cloud can be most succinctly defined as the radius all particles in the cloud would have, for a given set of cloud scattering properties, if the cloud had a monodisperse particle population (every particle within the cloud has the same radius). Mathematically, it is defined as the ratio between the third and second moments of the particle size distribution, like so:

$$r_e = \frac{\int r^3 n(r) dr}{\int r^2 n(r) dr} \tag{5}$$

Such a quantity is important, as it is a bulk microphysical footprint of the cumulative microphysical sub-processes which affect the radiative and scattering properties of a cloud.

Cloud Water Path

The cloud optical depth describes the extinction of radiation within a column by cloud, and the effective radius characterizes the idealized bulk size of the particles given a certain scattering signature, and so a combination of these two would yield a quantity which describes the integrated mass of the cloud within a given pixel. This quantity is called cloud water path (CWP), and is given by the (relatively) simple expression:

$$CWP = \frac{5}{9}\rho_l r_e \tau_c \tag{6}$$

where ρ_l is just the density of liquid water.

Cloud Height, Phase, Type

For brevity, the rest of the CWP variables will simply be summarized. A full table containing all variables used as a part of this work is provided in Table 2. Cloud phase describes the predominant phase of cloud particles within a sampled pixel (liquid, ice, supercooled liquid, mixed), and cloud height describes the top-of-cloud height of the sampled pixel. Both are fundamentally derived from cloud type as described in Pavolonis and Calvert (2020).

CLAVR-x Variable	Description	Units
Cloud Optical Depth	Vertical thickness between the top and bottom of the atmospheric column	Unitless
Cloud Effective Radius	Radius of cloud particles assuming a monodisperse population	μm
Cloud Water Path	Total cloud water mass in a cloud column	g m^{-2}
Cloud Phase	Phase of the cloud pixel near its top	N/A
Cloud Top Height	Estimated maximum height of the cloud pixel	meters

Table 2. List of CLAVR-x variables used, a brief description for each, and the subsequent units.

2.4. Colocation

2.4.1. Temporal Colocation

Once both the radar and satellite data are obtained, a process to match them up temporally ensued. Each piece of data had a different temporal frequency; the CLAVR-x data were taken from 5-minute GOES-16 CONUS scans. The radar data temporal frequency depends on the VCP, but ranges from ~5 to 10 minutes. Given that CLAVR-x has a more consistent temporal frequency, all available CLAVR-x timestamps were found within each Schneider (2024)-identified SNSQ hour. This approach equates to ~12 timestamps per hour, disregarding outages. Out of the 7319 potential snow squall timestamps, 68% had to be removed because they occurred when the satellite zenith angle was less than the 65° threshold, rendering the visible channel, and therefore CLAVR-x daytime products, unusable. The total usable count was thus reduced to 2336 files, which after some additional filtering for not generated and broken data, produced 2246 timestamps.

From this list of CLAVR-x times, the closest radar scan was found and colocated. Any time delta (the difference between the CLAVR-x and KCCX radar time) greater than 5 minutes was thrown out. The time delta between the datasets are shown in Figure 6. Generally, the temporal matchup stayed well within 5 minutes but a few outliers persisted. This filter left 2234 times. A few more timestamps were removed after this which were additional non-snow squall types, which left the final, usable timestamp number at 1894. Table 3 summarizes the resulting dataset, both in terms of the number of individual colocated timestamps and the number of events contained within each type.

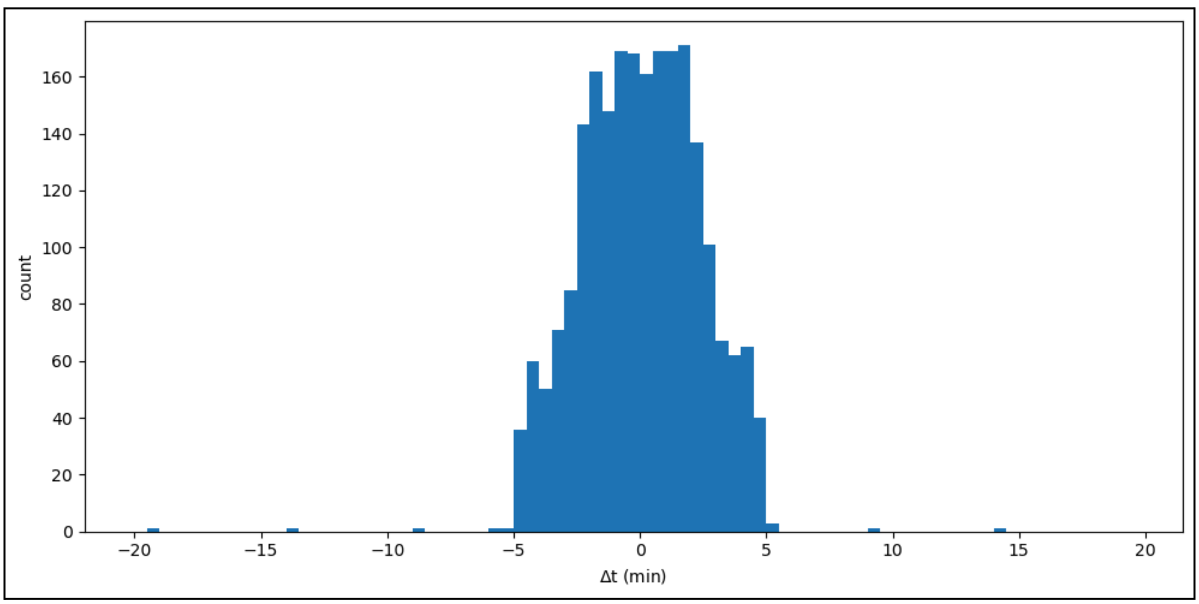

Figure 6. Histogram showing the CLAVR-x - radar colocated timestamp difference, in minutes.

Snow Squall Type	Number of Colocated Timestamps	Number of Events*
S1	73	3
S2	491	22
S3	1286	38
S4	44	4
Total:	1894	51*

Table 3. The number of colocated timestamps and snow squall events attributable to each snow squall type.

*Note that one unique snow squall event can have multiple snow squall types, so the total number of unique snow squall cases is 51, but the number within each row represents the number of events which had at least one occurrence of the type.

2.4.2. Spatial Colocation

After the final list of snow squall event times was obtained, radar and CLAVR-x data were spatially colocated. This colocation process was complicated by the fact that each dataset was contained on entirely separate grids. Therefore, a method to match these datasets onto a common, colocated grid was sought. The Earth System Modeling Framework (ESMF) created a universal regridding tool intended for geospatial data in python, called "xESMF" (Zhuang et al. 2025). The details of the tool can be found in Zhuang et al. (2025), but a brief summary follows. First, a set of weights is created for a linear transformation between the latitude and longitude grids of the two fields being colocated. This calculation needs to only be done once, after which the weights can be saved and reused. Several regridding methods are supported through xESMF, but the "conservative" method was chosen for this study, primarily because it was of utmost importance that the integrity of the reflectivity field was preserved as it was regridding to the CLAVR-x grid. Once this regridding is done, the radar and CLAVR-x dataset is on a common, shared grid of equal size to the CLAVR-x data (125 km square grid).

2.4.3. High-Level Data Filters

The combined dataset was further filtered by applying a 125-km upper-bound distance threshold, wherein any pixels greater than 125 km away from the KCCX radar were ignored. This step was accomplished by applying the following haversine distance equation to the latitude and longitude fields pixel by pixel:

$$d = 2R_E \arcsin\left(\sqrt{\sin^2(\frac{\phi_1 - \phi_2}{2}) + \cos(\phi_1) + \cos(\phi_2) + \sin^2(\frac{\theta_1 - \theta_2}{2})}\right) \quad (7)$$

wherein the distance between any two points on the Earth's surface can be calculated by just knowing each point's longitude and latitude, and the radius of the Earth R_E, which was approximated to be 6371 km. This process also allows the data to be analyzed based on the distance from the radar.

2.5. Analytical Methods

2.5.1. Feature Tracking

In order to further investigate the evolution of the macro- and microphysical cloud characteristics of snow squalls, a broad scene-by-scene view can only do so much. Therefore, the following question is posed; if one were to separate a scene by individual features for analysis, does this yield additional information of significance?

Such an analysis can be done with the Tracking and Object-Based Analysis of Clouds (TOBAC) Python module (Heikenfeld et al., 2019). This program automatically identifies and subsequently tracks features within any field, but is most commonly utilized within meteorological data. The parameters of this identification and tracking algorithm can be tuned to the user's specifications. The options used in this study can be seen in Table 4. The algorithm was applied to both the reflectivity and CWP fields, but only results from the CWP feature tracking will be shown.

Detection Parameter	**Radar Reflectivity**	**Cloud Water Path**
Threshold Value	5, 10, 15 dBZ	150, 300 g/m^2
Target	Maximum	Maximum
Position Threshold	Weighted Difference	Weighted Difference
"n" Erosion Threshold	2	2
Sigma Threshold	1	1
"n" Minimum Threshold	4	4
Segmentation Parameters		
Target	Maximum	Maximum
Method	Watershed	Watershed
Threshold	10 dBZ	150 g/m^2

Table 4. Specifications of the TOBAC feature identification and tracking algorithm

2.5.2. Gaussian Filtering and Averaging

Despite efforts to mitigate poor snow squall matchups, it is inevitable that radar/CLAVR-x features will not be perfectly matched. Although the radar-satellite matchups are typically within a few minutes of each other, the time delta between the two fields is enough that progressively higher mean wind speeds could induce higher spatial offsets. The satellite viewing geometry (from an oblique angle) can also induce a spatial radar-satellite offset, though this issue is likely minimized in this study due to the shallow vertical extent of snow squalls (Vicente et al. 2002). Still, this uncertainty exists, compounded by CWP fields that can be spatially inconsistent— a topic that will be explored in more depth later.

To mitigate these issues, Gaussian filter (or “blur” more qualitatively) sensitivity tests were applied to the dataset. This approach was facilitated by the SciPy Python package (Virtanen et al. 2020), through the “scipy.ndimage.gaussian_filter” module via a Gaussian convolutional kernel. The blur was applied to the raw CWP field (no cloud mask, no filters applied to the CWP field). After applying the blur, the field was then masked as usual. Several sigma blur thresholds were applied, which will be discussed in more detail in the results section.

Chapter 3: Results

3.1. RQ1: Spatial Match-up Results

This section will interrogate the rate at which CLAVR-x overlaps with the radar reflectivity field (i.e., to what extent can CLAVR-x identify precipitating pixels). CWP is chosen as the field to interrogate this overlap, due to its relevance with radar reflectivity described in Section 2.3.2. The results within this section will be shown within a distance range with respect to the radar; in other words, the dataset will be filtered based on whether it falls within the set distance from the location of the KCCX (within 125 km). The primary motivation behind this decision can be found in Figure 7. It is expected, as described in previous sections, that the quality of the radar data will considerably worsen with distance; either the radar will overshoot more snow squalls with distance, or beam sampling volumes get large enough that the radar cannot adequately resolve the squall. Therefore, the sensitivity of the radar-satellite matchups to distance from the radar site must be investigated. If obvious sensitivity is demonstrated, then a radar range threshold must be defined that yields optimal radar-satellite matchups. Two distance thresholds were therefore chosen: 0-75 km, and 25-100 km.

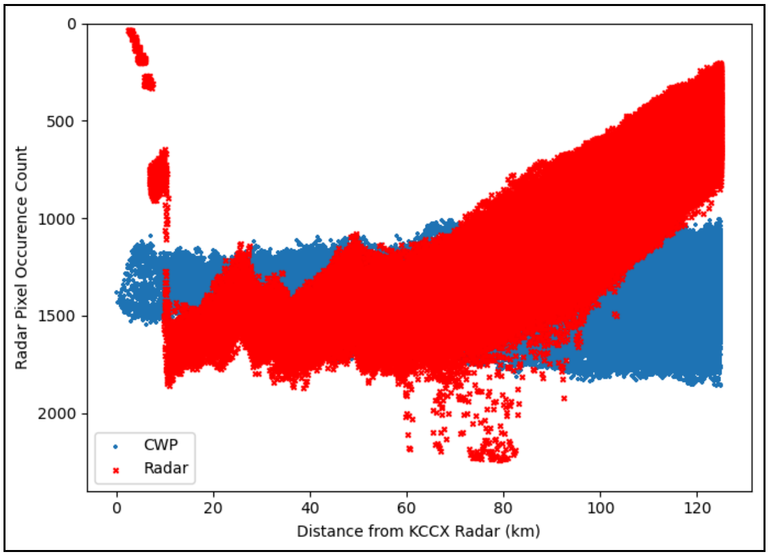

Figure 7. Scatterplot showing the total pixel occurrence count (y-axis) vs. distance from the KCCX radar (x-axis). Colored in red is this pixel occurrence for radar pixels, and in blue for CWP pixels.

These thresholds were chosen for a few reasons; Figure 7 shows that the majority of colocated points lie less than 100 km, and greater than ~10 km from the radar, though there is quite a bit of variability in the closest tens of kilometers most likely due to CAPPI variability from the different VCPs chosen. Therefore, the 0-75 km range encapsulates a sort of "limited range" thinking for colocation—that the best matchups will be the closest matchups (i.e., those within a range in which we are more confident that the radar fully samples the squall). The 25-100 km range offers a different, more data-driven approach. Based on Figure 7, colocation issues exist both very close to and far from the radar location; therefore, a "medium-range" (25-100 km in this case) set of collocation points were considered as a plausible optimal range.

It can be argued that these are subjective ranges which require more objective backing, and indeed it is true that the ranges chosen were based on qualitative reasoning rather than anything strictly empirical (and therefore less replicable). The crux of the uncertainty in this problem comes from the fact that it is difficult to ascertain whether the radar is "actually" underestimating or missing a squall, or whether the squall is actually weak or non-existent. Further, snow squalls and environmental conditions vary from case to case; it is indeed possible to have instances where snow squalls are deeper (> 3 km) or more shallow (< 2 km), or thermodynamic profiles more or less favorable for beam refraction, which therefore changes the distance at which squalls being to be underestimated/missed.

This uncertainty is acknowledged by the author, and therefore it is stated that these distance thresholds are initial best guesses, and a future methodology would involve a more thorough interrogation of 1) estimated snow squall heights, 2) beam refraction and the lowest beam height at each range ring, and 3) CAPPI creation which is more dynamic/robust by taking 1) and 2) into account, along with subsequent significance tests to determine which factor controls these optimal distance thresholds (if such a relationship exists).

In order to quantify how well the radar and CLAVR-x products matched up in space, a binarization method was applied to both products. From CLAVR-x, CWP was used given antecedent work showing its relationship with reflectivity (which will be shown within this dataset later), though theoretically other fields within the CLAVR-x dataset could be used as well, with thresholds adjusted accordingly. For the reflectivity, any values which exceeded -25 dBZ (the lowest dBZ value recorded in the dataset) were kept, and the rest discarded. In the CWP field, any value less than 1 was discarded. The dataset was then binarized based on those thresholds, and it was found where: 1) only the radar identified a pixel, 2) only the satellite identified a pixel, and 3) both radar and satellite identified a pixel. A representative plot of this new “overlap” field is shown in Figure 8, of which will be referred to as the "success ratio”.

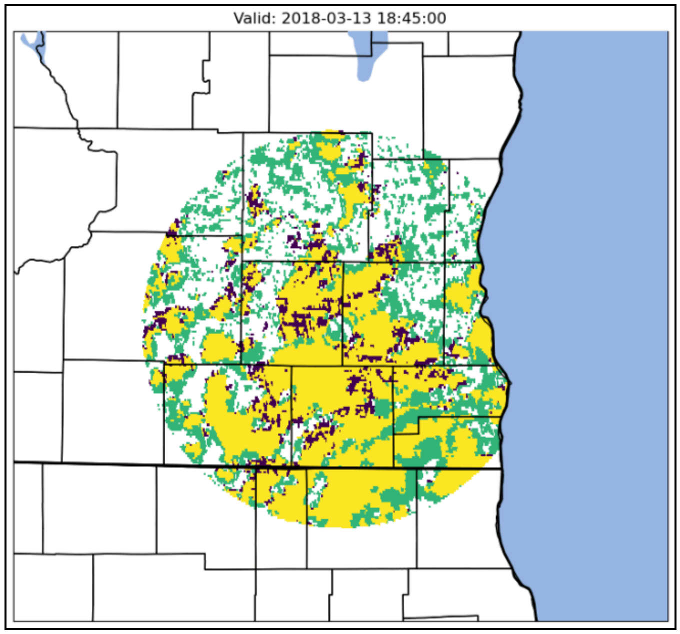

Figure 8. Plot centered on Milwaukee, Wisconsin demonstrating the "spatial overlap" index. Yellow pixels indicate radar and satellite have identified a pixel, green is satellite only, and purple is radar only.

This index allows, in a rudimentary way, the ability to quantify how well the CLAVR-x and radar are matched up. A simple expression used to show this is the following:

$$S = \frac{n_{p,ovp}}{n_{p,total}} \times 100\% \tag{8}$$

where S, or the so-called "success ratio" is defined as the number of "overlapped pixels" (i.e., the number of pixels wherein both radar and satellite have identified something) divided by the total number of pixels within the scene, and turned into a percentage. This percentage quantifies how much the reflectivity and CWP fields are "matched up" spatially. A score of 100% would indicate that every single reflectivity pixel has a match in the CWP field. This does not say anything about the relative magnitudes of the pixels, so it may very well be that the pixels are matched up spatially, but not in magnitude. Such analysis will be left to future sections.

Given this metric's ability to quantify match up quality, bulk statistics were calculated across the dataset. Other than quantifying the radar-satellite match up quality, the goal of this analysis was to determine whether success ratio values varied across distance-from-radar thresholds to address the concern brought up earlier in this section. As such, the results from different distance thresholds will be shown.

0-125 km

Figure 9 shows the spatial matchup rates between all snow squall types. Type 2 and 3 squalls have the highest (~40-60%) match up rates, with type 1 having the lowest (~10-30%) and type 4 having the highest (~70-80%). This generally makes sense; for example, figure Lambda shows a representative success ratio plot example from each snow squall type. Generally, type 1 snow squalls are associated with broader, synoptic scale forcing (often in the form of an arctic front or some equivalent boundary) so it is more likely that non-precipitating cloud cover populates any given scene. Conversely, type 4 squalls are more widespread and less intense. Therefore, the reflectivity and CWP fields are much more likely to be aligned.

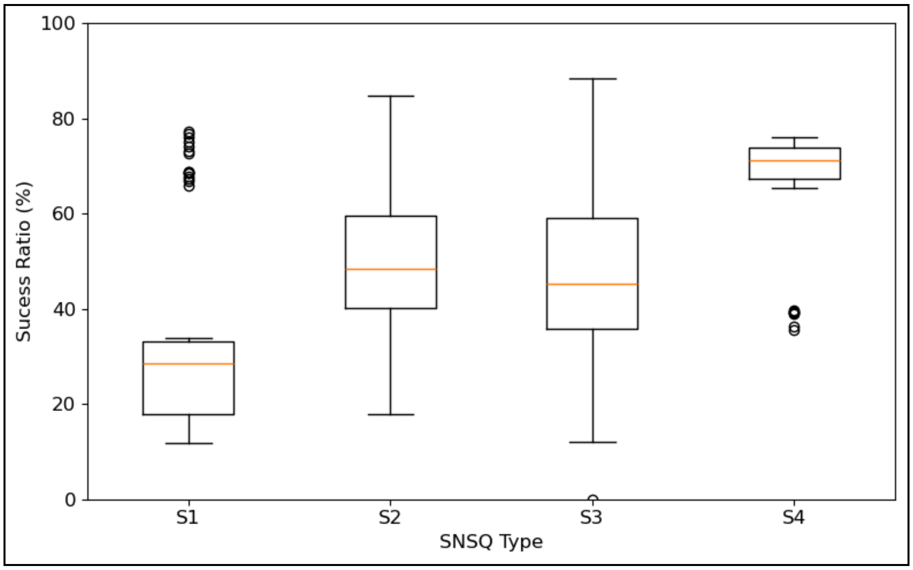

Figure 9. Box and whisker plot showing the success ratio distribution as a function of snow squall type within 0-125 km from the radar

To contextualize the results with the VCP and CAPPI issues discussed earlier, Figure 10 shows how the success ratio varies based on the VCP used. Clearly, VCP 215 has significantly lower success ratio values on average, with its median almost 20% lower than VCP 35. For the other VCPs, success ratio values are generally around ~50%, with values approaching 70-80% not uncommon. These VCP modes will be discussed in greater detail in Section 3.3.1.

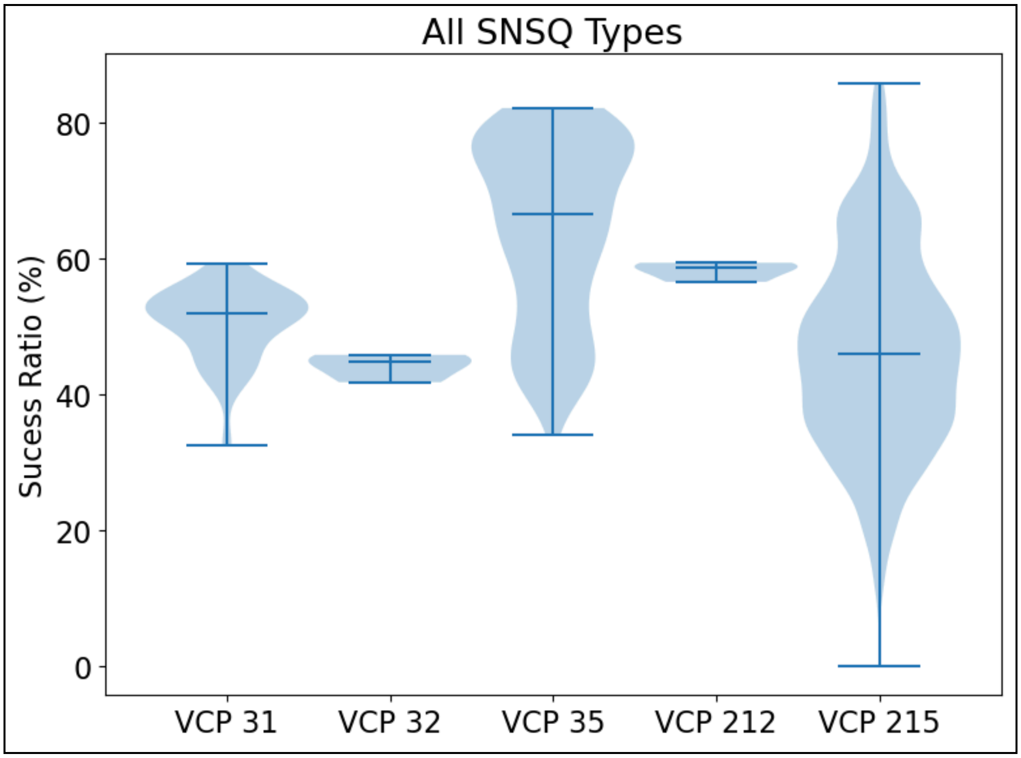

Figure 10. Violin plot showing the distribution of success ratio values based on the VCP used.

0-75 and 25-125 km Results

To see how introducing different distance thresholds impacted the spatial match up scores, Figure 11 shows the success ratio scores like those shown in Figure 9, but now including the 0-75 km and 25-125 km results. There is little difference between the 0-125 km and 25-125 km results, but 0-75 km shows significant improvement, with its median often ~10% above the other distance thresholds. These results may lead one to believe that using the 0-75 km threshold would yield higher spatial match-up scores, and therefore more accurate results. However, a key caveat is that a steadily increasing number of points with each radial ring farther from the radar exists; the implication being that the 0-75 km results are simply an underestimation, and not truly representative of any larger trends in match up accuracy.

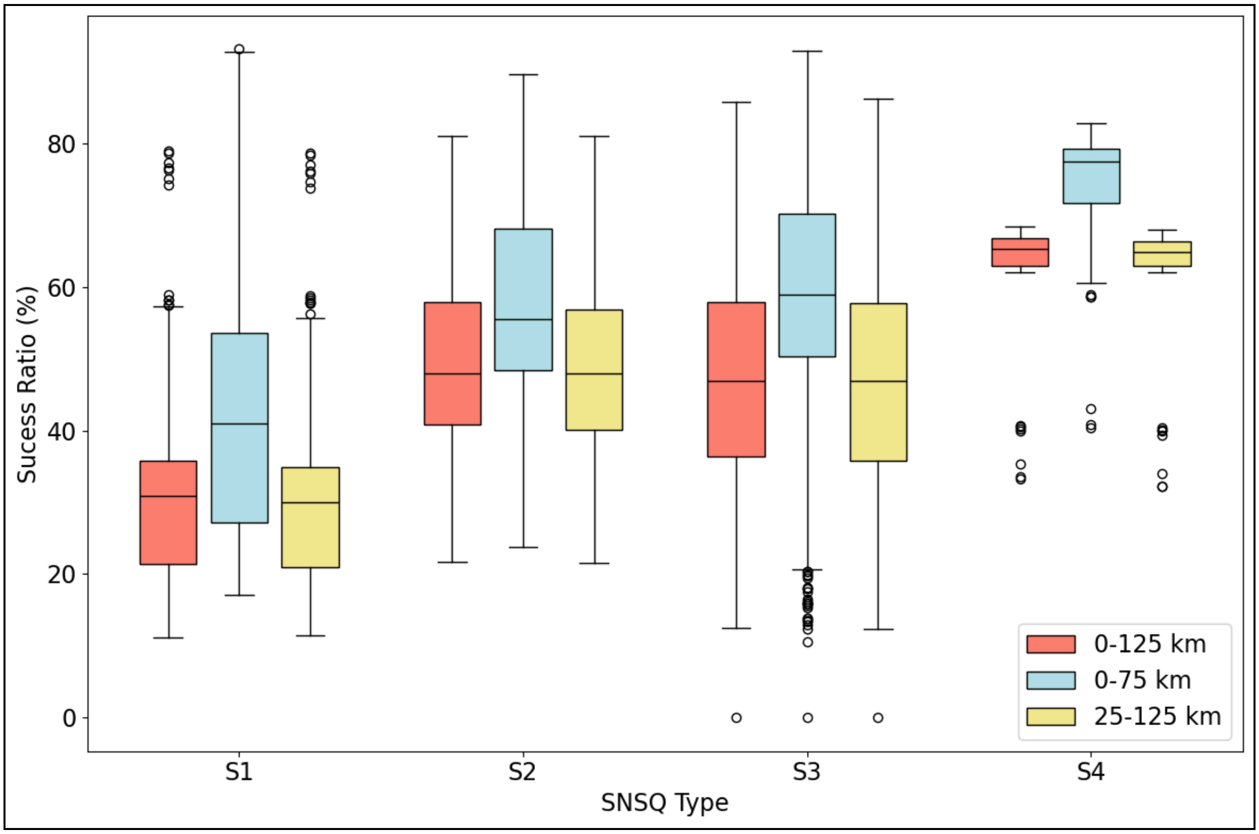

Figure 11. Box and whisker plot showing the success ratio distribution as a function of snow squall type within 0-75 km from the radar

Regardless of the difference among distance thresholds, success ratio values still fall within 40-70% for most snow squall cases, with notable exceptions on either end. This "success ratio" will be continually explored within the context of new results.

3.2. RQ2: Empirical Relationship

Now that a basis for the CLAVR-x/radar match up has been established (in other words, that sufficient evidence has arisen which corroborates the spatial validity of such a match up in the first place), a question of great interest arises; do the physical quantities observed between CLAVR-x and radar have any meaningful empirical relationship? This question will be interrogated within the context of all CLAVR-x data variables described within Table 2, but of principal interest is the potential relationship between CWP and radar reflectivity.

Recall that CWP is the measure of the total integrated cloud water content within a column (or in this case per-pixel, as defined by equation 4). It is related to the product of the optical depth of the column and the effective radius, the latter being a ratio between the third and second moment of the particle size distribution. Radar reflectivity is the integrated sixth moment of the drop size distribution. While ground-based radar and spaceborne satellite imagers fundamentally measure different quantities, it is theorized (for shallow, discrete convection) that the optical properties as viewed from satellites vary in proportion with radar reflectivity.

3.2.1. Initial Results

Wherever CWP and reflectivity met the binarized thresholds described in section 3.1, 2D density plots were created. Both the x-axis (radar reflectivity) and CWP (y-axis) are logged, and this is primarily because the two are expected to follow an exponential relationship (recall again the relationship each has to the PSD). This form would allow a linear regression equation to be drawn and an empirical relationship to be established. Figure 12 shows these density plots between CWP and radar reflectivity for all snow squall cases, grouped by type, and for all distances from the radar. One of the most prominent features in each subplot is a clustering of points along a mostly diagonal axis, the majority of which lie centered at 0 dBZ reflectivity and ~32 g/m^2 CWP values.

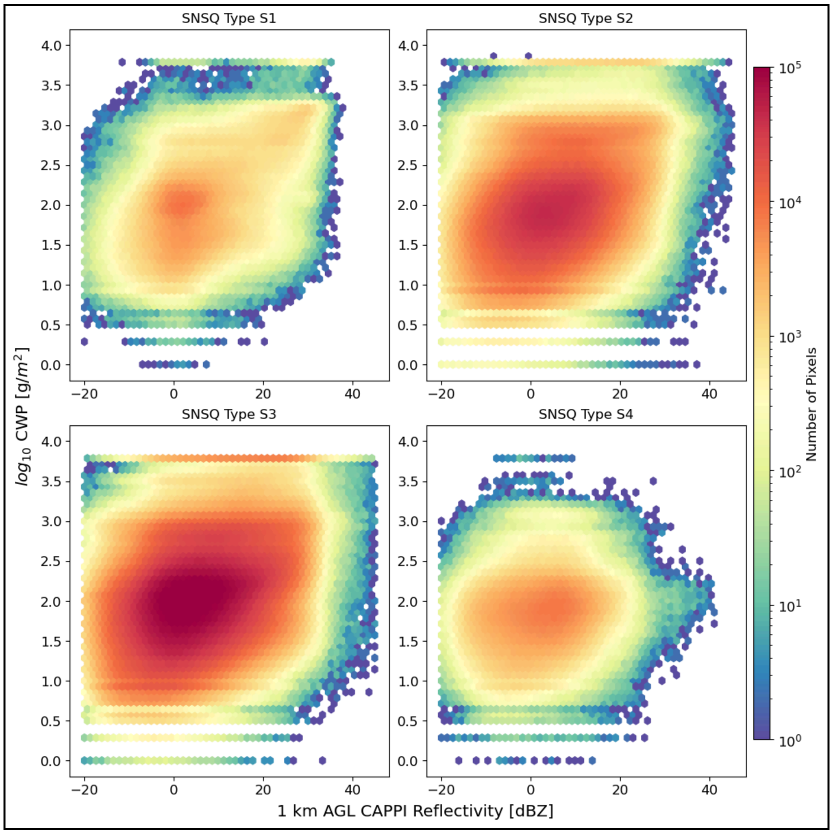

Figure 12. 2D Histograms of logged CWP (y-axis) vs. reflectivity (x-axis), broken down by snow squall type. Warmer colors indicate a higher density of points.

Qualitatively, a large amount of points tend to follow the rough orientation of that diagonal axis, but for any one reflectivity value there can be many different CWP values associated with it, spanning several orders of magnitude. This is further corroborated by analyzing every snow squall case in Figure 13. The CWP-reflectivity relationship varies significantly from case to case, with each one having different slope magnitude, point clustering patterns, concentrations, etc.

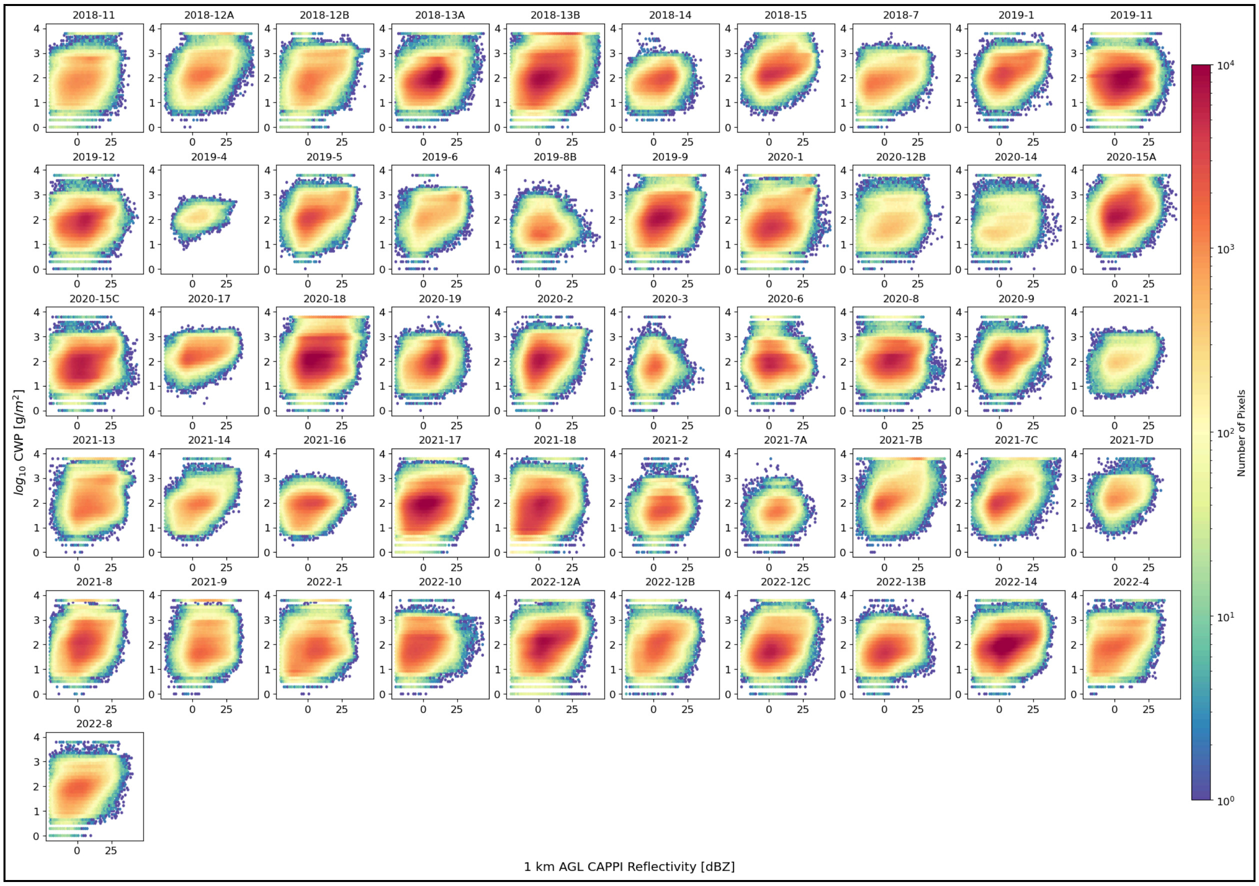

Figure 13. As in Figure 12, but instead broken down by each unique snow squall case.

Given the results of Figure 12, it comes as no surprise then that the R^2 score results from the dataset are generally very poor. Figure 14 shows these scores, again separated by snow squall type and distance from the radar. One of the more surprising elements of this figure is the relative uniformity of the results regardless of distance from the radar, with the exceptions perhaps being S1 and S4, whose 0-75 km range has slightly higher scores. Another element of note is the wide percentile range of R^2 scores, especially for snow squall type 1 whose 25th and 75th percentile values range from below 0.1 to past 0.3, respectively. The minimum and maximum whiskers extend even farther, from nearly 0 (min) to nearly 0.4 (max). S2 through S4 show less variability within the interquartile range, but have similar variability in their min/max whisker range.

Therefore, despite there being theorized a CWP vs. reflectivity relationship (wherein each increases in proportion), the data shows such a connection is not straightforward to make.

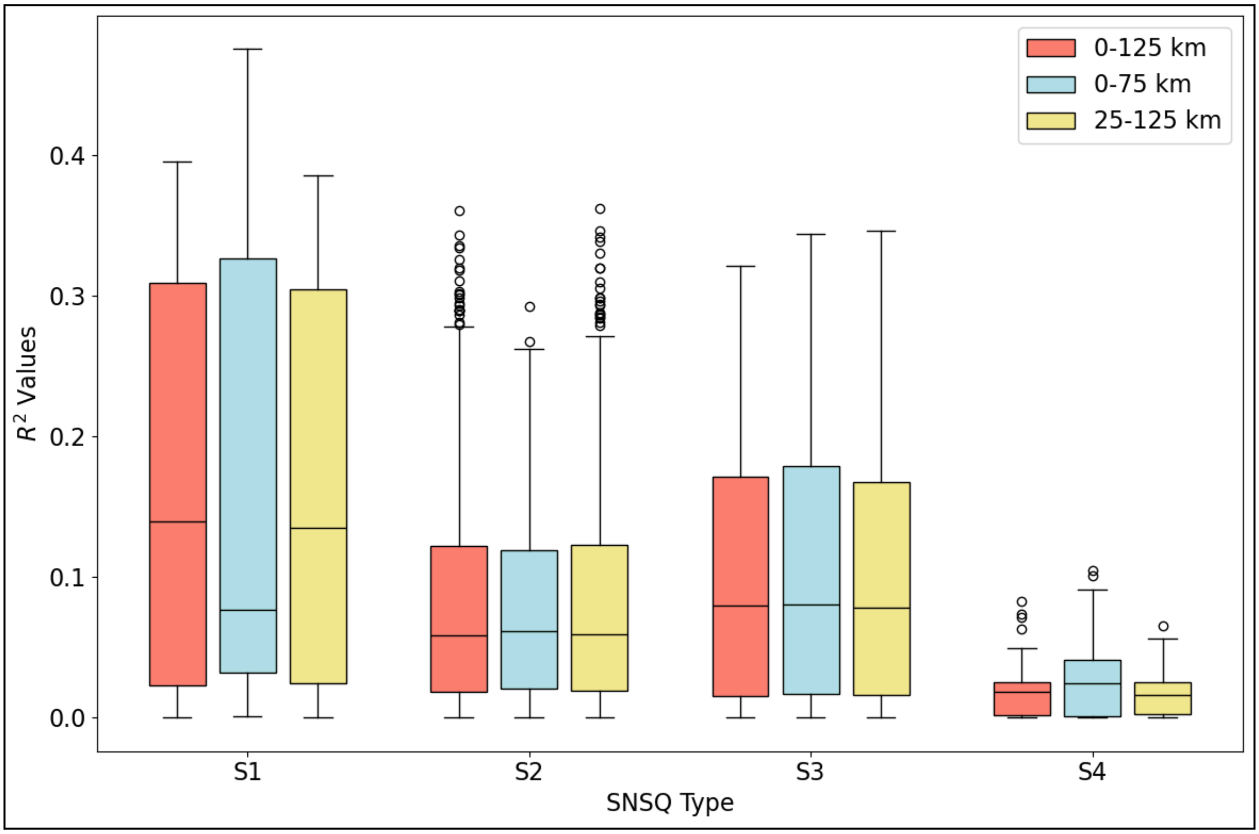

Figure 14. As in Figure 11, but now showing R^2 scores for each snow squall type.

Framework for Further Data Filtering

The process for determining a meaningful answer to this empirical match-up question is not straightforward; it is complicated by the numerous physical and technical challenges in 1) observing atmospheric phenomena with ground-based radar and geostationary satellites, and 2) matching those observations together in space (both horizontal and vertical) and time. In other words, it is not as easy as putting the matched reflectivity and CWP pixels together and creating a regression equation fit.

This will be contextualized in greater detail in the discussion section. For now, these challenges will frame how the results are discussed, broken down in the following manner:

1. **Satellite Retrieval:** involving CLAVR-x parameters like cloud top height, cloud phase, sensitivity of cloud effective radius to DCOMP mode used, etc.
2. **Viewing Parameters**: factors influencing the geometric and temporal specifications of the satellite and radar retrievals. This complication will primarily be addressed through gaussian filtering.
3. **Radar Retrieval:** dealing with the complications of the CAPPI method used, the underlying VCP, etc.

The goal for the points listed here is to serve as guidelines for filtering processes which aim to 1) address the scientific and data integrity concerns they bring, and 2) work towards improving the CWP vs. reflectivity relationship such that it is less diffuse, and therefore serves as a more accurate model.

3.3. Data Filtering

3.3.1. Radar Retrieval

As briefly mentioned in the methodology, there are several concerns when translating native radar data to a Cartesian grid. Many methods and procedures have been developed for various use cases, but for this study a CAPPI created with a nearest-neighbor interpolation scheme and 1km AGL height have been chosen. While the MRMS gridded radar dataset is the frequent choice for such analysis, there are several issues using it for this study: 1) MRMS data via official NOAA archived databases is limited to post-October 2020, with only a small subset of data available from non-NOAA sources before this time, and 2) the product which is available (hybrid-scan reflectivity, or HSR) tends to have a limited range of reflectivity values, especially on the lower (~0 dBZ). While the 1km native MRMS resolution is generally sufficient for snow squalls, the higher 0.5 km resolution via a custom CAPPI proved to be useful in initial analyses.

Although the author acknowledges that sensitivity tests to these chosen parameters would be ideal, this section will instead choose to focus on interrogating how breaking the data down by VCP affects the results, and a more comprehensive investigation would include both.

Sensitivity of Results to VCP

It was shown that the CWP vs. reflectivity relationship was diffuse. In order to quantify the factors which may lead to such diffusivity, a quality analysis of the gridded reflectivity dataset is imperative. One area of quantifiable uncertainty is the VCP employed by the radar operator, as discussed in the radar portion of the methodology section. Changing the VCP can affect the resulting CAPPI; this is shown in Figure 15, where changing the VCP results in more or less artifacts in the resultant grid depending on the VCP. Qualitatively, the issues seem to be most severe for VCP 31 or “clear-air mode”, where multiple “rings” around the radar can be seen. The other plots are not free of visual artifacts—they all maintain the so called “cone of silence” within the nearest few kilometers of the radar where values are unobtainable due to very low beam heights— but VCP 31 is different in that it does not do as many vertical scans as the other VCPs, and is thus less amenable to Cartesian gridding (NOAA 2022).

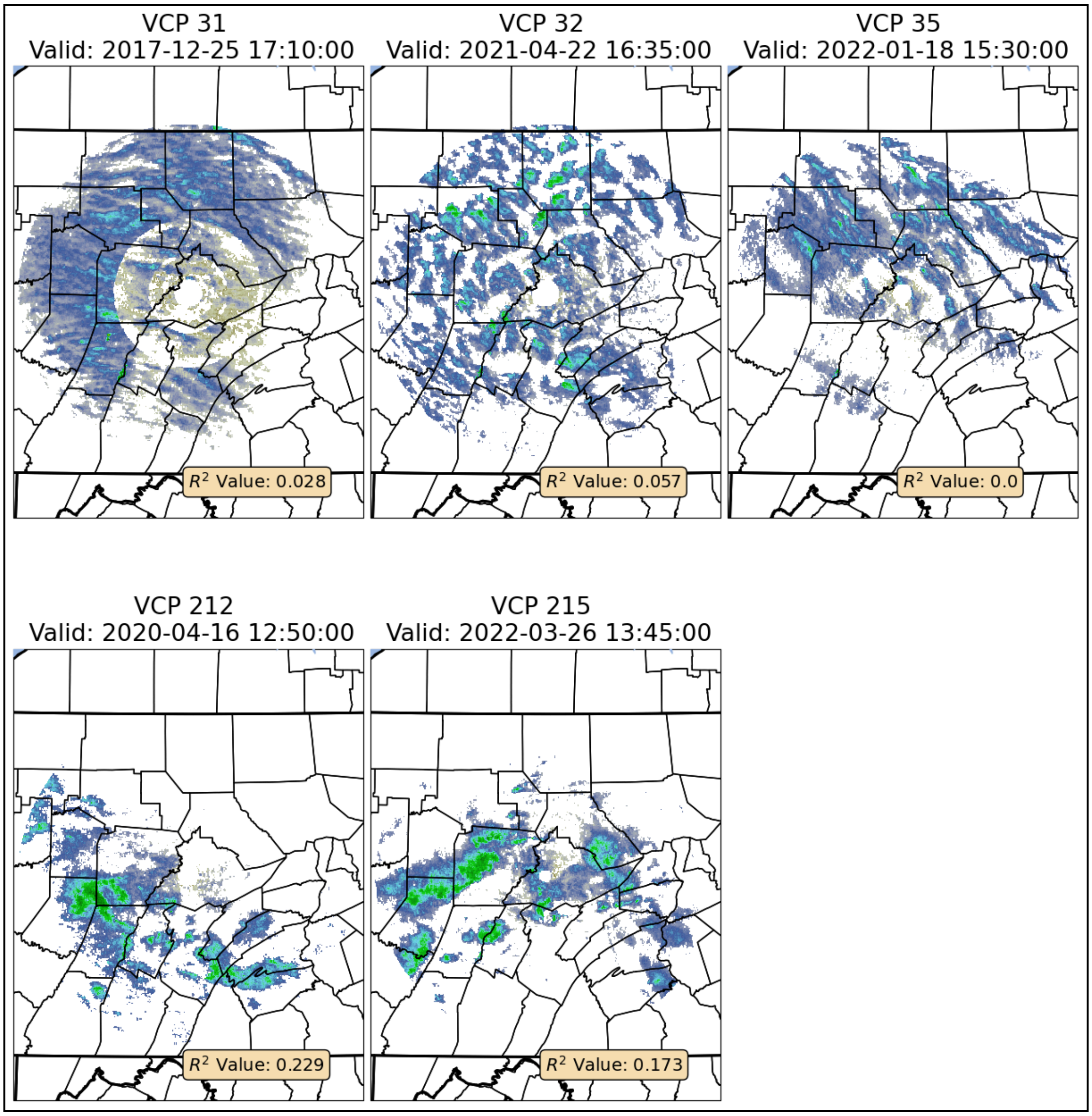

Figure 15. Multi-panel plot showing CAPPI reflectivity for various VCPs. The R^2 value shown is with respect to CWP (no gaussian filtering).

The impact of each VCP on the data is tangible. Figure 16 shows how the distribution of R^2 scores varies with the VCP chosen. What is immediately apparent is how low scores are for the clear-air mode VCPs (between 0 and 0.1 mainly for 31 and 32) and how much they increase for the precipitation-mode VCPs (between 0.05 and nearly 0.3 for 212 and 215). While scores are quite a bit lower than what one would like to see for creating a linear regression model, it is clear that the choice of VCP may have an effect on the quality of the regression fit. A caveat with these results is that filtering for VCP may also unintentionally be filtering for snow squall type/morphology, as the VCP used is usually based upon ancillary information (time of year, precipitation type, potential hazards, etc.). It's already been observed that type S4 squalls have the worst R^2 scores on average, and S1 has the best, so it must be made clear what the breakdown is of snow squall type for each VCP.

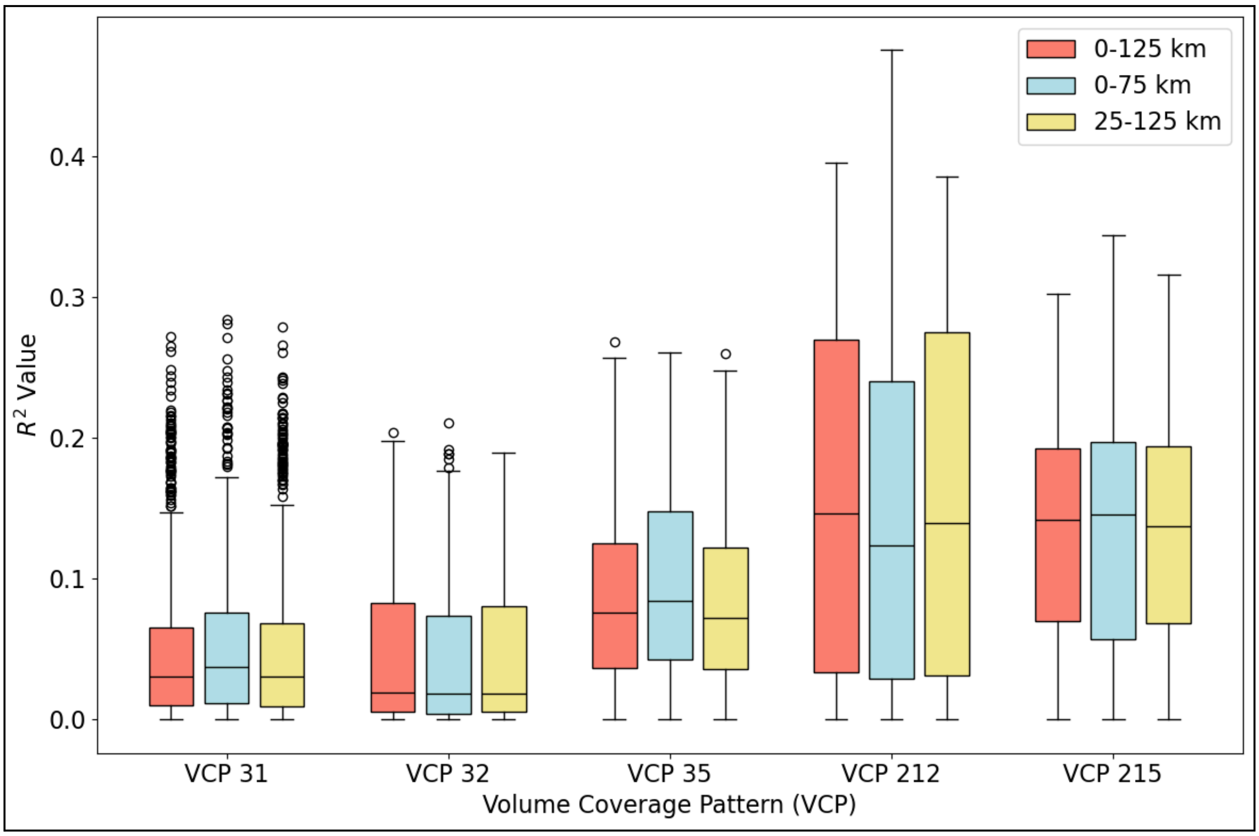

Figure 16. Boxplot showing R^2 score vs. VCP at three distance thresholds.

This breakdown is shown in Figure 17. For the most part, each snow squall type has representation within each VCP category, with the exception of S1 and S4. Still, this indicates that the R^2 plots in Figure 16 are not simply capturing the underlying morphological characteristics of the squall type, but rather the inherent sensitivity with the CAPPI regridding scheme.

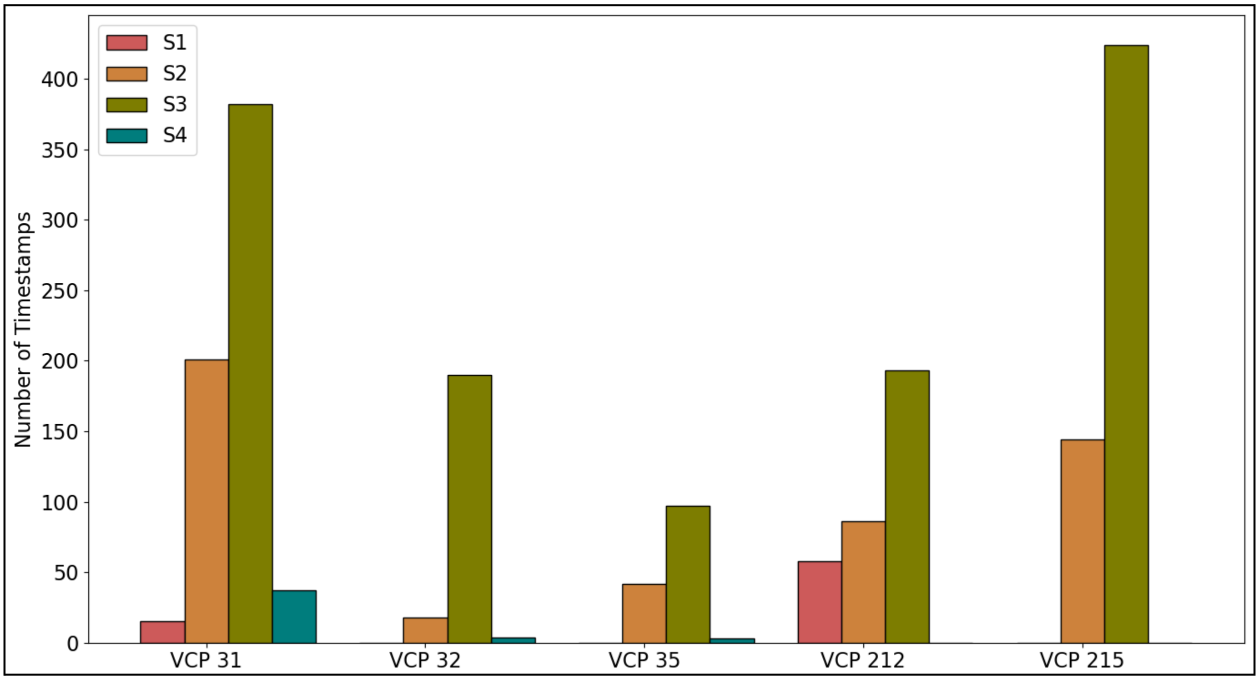

Figure 17. Histogram of VCP, broken down and colored by snow squall type.

3.3.2. Viewing Geometry

Spatial Averaging Through Gaussian Filtering

A potential source of uncertainty within the CLAVR-x dataset results from spatial heterogeneity and artifacting. The DCOMP algorithm utilizes data from multiple channels at multiple resolutions. As a result, the underlying CWP field can become noisy, especially near sunrise and sunset where enhanced contrasts between cloudy pixels and shadows can artificially enhance CWP. Employing a form of spatial averaging, such as a Gaussian filter or "blur", would seek to filter out some of this noise and retrieve signals within the dataset. While other forms of spatial averaging could achieve similar results, Gaussian filtering was found to be the most adaptable for this work and adept for spatial tasks.

Initial Thresholds

Three sigma thresholds were chosen: 0.5, 1, and 2. An example of each field versus the original CWP field is shown in Figure 18. The process for doing this was previously described in Section 2.5.2.

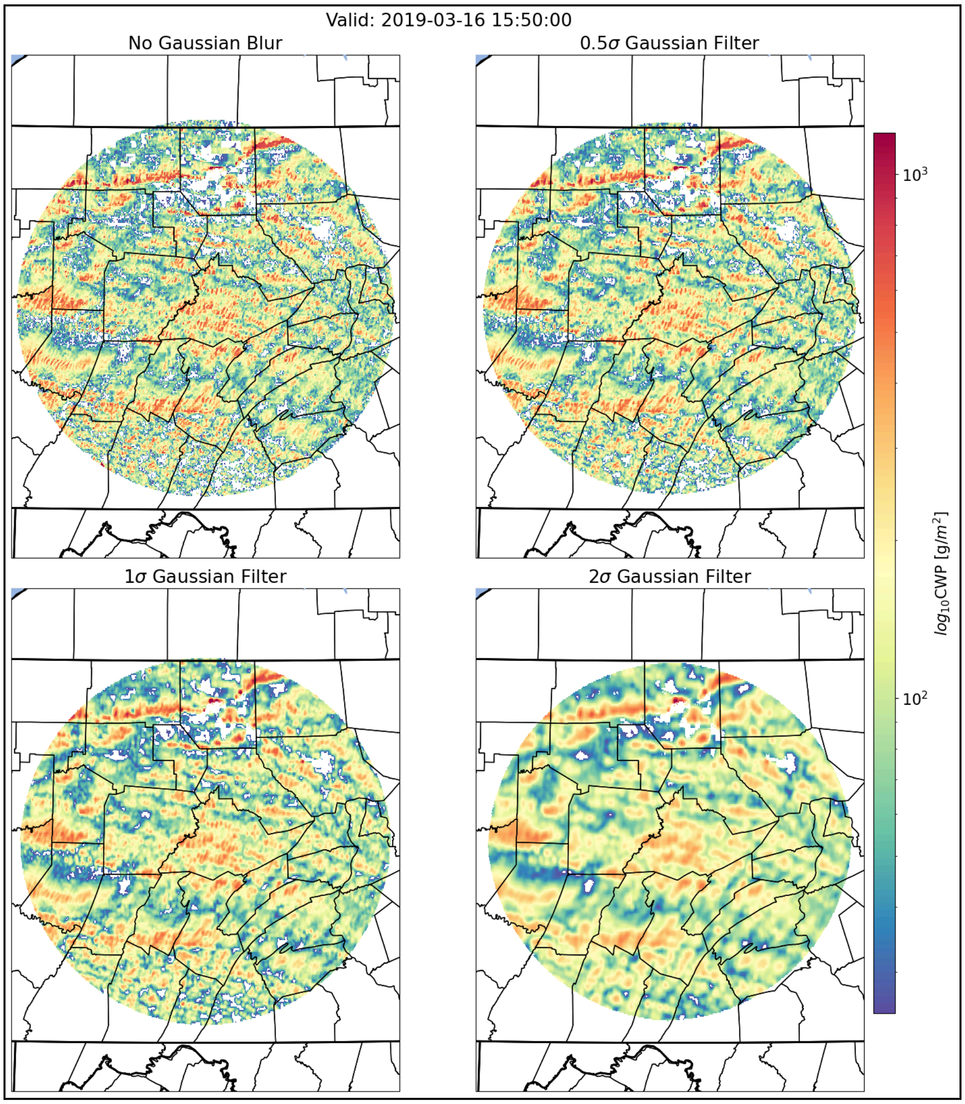

Figure 18. Four panel plot showing CWP (logged) with varying levels of Gaussian filtering applied.

Figure 19 shows how the distribution of CWP values changes with increasing Gaussian filters, choosing to only focus on the 0-125 km distance range here. As one might expect, the distribution smooths out considerably, creating a smoothly exponentially decreasing (which is actually linear given the logarithmic y-axis) curve from high to low CWP.

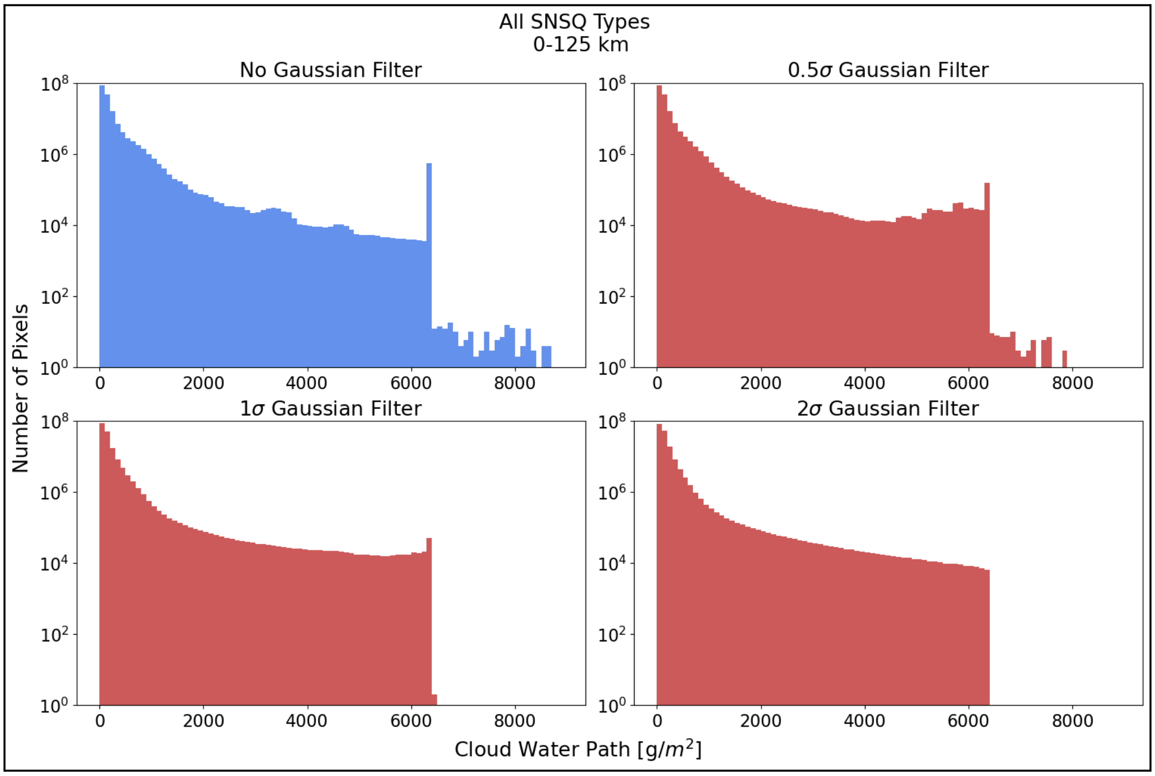

Figure 19. Histograms of CWP across the entire dataset, with progressively higher sigma Gaussian filters applied.

The impact on R^2 scores is clear; Figure 21 shows this result for points both 0-75 km (a) and 0-125 km (b) away from the radar. The interquartile range (including the median) increases quite substantially with increasing levels of Gaussian filter applied. As one might expect, this increase in R^2 score is also reflected in the CWP vs. reflectivity density plots, as shown in Figure 20.

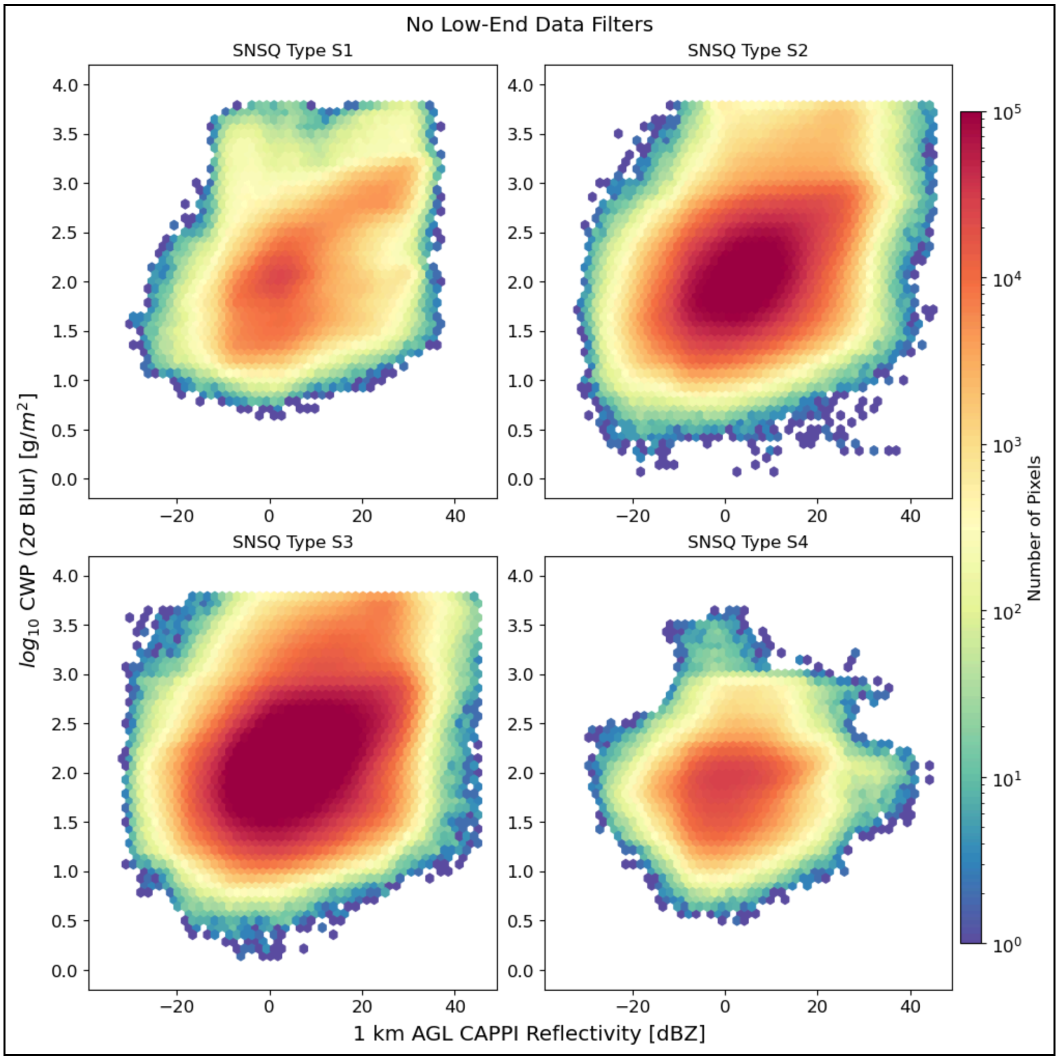

Figure 20. As in Figure 12, but with the Gaussian filtered CWP.

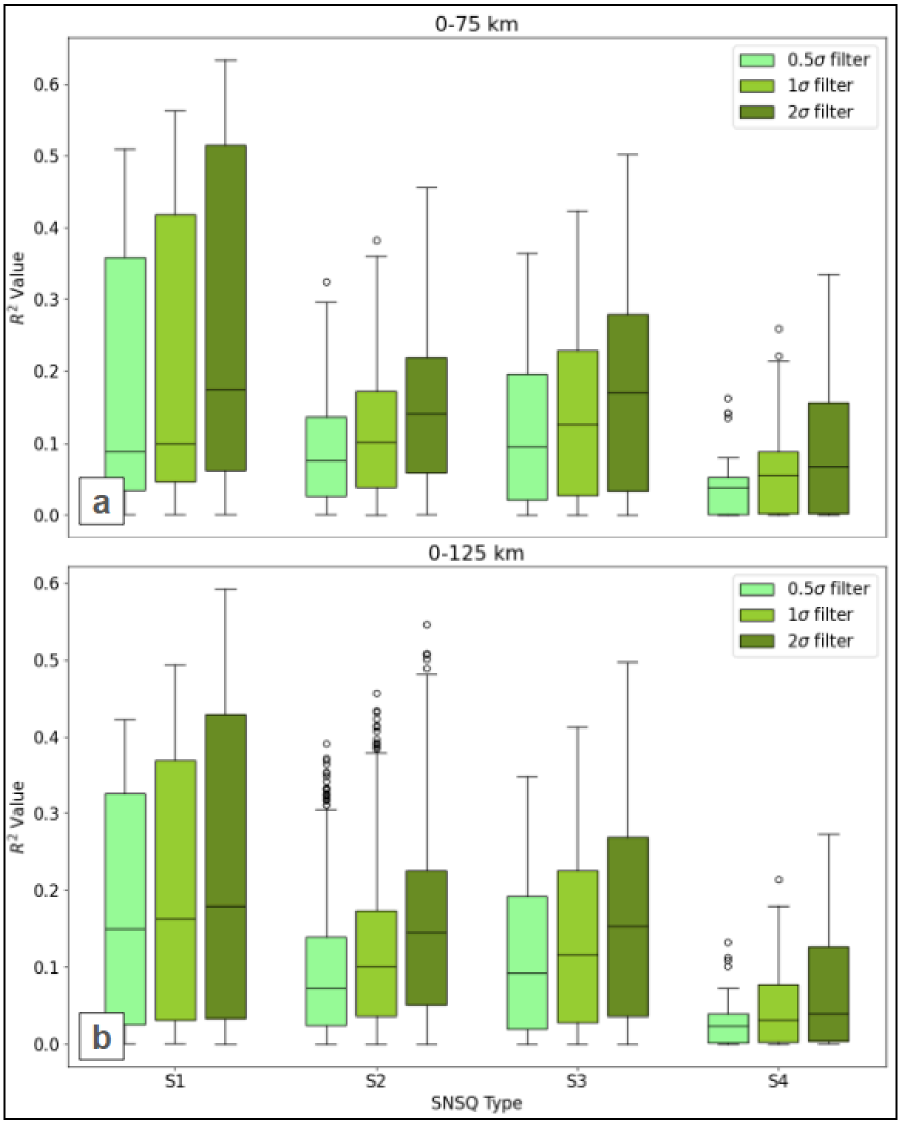

Figure 21. Box-and-whisker plots showing R^2 score variation based on snow squall type for all points a) 0-75 km and b) 0-125 km away from the radar.

Qualitatively, there is less spread in the distributions between CWP and reflectivity, much less spread than what was observed in Figure 12. This improvement is not just limited to the R^2 scores, as Figure 22 also shows notable (~0-5 %) improvement in success ratio scores, though the success ratio gains are difficult to discern.

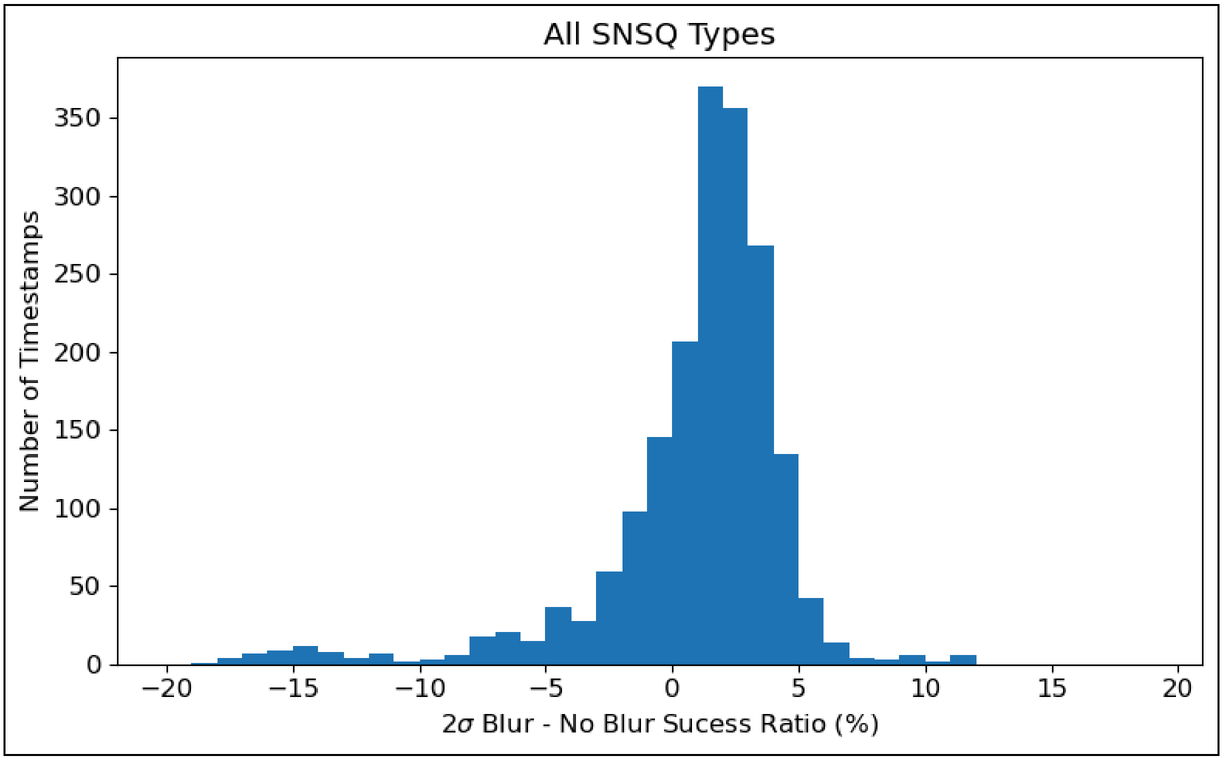

Figure 22. Histogram showing success ratio change for all timestamps.

Expanded Thresholds

The results from the previous sub-section showed marked improvement with increasing Gaussian filter thresholds, so then the question becomes whether these gains are improved with even higher thresholds? It is important to note that increasing the filter higher than 2 sigma yields data which increasingly deviates from the non-filtered field. Such data is questionable; after all, what use was the initial retrieval process in the first place if it is just going to be averaged with no real physical basis? However, this whole line of questioning began precisely because there are large retrieval uncertainties in any single pixel, but the aggregated consensus of a grouping of pixels lessens such concerns. This method of analysis can be likened to reducing random errors through averaging.

Regardless, increasing the Gaussian filtering threshold value does increase R^2 scores, but only to a certain extent. Figure 23 shows this improvement across all snow squall types. Only the

0-125 km distance threshold was tested. S1 cases show the most steady improvement across all sigma thresholds, increasing from 0.1 to nearly 0.3 until steadying about that value. S2 and S3 show similar improvements until leveling out at roughly the same value (about 0.2). S4 cases show little improvement overall, and actually begin showing a decrease in scores after the 3 sigma filtering threshold.

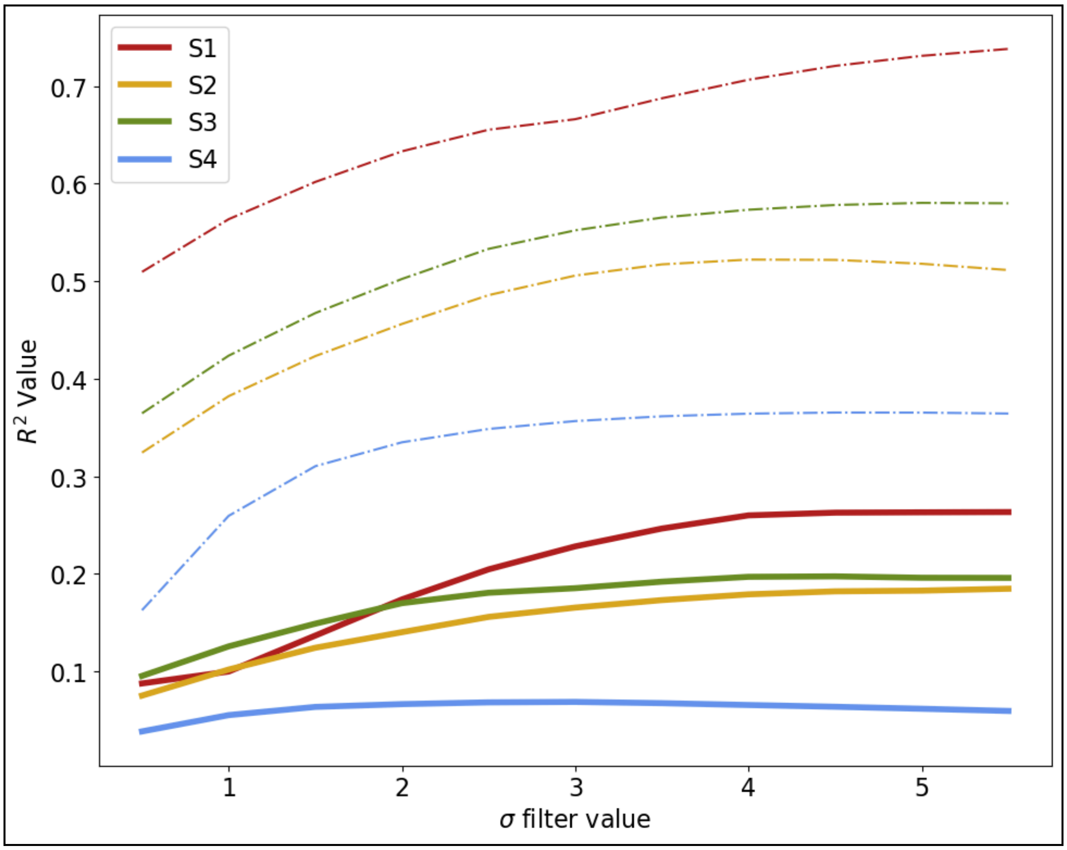

Figure 23. Plot showing how R^2 scores vary with increasing Gaussian filtering threshold values. The solid lines are the median for each snow squall type, and the dotted line is the maximum value.

3.3.3. Satellite Retrieval

While Gaussian filters proved to have at least some utility in reducing errors within the cloud water path field, such analysis is divorced from underlying interrogation on how the CLAVR-x variables both on their own and combined vary within the snow squall dataset; rather, it is just a post-processing technique attempting to correct suspected erroneous spatial heterogeneity in one specific field. Therefore, the goal of this section will be to interrogate how the variables within the CLAVR-x dataset 1) vary by snow squall morphological characteristics (type) and 2) whether or not they help to describe any portion of the variance in the CWP vs. reflectivity relationship. It is expected that the variables retrieved by CLAVR-x give some level of insight into the macrophysical characteristics of the snow squall clouds. Following the logic from the beginning of section 3.2, if indeed the characteristics of the clouds change with increasing reflectivity (or increasing intensity), then that should be reflected in the dataset variables which measure that change, making it possible to filter the dataset based upon those measured quantities. Each variable will be investigated individually, and then jointly if applicable.

Cloud Phase

A key characteristic measured by CLAVR-x is the cloud phase. Given that the clouds being measured are shallow, “cold” (below 0°C) cumuliform clouds, it is expected that the majority of clouds be composed of supercooled liquid water and ice (i.e., a “mixed-phase” cloud). As the cloud matures, it should gradually transition, or glaciate, from mixed-phase to primarily ice as supercooled liquid water is consumed to form frozen precipitation in the form of snow, graupel, etc. Therefore, by filtering for the phase of the cloud, one could potentially filter for snow squalls in different lifecycle stages, which could translate to changes in the efficacy of the CWP vs. reflectivity relationship.

Figure 24 shows the breakdown of cloud phase on a per-feature basis across the entire dataset, for all distances from the radar, using the tobac Python package. The phase mode was found for each feature (the phase type with the most amount of pixels within each feature) and then counted across the entire dataset. Features identified as predominately supercooled liquid water (SLW) exceed those identified as ice by nearly an order of magnitude, and exceed water by three orders of magnitude. Surprisingly, no explicitly “mixed” features were identified.

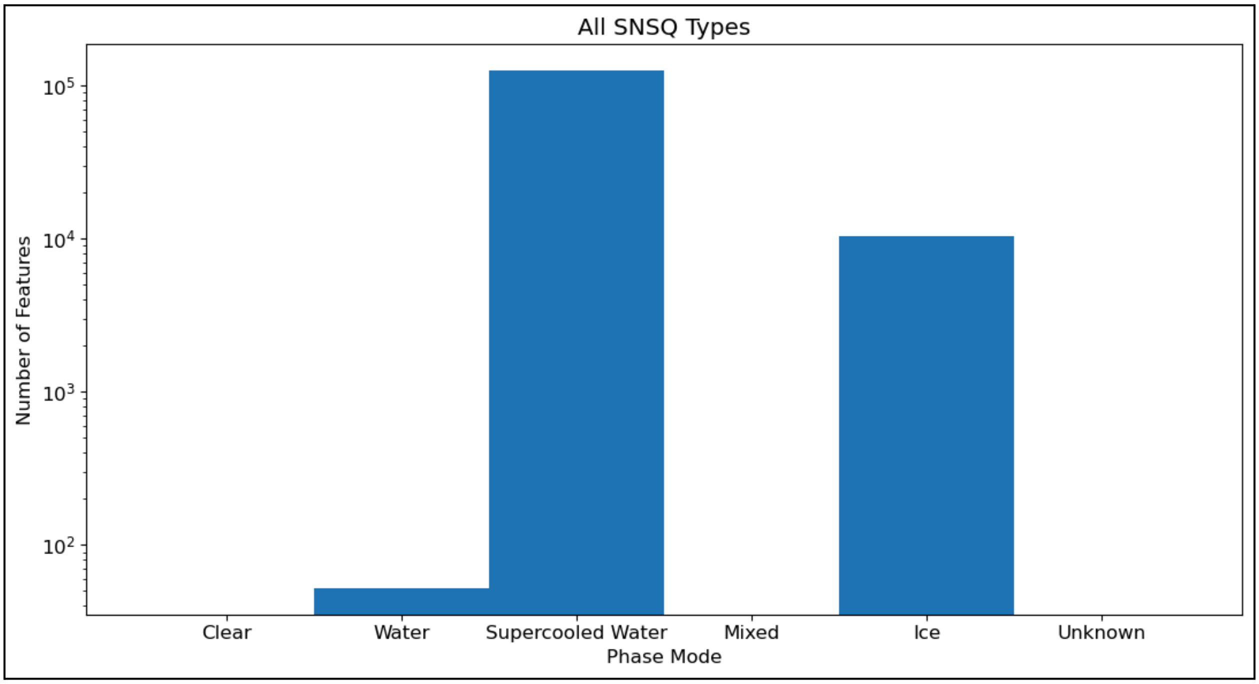

Figure 24. Histogram showing the number of features associated with the most common phase type within the feature.

Binary thresholds of “SLW” or “ice” do not account for the nuanced mix of SLW/ice that often occurs within a given snow squall scene. In order to represent this, a normalization method was applied to each cloud phase scene, shown with the following expression:

$$\frac{n_{p,ice} - n_{p,SLW}}{n_{p,total}} \tag{9}$$

where the number of SLW pixels within a scene is subtracted by the number of ice, and divided by the total number of valid pixels. If the resulting value is 1, then it is an entirely ice dominated scene. If it is -1, then it is an entirely SLW scene. A value of 0 could indicate either 1) that there is an equal proportion of ice and SLW pixels, and/or 2) that the scene is dominated by water pixels. Given the small number of pixels and scenes which have water pixels as shown by Figure 24, the former scenario is much more likely.

The results of this analysis are summarized in Table 5. Results here are similar to previous figures in that most scenes are mostly in the SLW category. However, it also reveals that these SLW scenes also tend to have some proportion of ice within them. In fact, qualitative analysis of numerous case studies reveals that snow squall scenes tend to evolve from SLW to ice frequently.

Category	Sample Size
Ice Only (100%)	13
Mostly Ice (50-99%)	19
Partially Ice (25-50%)	34
Mix (0-25% SLW or Ice)	96
Partially SLW (25-50%)	103
Mostly SLW (50-99%)	1468
All SLW (100%)	160

Table 5. Table showing the breakdown of SLW vs. ice pixels

An example of this is shown in Figure 25, wherein a snow squall scene evolves from being nearly exclusively SLW in a) to a mix of SLW and ice in b). This appears to be evidence for cloud glaciation processes as previously described.

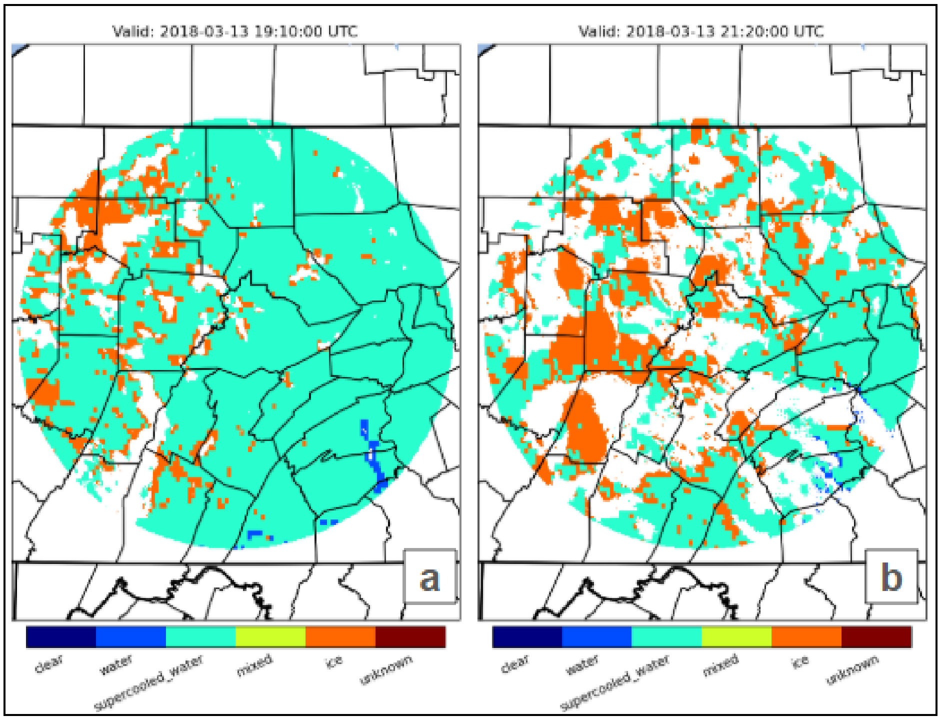

Figure 25. Cloud phase plots from two different times on the same day. (a) is at 19:10 UTC, and (b) is at 21:20 UTC.

Breaking down the previous analysis by type may hint at when these proposed glaciation processes occur most frequently. Figure 26 shows this using the phase normalization method and thresholds defined in the above equation. Type S2 squalls have a higher median number of timestamps with some proportion of ice pixels, followed by S1 and S3. S4 cases are identified as almost entirely SLW.

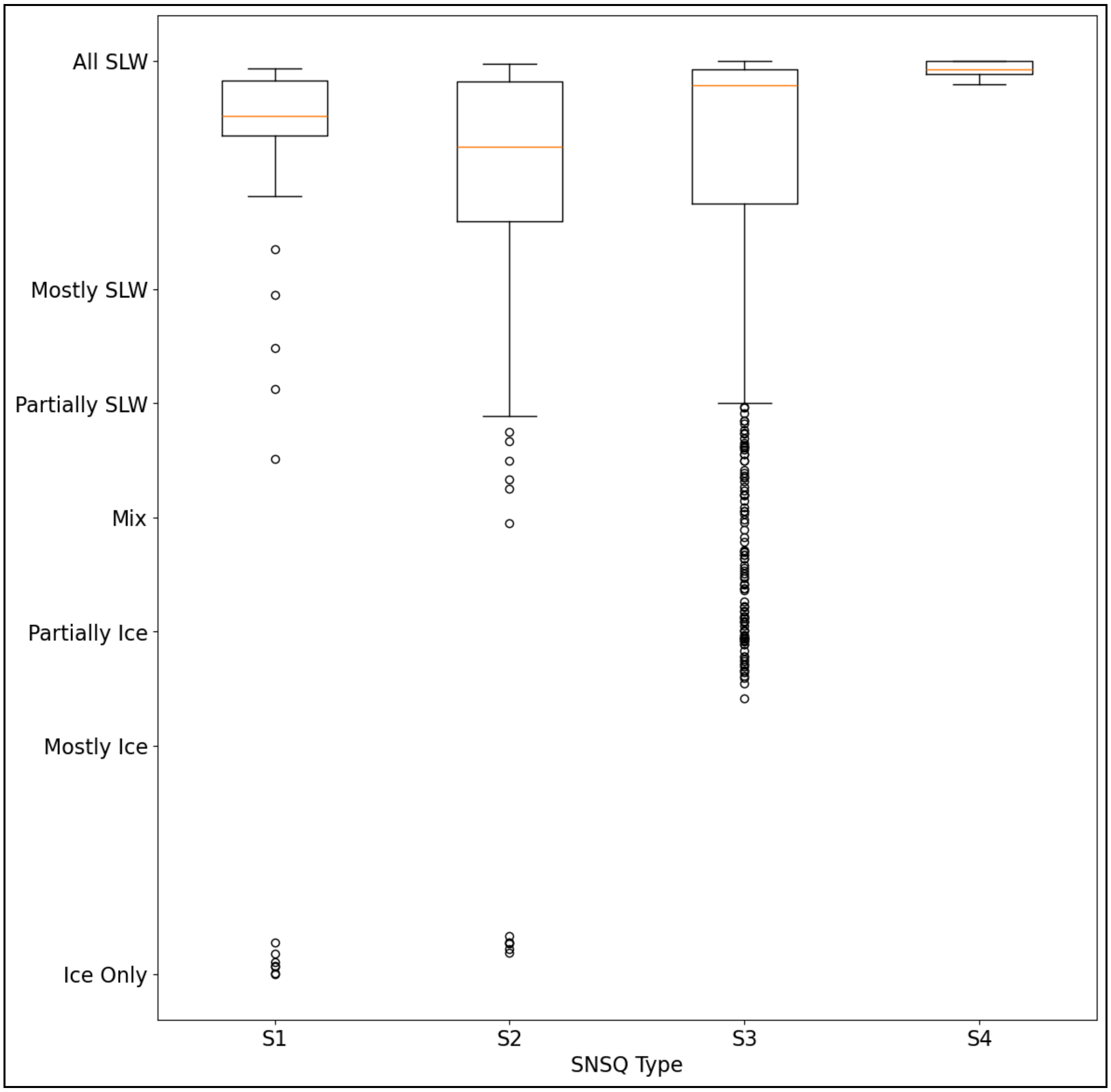

Figure 26. Box-and-whisker plots showing cloud phase proportions per timestamp by snow squall type.

Given that filtering for cloud phase has been corroborated to, at least partially, filter for particular stages of snow squall development, it then becomes of interest to ask whether those filters reflect changes to the CWP vs. reflectivity empirical relationship as hypothesized. In other words, do scenes with more/less ice tend to have higher/lower R^2 scores? Figure 27 shows the dataset R^2 score distributions, filtered into the cloud phase bins used above. A notable trend appears; scenes which are composed mostly or partially of ice pixels gave much higher median and interquartile R^2 scores than those with mostly or wholly supercooled liquid water. While intriguing, this result is hindered somewhat by the small sample size of the mostly and partially ice bins (53 timestamps) compared to the supercooled liquid water bins (1731).

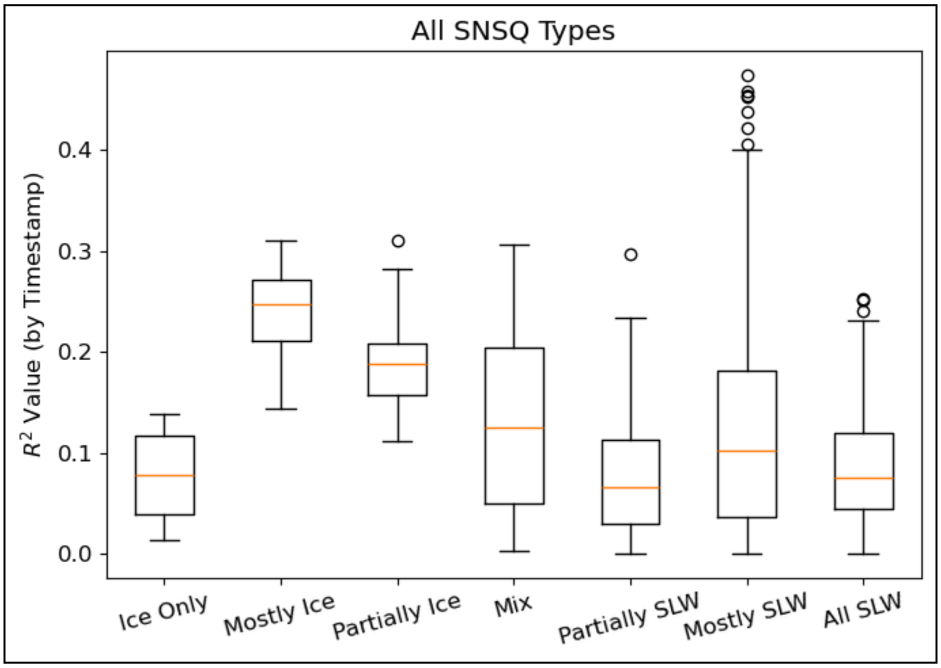

Figure 27. Box and whisker plot showing the R^2 values as a function of scene cloud phase proportion.

Despite the small sample size of the mostly and partially ice scenes, the score increase is notable, and may perhaps represent an increase in the skill of CLAVR-x to detect precipitating clouds when the glaciation process has been allowed to have occurred in earnest.

Cloud Top Height

Cloud top height is perhaps a more explicit variable to try and link with snow squall intensity. It is expected that, generally, the higher/deeper a cloud is within the atmosphere, the stronger updraft it must have taken to form it, and therefore the more intense/strong the resulting snow squall. It would also make sense if the highest cloud top heights occurred with clouds which are glaciating, since this would happen at the snow squall's mature stage, and therefore the point at which it has the highest cloud top heights. The important exception to this inference would be overlapping/cirrus clouds which obfuscate measurements of the shallow convection beneath.

CLAVR-x Cloud Height Overestimation

Before full analysis can be done, some preliminary analysis on CLAVR-x cloud top height output is needed. Figure 28 shows cloud top heights for the entire dataset, by pixel, partitioned based on snow squall type. Despite median cloud phase proportions being relatively different between snow squall types, cloud top height medians generally fall within ~3km, with type S2 and S3 having slightly higher medians than S1 or S4. However, every type has outliers significantly higher than the interquartile range, exceeding 6 km and nearly reaching 12 km for S3 squalls.

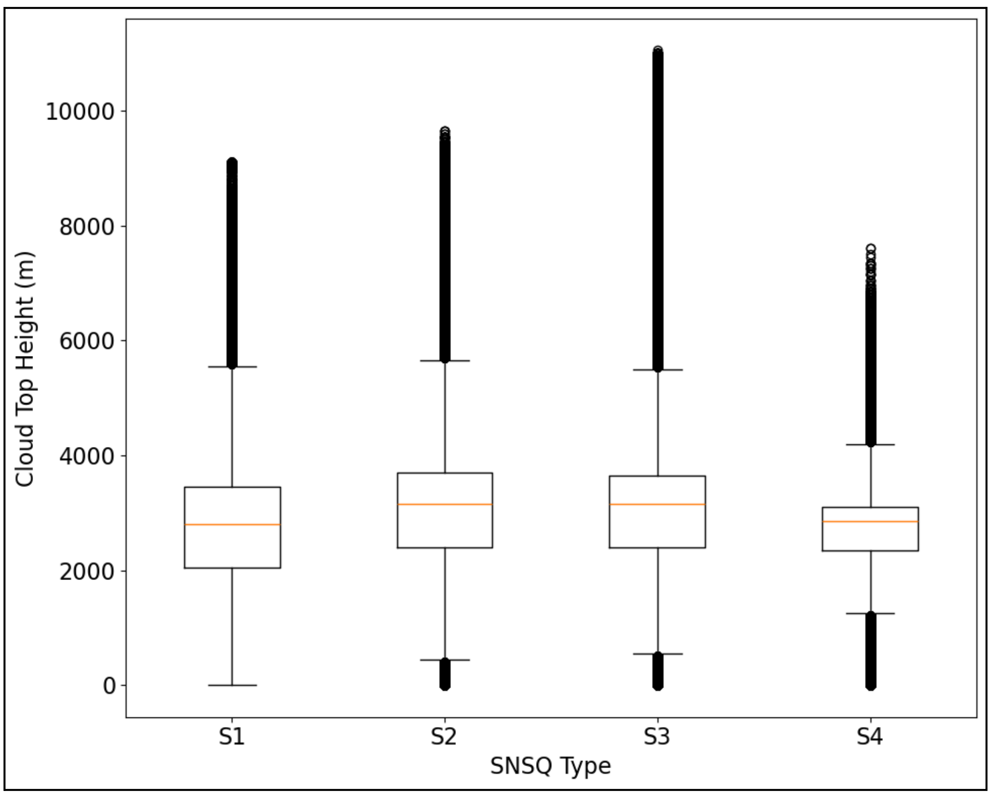

Figure 28. Box-and-whisker plot showing cloud top height per pixel based on snow squall type.

An immediate concern emerges from this result; cloud heights above 4-6 km are observed with some frequency, rather than 1-3 km cloud top heights. Previous studies, both specifically on overland CS and over-water CS, would allude to this being unphysical and generally unrealistic, with typical cloud top heights reaching ~2-3 km. This can be further corroborated for these specific cases by taking a cross-section through radar data, and comparing the vertical extent of the reflectivity core with the estimated CLAVR-x cloud top height along the azimuth. It can generally be assumed that the precipitating core and enveloping cloud structure encompass roughly the same region; this is supported by previous literature which found that the highest reflectivities in CS are often closest to the ground (Pettegrew et al. 2009).

Figure 29 combines both the reflectivity, cloud top height, and reflectivity cross-section into one plot. The snow squall scene chosen is from March 28th 2022, a date which produced several snow squalls—including one which caused a major accident mentioned in Section 1.1. (Medina 2022). Therefore, this scene should be one with squalls intense enough to cause impacts. The cloud top height product in b) shows cloud top heights exceeding 3-4 km, with maximum values around 6 km within the core of the snow squalls. Figure 29c directly contradicts b), showing the maximum vertical extent of the reflectivity cores extending only to ~2-3 km in depth. Although it is true that there is some portion of non-precipitating cloud above the precipitating portion observed by ground-based radar, this would need to account for 2-3 km of height. Both the literature and physical knowledge of these phenomena would deem this to generally be an unrealistic depth, as the capping inversion present within shallow convective regimes generally sets a hard limit on the vertical extent of all cloud particles. Still, this is an inherent uncertainty associated with estimating cloud top heights using imprecise methods (i.e., active ground-based radar and passive satellite imagery). A more effective analysis would involve examining these clouds with spaceborne active sensors (Cloudsat, CALIPSO, EarthCARE, etc.) which have been shown to more accurately characterize cloud top heights than other retrieval methods, though such analysis is outside the scope of this study.

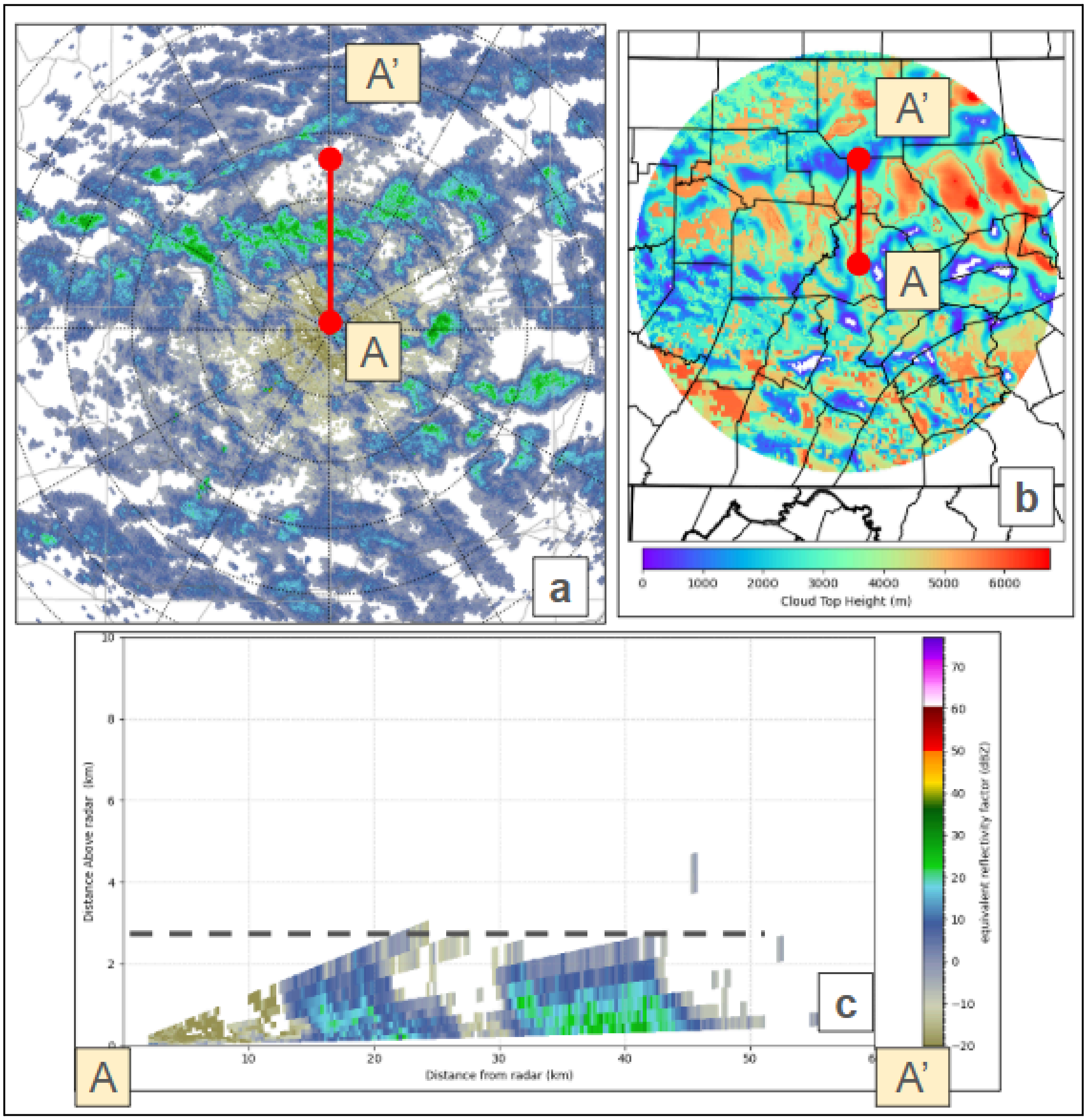

Figure 29. Plot showing a) radar reflectivity PPI at the 0.5 degree elevation angle, b) CLAVR-x cloud top height, and c) a psuedo-range height indicator plot from several PPIs, with an estimated cloud top height drawn with a dotted line. All images are valid at 17:11 UTC on March 28th, 2022.

Cloud Top Height vs. Phase

Despite the hypothesized proclivity of CLAVR-x to overestimate snow squall cloud top heights, it is still instructive to examine how this variable varied throughout the dataset, and if (as expected) higher cloud top heights result in more clouds identified as ice. Figure 30 shows the distribution of cloud top heights within the dataset, if one partitioned the data between all pixels, and only those which were deemed to be ice. The entire dataset has the majority of its distribution centered around ~3 km, while the ice-filtered distribution is much higher, ~6 km. This may reflect the development pattern described in the previous section, wherein cloud glaciation changes the radiative properties such that the DCOMP algorithm increases the cloud height.

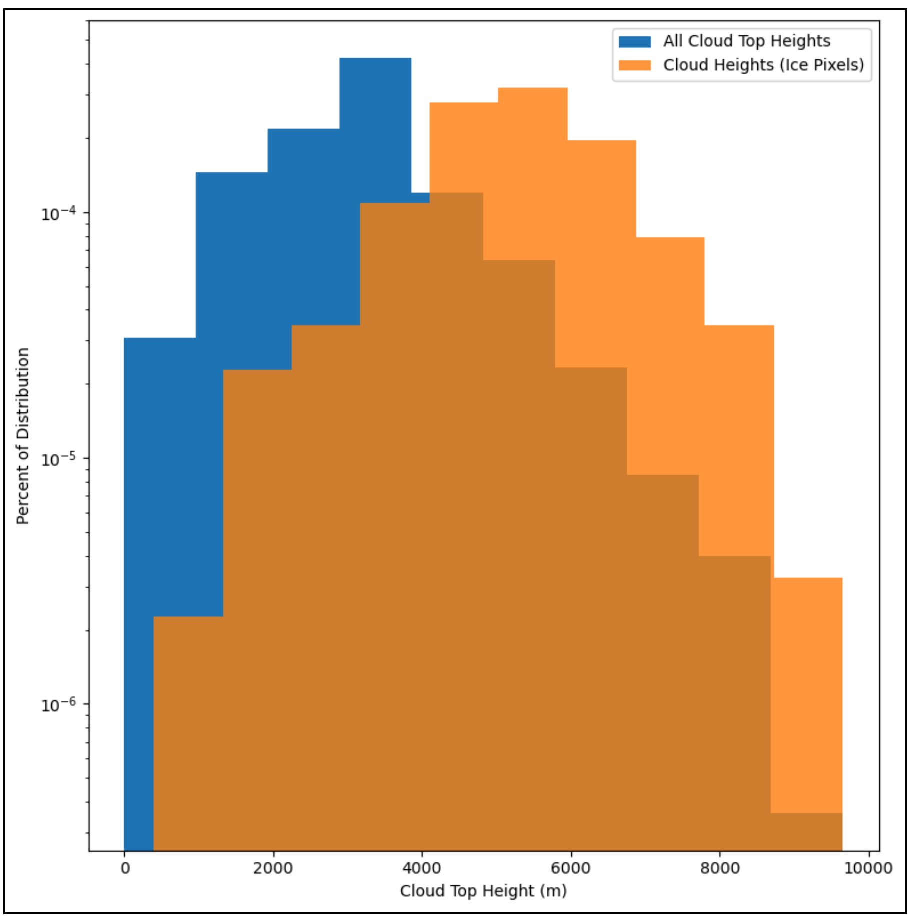

Figure 30. Histogram showing the distribution of cloud top heights.

Filtering by snow squall type yields similar results, perhaps even more dramatic than in Figure 30. Figure 31 shows that, when filtering for ice versus no ice pixels, ice pixels have much, much higher median cloud top heights than no-ice pixels. For type S1 through S3, this difference is ~2.5 km, while for S4 it is about 2 km.

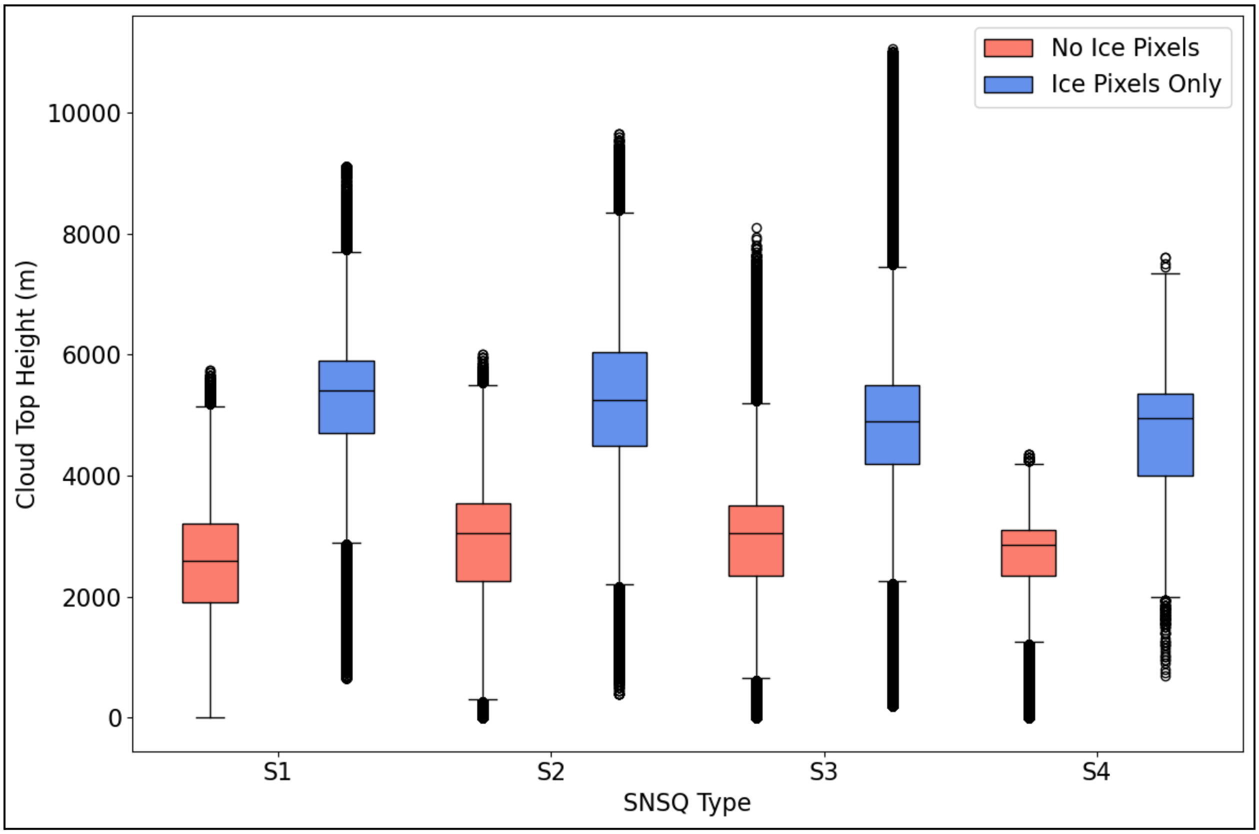

Figure 31. Box-and-whiskers plot showing cloud top height as a function of snow squall type, filtered by ice content.

This pattern is shown further in Figure 32, wherein cloud features whose most common phase classification is ice are associated with higher mean cloud top heights than those identified as supercooled liquid water, or just liquid water. This finding implies that the algorithm identities ice via glaciating processes, and subsequently “raises” the height of that cloud.

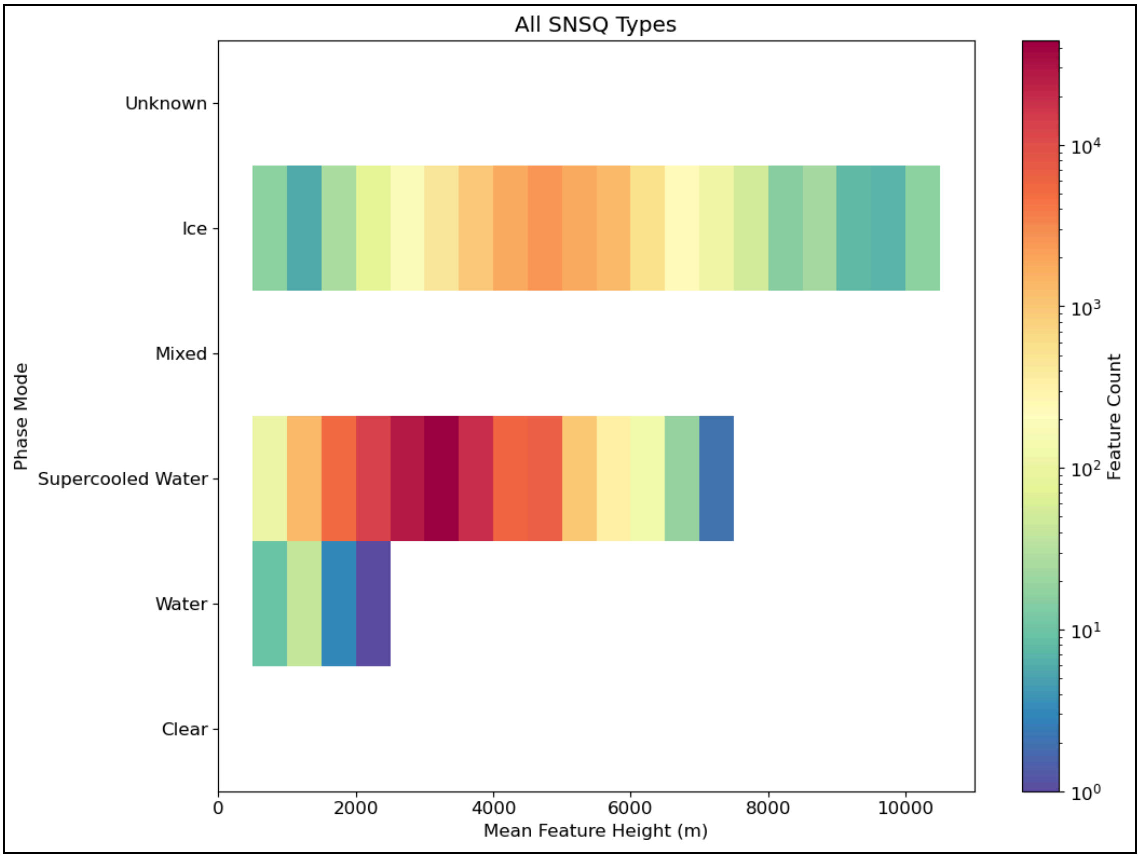

Figure 32. 2D Histogram showing the distribution of points based on mean tobac feature height and phase mode.

Cloud Top Height vs. CWP and Phase

Given that CLAVR-x can generally recognize pixels whose cloud phases are ice to be associated with higher clouds, the question becomes whether this has any relationship to CWP (and by proxy snow squall intensity). Figure 33 shows how CWP varies with cloud top height for each snow squall type. Generally, the cloud top height increases as the CWP increases, but there is large variation in CWP for any given cloud top height. As shown in Figure 28, S3 squalls have the highest cloud top heights of the entire dataset, exceeding 10 km and CWP values of ~10^2-10^3 g/m^2. S4 has the lowest heights, with the majority of points well within 5 km and spanning a wide range of CWP (10-10^3 g/m^2).

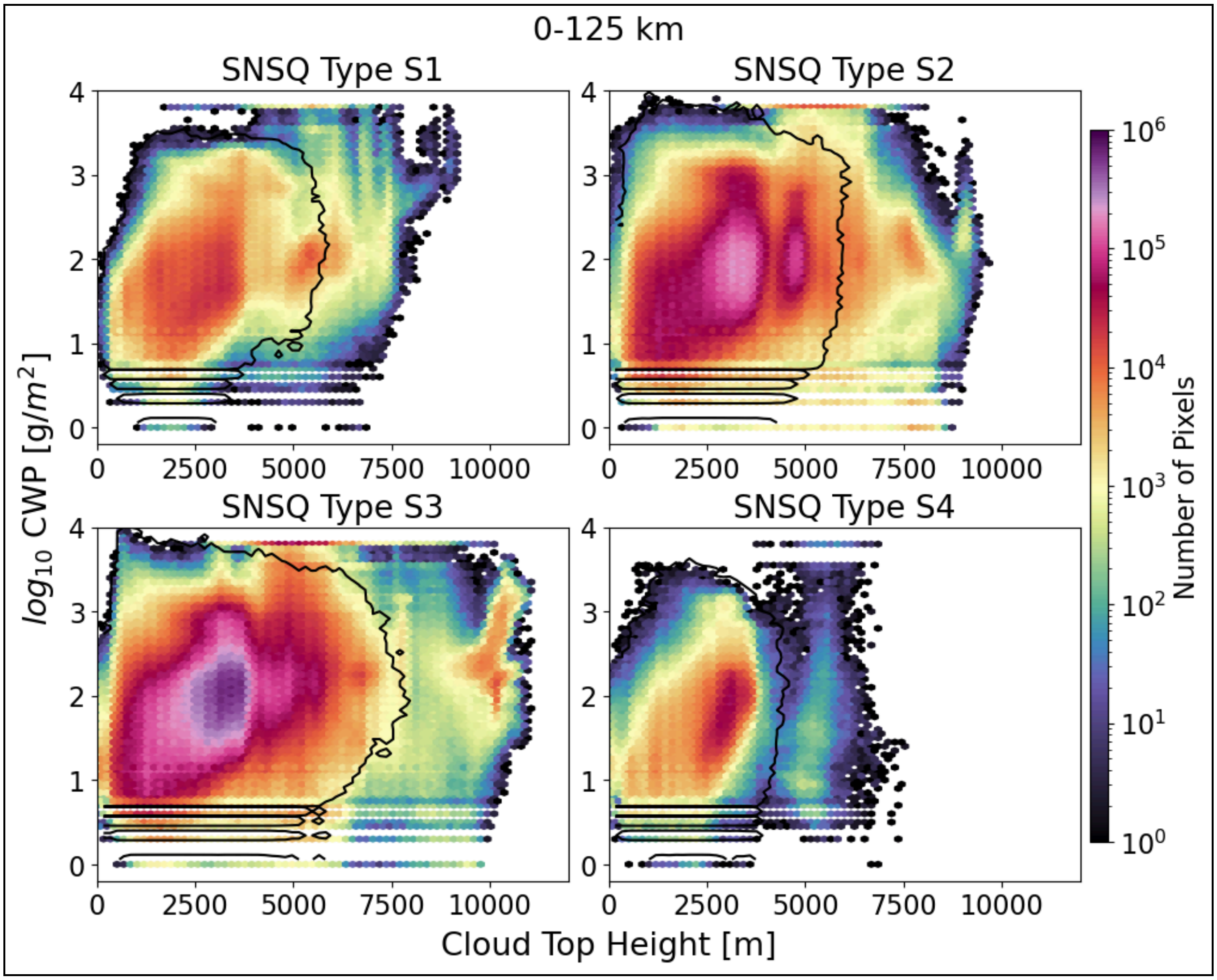

Figure 33. CWP versus Cloud Top Height by snow squall type. Black contours show all pixels which were identified as SLW.

Figure 33 also is contoured for all points which are identified to be SLW by the cloud phase. As expected from Figure 31, this isolates the lower end of cloud top height values, generally less than 5 km. The exception being S3 squalls, which regularly has SLW clouds with heights approaching 7.5 km. Filtering for just ice clouds does not have the same effect as with SLW (see Figure 34), indicating that while SLW clouds are relegated to only the lowest ~5 km, ice clouds can occupy nearly the entire spectrum of cloud top heights.

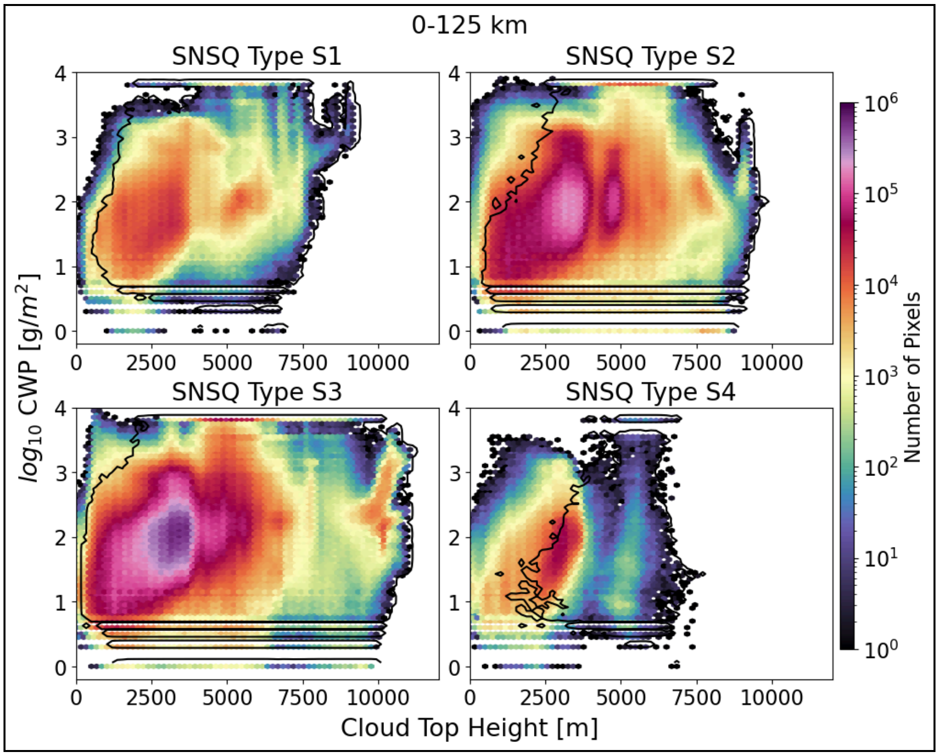

Figure 34. As in Figure 33, but contoured everywhere that ice clouds are identified.

Cloud Effective Radius

Overview

As described in Section 2.3.2, the effective radius is defined as the ratio between the third and second moment of the PSD (equation 5). Given that the effective radius is explicitly defined by the PSD (parameterized by equation 4), it will be sensitive to which PSD is used, and this is in part determined by the retrieval method the algorithm uses (DCOMP mode 1, 2, or, 3, which utilize different GOES-16 channels). Regardless, effective radius should serve as another proxy for snow squall development (similarly to cloud phase). From analogous lake-effect studies, particularly Barthold and Kristovich (2011), a developing snow squall would have a large population of small particles, which gradually decreases as aggregation and depositional processes create larger particles. Therefore, the integrated PSD will yield a larger value, and by extension a larger effective radius. Additionally, the precipitating core of the snow squalls should have larger effective radii from larger particles, and should decrease towards the edges of the cells. Therefore, filtering for effective radius should filter for pixels which are most likely to be precipitating.

The distribution of cloud effective radius values by snow squall type is plotted in Figure 35. For DCOMP mode 1, the effective radius distribution is much more broad than modes 2, and 3, with values exceeding 80 μm in the 95th percentile for S2 squalls, while modes 2 and 3 only reach ~30-40 μm. For all modes, S4 squalls have the smallest particles, with S1 and S3 generally having similarly-sized radii. S2 stands out as having the largest particles of any mode, though its interquartile range is generally consistent with S1 and S3 squalls. This makes sense; Schneider et al (2024) found that type S2 squalls have the highest values of CAPE on average, which means stronger updrafts, and therefore larger particles.

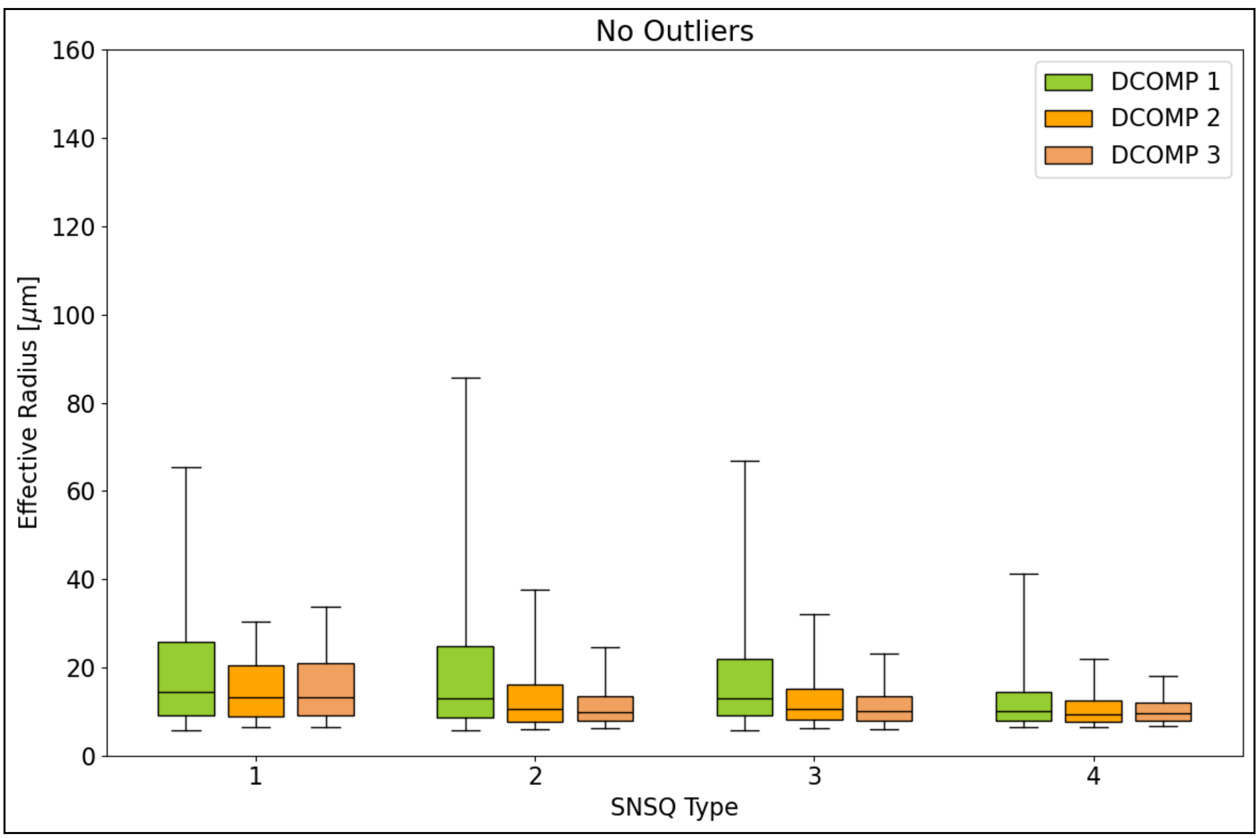

Figure 35. Boxplot of effective radius versus snow squall type for each DCOMP mode. Outliers past the 5th and 95th percentile whiskers are omitted.

Figure 36 shows what each of these effective radius products looks like in the data. It is immediately clear that DCOMP1 has the most artifacting (i.e., pixels which have unphysically high or low values) issues, with several patches of high (~100 μm) values. DCOMP2 has less artifacting, but appears to highlight the edges of some cloud features rather than the center. DCOMP3 is like DCOMP1, but with the artifacting largely removed.

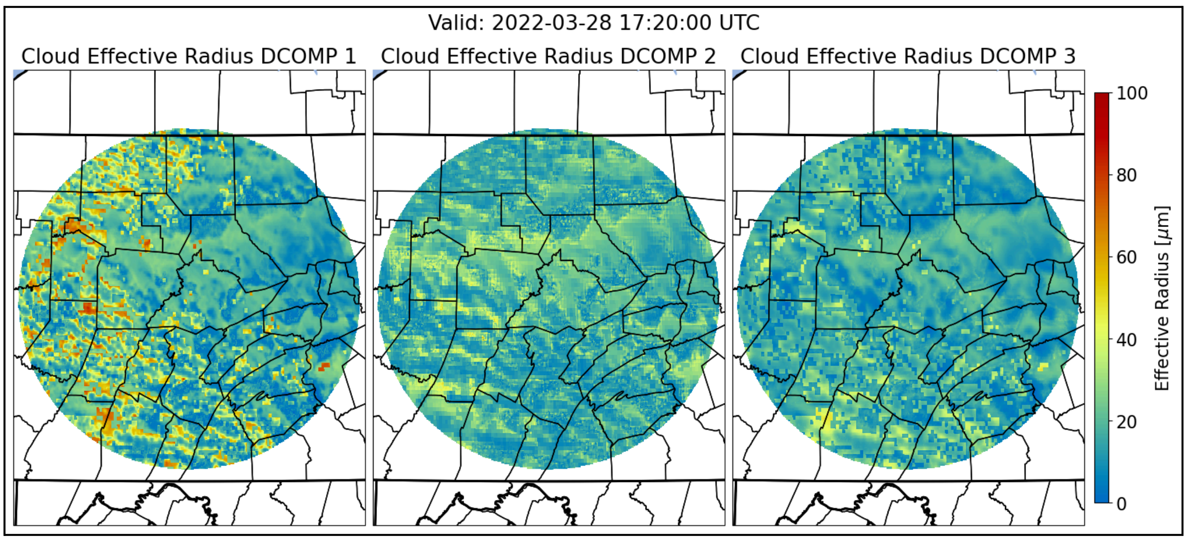

Figure 36. Cloud Effective Radius plots for each DCOMP mode centered over central Pennsylvania.

To get an idea of the range of the effective radius dataset, Figure 37 shows the same products as Figure 36, but for a different time. Here, the effective radius values are much higher, despite the change in colorbar. Values exceeding 80 μm are ubiquitous in DCOMP1, while DCOMP2 and 3 maintain much lower values. Notably, the artifacting issue present in Figure 36 is not present at all in this scene, though the pixelation is still present in DCOMP2.

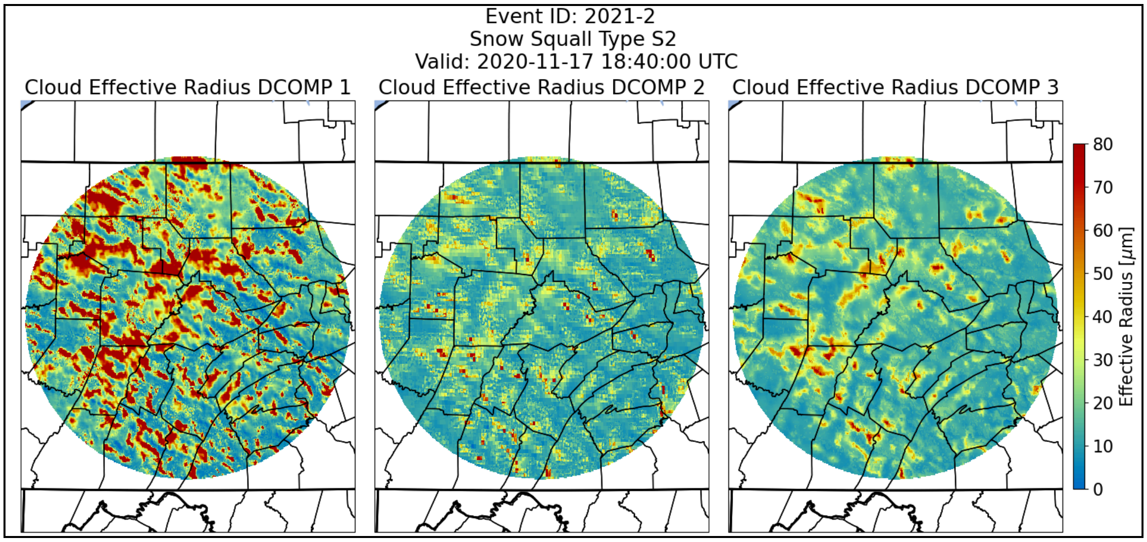

Figure 37. As in Figure 36, but for a different time.

Filtering by Effective Radius: Snowfall Intensity

Figure 38 shows how the effective radius varies with CWP depending on DCOMP mode. While the distribution for each mode is generally fixed about the same central point, the distribution of values beyond that varies with no discernable pattern (i.e., CWP generally does not vary in an organized manner with effective radius).

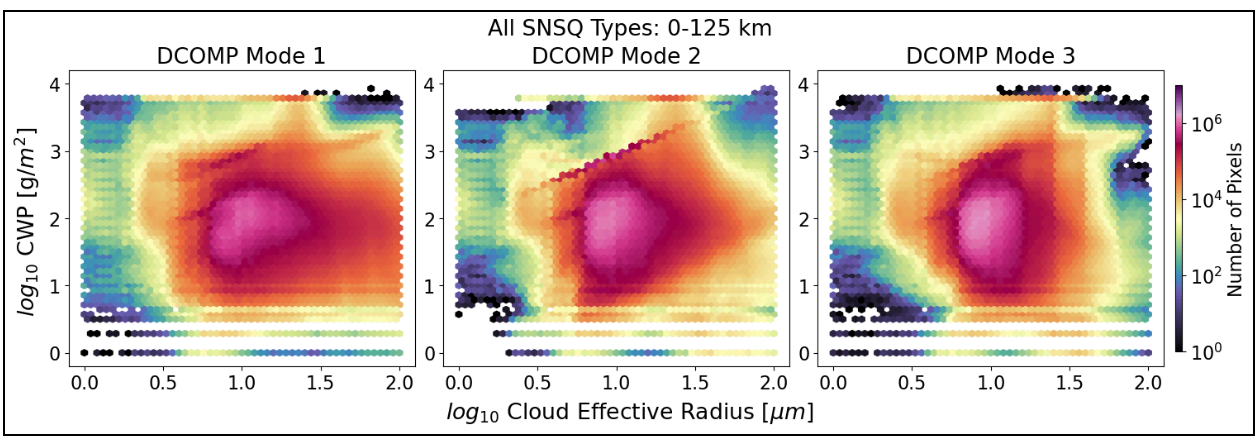

Figure 38. 2D histograms as in figure blue, but showing CWP against cloud effective radius for each DCOMP mode.

Of particular interest regarding the plots in Figure 38 was whether there existed a range of CWP and effective radius values which could delineate low, medium, and high reflectivity (and by extension snowfall intensity) values. Figure 39 and Figure 40 show this attempt for DCOMP modes 1 and 2, and is broken down by snow squall type. With the reflectivity threshold chosen (15 dBZ), no meaningful set of CWP/effective radius values stand out, with the slight exception of type S1 and S4. For these types, a set of points clustered around the most dense portion of their respective distributions is highlighted, and this holds true for both DCOMP modes 1 and 2. However, this result seems to not be revealing any particular physical process (i.e., some optical depth and radius property) but rather just highlighting the regions where most of the points reside. If such a result were to manifest, it would be expected that the region highlighted be displaced in some manner to the region where most points reside (since moderate/heavy precipitation should constitute only a small portion of the overall data).

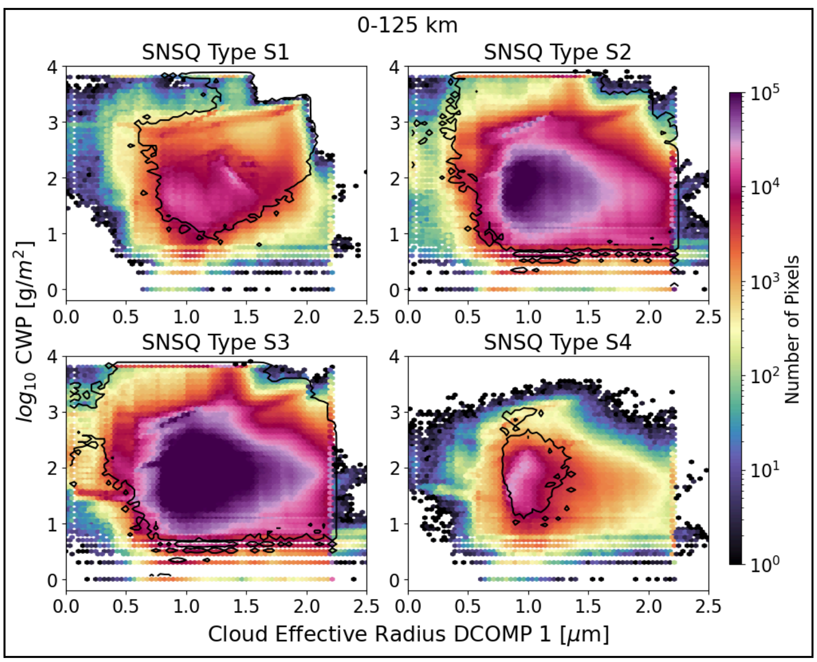

Figure 39. As in Figure 38, but broken down by snow squall type and contoured at points which reflectivity > 25 dBZ.

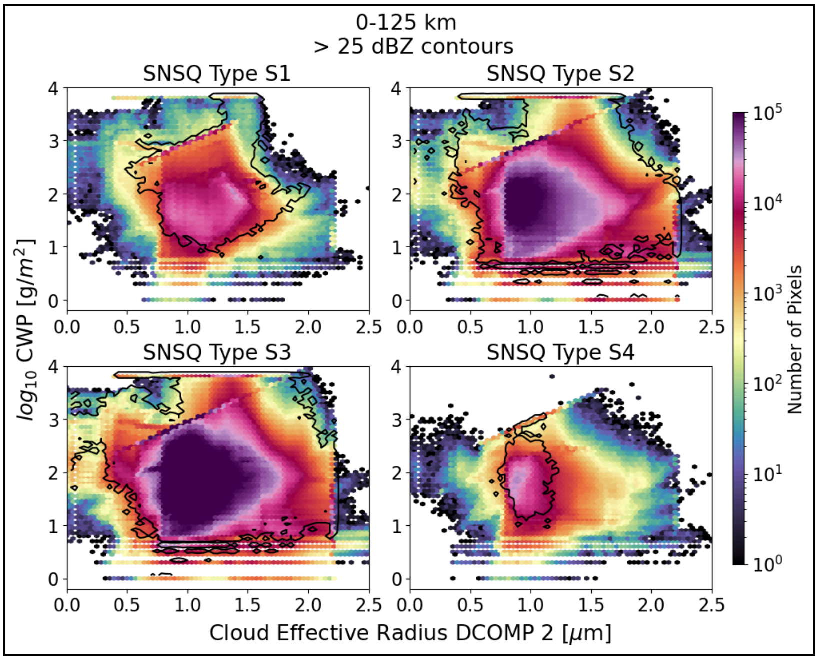

Figure 40. As in Figure 39, but for DCOMP mode 2.

Filtering by Effective Radius: CWP-Reflectivity

Although CWP-effective radius plots by themselves did not reveal any skill in identifying more intense snow squalls, the story changes when looking back at CWP-reflectivity plots from section 3.2.1. Since the satellite is able to see all cloudy pixels, it cannot natively discriminate which pixels are more or less likely to be precipitating, even with CWP greater than certain baseline threshold values (10, 50 g m^{-2}). So while the CWP-reflectivity density plots did show a relationship, it was diffused by this ambiguity. Filtering by effective radius helps to eliminate small radii particles on the edges where uncertainty is greatest, and focus in on the core of cells where CLAVR-x is most likely to correctly identify the CWP.

Figure 41 shows the same CWP vs. reflectivity plots as in section 3.2.1, but now contoured at points where DCOMP1 effective radius is greater than 50 μm. What immediately stands out is this filtering metric's ability to remove outlying areas of the distribution and come to a result which, qualitatively, should yield a much better R^2 score.

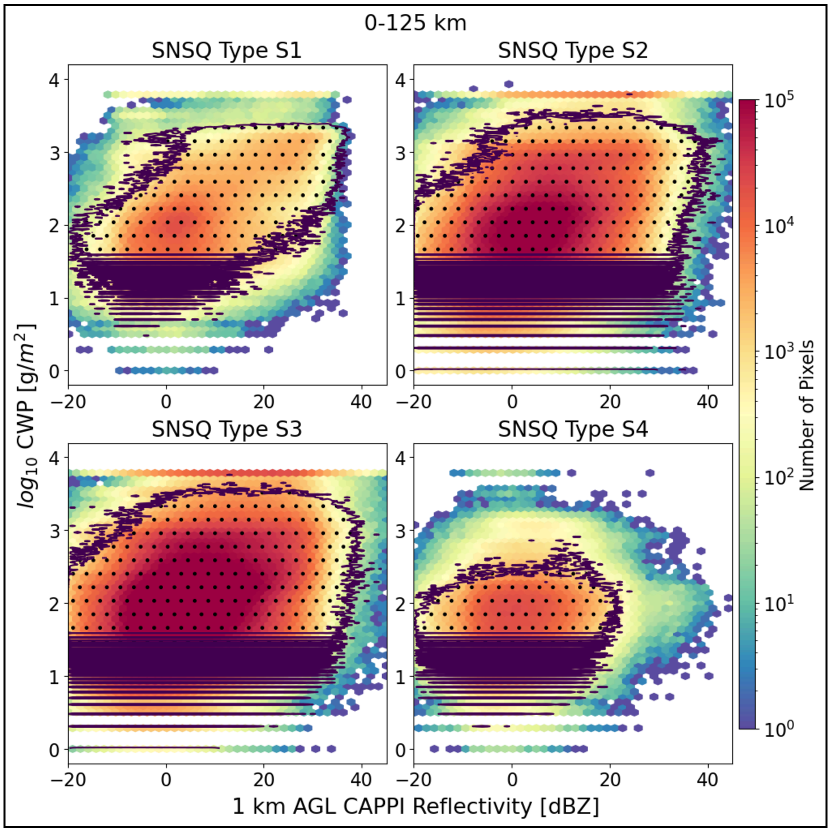

Figure 41. 4-panel plot as in figure blue, except points which had an effective radius DCOMP1 > 50 μm is contoured.

Quantitatively, Figure 42 shows the improvement this filtering has on the R^2 scores. Gains seem to be maximized at 50 μm, where S1 shows a near 0.2 median R^2 score improvement, with S2 and S3 showing a much more modest ~0.1 point improvement, though the inter-quartile range increases substantially. S4 shows little to no improvement at any threshold, with only the upper quartile ranges improving marginally. Perhaps unsurprisingly, this improvement gets even better when the underlying CWP field has a Gaussian filter applied, shown in Figure 43. This figure showcases by far the highest R^2 scores seen so far, with the inter-quartile range of type S1 snow squalls surpassing 0.6. One major caveat with these results is the sample size of each timestamp for these R^2 scores. As progressively higher effective radius filters are applied, less and less of a scene becomes usable for R^2 calculations. This may lead to biasing the data towards a small subset of favorable pixels while ignoring the other unfavorable ones. Another interesting result from Figure 43 is that the particle size which maximizes gains (i.e., the size which maximizes how much of the core of the squall gets represented, and therefore improve R^2 scores) changes with squall type; for example, S1 tends to have median scores stop increasing after 50 μm, but S2 sees scores stop increasing between 25 and 50 μm. It is possible that the exact effective radii threshold varies from event to event, or even scene to scene, depending on the exact values which correspond to the edge of the snow squalls (and therefore of the largest radii particles).

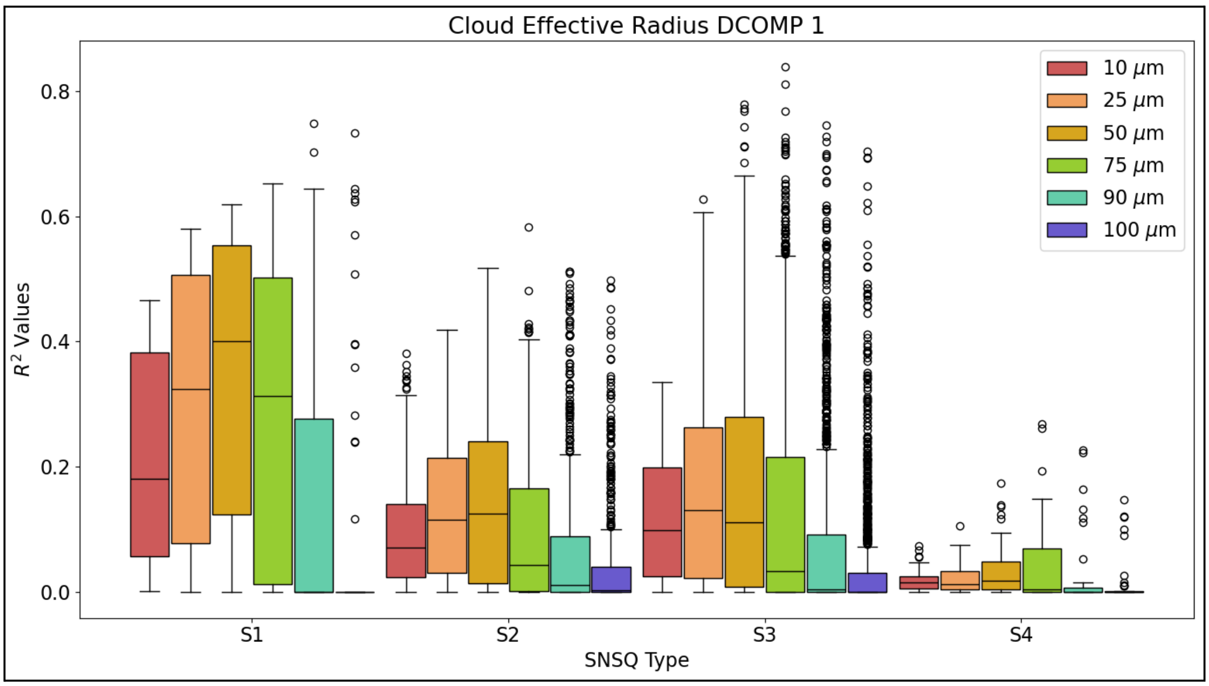

Figure 42. Box-and-whiskers plot showing R^2 values for each snow squall type, broken down by DCOMP1 effective radius filtering thresholds.

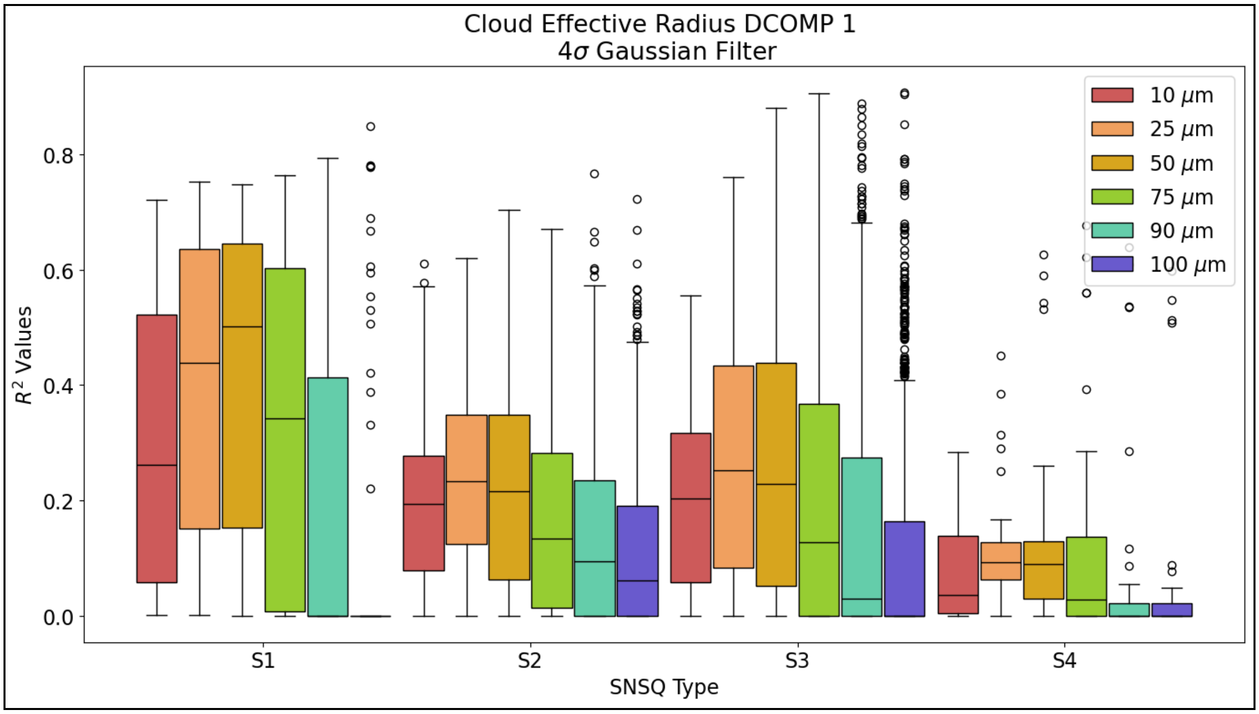

Figure 43. As in Figure 42, but with a gaussian filter applied to the underlying CWP data

Despite the dwindling sample sizes, there still are thousands of points in even the best scenes to work with; Figure 44 showcases this by picking out random scenes from type S1, S2, and S3 timestamps which had an R^2 score > 0.7, for those pixels whose effective radius DCOMP1 was greater than 50 μm. S4 did not have any scenes which met this criteria and therefore was not included for consideration. S1 and S3 have the highest number of sample pixels at over 2000, while S2 only has ~400 pixels. Despite the relative dearth of points for type S2, the density plots in Figure 44 are by far some of the best obtained in this study, which is reflected in their R^2 score.

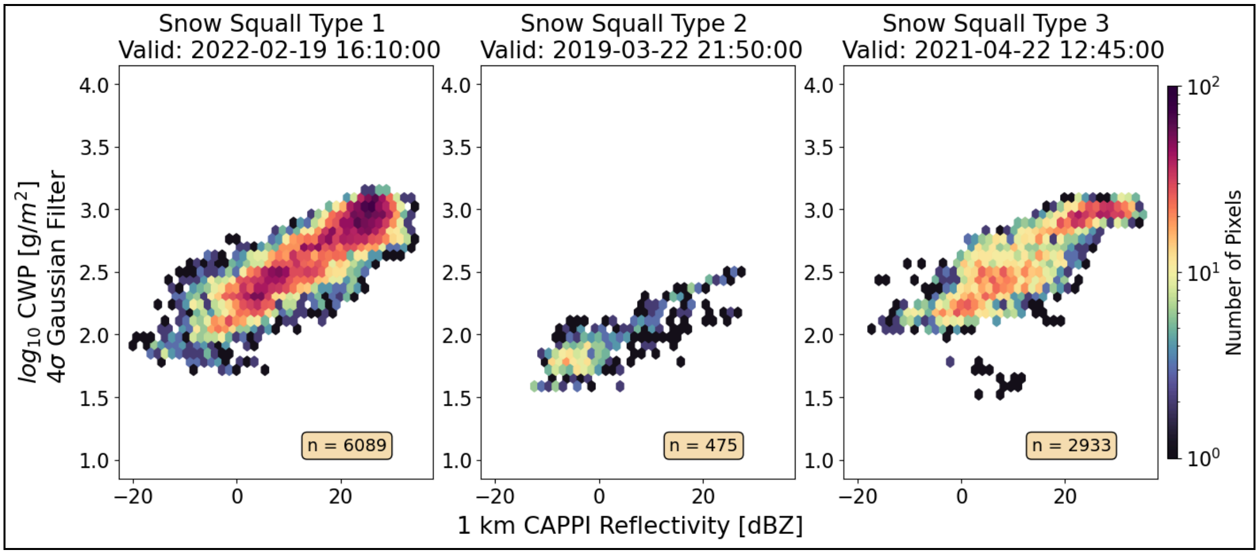

Figure 44. 2D histograms of CWP vs. CAPPI reflectivity, with a >50 μm effective radius DCOMP1 filter applied.

Chapter 4: Application: GREMLIN as an Example

As discussed in the introduction, GREMLIN is an operational synthetic radar product which seeks to predict radar reflectivity with a machine learning algorithm trained on colocated GOES-16/MRMS data. Given its relative similarity to the work done in this study, it is of interest to examine how well it performs for snow squalls compared to the methods presented in this study. It should be reiterated that GREMLIN was explicitly trained to work for deep convection - a fact explicitly acknowledged in both the original study and follow-up studies discussed in the introduction. Therefore, while it should be expected that it performs poorly for snow squall cases, it is still important to gauge its performance as a benchmark for future merged satellite-radar nowcasting tools that will be developed from the dataset compiled in this study.

Case Study: 28 March 2023

Data from GREMLIN are available online from 2021 through 2023. The snow squall climatology used in this study extends from late 2017 to early 2023. While every snow squall timestamp within the overlapping period could be matched with its GREMLIN equivalent, for the scope of this study a single case (28 March 2022) was chosen. This case was chosen in part because of its impact; this was an event with numerous snow squalls throughout the day, including a large accident as briefly mentioned in Section 1.1 (Medina 2022). S3 was the primary type, so several discrete snow squalls should be present in any given timestamp, serving as a useful benchmark in GREMLIN's ability to identify features at such scales (~50-100+ km^2).

GREMLIN data from 2022 (Hilburn 2023) was retrieved for the entire day, and matched with the nearest CLAVR-x timestamp during daylight hours. The GREMLIN composite reflectivity product was then interpolated to 0.5 km spatial resolution from its native 3 km grid conservatively using xESMF as described in the methodology. There may be some incredulity in such a method with respect to maintaining data integrity; Figure 45 should ameliorate such concerns. The conservative regridding method is designed not to change the structure of the data, but simply translate and increase its resolution as necessary. This allows direct pixel-for-pixel GREMLIN comparisons with the CLAVR-x and CAPPI reflectivity analysis that was previously described in this study.

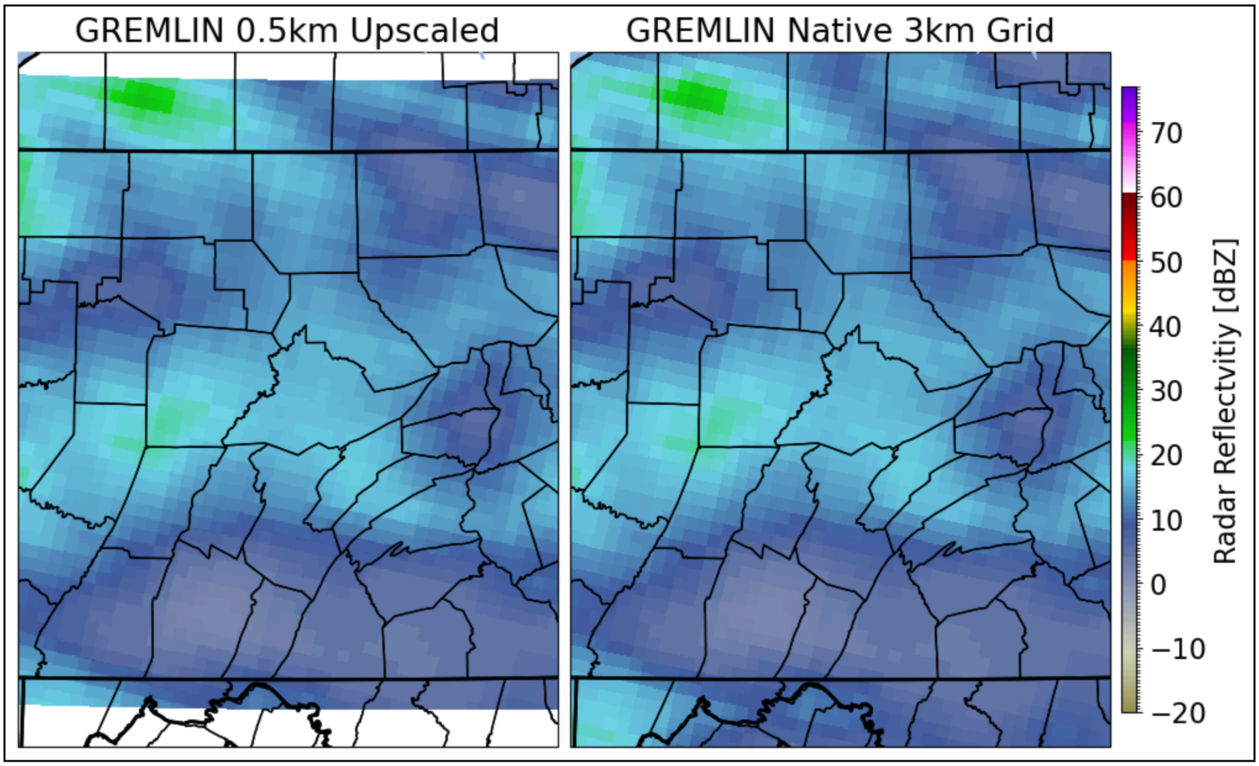

Figure 45. GREMLIN synthetic reflectivity centered over Central Pennsylvania

Results

Figure 46 showcases what the GREMLIN product looks like compared to the CAPPI reflectivity created for the same time. While a more fair comparison would be to use the MRMS composite reflectivity since that is what GREMLIN was trained on, the CAPPI reflectivity offers a more granular, high-resolution view on the scene. Regardless, GREMLIN struggles to capture any of the features seen in the CAPPI. The GREMLIN reflectivity is smoothed, indiscrete, and individual snow squall features are unable to be captured. GREMLIN does seem to capture the rough reflectivity values (~5-20 dBZ) but given the field's diffuse nature, the higher-end values are only seen in a select few regions, displaced from the core of the squalls seen in the CAPPI.

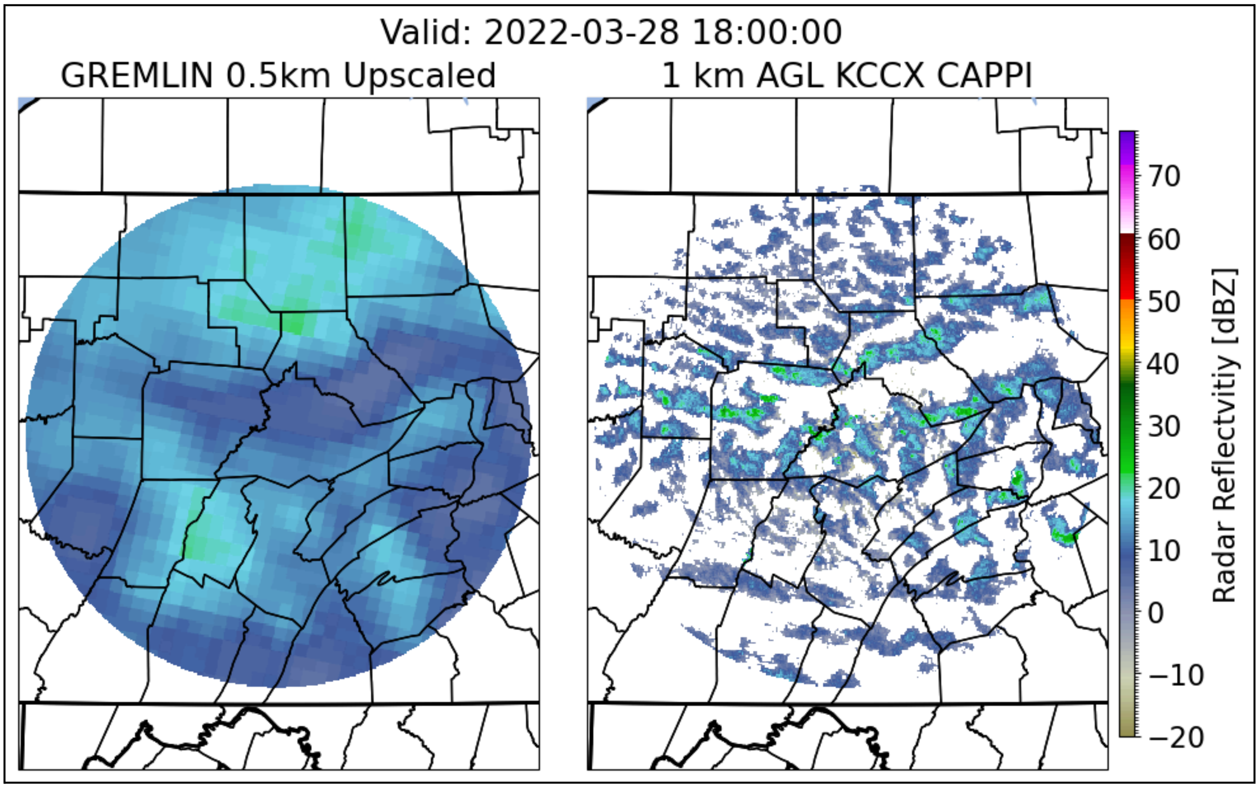

Figure 46. Radar reflectivity from GREMLIN (left) and KCCX (right) centred over Central Pennsylvania.

This is well represented quantitatively in Figure 47. GREMLIN systematically overpredicts reflectivity magnitudes compared to CAPPI values. The overprediction may be partially attributable to GREMLIN being unable to predict reflectivity lower than 0 dBZ, though even if one were to filter out CAPPI reflectivity values below 0 the issue would remain. In fact, the cluster of points in Figure 47 is nearly circular, indicating that GREMLIN is largely predicting the same "range" of values (~10-20 dBZ) for wide ranges of CAPPI reflectivity. This makes sense given the diffusivity mentioned before.

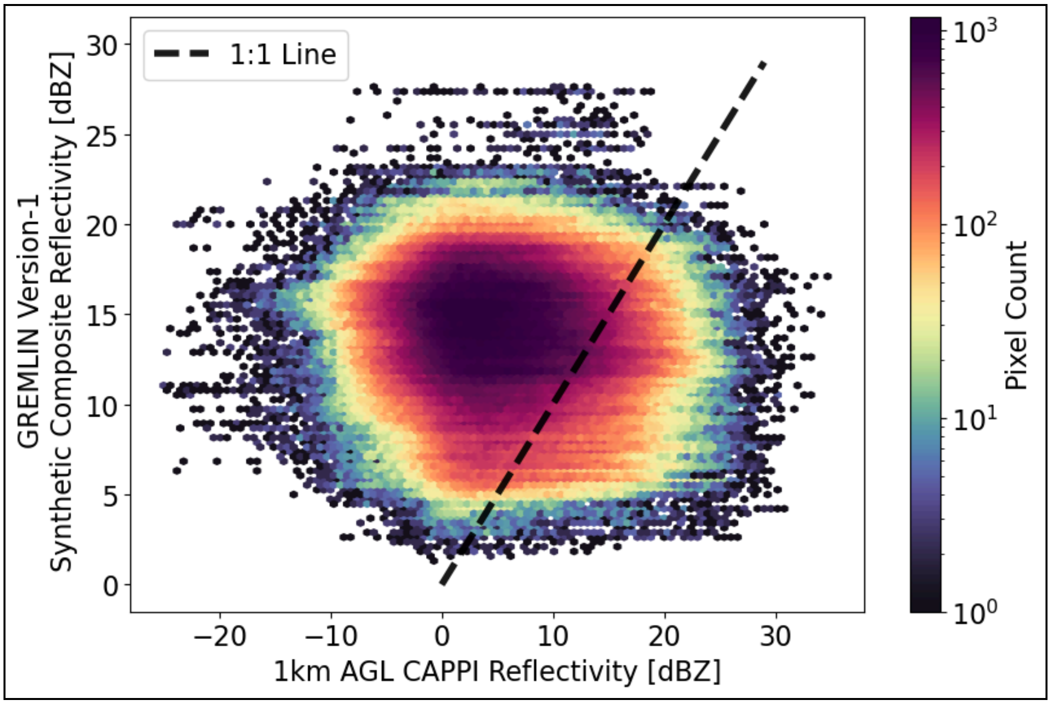

Figure 47. Density plot comparing CAPPI reflectivity to GREMLIN synthetic reflectivity.

Synthetic Reflectivity Using Cloud Water Path

Given how GREMLIN compares to CAPPI reflectivity, the next natural question would be how these results compare against reflectivity predicted by CLAVR-x products? Although it should be greatly emphasized that the intention of this work is not to simply make a synthetic reflectivity product, the creation of such a field has its utility. A crude method to do so would be the following:

1. Retrieve pixels which are matched up in both the CAPPI reflectivity and CWP field.
2. Take the logarithm of the CWP field and create a linear regression between the resultant field and the CAPPI reflectivity.
3. Use the regression equation on the CWP field to estimate reflectivity.

In this framework, no other filtering described in section 3.3 would be applied.

The result of this calculation for each timestamp is shown in Figure 48, along with the GREMLIN and CAPPI reflectivity.

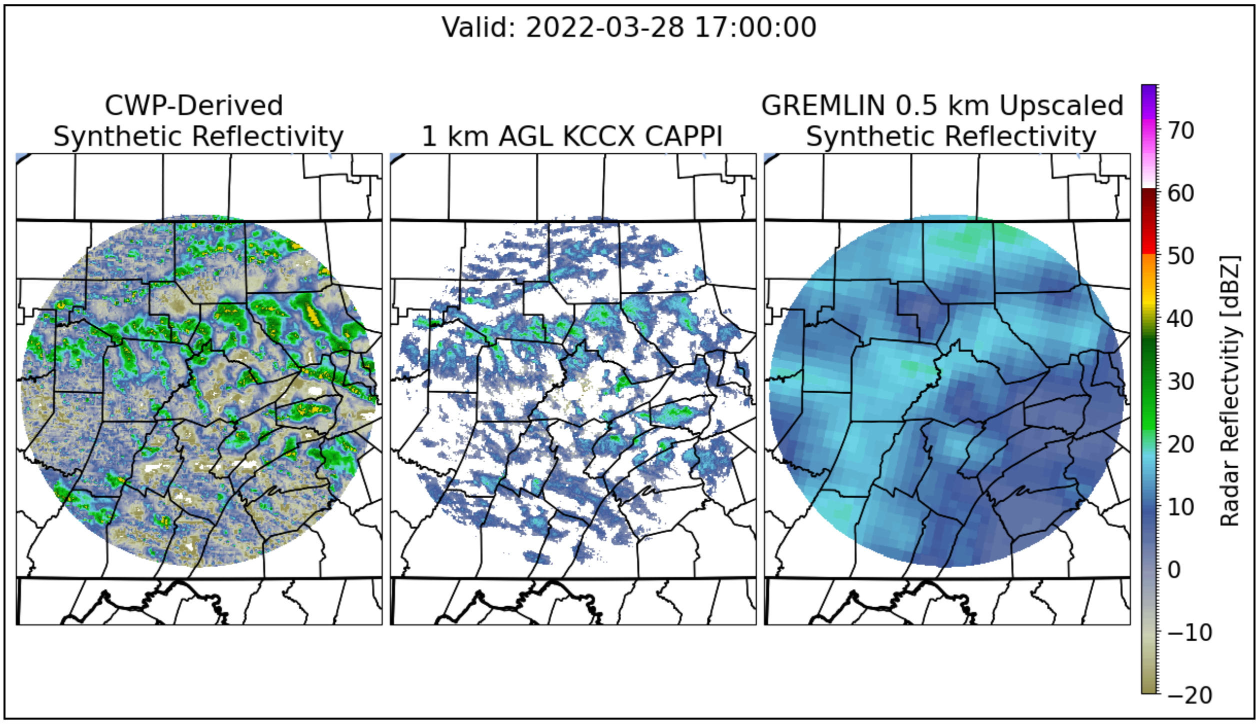

Figure 48. Radar reflectivity centered on Central Pennsylvania.

It is clear that increased resolution of the CWP-derived product impacts feature identification when compared to the GREMLIN. However, the CWP derived product does also overestimate reflectivity within the core of cells, often predicting values over 40 dBZ. This point is visualized in Figure 49, which is the same type of figure as Figure 48. The distribution of points is remarkably coherent, with the points clustering along a positively correlated linear trend between CAPPI and CWP-derived reflectivity. However, the slope of this line does not follow that of the 1:1 line; rather, the model tends to underestimate low reflectivities, and overestimates high reflectivities. The patterns of these biases are much more predictable than those observed in GREMLIN data.

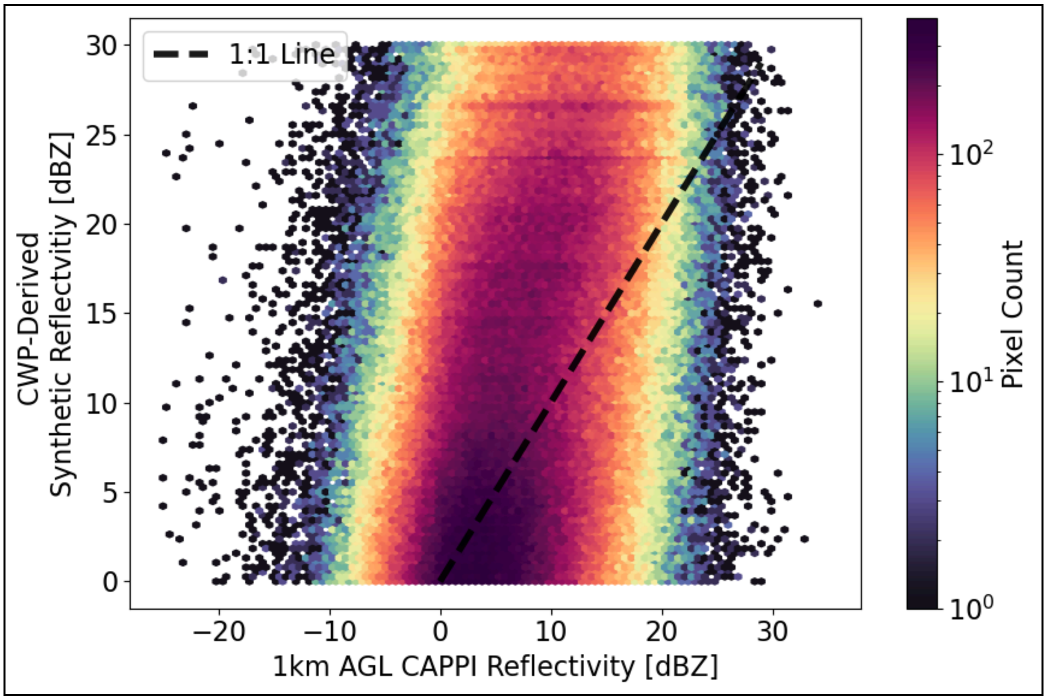

Figure 49. As in Figure 47, but with CWP-derived synthetic reflectivity.

Chapter 5: Discussion and Conclusion

The viability of CLAVR-x products colocated with WSR-88D reflectivity in the surveillance of convective snow over Northeast Pennsylvania was interrogated. The goal was to see whether CLAVR-x products had any utility in observing these snow squalls. To do this, two research questions were asked: 1) do these CLAVR-x products match up with radar products spatially? And if they do, 2) can what is observed between the two be related? Investigation into question 1 revealed that, indeed, CLAVR-x variable CWP and reflectivity do match up spatially, with match up rates generally around 50-70% for the bulk of cases. This rate depended somewhat on distance from radar, with matchup rates generally lower for farther distances.

It was then asked whether or not an empirical relationship existed between the radar and CWP, and generally the data was inconclusive. While a relationship could be made underneath certain circumstances, it was highly contextual and varied from case to case, and even timestamp to timestamp. It was then hypothesized that there must be some factor(s) which control the varying shape of these CWP-reflectivity plots, and that this could be found by looking at other variables in the CLAVR-x dataset which analyze the macrophysical characteristics of the snow squalls. Filtering the data by the effective radius of particles improved scores and reduced variability in CWP vs. reflectivity distributions. Results were mostly inconclusive with other CLAVR-x variables (cloud phase, height, etc.). As an additional test, a form of spatial averaging through a Gaussian convolutional filter was done, and R^2 scores did improve on average, sometimes by over 0.2. This says that there is some inherent uncertainty associated with the location and relative intensity of these CWP values. This may be tied back to uncertainties with the CLAVR-x algorithm itself, or it may be tied through a forest of assumptions and simplifications which are made to achieve the variables which are manipulated in this study, some of which was interrogated here.

5.1. Key Takeaways

The results show that, on average, CWP and reflectivity (at the thresholds set in this study) match up well spatially. Match up rates are regularly between 50-70% despite the distance thresholds used, and variance around this range is generally limited. The major caveat is that these results are framed by far fewer S1 and S4 cases as opposed to S2 and S3, which affects the statistical significance of these results . Still, these match up rates are seen qualitatively within the data, and generally there is less uncertainty governing understanding in this area.

The uncertainty compounds significantly for the development of an empirical relationship. Despite the promise that initial work had shown for this CWP-reflectivity relationship (and more broadly, relating CLAVR-x and radar variables together) the results from this study shows that such a relationship will require much more work. From the filtering attempts described in Section 3.3, a few key takeaways are:

1. Gaussian filtering improves R^2 scores (to a point), and it potentially works for a few reasons: 1) CLAVR-x fields can have large spatial heterogeneity, and averaging across pixels yields better results, 2) spatiotemporal parallax issues between radar and satellite get resolved.
2. Cloud effective radius (DCOMP1 in particular) works well to decrease the spread of CWP vs. reflectivity plots (and therefore improve R^2 scores), especially when in conjunction with 1).
3. Filtering by cloud phase can marginally improve scores, but CLAVR-x erroneously identifies the majority of clouds as supercooled liquid water and thus limits potential gains using this filtering method.
4. Radar VCP plays a role in matchup quality due to artifacting while creating CAPPI's. Future work should explore more contemporary radar gridding techniques to limit such error.

5.2. Future Work

This study has shown the potential of CLAVR-x to identify and quantitatively estimate precipitation from snow squalls, even if its utility is limited in its current capacity. From this work has arisen a large, multi-year dataset rich with nuance and heterogeneity, traits so often characteristic of real-world meteorological data. This study has shown that prior research into snow squalls is lacking, especially in the controls at the microphysical level. The heterogeneity displayed by the CWP-reflectivity relationship may seem impregnable only because there lacks existing contextual data to constrain understanding of them.

Therefore, future work should seek to 1) further constrain expected values of CWP by more thoroughly interrogating comparable studies on shallow cold season convection and/or using high-resolution large-eddy simulation models to simulate snow squalls. Such studies would also prove fruitful in assessing whether current microphysical parameterization schemes can replicate observed squalls—and therefore is a mirror to our microphysical understanding of them. In addition, 2) future work should utilize other satellite observations to cross compare with CLAVR-x to also further constrain expected values.

After such a time wherein the physics and the logistics of the problem have been thoroughly understood to the best of our ability, then it would be appropriate to leverage 3) a sophisticated machine learning model trained on the wealth of data within this dataset. Schneider et al's (2024) work could also be replicated for other comparable radars in the Northeast, thereby making an even larger dataset to both constrain CWP and serve as more training data. The output would be a probabilistic model which could, from only satellite data, predict snow squall intensity (a "nowcasting" tool). Such a tool has precedent; ProbSevere for thunderstorm intensity forecasts (Cintineo et al. 2018), and ThunderCast for lightning forecasts (Ortland et al. 2023). This study has taken the first steps towards the creation of such a tool; with more work, it may be fully realized.

References

Banacos, P., A. Loconto, and G. DeVoir, 2014: Snow squalls: Forecasting and hazard mitigation. *Journal of Operational Meteorology*, **2**, 130–151, https://doi.org/10.15191/nwajom.2014.0212.

Barthold, F. E., and D. A. R. Kristovich, 2011: Observations of the Cross-Lake Cloud and Snow Evolution in a Lake-Effect Snow Event. *Monthly Weather Review*, **139**, 2386–2398, https://doi.org/10.1175/mwr-d-10-05001.1.

Bender, L., and A. South, 2020: Environmental Controls on Banded Versus Cellular Organization of Mesoscale Snow Squalls in Western South Dakota 675. *18th Annual Student Conf., Phoenix, AZ, Amer. Meteor. Soc. S*, Vol. 254 of.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, **144**, 1669–1694, https://doi.org/10.1175/mwr-d-15-0242.1.

Braham, R. R., 1990: Snow Particle Size Spectra in Lake Effect Snows. *Journal of Applied Meteorology and Climatology*, **29**, 200–207, https://doi.org/10.1175/1520-0450(1990)029<0200:SPSSIL>2.0.CO;2.

Brook, J. P., A. Protat, J. S. Soderholm, R. A. Warren, and H. McGowan, 2022: A Variational Interpolation Method for Gridding Weather Radar Data. *Journal of Atmospheric and Oceanic Technology*, **39**, 1633–1654, https://doi.org/10.1175/jtech-d-22-0015.1.

Brown, R. A., T. A. Niziol, N. R. Donaldson, P. I. Joe, and V. T. Wood, 2007: Improved Detection Using Negative Elevation Angles for Mountaintop WSR-88Ds. Part III: Simulations of Shallow Convective Activity over and around Lake Ontario. *Weather and Forecasting*, **22**, 839–852, https://doi.org/10.1175/waf1019.1.

Campbell, L. S., and W. J. Steenburgh, 2017: The OWLeS IOP2b Lake-Effect Snowstorm: Mechanisms Contributing to the Tug Hill Precipitation Maximum. *Monthly Weather Review*, **145**, 2461–2478, https://doi.org/10.1175/mwr-d-16-0461.1.

Cintineo, J. L., and Coauthors, 2018: The NOAA/CIMSS ProbSevere Model: Incorporation of Total Lightning and Validation. *Weather and Forecasting*, **33**, 331–345, https://doi.org/10.1175/waf-d-17-0099.1.

Colby, F. P., D. Coe, M. Barlow, R. Brown, and E. Krajewski, 2022: A Climatology of Snow Squalls in Southern New England 1994–2018. *Monthly Weather Review*, **150**, 1541–1561, https://doi.org/10.1175/mwr-d-21-0082.1.

David, and Coauthors, 2017: The Ontario Winter Lake-Effect Systems Field Campaign: Scientific and Educational Adventures to Further Our Knowledge and Prediction of Lake-Effect Storms. *Bulletin of the American Meteorological Society*, **98**, 315–332, https://doi.org/10.1175/bams-d-15-00034.1.

Devoir, G., 2004: *10.2 high impact Sub-advisory snow events: the need to effectively communicate the threat of short duration high intensity snowfall*, accessed May 26, 2026, https://www.weather.gov/media/ctp/HISA/68261.pdf.

Hawkins, M., 2018: Service Change Notice 18-90 Updated. https://www.weather.gov/media/notification/pdfs/scn18-90swq_ dsw_oper_aaa.pdf.

Heidinger, A. K., and M. J. Pavolonis, 2009: Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology. *J. Appl. Meteor. Climatol.*, **48**, 1100–1116, https://doi.org/10.1175/2008jamc1882.1.

Heidinger, A. K., A. T. Evan, M. E. Foster, and A. Walther, 2012: A Naive Bayesian Cloud-Detection Scheme Derived from CALIPSO and Applied within PATMOS-x. *J. Appl. Meteor. Climatol.*, **51**, 1129–1144, https://doi.org/10.1175/jamc-d-11-02.1.

Heikenfeld, M., P. J. Marinescu, M. Christensen, D. Watson-Parris, F. Senf, Susan, and P. Stier, 2019: tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets. *Geoscientific Model Development*, **12**, 4551–4570, https://doi.org/10.5194/gmd-12-4551-2019.

Hilburn, K., 2023: GREMLIN CONUS3 Dataset for 2022 [Dataset]. *Dryad*.

Hilburn, K. A., I. Ebert-Uphoff, and S. D. Miller, 2021: Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations. *Journal of Applied Meteorology and Climatology*, **60**, 3–21, https://doi.org/10.1175/jamc-d-20-0084.1.

Kocin, P. J., L. W. Uccellini, J. W. Zack, and M. L. Kaplan, 1985: A Mesoscale Numerical Forecast of an Intense Convective Snowburst Along the East Coast. *Bulletin of the American Meteorological Society*, **66**, 1412–1424, https://doi.org/10.2307/26224594.

Kristovich, D. A. R., and Coauthors, 2000: The Lake—Induced Convection Experiment and the Snowband Dynamics Project. *Bulletin of the American Meteorological Society*, **81**, 519–542, https://doi.org/10.1175/1520-0477(2000)081<0519:tlceat>2.3.co;2.

——, 2017: The Ontario Winter Lake-Effect Systems Field Campaign: Scientific and Educational Adventures to Further Our Knowledge and Prediction of Lake-Effect Storms. *Bulletin of the American Meteorological Society*, **98**, 315–332, https://doi.org/10.1175/bams-d-15-00034.1.

Kulie, M. S., R. Bennartz, T. J. Greenwald, Y. Chen, and F. Weng, 2010: Uncertainties in Microwave Properties of Frozen Precipitation: Implications for Remote Sensing and Data Assimilation. *Journal of the Atmospheric Sciences*, **67**, 3471–3487, https://doi.org/10.1175/2010jas3520.1.

Kulie, M. S., L. Milani, N. B. Wood, S. A. Tushaus, R. Bennartz, and T. S. L'Ecuyer, 2016: A Shallow Cumuliform Snowfall Census Using Spaceborne Radar. *Journal of Hydrometeorology*, **17**, 1261–1279, https://doi.org/10.1175/jhm-d-15-0123.1.

Kumjian, M. R., O. P. Prat, K. J. Reimel, M. van Lier-Walqui, and H. C. Morrison, 2022: Dual-Polarization Radar Fingerprints of Precipitation Physics: A Review. *Remote Sensing*, **14**, 3706, https://doi.org/10.3390/rs14153706.

Lee, Y., and K. Hilburn, 2024: Validating GOES Radar Estimation via Machine Learning to Inform NWP (GREMLIN) Product over CONUS. *Journal of Applied Meteorology and Climatology*, **63**, 471–486, https://doi.org/10.1175/jamc-d-23-0103.1.

Liu, G., 2008: Deriving snow cloud characteristics from CloudSat observations. *Journal of Geophysical Research: Atmospheres*, **113**, https://doi.org/10.1029/2007jd009766.

Loisel, J., 1909: Squalls and Thunderstorms. *Monthly Weather Review*, **37**, 237–239, https://doi.org/10.1175/1520-0493(1909)37[237:sat]2.0.co;2.

Lundstedt, W., 1993: A Method to Forecast Wintertime Instability and non-lake Effect Snow Squalls across Northern New England. *noaa.gov*. Accessed May 26, 2026, https://repository.library.noaa.gov/view/noaa/6832.

Lyza, A. W., and Coauthors, 2026: The NOAA Hazardous Weather Testbed : FY2025 Activities Summary. *noaa.gov*, https://doi.org/10.25923/65kx-0a28.

Martin, F., J. Bernhardt, J. Banghoff, and S. Pesci, 2026: Development and evaluation of a virtual reality snow squall driving simulation. *The Visual Computer*, **42**, https://doi.org/10.1007/s00371-026-04409-x.

Medina, E., 2022: 80-Car Pileup on Pennsylvania Highway Leaves 6 Dead. *The New York Times*, March 30.

Milrad, S. M., J. R. Gyakum, E. H. Atallah, and J. F. Smith, 2010: A Diagnostic Examination of the Eastern Ontario and Western Quebec Wintertime Convection Event of 28 January 2010. *Weather and Forecasting*, **26**, 301–318, https://doi.org/10.1175/2010waf2222432.1.

Milrad, S. M., J. R. Gyakum, K. Lombardo, and E. H. Atallah, 2014: On the Dynamics, Thermodynamics, and Forecast Model Evaluation of Two Snow-Burst Events in Southern Alberta. *Weather and Forecasting*, **29**, 725–749, https://doi.org/10.1175/waf-d-13-00099.1.

NOAA, 2022: Volume Coverage Patterns (VCP). *Noaa.gov*. https://www.noaa.gov/jetstream/vcp_max.

NOAA National Severe Storms Laboratory, 2021: NEXRAD. *NOAA National Severe Storms Laboratory*. https://www.nssl.noaa.gov/tools/radar/nexrad/.

Ortland, S. M., M. J. Pavolonis, and J. L. Cintineo, 2023: The Development and Initial Capabilities of ThunderCast, a Deep Learning Model for Thunderstorm Nowcasting in the United States. *Artificial intelligence for the earth systems*, **2**, https://doi.org/10.1175/aies-d-23-0044.1.

Pavolonis, M., and C. Calvert, 2020: *Enterprise Algorithm Theoretical Basis Document For Cloud Type and Cloud Phase*, accessed May 26, 2026, https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Cloud_Type_v3_2020-06-01.pdf.

Peace, R. L., and R. B. Sykes, 1966: Mesoscale Study of a Lake Effect Storm. *Monthly Weather Review*, **94**, 495–507, https://doi.org/10.1175/1520-0493(1966)094<0495:msoale>2.3.co;2.

Pettegrew, B. P., P. S. Market, Raymond Herbert Wolf, R. L. Holle, and Nicholas, 2009: A Case Study of Severe Winter Convection in the Midwest. *Weather and Forecasting*, **24**, 121–139, https://doi.org/10.1175/2008waf2007103.1.

Reinking, R. F., and Coauthors, 1993: The Lake Ontario Winter Storms (LOWS) Project. *Bulletin of the American Meteorological Society*, **74**, 1828–1850, https://doi.org/10.1175/1520-0477-74-10-1828.

Rosenow, A. A., K. Howard, and José Meitín, 2018: Gap-Filling Mobile Radar Observations of a Snow Squall in the San Luis Valley. *Monthly Weather Review*, **146**, 2469–2481, https://doi.org/10.1175/mwr-d-17-0323.1.

Schneider, K., K. Lombardo, M. R. Kumjian, and K. Bowley, 2024: A Radar-Based 10-Year Climatology of Convective Snow Events in Central Pennsylvania. *Weather and Forecasting*, **39**, 965–994, https://doi.org/10.1175/waf-d-23-0187.1.

Schneider, R. S., 1990: Large-Amplitude Mesoscale Wave Disturbances Within the Intense Midwest Extratropical Cyclone of 15 December 1987. *Weather and Forecasting*, **5**, 533–558, https://doi.org/10.1175/1520-0434(1990)005<0533:LAMWDW>2.0.CO;2.

Schumacher, R. S., D. M. Schultz, and J. A. Knox, 2010: Convective Snowbands Downstream of the Rocky Mountains in an Environment with Conditional, Dry Symmetric, and Inertial Instabilities. *Monthly Weather Review*, **138**, 4416–4438, https://doi.org/10.1175/2010mwr3334.1.

Souto-Oliveira, C. E., M. de F. Andrade, P. Kumar, F. J. da S. Lopes, M. Babinski, and E. Landulfo, 2016: Effect of vehicular traffic, remote sources and new particle formation on the activation properties of cloud condensation nuclei in the megacity of São Paulo, Brazil. *Atmospheric Chemistry and Physics*, **16**, 14635–14656, https://doi.org/10.5194/acp-16-14635-2016.

Stephens, G. L., and Coauthors, 2002: The CloudSat Mission and the A-Train. *Bulletin of the American Meteorological Society*, **83**, 1771–1790, https://doi.org/10.1175/bams-83-12-1771.

Stowe, L. L., P. A. Davis, and E. P. McClain, 1999: Scientific Basis and Initial Evaluation of the CLAVR-1 Global Clear/Cloud Classification Algorithm for the Advanced Very High Resolution Radiometer. *Journal of Atmospheric and Oceanic Technology*, **16**, 656–681, https://doi.org/10.1175/1520-0426(1999)016<0656:SBAIEO>2.0.CO;2.

Vicente, G. A., J. C. Davenport, and R. A. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *International Journal of Remote Sensing*, **23**, 221–230, https://doi.org/10.1080/01431160010006935.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, **17**, 261–272.

von Lerber, A., D. Moisseev, D. A. Marks, W. Petersen, A.-M. Harri, and V. Chandrasekar, 2018: Validation of GMI Snowfall Observations by Using a Combination of Weather Radar and Surface Measurements. *Journal of Applied Meteorology and Climatology*, **57**, 797–820, https://doi.org/10.1175/jamc-d-17-0176.1.

Walther, A., and A. K. Heidinger, 2012: Implementation of the Daytime Cloud Optical and Microphysical Properties Algorithm (DCOMP) in PATMOS-x. *Journal of Applied Meteorology and Climatology*, **51**, 1371–1390, https://doi.org/10.1175/jamc-d-11-0108.1.

Walther, A., and W. Straka, 2020: *Algorithm Theoretical Basis Document For Daytime Cloud Optical and Microphysical Properties (DCOMP)*, accessed May 26, 2026, https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Daytime_Cloud_Optical_and_Microphysical_Properties(DCOMP)_v1.2_2020-10-09.pdf.

Wiggin, B. L., 1950: Great Snows of the Great Lakes. *Weatherwise*, **3**, 123–126, https://doi.org/10.1080/00431672.1950.9927065.

Wood, N. B., T. S. L'Ecuyer, F. L. Bliven, and G. L. Stephens, 2013: Characterization of video disdrometer uncertainties and impacts on estimates of snowfall rate and radar reflectivity. *Atmospheric Measurement Techniques*, **6**, 3635–3648, https://doi.org/10.5194/amt-6-3635-2013.

Zhuang, J., and Coauthors, 2025: pangeo-data/xESMF: v0.9.2 - Third time's a charm. *Zenodo (CERN European Organization for Nuclear Research)*, https://doi.org/10.5281/zenodo.4294774.