UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

OVERSHOOTING CONVECTION IN THE CONTIGUOUS UNITED STATES: SEVERE HAIL OCCURRENCE, STORM TOP CHARACTERISTICS, AND STRATOSPHERIC HYDRATION

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
Degree of
Doctor of Philosophy

By

ELISA MIKAELA MURILLO
Norman, Oklahoma
2022

OVERSHOOTING CONVECTION IN THE CONTIGUOUS UNITED STATES: SEVERE HAIL OCCURRENCE, STORM TOP CHARACTERISTICS, AND STRATOSPHERIC HYDRATION

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Cameron Homeyer, Chair

Dr. Elinor Martin

Dr. Naoko Sakaeda

Dr. Amy McGovern

Dr. Richard Zielinski

Acknowledgments

There is no way that the words written here will adequately thank all the people who have played such a vital role in me getting to this point, but I'm determined to try. I would not have completed this degree and accomplished what I have if it were not for SO many people there to support and encourage me throughout this wild journey. Having met my wonderful and brilliant husband, Dr. Brian Greene, in our first graduate class has given us the opportunity for deep understanding and empathy over the past 6 years together. Being quarantined together due to the COVID-19 pandemic for over a year immediately following our wedding only deepened our connection as we had to truly lean on each other in such a painful time. Although we lost our CeCe pup, she's proud of us for making it through, and now we can celebrate with our Rose pup! We love doing things together so much that we decided to defend our dissertations nearly simultaneously...which was MASSIVELY stressful, but so worth it. I am privileged to call you, Dr. Brian Greene, my partner, husband, and best friend.

Next, I would like to thank Dr. Cameron Homeyer, my advisor and "science dad". Your unconditional support for my interests, my struggles, and what is best for me has truly changed my life. Regardless of how I showed up, whether panicked, excited, defeated, motivated, depressed, determined, or overwhelmed, you made me feel safe and valued, and also celebrated my authenticity. You have consistently shown me how to leverage my strengths to improve my weaknesses, and helped me rethink what I had considered weaknesses. You've opened your home for us and accepted us as family when we had no means to be with other family. You trusted us to adopt Rose from y'all, and she has been the perfect addition to our little family. Because of you and your family, I've grown in so many ways and feel more excited to be myself in as many spaces as possible. I'm proud to be part of the Homeyer family and proud to have y'all as part of our Murillo-Greene family.

Speaking of family, I want to thank the massive group of people I am privileged to call my family. Again, there's no way I can properly thank each individual person for their love and support, so I'll have to be brief. Starting with my side of the family, my mom (Evelyn), dad (Gus), step-mom, (Tracey), brother (Gus), sister-in-law (Kelsie), sister (Ashlie), and brother-in-law (Will), thank you for putting up with my unpredictable schedule, frequent meltdowns, and ever-growing responsibilities. Most importantly, thank you for helping me remember why I did this in the first place, and for always believing in me. To my aunts, uncles, grandparents, and more, thank you for all the encouragement and understanding, especially as I wasn't able to spend as much time with y'all that I would have wanted. To all the younger members, thank you for always making me smile and giving me hope. I love getting to know each of you as you grow into your own person, and I'm excited to continue following y'all along your own path. On Brian's side, my mother-in-law (Diane), father-in-law (Jim), brother-in-law (Kevin), and sister-in-laws (Emily, Megan), thank you for accepting me and all my uniqueness into your family with open arms. To my aunts, uncles, cousins, and grandparents, thank you for embracing our 2 families together. I couldn't have chosen a better family to marry into if I tried. I also want to acknowledge just some of the friends and family whom we've lost along the graduate school journey (Miguel, Grandma Gayle, Great-grandma Methven, Mamistella, Uncle Tom, Aunt Nancy, Grandma Greene, Whitley, Joseph, Christian, Dominick, Shawn). You each have your own story that we will continue to share throughout our life, and you will not be forgotten.

Next, I want to thank all the incredible friends/chosen family I've made throughout my life, including my time at OU, ULM, and LSMSA, and from childhood. Thank y'all for the support, patience, excitement, and understanding throughout the many years of school. I also want to thank so many of my Tripp-C my research group members (Thea, Amanda, Emily, Andrea, Dan, Joel, Alex, Tao, Rachael, Lexy, Aodhan, Shawn, Russell) for the invaluable collaboration, support, encouragement, mutual wallowing, and so much more. Thank you 5900 office mates (Melanie, Bryony, Margaret, Alethia, Daniel, Allie, Veronica) for contributing to the "crying office" and making our area a safe place for so many, including myself. I also want to thank the CHALA family, especially those who led the group's creation and the BAMS publication. Getting to know other Latinos/as/es/x who were also meteorologists/geoscientists has played the

biggest role in me finally feeling like I belong in this field. Because of y'all, I am now thrilled to be a part of this community and continue showing up as my authentic self.

I also want to thank the many mentors/role models I've had throughout my education, which include previous professors, collaborators, administrative staff, and team leads. Thank you Dr. Todd Murphy for the support and understanding I so desperately needed during one of the hardest times of my life. Thank you Dr. Daphne LaDue for taking a chance on me, making me a part of the OU REU family, and standing up for me in ways I didn't know were possible, making me feel like I really do matter. Thank you to Dr. Jay and Ms. Mangum for your motivation, support, and genuine interest in me as a whole person. Without all of y'all, I wouldn't have believed I was capable of graduate school. Thank you to all the School of Meteorology staff and leadership throughout the years for reminding me that my voice matters and helping me use it. Thank you to my committee members (Drs. Elinor Martin, Naoko Sakaeda, Amy McGovern, Richard Zielinski) for your continual understanding and support throughout the PhD process. I don't believe I could have chosen a better committee. Next, I want to thank the awesome collaborators I've gotten to work directly with while in graduate school (Kris Bedka, Dr. John Allen, Dr. Matt Kumjian). I'm so grateful for the insight and support I receive from y'all and so proud of the work we have completed. I am a better scientist because of y'all, and I look forward to any opportunities to work together in the future. I also want to thank the full DCOTSS team for the unparalleled experience I've had as a member of the team. The leadership's overwhelming dedication for proper accommodations and mutual respect has truly reshaped my understanding of large field campaigns. Participation in this campaign has also provided me with the fantastic opportunity to work directly with Dr. Gretchen Mullendore after graduation as a Post-Doc. Thank you, Gretchen, for believing in me, and I cannot wait to see what these next few years have in store for us.

This whole section is in no way an exhaustive list of all the people who've supported and encouraged me over the years, but you all are on my mind and in my heart, as cheesy as that sounds. I also want to thank all of the people across the globe that have, by choice or not, sacrificed their own time and safety to support those in need during the COVID-19 pandemic. We would actually not be here if it were not for the tireless efforts of essential employees, physical and mental healthcare workers, and so many others. Thank you.

Lastly, I want to thank the funding sources that made this work possible: The NASA awards NNX15AV81G and 80NSSC19K0347, as well as the NOAA Weather Program Office Joint Technology Transfer Initiative grant NA20OAR4590356 to the University of Oklahoma.

Table of Contents

List Of Tables

List Of Figures

Abstract

Deep convection that overshoots the tropopause have been recognized for its relevance to stratospheric composition and prolific severe weather production for decades. Notably, tropopause-penetrating storms can inject cloud ice above the tropopause and the broader storm anvil, producing ice clouds called above-anvil cirrus plumes (AACPs). When transported into the upper troposphere and lower stratosphere (UTLS), the convectively-influenced air can modify the surrounding environment, resulting in short- and long-term impacts. Supercell storms, argued to be responsible for the majority of U.S. severe hail reports (especially for the largest sizes), also frequently produce AACPs due to the strong updrafts needed for supercell maintenance. However, U.S. hail report databases are plagued by serious limitations, including biases in reported sizes, occurrence time, and location, which has led to the development of various radar-based estimates of severe hail occurrence and hail size. Many established relationships between radar-observed storm characteristics and severe hail occurrence have been found using data for few storms and in isolation from other radar metrics. In addition to previous work being spatially and temporally limited, recent improvements to radar-based metrics warrant detailed comparisons between original and updated parameters.

While radar-based analyses of storm and associated severe report characteristics have a rich history, much less work has been done on these aspects of explicitly AACP-producing storms and the AACP itself. Recent studies of AACPs have shown large variability in their satellite-observed features, particularly in infrared (IR) imagery, although many of the causes remain unknown. The typical altitude range in relation to the storm anvil and tropopause height, in situ trace gas observations, and stratospheric hydration variability are all aspects of AACPs that have little-to-no existing research. The recently conducted NASA Earth Ventures Suborbital field project, Dynamics and Chemistry of the Summer Stratosphere (DCOTSS), was specifically designed to collect in situ observations and multi-satellite imagery of convectively-injected material

(i.e. AACPs) in the stratosphere to better understand their effects on the chemistry, dynamics, and radiative features that are characteristic of the upper troposphere and lower stratosphere (UTLS). By leveraging the different viewing geometry between two satellites simultaneously observing one region, stereoscopy (stereo) can estimate cloud-top heights for AACPs and other cloud features. Although a few studies have utilized stereo for older generations of satellites, there lacks a focused effort to update stereo methods for next-generation satellites.

Motivated by the vast knowledge gaps in AACP understanding, associated climate impacts, existing challenges for objective severe hail identification, and recent DCOTSS observations of AACPs, the four projects detailed herein aim to better understand deep, severe convective storms on climatological timescales and their potential impacts on the climate by utilizing long data records and/or recent, novel observations. The first study aims to quantify severe hail fall characteristics during a markedly longer time period than previous studies, using both radar observations and reanalysis data, and found that an environmentally-filtered radar-based U.S. hail climatology provided the greatest agreement with reports. The second study evaluates the reanalysis-represented environments of radar-identified supercell populations with different maximum reported hail size, and found that the greatest differences exist in the wind speed perpendicular to storm motion. The third study compares warm and cold AACP-producing storms (based on IR signatures) using ground-based radar observations and reanalysis data to better understand the different AACP IR signatures. This study found that cold AACPs tend to occur in tropical environments (higher, cold-point tropopauses), while warm AACPs tend to occur in mid-latitude environments (lower tropopauses and an isothermal layer in the lower stratosphere). Storm-relative environmental characteristics indicate that cold AACPs are predominantly tropospheric phenomena, while warm AACPs reside in the lower stratosphere. The fourth, and final study applies a recently developed stereoscopic cloud-top height retrieval algorithm for modern GOES visible imagery to DCOTSS research flight 13 (RF13) on 31 May 2022 for comparison with in situ water vapor observations. Preliminary stereo results show moderate agreement with various observational datasets but concerning nonphysical features.

Overall, these studies provide further evidence that radar-based hail identification and size estimation can benefit from incorporating environmental information, particularly in a storm-relative framework. These studies additionally highlight the importance of understanding trace gas composition and satellite observations of AACPs to accurately characterize their impacts on the UTLS.

Chapter 1

Introduction

Deep convection that overshoots the tropopause (or some other atmospheric stable layer, e.g., the equilibrium level or level of neutral buoyancy) can inject cloud ice above said layer and the broader storm anvil, producing ice clouds called above-anvil cirrus plumes (AACPs; Setvák and Doswell III, 1991; Wang, 2003; Levizzani and Setvák, 1996; Setvák et al., 2010; Homeyer et al., 2017). When transported into the upper troposphere and lower stratosphere (UTLS), the convectively-influenced air can modify the surrounding environment, resulting in short- and long-term impacts (Fujita, 1982; Holton et al., 1995; Wang, 2003; Stohl et al., 2003; Mullendore et al., 2005; Gettelman et al., 2011; Wang et al., 2016; Anderson et al., 2017; Seguchi et al., 2019; Smith, 2021; Zou et al., 2021). Additionally, tropopause-penetrating storms (e.g., Adler and Fenn, 1981; Bedka et al., 2015; Sandmæl et al., 2019), especially those that produce AACPs (Fujita, 1982; McCann, 1983; Brunner et al., 2007; Setvák et al., 2010, 2013; Bedka et al., 2015; Homeyer et al., 2017; Bedka et al., 2018b; Kunz et al., 2018; Mecikalski et al., 2021) have been recognized as prolific severe weather producers, and their utility for severe weather identification and forecasting has been documented for decades. Bedka et al. (2018b) evaluated the relationship between supercell and AACP-producing storms, finding that the majority of supercell storms produce AACPs; this is not surprising given the strong updrafts necessary for supercell and AACP development. Further, supercells are argued to be responsible for a majority of severe hail observed in the U.S., especially for the largest sizes ($\geq$2-in. (5 cm) maximum dimension; e.g.,

Thompson et al., 2003; Duda and Jr., 2010; Blair et al., 2011; Smith et al., 2012; Blair et al., 2017; Murphy et al., 2022).

Motivated by the vast knowledge gaps in AACP understanding, associated climate impacts, existing challenges for objective severe hail identification, and recent field campaign observations of AACPs, each of the four projects discussed herein focus on various aspects of hail, AACPs, and their storm and environmental characteristics. The objectives of this work are to better understand severe overshooting convective storms on climatological timescales, increase our knowledge of AACPs, and gain more insight into the potential impacts of such storms on the UTLS climate and at the surface. These goals are achieved by utilizing long records of numerous data sources and recent, novel observations and associated approaches.

1.1 Hail Occurrence: Remote Sensing and Storm Environments

Severe hail accounts for the vast majority of severe weather-induced property loss in the United States (U.S.), and supercell storms (those with mid-level rotation broadly collocated with their updraft) are argued to be responsible for a majority of severe hail observed in the U.S., especially for the largest sizes ($\geq$2-in. (5 cm) maximum dimension; e.g., Thompson et al., 2003; Duda and Jr., 2010; Blair et al., 2011; Smith et al., 2012; Blair et al., 2017; Murphy et al., 2022). However, our understanding of hail-producing storms and spatiotemporal variability of hail fall is limited due to numerous documented biases associated with hail reports (Kelly et al., 1985; Bardsley, 1990; Witt et al., 1998b; Fraile et al., 1992; Changnon, 1999; Schaefer et al., 2004; Doswell et al., 2005; Trapp et al., 2006; Ortega et al., 2009; Allen and Tippett, 2015; Blair et al., 2017; Allen et al., 2017; Witt et al., 2018; Murillo and Homeyer, 2019). There are a

variety of factors that can introduce errors to hail reports, including (but not limited to) population density near a storm, concurrent weather hazards, and the use of reference objects to estimate size (Kelly et al., 1985; Allen and Tippett, 2015; Allen et al., 2017; Witt et al., 2018). The largest and most quality-controlled U.S. hail report database produced by the Storm Prediction Center (SPC) is also limited by an exclusion of small hail (sizes below 19.05 mm [0.75 in.]) and an overall under-representation of severe hail frequency (Cecil, 2009; Ortega et al., 2009; Allen et al., 2015b; Blair et al., 2017). These limitations prevent testing of objective hail identification methods for the full range of possible hail sizes, which has motivated multiple efforts to fill this data gap (Ortega et al., 2009; Elmore et al., 2014). For a more extensive summary of the myriad limitations and errors associated with hail reports, see Allen and Tippett (2015).

Despite these limitations, hail reports provide one of the most confident ground-truths of hail fall. Thus, numerous studies have evaluated U.S. hail reports from different sources and time periods to establish U.S. spatiotemporal distributions of hail occurrence and size (e.g., Allen and Tippett (2015) and references therein; Changnon, 1999; Allen et al., 2015a,b; Childs and Schumacher, 2019). Overall, these studies have found a hail fall frequency maximum in the U.S. Great Plains, which shifts northward as Summer progresses, and a secondary frequency maximum in the Southeast U.S. (SEUS).

Alternative approaches to better understand and estimate severe hail occurrence, maximum hail size, and broader storm characteristics have increased during the past several decades, utilizing various remote sensing tools and statistical models driven by environmental parameters. Radar observations have been the primary remote-sensing tool used to study severe hail events and storms. Early work using single-polarization radar revealed a broad relationship between radar reflectivity (Z, which depends on

particle concentration and size) and hail size, as well as the maximum altitude (absolute and/or above the environmental 0°C level) reached by high-Z ($\geq$45 dBZ) radar echoes (e.g., Donaldson Jr., 1958, 1959; Geotis, 1963; Mather et al., 1976; Dye and Martner, 1978; Waldvogel et al., 1979). Although broad relationships exist between Z and hail size, there is no direct relationship due to complex backscattering behavior that depends on radar wavelength and hail size, which results from hail size approaching or exceeding the radar wavelength. These effects are often referred to as resonance scattering, which is unfortunately common in volumes containing severe hail (e.g., Herman and Battan, 1961; Atlas and Wexler, 1963; Bohren and Battan, 1982; Aydin et al., 1986; Kumjian et al., 2018; Jiang et al., 2019). Beginning shortly after the establishment of operational radar networks in the U.S. and elsewhere, various metrics involving vertical integration of Z have since been used for hail size estimation, with varying success.

A commonly used metric to estimate hail occurrence and size from radar is the single-polarization parameter "maximum expected size of hail" (MESH) from Witt et al. (1998a). Briefly, MESH represents the vertically integrated Z in a storm to estimate the presence and depth of hydrometeors that have the expected scattering characteristics of hail stones. Previous studies have used MESH as the main hail identification method and compared their results to existing hail reports (Cintineo et al., 2012; Nisi et al., 2016b; Allen et al., 2017; Schlie et al., 2018; Ortega, 2018; Nisi et al., 2016a; Murillo and Homeyer, 2019; Tang et al., 2019; Allen et al., 2020; Warren et al., 2020). For those focusing on U.S. hail fall patterns, a notable difference is found between the radar-based and report-based assessments; although most features are similar between the two methods, a local maximum in reports is seen the SEUS that is not replicated in radar-based analyses. Overall, key MESH-related results from several of these studies indicate that the parameter does not produce reliable hail size estimates

(e.g., see detailed analysis by Bunkers and Smith (2013), Ortega (2018), and Murillo and Homeyer (2019)), and is more skillful at identifying hail occurrence rather than hail size. Additionally, most studies using MESH have been carried out over a limited time period, lack extensive comparison with report distributions, or (in some cases) have only leveraged MESH as a secondary source of identification. Murillo and Homeyer (2019) recently evaluated numerous hail identification parameters to assess their ability to discriminate between severe hail producing and non-severe hail producing storms and found the original MESH, in addition to two revised MESH calculations, produced skillful storm discrimination. The revised MESH utilized a significantly larger number of reports than that used in the original MESH configuration Witt et al. (1998a) and could be utilized to assess broader spatiotemporal hail frequency in the U.S.

Kinematic techniques for estimating hail size were developed once Doppler (i.e. radial) radar velocity observations became operationally available. Studies often combine Doppler velocities with derived parameters, most frequently including storm-top divergence (e.g., Witt and Nelson, 1991; Boustead, 2008) and rotation within the hail growth region (between the 0°C and −30°C isotherms) inferred from azimuthal shear (e.g., Witt et al., 1998a; Blair et al., 2011; Witt et al., 2018; Gutierrez and Kumjian, 2021). These two metrics in particular, storm-top divergence and mid-altitude rotation, have shown to increase in magnitude with increasing hail size, which demonstrate a stronger updraft and mesocyclone, respectively, for supercells. These relationships are also broadly consistent with the strong updrafts required to support hailstone residence time in the hail growth region, especially for increasingly large hailstones (e.g., Johnson and Sugden, 2014; Gutierrez and Kumjian, 2021), and greater likelihood of favorable pathways for growth (e.g., Kumjian et al., 2021). However, despite radar-derived hail identification likely being the most documented of the hail report proxy methods, our knowledge regarding the best practices for objective hail identification,

and especially, size estimation using radar observations is still limited (Allen et al., 2020). This is ultimately due to the aforementioned complexity of observing hail with radar and the lack of a fully representative hail report database.

Satellite observations have also been leveraged for hail assessments, particularly useful for global hail climatology. Several studies have established hail occurrence metrics utilizing space-borne passive microwave imagery (e.g., Cecil, 2009; Cecil and Blankenship, 2012; Ferraro et al., 2015; Punge et al., 2017; Ni et al., 2017; Mroz et al., 2017, 2018; Bedka et al., 2018a; Bruick et al., 2019). Earlier work found similar distributions to those diagnosed from reports and radar over the U.S., though they noted that their methods likely produced overestimates in tropical regions and underestimates in mid-latitudes globally. These biases were a result of differing tropopause heights between the two environmental regimes. Mid-latitude overshooting convective storms often reside in the warmer, stratospheric environment as opposed to the colder, upper troposphere. More recent studies have made efforts to mitigate these biases and result in more consistent results with other hail climatologies (Bang and Cecil, 2019). Although satellite observations display realized convection and typically feature large spatial domains, the low temporal and horizontal resolutions, coupled with the 2-dimensional nature of microwave imagery make it difficult to discern vertical and small-scale characteristics of storms in these data. It is also difficult to establish a long-term record of satellite imagery due to the unique challenges associated with space-borne platforms. Specifically, it is uncertain how long a given platform remains operational, and there are often measurement differences between successive observing platforms.

Supercell physics and dynamics conducive to severe hail growth have been increasingly studied with numerical models over the past several decades. Such studies often leverage convection-resolving simulations (real or ideal) of supercell storms and/or a

trajectory model to elucidate hail growth pathways, source locations of hail embryos, and key dynamical features supporting hail growth to various sizes (e.g., Kumjian and Lombardo, 2020, and references therein). Recent findings include the importance of the vertical profile of storm-relative wind, which significantly affects the updraft width and mesocyclone structure that can lead to increased maximum hail size (e.g., Nelson, 1983; Dennis and Kumjian, 2017; Peters et al., 2019; Kumjian and Lombardo, 2020; Kumjian et al., 2021), and a more nuanced dependence of maximum hail size on convective available potential energy (CAPE) (Lin and Kumjian, 2022).

Due to the limitations of previously mentioned direct and indirect methods of hail identification, environmental parameters from proximity soundings and/or reanalyses have been leveraged as a proxy for hail occurrence and in combination with severe hail reports. Environments with favorable ingredients for convective storms and hail reaching the surface include various combinations of vertical wind shear, storm-relative winds, storm-relative helicity (SRH), and CAPE (e.g., Johnson and Sugden, 2014; Allen et al., 2015b; Barrett and Henley, 2015; Mohr et al., 2015a,b; Prein and Holland, 2018; Gensini and Tippett, 2019; Tang et al., 2019; Gutierrez and Kumjian, 2021). Allen et al. (2015b) developed the first model to predict hail frequency over the U.S. by combining convective precipitation, 0–3 km SRH, 180 hPa mixed-layer CAPE, and mean-specific humidity in the lowest 90 hPa of the atmosphere as hail identifiers and showed similar patterns to those seen in reports, albeit with some timing errors in the seasonal cycle. A more recent study by Prein and Holland (2018) focused on observable environmental parameters, including most unstable CAPE, melting level altitude, 0–3 km vertical wind shear and 0–3 km SRH. The overall U.S. spatial distribution and seasonal cycle of observed hail frequency was reproduced, however the SEUS and regions on the lee side of mountains were biased low, and the Central Plains was biased high.

Additionally, these studies have found a general increase in maximum reported hail size and CAPE (Edwards and Thompson, 1998; Johnson and Sugden, 2014; Taszarek et al., 2020), though most of the environmental parameters studied display significant overlap for adjacent hail-size categories. Reporting limitations and biases have been an important constraint on progress in understanding relationships between observed hail size and storm environments (e.g., Allen and Tippett, 2015; Blair et al., 2017). Moreover, the vast majority (if not all) of the extensive environmental analyses in past work have focused on relating bulk metrics such as CAPE, mid-level temperature lapse rates, 0–6 km wind shear, 0–3 km SRH, and others to observed severe hail size. Some studies leverage a combination of these and related fixed-layer measurements as composite parameters to assess hail potential, which have been more successful than evaluation of the bulk metrics alone (e.g., Johnson and Sugden, 2014). Thus, improved understanding of hail size and environment relationships may be possible with more comprehensive analysis of atmospheric profiles near well-characterized severe hail storms. Compared to satellite-based approaches, hail distributions (in both size and geographic location) derived from environmental proxies benefit from the similarly large spatial domains with the added benefit of higher temporal resolution. One unique limitation, however, is the lack of known realization of these environments (i.e., how often these environments result in severe hail occurrence and what their hail size maxima would be).

1.2 Overshooting Storm Tops: Characteristics and Environments

AACPs are primarily identified in visible (VIS) satellite imagery as a region within a storm anvil that is relatively smoother in texture when compared to the broader anvil

and casts a shadow on the underlying anvil, especially near sunset (Figs. 3.1a & 3.1b). Often, but not always, the VIS feature is co-located with warm brightness temperature anomalies in infrared (IR) imagery that are surrounded by colder brightness temperature anomalies characteristic of the broader storm anvil (Fig. 3.1c), a feature commonly referred to as the enhanced- or cold- U, V, or ring (Adler et al., 1983; McCann, 1983; Brunner et al., 2007; Setvák et al., 2010; Púčik et al., 2013; Homeyer, 2014). Most prior studies focus on such warm AACPs identified in IR imagery, occasionally supplemented with confirmation from VIS imagery.

In these studies, warm AACPs have been demonstrated to represent stratospheric injection of cloud ice, which then sublimates into water vapor. This convective UTLS transport can heavily influence the chemistry, dynamics, and radiative properties, thereby significantly impacting associated climate forcing (Fujita, 1982; Holton et al., 1995; Wang, 2003; Stohl et al., 2003; Mullendore et al., 2005; Gettelman et al., 2011; Wang et al., 2016; Anderson et al., 2017; Seguchi et al., 2019; Smith, 2021; Zou et al., 2021). In particular, anticyclonic circulation in summer over North America can trap convectively-transported air for long periods, potentially exacerbating their climate impacts (Solomon et al., 2016; Cooney et al., 2018). The recently conducted NASA Earth Ventures Suborbital field project, Dynamics and Chemistry of the Summer Stratosphere (DCOTSS), was specifically designed to sample convectively-injected material (i.e. AACPs) in the stratosphere to better understand their effects on the chemistry, dynamics, and radiative features that are characteristic of the UTLS. The NASA ER-2 high-altitude research aircraft, fitted with numerous instruments that measure trace gases and aerosol properties, was used to collect in situ observations up to approximately 69,500 feet. Based in Salina, KS, the project consisted of 23 research flights during the Summers of 2021 and 2022.

Satellite observations have also been routinely used to determine the altitude of various cloud tops, with particular focus on tropopause-overshooting storm tops (e.g., Berendes et al., 2008; Lindsey and Grasso, 2008; Rosenfeld et al., 2008; Bedka et al., 2010; Griffin et al., 2016). Single-satellite, cloud-top retrievals based on VIS, however, are less reliable at sunrise/sunset, miss frequent nighttime tropopause-overshooting storms, and do not consider convectively-modified tropopause characteristics (Bedka et al., 2010). Conversely, single-satellite IR-derived methods are more common, but several challenges hinder their wide-spread applicability. In addition to evidence suggesting that there is no direct relationship between IR brightness temperature and cloud-top height (Setvák et al., 2013), IR-based cloud-top height retrievals are subject to lower-than-VIS spatial resolution, variable cloud radiative properties, and inaccurate representations of cloud tops above the tropopause (due to continued cloud-top cooling or mixing). Even if the full environmental temperature profile and cloud emissivity were known, large deviations from thermodynamic equilibrium surrounding deep convection limit altitude accuracy (Hasler, 1981).

To minimize the limitations presented by single-satellite IR and VIS techniques, stereoscopy (stereo) leverages the viewing geometry of two satellites simultaneously observing the same region to retrieve cloud top heights. Early stereo analyses typically coupled geostationary and polar-orbiting satellite platforms and were thus constrained by short observing periods and coarse spatial resolution (Ondrejka and Conover, 1966; Kikuchi and Kasai, 1968; Shenk et al., 1975; Hasler, 1981). As more geostationary satellites became operational, particularly those tasked with viewing the same region, stereo cloud-top height retrievals leveraged the increased spatiotemporal resolution (Hasler et al., 1983; Wylie et al., 1998; Minzner et al., 1978; Seiz et al., 2007; Lee et al., 2020; Lee and Shin, 2021; Liu et al., 2021). The increased resolution motivated the development of automated techniques, as opposed to the manual binocular viewer

approach, for retrieving stereo cloud-top heights (Hasler et al., 1991). To date, the use of stereo to retrieve cloud top heights from observations of AACP-producing storms has been rather limited (Fujita, 1982; Mack et al., 1983).

While much of the early work on AACPs documented their satellite characteristics, recent efforts have focused on developing an understanding of mechanisms responsible for their formation. Numerous high-resolution numerical model simulations, complemented by observational analyses, have revealed that frequent gravity wave breaking borne out of strong storm-relative wind in the lower stratosphere in and near the overshooting top is responsible for AACP formation (Wang, 2003, 2007; Luderer et al., 2007; Wang et al., 2016; Homeyer et al., 2017; O'Neill et al., 2021). Earlier work from Wang (2007) identified two distinct AACP modes in numerical simulations that both formed as a result of gravity wave breaking. Recent very high-resolution simulations by O'Neill et al. (2021) emphasized that the continuous gravity wave breaking responsible for AACPs is accomplished through the establishment of a hydraulic jump downstream of the overshooting top. Although AACP formation through gravity wave breaking has been demonstrated repeatedly in existing work, the presence of two (or more) distinct AACP modes has not been routinely observed or documented in AACP studies.

Despite improved understanding of AACP formation and their significance to weather and climate, there are characteristics of AACPs that remain poorly understood. One such characteristic is that some AACPs exhibit typical features in VIS imagery while appearing cold (or colder) than the broader storm top (Figs. 3.1b & 3.1d). Heymsfield et al. (1983) attributed the warm region near the overshoot to subsidence from the descending portion of gravity wave breaking, which was also suggested by Fujita (1974, 1982), potentially indicating different formation mechanisms for warm and cold AACPs. Alternatively, other studies have suggested that these IR brightness temperature differences could instead be evidence of unique microphysics in AACPs.

Setvák and Doswell III (1991), Levizzani and Setvák (1996), Rosenfeld et al. (2008), Setvák et al. (2013), and Shou et al. (2019) utilized polar-orbiting satellite observations to show that AACPs often have unique microphysical signatures evidenced by higher radiance in shortwave IR imagery. However, it is exceedingly rare that these platforms encounter AACPs, as they do not sample over land during peak convection and overshooting periods. For example, Setvák et al. (2013) utilized observations from NASA's A-Train constellation to find global instances where AACPs were sampled and documented only five cases from 2006-2010. Rosenfeld et al. (2008) and Lindsey et al. (2006) leveraged geostationary satellite observations to assess cloud-top microphysical structures given their higher temporal sampling for a given spatial domain, but did not focus on AACPs. Prior to DCOTSS, in situ trace gas observations of AACPs had only recently been documented in Smith et al. (2017), highlighting elevated water vapor concentrations in the stratosphere from an aircraft observation during 2013 over the U.S. As such, these different microphysical characteristics could have implications for AACP temperature (e.g., due to differing optical thickness of AACPs).

Recently, Bedka et al. (2018b) acknowledged the greater complexity of AACP IR signatures, provided several examples, discussed potential explanations based on existing literature, and presented several hypotheses for variable AACP IR brightness temperature: 1) sedimentation of large ice crystals that reduces the cloud optical depth, allowing a colder tropospheric anvil beneath the AACP to dominate the radiative signal; 2) AACP injection into nearly isothermal UTLS environments or above-anvil layers that are cooling with height; 3) AACP subsidence into layers with colder temperature; and/or 4) cooling of the local UTLS temperature through AACP sublimation. These hypotheses from Bedka et al. (2018b), as well as potential alternative explanations for variable IR signatures, have not been evaluated to date.

Although AACP cloud-top heights would be valuable when investigating their unique IR signatures and depth of stratospheric hydration, it has historically been difficult to accurately estimate cloud heights given the aforementioned complexities of AACP appearance, frequent nocturnal overshooting, previous IR/VIS capabilities, and limited methods of validation. The majority of existing work developed stereoscopic algorithms for older NOAA/NASA Geostationary Operational Environmental Satellite (GOES) platforms (e.g., Wylie et al., 1998) and other new geostationary satellites (e.g., Merucci et al., 2016; Prata and Lynch, 2019; Lee et al., 2020; Lee and Shin, 2021; Liu et al., 2021). Only very recent work from Young (2021), built on similar framework established in (Hasler et al., 1991), has documented stereoscopic techniques capable of retrieving cloud-top heights from the newest generation of GOES platforms. Even with updated stereo techniques, the infrequency of very deep overshooting convection ($\geq$2 km into the stratosphere occurs a handful of times a year (Cooney et al., 2018)) combined with infrequent overlapping GOES–East and GOES–West (currently GOES-16 and GOES-17, respectively) domains results in exceedingly rare opportunities for their combined evaluation. However, two of the research flights during the DCOTSS field campaign sampled intense recent convection in a region of overlapping GOES-16 and GOES-17 mesoscale sectors: research flight 13 on 31 May 2022 (RF13) and research flight 19 on 24 June 2022 (RF19). These two flight in particular provide the unique, proper situation to analyze stereo-derived cloud top height retrievals of these intense convective events and leverage in situ observations for validation.

1.3 Study Objectives

In this section, research questions for all four studies are presented. The first two focus on hail identification and the second two on characteristics of overshooting storms.

Studies no. 1 & 3 have already been published in the refereed literature (Murillo et al., 2021; Murillo and Homeyer, 2022), while Study no. 2 is in final preparation for review and Study no. 4 summarizes ongoing preliminary research of field project observations.

Study no. 1 (Chapter 2.1):

When comparing the various methods for the objective of establishing a U.S. severe hail climatology with the greatest possible level of confidence, ground-based radar observations would appear to provide considerable strengths compared to other datasets that outweigh its weaknesses. Due to the extensive time period of data availability from the operational network in the U.S. and its ability to provide three-dimensional hydrometeor information of an existing storm in near real-time, we seek to quantify features of severe hail fall from radar using a markedly longer time period than previous studies to produce a climatology of 23 years. In particular, radar observations during the period 1995–2017 are used to create multiple MESH climatologies, which are compared to hail reports. In doing so, we seek to address three primary research questions: 1) How reliable is a radar-only hail climatology when compared to a report-based method? 2) What additional information (i.e., environmental variables) could improve accuracy of a radar climatology? and 3) What new information can we learn about severe hail fall over much of the contiguous U.S. (CONUS) from a long-term, well-calibrated, radar-based climatology? Potential severe hail occurrences are identified using radar observations, while environmental characteristics provide information regarding the likelihood that hail reaches the surface before melting.

Study no. 2 (Chapter 2.2):

With the increasingly extensive archive of CONUS Doppler radar data, higher temporal resolution reanalyses, and the more recent stability of hail reporting (Allen and

Tippett, 2015), it is now possible to identify severe hail-producing supercells in radar observations and comprehensively evaluate their environments based on their maximum reported hail size. As such, this study leverages hundreds of objectively-identified hail-producing supercell storms from 2010–2019 to answer the following research questions: 1) How useful are bulk environmental parameters for estimating the maximum hail size of supercells? and 2) Can any additional information be gained by leveraging full environmental profiles and storm-relative motions? To assess these questions, we group supercells into three maximum hail diameter categories based on the maximum diameter hail they produce: marginally severe (1–1.5-in.), significant severe (2–3-in.), and giant/gargantuan (≥4-in.). These hail size populations include a buffer between categories to account for uncertainty in the reported sizes used to define them. Profiles of the near-storm environment are extracted for a sample of each supercell population to conduct a detailed evaluation of the vertical distribution of buoyancy, humidity, wind shear, and their potential utility for estimating the maximum expected hail size. This work aims to both re-evaluate the utility of environmental proxy hail size indicators identified in previous studies from much smaller storm populations and reveal new insights that can be utilized in the development of future hail size algorithms.

Study no. 3 (Chapter 3.1):

In this study, we identify 89 warm and 89 cold AACPs from 1-minute GOES-16 satellite imagery, coupled with ground-based radar observations and environmental information from reanalysis to answer the following research questions: 1) Why do some AACPs exhibit a warm feature in IR imagery while others do not? and 2) What observable storm and environment differences exist between warm and cold AACPs? We have outlined 3 key hypotheses, which are illustrated in Fig. 3.2, and add contextual information to several of the points outlined in Bedka et al. (2018b).

For Hypothesis 1 (Fig. 3.2, top row), warm AACPs occur in environments with a single tropopause, reside in the lower stratosphere, and appear warm because the lower stratosphere environment is warmer than the tropopause below (this is consistent with warm AACP analyses from past studies). Cold AACPs for this hypothesis, however, occur in environments with a double tropopause (i.e., where temperature decreases significantly above the primary tropopause until eventually increasing again above the secondary tropopause), and reside in the layer between tropopauses where the lower stratosphere is colder than the primary tropopause. Double tropopauses occur most often during spring and early summer in the midlatitudes when AACPs are also frequently observed (Randel et al., 2007; Añel et al., 2008; Manney et al., 2017; Xian and Homeyer, 2019; Wilhelmsen et al., 2020).

Hypothesis 2 (Fig. 3.2, middle row) states that warm AACPs occur in UTLS environments with midlatitude characteristics, residing in the lower stratosphere, while cold AACPs occur in UTLS environments with tropical characteristics, residing in the upper troposphere where temperature is still decreasing with height. The wide variability in tropopause height over the North American midlatitudes, both zonally and meridionally (e.g., Hoinka, 1998; Li et al., 2017), suggests that such a difference could be responsible for the variable IR signatures so long as the depths of warm and cold AACP-producing storms do not differ substantially.

Lastly, Hypothesis 3 (Fig. 3.2, bottom row) indicates that warm and cold AACPs occur in similar UTLS environments, but differ microphysically. Warm AACPs in this case are optically thick, such that the satellite senses the temperature of the warm stratosphere environment (compared to the anvil and tropopause layer below). On the other hand, cold AACPs are optically thin in this hypothesis, such that the satellite is sensing temperatures consistent with the broader storm anvil below, rather than that characteristic of the AACP. Differences in apparent AACP translucence in

VIS imagery suggests that variability of AACP optical thickness and thus AACP IR brightness temperature could be a likely control.

Hypotheses 1 and 2 are explicitly tested in this study given the availability and utility of environmental reanalyses and ground-based radar observations in diagnosing tropopause structures and storm-relative characteristics. Thorough evaluations of Hypotheses 1 and 2, devoid of explicit Hypothesis 3 assessment due to the lack of available microphysical data for these AACPs, could provide sufficient evidence in favor of one explanation; if such results indicate neither Hypothesis 1 nor 2 are supported, then Hypothesis 3 or one of the more nuanced processes outlined in Bedka et al. (2018b) would prove the most likely alternative explanation.

Study no. 4 (Chapter 3.2):

Coordinated operations of GOES-16 and -17 mesoscale sectors during designated DCOTSS research flights provide a unique opportunity for detailed cloud-top height retrieval validation. This study leverages RF13, which took place on 31 May 2022, and the recent stereoscopic cloud-top height retrieval algorithm designed for modern GOES VIS imagery (Young, 2021) to answer the following research questions: 1) How do stereoscopic cloud-top height retrievals compare to radar-derived echo top products and in situ water vapor measurements? and 2) What new information can be gained from stereoscopic cloud-top height retrievals regarding AACPs? We focus on RF13 due to the flight track's proximity to recent AACP production (discussed further in Section 3.2.7). Given that the results presented here contain preliminary in situ observations with ongoing analyses, our results and conclusions are subject to change upon final analysis.

Chapter 2

Hail Occurrence

2.1 Study no. 1: A 23-Year Severe Hail Climatology using GridRad MESH Observations

As mentioned in Section 1.3, this study aims to answer three main research questions: 1) How reliable is a radar-only hail climatology when compared to a report-based method? 2) What additional information (i.e., environmental variables) could improve accuracy of a radar climatology? and 3) What new information can we learn about severe hail fall over much of the CONUS from a long-term, well-calibrated, radar-based climatology?

2.1.1 Radar observations

To identify potential hail events, this study employs hourly, four-dimensional space-time composite NEXRAD data known as GridRad Version 3.1 (Bowman and Homeyer, 2017) that are available over most of the CONUS (the west coast is not included, which rarely experiences severe hail), spanning the time period 1995–2017. Spatial and time weighted binning is used to merge individual NEXRAD scans in the common GridRad volume (described in greater detail in Homeyer and Bowman (2017) and references therein). These weights are such that shorter ranges and temporal proximity to the radar are prioritized, leading to retaining of data with the highest quality and

resolution. Volumes are available on a ∼2-km horizontal resolution and 1-km vertical resolution grid. The same or similar GridRad data has been used for hail detection in several recent studies (e.g., Murillo and Homeyer, 2019; Tang et al., 2019; Jeong et al., 2020).

Quality control is applied at several steps from the native GridRad volumes. One of the more substantial filters focuses on the removal of low-confidence echoes consistent with the recommendations outlined in Homeyer and Bowman (2017). These data quality measures ensure that GridRad data retained for analysis contain high-confidence observations of storms. In addition, manual identification of failed radar scans from individual radars (and consequently, erroneous or nonphysical echoes merged into the GridRad volumes) was carried out for the entire record to exclude such data from analyses in this study. Instances of uniform or linearly increasing radar reflectivity in range from a given radar are egregious examples of such erroneous scans. Subjective evaluation of maximum annual MESH was often the best indication of these events, which are more common in the earlier part of the record. In total, 200,694 hourly GridRad volumes are retained for analysis after completing these quality control measures (930 out of 201,624 volumes had erroneous scans that could impact MESH analysis, 78% of which were found in years prior to 2006).

MESH is a power law relationship of observed hail sizes to the radar reflectivity-based severe hail index (SHI), both of which are defined in Witt et al. (1998a). The original MESH relationship (hereafter $\text{MESH}_{\text{Witt}}$) was empirically fit to the 75^{th} percentile of 147 hail reports and has been broadly used in prior studies. However, Murillo and Homeyer (2019) refit the power law relationship between SHI and MESH to the 75^{th} and 95^{th} percentiles of observed hail size from nearly 6,000 hail reports using GridRad data (hereafter MESH_{75} and MESH_{95}, respectively). MESH_{75} and MESH_{95}

are more sensitive to lower SHI values, which results in overall smaller SHI values satisfying a 1-in size threshold (Murillo and Homeyer, 2019). In this study, we generate a hail climatology using all three MESH configurations to test their performance with respect to hail reports. Based on the results of Murillo and Homeyer (2019), we expect that $\text{MESH}_{\text{Witt}}$ might underestimate hail occurrence in comparison to the new configurations because it was not calibrated using smoothed radar observations and was developed using a small sample of 147 hail reports, which were geographically confined to Oklahoma and Florida. MESH calculations using any of these fits may be affected by radar reflectivity biases from radar miscalibration, which have been documented to commonly reach 1 dB for NEXRAD radars (OFCM, 2005) and may lead to an overestimate of MESH by up to 20% (Warren et al., 2020).

2.1.2 Environmental data

Numerical model analyses provide environmental quantities for radar metric calculations and statistical analyses carried out in this study, for which we employ the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2; Gelaro et al., 2017) for the 23-year period. MERRA-2 is available from 1979 to present with a 3-hourly temporal resolution, a 0.5° x 0.625° longitude-latitude resolution, and with 72 vertical model levels with a model top of 0.01 hPa. Environmental parameters used for analysis include the altitudes of multiple isotherms, relative humidity (RH), low- to mid-tropospheric temperature lapse rates, vertical wind shear, and precipitable water (PWAT). It is also worth noting that we tested the sensitivity of isotherm altitudes used in MESH calculations to different reanalysis temperature fields and found that MESH was insensitive to these changes.

2.1.3 Hail reports

The SPC Severe Weather Database (available at http://www.spc.noaa.gov/wcm/) provides the time, location, and size of severe hail reports in the U.S. (NOAA/NWS/SPC, 1955). Though the severe hail criteria changed in 2010 from a diameter of $\geq$ 19.05 mm [0.75 in.] to $\geq$ 25.4 mm [1 in.], comparisons between between the two severe hail day frequencies showed small differences in spatiotemporal distributions (Allen and Tippett, 2015). Therefore, we utilize the current severe hail criteria of a diameter $\geq$ 25.4 mm [1 in.] for the full time period evaluated in this study. The criteria for significant severe hail for this period remained consistent and is defined as a diameter of $\geq$ 50.8 mm [2.0 in.] In total, 272,921 reports were used in this study.

2.1.4 Data synthesis and analysis techniques

All datasets are analyzed on an 80 x 80 km grid with 1-sigma Gaussian smoothing applied spatially. The resolution was chosen based on the lowest resolution dataset, the hail report database, following the same procedure as Brooks et al. (2003) and Doswell et al. (2005). This allows for spatial inhomogeneities of the population biased reports to be somewhat smoothed out for a climatological evaluation. This methodology allows for all data to be evaluated on a common grid, while appropriately utilizing the report observations. We then structure all data such that days are evaluated for the 24-hr period from 12 UTC through 11 UTC to be consistent with the SPC definition of a severe weather day, allowing for direct comparison between MESH and report distributions. For analysis, we retain the climatological maximum MESH size, and counts of severe and significant severe hail days, along with the hour of each severe and significant severe hail occurrence. These distributions are calculated over the full GridRad domain with reports and each MESH configuration to produce yearly and

seasonal distributions. A report hail day is defined as at least one severe or significant severe hail report in a grid box during a 24-hr SPC severe weather day. A MESH hail day is only defined when there is at least one severe or significant severe MESH observation in a grid box during a 24-hr SPC severe weather day, and GridRad data is available for at least 20 of those hours to ensure that the comparison between radar and reports is fair. MESH days identified over the ocean are discarded due to reduced radar coverage and lack of validation (report) data. To be consistent with the radar analysis, report data analyzed was limited to ± 10 minutes from each GridRad analysis time to account for potential/expected errors in report timing. There are two types of thresholds that we use to define hail days for MESH: 1) conventional thresholds for severe and significant severe hail days, and 2) threshold values corresponding to the peak critical success index (CSI) for detecting severe and significant severe hail-producing storms, as found in Murillo and Homeyer (2019) (see their Figs. 9a and 11a). This second method removes sensitivity of the result to the MESH relationship, essentially relaxing the identification to the integrated radar reflectivity (i.e., SHI value) - the resulting climatologies should be insensitive to the MESH fit used. Table 2.1 lists the MESH thresholds for all methods used in this study, for which a hail event is identified as a MESH value that reaches or exceeds the given threshold.

To address the second research question, we evaluate the ability of several environmental parameters in correctly distinguishing identified and false alarm hail days (see Section 2.1.5 for the motivation to incorporate environmental characteristics). We focus on environmental discriminators that could produce storms with similar MESH values to those observed during hail events but dissimilar hail likelihood due to increased opportunity for melting before reaching the ground. The parameters we test are as follows: melting level (0°C) altitude, low-level (altitudes below the melting level) RH, low-level (altitudes below the melting level) lapse rate, 2–4 km lapse rate, PWAT,

and 0–6 km bulk wind shear. To do so, we first group hail events into two categories: small hail (sizes of at least 0.75 in. and less than 1.5 in.) and large hail (sizes of at least 1.5 in.). Each of these two categories are further divided into two populations: 1) the "correct" hail identification group, defined at a grid box level as observing both a MESH value and a hail report within the given size range within 10 minutes of each hourly GridRad analysis and 2) the "false alarm" group, defined as observing a MESH value within the given size range but no corresponding hail report. These size categories were chosen because false alarms are considered most likely in the lower range (given their relatively slow fall speeds – for hail – and as a result, greater times spent in the warm cloud layer (Heymsfield et al., 2014)). Note that due to the aforementioned limitations of hail reports, these populations are not expected to be solely comprised of true hits or false alarms, but it is the best available approach for statistical evaluation. The correct and false alarm populations of MESH events are then used to: 1) identify environmental factors that discriminate best between correct identifications and false alarms, and 2) use the leading discriminants as input for Fisher's Linear Discriminant Analysis (LDA) for two groups (Wilks, 2006). Briefly, Fischer's LDA solves for a function (or two-dimensional boundary in discriminant space) that best separates

Table 2.1: MESH thresholds for the various hail day analyses.

Hail Day Criteria Type	$MESH_{Witt}$ mm [in.]	$MESH_{75}$ mm [in.]	$MESH_{95}$ mm [in.]
Conventional Severe	25.4 [1.00]	25.4 [1.00]	25.4 [1.00]
Max CSI Adjusted Severe	35.56 [1.14]	41.91 [1.65]	63.25 [2.49]
Conventional Sig. Severe	50.8 [2.00]	50.8 [2.00]	50.8 [2.00]
Max CSI Adjusted Sig. Severe	45.72 [1.80]	50.55 [1.99]	76.71 [3.02]

potential hail events into the two target populations (correct and false alarm) using linear combinations of the corresponding environmental quantities.

2.1.5 Results: Report and radar-based climatology

Though hail size comparisons can be useful for identifying large hail risk areas, there are several other limitations when diagnosing hail size extrema from radar or hail report observations. In addition to the reporting biases previously mentioned in Section 1.1, the rarity of large hail, coupled with the an unknown hail size distribution produced from a storm, pose unique challenges when interpreting hail size extrema (Cheng et al., 1985; Bardsley, 1990; Fraile et al., 1992; Blair and Leighton, 2012; Allen and Tippett, 2015; Blair et al., 2017; Allen et al., 2017). The sensitivity to errors in radar observations, though minimized by the extensive quality control, could still result in an overestimation or underestimation of the maximum hail size. Given these limitations, hail magnitude comparisons are focused on assessing broad spatial maximums, rather than event specific hail sizes. The climatology from reports indicates that the largest hail sizes occur over the Central Plains (Fig. 2.1), as routinely documented in previous work. Using the 23-year GridRad dataset, $\text{MESH}_{\text{Witt}}$ broadly underestimates the reported hail sizes, which are known to generally under-represent true sizes (Fig. 2.2a). MESH_{75} better represents reported hail sizes in most areas excluding the Central Plains, where it likely underestimates hail size with respect to reports (Fig. 2.2b). MESH_{95} provides more reliable estimates over the Central Plains, but generally is calibrated slightly high otherwise, with more extreme overestimates across the Ohio Valley and in southern Florida (Fig. 2.2c).

Rather than evaluating individual hail reports, we analyze the frequency of hail days as suggested in Allen and Tippett (2015), which reduces the sensitivity to report errors, mitigates duplicate reporting, and is more relevant to impacts and losses. There

is a maximum in severe hail report days per year over the Central Plains (Fig. 2.3). When using the severe threshold corresponding to the maximum skill (Table 2.1), we see similar distributions and frequencies for all MESH identifications, which broadly agree with the reports (Fig. 2.4). However, as mentioned above, hail reports likely underestimate true hail occurrence, which indicates that this approach is also likely underestimating severe hail frequency. Thus, a more appropriate approach for capturing the majority of severe hail events and identifying differences in hail climatology due to varying MESH configuration is to apply the conventional threshold (i.e., diameter $\geq$ 25.4 mm [1 in.]) to each (Fig. 2.5). This conventional threshold approach reveals considerable differences in severe hail frequency between the MESH configurations. The conventional $\text{MESH}_{\text{Witt}}$ hail day climatology (Fig. 2.5a) shows similar geographic distribution and magnitude to that seen in the report climatology, with slightly higher frequencies than found in the climatology using the maximum skill threshold. The conventional MESH_{75} and MESH_{95} hail day climatologies result in 2–4 times more hail days per year than the corresponding max skill climatologies, with peak occurrence found over the Central Plains and SEUS (Fig. 2.5b & c).

The considerable differences seen between $\text{MESH}_{\text{Witt}}$ and new MESH configurations in the SEUS is likely due to the reduced slopes of the new power law fits. Ultimately, the changes in the new MESH fits result in a greater sensitivity to lower SHI values and smaller sizes than $\text{MESH}_{\text{Witt}}$, resulting in more severe hail days as the 25.4-mm [1-in.] diameter conventional severe hail threshold is more easily reached. When evaluating seasonality (Fig. 2.6), we see the largest discrepancy between the reports and MESH_{75}/MESH_{95} in the SEUS occurs during the summer (JJA). This large (and presumably false) signal in the SEUS is likely a result of deep, tropical summer storms that achieve similarly high Z_{H} aloft as compared continental hailstorms. However, these storms typically lack the necessary thermodynamic and dynamic conditions for

severe hail production and/or maintenance. From a radar perspective, storms with a large concentration of small, sub-severe hail or a small concentration of large hail can produce similarly high radar reflectivity but unique dual-polarization scattering characteristics, which is impossible to assess using MESH-only identification approaches (e.g., see emerging work on storms with large accumulations of small hail; Kumjian et al., 2019). In the SEUS, many storms occur in a dynamic regime that is not supportive for severe hail growth but may be capable of producing small hail aloft. Due to the large warm layer depth (or similarly, high melting level height) in the SEUS during summer, these small hail stones melt before reaching the surface. Thus, leveraging environmental information to establish confidence in a diagnosed MESH severe hail day would help to limit impacts from non-severe hail producing storms in the analysis, in the SEUS and elsewhere.

2.1.6 Results: Radar climatology with parameter filtering

The results summarized in Section 2.1.5 highlight that a severe hail climatology based on MESH alone does not adequately represent (or agree with) the frequency and size of severe hail as diagnosed from reports. Here, we address the second research question: What additional information could improve accuracy of radar climatology? Namely, we evaluate the utility of incorporating certain environmental quantities into severe hail diagnosis that have been previously identified as relevant to hail potential: melting level altitude, low-level RH, low-level lapse rate, 2–4 km lapse rate, PWAT, and 0–6 km bulk wind shear (Johnson and Sugden, 2014; Allen et al., 2015a, 2020). To do this, we considered two criteria: 1) separation between the correct identifications

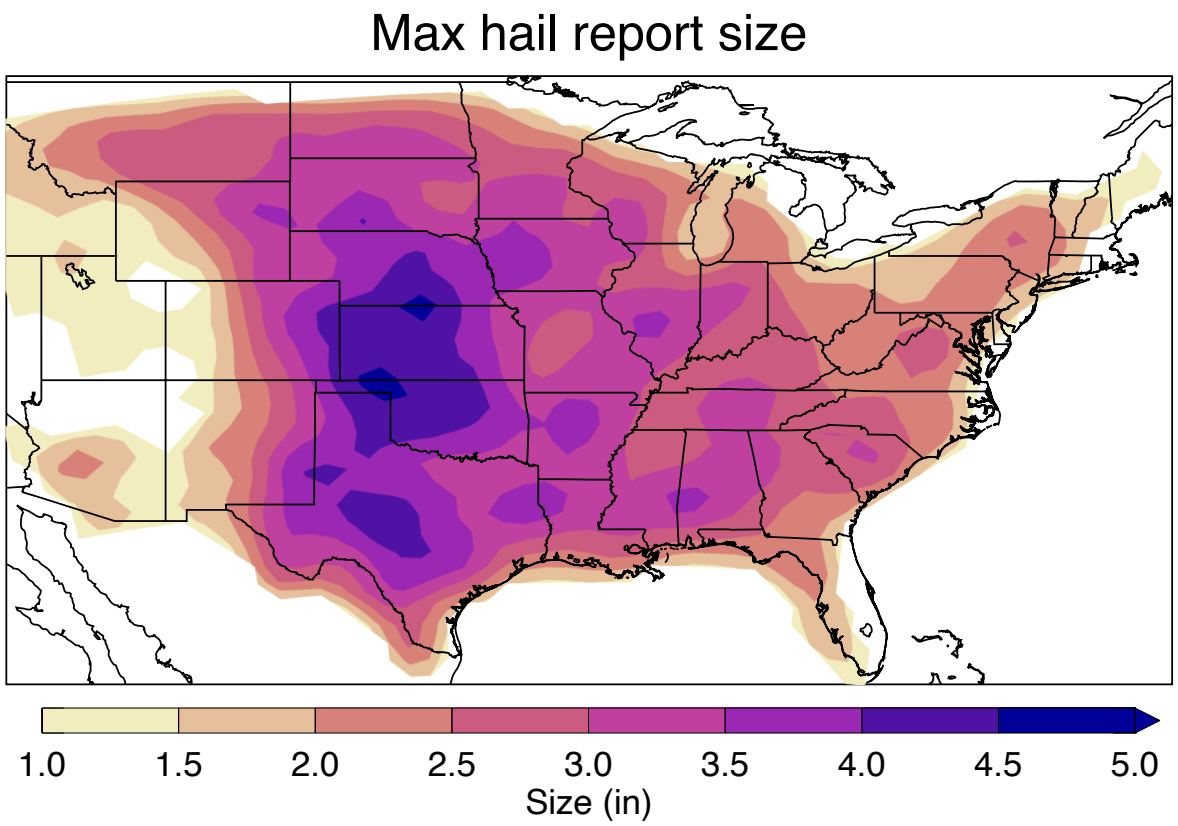

Figure 2.1: The maximum reported hail size over the full domain.

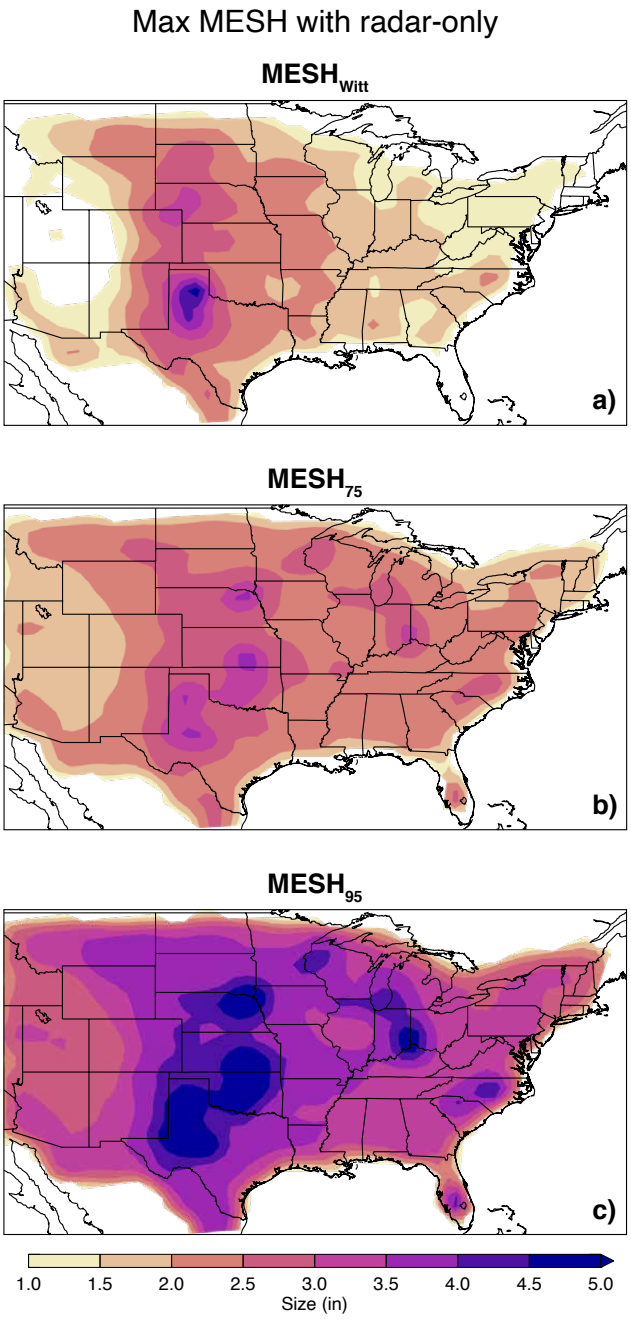

Figure 2.2: The maximum MESH value computed with a) $\text{MESH}_{\text{Witt}}$, b) MESH_{75}, and c) MESH_{95} over the full domain.

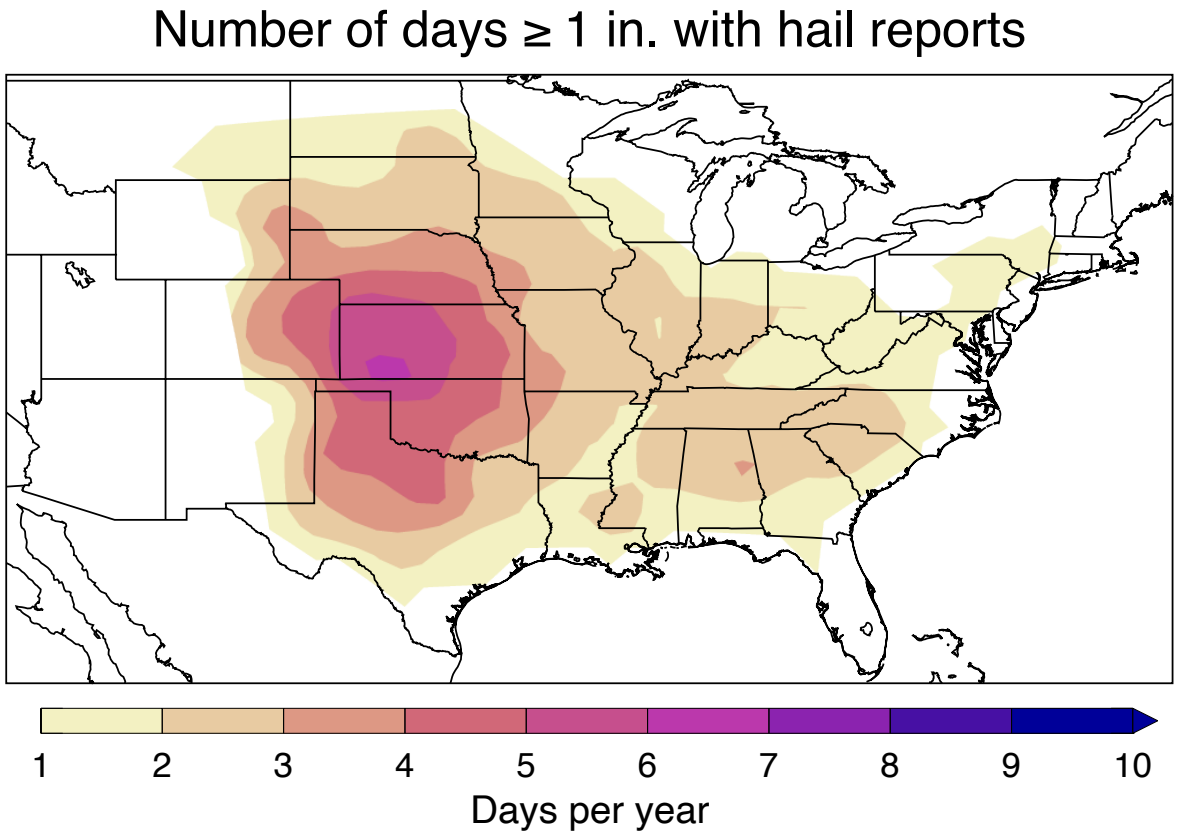

Figure 2.3: The number of reported severe (diameter $\geq$ 25.4 mm [1 in.]) hail days per year.

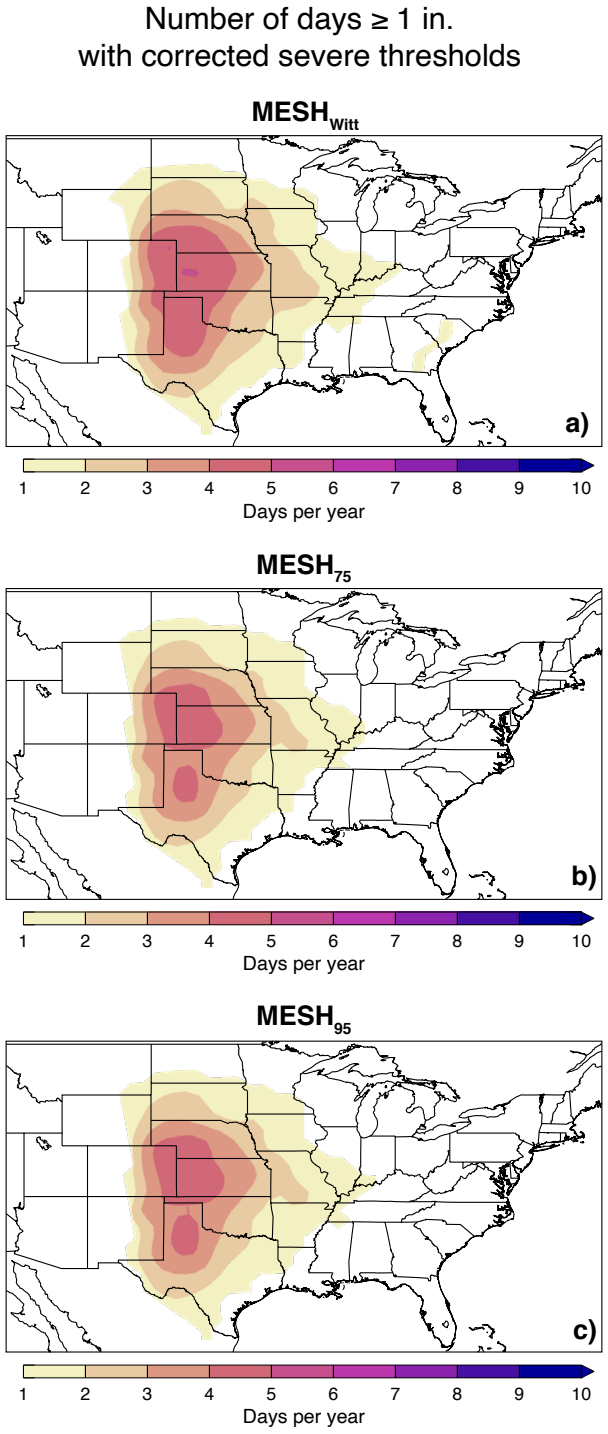

Figure 2.4: The number of severe hail days per year as defined by the corrected threshold sizes for a) $\text{MESH}_{\text{Witt}}$ (diameter $\geq$ 35.56 mm [1.14 in.]), b) MESH_{75} (diameter $\geq$ 41.91 mm [1.65 in.]), and c) MESH_{95} (diameter $\geq$ 63.25 mm [2.49 in.].

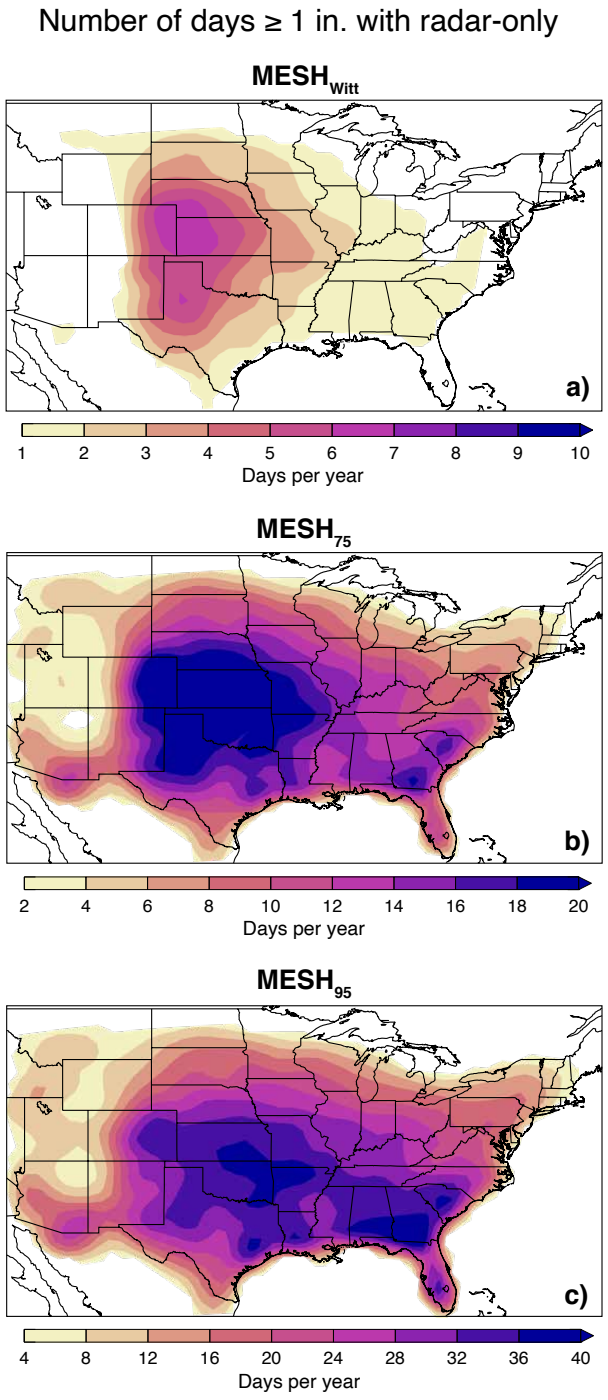

Figure 2.5: The number of severe hail days per year using the conventional threshold (diameter $\geq$ 25.4 mm [1 in.]) for a) $\text{MESH}_{\text{Witt}}$, b) MESH_{75}, and c) MESH_{95}. Note that the range in b) and c) are 2x and 4x as large, respectively, as that seen in a).

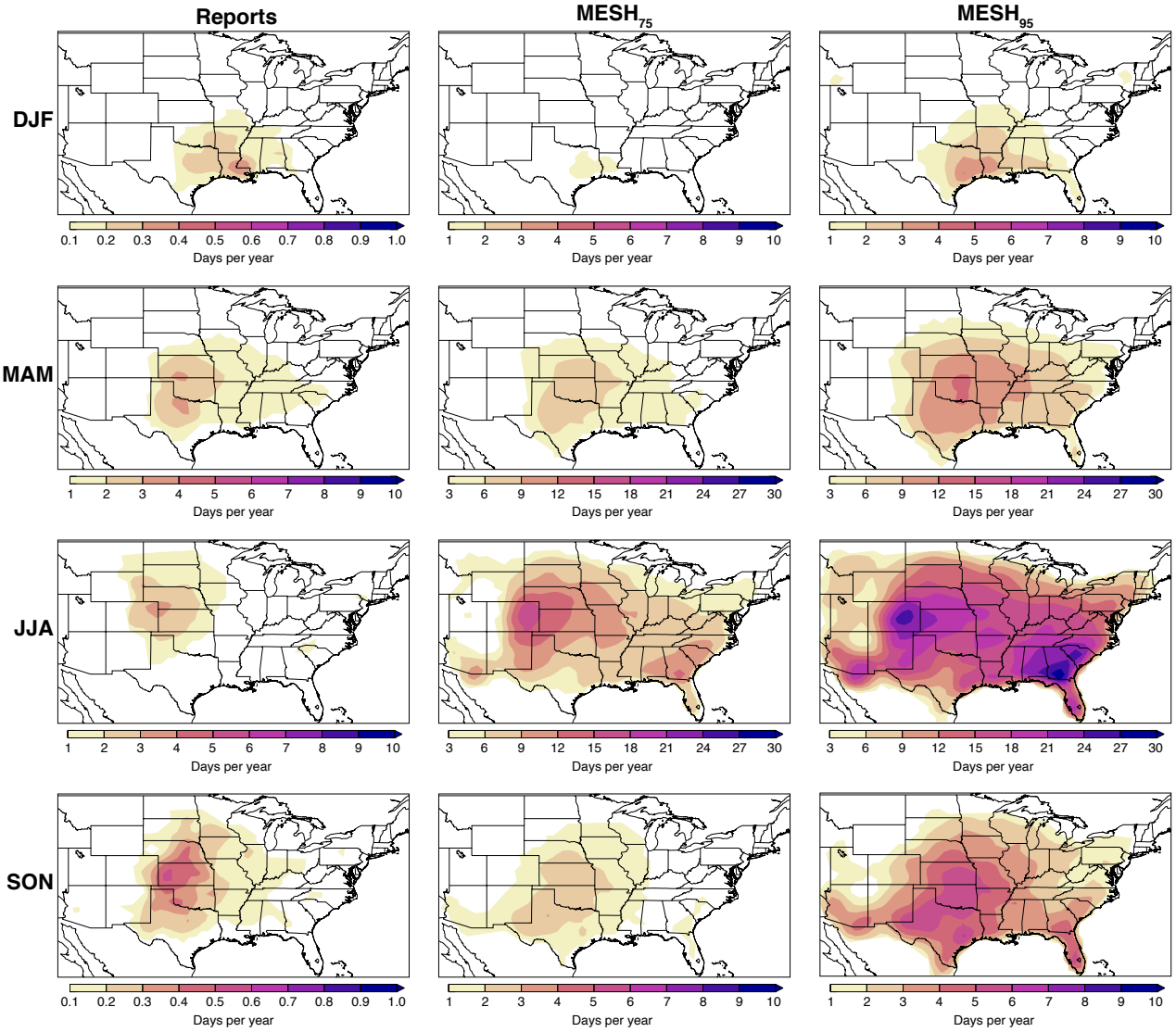

Figure 2.6: The number of severe hail days per year using the conventional threshold (diameter $\geq$ 25.4 mm [1 in.]) during each season for reports (left column), MESH_{75} (middle column), and MESH_{95} (right column). Seasons are represented by rows and labelled with each season's month abbreviations. Note that the range is different depending on the season and type of hail identification.

and "false" alarms in a 2-dimensional parameter space and 2) sensitivity to diurnal variability. The environmental parameters that show the greatest separation between the two populations and little sensitivity to the diurnal cycle were PWAT and 0–6 bulk wind shear (Fig. 2.7d). Other parameter combinations were tested (Fig. 2.7, a–c and others not shown) but either displayed larger overlap between the two populations or were sensitive to diurnal variations and produced vastly different peak timing of hail frequency.

Performing the LDA using PWAT and 0–6 bulk wind shear produced greater separation between the two populations, further indicating "false" alarms tend to occur in regions with higher PWAT and lower 0–6 bulk wind shear. We note that the two populations seen in both the pre- and post- LDA distributions are not expected to be completely composed of true hits or false alarms due to the numerous report limitations mentioned above. However, utilizing the large sample size and LDA provide a higher confidence in these distributions than that possible using radar data alone.

Applying LDA environmental filtering to the MESH climatologies leads to minimal differences in the maximum sizes. $\mathrm{MESH_{Witt}}$ is nearly equivalent to the radar-only climatology (not shown). The large maximum for southern Florida in $\mathrm{MESH_{75}}$ and $\mathrm{MESH_{95}}$ in the radar-only climatology are suppressed after environmental filtering (not shown). Larger climatological differences are seen in the frequency analysis, with an environmentally filtered $\mathrm{MESH_{Witt}}$ producing significant underestimates in hail days compared to reports (Fig. 2.8a). $\mathrm{MESH_{75}}$ and $\mathrm{MESH_{95}}$ severe hail day frequencies in the corrected climatology better represent the overall spatial distribution of severe hail days, with the previous presumably false maximum in the SEUS now absent (Fig. 2.8b & c), a result that was also found in Cintineo et al. (2012). Additionally, the resulting magnitudes are slightly higher than the reports, which is expected and is likely more representative of true hail days given the previously mentioned report

biases. When evaluating hail days by season (Fig. 2.9), we see the expected Central Plains maximum during the spring and summer, which is consistent with the report climatology. $MESH_{95}$ better captures extremes, especially in the cool season and for regions with non-zero but rare hail occurrence (e.g. Arizona).

Using this filtered hail climatology, we compare $MESH_{75}$ and $MESH_{95}$ to the original report distribution to assess the overall spatiotemporal variability and address our third research question: What new information can we learn about U.S. hail fall from a long-term, well-calibrated, radar-based climatology? Apart from the previously mentioned differences in frequency between the annual filtered MESH and report distributions, there are notable differences in locations of frequency maxima both in the annual and seasonal geographic distributions (Figs. 2.3 & 2.8, b–c). First, the SEUS frequency maximum (stretching mostly in the east-west direction from Northeast Mississippi to the western portions of the Carolinas) is only present in the report climatology, which indicates possible over-reporting in this region noted in several previous studies (Cintineo et al., 2012; Allen et al., 2015b; Allen and Tippett, 2015). This high bias could be due to various factors such as the high population density in the SEUS, resulting in a higher likelihood of severe hail being reported compared to adjacent locations. Non-stationarities identified by Allen and Tippett (2015) in this region in the spatial frequency of reporting also suggest that this region may be subject to a greater fraction of reports being driven by warning verification. Perceptual biases as to the size of SEUS hail may also be a factor, because this region infrequently receives hail much larger than 25.4 mm [1 in.]. Second, the Central Plains frequency maximum extends slightly further into SW Texas and E Colorado in the filtered MESH than the report climatology indicated. These shifts could be a result of better capturing hail occurrence in regions with low population density. The independence of the filtered radar-based climatology

of this bias indicates greater severe hail occurrence than previously estimated via radar methods.

The filtered seasonal hail climatology seems to indicate that filtered MESH_{75} produces a more realistic climatological frequency distribution, while filtered MESH_{95} generates frequencies closer to a worst case scenario (Fig. 2.9). There is also a clear seasonal shift in the peak frequency of hail. The winter is characterized by predominantly SEUS events, while the rest of the year sees a shift towards different regions of the Central Plains. There are also frequency maxima in Idaho and upstate New York during winter when using MESH_{95}, which are mostly a result of false detections of a few heavy winter precipitation events. The higher sensitivity to lower hail sizes in MESH_{95}, combined with bright band contamination during these winter events, likely contributes to the false signal over these regions.

To briefly evaluate the yearly changes of the multiple MESH configurations, we accumulate all the grid points with severe hail days for each year (Fig. 2.10). Data were quite sparse prior to 2000 and showed increasing trends in frequency until 2006, after which data availability is roughly constant from the NEXRAD network. From that point forward, natural variability dominates, and frequency on the national scale appears to be fairly stable, consistent with findings from Allen and Tippett (2015). The relative severe hail day frequency between each unfiltered and filtered MESH also stays rather constant, indicating that the filtering process works uniformly. When assessing each MESH configuration, we find that $\text{MESH}_{\text{Witt}}$ is the least useful, since it is insensitive to filtering, does not adequately demonstrate observed spatial patterns, and underestimates the frequency of severe hail days. Conversely, both MESH_{95} configurations likely overestimate the frequency, given the considerable increase from that seen in reports. MESH_{75} is likely the most useful for a climatological approach, as it displays a moderate increase in frequency with respect to reports and demonstrates

reasonably consistent interannual variability (after 2006). Because of the lower radar data quality/completeness before 2006, frequency analyses (Figs. 2.8-2.9) likely under-represent the overall true frequency of hail events diagnosed by radar over the 23-year period. However, restricting MESH analyses to 2006–2017 where data quality are much more consistent reveals consistent spatial distributions compared to the full period (not shown).

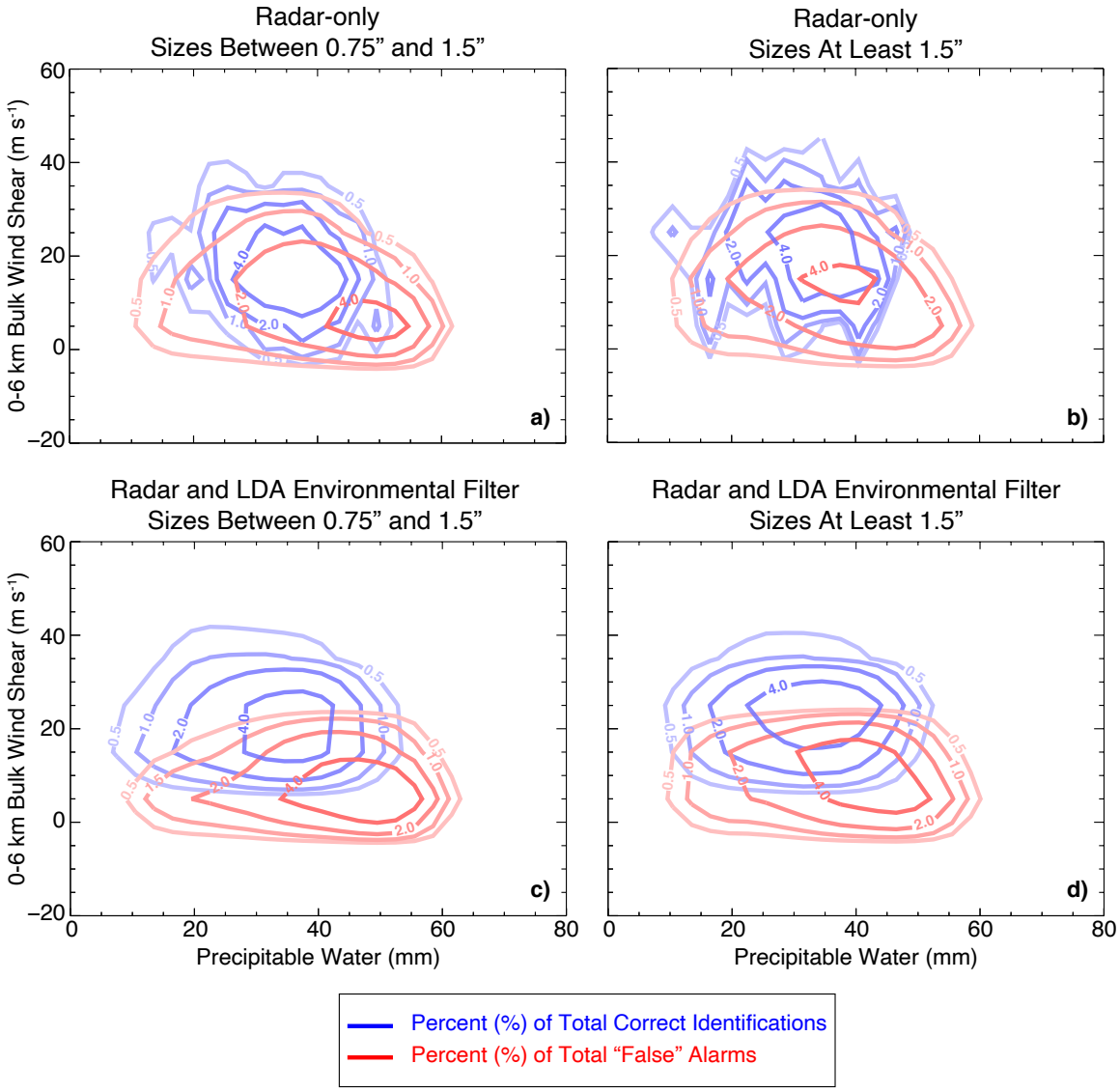

Figure 2.7: 2-dimensional joint frequency distributions of correct identifications (blue) and "false" alarms (red) for precipitable water and 0–6 bulk wind shear. Populations are split into two size groups: a) and c) are limited to sizes between 0.75 and 1.5 in., while b) and d) are limited to sizes greater than 1.5 in. Populations in a)–b) are generated through report verification, and those in c)–d) are generated through LDA environmental filtering. The contour values represent the percentage of all correct identifications (blue) or "false" alarms (red) that exist within the parameter space.

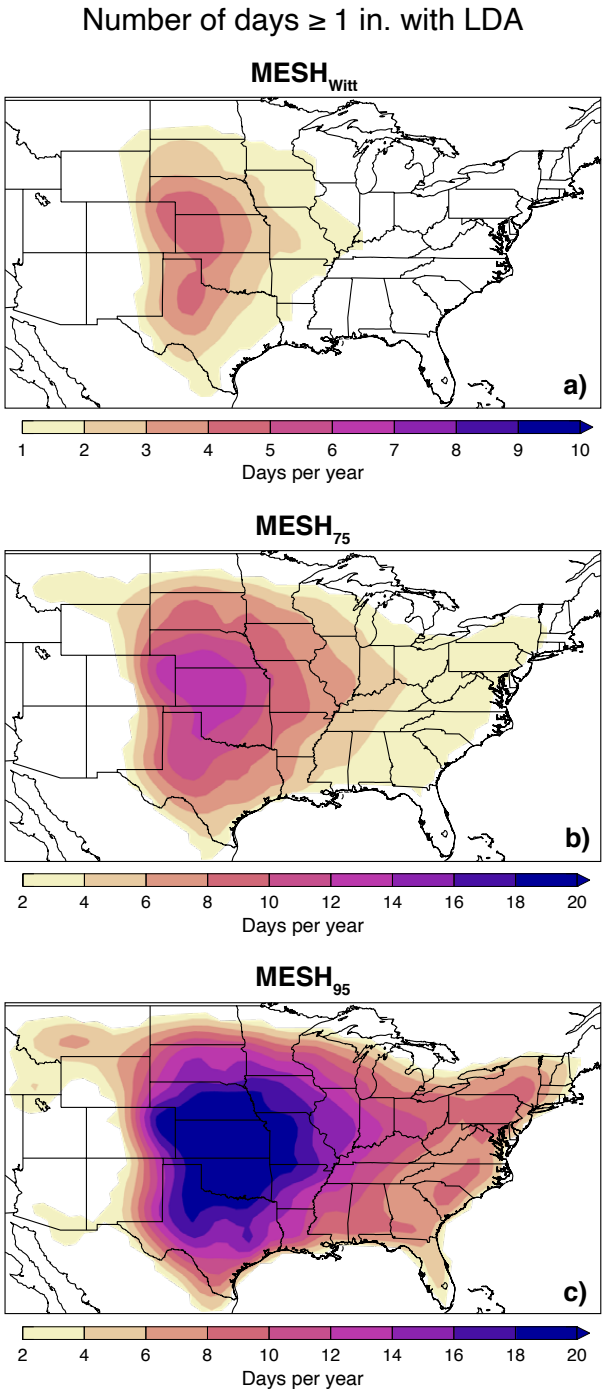

Figure 2.8: The number of severe hail days per year, after environmental filtering, using the conventional threshold (diameter $\geq$ 25.4 mm [1 in.]) for a) $\text{MESH}_{\text{Witt}}$, b) MESH_{75}, and c) MESH_{95}. Note that the range in b) and c) are 2x and 4x as large, respectively, as that seen in a).

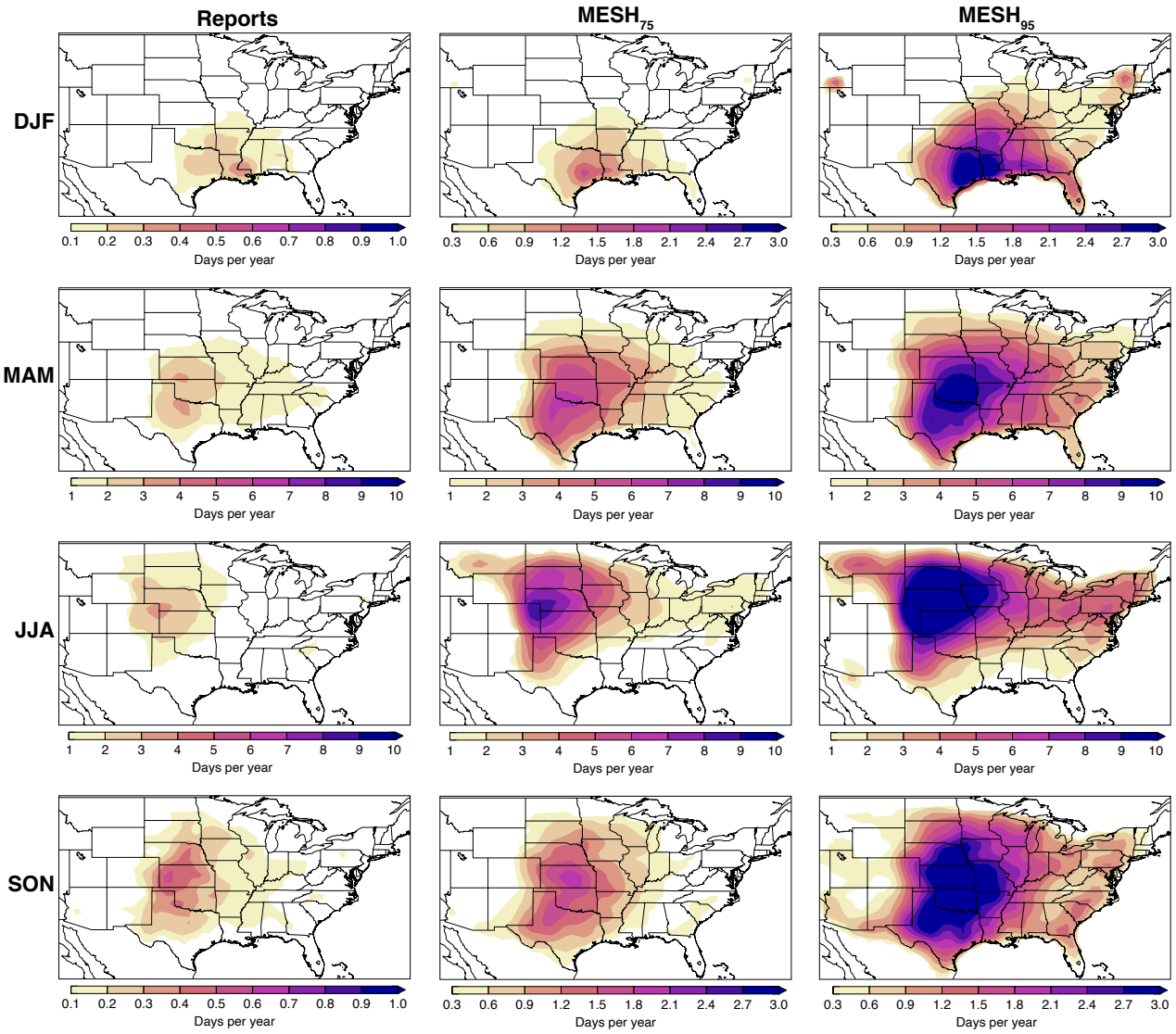

Figure 2.9: The number of severe hail days per year, after environmental filtering, using the conventional threshold (diameter $\geq$ 25.4 mm [1 in.]) during each season for reports (left column), MESH_{75} (middle column), and MESH_{95} (right column). Seasons are represented by rows and labelled with each season's month abbreviations. Note that the range is different depending on the season and type of hail identification.

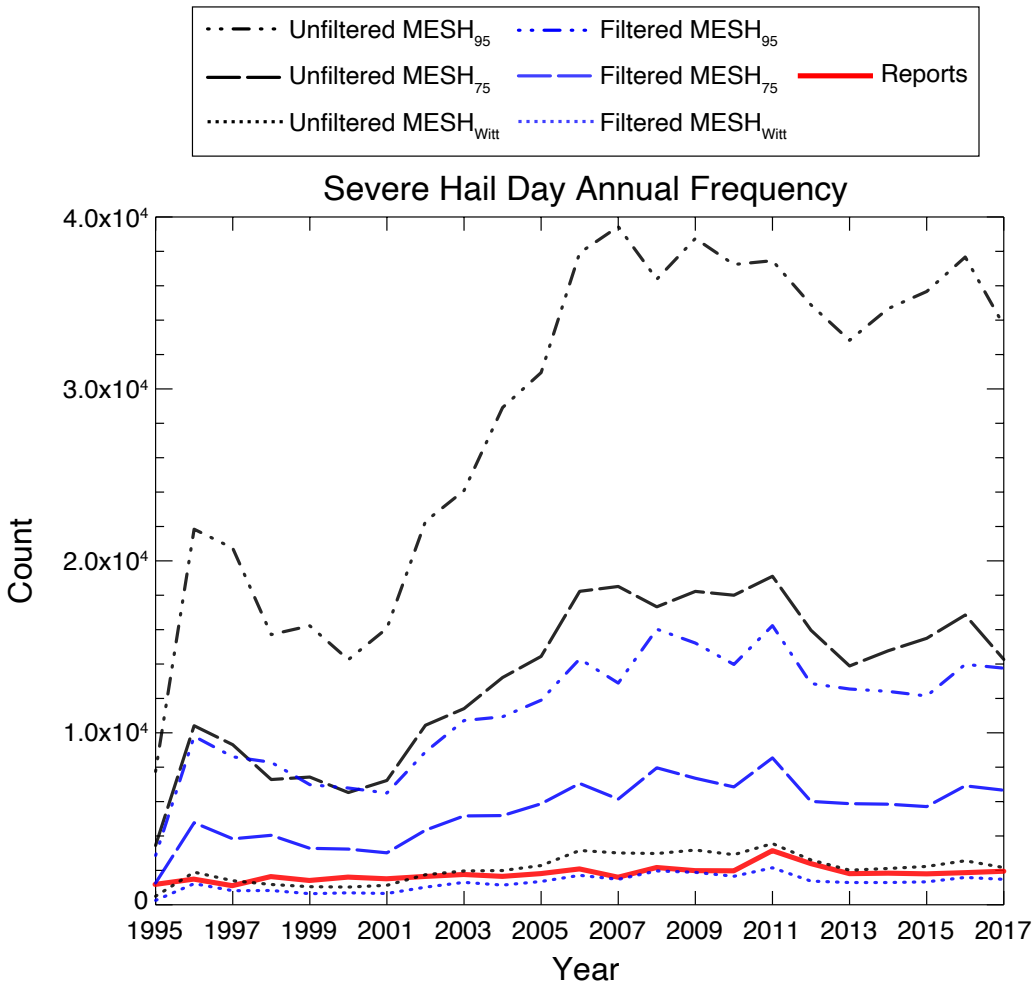

Figure 2.10: Annual total severe hail day grid boxes using the conventional threshold (diameter $\geq$ 25.4 mm [1 in.]) over the full spatial domain for each year used in the study. The black (blue) axes on the left (right) correspond to the black (blue) curves.

2.2 Study no. 2: A 10-Year Proximity Sounding Analysis of Severe Hail-Producing Supercells

The second hail identification study addresses two research questions, as listed in Section 1.3: 1) How useful are bulk environmental parameters for estimating the maximum hail size of supercells? and 2) Can any additional information be gained by leveraging full environmental profiles and storm-relative motions?

2.2.1 Radar observations

To first identify supercells, we leverage the recently created GridRad-Severe dataset (School of Meteorology/University of Oklahoma, 2021). GridRad-Severe was curated using GridRad Version 4.2 methods (Homeyer and Bowman, 2022) to develop a 10-year record (2010–2019) of data for ∼100 of the most severe events (12 UTC – 12 UTC days) per year in the CONUS (Murphy et al., 2022). The longitude, latitude, and time bounds of GridRad-Severe events vary to encompass ∼90% of tornado, severe hail, and/or severe wind reports during the event day. Radar observations feature a vertical resolution of 0.5 km for altitudes below 7 km above mean sea level (AMSL) and 1 km for altitudes between 7 and 22 km, as well as a horizontal grid with ∼0.02° × ∼0.02° longitude-latitude resolution. The 5-min radar volumes are then used to generate objective storm tracks from a $Z_{\mathrm{H}} = 30$ dBZ echo-top tracking algorithm (following Homeyer et al. (2017), with revisions as in Lagerquist et al. (2020)).

Table 2.2: The number of hail-producing supercell storms identified and analyzed in this study, sorted by maximum reported hail size linked with each storm.

Storm Category	**No. Storms Identified**	**No. Storms Used**
Marginally Severe (1–1.5 in.)	3262	235
Significant Severe (2–3 in.)	882	235
Giant/Gargantuan (4+ in.)	235	235

2.2.2 Environmental data

ERA5 is the fifth and most recent generation of the global European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis and is available from 1979 to present (Hersbach et al., 2020). ERA5 features 137 vertical hybrid sigma/pressure (model) levels from the surface up to 0.01 hPa, with approximately ∼20 m resolution near the ground and 400–500 m resolution in the UTLS. ERA5 has 0.28125° native horizontal resolution, but here the 1° horizontal resolution data are utilized. We use vertical profiles of temperature, pressure, humidity, and winds.

2.2.3 Hail reports

The National Centers for Environmental Information hosts the storm event database (SED) which encompasses the time and/or duration, location, magnitude, and reporting source for all U.S. severe weather reports (NCEI/NOAA, 2022). A severe weather event, following the NOAA Storm Prediction Center (SPC) definition, is defined as wind gusts $\geq$ 50 kts, hail $\geq$ 1 inch in diameter, or any tornado report. To reduce limitations from population-based reporting, we follow methods from Murillo and Homeyer (2019), such that only severe weather reports that occurred in a population dense region ($>$ 25 people per square mile) are retained for analysis, according to version 4 of

the Gridded Population of the World dataset (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018). This provides the best chance to analyze hail-producing supercells at a time where reliable maximum hail size reports exist.

Tornado, wind, and hail SED reports were leveraged here, as opposed to the commonly used alternative quality-controlled version released by the SPC, due to the additional tornado report information available in the original SED reports that is excluded when incorporated in the SPC. The inherent filtering that results from objective storm-matching techniques (see Section 2.2.4) result in a similar quality dataset to that of the SPC.

2.2.4 Data synthesis and analysis techniques

To identify hail-producing supercell storms with differing maximum hail size, we first add several parameters from the GridRad volumes to the storm tracks, considering only observations within 30 km of storm center, and population density. The GridRad parameters extracted along each storm track are used to objectively identify right-moving supercell storms, following that done previously and well-demonstrated to be reliable in Sandmæl et al. (2019) and Homeyer et al. (2020). The only exception to our application of the supercell classification is that we remove a previously used echo top criterion to avoid exclusion of shallow supercell storms, which we found to be common in months outside of May-September. We objectively define storms as supercells if the following criteria are met: 1) maximum mid-level (4–7 km AMSL) azimuthal shear $> 4{\times}10^{-3}$ s^{-1} for at least 40 min, 2) storm-maximum mid-level azimuthal shear $\geq 5{\times}10^{-3}$ s^{-1} and column-maximum azimuthal shear $7{\times}10^{-3}$ s^{-1}, 3) storm-maximum radial divergence at any altitude $\geq 1{\times}10^{-2}$ s^{-1}, and 4) storm-maximum σ_V at any altitude ≥ 13 m s^{-1}.

Once supercells are identified, they are examined for hail reports and grouped into three categories based on their maximum hail diameter: marginally severe (1–1.5-in.), significant severe (2–3-in.), and giant/gargantuan (≥4-in.). For marginally and significant severe hail-producing supercell storms, we only consider times at which a storm was over a population-dense location and extract the radar observations for analysis at the time the maximum hail size was reported in such locations. For giant/gargantuan hail-producing supercells, no population density requirement is applied, given the smaller sample size, and therefore, encompasses all hail sizes ≥4 in in maximum dimension. The number of hail-producing supercell storms identified in each category are summarized in Table 2.2, and their locations are shown in Figure 2.11. Given the different population sizes, with the 235 giant/gargantuan storm population being the smallest (Table 2.2), we randomly select 235 profiles from each smaller size category to normalize the sample sizes. This sub-sampling helps to identifying meaningful differences between the storm categories.

We extract volumes at the 5-min radar analysis time closest to the time of maximum reported hail size, centered on the objectively tracked storm location, and rotate the volume so that the 30-min average storm motion vector is aligned with the positive x-dimension, such that we now have supercells on a storm-relative grid. Note that extracting the closest 5-min GridRad volume to the report time introduces a timing uncertainty of ±2.5 min in the radar products, but the reports typically have a larger time uncertainty of up to ±5 min (Witt et al., 1998b). Environmental profiles are extracted from the closest hour preceding and ≥30 min prior to each storm analysis time and one grid point removed to the east. Although the results were insensitive to the precise grid point extracted within one grid point, these spatial and temporal offsets minimize the chances of outflow contamination in the lower troposphere.

To assess our first research question, we calculate bulk parameters that are commonly used to assess differences in hail environments, including most unstable CAPE (MUCAPE), 0–6 km above ground level (AGL) bulk wind shear, and PWAT. We then examine profiles of environmental temperature, most unstable parcel temperature, humidity, and wind speed to assess purely environmental influences, including conventional hodographs calculated based on the u- and v-components of the ground-relative environmental wind. Last, we investigate distributions of storm-relative winds by leveraging radar-derived storm motions. In particular, we construct a hodograph that is calculated using the storm-relative u- and v-component winds, with the entire profile rotated based on the observed direction of storm motion so that it aligns with the positive x-dimension prior to computing the mean. Rotating the profiles prior to calculating the average storm-relative component wind speeds ensures that the means represent purely parallel and perpendicular storm-relative flow. It is important to note that this is different from the rotation relative to the deep-layer shear vector performed in other studies (e.g., Parker, 2014; Kumjian et al., 2019; Gutierrez and Kumjian, 2021). The approach here assesses what the observed storms experience in a more easily understood, physical framework (parallel/perpendicular to storm motion), made possible by the availability of observed storm motions.

2.2.5 Results

Focusing first on frequency distributions of MUCAPE, 0–6 km AGL bulk wind shear, and PWAT, we find broad overlap between these distributions with slight indications of increased MUCAPE, increased bulk shear, and decreased PWAT with increasing hail size, similar to prior work (Fig. 2.12). These bulk parameters, however, provide limited understanding on how exactly such differences arise (i.e., where in the vertical

profile differences are most substantial) and ultimately contribute little-to-no practical utility for forecasting or nowcasting applications.

To best assess where differences exist between environments of supercell storms with differing hail potential, we evaluate distributions as a function of altitude (Fig. 2.13). Temperature, humidity, and wind speeds at all altitudes are mostly identical among the three storm populations (Figs. 2.13a–c). However, comparing profiles of the difference between most unstable parcel and environmental temperature (Fig. 2.13d) and storm-relative wind speed (Fig. 2.13e) leads to the emergence of clear incremental increases in both quantities throughout much of the troposphere with increasing hail size (also consistent with prior work). For example, low-level storm-relative wind speeds for giant/gargantuan cases increase to at least 1 m/s greater than other hail-size categories by just a few hundred meters AGL. Such an increase in storm-relative wind speeds in the inflow layer lead to wider updrafts (e.g., Peters et al., 2019). The storm-relative wind profiles also demonstrate that low-level (0–2 km AGL) wind shear is $\sim$0.5 m s^{-1} km^{-1} weaker in the giant/gargantuan hail-producing supercell environment. Substantial overlap in the environments remain for these quantities (as for the bulk parameters), but differences in the hail growth layer (4–10 km AGL or temperatures of $\sim$0°C to $-$40°C) are most evident. Note that these profiles were also made in altitude relative to the 0°C level, but results found are broadly insensitive to this choice since the altitudes of a given tropospheric isotherm typically differ by $\leq$1 km in these environments (identifiable in Fig. 2.13a as the altitude span of the distributions at a select threshold temperature). In summary, the practical use of detailed profile information likely remains as low as the bulk parameters given the overlap between hail size categories.

Because the environmental ground-relative wind profiles differ little between the supercell storm populations, whereas the storm-relative wind speed shows incremental

differences, statistics on the speed and direction of radar-observed storm motion were investigated (Fig. 2.14). This analysis demonstrates that both the speed and direction of supercell motion change with increasing hail size potential and are the primary source of identified differences in storm-relative wind. There is a slight slowing of storm motion and more right-ward (i.e., deviant) motion with increasing hail size, especially at the most extreme hail sizes analyzed. Thus, both elements of storm motion contribute to the incremental increases in storm-relative wind speed found with increasing hail size, in agreement with recent studies by Bunkers (2018), Gutierrez and Kumjian (2021), and Lin and Kumjian (2022).

Lastly, we find the greatest differences between supercell environments from the hodograph analysis (Fig. 2.15). The conventional hodographs suggest slightly greater 0–3 km directional shear and slightly greater deep-layer speed shear for increasing hail size potential (Fig. 2.15a). Conversely, the storm-relative, storm-rotated hodographs partially de-emphasize the importance of differences in low-level directional shear and more substantially indicate that storm-relative wind speed perpendicular to storm motion above 2 km AGL is an important environmental constraint on hail size (Fig. 2.15b). This result is consistent with recent machine learning efforts by Gensini et al. (2021) that identified mid-level SRH as the most important environmental predictor of severe hail size.

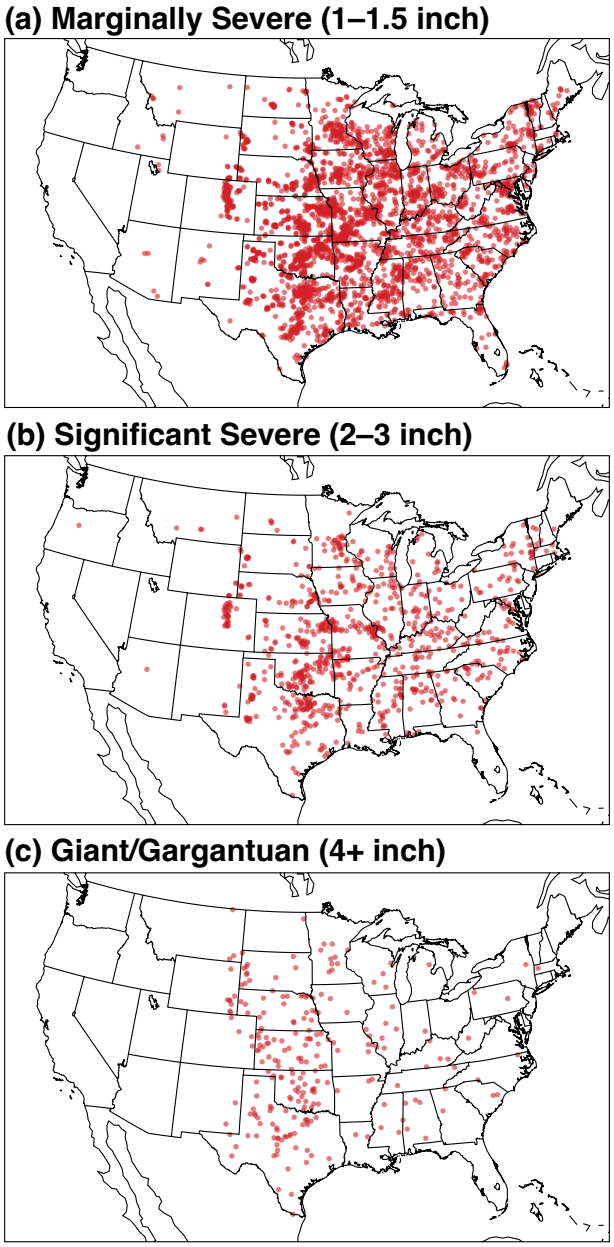

Figure 2.11: Maps of analyzed hail-producing supercell storm locations that are (a) marginally severe, (b) significant severe, and (c) giant/gargantuan.

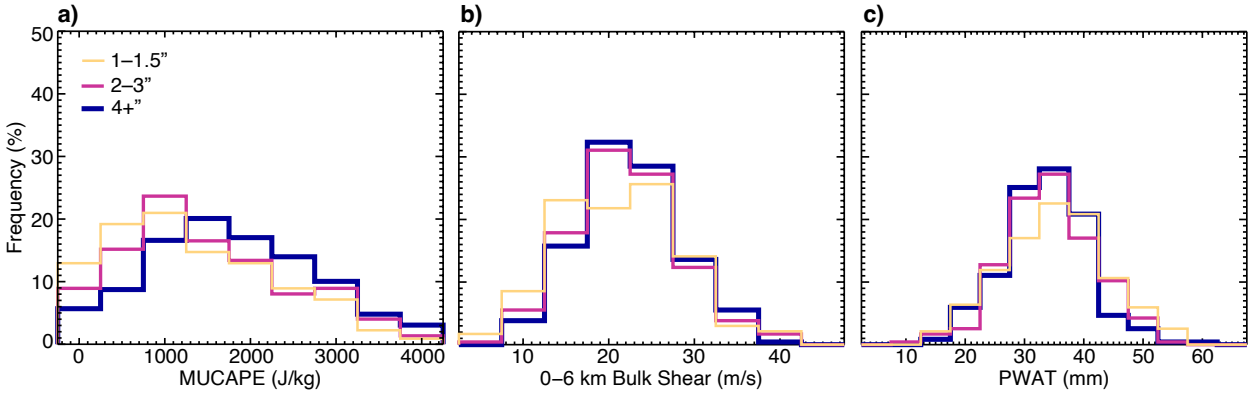

Figure 2.12: Frequency distributions of (a) MUCAPE, (b) 0–6 km AGL bulk wind shear, and (c) PWAT from near-storm environments. Tan, pink, and dark blue lines of increasing thickness show distributions for marginally severe, significant severe, and giant/gargantuan hail-producing supercell storm populations, respectively.

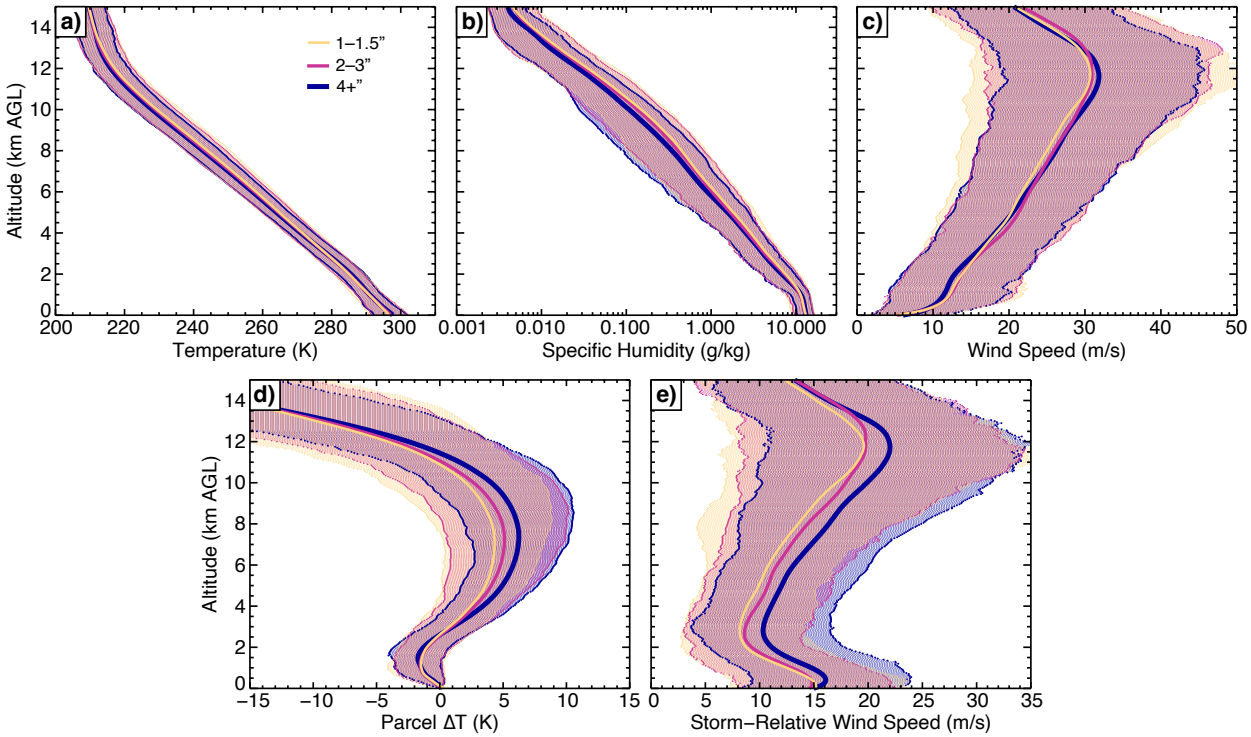

Figure 2.13: Ground-relative vertical profiles of near-storm environments. Panels a–e show profiles of temperature, specific humidity, wind speed, the difference between most unstable parcel temperature and the environmental temperature (i.e., potential buoyancy), and storm-relative wind speed, respectively. Solid lines show the average profile, while dashed lines with dotted ends show the 10^{th} to 90^{th} percentile range of values. Colors as in Fig. 2.12.

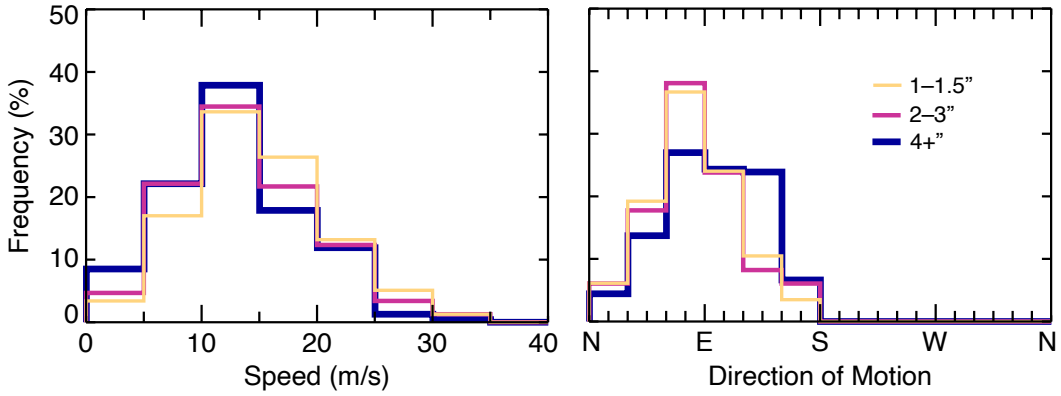

Figure 2.14: As in Fig. 2.12, but for the (left) speed and (right) direction of storm motion.

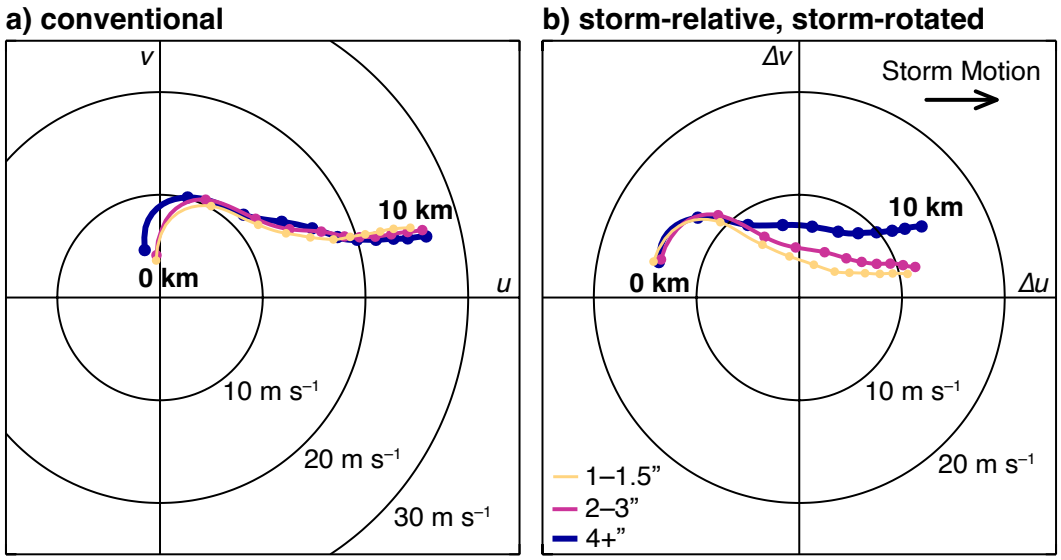

Figure 2.15: Mean 0–10 km AGL hodographs of ERA5 winds computed (a) conventionally such that u- and v-component winds are aligned with the x- and y-dimensions, respectively, and (b) in magnitude relative to storm motion and rotated prior to computing the mean such that the storm motion vector is aligned with the positive x-dimension. Colors indicate the corresponding hail size population, as in Fig. 2.12, and circles superimposed along each storm profile indicate altitudes every 1 km AGL beginning at 0 km.

Chapter 3

Overshooting Storm Tops

3.1 Study no. 3: What Determines Above-Anvil Cirrus Plume Infrared Temperature?

As previously described in Section 1.3, this study seeks to answer the following research questions: 1) Why do some AACPs exhibit a warm feature in IR imagery while others do not? and 2) What observable storm and environment differences exist between warm and cold AACPs? Hypothesis, outline in Section 1.3, are depicted in Fig. 3.2.

3.1.1 Satellite observations

The latest GOES generation, including GOES-16, features the Advanced Baseline Imager (ABI; Schmit et al., 2017, 2018). The ABI samples 16 spectral bands at 2 km horizontal resolution every 5 min for the CONUS domain. Additionally, the ABI can perform more frequent sampling over two fixed areas, each approximately 1000 km^2, that can provide more detailed observations for areas of interest (e.g., severe storms, blizzards, or tropical cyclones). When the ABI is operating in this mesoscale mode, scans are completed every 60 s, such that if the two mesoscale domains overlap, the overlapping area would be scanned every 30 s. Given the manual and highly complex nature of AACPs and their identification (Bedka et al., 2018b), we only evaluated mesoscale mode imagery between March and July from years 2017-2020 in this study

to increase confidence in AACP identification and focus on times when AACPs are common.

AACPs are manually identified using visible imagery (0.64 μm band) and classified as either warm or cold based on their IR imagery characteristics (10.3 μm band). We first searched for cold AACPs, given their relative infrequency, recording the latitude and longitude of the corresponding overshoot at hourly intervals. We required that all AACPs used in this study were sustained (continuously emitted) for at least one hour to compare only long-lived AACPs and to ensure that hourly model analyses were representative. Once 100 cold AACPs were identified, we repeated the process for warm AACPs, but only searching during the months and years that cold AACPs had been identified to mitigate biases associated with seasonality, thus initially resulting in 100 warm and 100 cold manually-identified AACP storms for analysis (reduced slightly to 89 storms each for final analysis based on additional criteria outlined in Section 3.1.5). It is important to note that regardless of their IR signatures, AACPs in this study had no discernible differences in their visible signatures.

3.1.2 Radar observations

To evaluate absolute and storm-relative warm and cold AACP-producing storm characteristics and to quantify their degree of tropopause-overshooting, we utilize ground-based radar observations. This study leveraged identical methods of GridRad creation (GridRad Version 4.2), objective storm tracking, and severe weather report matching as described in Section 2.2.1. The resulting multi-radar volumes were then used to calculate the level of maximum detrainment (LMD, see Section 3.1.5). GridRad echo top heights and intra-storm features typically have near-zero bias and an uncertainty of $\pm$1 km (e.g., Homeyer and Bowman, 2022, and references therein). The radar tracks

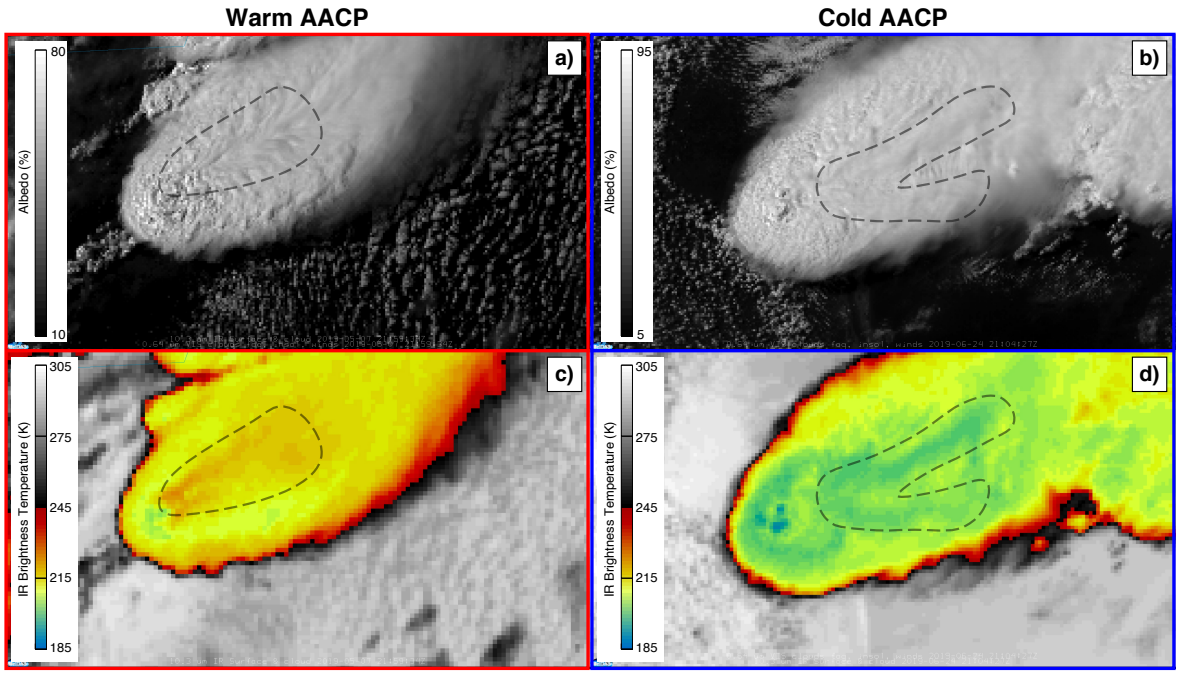

Figure 3.1: Example satellite imagery from 0.64 μm channel 2 (a and b) and 10.3 μm channel 13 (c and d) for a warm AACP event (left) and a cold AACP event (right). The warm AACP is from 21:59Z on 7 May 2019 (a and c). The cold AACP is from 21:04Z on 24 June 2019 (b and d). The outlines in each panel enclose the AACP features.

include 5-min latitude/longitude position, LMD heights, echo-top heights for multiple radar reflectivity thresholds, and 30-min mean storm motion.

3.1.3 Environmental data

This study used similar ERA5 profiles to that described in Section 2.2.2, featuring approximately 400–500 m resolution in the UTLS. We extracted the closest ERA5 grid point to the storm location every hour for the environmental analysis. We tested the sensitivity of the ERA5 results to the precise model grid box chosen by performing similar analyses with a randomly-selected nearby grid point and found little to no impact on the results summarized below. Vertical profiles of temperature, pressure, geopotential height and winds are leveraged to examine tropopause characteristics

and environmental differences between those associated with warm or cold AACPs. When defining the tropopause in this study, we use the lapse rate tropopause definition (World Meteorological Organization, 1957) given its capability for multiple tropopause identification in a profile and its well-demonstrated global reliability (Gettelman et al., 2011; Hoffmann and Spang, 2021; Pan et al., 2018).

3.1.4 Severe weather reports

These analyses use the same severe report dataset and filtering process discussed in Section 2.2.1, that leverages the SED (NCEI/NOAA, 2022) and the Gridded Population of the World dataset (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018).

3.1.5 Data synthesis and analysis techniques

The 100 warm and 100 cold AACPs identified using satellite imagery were manually matched with their corresponding radar storm track to enable comprehensive analysis. Warm and cold AACP storms that were not fully captured in radar observations due to data coverage limitations were removed from final analysis. After reevaluating storm populations so that equal seasonal distributions for both storm sets remained, 89 warm and 89 cold AACPs were retained. The full radar storm tracks are shown in Figure 3.3, with storms exhibiting warm (cold) AACPs in red (blue). Storm-relative winds, defined as the difference between ERA5 winds and the radar-tracked 30-min storm motion, are computed for warm and cold AACPs to characterize the potential for gravity wave breaking and thus AACP formation in the UTLS (e.g., Homeyer et al., 2017; O'Neill et al., 2021, and references therein).

Severe weather reports were matched to the radar-based storm tracks if they occurred at a time the storm was tracked and within 30 km of storm center. To complement the report analyses, we evaluate one of the recently revised hail parameters from Murillo and Homeyer (2019) and Murillo et al. (2021), the Linear Discriminant Analysis (LDA)-filtered 75$^{\text{th}}$ percentile of the maximum estimated size of hail (LDA-filtered MESH_{75}, see Murillo et al. (2021) for more details). We also calculate the level of maximum detrainment (LMD) in each storm to compare transport characteristics, defined here as the altitude of the column-maximum anvil ice water content, utilizing methods similar to those of Mullendore et al. (2009), Carletta et al. (2016), and Starzec et al. (2020). Several steps were taken to minimize errors in LMD height due to limitations of the radar observations and downward settling of precipitation-sized hydrometeors in the anvil separate from the convective core (e.g., Homeyer, 2014; Homeyer and Bowman, 2022), as was done in previous studies using GridRad data (e.g., Starzec et al., 2020). Namely, we limited our search radius to 30 km from storm center to be consistent across the analysis and applied the Storm Labeling in 3 Dimensions (SL3D; Starzec et al., 2017) algorithm to identify anvil grid boxes adjacent to convective grid boxes for ice water content analysis. We only considered convective regions where reflectivity observations exist below 4 km. Additionally, we only considered anvil regions that featured $\geq$ 15 dBZ above 6 km, $\geq$ 5 grid boxes of observed reflectivity at individual heights, and $\geq$ 25 grid boxes of observed reflectivity within the full search volume. Carletta et al. (2016) demonstrate that an LMD based on anvil ice water content is generally biased $\sim$750 m low compared to that derived from 3-D winds. The uncertainty of GridRad-retrieved LMDs is expected to be comparable to the vertical grid spacing of the data in anvil regions, $\pm$1 km.

The two-sample Kolmogorov–Smirnov (KS) test is used in this study to determine statistical significance with 99% confidence ($\alpha = 0.01$) for one-dimensional radar and

environmental metrics of warm and cold AACP storms. The null hypothesis of the two-sample KS test states that the two populations originate from the same distribution. Statistically significant differences between samples are found when the p-value $< \alpha$ and thus, the null hypothesis is rejected. Although warm and cold AACP vertical profile analyses are not explicitly evaluated for statistical significance, vertical layers where there is clear separation (a lack of overlap) between the full distribution of values for each AACP population likely indicate statistically significant differences. Moreover, the approximate coincidence in the locations of warm and cold AACP storm populations implies that radar data coverage (and therefore, quality) is similar such that statistically significant differences between the two populations are robust, especially given the relatively large uncertainty ($\pm$ 1 km) of the GridRad echo tops and LMD heights.

3.1.6 Results

Although warm and cold AACP environments in the lower and middle troposphere are quite similar, the largest differences between them are found in the UTLS region. When looking at vertical profiles of temperature, we find that warm AACPs are associated with lower tropopause heights and warmer UTLS temperatures, as opposed to cold AACPs environments that feature higher tropopause heights and colder UTLS temperatures (Fig. 3.4, left). Namely, primary tropopause altitudes in warm AACP environments are most frequently found near 13 km, consistently lower than cold AACP environments, which commonly occur near 15 km (Fig. 3.6, left). We also find that temperature starts to increase immediately above the tropopause in cold AACP environments, while an inversion/isothermal layer exists above the tropopause in warm AACP environments (Fig. 3.4, right). These characteristics indicate that warm AACPs are associated with midlatitude environments, given the low tropopause height and

tropopause inversion layer; meanwhile, cold AACPs are associated with tropical environments, given the resemblance to the cold-point tropopause and transition layer that is characteristic of tropical environments.

We also independently evaluate warm (cold) AACPs that feature double tropopause environments, which includes 38 (13) of the 89 storms. These environments exhibit similar primary and secondary tropopause heights, but feature unique lapse rate characteristics above the primary tropopause (Fig. 3.5, left). In cold AACP environments, temperature continues to decrease with height, however at a slower rate, and satisfies the WMO second tropopause condition several kilometers higher (Fig. 3.5, right). Above the primary tropopause in warm AACP environments, however, there are clear inversion and isothermal segments up to the secondary tropopause, which is found at even higher altitudes than that in cold AACP environments. Such separation between tropopause heights is commonly ∼4 km in warm AACP environments and ∼3 km in cold AACP environments, despite similar secondary tropopause altitudes (Fig. 3.6, right). Thus, these results demonstrate that in a select few cold AACP cases (∼15%), the UTLS contains a double tropopause and is characterized as more of a subtropical environment, with no inversion layer above the primary tropopause and relatively small separation between tropopauses. Warm AACP environments with a double tropopause are far more common (∼43%) and, similar to the analysis for all cases, exhibit a strong tropopause inversion layer above the primary tropopause.

Next, we evaluate radar-observed characteristics of each storm population in the context of associated environmental factors. When assessing storm-relative wind speed, we find comparable structure throughout the troposphere (Fig. 3.7, left), but key differences are revealed in tropopause-relative analyses (Fig. 3.7, right). Most notably, warm AACP storms have higher storm-relative wind speeds in the first 3 km above the

tropopause, while storm-relative wind in cold AACP storms remains relatively weak (below 15 m s^{-1}) up to ∼4 km above tropopause level (Fig. 3.7, right). Although there are somewhat larger overall differences between warm and cold AACP storm-relative wind profiles within double tropopause environments, tropopause-relative analyses exhibit similar above-tropopause features to that for all environments (Fig. 3.8). Given that previous work indicates the necessity of storm-relative wind speeds ≥15 m s^{-1} to drive frequent gravity wave breaking and AACP development (e.g., Homeyer et al., 2017; O'Neill et al., 2021, and references therein), these results suggest that AACP development above the tropopause in cold AACP environments is unlikely.

To better assess the joint relationships between AACP storms and their environments, we examine the observed depth and detrainment levels of each storm here. Absolute LMD altitudes are essentially identical in warm and cold AACP storms (Figs. 3.9a & 3.9c), commonly spanning altitudes from 9-12 km. However, tropopause-relative analysis shows that LMD heights are significantly closer to the tropopause in warm AACP storms than cold AACP storms (Figs. 3.9b & 3.9d), as expected based on the environmental results discussed previously. 10-dBZ echo top altitudes in warm and cold AACP storms also overlap considerably (Figs. 3.10a & 3.10c). Consistent with LMD analyses, we find that tropopause-relative 10-dBZ echo top altitudes within warm AACP storms are significantly higher than cold AACP storms, commonly reaching 4 km above the tropopause (Figs. 3.10b & 3.10d). Tropopause-relative 10-dBZ echo tops within double tropopause environments follow similar patterns to the full distributions, with slightly less overlap (Fig. 3.11).

Lastly, we assess the severity of each AACP storm population using severe weather reports and $MESH_{75}$. To adequately account for the different severe storm population

sizes, storm lifetimes, and additional limitations in the report database, we compute the number of severe weather reports for every 5 min that storms were within a population dense area (Table 3.1). Warm AACP storms produced approximately twice as many severe wind and hail reports per 5 min than cold AACP storms and comparable tornadoes. However, warm and cold AACP storms exhibit similar lifetimes with severe $MESH_{75}$ (Fig. 3.12). Given the extensively documented limitations of hail reports, radar products, such as $MESH_{75}$, are often viewed as equally reliable as severe hail reports after appropriate methods have been implemented to the reports (e.g., Allen and Tippett, 2015; Murillo and Homeyer, 2019; Allen et al., 2020; Murillo et al., 2021). As such, these $MESH_{75}$ results cast reasonable doubt on the reliability of the severe hail report results.

AACP Storm Statistics	**Warm AACPs**	**Cold AACPs**
No. all storms	89	89
No. storms within population dense area at any point	44	36
Mean storm time duration within population dense area	100.2 min (63.1%)	83.7 min (51.4%)
No. all severe reports per 5min storm duration (No. severe storms)	9.1 (31)	4.5 (23)
No. tornado reports per 5min storm duration (No. tornadic storms)	0.6 (6)	0.7 (7)
No. severe wind reports per 5min storm duration (No. severe wind storms)	2.2 (16)	0.7 (8)
No. severe hail reports per 5min storm duration (No. severe hail storms)	6.2 (30)	3.2 (19)

Table 3.1: Severe weather report statistics for warm and cold AACP storms.

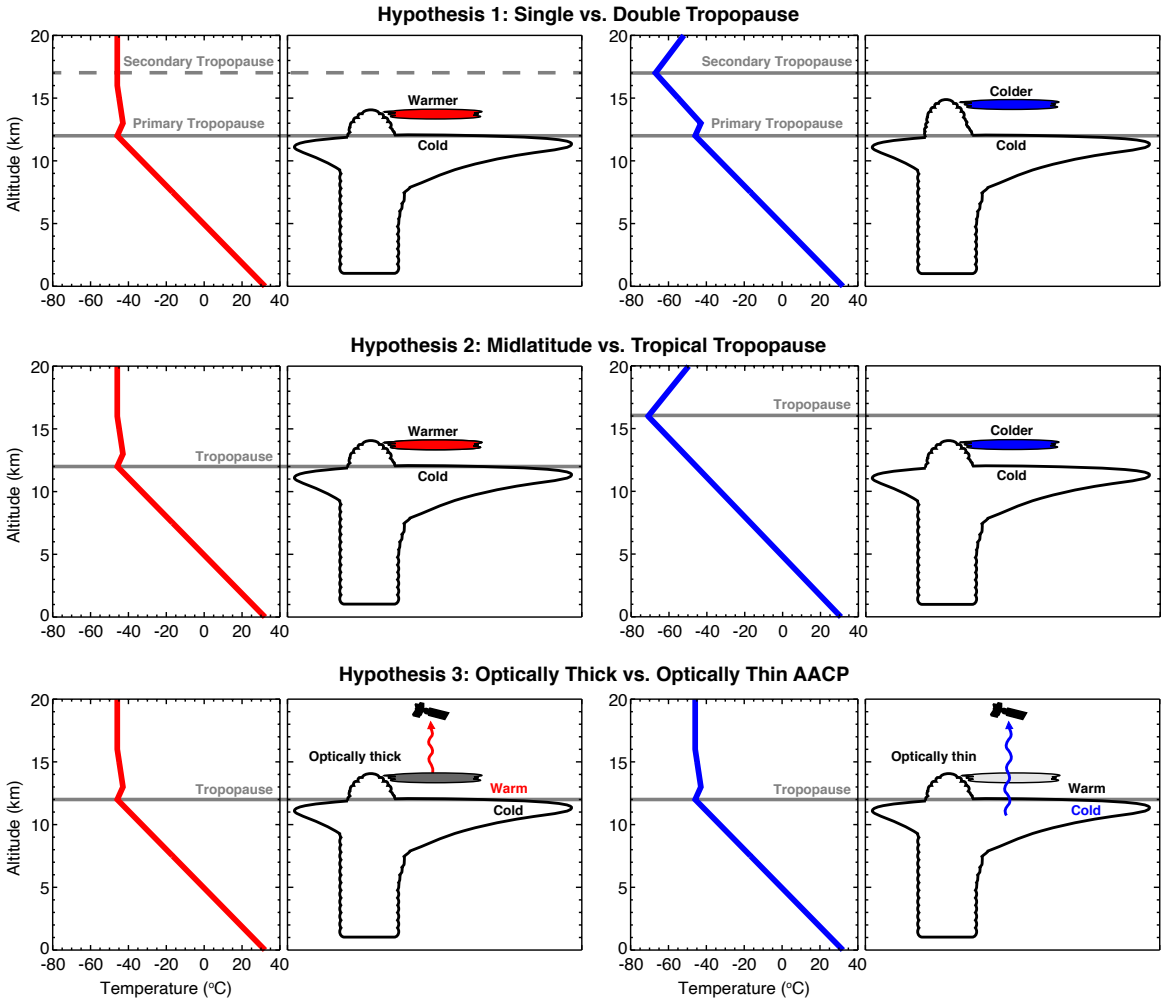

Figure 3.2: Conceptual illustrations of the 3 hypotheses proposed to explain the AACP IR temperature differences that are the focus of this study. Hypothetical temperature profiles, tropopause heights, and cloud positions are shown for warm (red) and cold (blue) AACPs.

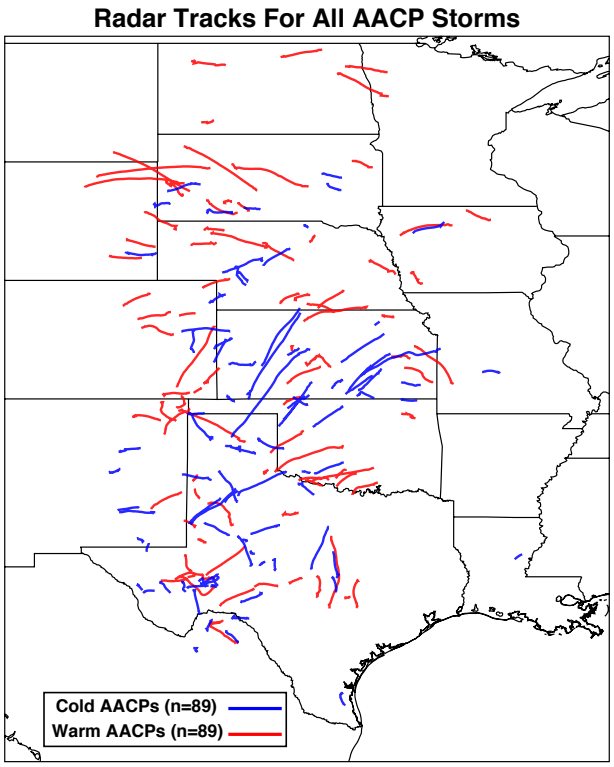

Figure 3.3: Radar-based storm tracks for all warm (red) and cold (blue) AACP storms analyzed in this study.

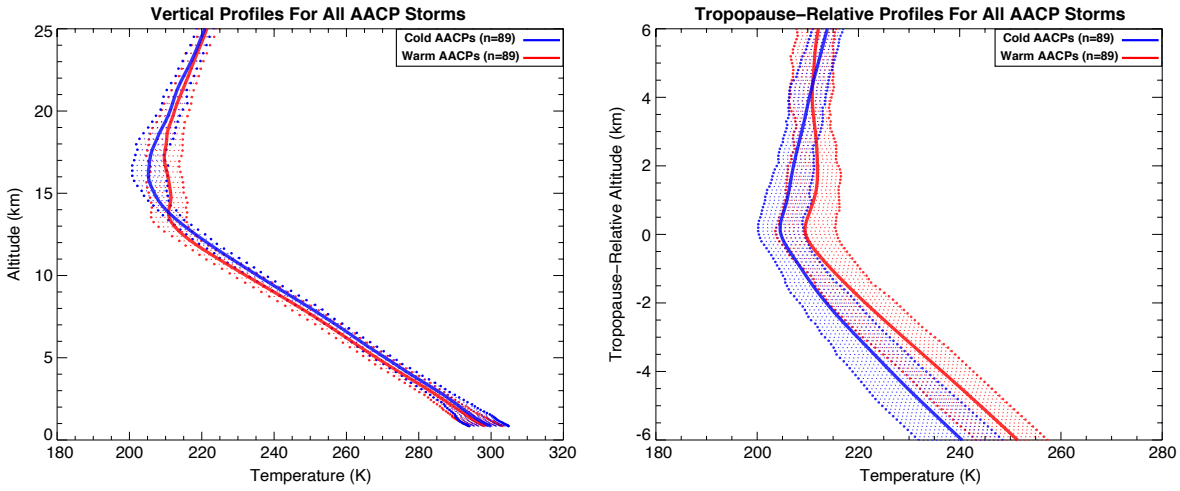

Figure 3.4: Vertical profiles of environmental temperature for warm (red) and cold (blue) AACPs in absolute altitude (left) and tropopause-relative altitude (right). Solid curves indicate mean values at each level. The dots and dashed lines indicate the 10^{th} to 90^{th} percentile range at each level.

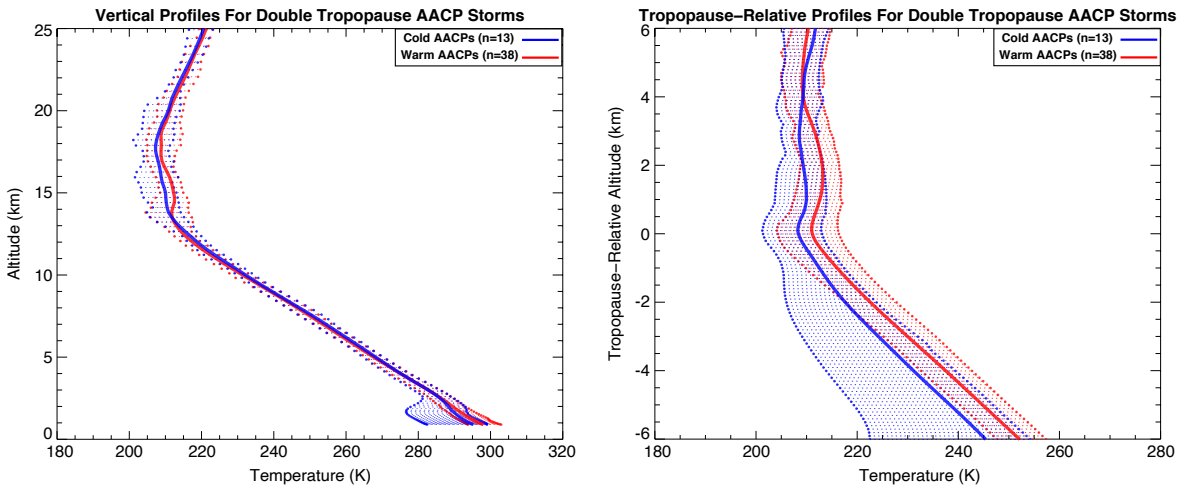

Figure 3.5: As in Fig. 3.4, but only for double tropopause cases.

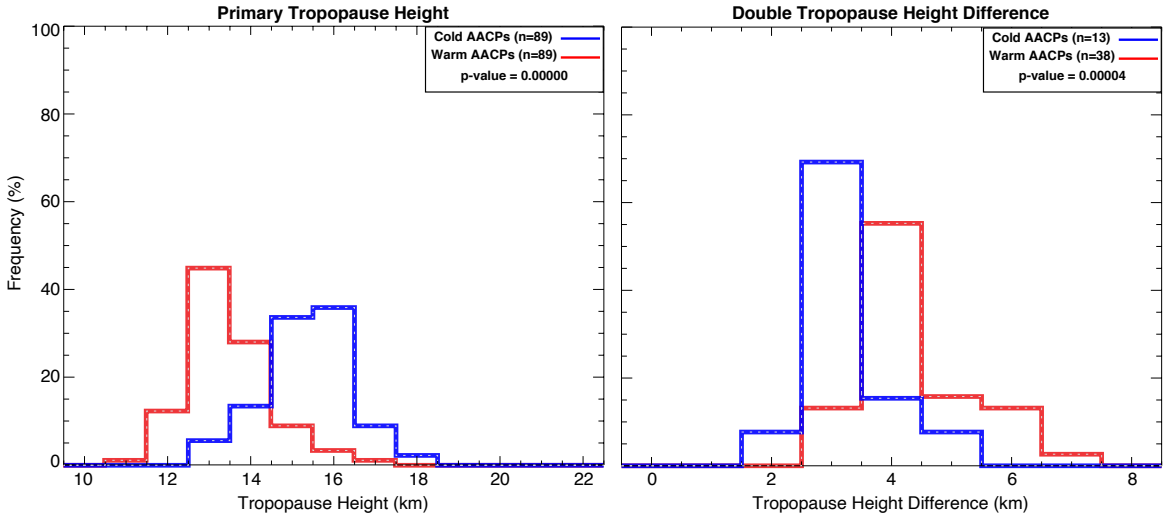

Figure 3.6: Frequency distributions of primary (left) and secondary (right) tropopause heights for warm (red) and cold (blue) AACPs. Distributions that have significant separation (based on the K-S test) feature bolded p-values in the key of each panel and dotted lines within thicker curves.

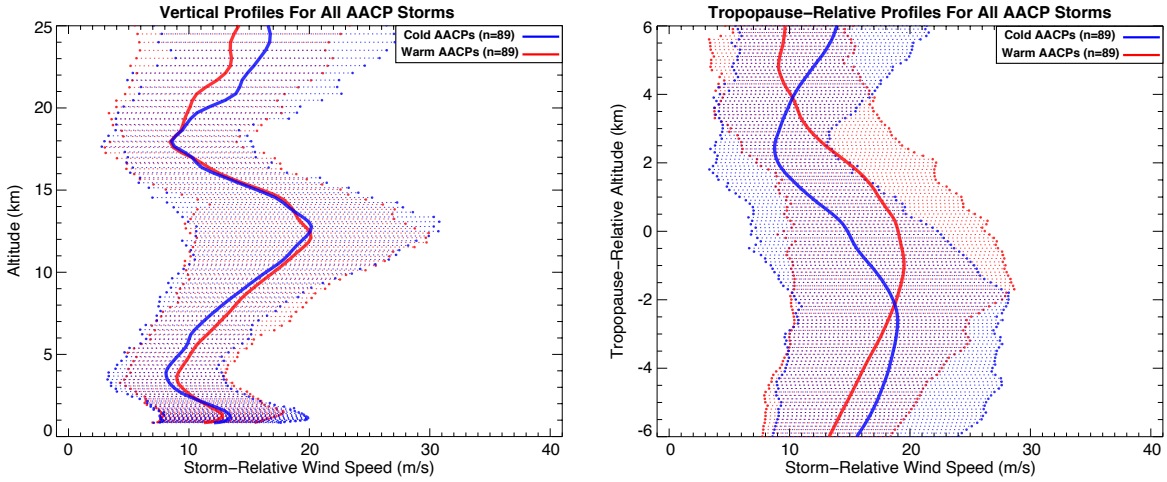

Figure 3.7: As in Fig. 3.4, but for storm-relative wind speed.

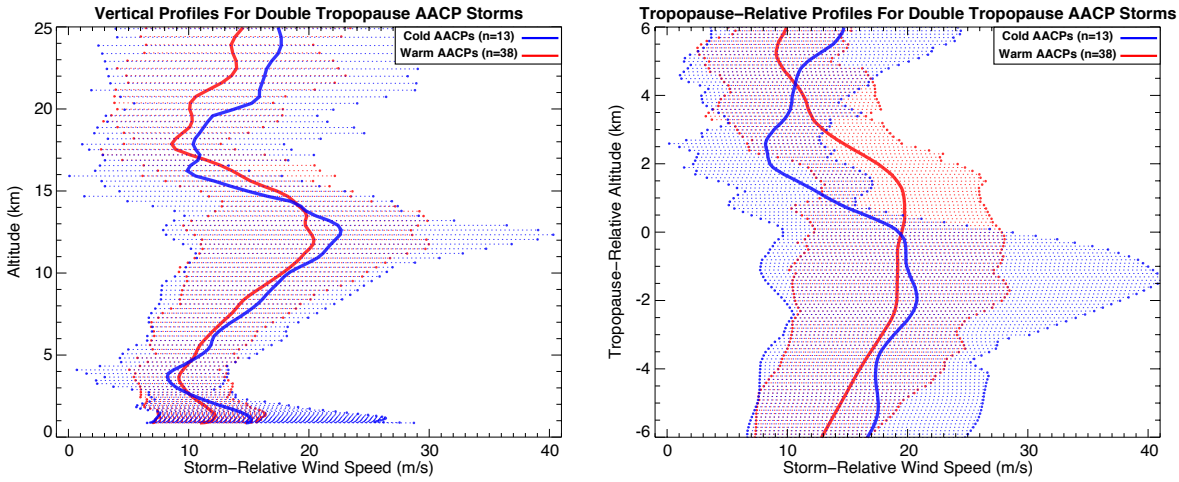

Figure 3.8: As in Fig. 3.7, but only for double tropopause cases.

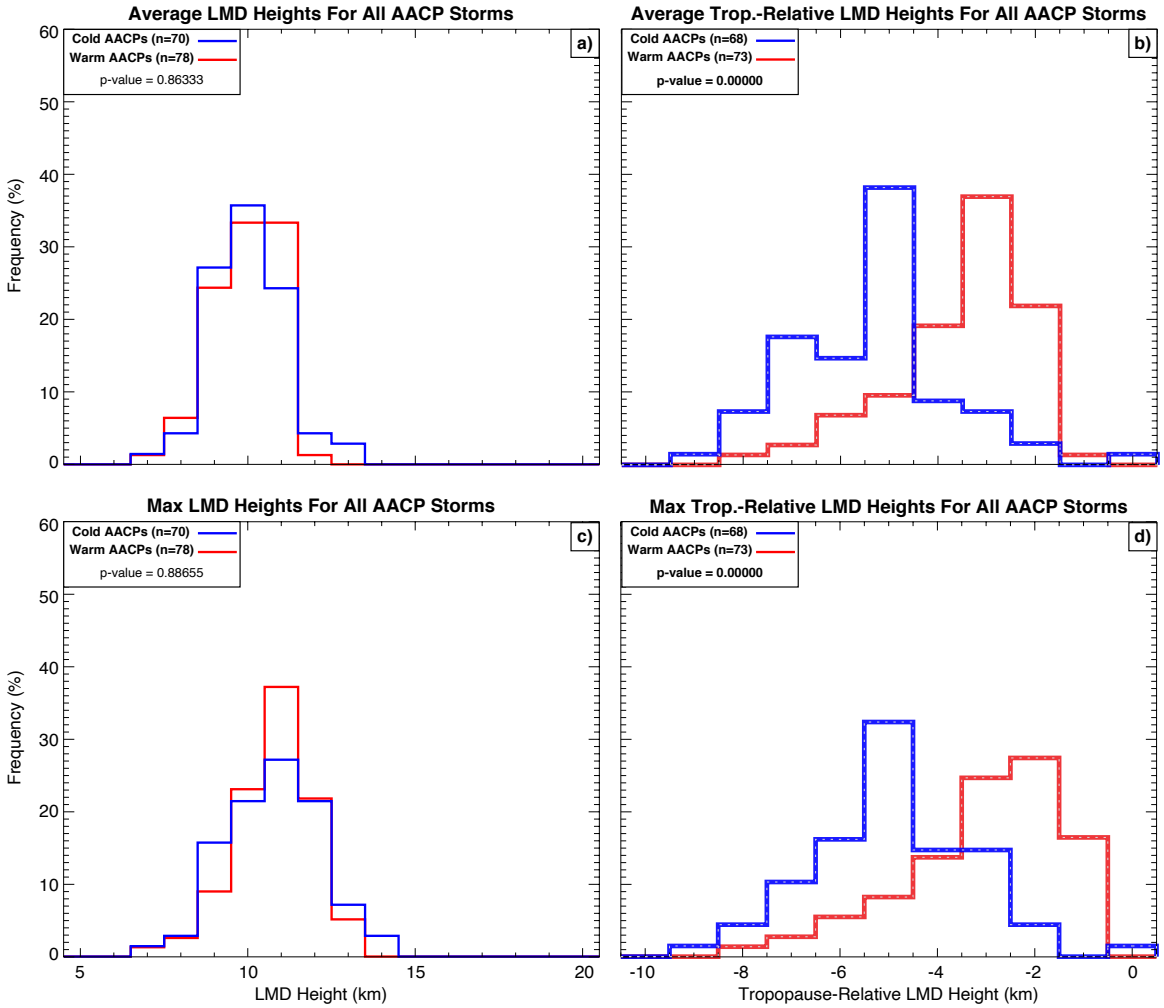

Figure 3.9: Frequency distributions of average (a and b) and maximum (c and d) level of maximum detrainment (LMD) heights for warm (red) and cold (blue) AACPs. Distributions on the left show absolute heights, with tropopause-relative LMD heights on the right. Distributions that have significant separation (based on the K-S test) feature bolded p-values and dotted lines within thicker curves.

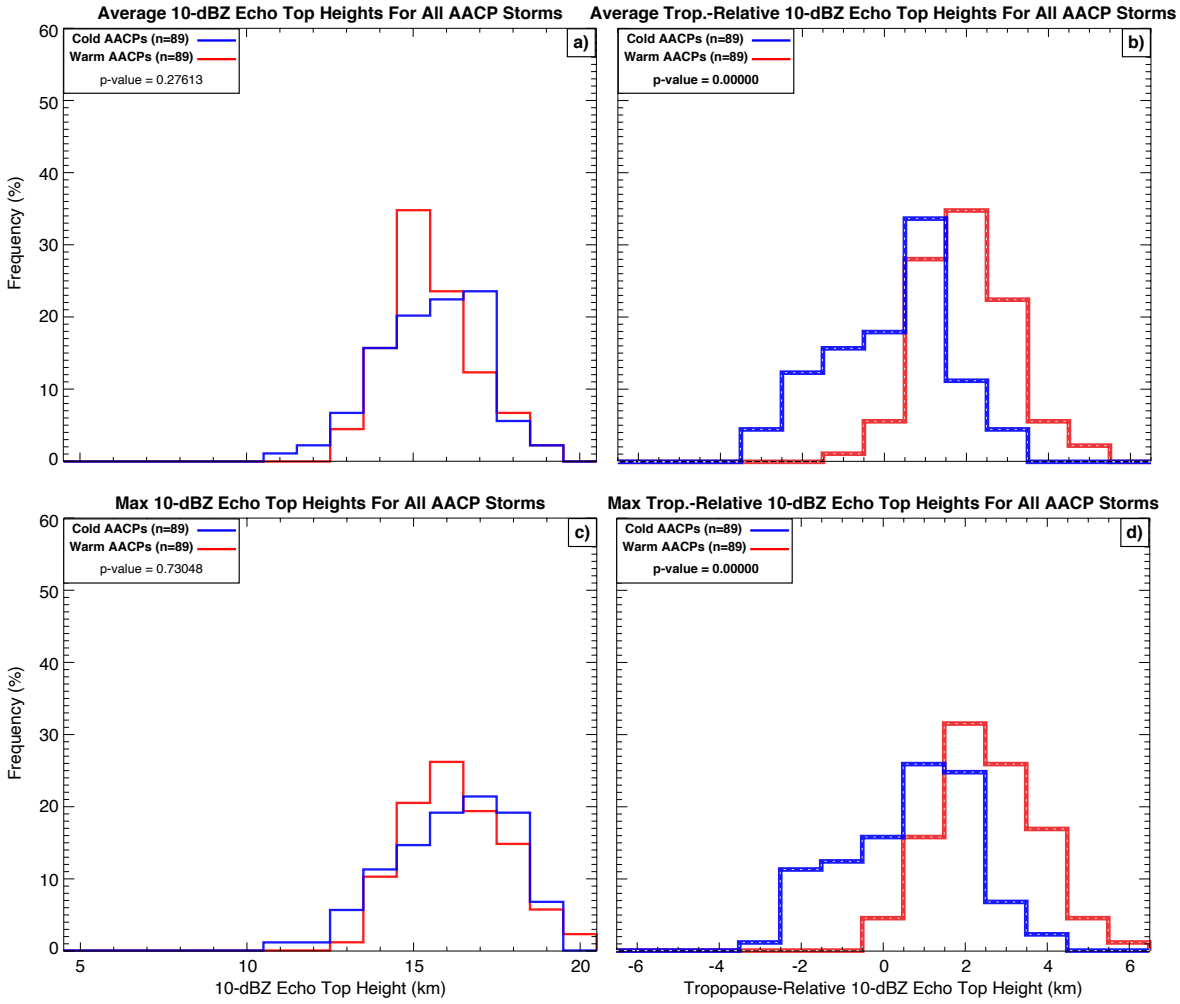

Figure 3.10: As in Fig. 3.9, but for 10-dBZ echo top heights.

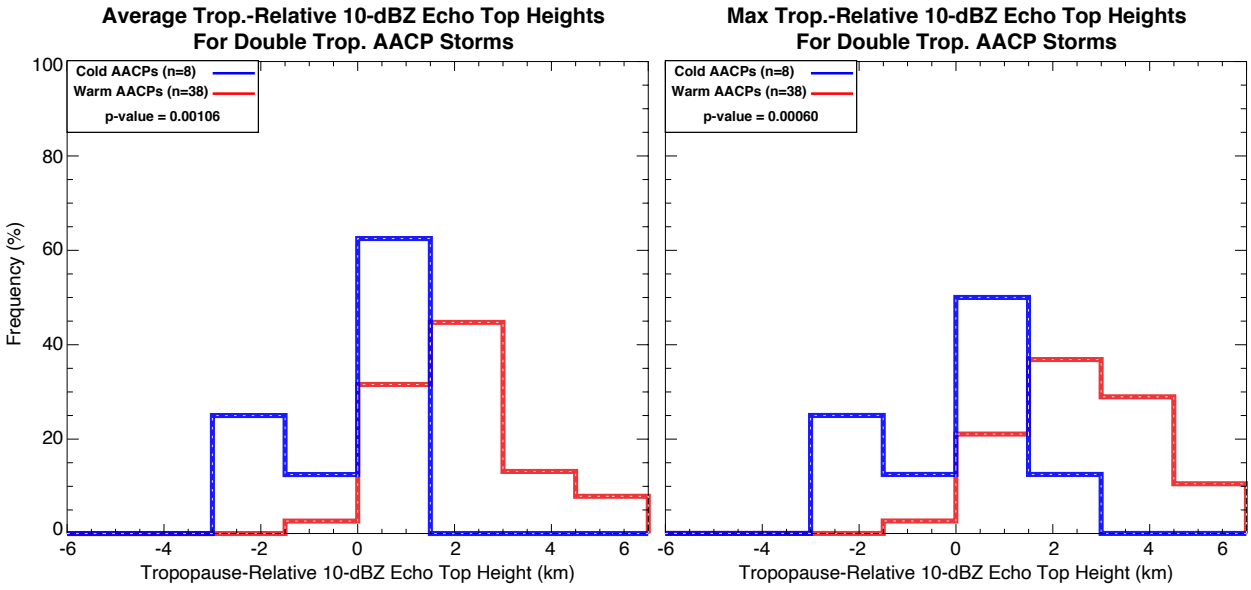

Figure 3.11: As in Fig. 3.10 (right), but only for double tropopause cases.

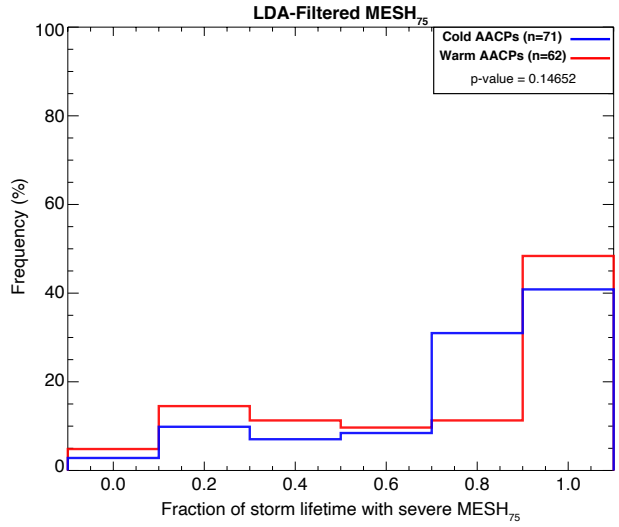

Figure 3.12: Frequency distributions of storm lifetime fractions with severe LDA-filtered $MESH_{75}$ for warm (red) and cold (blue) AACPs. Distributions that have significant separation (based on the K-S test) feature bolded p-values and dotted lines within thicker curves.

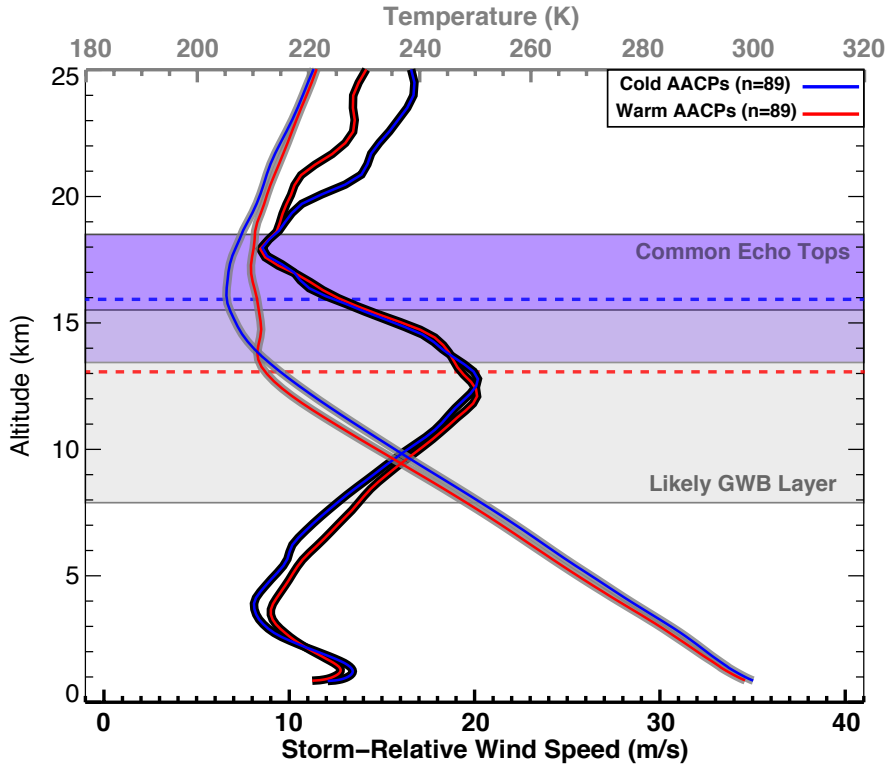

Figure 3.13: A composite diagram summarizing the results of this study. Vertical profiles of environmental temperature (gray, top x-axis) and storm-relative wind speeds (black, bottom x-axis) show the mean values for all warm (red) and cold (blue) AACPs, as previously shown in Figs. 3.4 & 3.7, respectively. Dashed horizontal lines represent mean tropopause altitudes for each AACP type. The gray shading indicates the altitude range that is most supportive of frequent gravity wave breaking (GWB), and the purple shading represents the common altitude range of overshooting tops (10-dBZ echo tops, from Fig. 3.10).

3.2 Study no. 4: Stereoscopic Cloud Top Height Retrievals for 31 May 2022

Introduced in Section 1.3, the last study is comprised of preliminary observational analyses aiming to answer the two research questions: 1) How do stereoscopic cloud-top height retrievals compare to radar-derived echo top products and in situ water vapor measurements? and 2) What new information can be gained from stereoscopic cloud-top height retrievals regarding AACPs?

3.2.1 Satellite observations

GOES-16 and GOES-17 platforms, both stationed over the equator, are longitudinally separated at 75° W and 135° W, respectively, such that they can provide sufficiently unique viewing angles for stereoscopy. Mesoscale domain observations from both satellites used in this study, similar to that described in Section 3.1.1, include scans every 60 sec, VIS (0.64 μm band) and IR imagery (10.3 μm band) with 0.5 km and 2 km horizontal resolution, respectively. Stereoscopic analyses performed here are limited to daylight hours and simultaneous GOES-16 and -17 sampling within a mesoscale domain; thus, we obtain observations from NOAA (1994) and evaluate the period between 22 UTC 31 May through 01 UTC 1 June 2022, providing a 3-hr period for stereoscopy.

3.2.2 Radar observations

This study utilizes identical methods of GridRad creation (GridRad Version 4.2) and echo top height calculations to that described in Section 3.1.2. We calculate echo top altitude fields (following methods in Cooney et al. (2018) to ensure storm continuity) for backward trajectory matching (described in Section 3.2.6). Echo top heights usually

feature near-zero bias and an uncertainty of ±1 km (e.g., Homeyer and Bowman, 2022, and references therein).

3.2.3 In Situ observations

All in situ observations presented here are preliminary from the DCOTSS field campaign and therefore, not for public distribution. Water vapor measurements from the most current version of the Harvard Water Vapor Herriott Hygrometer (HWV-HHH; hereafter, HHH) instrument are used, in conjunction with backward trajectories (see Section 3.2.6 for details), to validate stereo-derived cloud-top heights. Aboard the ER-2, HHH consists of a tunable diode laser that directly detects absorption due to water vapor (Weinstock et al., 1994; Sargent et al., 2013). Measurement accuracy and precision are determined through various in-flight and laboratory testing of independent detection methods, with predicted ±10% accuracy and ±0.1 ppmv water vapor precision for 1 second observations in the lower stratosphere (Weinstock et al., 2009; Smith, 2012; Sargent et al., 2013). These observations use the Meteorological Measurement System (MMS; on the NASA ER-2 Aircraft, 1990; Gaines et al., 1992) as a reference for GPS position and time synchronization.

Due to their preliminary nature, there may be small errors related to bias corrections, quality control methods yet to be performed, aircraft performance, or timing offsets. While these errors might affect exact concentration values, their timing and placements, the errors would likely have minimal impacts on the conclusions presented here due to the larger physical and temporal scales of the associated phenomena.

3.2.4 Environmental data

ERA5 fields used in this study, available at for Medium-Range Weather Forecasts (2019), feature 37 vertical pressure levels on a 0.25° regular horizontal grid (re-gridded from the 0.28125° native model grid (Hersbach et al., 2020)). Tropopause heights, as similarly defined by the lapse rate definition in Section 3.1.3, environmental temperature and winds are used to generate the backwards trajectories (Bowman et al., 2013, see Section 3.2.6 for details) for measurement attributions.

3.2.5 Severe weather reports

For a qualitative assessment of storm severity, the severe report dataset discussed in Section 2.1.3 was also utilized, excluding the population-based filter given the limited spatiotemporal domain.

3.2.6 Data synthesis and analysis techniques

The iterative-based stereoscopic algorithm and associated quality control methods detailed in Young (2021) are used here, which leverage VIS satellite observations to derive cloud-top heights. While IR-based methods for stereoscopic cloud top height retrievals are possible, IR imagery is only used here for quality control. Briefly, the stereo algorithm defines the minimum difference between the disparity value and parallax shift as the cloud top height for each pixel. Disparity characterizes the pixel-by-pixel difference (east-west offset) between each GOES-16 and -17 image, resulting in a value proportional to the observed cloud-top height, while "parallax" is the satellite-observed displacement of high-altitude objects, with displacements increasing as altitudes increase (see Fig. 2.3 in Sandmæl (2017)).

Next, water vapor enhancements are subjectively defined as at least 30 seconds of increased water vapor above local background levels coincident with real-time GOES and GridRad overshoot forecast trajectories (independent and unrelated to those calculated in this study and described below). Here, "local background levels" are defined as low-variability trace gas concentrations at comparable altitudes surrounding the water vapor enhancement along the flight track (as available). Isentropic trajectories are then initialized at the time and location of the aircraft-sampled water vapor enhancements, and advected by ERA5 wind fields backward for 12 hours using the TRAJ3D model (Bowman, 1993; Bowman and Carrie, 2002; Bowman et al., 2013). Particle positions are saved every 10 s along the trajectory path, with their approximate altitudes later interpolated to the track location. To validate the stereo-derived cloud-top heights, echo top and cloud top altitudes are also interpolated along each track. Lastly, the relationship between the spatiotemporal distribution of severe weather reports in our domain and cloud-top heights is assessed.

3.2.7 Results

The main objective for RF13 was to sample fresh outflow from ongoing overshooting convection over central and southwest Oklahoma. Convection initiated in west Oklahoma at approximately 1930 UTC, but unfortunately, GOES-17 did not begin observations in this region until 2200 UTC when AACP-production in multiple storms was ongoing. With takeoff at 2259 UTC on 31 May 2022 and landing at 0407 UTC on 1 June 2022, the 5-hr flight first sampled convective outflow at various altitudes just east of the ongoing convection (Fig. 3.14, right). After completing stacked, level legs between P1 and P2, the ER-2 was then able to sample along the AACP axis between P2 and P3 following radar indications that the corresponding storm had collapsed. Near P3 at around 0230 UTC, the ER-2 experienced moderate turbulence, which is unsafe

for this aircraft and thus ascended until the turbulence subsided. The ER-2 ascended from 51,000 ft (∼15.5 km) to 55,000 ft (∼16.8 km), turned back to P2, and was soon able to safely descend to 51,500 ft for the level leg back to P2. Upon approaching P2, the ER-2 ascended to a maximum altitude of 66,000 ft (∼20 km) and proceeded to return to Salina.

During the period of AACP sampling between P2 and P3, four water vapor enhancement features were identified (Figs. 3.14 & 3.15, segments enlarged or highlighted). First, feature 09, between 02:07:22 - 02:25:47 UTC 1 June, contains most of the westward P2-P3 leg, with water vapor enhancements measured throughout the 15.6–16.5 km layer. Next, feature 10 was observed within the 16–16.5 km layer during the descent around P3 between 02:27:27 - 02:28:47 UTC 1 June. Then, feature 11 contains the latter portion of the eastward P3-P2 leg between 02:31:27 - 02:39:32 UTC 1 June at ∼15.7 km. Finally, feature 12 was observed at 17 km between 02:42:22 - 02:43:07 UTC 1 June during the final ascent. The storm targeted during the flight, located in western OK (circled in Fig. 3.16), was ongoing with AACP production at 2200 UTC 31 May and dissipated between 0100–0130 UTC 1 June. This storm reached it's maximum strength around 2300 UTC 31 May, where it exhibited supercell characteristics and 30-dBZ echo top heights of approximately 17 km (Figs. 3.16a & 3.16b). Simultaneously, a clear and prominent AACP in both GOES-16 and GOES-17 VIS imagery (Figs. 3.16c & 3.16d) was collocated with relatively warm brightness temperatures (Figs. 3.16e & 3.16f). As discussed in 3.1.6, this is most indicative of AACP presence above the local tropopause altitude.

When evaluating swaths of maximum stereo-derived cloud-top heights, we find appropriately differentiated storm-top features, including the satellite-observed updraft, anvil, and AACP characteristics throughout the period (Fig. 3.17). It is important to note that there are erroneous cloud-top heights along portions of the domain boundary

(lower and upper left corners) that are disregarded for analysis. We focus on the cloud features associated with the targeted storm, which exhibits maximum updraft cloud-top heights of ∼21 km and AACP heights between 18–20 km. Updraft heights from 10-dBZ echo tops, however, indicate maximum altitudes of ∼18 km, with decreasing maximum altitudes at increasing dBZ thresholds (as expected, Fig. 3.18). Although updraft magnitudes differ, it appears that areas of relatively strong updrafts in echo tops and cloud top heights are well correlated with severe weather reports (Figs. 3.17 & 3.18). While some discrepancy is anticipated between cloud-top height retrievals and 10-dBZ echo top heights due to the different particle types/sizes observed with each platform, a difference of 3 km warrants further investigation. Tropopause heights near the storm range from 14–15.5 km (Fig. 3.19), producing typically observed (1–3 km) overshooting depths for 10-dBZ echo tops but excessive (5+ km) overshooting for stereo-derived cloud tops. Upon broader inspection, there are additional issues with stereo-derived cloud tops, as seen by the nonphysical, diagonal striping throughout the analysis period. **As such, we note that there are errors within the stereoscopy algorithm that results in nonphysical features in the cloud-top height fields.**

Next, trajectory analyses are used to spatially and temporally identify potential sources (i.e., individual storms) for each water vapor feature. Through this, we are able to provide greater insight into which storms were most likely contributing to lower stratospheric water vapor enhancements. However, specific and confident storm attributions are not possible here due to various errors associated with ERA5 winds and, ultimately, trajectories. Specifically, wind field magnitudes and associated particle displacement derived from any reanalysis contain larger errors in the stratosphere (compared to upper tropospheric winds) (Hersbach et al., 2020) and typically have displacement errors of at least a few kilometers per day (Bowman et al., 2013). It is also likely that hourly ERA5 analyses do not fully capture the large-scale flow in

this region, which — combined with storm-scale flow perturbations not captured by the reanalysis — can produce a misrepresentation of true pathways taken by sampled material.

With these limitations in mind, we can evaluate trajectory locations and any associated cloud-top and echo top extrema potentially responsible for each water vapor feature (Fig. 3.20a). Although the aforementioned storm (circled in Fig. 3.16) was the desired sampling target for RF13, these analyses introduce the possibility that multiple storms, or a different storm entirely, contributed to the water vapor enhancements sampled during these segments. The clear presence of a broad AACP in VIS emanating from the targeted storm's updraft suggests that all four water vapor features could have originated from the targeted storm. However, cloud-top heights along trajectories for features 09–11 are maximized near the location of the targeted storm and further upstream at the more southwestern storm at 2200 UTC 31 May (Fig. 3.20b), indicating that either or both storms could be responsible for the observed water vapor enhancement. Further, areas of ∼16.5 km 10-dBZ echo top maxima along trajectories for features 09–11, consistent with the altitude range (15.6–16.5 km) of sampled water vapor enhancements, are only present at the earlier times when the material is projected to be in the Texas panhandle, coincident with the southwestern storm at 2200 UTC 31 May (Fig. 3.20c). Higher thresholds of echo top heights follow similar, more pronounced trends, providing a consistent indication that these features are more likely to have originated from the southwestern storm as opposed to the targeted storm (Fig. 3.20, d–f). Trajectories from the last and highest water vapor enhancement, feature 12, exhibit stereo-derived cloud-top heights (∼16–18) consistent with the measured enhancement at 17 km, but markedly lower echo top heights. When comparing instantaneous locations of echo top maxima throughout the period of trajectory integration,

feature 12 trajectories were best correlated with the targeted storm, while features 09–11 were best correlated with the southwestern storm (not shown).

Overall, 10-dBZ echo-top height maxima provided the greatest consistency with water vapor enhancement altitude maxima, with only feature 12 showing much lower altitudes. On the other hand, stereo-derived cloud-top height maxima were only consistent within feature 12 trajectories, overestimating water vapor enhancement altitude maxima by ∼3–4 km for the remaining three features. These results suggest that for this case, echo-top heights produce more reliable altitude estimates for AACP-driven water vapor enhancements, given our confidence in the observed altitude ranges. **However, it is important recall the obviously erroneous, nonphysical phenomena in stereo-derived cloud-top heights; corrections to the stereo algorithm are needed to properly assess consistency with sampled altitudes.**

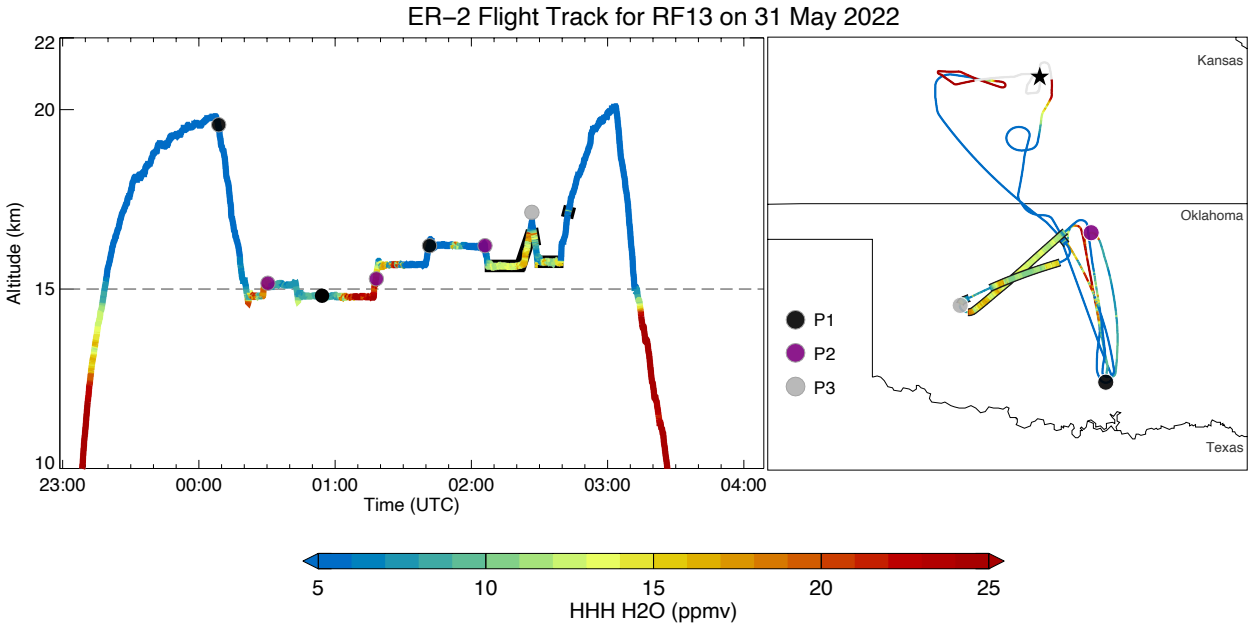

Figure 3.14: Time-height curtain (left) and planar map (right) of the flight path and corresponding water vapor measurements taken during RF13 on 31 May 2022. Points of interest, P1 (black), P2 (purple), and P3 (gray) are overlaid in both panels. Portions of the flight track outlined in black and enlarged indicate water vapor features associated with AACP sampling. Dashed gray line in left panel represents the estimated tropopause altitude from ERA5. The star in right panel indicates the location of Salina, KS.

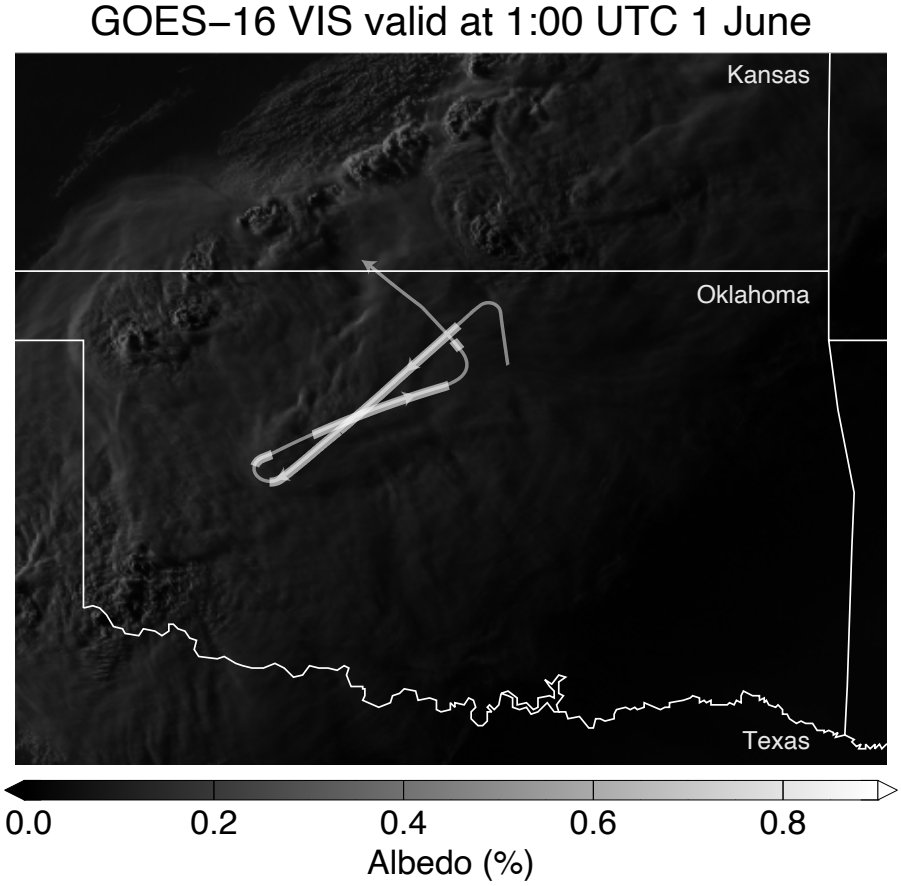

Figure 3.15: GOES -16 0.64 μm channel 2 imagery valid at 0000 UTC 1 June 2022 (last VIS image available where AACP is visible) with the along-AACP sampling portion of the flight track overlaid in gray. Highlighted portions of the flight track indicate water vapor features associated with AACP sampling.

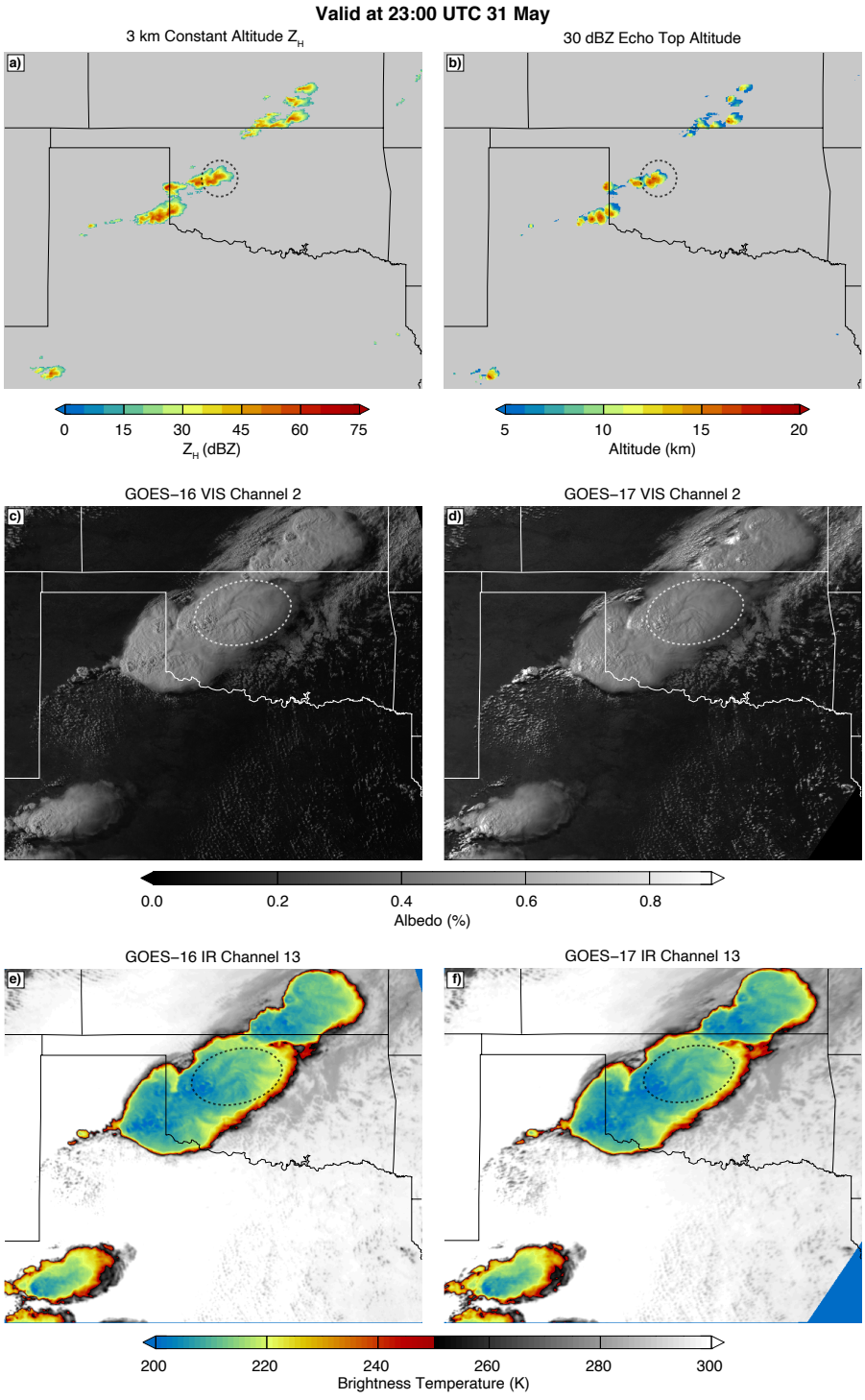

Figure 3.16: Maps of GridRad radar reflectivity at a constant altitude of 3 km AMSL (a), GridRad 30-dBZ echo top altitudes (b), GOES -16 and GOES -17 VIS (c & d), and IR (e & f) imagery, all valid at 2300 UTC 31 May 2022. The storm sampled by in situ observations and its corresponding AACP are circled in each panel.

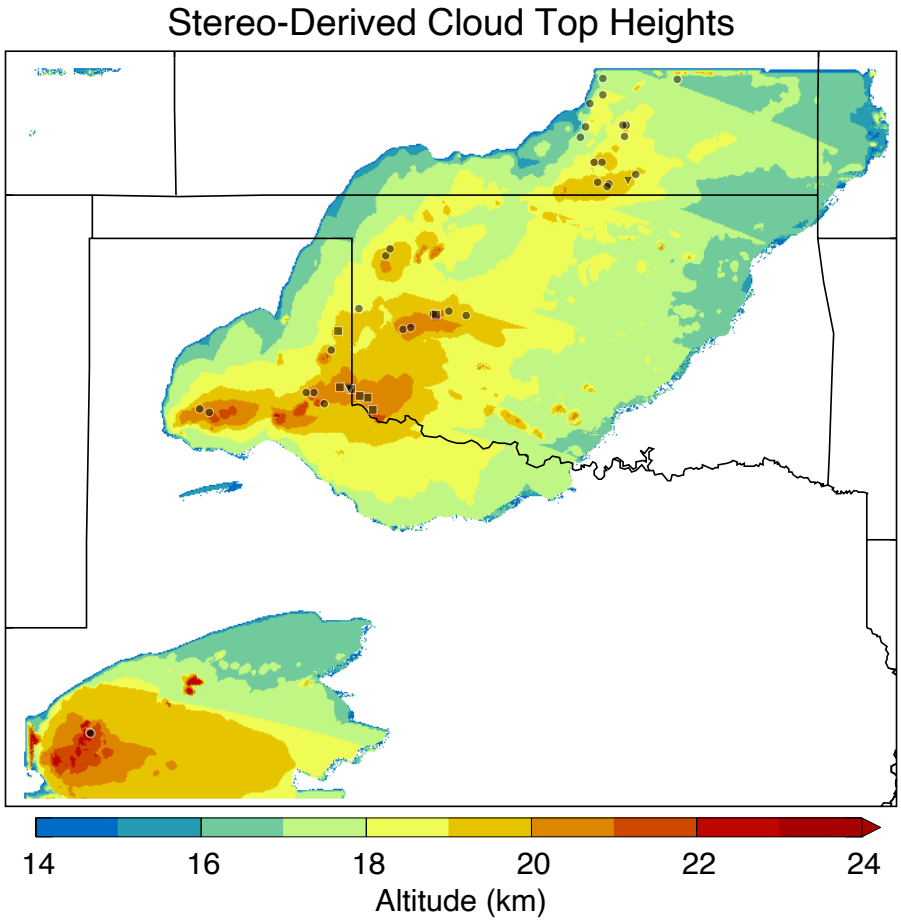

Figure 3.17: Maximum stereo-derived cloud-top altitudes accumulated throughout the 3-hr analysis period between 2200 UTC 31 May and 0100 UTC 1 June 2022. Severe hail, wind, and tornado reports are shown as circles, squares, and upside-down triangles, respectively.

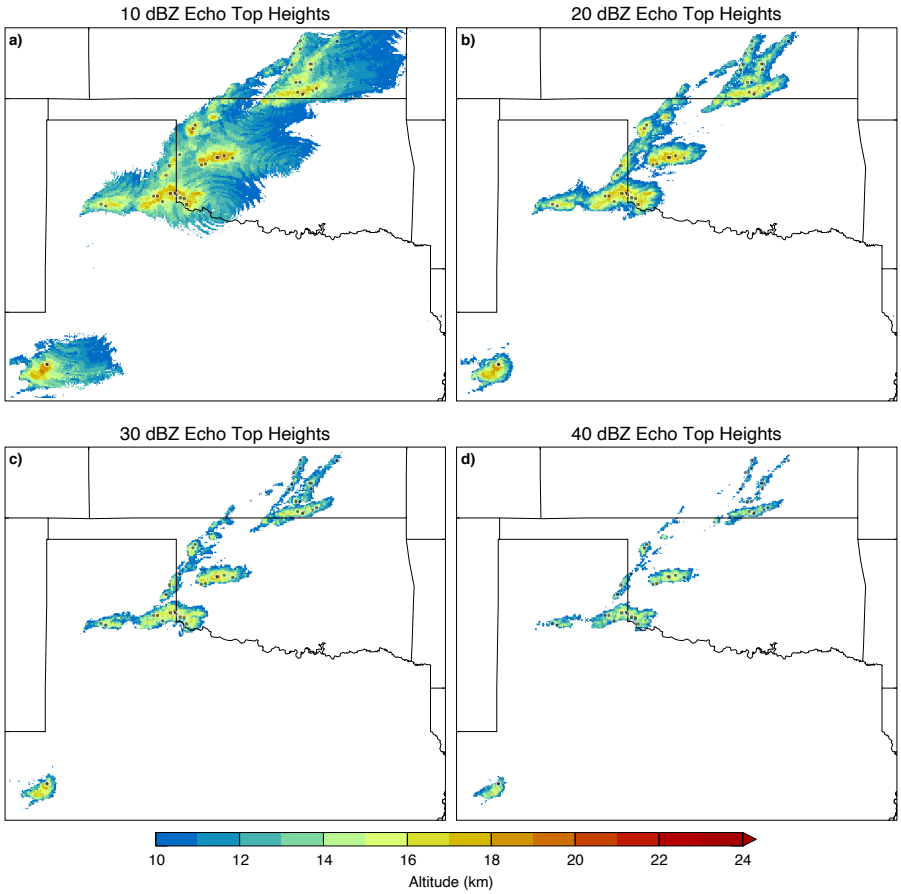

Figure 3.18: Same as Fig. 3.17, but for 10 dBZ (a), 20 dBZ (b), 30 dBZ (c), and 40 dBZ (d) GridRad echo top altitudes.

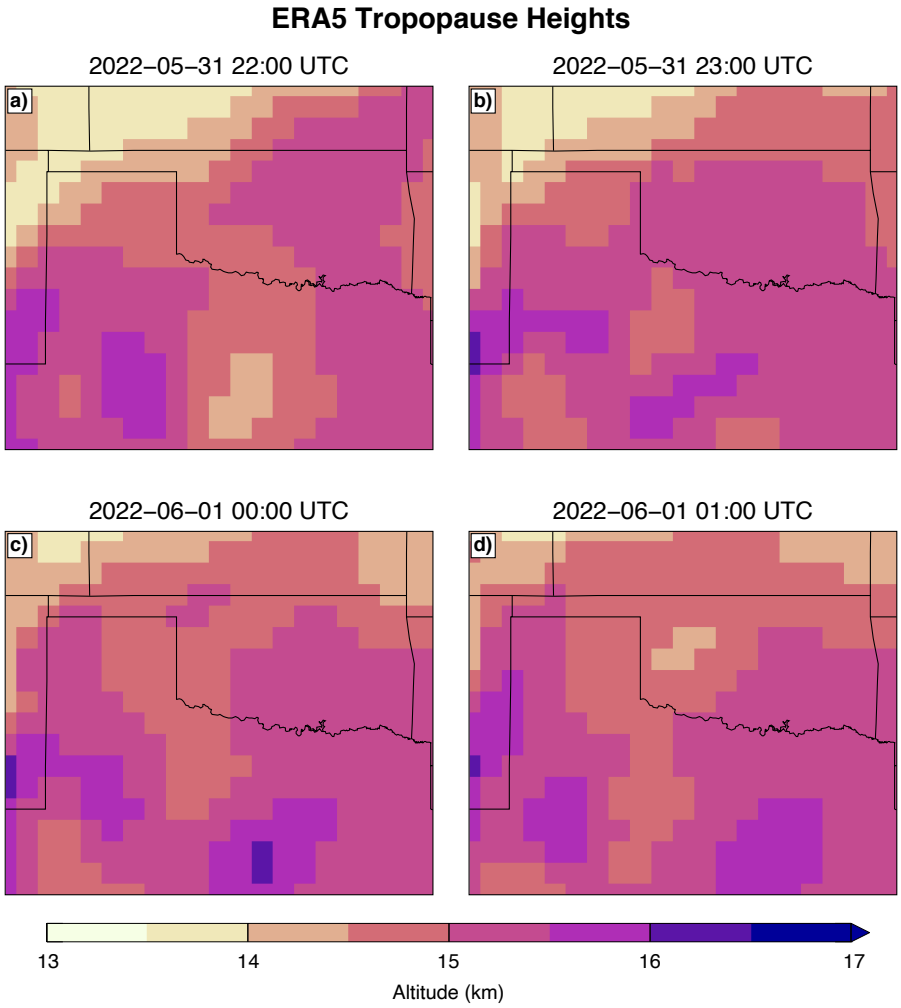

Figure 3.19: Lapse rate tropopause altitudes derived from ERA5 temperature profiles during the stereo analysis time frame: valid at 2200 UTC 31 May (a), 2300 UTC 31 May (b), 0000 UTC 1 June (c), and 0100 UTC 1 June (d).

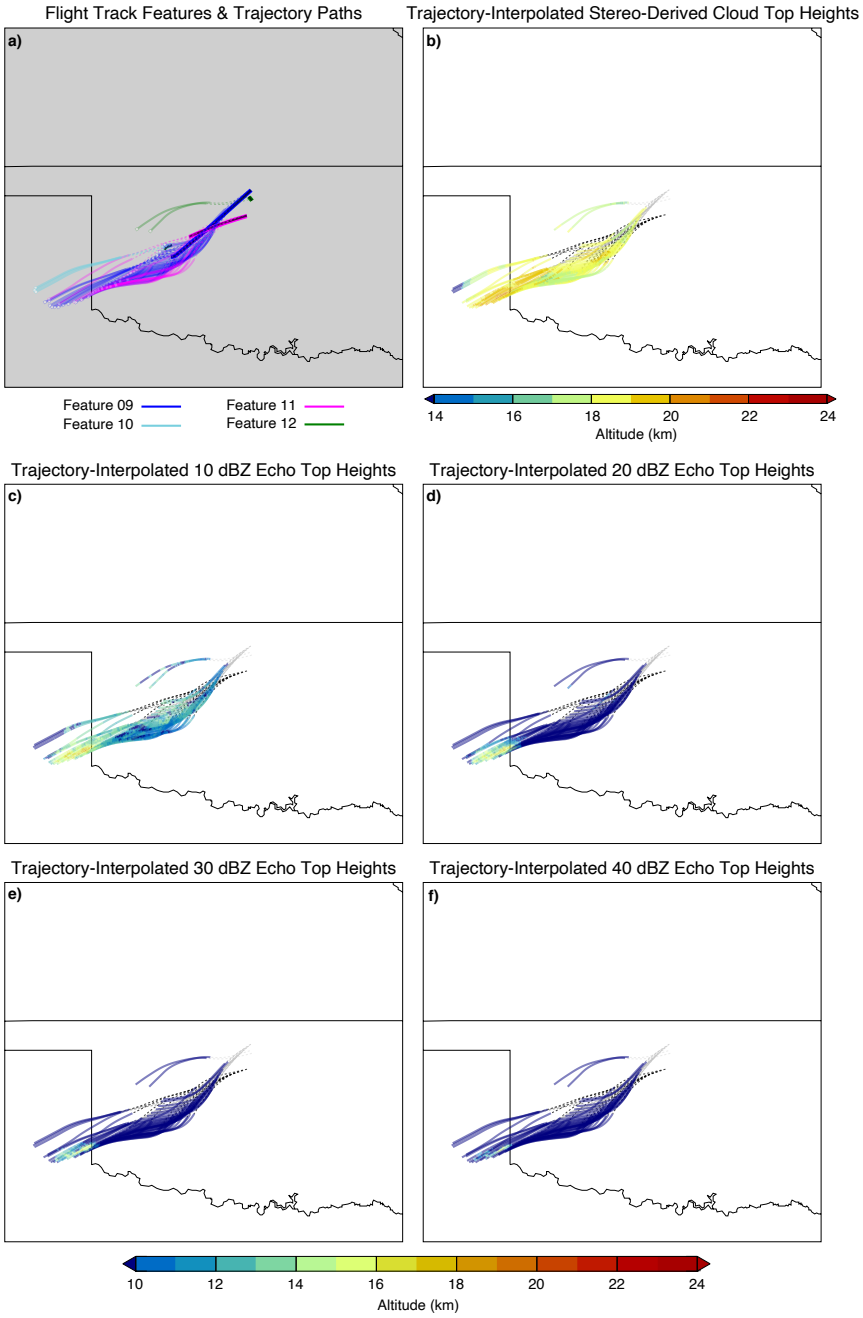

Figure 3.20: Paths for all backward trajectories from initialization time, through 2200 UTC 31 May. In panel (a), each color represents individual water vapor features for which trajectories were initialized; flight track segments of each feature are shown in black and outlined by each respective feature color, with final trajectory locations marked by white boxes. Values of cloud-top altitudes (b), 10 dBZ (c), 20 dBZ (d), 30 dBZ (e), and 40 dBZ (f) GridRad echo top altitudes interpolated to trajectory paths are contoured. Dashed lines indicate portions of trajectory paths prior to 0100 UTC 1 June.

Chapter 4

Conclusions

4.1 Summary of Findings

Study no. 1:

This study developed a new U.S. severe hail climatology based on 23 years of ground-based radar observations and environmental characteristics from reanalysis data in an effort to answer our three research questions: 1) How reliable is an radar-only hail climatology when compared to a report-based method? 2) What additional information (i.e., environmental variables) could improve accuracy of a radar climatology? and 3) What new information can we learn about CONUS severe hail fall from a long-term, well-calibrated, radar-based climatology? Our conclusions are as follows:

1. Hail climatologies based on radar observations alone are insufficient for reproducing the broad characteristics of hail report distributions, and likely also true hail fall distributions as well.

2. MESH configurations developed to represent a large number of reports are more sensitive to marginal severe hail-producing storms. This results in large frequency maxima across the SEUS, particularly during JJA, which likely indicate strong, tropical summer convection that have similar Z_{H} to continental hailstorms, but lack environmental conditions necessary for severe hail occurrence.

3. PWAT and 0-6 km bulk wind shear show the best discrimination between correct identifications and "false" detections of severe hail occurrence.

4. Performing LDA to filter out environments that are not favorable for severe hail occurrence improves the agreement between radar climatology and report climatology.

5. U.S. severe hail climatology based on well-calibrated radar and environmental characteristics shows overall larger frequencies and one single hail frequency maxima through the Plains and Ozarks.

Study no. 2:

This study leveraged hundreds of supercell storms throughout 10 years of recent NEXRAD WSR-88D data, facilitated by the GridRad-Severe record, to answer our two research questions: 1) How useful are bulk environmental parameters for estimating maximum hail size for supercell storms? and 2) Is any additional information gained by leveraging full vertical distributions and storm motions? Our findings of the present study include:

1. Bulk parameters evaluated in this study contribute little-to-no practical utility for determining maximum hail size in supercell storms (Fig. 2.12).

2. Substantial overlap in the full environmental profiles remain, but differences in the hail growth layer in most unstable parcel-environmental temperature difference and storm-relative wind speeds are most evident (Fig. 2.13).

3. The storm-relative, storm-rotated hodograph analysis indicates increasing storm-relative environmental wind speed perpendicular to storm motion above 2 km AGL with increasing hail size (Fig. 2.15b).

Study no. 3:

This study evaluated the UTLS environments and storm characteristics of 89 warm and 89 cold AACP storms to answer two main research questions: 1) Why do some AACPs exhibit a warm feature in IR imagery while others do not? and 2) What observable storm and environment differences exist between warm and cold AACPs? Our conclusions are as follows:

1. Warm AACPs frequently occur in low-tropopause, midlatitude environments, while cold AACPs frequently occur in high-tropopause, tropical environments, thus supporting Hypothesis 2.

2. Stratospheric AACP formation necessitates consistently strong storm-relative winds within the lower stratosphere, which are observed in warm AACP environments but not in cold AACP environments.

3. Storm characteristics (LMD and echo top heights) are largely consistent between warm and cold AACP storms.

Study no. 4:

This study leverages the recently developed stereoscopic cloud-top height retrieval algorithm designed for modern GOES imagery from Young (2021) to answer the following research questions: 1) How do stereoscopic cloud-top height retrievals compare to radar-derived echo top products and in situ water vapor measurements? and 2) What new information can be gained from stereoscopic cloud-top height retrievals regarding AACPs? Our conclusions are as follows:

1. **Existing errors in the stereo algorithm must be corrected in order to accurately assess their representation of observations and to consider potential new insights into AACPs. These stereo results should not**

be used for future reference, and later studies should instead reference the most recent peer-reviewed publication containing the completed study.

2. Stereo-derived cloud-top heights frequently overestimated the observed altitudes of in situ water vapor enhancements by ∼3–4 km.

3. 10 dBZ echo-top heights produced the most consistent altitude range estimates to those of observed water vapor enhancements.

4. AACPs observed by modern GOES VIS imagery can be sufficiently prominent to be differentiated from other storm-top features.

4.2 Hail Occurrence Discussion

The hail studies discussed here have shown that environmental information can be helpful in identifying severe hail occurrence and estimating size. Though various methods exist for using radar observations to identify likely areas of severe hail fall, they alone do not provide enough information to accurately establish a severe hail climatology. $\text{MESH}_{\text{Witt}}$, the commonly used radar-derived parameter, was fit to only 147 reports, which resulted in an unreliable relationship between hail size and SHI (See Fig. 13 of Murillo and Homeyer (2019)). The increased sensitivity of MESH_{75} and MESH_{95} to lower SHI values resulted in greater "false" detection in regions of deep, tropical convection that were removed from the climatology via LDA of environmental parameters. Additionally, our analysis of hail-producing supercells found minor variations in bulk metrics (particularly in MUCAPE; see Fig. 2.12) and in profiles of most unstable parcel−environment temperature differences and storm-relative wind speed (Figs.

2.13d & 2.13e) between hail size categories, similar to prior storm environment analyses. In particular, the greater separation in storm-relative, storm-rotated hodographs is an encouraging finding and differs in nature from the storm-relative wind sensitivities explored in prior work (i.e., increasing storm-relative wind perpendicular to storm motion at low levels or deep-layer wind shear parallel to storm motion; see, e.g., Johnson and Sugden, 2014; Dennis and Kumjian, 2017; Kumjian and Lombardo, 2020). However, substantial differences in deep-layer shear for supercell storms with differing maximum hail size were not found in Study no. 2.

As these studies necessitate the use of hail reports, they are still limited by the need for more reliable and representative size reporting. Unlike the hail-producing supercell study which leveraged population-filtered reports, the climatology developed in Study no. 1 provided hail frequency information where reports did not exist. As such, the uncertainty associated with "false" detection continues to be one of the greatest challenges for objective severe hail identification, limiting the ability to accurately measure the skill of any parameter, given the unknown number of hail occurrences left unreported. However, certain techniques can be adopted to help minimize potential bias, such as LDA, objective storm tracking, filtering hail reports, and excluding earlier, less reliable time periods as was done in Study no. 2.

There are additional challenges associated with radar-derived products, including the variable number of operational radars throughout the time period, potential erroneous radar scans missed during quality control, the unknown time- and space-varying calibration biases of the radars, and the resolution of the radar observations. For example, high MESH values could be spatially overestimated due to non-uniform beam filling of high reflectivity, which was also discussed in Cintineo et al. (2012). However, the overall observational record length, stability in radar data availability since 2006, and reduction in spatial resolution are all aspects of the data that serve to mitigate

these issues. It is worth noting that while the application of Gaussian smoothing to re-gridded hail climatology might introduce additional errors, we found the results were insensitive to such smoothing. Radar errors are also less impactful to the findings from Study no. 2, given that the only radar parameters utilized are storm locations and motions. The radar-derived supercell characteristics are evaluated in the complementary study by Homeyer et al. (2022).

As is typical for storm-environment studies utilizing reanalyses, uncertainty exists around the degree to which these environments are accurately represented in reanalyses. Recently, Taszarek et al. (2021) evaluated how MERRA-2 and ERA5 convective parameters compared to those from observed soundings. While they did find that MERRA-2 was overall less representative than ERA5 analyses, the two parameters used for LDA (precipitable water and 0-6 bulk wind shear) were still well-represented. Future work could instead use ERA5 for LDA to produce an updated U.S. hail climatology with slightly greater confidence. As Taszarek et al. (2021) also found that ERA5 underestimated parcel temperatures over the U.S., with mixed-layer profiles being less erroneous than most unstable profiles, future work could evaluate mixed-layer parcels within spatiotemporal domains where such errors are minimized (see Taszarek et al. (2021) for details). Despite the aforementioned caveats and broad environmental overlap between supercell storms with varying hail size potential, diagnosed hodograph differences from ERA5 near-storm environments can be applied to operations with minimal training. Such an application could provide forecasters with greater confidence for the potential of individual storms to produce increasingly large hail.

The U.S. severe hail climatology established in Study no. 1 produces a solid foundation that can be leveraged for a variety of applications, including its utility as a validation dataset, the ability to evaluate how hail events co-vary with climactic variability, and the opportunity to apply similar LDA evaluations to improve model output-based

hail proxy methods. Although we only utilized single-polarization observations to establish a U.S. severe hail climatology, the findings provide valuable information for the next decade as a sufficient database of dual-polarization observations becomes established. While the new MESH configurations performed well at severe hail sizes, additional work and higher quality reports are needed to produce a more representative sub-severe and significantly severe hail climatology (Wendt and Jirak, 2021; Elmore et al., 2022). A comprehensive environmental evaluation of other hail-producing storm modes (e.g., multicell, quasi-linear convective systems, etc.) could also provide additional insight into environmental controls for hail production, and most importantly, better capture sub-severe hail occurrence, given the frequent severe hail production in supercell storms. Another immediate path could be similar analyses to those conducted or proposed here using radar and environmental data from other countries/regions.

4.3 Overshooting Storm Top Discussion

Analysis of tropopause characteristics in warm and cold AACP storms reveal lower tropopause heights, broad isothermal regions above, and more frequent double tropopauses for warm AACPs compared to higher tropopause heights with inflection point characteristics and infrequent double tropopauses for cold AACPs (Fig. 3.13). Similar storm-relative wind profiles and 10-dBZ echo top altitudes for warm and cold AACP storms, despite different tropopause-relative characteristics, suggest similar storm updrafts within different environments. We have high confidence that these warm AACPs occur in the lower stratosphere, evidenced by ∼95% of warm AACP 10-dBZ echo tops exceeding the tropopause, where sufficient dynamical support for AACP formation exists. Conversely, only ∼50% of cold AACP 10-dBZ echo tops exceed the tropopause, suggesting that sustained AACP production, which is needed for Study no. 3 cases, in

the lower stratosphere is unlikely. Given that there is also a lack of dynamical support (strong storm-relative wind) for gravity wave breaking in the lower stratosphere of cold AACP environments, when tropopause-overshooting did occur, it was likely intermittent and unlikely to produce the observed sustained AACPs in the lower stratosphere (Fig. 3.13). Thus, cold AACPs most likely reside in the upper troposphere (i.e. below the tropopause) where sufficient storm-relative winds are present.

The aforementioned result that warm AACPs reside in the stratosphere and cold AACPs reside in the troposphere has important implications for studies that aim to assess the stratospheric water vapor impact of AACP-producing storms. While it is still possible that Hypothesis 3 is also valid in some (or many) instances, it is likely a minimal contribution to AACP IR characteristics in comparison to Hypothesis 2 given the clear differences in the storm-relative environments in warm and cold AACP storms found in Study no. 3. Although the microphysical characteristics of warm and cold AACPs were not evaluated, the work carried out in Study no. 4, as part of an ongoing effort to investigate observations obtained during DCOTSS field operations, presents a valuable first look at sampled AACP altitude ranges and associated microphysics. Given the unique objectives of and resources utilized for DCOTSS, mentioned in Section 1.2, RF13 was the first flight of its kind, targeting very fresh convective outflow to sample trace gas and particle enhancements relative to the lower stratospheric background. Such operations provide the necessary observations to validate new stereo algorithms for cloud-top height retrievals.

While errors in the stereo algorithm are, at least in part, responsible for the inconsistencies between echo-top heights and stereo-derived cloud-top heights, discrepancy to some extent is expected given the inherently different phenomena represented by each parameter. Echo-top heights can only represent altitudes where precipitation-sized hydrometeors are large enough to be observed by S-band radar, while stereo

products can estimate the vertical extent of clouds consisting of much smaller particles. **Thus, we do expect that stereo-derived heights will be in general higher than echo-top heights, but the erroneous stereo algorithm limits our ability to characterize the expected inconsistency.**

As VIS is currently the foundation for AACP identification, these studies inherently neglect AACPs during nighttime hours. As such, stereo-derived cloud-top heights were only available prior to 0100 UTC 1 June, and with AACP-produced water vapor enhancements observed after 0200 UTC 1 June, it was not possible to compare stereo-derived heights and water vapor enhancements simultaneously. Since the targeted storm collapsed during sunset, however, it is likely that the stereo analyses captured all overshooting/AACP injections relevant to the sampled features. Nonetheless, we do not expect there to be nighttime-dependent warm or cold AACP characteristics, as the tropopause environments lack a substantial dependence on the diurnal cycle. Although, there is still the possibility of missing relevant information about AACPs in general between sunset and sunrise.

Additional limitations are presented by the inclusion of reanalysis environments, particularly in the UTLS. Combined with infrequent sampling above 100 hPa, tropopause-penetrating convection can introduce smaller-scale motions within the lower stratospheric flow that, together, can create higher than expected reanalysis errors in this layer (Hersbach et al., 2020). Such missing dynamics in the reanalysis wind fields is most problematic to trajectory analysis. Insentropic trajectories utilized for storm source attribution in Study no. 4, as opposed to fully 3-dimensional, are typically preferred in the UTLS, where diabatic heating rates are small and pressure vertical coordinates are more easily influenced by small-scale motions. However, whether isentropic or fully 3-dimensional, these trajectory analyses are susceptible to errors in the stratospheric wind, given the greater presence and importance of storm-induced small

scale motions not resolved by the reanalysis (Bowman et al., 2013). It is possible there are "storm-specific" signatures in other trace gas measurements (e.g., surface pollutants) measured by the ER-2 that could provide further evidence as to which storm(s) contributed to each of the water vapor enhancement features. Conversely, it is unlikely that these ERA5 wind errors substantially impact the conclusions on environment differences from Study no. 3, because differences between the two AACP-producing storm populations were only found for storm-relative winds, demonstrating the greater dependence on storm motion than true environmental winds.

Brief evaluations of storm severity in these two studies found that many AACP-producing storms featured severe reports of various types or radar indications of severe weather (hail), consistent with prior work. While some apparent differences in storm severity (hail and wind) were found using reports for warm and cold AACP storm populations, with warm AACP storms being more prolific severe weather producers, independent evaluation of storm characteristics from radar suggest that such differences (at least for hail) are an artifact of the limited report data (compare Table 3.1 and Fig. 3.12). For Study no. 4, spatiotemporal distributions of severe weather reports were broadly coincident with the highest echo-top and stereo-derived cloud top heights, of which most storms produced an AACP at some point during their life (not shown). Nonetheless, a more comprehensive and reliable report database (e.g., as in Blair et al. (2017)) is needed to adequately capture storm severity and meaningful differences between storms, particularly of warm and cold AACP-producing storms.

Finally, future assessments that can utilize the quality-controlled in situ measurements, as those in Study no. 4 are only preliminary (see Section 3.2.3), will increase confidence in the altitude range and water vapor enhancement magnitudes, **especially in conjunction with the corrected stereo algorithm.** Once these improvements/corrections have been completed, stereo analysis of the remaining DCOTSS

flight, RF19 on 24 June 2022 (see Section 1.2), could then be appropriately carried out. Additional confidence in stereo-derived altitudes is possible by combining IR-based stereo analyses with VIS-only methods to produce cloud-top altitudes. Such IR-based stereo algorithms would also provide nighttime cloud-top heights that are impossible with VIS-only stereo methods.

4.4 Looking Forward

Together, the projects summarized here have increased our knowledge of AACPs, developed a variety of severe overshooting convective storm climatologies, and contributed to our understanding of how these storms can influence UTLS composition and climate. The DCOTSS field campaign produced the first documented opportunity to analyze in situ observations of AACPs in conjunction with 1-minute stereoscopy and 5-m radar observations. Additionally, we have demonstrated the challenges related to connecting convectively-influenced air (i.e. AACPs) to it's particular source (even in a relatively isolated event) and to discriminating hail sizes from radar or environmental information in a reliable, broadly applicable manner. As routine dual-polarization radar observations are collected over the next few years, it will soon be possible to assess their utility for improving hail size estimation and climatology. While such a collection would likely provide additional discrimination, severe weather climatologies built on trustworthy, impact-based observations, such as insurance claim data, could be the most valuable for assessing hazard magnitudes, risk, and damage. Although there are many unresolved questions regarding AACPs, hail-producing and supercell storm climatologies established here (and future, more representative iterations) could be adapted to advance our understanding of AACP spatiotemporal variability. Nonetheless, the time-consuming, manual process of AACP identification warrants the development of an automated

detection method that would enable comprehensive evaluation of their characteristics, frequency, and geographic distribution, particularly at night (likely not possible for cold AACPs).

Reference List

Adler, R. F., and D. D. Fenn, 1981: Satellite-observed cloud-top height changes in tornadic thunderstorms. *J. Appl. Meteor. Climatol.*, **20 (11)**, 1369–1375, https://doi.org/10.1175/1520-0450(1981)020h1369:SOCTHCi2.0.CO;2.

Adler, R. F., M. J. Markus, D. D. Fenn, G. Szejwach, and W. E. Shenk, 1983: Thunderstorm top structure observed by aircraft overflights with an infrared radiometer. *J. Appl. Meteor. Climatol.*, **22 (4)**, 579–593, https://doi.org/10.1175/1520-0450(1983)022⟨0579:TTSOBA⟩2.0.CO;2.

Allen, J. T., I. M. Giammanco, M. R. Kumjian, H. J. Punge, Q. Zhang, P. Groenemeijer, M. Kunz, and K. Ortega, 2020: Understanding hail in the earth system. *Rev. Geophys.*, **58 (1)**, e2019RG000 665, https://doi.org/10.1029/2019RG000665.

Allen, J. T., and M. Tippett, 2015: The characteristics of United States hail reports: 1955-2014. *Electronic J. Severe Storms Meteor.*, **10 (3)**, 1–31.

Allen, J. T., M. Tippett, and A. Sobel, 2015a: Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States. *Nat. Geosci.*, **8**, 278–283, https://doi.org/10.1038/ngeo2385.

Allen, J. T., M. K. Tippett, Y. Kaheil, A. H. Sobel, C. Lepore, S. Nong, and A. Muehlbauer, 2017: An extreme value model for U.S. hail size. *Mon. Wea. Rev.*, **145 (11)**, 4501–4519, https://doi.org/10.1175/MWR-D-17-0119.1.

Allen, J. T., M. K. Tippett, and A. H. Sobel, 2015b: An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J. Adv. Model. Earth Syst.*, **7 (1)**, 226–243, https://doi.org/10.1002/2014MS000397.

Anderson, J. G., and Coauthors, 2017: Stratospheric ozone over the united states in summer linked to observations of convection and temperature via chlorine and bromine catalysis. *Proc. Natl. Acad. Sci. (USA)*, **114 (25)**, E4905–E4913, https://doi.org/10.1073/pnas.1619318114.

Atlas, D., and R. Wexler, 1963: Back-scatter by oblate ice spheroids. *J. Atmos. Sci.*, **20**, 48–61, https://doi.org/10.1175/1520-0469(1963)020⟨0048:BSBOIS⟩2.0.CO;2.

Aydin, K., T. A. Seliga, and V. Balaji, 1986: Remote sensing of hail with a dual linear polarization radar. *J. Climate Appl. Meteor.*, **25 (10)**, 1475–1484, https://doi.org/10.1175/1520-0450(1986)025⟨1475:RSOHWA⟩2.0.CO;2.

Añel, J. A., J. C. Antuña, L. de la Torre, J. M. Castanheira, and L. Gimeno, 2008: Climatological features of global multiple tropopause events. *J. Geophys. Res. Atmos.*, **113**, https://doi.org/https://doi.org/10.1029/2007JD009697.

Bang, S. D., and D. J. Cecil, 2019: Constructing a multifrequency passive microwave hail retrieval and climatology in the GPM domain. *J. Appl. Meteor. Climatol.*, **58 (9)**, 1889–1904, https://doi.org/10.1175/JAMC-D-19-0042.1.

Bardsley, W. E., 1990: On the maximum observed hailstone size. *J. Appl. Meteor.*, **29 (11)**, 1185–1187, https://doi.org/10.1175/1520-0450(1990)029⟨1185:OTMOHS⟩2.0.CO;2.

Barrett, B. S., and B. N. Henley, 2015: Intraseasonal variability of hail in the contiguous United States: Relationship to the Madden-Julian Oscillation. *Mon. Wea. Rev.*, **143 (4)**, 1086–1103, https://doi.org/10.1175/MWR-D-14-00257.1.

Bedka, K., J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, 2010: Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients. *J. Appl. Meteor. Climatol.*, **49 (2)**, 181–202, https://doi.org/10.1175/2009JAMC2286.1.

Bedka, K. M., J. T. Allen, H. J. Punge, M. Kunz, and D. Simanovic, 2018a: A long-term overshooting convective cloud-top detection database over Australia derived from MTSAT Japanese Advanced Meteorological Imager Observations. *J. Appl. Meteor. Climatol.*, **57 (4)**, 937–951, https://doi.org/10.1175/JAMC-D-17-0056.1.

Bedka, K. M., E. M. Murillo, C. R. Homeyer, B. Scarino, and H. Mersiovsky, 2018b: The above anvil cirrus plume: An important severe weather indicator in visible and infrared satellite imagery. *Wea. Forecasting*, **33 (5)**, 1159–1181, https://doi.org/10.1175/WAF-D-18-0040.1.

Bedka, K. M., C. Wang, R. Rogers, L. D. Carey, W. Feltz, and J. Kanak, 2015: Examining deep convective cloud evolution using total lightning, wsr-88d, and goes-14 super rapid scan datasets. *Wea. Forecasting*, **30 (3)**, 571–590, https://doi.org/10.1175/WAF-D-14-00062.1.

Berendes, T. A., J. R. Mecikalski, W. M. M. Jr., K. M. Bedka, and U. S. Nair, 2008: Convective cloud identification and classification in daytime satellite imagery using standard deviation limited adaptive clustering. *J. Geophys. Res. Atmos.*, **113 (D20)**, https://doi.org/10.1029/2008JD010287.

Blair, S. F., D. R. Deroche, J. M. Boustead, J. W. Leighton, B. L. Barjenbruch, and W. P. Gargan, 2011: A radar-based assessment of the detectability of giant hail. *Electronic J. Operational Meteor.*, **6 (7)**, 1–30, https://doi.org/10.55599/ejssm.v6i7.34.

Blair, S. F., and J. W. Leighton, 2012: Creating high-resolution hail datasets using social media and post-storm ground surveys. *Electronic J. Operational Meteor.*, **13**, 32–45.

Blair, S. F., and Coauthors, 2017: High-resolution hail observations: Implications for NWS warning operations. *Wea. Forecasting*, **32 (3)**, 1101–1119, https://doi.org/10.1175/WAF-D-16-0203.1.

Bohren, C. F., and L. J. Battan, 1982: Radar backscattering of microwaves by spongy ice spheres. *J. Atmos. Sci.*, **39**, 2623–2628, https://doi.org/10.1175/1520-0469(1982)039⟨2623:RBOMBS⟩2.0.CO;2.

Boustead, J. M., 2008: Using maximum storm-top divergence and the vertical freezing level to forecast hail size. Conf. on Severe Local Storms, Savannah, GA, P6.6, URL https://ams.confex.com/ams/pdfpapers/142145.pdf.

Bowman, K. P., 1993: Large-scale isentropic mixing properties of the antarctic polar vortex from analyzed winds. *J. Geophys. Res.*, **98 (D12)**, 23 013—-23 027, https://doi.org/10.1029/93JD02599.

Bowman, K. P., and G. D. Carrie, 2002: The mean-meridional transport circulation of the troposphere in an idealized GCM. *J. Atmos. Sci.*, **59 (9)**, 1502–1514, https://doi.org/10.1175/1520-0469(2002)059⟨1502:TMMTCO⟩2.0.CO;2.

Bowman, K. P., and C. R. Homeyer, 2017: GridRad - Three-Dimensional Gridded NEXRAD WSR-88D Radar Data. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, http://rda.ucar.edu/datasets/ds841.0/, https://doi.org/10.5065/D6NK3CR7.

Bowman, K. P., J. C. Lin, A. Stohl, R. Draxler, P. Konopka, A. Andrews, and D. Brunner, 2013: Input data requirements for lagrangian trajectory models. *Bull. Amer. Meteor. Soc.*, **94 (7)**, 1051–1058, https://doi.org/10.1175/BAMS-D-12-00076.1.

Brooks, H. E., C. A. Doswell, and M. P. Kay, 2003: Climatological estimates of local daily tornado probability for the United States. *Wea. Forecasting*, **18 (4)**, 626–640, https://doi.org/10.1175/1520-0434(2003)018⟨0626:CEOLDT⟩2.0.CO;2.

Bruick, Z. S., K. L. Rasmussen, and D. J. Cecil, 2019: Subtropical South American hailstorm characteristics and environments. *Mon. Wea. Rev.*, **147 (12)**, 4289–4304, https://doi.org/10.1175/MWR-D-19-0011.1.

Brunner, J. C., S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin, 2007: A quantitative analysis of the enhanced-V feature in relation to severe weather. *Wea. Forecasting*, **22 (4)**, 853–872, https://doi.org/10.1175/WAF1022.1.

Bunkers, M. J., 2018: Observations of right-moving supercell motion forecast errors. *Wea. Forecasting*, **33**, 145–159, https://doi.org/10.1175/WAF-D-17-0133.1.

Bunkers, M. J., and P. L. Smith, 2013: Comments on "An Objective High-Resolution Hail Climatology of the Contiguous United States". *Wea. Forecasting*, **28 (3)**, 915–917, https://doi.org/10.1175/WAF-D-13-00020.1.

Carletta, N. D., G. L. Mullendore, M. Starzec, B. Xi, Z. Feng, and X. Dong, 2016: Determining the best method for estimating the observed level of maximum detrainment based on radar reflectivity. *Mon. Wea. Rev.*, **144**, 2915–2926, https://doi.org/10.1175/MWR-D-15-0427.1.

Cecil, D. J., 2009: Passive microwave brightness temperatures as proxies for hailstorms. *J. Appl. Meteor. Climatol.*, **48 (6)**, 1281–1286, https://doi.org/10.1175/2009JAMC2125.1.

Cecil, D. J., and C. B. Blankenship, 2012: Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *J. Climate*, **25**, 687–703, https://doi.org/10.1175/JCLI-D-11-00130.1.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2018: Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H49C6VHW. Accessed 8 September 2021.

Changnon, S. A., 1999: Data and approaches for determining hail risk in the contiguous United States. *J. Appl. Meteor.*, **38 (12)**, 1730–1739, https://doi.org/10.1175/1520-0450(1999)038⟨1730:DAAFDH⟩2.0.CO;2.

Cheng, L., M. English, and R. Wong, 1985: Hailstone size distributions and their relationship to storm thermodynamics. *J. Climate Appl. Meteor.*, **24 (10)**, 1059–1067, https://doi.org/10.1175/1520-0450(1985)024⟨1059:HSDATR⟩2.0.CO;2.

Childs, S. J., and R. S. Schumacher, 2019: An updated severe hail and tornado climatology for Eastern Colorado. *J. Appl. Meteor. Climatol.*, **58 (10)**, 2273–2293, https://doi.org/10.1175/JAMC-D-19-0098.1.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea. Forecasting*, **27 (5)**, 1235–1248, https://doi.org/10.1175/WAF-D-11-00151.1.

Cooney, J. W., K. P. Bowman, C. R. Homeyer, and T. M. Fenske, 2018: Ten year analysis of tropopause-overshooting convection using gridrad data. *J. Geophys. Res. Atmos.*, **123 (1)**, 329—-343, https://doi.org/10.1002/2017JD027718.

Dennis, E. J., and M. R. Kumjian, 2017: The impact of vertical wind shear on hail growth in simulates supercells. *J. Atmos. Sci.*, **74 (3)**, 641–663, https://doi.org/10.1175/JAS-D-16-0066.1.

Donaldson Jr., R. J., 1958: Analysis of severe convective storms observed by radar. *J. Atmos. Sci.*, **15**, 44–50, https://doi.org/10.1175/1520-0469(1958)015⟨0044:AOSCSO⟩2.0.CO;2.

Donaldson Jr., R. J., 1959: Analysis of severe convective storms observed by radar – II. *J. Atmos. Sci.*, **16**, 281–287, https://doi.org/10.1175/1520-0469(1959)016⟨0281:AOSCSO⟩2.0.CO;2.

Doswell, C. A. I., H. E. Brooks, and M. P. Kay, 2005: Climatological estimates of daily local nontornadic severe thunderstorm probability for the United States. *Wea. Forecasting*, **20 (4)**, 577–595, https://doi.org/10.1175/WAF866.1.

Duda, J. D., and W. A. G. Jr., 2010: Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Wea. Forecasting*, **25 (1)**, 190–206, https://doi.org/10.1175/2009WAF2222338.1.

Dye, J. E., and B. E. Martner, 1978: The relationship between radar reflectivity factor and hail at the ground for northeast colorado thunderstorms. *J. Appl. Meteor. Climatol.*, **17**, 1335–1341, https://doi.org/10.1175/1520-0450(1978)017⟨1335:TRBRRF⟩2.0.CO;2.

Edwards, R., and R. L. Thompson, 1998: Nationwide comparisons of hail size with wsr-88d vertically integrated liquid water and derived thermodynamic sounding data. *Wea. Forecasting*, **13 (2)**, 277–285, https://doi.org/10.1175/1520-0434(1998)013⟨0277:NCOHSW⟩2.0.CO;2.

Elmore, K. L., J. T. Allen, and A. E. Gerard, 2022: Sub-severe and severe hail. *Wea. Forecasting*, **37 (8)**, 1357–1369, https://doi.org/10.1175/WAF-D-21-0156.1.

Elmore, K. L., Z. L. Flamig, V. Lakshmanan, B. T. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz, 2014: MPING: Crowd-sourcing weather reports for research. *Bull. Amer. Meteor. Soc.*, **95 (9)**, 1335–1342, https://doi.org/10.1175/BAMS-D-13-00014.1.

Ferraro, R., J. Beauchamp, D. Cecil, and G. Heymsfield, 2015: A prototype hail detection algorithm and hail climatology developed with the advanced microwave sounding unit (AMSU). *Atmos. Res.*, **163**, 24–35, https://doi.org/https://doi.org/10.1016/j.atmosres.2014.08.010.

for Medium-Range Weather Forecasts, E. C., 2019: ERA5 reanalysis (0.25 degree latitude-longitude grid). Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder CO, URL https://doi.org/10.5065/BH6N-5N20.

Fraile, R., A. Castro, and J. Sánchez, 1992: Analysis of hailstone size distributions from a hailpad network. *Atmos. Res.*, **28 (3)**, 311–326, https://doi.org/https://doi.org/10.1016/0169-8095(92)90015-3.

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

Fujita, T. T., 1982: Principle of stereographic height computations and their application to stratospheric cirrus over severe thunderstorms. *J. Meteor. SocJapan*, **60**, 355–368, https://doi.org/https://doi.org/10.2151/jmsj1965.60.1_355.

Gaines, S. E., S. W. Bowen, R. S. Hipskind, T. P. Bui, and K. R. Chan, 1992: Comparisons of the NASA ER-2 meteorological measurement system with radar tracking and radiosonde data. *J. Atmos. Oceanic Technol.*, **9**, 210–225, https://doi.org/10.1175/1520-0426(1992)009⟨0210:COTNEM⟩2.0.CO;2.

Gelaro, R., and Coauthors, 2017: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Climate*, **30 (14)**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gensini, V. A., C. Converse, W. S. Ashley, and M. Taszarek, 2021: Machine learning classification of significant tornadoes and hail in the united states using ERA5 proximity soundings. *Wea. Forecasting*, **36**, 2143–2160, https://doi.org/10.1175/WAF-D-21-0056.1.

Gensini, V. A., and M. K. Tippett, 2019: Global ensemble forecast system (GEFS) predictions of days 1-15 U.S. tornado and hail frequencies. *Geophys. Res. Lett.*, **46 (5)**, 2922–2930, https://doi.org/10.1029/2018GL081724.

Geotis, S. G., 1963: Some radar measurements of hailstorms. *J. Appl. Meteor. Climatol.*, **2**, 270–275, https://doi.org/10.1175/1520-0450(1963)002⟨0270:SRMOH⟩2.0.CO;2.

Gettelman, A., P. Hoor, L. L. Pan, W. J. Randel, M. I. Hegglin, and T. Birner, 2011: The extratropical upper troposphere and lower stratosphere. *Rev. Geophys.*, **49**, RG3003, https://doi.org/https://doi.org/10.1029/2011RG000355.

Griffin, S. M., K. M. Bedka, and C. S. Velden, 2016: A method for calculating the height of overshooting convective cloud tops using satellite-based IR imager and CloudSAT cloud profiling radar observations. *J. Appl. Meteor. Climatol.*, **55 (2)**, 479–491, https://doi.org/10.1175/JAMC-D-15-0170.1.

Gutierrez, R. E., and M. R. Kumjian, 2021: Environmental and radar characteristics of gargantuan hail-producing storms. *Mon. Wea. Rev.*, **149**, 2523–2538, https://doi.org/10.1175/MWR-D-20-0298.1.

Hasler, A. F., 1981: Stereographic observations from geosynchronous satellites: An important new tool for the atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **62 (2)**, 194–212, https://doi.org/10.1175/1520-0477(1981)062h0194:SOFGSAi2.0.CO;2.

Hasler, A. F., R. Mack, and A. Negri, 1983: Stereoscopic observations from meteorological satellites. *Advances in Space Research*, **2 (6)**, 105–113, https://doi.org/10.1016/0273-1177(82)90130-2.

Hasler, A. F., J. Strong, R. H. Woodward, and H. Pierce, 1991: Automatic analysis of stereoscopic satellite image pairs for determination of cloud-top height and structure. *J. Appl. Meteor. Climatol.*, **30 (3)**, 257–281, https://doi.org/10.1175/1520-0450(1991)030h0257:AAOSSIi2.0.CO;2.

Herman, B. M., and L. J. Battan, 1961: Calculations of mie back-scattering from melting ice spheres. *J. Atmos. Sci.*, **18**, 468–478, https://doi.org/10.1175/1520-0469(1961)018⟨0468:COMBSF⟩2.0.CO;2.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146 (730)**, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803.

Heymsfield, A. J., I. M. Giammanco, and R. Wright, 2014: Terminal velocities and kinetic energies of natural hailstones. *Geophys. Res. Lett.*, **41 (23)**, 8666–8672, https://doi.org/10.1002/2014GL062324.

Heymsfield, G. M., R. H. Blackmer, and S. Schotz, 1983: Upper-level structure of oklahoma tornadic storms on 2 may 1979. i: Radar and satellite observations. *J. Atmos. Sci.*, **40**, 1740–1755, https://doi.org/10.1175/1520-0469(1983)040⟨1740:ULSOOT⟩2.0.CO;2.

Hoffmann, L., and R. Spang, 2021: An assessment of tropopause characteristics of the era5 and era-interim meteorological reanalyses. *Atmos. Chem. Phys.*, **2021**, 1–44, https://doi.org/10.5194/acp-2021-961.

Hoinka, K. P., 1998: Statistics of the global tropopause pressure. *Mon. Wea. Rev.*, **126**, 3303–3325, https://doi.org/10.1175/1520-0493(1998)126⟨3303:SOTGTP⟩2.0.CO;2.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33 (4)**, 403–439, https://doi.org/10.1029/95RG02097.

Homeyer, C. R., 2014: Formation of the enhanced-V infrared cloud-top feature from high-resolution three-dimensional radar observations. *J. Atmos. Sci.*, **71 (1)**, 332–348, https://doi.org/10.1175/JAS-D-13-079.1.

Homeyer, C. R., and K. P. Bowman, 2017: Algorithm Description Document for Version 3.1 of the Three-Dimensional Gridded NEXRAD WSR-88D Radar (GridRad) Dataset. Available online at: http://gridrad.org/pdf/GridRad-v3.1-Algorithm-Description.pdf.

Homeyer, C. R., and K. P. Bowman, 2022: Algorithm Description Document for Version 4.2 of the Three-Dimensional Gridded NEXRAD WSR-88D Radar (GridRad) Dataset. Technical Report.

Homeyer, C. R., J. D. McAuliffe, and K. M. Bedka, 2017: On the development of above-anvil cirrus plumes in extratropical convection. *J. Atmos. Sci.*, **74 (5)**, 1617–1633, https://doi.org/10.1175/JAS-D-16-0269.1.

Homeyer, C. R., E. Murillo, and M. R. Kumjian, 2022: Relationships between 10 years of radar-observed supercell characteristics and hail potential. to be submitted in Mon. Wea. Rev.

Homeyer, C. R., T. N. Sandmæl, and C. K. Potvin, 2020: Distinguishing characteristics of tornadic and nontornadic supercell storms from composite mean analyses of radar observations. *Mon. Wea. Rev.*, **148**, 5015–5040, https://doi.org/10.1175/MWR-D-20-0136.1.

Jeong, J.-H., J. Fan, C. R. Homeyer, and Z. Hou, 2020: Understanding hailstone temporal variability and contributing factors over the U.S. Southern Great Plains. *J. Climate*, **33**, 3947–3966, https://doi.org/10.1175/JCLI-D-19-0606.1.

Jiang, Z., M. R. Kumjian, R. S. Schrom, I. Giammanco, T. Brown-Giammanco, H. Estes, R. Maiden, and A. J. Heymsfield, 2019: Comparisons of electromagnetic scattering properties of real hailstones and spheroids. *J. Appl. Meteor. Climatol.*, **58**, 93–112, https://doi.org/10.1175/JAMC-D-17-0344.1.

Johnson, A., and K. Sugden, 2014: Evaluation of sounding-derived thermodynamic and wind-related parameters associated with large hail events. *Electronic J. Severe Storms Meteor.*, **9 (5)**, 1–42.

Kelly, D. L., J. T. Schaefer, and C. A. Doswell, 1985: Climatology of nontornadic severe thunderstorm events in the United States. *Mon. Wea. Rev.*, **113 (11)**, 1997–2014, https://doi.org/10.1175/1520-0493(1985)113⟨1997:CONSTE⟩2.0.CO;2.

Kikuchi, K., and T. Kasai, 1968: Stereoscopic analysis of photographs taken by nimbus II APT system. *J. Meteor. Soc. Japan*, **46 (1)**, 60–67, https://doi.org/10.2151/jmsj1965.46.160.

Kumjian, M. R., Z. J. Lebo, and A. M. Ward, 2019: Storms producing large accumulations of small hail. *J. Appl. Meteor. Climatol.*, **58**, 341–364, https://doi.org/10.1175/JAMC-D-18-0073.1.

Kumjian, M. R., and K. Lombardo, 2020: A hail growth trajectory model for exploring the environmental controls on hail size: Model physics and idealized tests. *J. Atmos. Sci.*, **77**, 2765–2791, https://doi.org/10.1175/JAS-D-20-0016.1.

Kumjian, M. R., K. Lombardo, and S. Loeffler, 2021: The evolution of hail production in simulated supercell storms. *J. Atmos. Sci.*, **78**, 3417–3440, https://doi.org/10.1175/JAS-D-21-0034.1.

Kumjian, M. R., Y. P. Richardson, T. Meyer, K. A. Kosiba, and J. Wurman, 2018: Resonance scattering effects in wet hail observed with a dual-x-band-frequency, dual-polarization doppler on wheels radar. *J. Appl. Meteor. Climatol.*, **57**, 2713–2731, https://doi.org/10.1175/JAMC-D-17-0362.1.

Kunz, M., U. Blahak, J. Handwerker, M. Schmidberger, H. J. Punge, S. Mohr, E. Fluck, and K. M. Bedka, 2018: The severe hailstorm in southwest germany on 28 july 2013: characteristics, impacts and meteorological conditions. *Quart. J. Roy. Meteor. Soc.*, **144 (710)**, 231–250, https://doi.org/10.1002/qj.3197.

Lagerquist, R., A. McGovern, C. R. Homeyer, D. J. G. II, and T. Smith, 2020: Deep learning on three-dimensional multiscale data for next-hour tornado prediction. *Mon. Wea. Rev.*, **148**, 2837–2861, https://doi.org/10.1175/MWR-D-19-0372.1.

Lee, J., and D.-B. Shin, 2021: Algorithm for improved stereoscopic cloud-top height retrieval based on visible and infrared bands for Himawari-8 and FY-4A. *Remote. Sens.*, **13**, 4993, https://doi.org/10.3390/rs13244993.

Lee, J., D. B. Shin, C. Y. Chung, and J. Kim, 2020: A cloud top-height retrieval algorithm using simultaneous observations from the Himawari-8 and FY-2e satellites. *Remote. Sens.*, **12 (12)**, https://doi.org/10.3390/rs12121953.

Levizzani, V., and M. Setvák, 1996: Multispectral, high-resolution satellite observations of plumes on top of convective storms. *J. Atmos. Sci.*, **53**, 361–369, https://doi.org/10.1175/1520-0469(1996)053⟨0361:MHRSOO⟩2.0.CO;2.

Li, W., Y. bin Yuan, Y.-J. Chai, Y.-A. Liou, J. kun Ou, and S. ming Zhong, 2017: Characteristics of the global thermal tropopause derived from multiple radio occultation measurements. *Atmos. Res.*, **185**, 142–157, https://doi.org/https://doi.org/10.1016/j.atmosres.2016.09.013.

Lin, Y., and M. R. Kumjian, 2022: Influences of CAPE on hail production in simulated supercell storms. *J. Atmos. Sci.*, **79**, 179–204, https://doi.org/10.1175/JAS-D-21-0054.1.

Lindsey, D. T., and L. Grasso, 2008: An effective radius retrieval for thick ice clouds using GOES. *J. Appl. Meteor. Climatol.*, **47 (4)**, 1222–1231, https://doi.org/10.1175/2007JAMC1612.1.

Lindsey, D. T., D. W. Hillger, L. Grasso, J. A. Knaff, and J. F. Dostalek, 2006: Goes climatology and analysis of thunderstorms with enhanced 3.9-μm reflectivity. *Mon. Wea. Rev.*, **134 (9)**, 2342–2353, https://doi.org/10.1175/MWR3211.1.

Liu, B., J. Huo, D. Lyu, and X. Wang, 2021: Assessment of FY-4A and Himawari-8 cloud top height retrieval through comparison with ground-based millimeter radar at sites in Tibet and Beijing. *Advances in Atmospheric Sciences*, **38**, 1334–1350, https://doi.org/10.1007/s00376-021-0337-2.

Luderer, G., J. Trentmann, K. Hungershöfer, M. Herzog, M. Fromm, and M. O. Andreae, 2007: Small-scale mixing processes enhancing troposphere-to-stratosphere transport by pyro-cumulonimbus storms. *Atmos. Chem. Phys.*, **7**, 5945–5957, https://doi.org/10.5194/acp-7-5945-2007.

Mack, R. A., A. F. Hasler, and R. F. Adler, 1983: Thunderstorm cloud top observations using satellite stereoscopy. *Mon. Wea. Rev.*, **111**, 1949–1964, https://doi.org/10.1175/1520-0493(1983)111⟨1949:TCTOUS⟩2.0.CO;2.

Manney, G., and Coauthors, 2017: Reanalysis comparisons of upper tropospheric–lower stratospheric jets and multiple tropopauses. *Atmos. Chem. Phys.*, **17**, 11 541–11 566, https://doi.org/10.5194/acp-17-11541-2017.

Mather, G. K., D. Treddenick, and R. Parsons, 1976: An observed relationship between the height of the 45 dBZ contours in storm profiles and surface hail reports. *J. Appl. Meteor.*, **15 (12)**, 1336–1340, https://doi.org/10.1175/1520-0450(1976)015⟨1336:AORBTH⟩2.0.CO;2.

McCann, D. W., 1983: The enhanced-V: A satellite observable severe storm signature. *Mon. Wea. Rev.*, **111 (4)**, 887–894, https://doi.org/10.1175/1520-0493(1983)111⟨0887:TEVASO⟩2.0.CO;2.

Mecikalski, J. R., T. N. Sandmæl, E. M. Murillo, C. R. Homeyer, K. M. Bedka, J. M. Apke, and C. P. Jewett, 2021: A random-forest model to assess predictor importance and nowcast severe storms using high-resolution radar–goes satellite–lightning observations. *Mon. Wea. Rev.*, **149**, 1725–1746, https://doi.org/10.1175/MWR-D-19-0274.1.

Merucci, L., K. Zaksek, E. Carboni, and S. Corradini, 2016: Stereoscopic estimation of volcanic ash cloud-top height from two geostationary satellites. *Remote. Sens.*, **8**, 206, https://doi.org/10.3390/rs8030206.

Minzner, R. A., W. E. Shenk, R. D. Teagle, and J. Steranka, 1978: Stereographic cloud heights from imagery of SMS/GOES satellites. *Geophys. Res. Lett.*, **5 (1)**, 21–24, https://doi.org/10.1029/GL005i001p00021.

Mohr, S., M. Kunz, and B. Geyer, 2015a: Hail potential in Europe based on a regional climate model hindcast. *Geophys. Res. Lett.*, **42 (24)**, 10–904, https://doi.org/10.1002/2015GL067118.

Mohr, S., M. Kunz, and K. Keuler, 2015b: Development and application of a logistic model to estimate the past and future hail potential in Germany. *J. Geophys. Res. Atmos.*, **120 (9)**, 3939–3956, https://doi.org/10.1002/2014JD022959.

Mroz, K., A. Battaglia, T. J. Lang, D. J. Cecil, S. Tanelli, and F. Tridon, 2017: Hail-detection algorithm for the GPM core observatory satellite sensors. *J. Appl. Meteor. Climatol.*, **56 (7)**, 1939–1957, https://doi.org/10.1175/JAMC-D-16-0368.1.

Mroz, K., A. Battaglia, T. J. Lang, S. Tanelli, and G. F. Sacco, 2018: Global precipitation measuring dual-frequency precipitation radar observations of hailstorm vertical structure: Current capabilities and drawbacks. *J. Appl. Meteor. Climatol.*, **57 (9)**, 2161–2178, https://doi.org/10.1175/JAMC-D-18-0020.1.

Mullendore, G. L., D. R. Durran, and J. R. Holton, 2005: Cross-tropopause tracer transport in midlatitude convection. *J. Geophys. Res. Atmos.*, **110 (D6)**, https://doi.org/https://doi.org/10.1029/2004JD005059.

Mullendore, G. L., A. J. Homann, K. Bevers, and C. Schumacher, 2009: Radar reflectivity as a proxy for convective mass transport. *J. Geophys. Res. Atmos.*, **114 (D16)**, https://doi.org/https://doi.org/10.1029/2008JD011431.

Murillo, E. M., and C. R. Homeyer, 2019: Severe hail fall and hailstorm detection using remote sensing observations. *J. Appl. Meteor. Climatol.*, **58 (5)**, 947–970, https://doi.org/10.1175/JAMC-D-18-0247.1.

Murillo, E. M., and C. R. Homeyer, 2022: What determines above-anvil cirrus plume infrared temperature? *J. Atmos. Sci.*, https://doi.org/10.1175/JAS-D-22-0080.1.

Murillo, E. M., C. R. Homeyer, and J. T. Allen, 2021: A 23-year severe hail climatology using gridrad mesh observations. *Mon. Wea. Rev.*, **149**, 945–958, https://doi.org/10.1175/MWR-D-20-0178.1.

Murphy, A. M., C. R. Homeyer, and K. Q. Allen, 2022: Development and investigation of gridrad-severe, a multi-year severe event radar dataset. to be submitted in Mon. Wea. Rev.

NCEI/NOAA, 2022: NOAA's Storm Events Database. National Centers for Environmental Information, URL https://www.ncdc.noaa.gov/stormevents/.

Nelson, S. P., 1983: The influence of storm flow structure on hail growth. *J. Atmos. Sci.*, **40**, 1965–1983, https://doi.org/10.1175/1520-0469(1983)040⟨1965:TIOSFS⟩2.0.CO;2.

Ni, X., C. Liu, D. J. Cecil, and Q. Zhang, 2017: On the detection of hail using satellite passive microwave radiometers and precipitation radar. *J. Appl. Meteor. Climatol.*, **56 (10)**, 2693–2709, https://doi.org/10.1175/JAMC-D-17-0065.1.

Nisi, L., A. Hering, U. Germann, and O. Martius, 2016a: A 15-year hail streak climatology for the Alpine region. *Quart. J. Roy. Meteor. Soc.*, **144 (714)**, 1429–1449, https://doi.org/10.1002/qj.3286.

Nisi, L., O. Martius, A. Hering, M. Kunz, and U. Germann, 2016b: Spatial and temporal distribution of hailstorms in the Alpine region: a long-term, high resolution, radar-based analysis. *Quart. J. Roy. Meteor. Soc.*, **142 (697)**, 1590–1604, https://doi.org/10.1002/qj.2771.

NOAA, 1994: NOAA geostationary operational environmental satellite imager data. NOAAcomprehensive large array data stewardship system, https://www.avl.class.noaa.gov/saa/products/welcome.

NOAA/NWS/SPC, 1955: Severe Weather Database. NOAA/NWS/SPC, https://www.spc.noaa.gov/wcm/.

OFCM, 2005: Federal Meterological Handbook No. 11 – Doppler Radar Meteorological Observations, Part B: Doppler Radar Theory and Meteorology. FCM-H11B-2005 (Available online at http://www.ofcm.gov/publications/fmh/allfmh2.htm).

on the NASA ER-2 Aircraft, T. M. M. S., 1990: Stan g. scott and t. paul bui and k. roland chan and stuart w. bowen. *J. Atmos. Oceanic Technol.*, **7**, 525–540, https://doi.org/10.1175/1520-0426(1990)007⟨0525:TMMSOT⟩2.0.CO;2.

Ondrejka, R. J., and J. H. Conover, 1966: Note on the stereo interpretation of nimbus II APT photography. *Mon. Wea. Rev.*, **94 (10)**, 611–614, https://doi.org/10.1175/1520-0493(1966)094h0611:NOTSIOi2.3.CO;2.

Ortega, K., 2018: Evaluating multi-radar, multi-sensor products for surface hailfall diagnosis. *Electronic J. Severe Storms Meteor.*, **13 (1)**, URL http://www.ejssm.org/ojs/index.php/ejssm/article/view/163/113.

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: The severe hazards analysis and verification experiment. *Bull. Amer. Meteor. Soc.*, **90 (10)**, 1519–1530, https://doi.org/10.1175/2009BAMS2815.1.

O'Neill, M. E., L. Orf, G. M. Heymsfield, and K. Halbert, 2021: Hydraulic jump dynamics above supercell thunderstorms. *Science*, **373 (6560)**, 1248–1251, https://doi.org/10.1126/science.abh3857.

Pan, L. L., S. B. Honomichl, T. V. Bui, T. Thornberry, A. Rollins, E. Hintsa, and E. J. Jensen, 2018: Lapse rate or cold point: The tropical tropopause identified by in situ trace gas measurements. *Geophys. Res. Lett.*, **45 (19)**, 10,756–10,763, https://doi.org/https://doi.org/10.1029/2018GL079573.

Parker, M. D., 2014: Composite VORTEX2 supercell environments from near-storm soundings. *Mon. Wea. Rev.*, **142**, 508–529, https://doi.org/10.1175/MWR-D-13-00167.1.

Peters, J. M., C. J. Nowotarski, and H. Morrison, 2019: The role of vertical wind shear in modulating maximum supercell updraft velocities. *J. Atmos. Sci.*, **76**, 3169–3189, https://doi.org/10.1175/JAS-D-19-0096.1.

Prata, F., and M. J. Lynch, 2019: Passive earth observations of volcanic clouds in the atmosphere. *Atmosphere*, **10**, 199, https://doi.org/10.3390/atmos10040199.

Prein, A. F., and G. J. Holland, 2018: Global estimates of damaging hail hazard. *Weather Clim. Extremes*, **22**, 10–23, https://doi.org/https://doi.org/10.1016/j.wace.2018.10.004.

Punge, H. J., K. M. Bedka, and A. Reinbold, 2017: Hail frequency estimation across Europe based on a combination of overshooting top detections and the ERA-INTERIM reanalysis. *Atmos. Res.*, **198**, 34–43, https://doi.org/https://doi.org/10.1016/j.atmosres.2017.07.025.

Púčik, T., M. Valachová, and P. Zacharov, 2013: Upper tropospheric conditions in relation to the cloud top features of 15 august 2010 convective storms. *Atmos. Res.*, **123**, 249–267, https://doi.org/https://doi.org/10.1016/j.atmosres.2012.10.006.

Randel, W. J., D. J. Seidel, and L. L. Pan, 2007: Observational characteristics of double tropopauses. *J. Geophys. Res. Atmos.*, **112**, https://doi.org/https://doi.org/10.1029/2006JD007904.

Rosenfeld, D., W. L. Woodley, A. Lerner, G. Kelman, and D. T. Lindsey, 2008: Satellite detection of severe convective storms by their retrieved vertical profiles of cloud particle effective radius and thermodynamic phase. *J. Geophys. Res. Atmos.*, **113 (D4)**, https://doi.org/https://doi.org/10.1029/2007JD008600.

Sandmæl, T. N., 2017: An evaluation of radar- and satellite-data based products to discriminate between tornadic and non-tornadic storms. M.S. thesis, School of Meteorology, University of Oklahoma, 98 pp., https://hdl.handle.net/11244/52775.

Sandmæl, T. N., C. R. Homeyer, K. M. Bedka, J. M. Apke, J. R. Mecikalski, and K. Khlopenkov, 2019: Evaluating the ability of remote sensing observations to identify significantly severe and potentially tornadic storms. *J. Appl. Meteorol. Climatol.*, **58 (12)**, 2569–2590, https://doi.org/10.1175/JAMC-D-18-0241.1.

Sargent, M. R., and Coauthors, 2013: A new direct absorption tunable diode laser spectrometer for high precision measurement of water vapor in the upper troposphere and lower stratosphere. *Rev. Sci. Instrum*, **84 (7)**, 074 102, https://doi.org/10.1063/1.4815828.

Schaefer, J., J. J. Levit, S. J. Weiss, and D. W. McCarthy, 2004: The frequency of large hail over the contiguous United States. *14th Conf. on Applied Climatology*, Seattle, WA, Amer. Meteor. Soc., 3.3, [Available online at https://ams.confex.com/ams/pdfpapers/69834.pdf].

Schlie, E. E.-J., D. Wuebbles, S. Stevens, R. Trapp, and B. Jewett, 2018: A radar-based study of severe hail outbreaks over the contiguous United States for 2000-2011. *Int. J. Climatol.*, **39 (1)**, 278–291, https://doi.org/10.1002/joc.5805.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, https://doi.org/https://doi.org/10.1175/BAMS-D-15-00230.1.

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

School of Meteorology/University of Oklahoma, 2021: GridRad-Severe - Three-Dimensional Gridded NEXRAD WSR-88D Radar Data for Severe Events. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO, http://dx.doi.org/10.5065/2B46-1A97.

Seguchi, T., S. Iwasaki, M. Kamogawa, T. Ushiyama, and H. Okamoto, 2019: Observation of jumping cirrus with ground-based cameras, radiosonde, and himawari-8. *J. Meteor. Soc. Japan*, https://doi.org/10.2151/JMSJ.2019-033.

Seiz, G., S. Tjemkes, and P. Watts, 2007: Multiview cloud-top height and wind retrieval with photogrammetric methods: Application to meteosat-8 hrv observations. *J. Appl. Meteor. Climatol.*, **46 (8)**, 1182–1195, https://doi.org/10.1175/JAM2532.1.

Setvák, M., K. M. Bedka, D. T. Lindsey, A. Sokol, Z. Charvát, J. Šťástka, and P. K. Wang, 2013: A-train observations of deep convective storm tops. *Atmos. Res.*, **123**, 229–248, https://doi.org/https://doi.org/10.1016/j.atmosres.2012.06.020.

Setvák, M., and C. A. Doswell III, 1991: The AVHRR channel 3 cloud top reflectivity of convective storms. *Mon. Wea. Rev.*, **119 (3)**, 841–847, https://doi.org/10.1175/1520-0493(1991)119⟨0841:TACCTR⟩2.0.CO;2.

Setvák, M., and Coauthors, 2010: Satellite-observed cold-ring-shaped features atop deep convective clouds. *Atmos. Res.*, **97 (1)**, 80–96, https://doi.org/https://doi.org/10.1016/j.atmosres.2010.03.009.

Shenk, W. E., R. J. Holub, and R. A. Neff, 1975: Stereographic cloud analysis from Apollo 6 photographs over a cold front. *Bull. Amer. Meteor. Soc.*, **56 (1)**, 4–16, https://doi.org/10.1175/1520-0477-56.1.4.

Shou, Y.-X., F. Lu, H. Liu, P. Cui, S. W. Shou, and J. Liu, 2019: Satellite-based observational study of the Tibetan Plateau Vortex: Features of deep convective cloud tops. *Adv. Atmos. Sci.*, **36 (2)**, 189–205, https://doi.org/10.1007/s00376-018-8049-y.

Smith, B. T., R. L. Thompson, J. S. Grams, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous united states. part i: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Smith, J. B., 2012: The sources and significance of stratospheric water vapor: Mechanistic studies from equator to pole. Ph.D. thesis, Harvard University.

Smith, J. B., 2021: Convective hydration of the stratosphere. *Science*, **373 (6560)**, 1194–1195, https://doi.org/10.1126/science.abl8740.

Smith, J. B., and Coauthors, 2017: A case study of convectively sourced water vapor observed in the overworld stratosphere over the united states. *J. Geophys. Res. Atmos.*, **122 (17)**, 9529–9554, https://doi.org/10.1002/2017JD026831.

Solomon, D. L., K. P. Bowman, and C. R. Homeyer, 2016: Tropopause-penetrating convection from three-dimensional gridded nexrad data. *J. Appl. Meteor. Climatol.*, **55 (2)**, 465–478, https://doi.org/10.1175/JAMC-D-15-0190.1.

Starzec, M., C. R. Homeyer, and G. L. Mullendore, 2017: Storm labeling in three dimensions (SL3D): A volumetric radar echo and dual-polarization updraft classification algorithm. *Mon. Wea. Rev.*, **145 (3)**, 1127–1145, https://doi.org/10.1175/MWR-D-16-0089.1.

Starzec, M., G. L. Mullendore, and C. R. Homeyer, 2020: Retrievals of convective detrainment heights using ground-based radar observations. *J. Geophys. Res. Atmos.*, **125 (5)**, e2019JD031 164, https://doi.org/https://doi.org/10.1029/2019JD031164.

Stohl, A., and Coauthors, 2003: Stratosphere-troposphere exchange: A review, and what we have learned from STACCATO. *J. Geophys. Res. Atmos.*, **108 (D12)**, https://doi.org/10.1029/2002JD002490.

Tang, B., V. A. Gensini, and C. R. Homeyer, 2019: Trends in United States large hail environments and observations. *npj Clim. Atmos. Sci.*, **2 (45)**, https://doi.org/10.1038/s41612-019-0103-7.

Taszarek, M., J. T. Allen, T. Pucik, K. A. Hoogewind, and H. E. Brooks, 2020: Severe convective storms across europe and the united states. part II: ERA5 environments associated with lightning, large hail, severe wind, and tornadoes. *J. Clim.*, **33**, 10 263–10 286, https://doi.org/10.1175/JCLI-D-20-0346.1.

Taszarek, M., J. T. Allen, T. Pucik, K. A. Hoogewind, and H. E. Brooks, 2021: Severe convective storms across europe and the united states. part II: ERA5 environments associated with lightning, large hail, severe wind, and tornadoes. *J. Clim.*, **34 (8)**, 3211–3237, https://doi.org/10.1175/JCLI-D-20-0484.1.

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the rapid update cycle. *Wea. Forecasting*, **18**, 1243–1261, https://doi.org/10.1175/1520-0434(2003)018⟨1243:CPSWSE⟩2.0.CO;2.

Trapp, R. J., D. M. Wheatley, N. T. Atkins, R. W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21 (3)**, 408–415, https://doi.org/10.1175/WAF925.1.

Waldvogel, A., B. Federer, and P. Grimm, 1979: Criteria for the detection of hail cells. *J. Appl. Meteor.*, **18 (12)**, 1521–1525, https://doi.org/10.1175/1520-0450(1979)018⟨1521:CFTDOH⟩2.0.CO;2.

Wang, P. K., 2003: Moisture plumes above thunderstorm anvils and their contributions to cross-tropopause transport of water vapor in midlatitudes. *J. Geophys. Res.*, **108 (D6)**, https://doi.org/https://doi.org/10.1029/2002JD002581.

Wang, P. K., 2007: The thermodynamic structure atop a penetrating convective thunderstorm. *Atmos. Res.*, **83 (2)**, 254–262, https://doi.org/https://doi.org/10.1016/j.atmosres.2005.08.010.

Wang, P. K., K.-Y. Cheng, M. Setvák, and C.-K. Wang, 2016: The origin of the gullwing-shaped cirrus above an argentinian thunderstorm as seen in CALIPSO images. *J. Geophys. Res. Atmos.*, **121 (7)**, 3729–3738, https://doi.org/https://doi.org/10.1002/2015JD024111.

Warren, R. A., H. A. Ramsay, S. T. Siems, M. J. Manton, J. R. Peter, A. Protat, and A. Pillalamarri, 2020: Radar-based climatology of damaging hailstorms in brisbane and Sydney, Australia. *Quart. J. Roy. Meteor. Soc.*, **146 (726)**, 505–530, https://doi.org/https://doi.org/10.1002/qj.3693.

Weinstock, E. M., and Coauthors, 1994: New fast response photofragment fluorescence hygrometer for use on the NASA ER-2 and the perseus remotely piloted aircraft. *Rev. Sci. Instrum*, **65 (11)**, 3544–3554, https://doi.org/10.1063/1.1144536.

Weinstock, E. M., and Coauthors, 2009: Validation of the harvard lyman-α in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor. *J. Geophys. Res. Atmos.*, **114 (D23)**, https://doi.org/https://doi.org/10.1029/2009JD012427.

Wendt, N. A., and I. L. Jirak, 2021: An hourly climatology of operational MRMS MESH-diagnosed severe and significant hail with comparisons to *Storm Data* hail reports. *Wea. Forecasting*, **36**, 645–659, https://doi.org/10.1175/WAF-D-20-0158.1.

Wilhelmsen, H., F. Ladstadter, T. Schmidt, and A. K. Steiner, 2020: Double tropopauses and the tropical belt connected to enso. *Geophys. Res. Lett.*, **47**, e2020GL089 027, https://doi.org/10.1029/2020GL089027.

Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences*, International Geophysics Series, Vol. 59. 2nd ed., Academic Press, 627 pp.

Witt, A., D. Burgess, A. Seimon, J. T. Allen, J. Snyder, and H. Bluestein, 2018: Rapid-scan radar observations of an Oklahoma tornadic hailstorm producing giant hail. *Wea. Forecasting*, **0 (0)**, null, https://doi.org/10.1175/WAF-D-18-0003.1.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998a: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13 (2)**, 286–303, https://doi.org/10.1175/1520-0434(1998)013⟨0286:AEHDAF⟩2.0.CO;2.

Witt, A., M. D. Eilts, G. J. Stumpf, E. D. W. Mitchell, J. T. Johnson, and K. W. Thomas, 1998b: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13 (2)**, 513–518, https://doi.org/10.1175/1520-0434(1998)013⟨0513:ETPOWS⟩2.0.CO;2.

Witt, A., and S. P. Nelson, 1991: The use of single-doppler radar for estimating maximum hailstone size. *J. Appl. Meteorol. Climatol.*, **30**, 425–431, https://doi.org/10.1175/1520-0450(1991)030⟨0425:TUOSDR⟩2.0.CO;2.

World Meteorological Organization, 1957: Meteorology –a three-dimensional science: Second session of the commission for aerology. *WMO Bull.*, **IV (5)**, 134–138.

Wylie, D. P., D. A. Santek, and D. Starr, 1998: Cloud-top heights from GOES-8 and GOES-9 stereoscopic imagery. *J. Appl. Meteor.*, **37**, 405–413, https://doi.org/10.1175/1520-0450(1998)037⟨0405:CTHFGA⟩2.0.CO;2.

Xian, T., and C. R. Homeyer, 2019: Global tropopause altitudes in radiosondes and reanalyses. *Atmos. Chem. Phys.*, **19**, 5661–5678, https://doi.org/10.5194/acp-19-5661-2019.

Young, T., 2021: GOES-R stereoscopic cloud top retrieval algorithm using normalized cross-correlation for 1-min overshoot analysis. M.S. thesis, School of Meteorology, University of Oklahoma, 59 pp., https://hdl.handle.net/11244/330215.

Zou, L., L. Hoffmann, S. Griessbach, R. Spang, and L. Wang, 2021: Empirical evidence for deep convection being a major source of stratospheric ice clouds over north america. *Atmos. Chem. Phys.*, **21**, 10 457–10 475, https://doi.org/10.5194/acp-21-10457-2021.