

Trajectory-Enhanced AIRS Observations of Environmental Factors Driving Severe Convective Storms

PETER KALMUS AND BRIAN H. KAHN

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

SEAN W. FREEMAN AND SUSAN C. VAN DEN HEEVER

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 16 February 2018, in final form 6 February 2019)

ABSTRACT

We investigate environmental factors of severe convective weather using temperature and moisture retrievals from the Atmospheric Infrared Sounder (AIRS) that lie along parcel trajectories traced from tornado, large hail, and severe wind producing events in the central United States. We create AIRS proximity soundings representative of the storm environment by calculating back trajectories from storm times and locations at levels throughout the troposphere, using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model forced with the 32-km North American Regional Reanalysis (NARR) and 12-km North American Mesoscale Forecast System (NAM12). The proximity soundings are calculated for severe weather events including tornadoes, hail ≥ 2 in. diameter, and wind gusts >65 mph (29 m s^{-1}) specified in the NCEI Storm Events database. Box-and-whisker diagrams exhibit more realistic values of enhanced convective available potential energy (CAPE) and suppressed convective inhibition (CIN) relative to conventional “nearest neighbor” (NN) soundings; however, differences in lifting condensation level (LCL), level of free convection (LFC), and significant tornado parameter (STP) from the HYSPLIT-adjusted back traced soundings are more similar to NN soundings. This methodology should be extended to larger swaths of soundings, and to other operational infrared sounders, to characterize the large-scale environment in severe convective events.

1. Introduction

There is an extensive literature that describes proximity soundings of temperature T , humidity, vertical wind shear, and derived convective indices from radiosondes or model analyses (e.g., Brooks et al. 1994; Rasmussen and Blanchard 1998; Doswell and Evans 2003; Thompson et al. 2003, 2012; and references therein). Brooks et al. (1994) highlight the importance of low-level moisture in discriminating types of convection, and that temporally and spatially variable low-level moisture fields are more common than quasi-uniform moisture fields. Doswell and Evans (2003) describe proximity soundings for tornadic and nontornadic supercells and the environmental differences among them, finding (cf. their Fig. 3) higher relative humidity (RH) in the boundary layer and lower troposphere for

supercells with significant tornadoes, compared to supercells with either weak tornadoes or no tornadoes. The statistical properties of the environment obtained from proximity soundings can be ambiguous because of time and space mismatches to convective storms (Potvin et al. 2010).

The processes contributing to the tornado life cycle within supercells and other types of convective storms remain enigmatic (Markowski and Richardson 2009). Advances in our theoretical (e.g., Davies-Jones et al. 2001) and observational (e.g., Wakimoto and Cai 2000; Markowski et al. 2002; Wurman et al. 2007) understanding, and the numerical representation of convection (e.g., Rotunno and Klemp 1985; Rotunno et al. 1988; Weisman et al. 2008) and of tornado-like vortices (e.g., Grasso and Cotton 1995; Wicker and Wilhelmson 1995; Markowski et al. 2003; Lerach and Cotton 2012; Coffey and Parker 2017) illustrate the importance of a rapidly evolving thermodynamic environment in proximity to convective storms.

Corresponding author: Peter Kalmus, peter.m.kalmus@jpl.nasa.gov

DOI: 10.1175/MWR-D-18-0055.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

Large networks of surface-based mesoscale observations and twice-daily radiosonde launches at 0000 and 1200 UTC over the continental United States are unable to resolve time-varying and spatial-scale-dependent gradients of T and specific humidity q ; radiosonde launches during morning and late evening miss the “prime time” of convective initialization (e.g., [Kain et al. 2008](#)). As a result, T and q are frequently poorly represented in model analyses used in operations and research. This may be true in coarsely gridded analyses such as the 40-km resolution Rapid Update Cycle (RUC-2), the finely gridded 3-km resolution variational version of the Local Analysis and Prediction System (vLAPS; [Jiang et al. 2015](#)), and 3-km High-Resolution Rapid Refresh (HRRR; [Benjamin et al. 2016](#)). [Koch et al. \(2016\)](#) show that very strong spatial gradients and tendencies in moisture, stability, and cloudiness were responsible for an unusual EF3 tornado near Windsor, Colorado, on 22 May 2008. A notably improved representation of these features was observed in vLAPS compared to the RUC-2.

There is a clear need for a dense time- and space-resolved network of sounding observations throughout the troposphere. [Dabberdt et al. \(2005\)](#) specifically call out the potential capabilities of hyper-spectral infrared sounders such as the Atmospheric Infrared Sounder (AIRS, with 45 km footprints and twice-daily overpasses at 0130 and 1330 local time; [Chahine et al. 2006](#)) to meet some of these needs. [Feltz and Mecikalski \(2002\)](#) use retrievals of T and q obtained from the upward-looking Atmospheric Emitted Radiance Interferometer (AERI) at the Southern Great Plains (SGP) Department of Energy (DOE) site, with a temporal resolution of 10 min, to show a rapid evolution of convective parameters in proximity to severe convective weather. Substantial differences between convective available potential energy (CAPE) and convective inhibition (CIN) among three AERI instrument sites, situated approximately in an equilateral triangle with sides of length 200 km, illustrate the inherent variability of T and q in the prestorm convective environment. [Wagner et al. \(2008\)](#) show the time evolution of CAPE and CIN obtained from AERI retrievals of T and q in tornadic versus nontornadic environments. CAPE is shown to be larger in proximity to tornadic convective storms. However, CIN is also reported to be larger in tornadic convective events.

While the upward-looking infrared AERI observations complement radiosonde observations, and provide excellent temporal resolution, their limitations—namely, poor spatial coverage, fixed locations, and inability to probe above 3 km—suggest that downward-looking satellite-based hyper-spectral infrared sounders may

have potential to characterize the local storm environment. The dense spatial sampling of T and q vertical profiles from AIRS during the afternoon observation time provides a unique opportunity to assess the developing storm environment a few hours before convective initiation, complementing a glaring temporal gap in the operational radiosonde record. [Botes et al. \(2012\)](#) calculated atmospheric stability indices from AIRS, RUC, and radiosondes matched to nearest neighbors in space and within two hours of each other, to probe preconvective environments for convective events occurring no more than six hours after the AIRS overpass. CAPE and CIN calculated using AIRS profiles were found to be generally underestimated relative to CAPE and CIN calculated using the radiosonde measurements. [Gartzke et al. \(2017\)](#) showed that the correlation of CAPE between AIRS and radiosondes launched at the SGP ARM site is $r = 0.34$; however, if the AIRS retrieved surface temperature is replaced with a surface observation before CAPE is calculated, the comparison is vastly improved with $r = 0.95$.

[Jones and Stensrud \(2012\)](#) showed that the assimilation of AIRS thermodynamic profiles led to improvements in the representation of dewpoint profiles in proximity to convective features and convective initiation within a numerical model using radar as ground truth for validation. This shows that the assimilation of larger-scale environmental thermodynamic variability that is resolved by AIRS can have marked improvements in the 1–4-h time frame at the smaller convective scales. However, the optimal locations and spatial density of soundings, and their proximity to boundaries and thermodynamic gradients, remain highly uncertain. [Coniglio et al. \(2016\)](#) performed data assimilation of radiosonde observations launched during the Mesoscale Predictability Experiment (MPEX) to show that improvements in convective forecasts using WRF-ARW depended on optimal positioning of the radiosondes. One implication is that a dense network of satellite soundings may be useful because the most relevant locations for data assimilation can be customized on a case-by-case basis. Furthermore, a higher density of time-resolved soundings is beneficial for observing an evolving preconvective environment. [Weisz et al. \(2015\)](#) used soundings from AIRS, the Infrared Sounding Atmospheric Interferometer (IASI), and the Cross-track Infrared Sounder (CrIS) to show that increased low-level moistening and decreased lifted index (LI) preceded the 20 May 2013 tornado outbreak near Moore, Oklahoma.

The promising early results from surface-based AERI observations, the successful predictive power of AIRS data assimilation in a numerical prediction of convection by [Jones and Stensrud \(2012\)](#), and the obvious

shortcomings from a lack of temporal resolution despite multiple polar-orbiting hyperspectral infrared sounders, motivate a novel approach for exploiting AIRS as proximity soundings for severe convection. While not as precise as conventional radiosondes, there are potentially orders of magnitude more AIRS soundings available. We describe a combined approach that uses the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT; Stein et al. 2015) model to search for the AIRS pixels that best match in time and space the National Centers for Environmental Information (NCEI) Storm Events database. We find this methodology enhances the interpretation of time evolution of severe convective storm environments. We also speculate that this approach may be useful in an operational “nowcasting” context and could offer additional guidance to forecasters in the 1–4-h time frame.

The rest of this paper is organized as follows. Section 2 describes the methodology. Section 3 presents the results, including ensemble studies of thermodynamic variables relevant to convective environments and case studies of individual storms. Section 4 provides additional context using sensitivity studies. Section 5 summarizes, concludes, and provides a detailed outlook on next steps in future research and applications including those in operational forecasting environments.

2. Data and methods

The latitudes, longitudes, times, and classification of tornadoes, hail, and wind events from severe thunderstorms over the continental United States are obtained from NCEI’s Storm Events database between 2003 and 2016.

Level 2 (L2) swath data of T and q profiles from the AIRS instrument on board the *Aqua* satellite (Chahine et al. 2006), launched on 4 May 2002, are obtained for all days and locations that spatially and temporally coincide with the NCEI Storm Events. The AIRS Level 2 Support products contain T and q profiles that are sampled vertically at a spacing of about 26 hPa or 250 m (Maddy and Barnet 2008) and are archived starting in September 2002. While the vertical grid resolution is high, the information content of the retrievals implies retrieval resolution on the order of ~ 2 km, depending on the vertical structure of T and q . The AIRS retrievals incorporate information from the Advanced Microwave Sounding Unit (AMSU) instrument collocated with AIRS on board the *Aqua* satellite.

The longitudinal range of the analysis is limited to between 83° and 110°W, focusing on the plains east of the U.S. Rockies. This region of the United States contains the highest frequencies of severe weather reports (Smith

et al. 2012) and is east of the highly variable topography of the Rockies that are much more problematic for lower-tropospheric infrared sounding. We divide the storm events into six categories: EF5, EF4, and EF3 tornadoes (“TOR2”); EF2, EF1, and EF0 tornadoes (“TOR1”); hail with diameter $d \geq 3$ in. (“HAIL2”); hail with $2 \leq d < 3$ in. (“HAIL1”); thunderstorm winds with gust speeds ≥ 100 miles per hour (mph) (“WND2”); and thunderstorm winds with speeds between 65 and 100 mph ($29\text{--}45 \text{ m s}^{-1}$) (“WND1”). Wind events that are reported without speeds are not included in this investigation. The wind events in Storm Data should not be treated as precise estimates of wind speed as there are many challenges to using these data quantitatively (Trapp et al. 2006). Wind gusts are estimated from human observers, observed damage, and anemometers; each of these observational methodologies has limitations in precision, accuracy, and sampling. Furthermore, many severe wind gusts go unreported or do not lead to damage. In this investigation, we include two wind categories for the strongest wind gusts to explore differences with EF-scale and hail categories.

Ideally the analysis would account for each tornadic supercell, nontornadic supercell, and nonsupercellular thunderstorm within our spatiotemporal domain separately. However, this would require knowing the underlying mapping between the storm events and the actual storm modes. For example, a single supercell may last for several hours, or evolve into another mode and affect multiple municipalities, resulting in a one-to-many mapping in the database. For another example, if a single supercell triggers both tornado and hail reports, these will remain as separate entries in the NCEI database. We henceforth make no attempt to discriminate by convective mode and will group convective indices by the type of severe weather reported. Every storm report within the spatiotemporal domain is used in the analysis, except for the cases of EF0 and WND1 storm reports, for which every other storm report is used due to their high frequency of occurrence. Some degree of overlap between the defined bins within each severe weather category is likely to contribute to reduced discrimination of convective indices between the bins. The possibility of co-occurrence of different events (e.g., hail and tornado) remains a potential impact on the results.

The HYSPLIT model is forced with winds using the North American Regional Reanalysis (NARR, with horizontal resolution of 32 km, temporal resolution of 3 h, and 24 pressure levels) and the North American Mesoscale Forecast System (NAM12; available from 2008, with resolution of 12 km, temporal resolution of 1 h, and 26 pressure levels) to calculate four-dimensional back trajectories from a storm report’s

TABLE 1. Sequential attrition of storm reports within the analysis pipeline as described in the text, from an input of 26 993 initial storm reports, for the NARR winds over the entire AIRS record.

Reason for exclusion	Percent of input
Storm before overpass	28
HYSPLIT lost backtrace level(s)	10
No NN matchup	30
No matchup for at least one level	8
Poor data quality, at least one level	15
CAPE = 0, NN or backtrace	3

beginning latitude, beginning longitude, and beginning time at one of 40 starting altitude levels evenly spaced between 50 m and 13.5 km above model ground level. This enables matching of back trajectories to the nearest AIRS retrieved sounding within the *Aqua* overpass at approximately 1330 local time. The use of both NARR and NAM12 inputs into HYSPLIT is intended to provide two independent estimates of the wind fields (speed and direction) that may arise from differences in model initialization, data assimilation, model subgrid physics, model resolution, and model output reported at different temporal frequencies. Back trajectories are run for storm reports with beginning times later than the AIRS overpass but occurring that same day. There is no consideration of the end time of storm reports. Eliminating reports earlier than 1330 LT cuts about a quarter of the total number of storm reports (Table 1).

Before the back trajectories are matched to AIRS soundings, the storm report's beginning time and location are matched to the nearest AIRS footprint; this AIRS footprint is referred to as the nearest neighbor (NN) matchup. As per convention in NN analyses (e.g., Botes et al. 2012) the storm report time is required to be no more than 6 h after the overpass (Fig. 1), and to reside

within an AIRS footprint. In approximately a third of cases, no such NN matchup exists (Table 1). To facilitate comparison between the conventional NN method and the new backtrace method described below, these cases are excluded from the analysis. In future analyses or operational deployments of this methodology relying only on the backtrace, such losses from lack of NN matchups would not be incurred.

Next, back trajectories are matched to AIRS soundings. Our HYSPLIT runs produce back-trajectory estimates at the top of any given hour. The back-trajectory locations for matching up to the AIRS overpass are chosen by linearly interpolating the backtraced latitude, longitude, and altitude to the NN matchup time (i.e., the time of the AIRS overpass near the storm location). For back-trajectory AIRS matchups, the back-trajectory location is required to reside within an AIRS footprint and to be within 30 min of the overpass time; in practice, the time interpolation ensures that AIRS and backtrace matchups match to within one minute. For events with fewer than 40 backtrace matchups, that is, if some altitudes fall outside of the AIRS swath, the storm is excluded from the analysis (Table 1). This occurs in ~8% of cases.

Soundings of temperature and specific humidity are then created from the level-by-level matched AIRS retrievals. Every datum of a backtrace profile, both T and q , must have an AIRS data quality of either 0 (best) or 1 (good) (e.g., Yue et al. 2013) or else that storm is excluded from the analysis, which occurs in 15% of cases (Table 1). Up to two quality-control marked pressure levels for each sounding are reported, one for 0 (best) and the other for 1 (good), and demarcate the highest pressure (lowest altitude) for which levels can fall within "best" or "good" categories. Most individual soundings are typically only categorized as "best," "good," or "do not use" (all levels below "good"). Since the backtrace

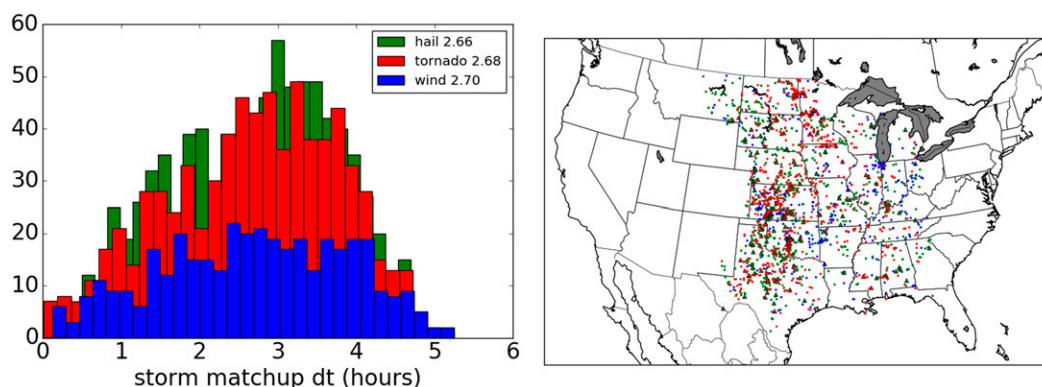


FIG. 1. (left) Histogram of storm beginning time minus AIRS overpass time for nearest neighbor (NN) matchups remaining in the NARR analysis (dt): grouped as hail storms (green), tornadic storms (red), and thunderstorm wind events (blue). Mean dt is given in the legend for each group. (right) Map of these storms, with large triangles denoting TOR2, HAI2 categories, and WND2 categories; and dots denoting the remaining categories.

methodology by design uses multiple adjacent soundings, the quality control flagging can be highly variable within a backtrace profile. Altitudes corresponding to pressure levels are obtained from HYSPLIT for backtrace profiles, and from the AIRS geopotential height product for NN profiles.

Backtraced parcels in general experience vertical motions as they converge to the space–time location of the storm. The parcel temperatures are adjusted along both dry ($RH < 100\%$) and moist ($RH \geq 100\%$ if ascending) adiabats as appropriate, and the lifting condensation level (LCL) separating the dry and moist adjustment regimes is calculated using the method of Romps (2017). The value of specific humidity corresponding to saturation at each parcel's ending altitude and adjusted temperature is calculated, and the saturation-adjusted q value is used if less than the parcel's starting q value. This saturation adjustment occurs infrequently and was found to have a negligible effect on results. Phase changes and precipitation are assumed to have no further impact on the parcel mixing ratios (e.g., Doswell 1987). Presumably parcels that undergo these changes are typically masked out in the AIRS data quality control, since infrared sounders cannot penetrate opaque cloud. Errors from this assumption will be investigated in future work. Finally, for parity in subsequent calculations, the NN T and q profiles are logarithmically interpolated to the backtrace profile pressure levels. Figure 2 illustrates the backtrace methodology for a single storm in the analysis, the EF5 Joplin, Missouri, tornado (section 3).

For each surviving storm report (Fig. 1) we thus have T and q profiles from the NN AIRS retrieval, as well as profiles reconstructed from AIRS retrievals at each of the 40 backtraced locations and levels. From these T and q profiles, the SHARPPy python package (Blumberg et al. 2017) is used to calculate thermodynamic quantities of interest: CAPE, CIN, LCL, and LFC from most unstable (MU) parcels. For backtraced profiles, shear is calculated using the compass bearing direction from the backtraced matchup location to the storm location, and a mean wind speed calculated from the distance between these two locations and the difference in their timestamps. Occasionally CAPE is calculated to be zero; when this occurs for either the NN profile or the backtrace profile, the corresponding storm report is excluded from the analysis (Table 1).

3. Results

a. Two contrasting supercells

We demonstrate the methodology on two well-known and consequential supercells, one tornadic and one

nontornadic, in the context of their respective environments. The first supercell produced a destructive EF5 tornado in Joplin, Missouri, on 22 May 2011 that led to numerous fatalities (Karstens et al. 2013). Back trajectories and q and T profiles are shown for the Joplin supercell in Fig. 2. The direction of the back trajectories veers with height indicating strong vertical wind shear. Enhancement of low-level moisture is apparent in the backtrace profile relative to the NN profile. The brightness temperature at 1231 cm^{-1} (T_{b1231}), q at 925 and 850 hPa, T at 850 and 700 hPa, and the 925–700 hPa temperature lapse rate ($\Gamma_{925-700}$) are shown in Fig. 3. The q at 925 and 850 hPa show a few approximately north–south-oriented bands of high values of q that appear to be moistening the low levels of the backtrace sounding in the few hours leading up to the tornado. In the case of T , the east to west gradient of increasing T at 850 hPa is stronger than at 700 hPa which implies more rapid warming at low levels than aloft. The $\Gamma_{925-700}$ also shows about 2 K km^{-1} of steepening to the south and west of the Joplin supercell, which is indicative of an increasingly unstable boundary layer with time.

The second example supercell produced 4+ in. diameter hail in and around Oklahoma City, Oklahoma, on 16 May 2010. The T_{b1231} , 925 and 700 hPa q , 925 and 850 hPa T , and $\Gamma_{925-700}$ are shown in Fig. 4. Very strong and complex horizontal and vertical gradients in q are observed. A strong moistening from the west at 700 hPa simultaneously occurred along with moistening from the east at 925 hPa. At 600 hPa, there is a sharp reduction in q to the south of the supercell, while moist air is found to the north (not shown). The vertical structure of q exhibits a moist layer up to 700 hPa and an extensive dry layer above it. The T gradients in the horizontal were strongest outside of the region containing the back trajectories, however, some warming was observed at 850 hPa with some slight cooling at 925 hPa that implies a slight stabilization with time. The $\Gamma_{925-700}$ map shows a stabilization in southeast Oklahoma near the region where the supercell eventually dissipated.

The horizontal gradients in lapse rates, and the time tendencies implied by the back trajectories, are substantially different between Figs. 3 and 4. Multiple bands of high q in Fig. 3 and height-dependent horizontal gradients of q in Fig. 4 illustrate AIRS' ability to distinguish case-by-case differences in the near-storm environment. These two cases uniquely depicted by AIRS illustrate near-storm environmental differences that are consistent with the severe weather generated by the two supercells.

b. Environmental convective indices

The storm dataset is now examined as an ensemble divided into the six storm categories previously defined.

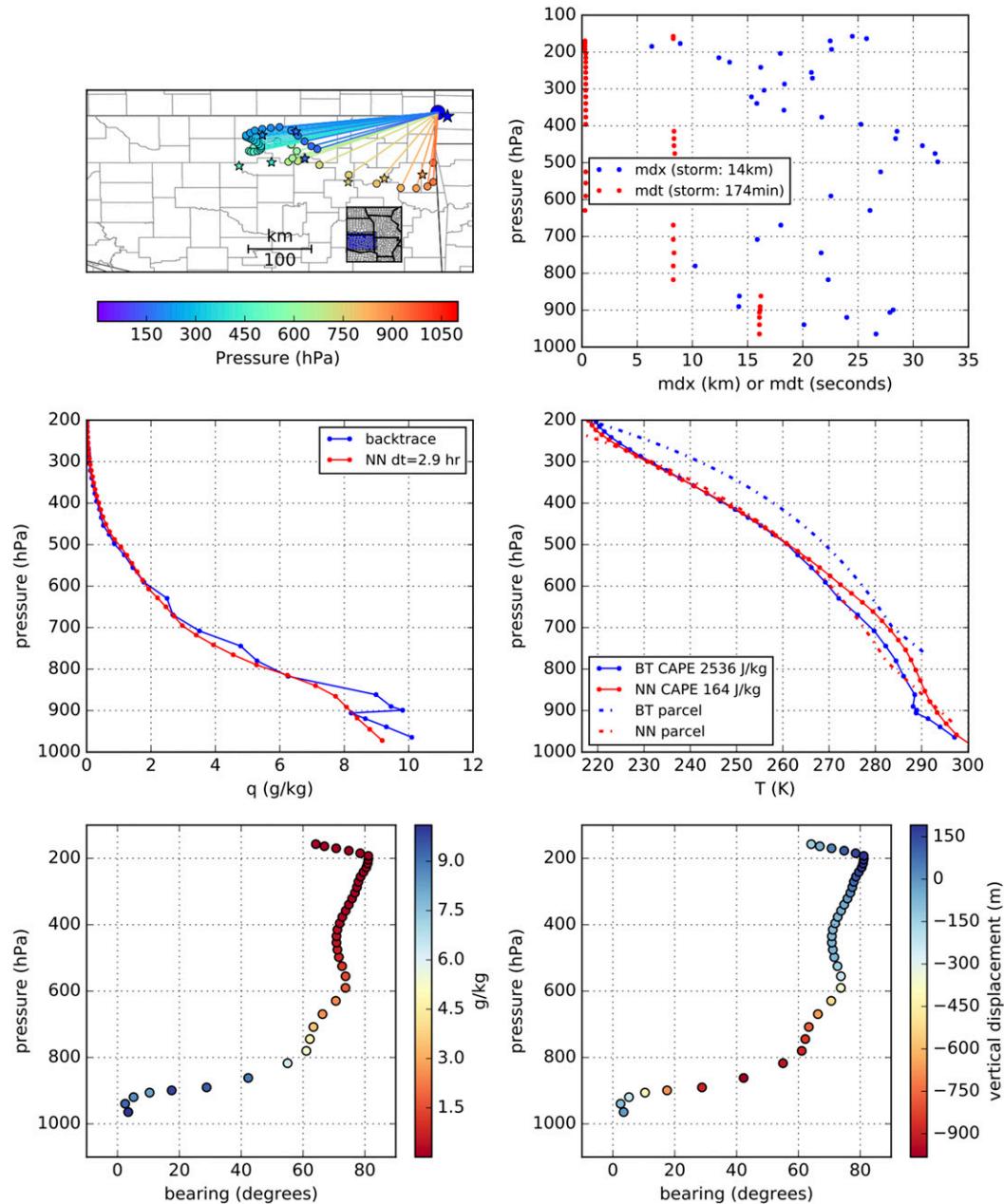


FIG. 2. Case study illustrating the backtrace methodology, in this case 6 aspects of the Joplin EF5 tornadic supercell (see section 3). (top left) Back trajectories using the NAM12 winds (blue 150 hPa to orange 1000 hPa) and AIRS matchups (stars). The length scale indicates 100 km. (top right) Time and space differences [matchup time difference in seconds (mdt) in red and matchup spatial distance in km (mdx) in blue] for each matchup. (middle left) Specific humidity (q) profiles for backtrace and nearest neighbor. (middle right) As in (middle left), but for T , and including most unstable parcel lifting profiles (dashed lines). (bottom left) q vs compass bearing and pressure level. (bottom right) As in (bottom left), but for parcel vertical displacement from the storm location back to the AIRS footprint.

For simplicity, the discussion is focused on thermodynamic variables (CAPE, CIN, LCL, and LFC calculated using SHARPPy, and low-level moisture) around the NARR backtrace results, which unlike the NAM12 results, span

the entire AIRS record. However, the higher-resolution NAM12 backtrace results are used to infer dynamic variables (shear, storm relative helicity) because of the poorer representation of dynamics in the NARR (shown later).

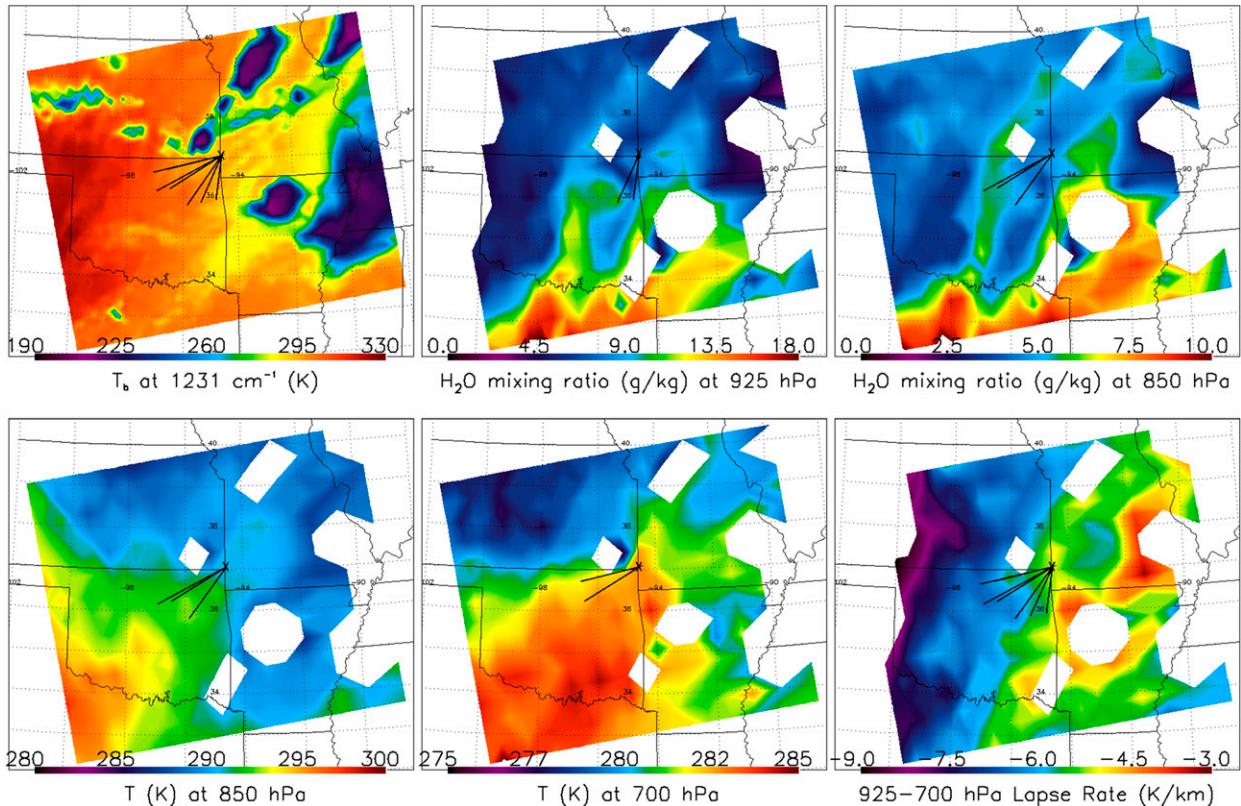


FIG. 3. AIRS granule measurements for the Joplin tornado case study. (top left) 1231 cm^{-1} brightness temperature, with “X” showing tornado report location and lines showing backtrace trajectories. (top center) q at 925 hPa. (top right) q at 850 hPa. (bottom left) T at 850 hPa. (bottom center) T at 700 hPa. (bottom right) Lapse rate between 925 and 700 hPa. The back trajectories calculated to the nearest AIRS pixel, from the surface to 600 hPa, are shown as black segments with the “X” denoting the tornado report at 1640 local time at 37.06°N , 94.57°W . The white areas indicate retrievals of T and q that did not pass the quality control at that particular pressure level (i.e., data quality was not “good” or “best,” 1 or 0, respectively).

1) CAPE

The AIRS backtrace analysis demonstrates an enhancement in CAPE from most unstable parcel lifting (MUCAPE) across storm categories relative to the conventional NN analysis (Tables 2 and 3, Fig. 5), with mean values enhanced by a factor of about 1.5 and somewhat greater enhancement for tornadoes than hail and winds.

The results of the backtracing approach are generally in good agreement with previous studies sounding the storm environment via various methods, although these previous works mostly relate to supercells. To facilitate comparison, Fig. 5 annotates estimates from these previous studies using green text. In general, backtracing brings the space-based CAPE estimates into much better agreement with the literature than the traditional NN approach, and could be thought of as a correction for systematic error introduced by the multihour time gap between observations of storm environments and convective initiation. Furthermore, backtracing from

higher-resolution NAM12 winds yields clear improvement over backtracing from NARR winds, suggesting that estimates might improve further as additional resolution becomes available.

Parker (2014), Rasmussen and Blanchard (1998), Grzych et al. (2007), and Wagner et al. (2008) describe proximity radiosondes, mesonet, and AERI observations made near tornadic and nontornadic supercells. These observations clearly demonstrate large variations at spatial scales not resolved by the AIRS operational (45 km footprint) retrieval, but Fig. 5 shows that mean values of CAPE are similar to AIRS. The results of the backtracing approach are furthermore in general agreement with previous studies (Thompson et al. 2003, 2012; Grams et al. 2012; Schultz and Askelson 2012) using MUCAPE and MLCAPE derived from the RUC analysis, although some differences are observed.

Using numerical simulations, Crook (1996) determined that CAPE is generally more impacted by q than T , while CIN is more impacted by T than q .

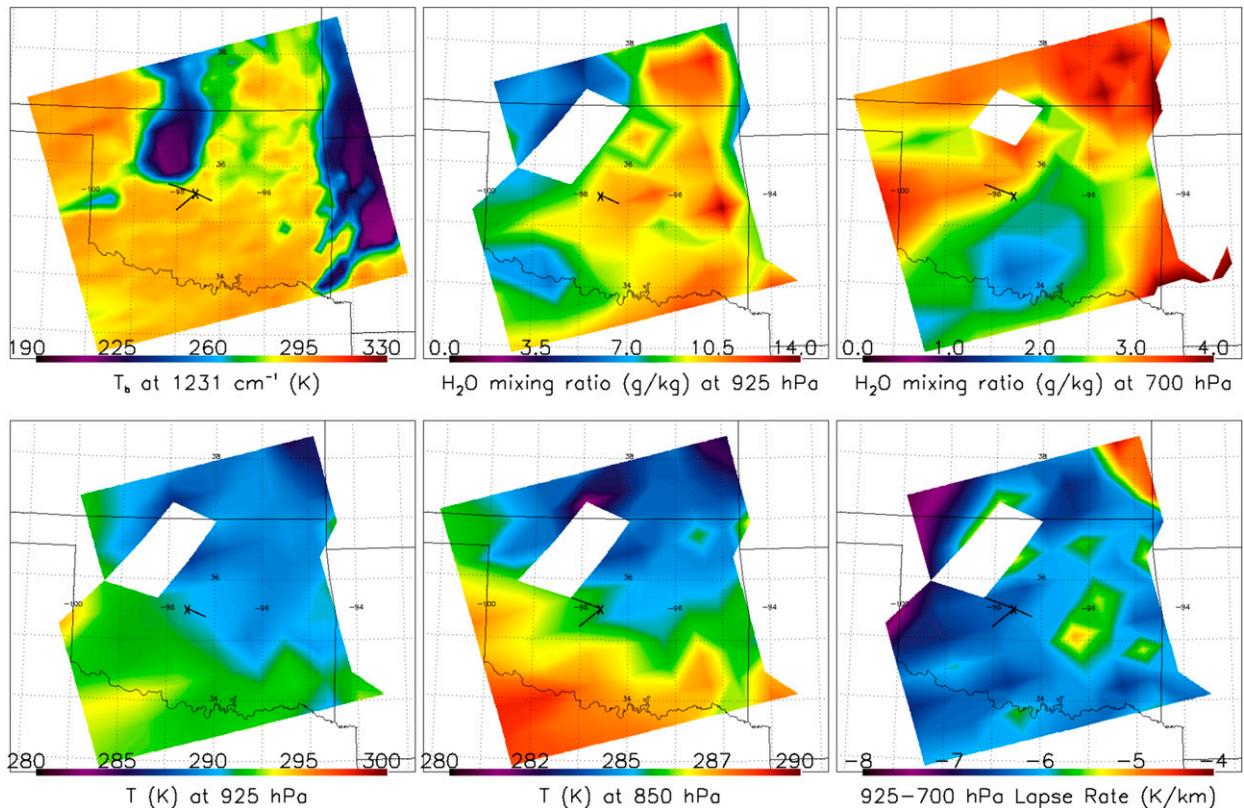


FIG. 4. AIRS granule measurements for the Oklahoma hail case study. (top left) 1231 cm^{-1} brightness temperature, with “x” showing first hail report location and lines showing backtrace trajectories. (top center) q at 925 hPa. (top right) q at 700 hPa. (bottom left) T at 925 hPa. (bottom center) T at 850 hPa. (bottom right) Lapse rate between 925 and 700 hPa. The back trajectories calculated to the nearest AIRS pixel, from the surface to 600 hPa, are shown as black segments with the “x” denoting the first 4.25 in. hail report at 1358 LT at 36.11°N , 98.35°W . The white areas indicate retrievals of T and q that did not pass the quality control at that particular pressure level.

Once convection is well developed, the variance of q near the surface plays a much larger role in modulating convective intensity. McCaul and Weisman (2001) showed with numerical simulations that the shape of the CAPE profile, even for fixed CAPE values, has a profound impact on convective evolution but this effect is stronger for lower values of CAPE. The backtrace methodology is able to more realistically sample fluctuations in environmental CAPE derived from AIRS soundings and therefore it is no surprise that improved representation of CAPE is obtained in the backtrace results relative to the NN results.

2) CIN

The AIRS backtrace method shows a striking difference in CIN estimates from the conventional NN approach (Tables 2 and 3, Fig. 6), with mean CIN reduced by a factor of ~ 1.5 using the backtrace method. Using NAM12, a much lower CIN is also observed for the distribution of violent EF4 and EF5 tornado events relative to weaker tornado events (the appendix), the

strongest hailstorm events with hailstone diameters larger than 4 in. relative to weaker hail events (not shown), and the strongest wind gusts relative to weaker subcategories using the backtrace method, whereas the opposite trend is observed using the NN method.

As with CAPE, CIN calculated with the backtrace method is in much better agreement with the prior literature than CIN calculated from NN soundings. The preponderance of literature results show a substantial decrease in CIN for tornadic versus nontornadic events (Rasmussen and Blanchard 1998; Grzych et al. 2007; Parker 2014; Klees et al. 2016; Davies 2004) although a large spread in CIN magnitude is noted. However, Wagner et al. (2008) present opposing results using AERI profiles, finding that environments up to three hours before tornadic and nontornadic supercells have CIN of $50\text{--}80$ and $0\text{--}20\text{ J kg}^{-1}$, respectively. The large spread in CIN estimates obtained in the studies cited above and the results from backtraced AIRS profiles, are consistent with a lack of an obvious physical connection of CIN to tornado formation in the near-storm

TABLE 3. Mean environmental indices tabulated by category, from backtraces calculated with NAM12 winds from 2008 to 2016.

Key	CAPE (J kg ⁻¹)	CIN (J kg ⁻¹)	LCL (km)	LFC (km)	q (g kg ⁻¹)	Shear magnitude (kt)	Shear direction (°)	SRH (m ² s ⁻²)	STP
TOR	1697	78	1.6	2.8	11.0	10.8	80	76	0.19
HAIL	1741	79	1.7	2.9	10.8	7.5	74	44	0.07
TOR2	1874	72	1.7	2.8	11.1	13.6	85	104	0.19
TOR1	1682	78	1.6	2.8	11.0	10.6	80	73	0.19
HAIL2	2184	78	1.6	2.7	11.6	7.4	67	47	0.13
HAIL1	1663	80	1.7	2.9	10.7	7.5	75	44	0.07
WND2	2257	55	1.6	2.5	12.1	8.3	67	55	0.15
WND1	1603	77	1.7	2.8	12.3	6.5	101	35	0.06

environment. Reliable CIN fields resolved at the storm-scale, especially those near the rear flank downdraft using available satellite soundings, may not be possible with a nominal AIRS resolution of 45 km. However, this topic should be revisited when an improved 15-km resolution sounding is made widely available (Irion et al. 2018).

Last, Davies (2004) showed that CIN is lowest for F2–F4 tornadoes (31 J kg⁻¹), slightly elevated for F0–F1 tornadoes (38 J kg⁻¹), and even higher for nontornadic supercells (72 J kg⁻¹). In stark contrast to the AIRS NN results, the AIRS backtracing with NAM12 are consistent with the results of Davies (2004) despite the fact that the AIRS profile is inherently smoothed compared to radiosondes (Botes et al. 2012). This implies that CIN created from variations in stabilization can be largely resolved by AIRS soundings.

3) LCL

Figure 7 shows that backtracing increases the lifting condensation level (LCL) estimates from AIRS relative to NN analysis by 200–500 m across storm categories with NARR typically 100–200 m higher than NAM12. From the AIRS record, violent tornadic environments have a somewhat lower LCL relative to weak and strong tornadic environments independent of analysis method (see the appendix). LCL from the AIRS record does not clearly distinguish between the mean tornadic and hail environments.

Parker (2014) and Rasmussen and Blanchard (1998) found the LCL to be lower in tornadic supercells than in nontornadic supercells; this trend is shown in both the NN results and the NAM12 backtrace results. We note that for both the NN and NAM12 backtrace methods, the LCL distributions for tornadic versus hail events are statistically distinct at 95% confidence, as described in detail below. However, Thompson et al. (2003) showed a substantial decrease in MLLCL for strong versus weaker tornadoes, a trend not exhibited by the AIRS backtrace MULCL results. Bunkers et al. (2006) found that the longer-lived supercells generally had lower LCLs.

In contrast to CAPE and CIN, the AIRS backtraced estimates are generally in worse agreement with the literature than the AIRS NN estimates. We speculate that the reduced differentiation in AIRS-derived LCL could be related to the coarse AIRS spatial resolution (~45 km near-nadir view), the coarse gridding of NARR (32 km) or NAM (12 km), rendering these unable to resolve storm-scale features in q variability. This is supported by slightly larger separation observed with the NAM compared to NARR; efforts using the 3-km

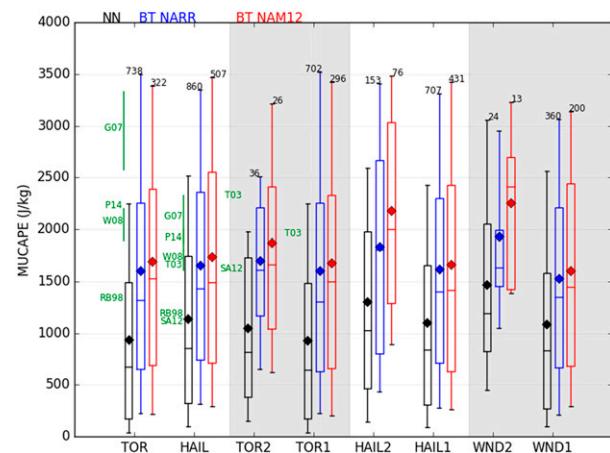


FIG. 5. CAPE (from most unstable parcel) from nearest neighbor (black), NARR backtrace soundings (blue), and NAM12 backtrace soundings (red). Here, as in the rest of the paper, box-and-whisker plots give means (stars), medians (lines), interquartile ranges (boxes), and 10%–90% ranges (whiskers). The black number on the left in each group of three boxes, located horizontally between the NN and NARR whiskers, gives the count of NN and NARR backtrace soundings; the black number on the right in each group of three boxes, located horizontally aligned with the NAM whisker, gives the count of NAM backtrace soundings. Estimates from references are shown in green (vertical green lines, when present, represent interquartile ranges) and are abbreviated here and in following figures as follows: RB98 (Rasmussen and Blanchard, 1998); G07 (Grzych et al. 2007); W08 (Wagner et al. 2008); T03 (Thompson et al. 2003); SA12 (Schultz and Askelson, 2012); D04 (Davies, 2004); DE03 (Doswell and Evans, 2003); and P14 (Parker, 2014).

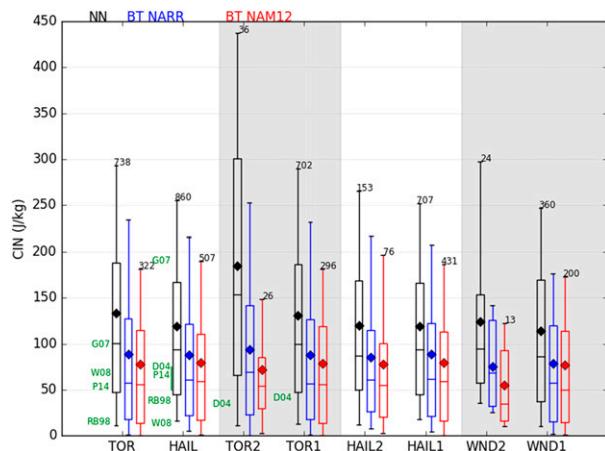


FIG. 6. As in Fig. 5, but for CIN (from most unstable parcel) from nearest neighbor and backtrace soundings.

HRRR model will be made in the near future. Furthermore, the lowest layers within the profile and the surface itself are not extrapolated from the existing layers in the profile. Future work will consider ways in which to couple together the surface and boundary layer in a physically consistent manner (e.g., Gartzke et al. 2017).

4) LFC

The NN and backtrace methods produce similar magnitudes of level of free convection (LFC), indicating that this parameter is not very sensitive to the backtrace methodology. Results from higher-resolution NAM12 are consistently lower than results from NARR. Results from Schultz and Askelson (2012) from RUC proximity soundings for significant tornadoes and nontornadic supercells are very similar to the corresponding AIRS estimates (Tables 2 and 3, Fig. 8). We note that for the NN method only, the LFC distributions for tornadic versus hail events are statistically distinct at 95% confidence, as described below.

5) LOW-LEVEL MOISTURE

The low-level moisture is quantified by taking the maximum value of q between 850 hPa and the surface. Figure 9 shows the results using the NN and backtrace methods. The backtrace methodology shows moistening over the NN method for all categories, although the change is small for the wind categories. The mean value of q between 850 hPa and the surface was also examined (not shown); the results were generally consistent with Fig. 9, although, unsurprisingly, the enhancement was less pronounced. We note that for the backtrace method with NAM12 winds, the maximum q distributions for tornadic versus hail events are statistically distinct at 95% confidence, as described below.

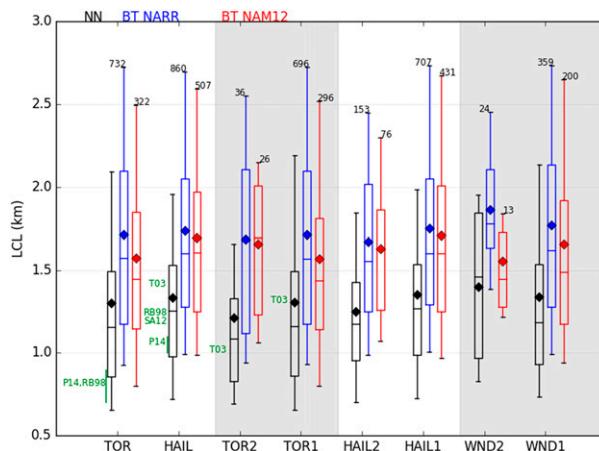


FIG. 7. As in Fig. 5, but for LCL from nearest neighbor and backtrace soundings.

6) SHEAR, STORM RELATIVE HELICITY

The wind direction and magnitude shear are estimated from the start and stop of backtrace trajectories at each of the 40 pressure levels. A backtraced proxy to 1 km directional and magnitude shear (Fig. 10) and directional shear (Fig. 11) are calculated. Shear magnitude shows a clear distinction between tornadic and nontornadic storm-event categories, as well as a clear increase with increasing tornado intensity. Using NAM12 winds consistently and substantially enhances values of shear magnitude relative to NARR winds, again suggesting that the analysis remains resolution-limited. The magnitudes with respect to the severe event categories, for the NAM12 winds, shows good agreement in shear magnitude with Schultz and Askelson (2012); but these estimates from the higher-resolution winds are low compared to estimates from Doswell and Evans (2003), a discrepancy that might be reduced with improved resolution. Directional shear (Fig. 11) did not discriminate between tornadic and nontornadic environments.

Storm relative helicity (SRH) at 1 km measures the potential for cyclonic updraft rotation and is shown in Fig. 12. Thompson et al. (2003), Schultz and Askelson (2012), and Parker (2014) report substantially higher estimates of SRH for tornadic and nontornadic supercells. This discrepancy may be partly due to limitations in spatial and temporal resolution of the meteorological data used by HYSPLIT.

While shear and SRH are well described in the literature and AIRS does not provide an estimate of either quantity, Figs. 10–12 are useful for validating the methodology and assumed meteorological forcing, which may vary depending on the convective index of interest. We note that for the backtrace method with NAM12 winds, the shear magnitude and SRH distributions for both

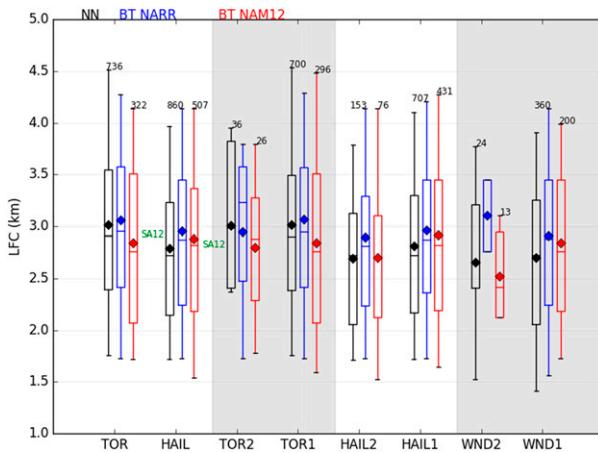


FIG. 8. As in Fig. 5, but for LFC from nearest neighbor and backtrace soundings.

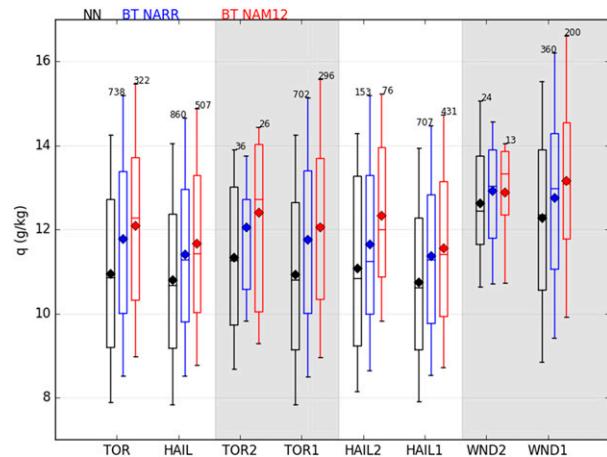


FIG. 9. As in Fig. 5, but for maximum low-level q (between the ground and 850 hPa).

tornadic versus hail events and TOR2 versus TOR1 events are statistically distinct at 95% confidence, as described below.

7) SIGNIFICANT TORNADO PARAMETER

Figure 13 shows significant tornado parameter (STP), a compound index that takes into consideration both thermodynamic and kinematic information (MLCAPE, MLLCL, MLCIN, storm relative helicity, and magnitude shear) in an attempt to maximize discrimination between tornadic and nontornadic supercells (Thompson et al. 2012). When using the NAM12 winds, STP estimates in the tornado category via the backtrace method are enhanced by a factor of ~ 2 relative to corresponding estimates via the NN method, although they are still significantly below values in the literature. As with SRH, the STP is apparently limited by resolution in winds, as the NARR analysis shows little increase in STP relative to NN. The backtrace analysis might be further limited by the spatial resolution of the AIRS soundings.

8) TESTING STATISTICAL SIGNIFICANCE OF DIFFERENCES AMONG STORM REPORTS

The statistical significance of differences for each convective index was quantified with respect to storm report type. The statistical significance is tested separately for two combinations of storm report types: (i) all tornado events versus all hail events, and (ii) between the two different tornado event categories (TOR2 and TOR1). For each convective index depicted in Figs. 5–13, the samples are tested to determine whether each combination arises or does not arise from the same population. Using the D’Agostino and Pearson test (D’Agostino and Pearson 1973), we show that none of

the samples are normally distributed (95% confidence). Therefore, the nonparametric Kruskal–Wallis test (Kruskal and Wallis 1952) is used to test the null hypothesis that the sample population median of all distributions in the combinations is equal (i.e., whether the samples are drawn from the same underlying population). The results of the tests are presented in Table 4 for both the NN analysis and the NAM12 backtrace analysis and are reported at the 95% confidence level.

The differences in CAPE are significant between tornadoes and hail only when using the NN method and not when using the backtrace method, while differences in CAPE are not significant for the tornado EF-scale in either NN or backtrace. However, it is important to stress that the distinction between the NN tornado and hail distributions runs contrary to the values in the literature (hence the designation “CONTRA LIT” in Table 4); in addition, CAPE magnitudes themselves are more in line with the literature using the backtrace method. For CIN, the differences among backtrace and NN are not significant in any category, but the CIN magnitude is greatly improved with the backtrace. The LCL discriminates between tornadoes and hail for NN and backtrace at the 95% confidence level, but not within the EF scale for tornadoes. While the LFC only discriminates for NN between tornadoes and hail, the larger value for hail does not agree with prior literature. The differences in q_{\max} are significant between tornadoes and hail when using the backtrace method, but not the NN method. We emphasize that q_{\max} is not a proxy for relative humidity in which substantial differences have been found previously (e.g., Brooks et al. 1994). The three dynamical indices (shear magnitude, shear direction, and SRH) are derived from the backtrace

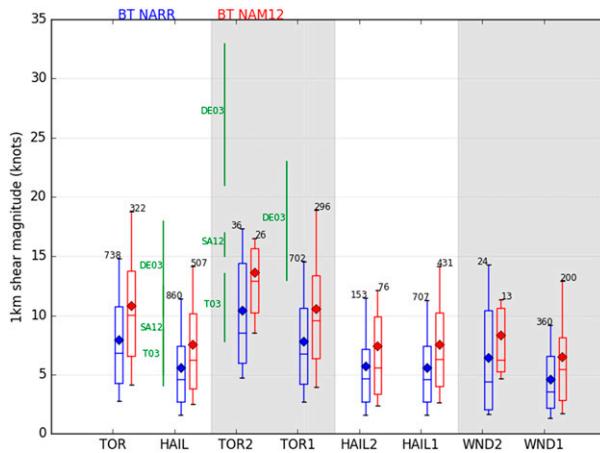


FIG. 10. As in Fig. 5, but for the magnitude of 1 km wind shear. Shear is calculated from wind directions and magnitudes estimated from the beginning and endpoints of backtrace trajectories; there is no NN shear estimate.

only. While the composite STP index is not significant for NN or backtrace in both storm report difference cases, there is a notable improvement in the separation for the NAM12 over the NARR for the backtrace in STP.

c. Parcel vertical ascent

Figure 14 shows mean parcel vertical motions by storm category, from the location at the AIRS overpass time to the location of the storm event, for NARR winds (left) and NAM12 winds (right). Except at the ground level, mean values of parcel lifting are observed at every other level and every storm category. A 2–3-h time interval in between the AIRS overpass and storm event (Fig. 1), and a typical ascent of ~ 300 m, translates to vertical velocities of $\sim 3\text{--}5\text{ cm s}^{-1}$ that are typical of synoptic-scale lift. Large-scale lifting is essential for overcoming CIN and increasing environmental favorability for deep moist convection (Doswell 1987). An increase in ascent with height also increases the temperature lapse rate. In the absence of diabatic processes a typical ascent of 300 m at 2 km translates to a steepening of the 0–2 km lapse rate by 1.5 K km^{-1} . The variance of parcel ascent across storm events within each category (not shown) is large: level-dependent standard deviations are as large as 400 m for the tornado and hail categories with NAM12 winds, and even higher for the wind categories. Because of substantial differences in vertical structure in lifting between the lower-resolution NARR winds and higher-resolution NAM12, we speculate that the estimated synoptic-scale lifting profiles remain model resolution-limited, and future work with higher-resolution models (e.g., HRRR) is warranted.

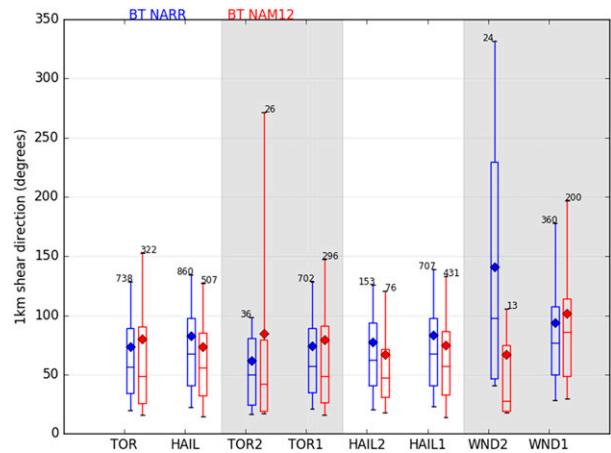


FIG. 11. As in Fig. 5, but for the directional component of 1 km wind shear.

4. Sensitivity of results to methodology

An unknown magnitude of uncertainty arises from a substantial reduction to 6%–7% of the original sample size (Table 1). An estimate of sampling bias uncertainty in the categorical mean and variance estimates of environmental properties at the storm scale may not offer as much insight as the quantification of the larger-scale environment. Below the focus is instead on the sensitivity of environmental estimates to the choice of meteorological forcing and the use of a mixed layer versus most unstable parcel assumption.

a. NARR 32 km versus NAM 12 km

Backtraced estimates of environmental properties use two distinct meteorological forcings, NARR and NAM12 are shown in Fig. 15 for a randomly chosen storm. The variations in back trajectories and moisture profiles using NARR and NAM12 can be significant on a storm-by-storm basis. With its coarser spatial (32 km) and temporal (3 h) resolution, the NARR is missing higher-frequency and smaller spatial variability in the flow field relative to the NAM12 with finer spatial (12 km) and temporal (1 h) resolution. Figure 16 shows MUCAPE using both the NARR and NAM 12 km across all storm categories. The set of storms remaining in the analysis differs substantially between the two meteorological forcings, due to the different back trajectories sampling different portions of the AIRS swath and are matched to different AIRS retrievals, not all of which pass our quality control criteria. Out of 1983 NARR events and 1041 NAM12 events, the intersection includes 757 storm events. Over those 757 storm events, the mean value of NAM12 MUCAPE is 7% higher than the mean value of NARR MUCAPE, and the correlation coefficient is $r = 0.77$. For CIN and LCL, the mean

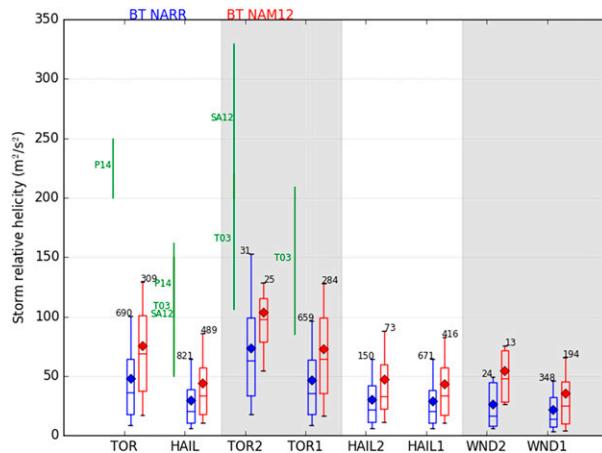


FIG. 12. As in Fig. 5, but for storm relative helicity.

value of NAM12 across all categories is 9% larger and 4% smaller than the mean value of NARR, respectively (not shown).

b. Mixed layer versus most unstable parcel

Thermodynamic indices from NAM12 backtraces calculated via the most unstable parcel method and the mixed layer parcel method (up to 100 hPa, using the SHARPPy) were compared (Fig. 17). While the mean CAPE varies by up to a factor of about 2, the mean percent difference (the difference of the means divided by the mean of the means) across all eight subcategories excluding the two composite categories is 32% with the most unstable parcel method giving higher mean values in every category. This is consistent with previous layer-based calculations (e.g., Craven et al. 2002). For CIN, parcel choice also makes a critical difference; the mean percent difference across all eight subcategories is 84%, with mixed layer parcels always giving larger values of CIN (not shown). For LCL, the mean percent difference across all eight subcategories is 12% with the most unstable parcel method being higher in every category (not shown).

c. Thermodynamic indices versus overpass/storm initiation time delay

We examined how the thermodynamic indices, as estimated via the NN method and the backtrace method, varied as a function of the time difference between the AIRS overpass (at approximately 1330 local time) and the reported storm initiation time. Figure 18 shows that estimates from the two methods tend to converge at small time differences, as expected, and tend to diverge as the time difference increases, especially for CAPE and CIN. Much of the benefit of the backtrace method comes from correcting estimates

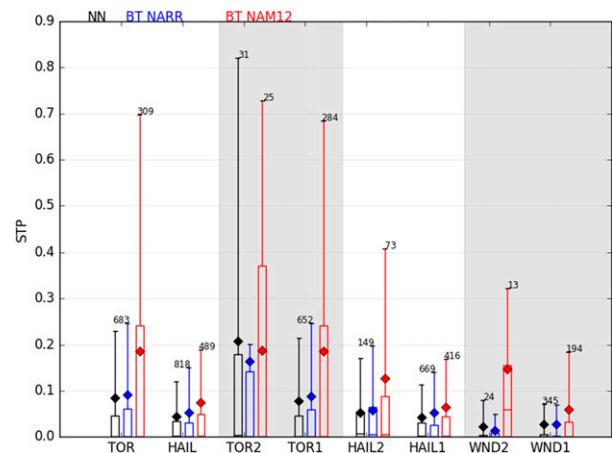


FIG. 13. As in Fig. 5, but for significant tornado parameter (STP).

for storms that occur more than an hour or two after the AIRS overpass.

5. Conclusions and outlook

A novel methodology is described that couples back trajectories to satellite soundings of temperature and specific humidity. This methodology enables more accurate environmental estimates of satellite soundings of temperature T and specific humidity q in the proximity of rapidly evolving severe convective storms relative to the more conventional nearest neighbor approach, by accounting for changes in the environment that occur between the time of satellite overpass and the time of interest (e.g., tornado formation). The new method is applied to the Atmospheric Infrared Sounder (AIRS) instrument on the EOS *Aqua* satellite and uses the NOAA Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT) with meteorological forcing from the North American Regional Reanalysis (NARR) and the 12-km resolution North American Mesoscale Forecast System (NAM12). Nearest neighbor without backtracing (NN), as well as backtraced soundings of T and q are estimated in proximity to NCEI Storm Event tornadoes, large hail ≥ 2 in. diameter, and thunderstorm wind gusts ≥ 65 mph. While parcel T is adjusted adiabatically for vertical motions, parcel q is not adjusted for possible precipitation. Verification of the storm mode (i.e., tornadic supercell, nontornadic supercell, or nonsupercellular thunderstorm) underlying the reported event types was not attempted; therefore, the categories reported likely contain mixing between storm modes, decreasing our ability to discriminate between the categories in subsequent analysis. Future work could mitigate this limitation by using additional datasets to categorize events by storm mode.

TABLE 4. Shown are the results for the nonparametric Kruskal-Wallis test (Kruskal and Wallis, 1952). “NO” symbolizes that the test did not pass (p value >0.05), while “YES” symbolizes that the test did pass (p value <0.05). The bold italic font denotes instances where one method (either NN or backtrace) allows for distinguishing between the samples while the other method does not. “CONTRA LIT” indicates that the test did pass, but that the sign of the separation between the distributions is in clear contradiction to the body of previous estimates in the literature. All p values are reported in parentheses. The backtrace results are obtained with the NAM12.

Convective parameter	NN tornado vs hail	NN TOR1 vs TOR2	Backtrace tornado vs hail	Backtrace TOR1 vs TOR2
CAPE	CONTRA LIT (3×10^{-6})	NO (0.1)	NO (0.7)	NO (0.2)
CIN	NO (0.05)	NO (0.1)	NO (0.5)	NO (0.9)
LCL	YES (3×10^{-4})	NO (0.4)	YES (0.002)	NO (0.2)
LFC	YES (6×10^{-6})	NO (0.7)	NO (0.5)	NO (0.6)
q_{\max}	NO (0.1)	NO (0.3)	YES (0.009)	NO (0.5)
Shear magnitude	—	—	YES (4×10^{-19})	YES (2×10^{-3})
Shear direction	—	—	NO (0.5)	NO (0.5)
SRH	—	—	YES (5×10^{-23})	YES (7×10^{-4})
STP	NO (0.5)	NO (0.3)	NO (0.3)	NO (0.2)

The backtracing methodology unlocks a previously untapped potential of the utility of polar-orbiting satellites in quantifying the favorability of the thermodynamic environment for convective initiation and storm evolution. This utility is afforded by accounting for temporal gaps between the satellite overpasses, and furthermore by providing a continuous record of the evolution of T and q during these temporal gaps which, in future work, could yield insight into storm evolution. The 1330 local time overpass of the AIRS instrument is ideal for observing subsequent severe weather outbreaks over the North American continent that initiate after the satellite overpass. More realistic values of convective available potential energy (CAPE) and

convective inhibition (CIN) are derived for backtracing than NN soundings. This method may also offer a pathway forward to determine optimal satellite sounding locations in proximity to severe convective storms. The results show some sensitivity to the choice of winds (NARR versus NAM12) for HYSPLIT, and suggest that the method remains resolution-limited even with the 12-km winds.

There is potential in the near future to further enhance the spatial (and possibly vertical) resolution and to otherwise enhance the information content of backtraced profiles via 1) AIRS soundings at 13.5 km resolution that are produced for targeted studies (Irion et al. 2018); 2) 3-km HRRR winds to drive HYSPLIT;

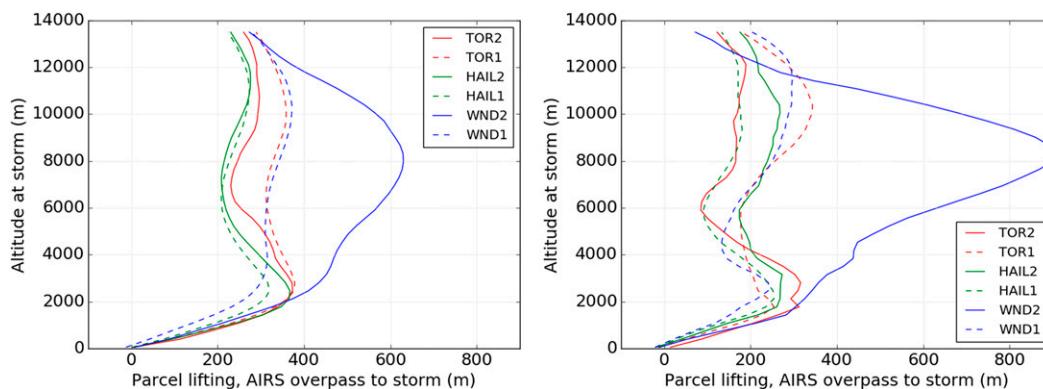


FIG. 14. Mean parcel vertical motions by storm category, from the site of the AIRS overpass to the site of the storm event, for (left) NARR winds and (right) NAM12 winds.

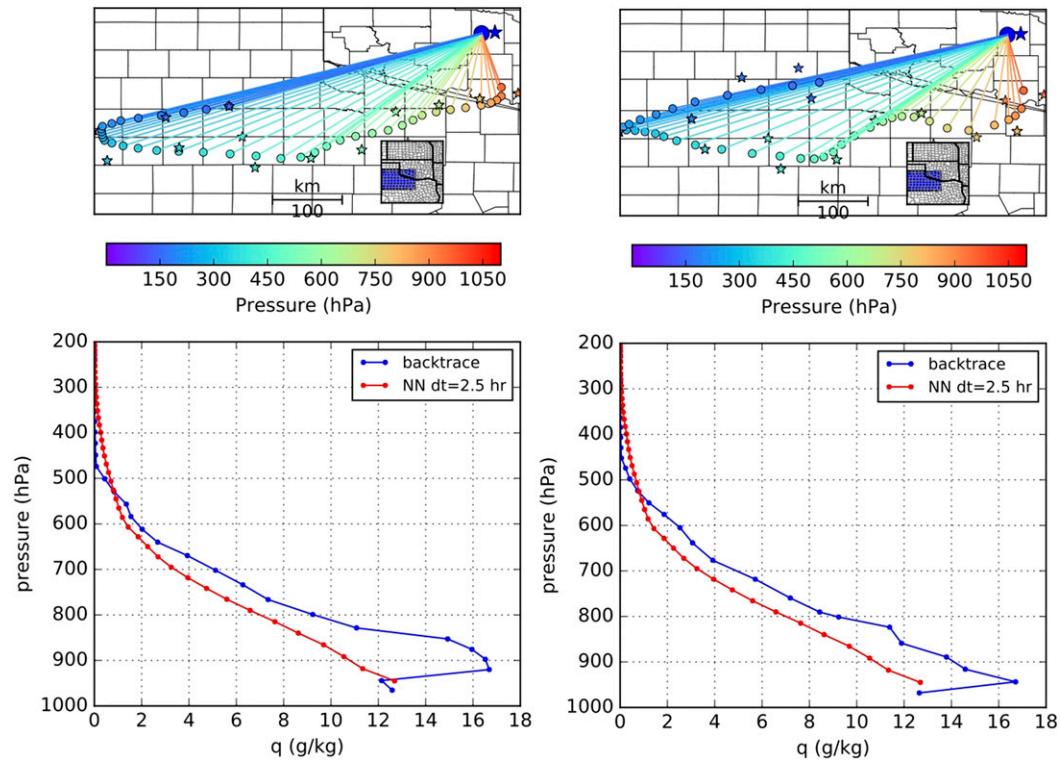


FIG. 15. A single storm, an EF4 tornado on 24 May 2011, analyzed with (left) NARR winds and (right) NAM12 winds. (top) Backtrace contours, with color indicating height (orange is near ground, blue is near 200 hPa). (bottom) Water vapor profiles; cyan are from the AIRS NN retrieval, while blue is from the backtraced parcels. The dt notations in the legends indicate the time difference between the NN matchup and the storm event in hours.

3) spatial kriging (interpolation via Gaussian process regression) between AIRS soundings to extract the information contained in spatial gradients (Cressie and Johannesson 2008); and 4) coupling the surface and boundary layer in a physically consistent manner (Agard and Emanuel 2017) to further enhance the realism of convective indices in the presence of soil moisture variability (e.g., Schär et al. 1999; Hohenecker et al. 2009). Furthermore, the accuracy of convective indices could be improved by using surface values of T and q from nearby ground stations in place of surface AIRS retrievals of T and q as demonstrated by Gartzke et al. (2017). There is also the potential to quantify the time evolution of environments in proximity to long-lived supercells or organized convective systems, either by calculating a series of backtraces with time along the supercell or convective system trajectory or using soundings from multiple polar-orbiting satellites with varying overpass times. The diurnal cycle is a key aspect to the evolution of severe weather outbreaks (e.g., Agard and Emanuel 2017) and a time series of backtraces from one or more hyperspectral infrared sounding instruments would yield new insight on the diurnal variability of convective indices.

This new methodology could also be utilized to develop semicontinuous spatial maps of time-dependent soundings and derived convective indices in regions suspected of having convective potential by applying forward tracing in time to hundreds or thousands of neighboring satellite soundings. This approach is a type

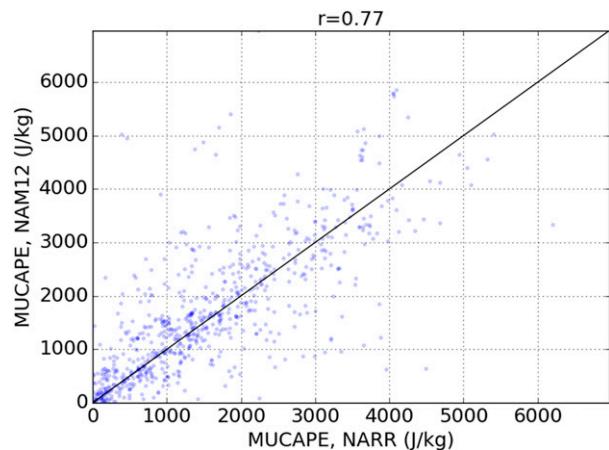


FIG. 16. MUCAPE calculated from two meteorological analyses, NAM12 and NARR (using SHARPPy).

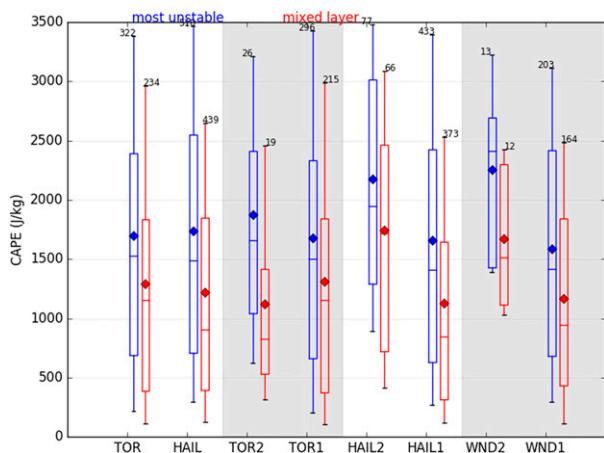


FIG. 17. Ensemble CAPE calculated using most unstable parcel method (blue) and mixed layer parcel method (red).

of “nowcasting” that may offer additional guidance to forecasters in the 1–4-h time frame. Future research will focus on using the HRRR wind fields as input to HYSPLIT applied to swaths of satellite soundings using forward trajectories. Rather than use the current approach that differentiates the environment among

different types of severe convective weather, the forward tracing will be applied over swath data to obtain insight on convective initiation, convective mode preference, and convective evolution and subsequent potential for severe weather. This method is also likely to be useful for assessing the processes that contribute to extreme precipitation events.

The RAPv4/HRRRv3 system does assimilate some clear-sky hyperspectral infrared sounder radiances. As the derived satellite sounding products used in this investigation are obtained in both clear and partly cloudy scenes, there is additional information regarding T/q gradients and features within the AIRS L2 satellite retrievals that is absent from the clear-sky AIRS radiance assimilation. Additionally, the HRRR errors in relative humidity are notably higher during the warm season when convective regimes are more dominant (Benjamin et al. 2016). Thus, we assert that there is potentially added value in using derived T/q satellite soundings until data assimilation systems are unambiguously shown to be superior to satellite retrievals.

The satellite infrared sounder record will be extended far into the foreseeable future with missions already planned to return data through the mid-2040s.

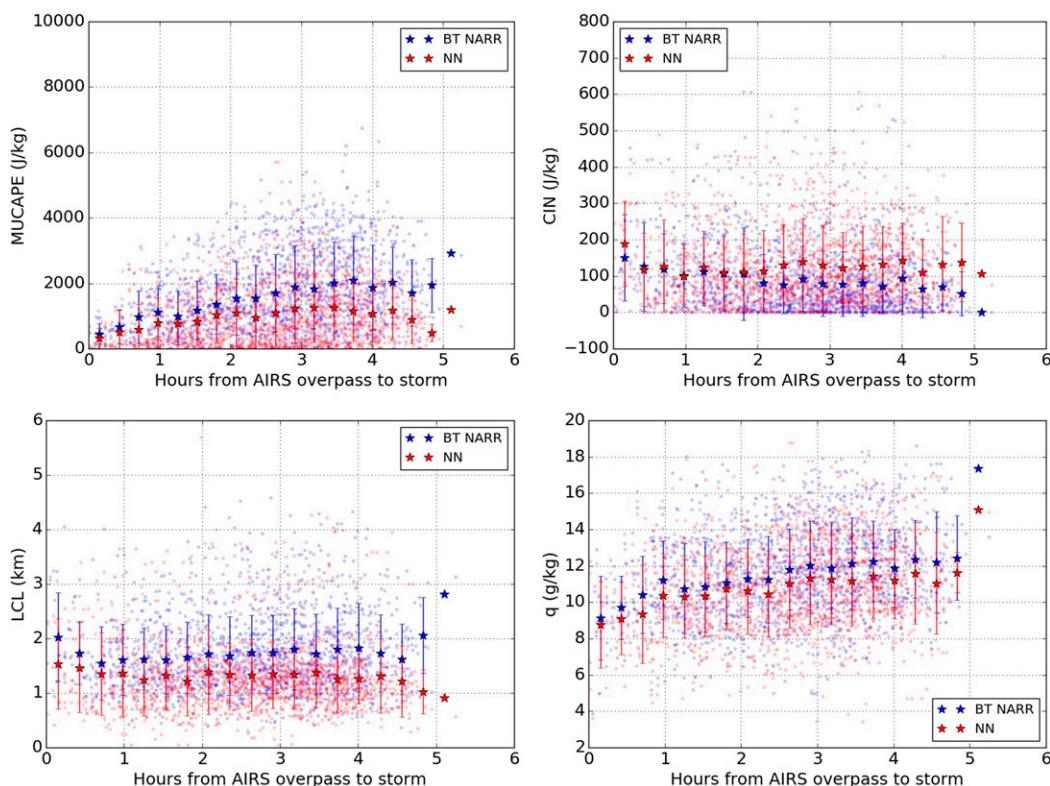


FIG. 18. Thermodynamic indices, estimated from both nearest neighbor (red) and NARR backtrace soundings (blue), as a function of time between the AIRS overpass (at approximately 1330 local time) and the reported storm initiation time.

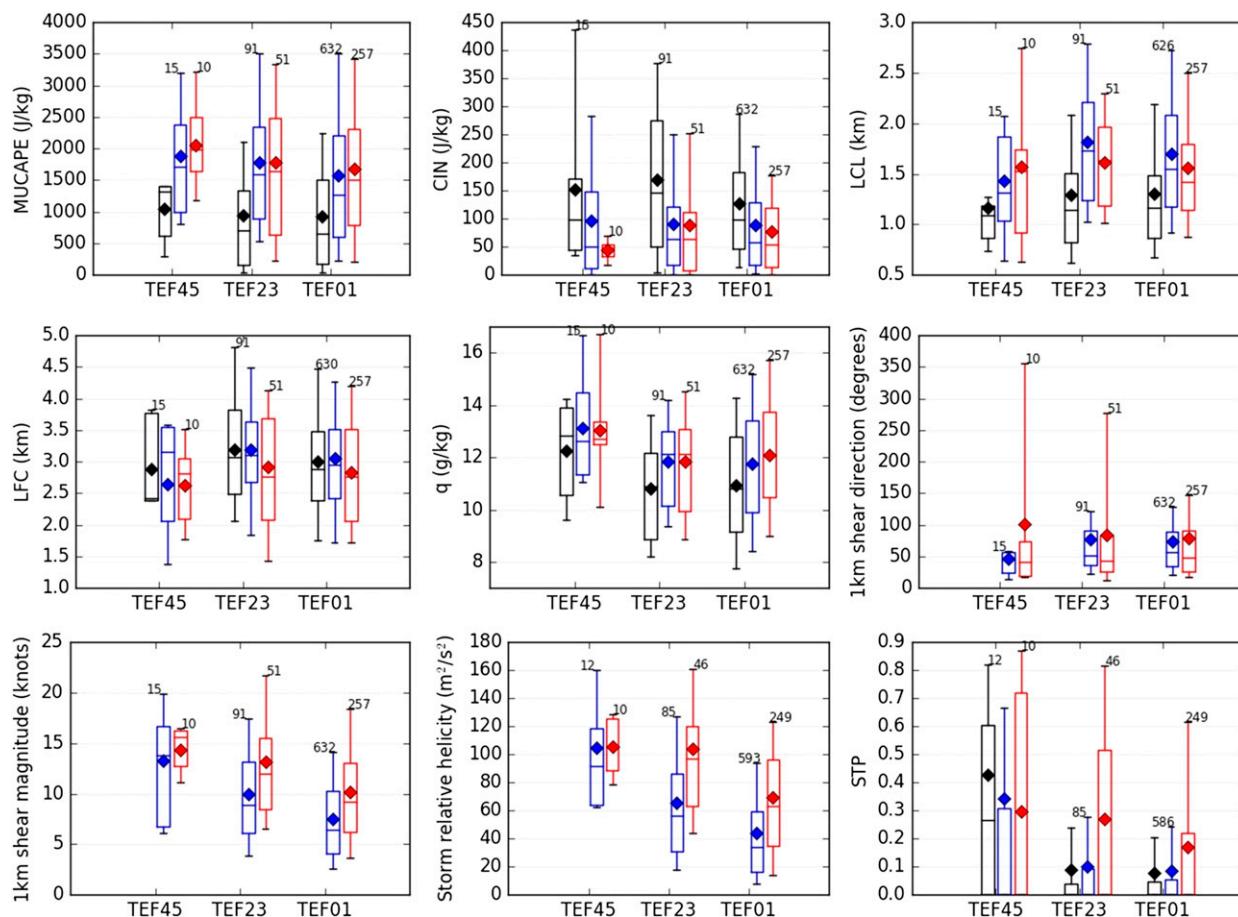


FIG. A1. Convective indices depicted in Figs. 5–13 for EF0–1, EF2–3, and EF4–5 categorizations. Black indicates NN analysis, blue indicates backtrace analysis with NARR, and red indicates backtrace analysis with NAM.

The availability of a continuous record of high-spectral-resolution infrared sounders starting with AIRS on the NASA *Aqua* satellite in 2002, following with CrIS on the *Suomi National Polar-Orbiting Partnership (SNPP)* satellite in 2011, in the near future with CrIS on the Joint Polar Satellite System (JPSS) series of satellites (*JPSS-1–JPSS-4*) with continuous coverage planned through 2036, and IASI on the MetOp satellite series since 2006 with continuous coverage planned through 2020 and next-generation coverage [(IASI-NG) on the MetOp Second Generation (MetOP-SG-A) series] planned to carry the IASI record continuously through the mid-2040s, will provide a multidecadal record in which this methodology could be extended.

This investigation demonstrates the feasibility and promise of the backtracing methodology, which ultimately will be applied to longer records, to other regions around the world with severe convective storms such as Europe, South America, South Asia, Australia, and elsewhere. As the climate change imprint will lead to changes in atmospheric temperature, humidity, and

vertical and horizontal wind shear over a range of spatial and temporal scales (Trapp et al. 2007; Diffenbaugh et al. 2013; Hoogewind et al. 2017), which ultimately will impact convective processes, this methodology will facilitate a more rigorous approach to assess rapid mesoscale thermodynamic environmental changes from satellite remote sensing datasets.

Acknowledgments. A portion of this research was carried out at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The NASA Severe Storm Research program under Grant NNN13D497T supported all authors in this research, while NASA Grant NNN12AA01C supported SWF and SCV. We express our gratitude to Ramesh Kakar for supporting this effort. We thank three anonymous reviewers and the editor for constructive comments that led to an improved manuscript, and to Chris Barnet and Amy Braverman for fruitful discussions. The NCEI Storm Events database was accessed through

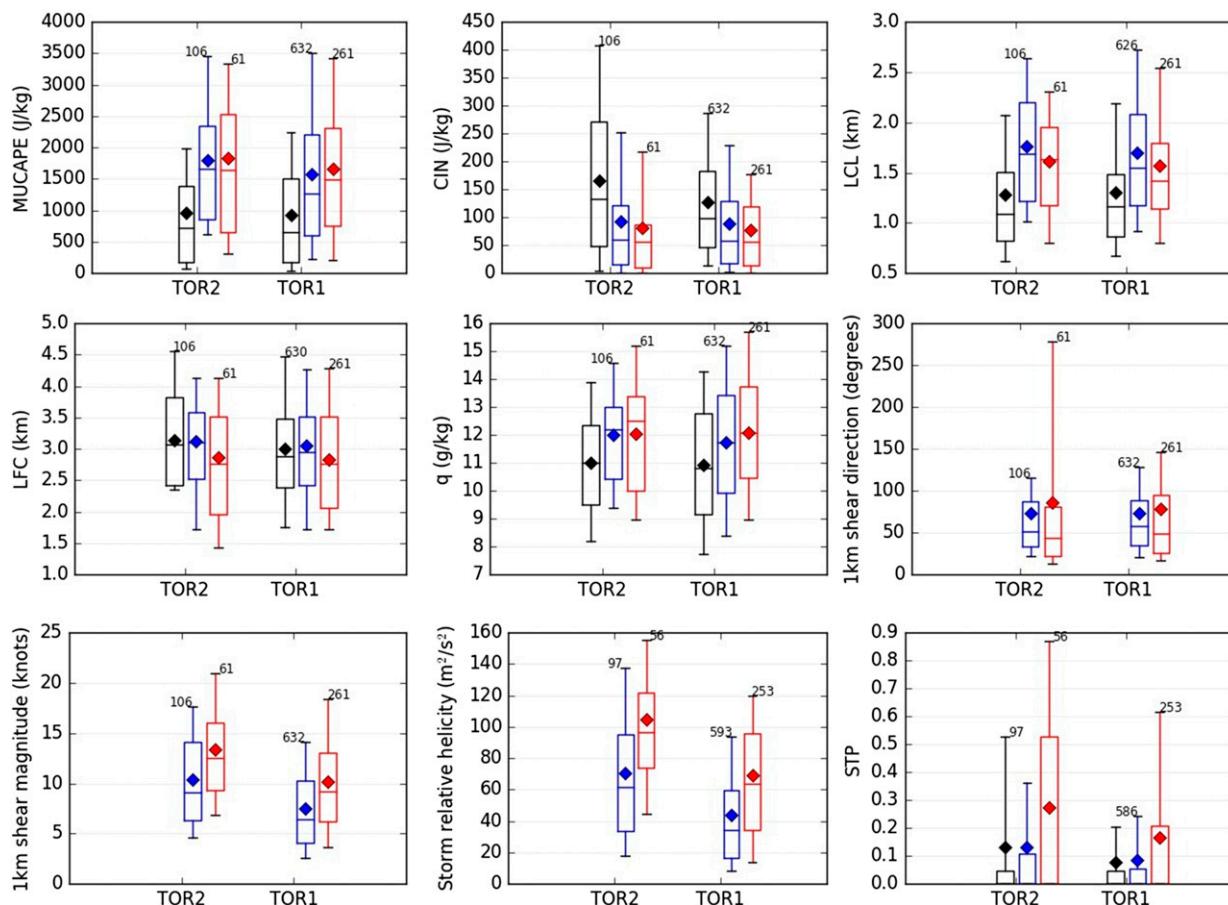


FIG. A2. Convective indices depicted in Figs. 5–13 for EF0–1 (“TOR1”) and EF2–5 (“TOR2”) categorizations. Black indicates NN analysis, blue indicates backtrace analysis with NARR, and red indicates backtrace analysis with NAM.

<https://www.ncdc.noaa.gov/stormevents/>. We gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model (<http://www.ready.noaa.gov>). NARR data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at <http://www.esrl.noaa.gov/psd/>, and NAM12 data were provided by NOAA/NCEP/NCEI (<https://www.ncdc.noaa.gov/data-access>). AIRS data were obtained through the Goddard Earth Services Data and Information Services Center (<http://daac.gsfc.nasa.gov/>).

indices depicted in Figs. 5–13. While the counts for EF2–3 are 5–6 times larger than EF4–5, the violent tornado category shows additional increases in MUCAPE, q , shear magnitude, SRH, STP, and reductions in CIN and LCL that are consistent with literature on violent tornadoes. The differences between EF2–3 and EF2–5 are small (not shown) because of small sample sizes for EF4–5 tornadoes. We also report results from the EF0–1 and EF2–5 categorization (Fig. A2), which demonstrates relatively low sensitivity to placing the category boundary between EF1 and EF2 versus between EF2 and EF3.

APPENDIX

Convective Indices for Three EF-Scale Categories

The choice for the EF0–2 and EF3–5 tornado categorizations was motivated by small sample sizes for EF4–5 tornadoes but was otherwise arbitrary. In this appendix, the EF0–1, EF2–3, and EF4–5 categorization is reported separately (Fig. A1) for the nine convective

REFERENCES

Agard, V., and K. Emanuel, 2017: Clausius–Clapeyron scaling of peak CAPE in continental convective storm environments. *J. Atmos. Sci.*, **74**, 3043–3054, <https://doi.org/10.1175/JAS-D-16-0352.1>.
 Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, <https://doi.org/10.1175/MWR-D-15-0242.1>.
 Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L. Thompson, and J. A. Hart, 2017: SHARPPy: An Open-Source

- Sounding Analysis Toolkit for the Atmospheric Sciences. *Bull. Amer. Meteor. Soc.*, **98**, 1625–1636, <https://doi.org/10.1175/BAMS-D-15-00309.1>.
- Botes, D., J. R. Mecikalski, and G. J. Jedlovec, 2012: Atmospheric Infrared Sounder (AIRS) sounding evaluation and analysis of the pre-convective environment. *J. Geophys. Res.*, **117**, D09205, <https://doi.org/10.1029/2011JD016996>.
- Brooks, H. E., C. A. Doswell, and J. Cooper, 1994: On the environments of tornadic and nontornadic supercells. *Wea. Forecasting*, **9**, 606–618, [https://doi.org/10.1175/1520-0434\(1994\)009<0606:OTEOTA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0606:OTEOTA>2.0.CO;2).
- Bunkers, M. J., J. S. Johnson, L. J. Czephyha, J. M. Grzywacz, B. A. Klimowski, and M. R. Hjelmfelt, 2006: An observational examination of long-lived supercells. Part II: Environmental conditions and forecasting. *Wea. Forecasting*, **21**, 689–714, <https://doi.org/10.1175/WAF952.1>.
- Chahine, M. T., and Coauthors, 2006: AIRS: Improving weather forecasting and providing new data on greenhouse gases. *Bull. Amer. Meteor. Soc.*, **87**, 911–926, <https://doi.org/10.1175/BAMS-87-7-911>.
- Coffer, B. E., and M. D. Parker, 2017: Simulated supercells in nontornadic and tornadic VORTEX2 environments. *Mon. Wea. Rev.*, **145**, 149–180, <https://doi.org/10.1175/MWR-D-16-0226.1>.
- Coniglio, M. C., S. M. Hitchcock, and K. H. Knopfmeier, 2016: Impact of assimilating preconvective upsonde observations on short-term forecasts of convection observed during MPEX. *Mon. Wea. Rev.*, **144**, 4301–4325, <https://doi.org/10.1175/MWR-D-16-0091.1>.
- Craven, J. P., R. E. Jewell, and H. E. Brooks, 2002: Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. *Wea. Forecasting*, **17**, 885–890, [https://doi.org/10.1175/1520-0434\(2002\)017<0885:CBOCCB>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0885:CBOCCB>2.0.CO;2).
- Cressie, N., and G. Johannesson, 2008: Fixed rank kriging for very large spatial data sets. *J. Roy. Stat. Soc. B*, **70**, 209–226, <https://doi.org/10.1111/j.1467-9868.2007.00633.x>.
- Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785, [https://doi.org/10.1175/1520-0493\(1996\)124<1767:SOMCFB>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1767:SOMCFB>2.0.CO;2).
- Dabberdt, W. F., and Coauthors, 2005: Multifunctional mesoscale observing networks. *Bull. Amer. Meteor. Soc.*, **86**, 961–982, <https://doi.org/10.1175/BAMS-86-7-961>.
- D’Agostino, R., and E. S. Pearson, 1973: Tests for departure from normality. *Biometrika*, **60**, 613–622.
- Davies, J. M., 2004: Estimations of CIN and LFC associated with tornadic and nontornadic supercells. *Wea. Forecasting*, **19**, 714–726, [https://doi.org/10.1175/1520-0434\(2004\)019<0714:EOCALA>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0714:EOCALA>2.0.CO;2).
- Davies-Jones, R., R. J. Trapp, and H. B. Bluestein, 2001: Tornadoes and tornadic storms. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 167–222, <https://doi.org/10.1175/0065-9401-28.50.167>.
- Diffenbaugh, N. S., M. Schruder, and R. J. Trapp, 2013: Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proc. Natl. Acad. Sci. USA*, **110**, 16 361–16 366, <https://doi.org/10.1073/pnas.1307758110>.
- Doswell, C. A., 1987: The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Wea. Forecasting*, **2**, 3–16, [https://doi.org/10.1175/1520-0434\(1987\)002<0003:TDBLSA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1987)002<0003:TDBLSA>2.0.CO;2).
- , and J. S. Evans, 2003: Proximity sounding analysis for derechos and supercells: An assessment of similarities and differences. *Atmos. Res.*, **67–68**, 117–133, [https://doi.org/10.1016/S0169-8095\(03\)00047-4](https://doi.org/10.1016/S0169-8095(03)00047-4).
- Feltz, W. F., and J. R. Mecikalski, 2002: Monitoring high-temporal-resolution convective stability indices using the ground-based Atmospheric Emitted Radiance Interferometer (AERI) during the 3 May 1999 Oklahoma–Kansas tornado outbreak. *Wea. Forecasting*, **17**, 445–455, [https://doi.org/10.1175/1520-0434\(2002\)017<0445:MHTRCS>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0445:MHTRCS>2.0.CO;2).
- Gartzke, J., R. Knuteson, G. Przybyl, S. Ackerman, and H. Revercomb, 2017: Comparison of satellite-, model-, and radiosonde-derived convective available potential energy in the Southern Great Plains region. *J. Appl. Meteor. Climatol.*, **56**, 1499–1513, <https://doi.org/10.1175/JAMC-D-16-0267.1>.
- Grams, J. S., R. L. Thompson, D. V. Snively, J. A. Prentice, G. M. Hodges, and L. J. Reames, 2012: A climatology and comparison of parameters for significant tornado events in the United States. *Wea. Forecasting*, **27**, 106–123, <https://doi.org/10.1175/WAF-D-11-00008.1>.
- Grasso, L. D., and W. R. Cotton, 1995: Numerical simulation of a tornado vortex. *J. Atmos. Sci.*, **52**, 1192–1203, [https://doi.org/10.1175/1520-0469\(1995\)052<1192:NSOATV>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1192:NSOATV>2.0.CO;2).
- Grzych, M. L., B. D. Lee, and C. A. Finley, 2007: Thermodynamic analysis of supercell rear-flank downdrafts from project ANSWERS. *Mon. Wea. Rev.*, **135**, 240–246, <https://doi.org/10.1175/MWR3288.1>.
- Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär, 2009: The soil moisture–precipitation feedback in simulations with explicit and parameterized convection. *J. Climate*, **22**, 5003–5020, <https://doi.org/10.1175/2009JCLI2604.1>.
- Hoogewind, K. A., M. E. Baldwin, and R. J. Trapp, 2017: The impact of climate change on hazardous convective weather in the United States: Insight from high-resolution dynamical downscaling. *J. Climate*, **30**, 10 081–10 100, <https://doi.org/10.1175/JCLI-D-16-0885.1>.
- Irion, F. W., and Coauthors, 2018: Single-footprint retrievals of temperature, water vapor, and cloud properties from AIRS. *Atmos. Meas. Tech.*, **11**, 971–995, <https://doi.org/10.5194/amt-11-971-2018>.
- Jiang, H., and Coauthors, 2015: Real-time applications of the variational version of the Local Analysis and Prediction System (vLAPS). *Bull. Amer. Meteor. Soc.*, **96**, 2045–2057, <https://doi.org/10.1175/BAMS-D-13-00185.1>.
- Jones, T. A., and D. J. Stensrud, 2012: Assimilating AIRS temperature and mixing ratio profiles using an ensemble Kalman filter approach for convective-scale forecasts. *Wea. Forecasting*, **27**, 541–564, <https://doi.org/10.1175/WAF-D-11-00090.1>.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Karstens, C. D., W. A. Gallus, B. D. Lee, and C. A. Finley, 2013: Analysis of tornado-induced tree-fall using aerial photography from the Joplin, Missouri, and Tuscaloosa–Birmingham, Alabama, tornadoes of 2011. *J. Appl. Meteor. Climatol.*, **52**, 1049–1068, <https://doi.org/10.1175/JAMC-D-12-0206.1>.
- Klees, A. M., Y. P. Richardson, P. M. Markowski, C. Weiss, J. M. Wurman, and K. K. Kosiba, 2016: Comparison of the tornadic and nontornadic supercells intercepted by VORTEX2 on 10 June 2010. *Mon. Wea. Rev.*, **144**, 3201–3231, <https://doi.org/10.1175/MWR-D-15-0345.1>.
- Koch, S. E., R. Ware, H. Jiang, and Y. Xie, 2016: Rapid mesoscale environmental changes accompanying genesis of an unusual

- tornado. *Wea. Forecasting*, **31**, 763–786, <https://doi.org/10.1175/WAF-D-15-0105.1>.
- Kruskal, W. H., and W. W. Wallis, 1952: Use of ranks in one-criterion variance analysis. *J. Amer. Stat. Assoc.*, **47**, 583–621, <https://doi.org/10.1080/01621459.1952.10483441>.
- Lerach, D. G., and W. R. Cotton, 2012: Comparing aerosol and low-level moisture influences on supercell tornadogenesis: Three-dimensional idealized simulations. *J. Atmos. Sci.*, **69**, 969–987, <https://doi.org/10.1175/JAS-D-11-043.1>.
- Maddy, E., and C. Barnet, 2008: Vertical resolution estimates in version 5 of AIRS operational retrievals. *IEEE Trans. Geosci. Remote Sens.*, **46**, 2375–2384, <https://doi.org/10.1109/TGRS.2008.917498>.
- Markowski, P. M., and Y. P. Richardson, 2009: Tornadogenesis: Our current understanding, forecasting considerations, and questions to guide future research. *Atmos. Res.*, **93**, 3–10, <https://doi.org/10.1016/j.atmosres.2008.09.015>.
- , J. M. Straka, and E. N. Rasmussen, 2002: Direct surface thermodynamic observations within the rear-flank downdrafts of nontornadic and tornadic supercells. *Mon. Wea. Rev.*, **130**, 1692–1721, [https://doi.org/10.1175/1520-0493\(2002\)130<1692:DSTOWT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1692:DSTOWT>2.0.CO;2).
- , —, and —, 2003: Tornadogenesis resulting from the transport of circulation by a downdraft: Idealized numerical simulations. *J. Atmos. Sci.*, **60**, 795–823, [https://doi.org/10.1175/1520-0469\(2003\)060<0795:TRFTTO>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0795:TRFTTO>2.0.CO;2).
- McCaul, E. W., and M. L. Weisman, 2001: The sensitivity of simulated supercell structure and intensity to variations in the shapes of environmental buoyancy and shear profiles. *Mon. Wea. Rev.*, **129**, 664–687, [https://doi.org/10.1175/1520-0493\(2001\)129<0664:TSOSSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0664:TSOSSS>2.0.CO;2).
- 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>.
- Potvin, C. K., K. L. Elmore, and S. J. Weiss, 2010: Assessing the impacts of proximity sounding criteria on the climatology of significant tornado environments. *Wea. Forecasting*, **25**, 921–930, <https://doi.org/10.1175/2010WAF2222368.1>.
- Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164, [https://doi.org/10.1175/1520-0434\(1998\)013<1148:ABCOSD>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2).
- Romps, D. M., 2017: Exact expression for the lifting condensation level. *J. Atmos. Sci.*, **74**, 3891–3900, <https://doi.org/10.1175/JAS-D-17-0102.1>.
- Rotunno, R., and J. B. Klemp, 1985: On the rotation and propagation of simulated supercell thunderstorms. *J. Atmos. Sci.*, **42**, 271–292, [https://doi.org/10.1175/1520-0469\(1985\)042<0271:OTRAPO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1985)042<0271:OTRAPO>2.0.CO;2).
- , —, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines. *J. Atmos. Sci.*, **45**, 463–485, [https://doi.org/10.1175/1520-0469\(1988\)045<0463:ATFSL>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<0463:ATFSL>2.0.CO;2).
- Schär, C., D. Lüthi, U. Beyerle, and E. Heise, 1999: The soil-precipitation feedback: A process study with a regional climate model. *J. Climate*, **12**, 722–741, [https://doi.org/10.1175/1520-0442\(1999\)012<0722:TSPFAP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0722:TSPFAP>2.0.CO;2).
- Schultz, C. J., and M. A. Askelson, 2012: Vertical variations of boundary layer potential buoyancy in tornadic and nontornadic near-storm environments. *Wea. Forecasting*, **27**, 1489–1506, <https://doi.org/10.1175/WAF-D-11-00097.1>.
- Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 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>.
- Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan, 2015: NOAA’s HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Amer. Meteor. Soc.*, **96**, 2059–2077, <https://doi.org/10.1175/BAMS-D-14-00110.1>.
- 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](https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2).
- , B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, <https://doi.org/10.1175/WAF-D-11-00116.1>.
- 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**, 408–415, <https://doi.org/10.1175/WAF925.1>.
- , N. S. Diffsenbaugh, H. E. Brooks, M. E. Baldwin, E. D. Robinson, and J. S. Pal, 2007: Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proc. Natl. Acad. Sci. USA*, **104**, 19719–19723, <https://doi.org/10.1073/pnas.0705494104>.
- Wagner, T. J., W. F. Feltz, and S. A. Ackernam, 2008: The temporal evolution of convective indices in storm-producing environments. *Wea. Forecasting*, **23**, 786–794, <https://doi.org/10.1175/2008WAF2007046.1>.
- Wakimoto, R. M., and H. Cai, 2000: Analysis of a nontornadic storm during VORTEX95. *Mon. Wea. Rev.*, **128**, 565–592, [https://doi.org/10.1175/1520-0493\(2000\)128<0565:AOANSD>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0565:AOANSD>2.0.CO;2).
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437, <https://doi.org/10.1175/2007WAF2007005.1>.
- Weisz, E., N. Smith, and W. L. Smith Sr., 2015: The use of hyperspectral sounding information to monitor atmospheric tendencies leading to severe local storms. *Earth Space Sci.*, **2**, 369–377, <https://doi.org/10.1002/2015EA000122>.
- Wicker, L. J., and R. B. Wilhelmson, 1995: Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm. *J. Atmos. Sci.*, **52**, 2675–2703, [https://doi.org/10.1175/1520-0469\(1995\)052<2675:SAOTD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<2675:SAOTD>2.0.CO;2).
- Wurman, J., Y. Richardson, C. Alexander, S. Weygandt, and P. F. Zhang, 2007: Dual-doppler and single-doppler analysis of a tornadic storm undergoing mergers and repeated tornadogenesis. *Mon. Wea. Rev.*, **135**, 736–758, <https://doi.org/10.1175/MWR3276.1>.
- Yue, Q., E. J. Fetzer, B. H. Kahn, S. Wong, G. Maniyan, A. Guillaume, and B. Wilson, 2013: Cloud-state dependent sampling in AIRS observations based on *CloudSat* cloud classification. *J. Climate*, **26**, 8357–8377, <https://doi.org/10.1175/JCLI-D-13-00065.1>.