Impact of WSR-88D Intra-Volume Low-Level Scans on Severe Weather Warning Performance®

JOHN Y. N. CHO,^a JAMES M. KURDZO,^a BETTY J. BENNETT,^a MARK E. WEBER,^a JOSEPH W. DELLICARPINI,^b
ANDREW LOCONTO.^b AND HAYDEN FRANK^b

^a Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts
^b NOAA/National Weather Service, Norton, Massachusetts

(Manuscript received 27 September 2021, in final form 8 April 2022)

ABSTRACT: The statistical relationship between supplemental adaptive intra-volume low-level scan (SAILS) usage on the Weather Surveillance Radar-1988 Doppler and National Weather Service severe storm warning performance during 2014–20 is analyzed. Results show statistically significant improvement in severe thunderstorm (SVR), flash flood (FF), and tornado (TOR) warning performance associated with SAILS-on versus SAILS-off. Within the three possible SAILS modes of one (SAILSx1), two (SAILSx2), and three (SAILSx3) additional base scans per volume, for SVR, SAILSx2 and SAILSx3 are associated with better warning performance compared to SAILSx1; for FF and TOR, SAILSx3 is associated with better warning performance relative to SAILSx1 and SAILSx2. Two severe storm cases (one that spawned a tornado, one that did not) are presented where SAILS usage helped forecasters make the correct TOR warning decision, lending real-life credence to the statistical results. Furthermore, a statistical analysis of automated volume scan evaluation and termination effects, parsed by SAILS usage and mode, yield a statistically significant association between volume scan update rate and SVR warning lead time.

KEYWORDS: Flood events; Severe storms; Tornadoes; Radars/Radar observations; Operational forecasting; Communications/decision making

1. Introduction

Weather forecasters consult many types of information (satellite imagery, lightning data, spotter reports, near-term forecasts, probabilistic warning guidance, etc.) while making severe weather warning decisions. For National Weather Service (NWS) forecasters in the United States, data from the Weather Surveillance Radar-1988 Doppler (WSR-88D; Crum and Alberty 1993) often play a leading role in the severe weather warning decision process (e.g., Burgess and Lemon 1990; Polger et al. 1994). The quality and timeliness of the radar data are, of course, important to the forecaster. For example, long-term statistical analyses have shown that higher operational warning performance for tornadoes (TOR; Cho and Kurdzo 2019a,b), flash floods (FF; Cho and Kurdzo 2020a), and non-tornadic severe thunderstorms (SVR; Cho and Kurdzo 2020b) are associated with better radar coverage and spatial resolution. Severe thunderstorms generate hail greater than or equal to 1 inch (2.54 cm) in diameter and/or convective wind gusts greater than or equal to 50 kt (25.7 m s⁻¹) or convective wind damage. A flash flood is a flood occurring generally less than 6 h from a causative event (e.g., heavy rainfall, dam failure, ice jam); in this study, the causative events were restricted to heavy rainfall. Studies conducted within testbed settings utilizing experimental, non-operational radar data have

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WAF-D-21-0152.s1.

Corresponding author: John Y. N. Cho, jync@ll.mit.edu

shown that faster radar data updates can also lead to improved TOR and SVR warning performance (Heinselman et al. 2015; Wilson et al. 2017).

Ideally, in terms of timeliness, weather radar observations from every point in space would be available continuously and instantaneously. This is impossible, because it takes the radar a finite amount of time to obtain high quality data from each resolution volume (e.g., Doviak and Zrnić 1993). In fact, there are interdependent trade-offs between spatial coverage, observation update frequency, and data quality. With the WSR-88D, the NWS forecaster manages these trade-offs in real time by selecting from a menu of available volume coverage patterns (VCPs; https://training.weather.gov/wdtd/buildTraining/build19/presentation/presentation_content/external_files/Quick_ref_VCP_comparison.png), based on which VCP best suits the weather phenomenon of immediate interest (e.g., Brown et al. 2000).

The WSR-88D VCPs are defined as automated sets of 360° azimuthal antenna rotations, each at a constant elevation angle that make up an entire volume of scans. These have evolved over time in response to research outcomes and forecaster input. In the past decade, options to the basic VCPs were introduced that further expanded the WSR-88D's observational flexibility: 1) automated volume scan evaluation and termination (AVSET; Chrisman 2013), 2) supplemental adaptive intra-volume low-level scan (SAILS; Chrisman 2014), and 3) mid-volume rescan of low-level elevations (MRLE; Chrisman 2016). Briefly, AVSET shortens the volume update time whenever possible by adaptively skipping highelevation-angle scans that contain no precipitation returns. SAILS inserts one to three extra lowest-elevation-angle (base) scan(s) dispersed evenly in time throughout the VCP

DOI: 10.1175/WAF-D-21-0152.1

cycle. MRLE is similar to SAILS, except other low-level elevations scans (e.g., 0.9°, 1.3°, and 1.8°) are updated more frequently as well. After an initial trial period, AVSET has been on by default at all sites since 2012, although it can be manually turned off when desired. SAILS and MRLE are options that are selected in real time by NWS weather forecast office (WFO) forecasters, and have been available since 2014–15 and 2018–19, respectively, depending on the date that each site was updated with the corresponding software build.

Thus, it is important to quantify the efficacy of these recent VCP enhancements to 1) provide better guidance for their usage, 2) aid in the development of future VCPs and options, 3) help inform scanning requirements for the eventual WSR-88D replacement, and 4) document benefits gained from optimizing scan strategies. Since the WSR-88D has a mechanically steered antenna that places physical constraints on scanning methodologies, electronically scanned phased-array radars (PARs) are being considered as a potential future alternative (Weber et al. 2021). Points 3 and 4 inform the research-to-operations plan (NOAA 2020) for this alternative radar system.

This paper focuses on the effects of SAILS on SVR, FF, and TOR warning performance. We hypothesized that faster base scan updates may help forecasters make more accurate and timely warning decisions. SAILS is further subdivided into SAILSxN, where N = 1, 2, or 3 corresponds to the number of additional base scan(s) inserted into the VCP. (The N > 1 cases are referred to as MESO-SAILS, where MESO stands for "multiple elevation scan option.") There are now barely enough data on MRLE usage for statistically significant results to begin emerging, and this scanning mode will be examined at a later time. (Since the national roll-out of MRLE began in May 2018 to the end of 2020, MRLE has been used on just 0.14% of volume scans, including only the post-deployment period at each site.) As AVSET was in effect during the study period, we analyzed its influence as well (see the appendix). We conclude with two case studies that illustrate the utility of SAILS in data to assist forecasters making TOR warning decisions.

2. Analysis methodology

The basic idea behind our analysis was to find the scan mode that was operating on the primary WSR-88D used by the forecaster while making the decision to issue or not to issue a severe weather warning. The warning performance metrics—probability of detection (POD), mean lead time (MLT), and false alarm ratio (FAR)—were then computed, parsed by VCP type. We chose mean lead time over median lead time, because confidence intervals for the mean are more straightforwardly computed and more commonly agreed upon than confidence intervals for the median, and confidence intervals are crucial in establishing statistically significant differences. Warnings were further categorized as being associated with the same event (leading and trailing warnings) or not (solo warnings). This categorization is explained at the beginning of section 3.

The data needed were the following: 1) storm event data from the National Oceanic and Atmospheric Administration (NOAA) National Center for Environmental Information (NCEI; https://www.ncdc.noaa.gov/stormevents/), 2) storm warning data from the Iowa Environmental Mesonet NWS Watch/Warnings archive (https://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml), and 3) WSR-88D Archive III Status Product (ASP). The ASP data, which contain pervolume-scan information on time, VCP number, SAILS and MRLE status, and volume scan duration, are available on NCEI (https://www.ncdc.noaa.gov/nexradinv/) as well as on Google Cloud (https://console.cloud.google.com/storage/browser/gcp-public-data-nexrad-13/2019/12? authuser=0&prefix).

On one hand, input data volume should be maximized to reduce statistical uncertainty in the results. On the other hand, background circumstances should be kept constant to reduce unintended biases. Given these opposing exigencies, the analysis period was selected to start on the SAILS deployment date at each WSR-88D site. This varied from 28 February 2014 at KRAX (Raleigh–Durham, North Carolina) to 15 May 2015 at KBYX (Key West, Florida). The analysis period end date was 31 December 2020. We set the geographic coverage to be the contiguous United States (CONUS).

To determine the scan mode used at the time of a warning decision, we first matched each severe weather event and warning polygon to the nearest WSR-88D. This procedure assumes that the closest WSR-88D was the one being relied upon most by the forecaster when the warning decision was made. This is not necessarily true, as the closest radar may have been not operating or perhaps a radar farther away had less terrain blockage; also, data from multiple radars may have been consulted. It is impossible to be certain of the "primary" radar that was used without forecasters' logs recording this information. However, because of the large data quantity that we processed, we expect deviations from the assumption (nearest radar being most important) would be fairly small "noise" components in the statistical results.

We then found the ASP volume scan timestamp that was closest to and before the warning issuance time. However, if the difference between the warning issuance time and the ASP timestamp was greater than 11 min, the match was discarded. This was necessary to filter out instances where the nearest radar was not in operation or its ASP record was not available (11 min marks the longest possible normal volume scan duration, corresponding to VCPs 31 and 32). Less than 0.7% of the data were eliminated by this criterion. For a missed detection, there was no associated warning; it is impossible to know when the decision not to issue a warning was made. In fact, it may be more accurate to say that there was a range of time during which the decision was made not to issue a warning. To account for this spread of time during the assessment period, we subtracted the median warning lead time for that event type (SVR, FF, or TOR) from the initial eventoccurrence time, with the idea that this would approximate when the decision not to issue a warning was made. The medians were computed over detected events parsed by warning category (lead, trailing, solo) and EF number groupings (EF0-1, EF2, EF3-5). The median was chosen for this purpose rather than the mean, because the median is less sensitive to outliers. Although median lead times also vary with WFO, we did not subdivide the data further in this way, since this would have led to significantly higher variance in the lead time estimates. This approximate decision time was then used for matching with the ASP scan timestamp.

For FF events, instead of using the event locations directly to find the nearest radar, we first mapped the locations to the corresponding catchment basin polygons—see Cho and Kurdzo (2020a) for details on how this was done. This process more accurately places the radar observations most relevant for FF warning decisions, i.e., where the rain was falling, not where the flooding occurred.

Since a warning decision is not made instantly, one may wonder whether the single scan mode that was being used shortly before the warning issuance time was truly representative of the scan modes operating throughout the decision process. To probe this question, we computed how often the scan modes were changed during the last hour leading up to the warning issuance time. The results are shown in Fig. 1. (The data used to generate all the plots in this paper are available in the supplemental material.) We see that, for the vast majority of cases, the VCP number, SAILS on/off status, and SAILSxN were held constant during that hour. This gives us confidence in the validity of the single, matched scan mode in computing warning performance dependence on scan usage statistics.

SAILS usage is restricted to five VCPs: 12, 35, 112, 212, and 215. VCP 35 is a "clear air" mode with a volume update period of ~7 min that is rarely employed for non-winter severe weather observation. VCP 215 is a general precipitation surveillance mode with dense upper-elevation sampling and a volume update period of about 6 min. VCP 112, which replaced VCP 121 in 2020, is used for large-scale systems with widespread high velocities, and has a volume update period of about 5.5 min. VCPs 12 and 212 trade off upper-elevation sampling density in favor of faster volume update periods (about 4 and 4.5 min, respectively); the only difference between these two VCPs is that VCP 212 uses pulse-phase coding and processing for second-trip recovery (Sachidananda and Zrnić 1999), whereas VCP 12 does not. The second-trip recovery scheme requires slightly longer scan times, but increases Doppler data coverage to longer ranges. (The approximate volume update periods given above are for when AVSET does not skip any high-elevation scans and SAILS is off. AVSET effects are analyzed apart from SAILS effects in the appendix.) Furthermore, VCPs 112 and 215 allow only SAILSx1. In order to maintain uniform conditions as much as possible while not discarding too much data, we decided to keep only VCPs 12 and 212 for the SAILS analysis. These two VCPs were used in the overwhelming fraction of warning decisions for all event types (Fig. 2): 94% for SVR, 87% for FF, and 96% for TOR.

Finally, we know from our past studies that radar coverage quality around the storm location is statistically related to warning performance for TOR (Cho and Kurdzo 2019a,b), FF (Cho and Kurdzo 2020a), and SVR (Cho and Kurdzo 2020b). If there is a strong positive correlation between SAILS usage and radar coverage, e.g., fraction of vertical space observed (FVO), around the storm of interest, then the

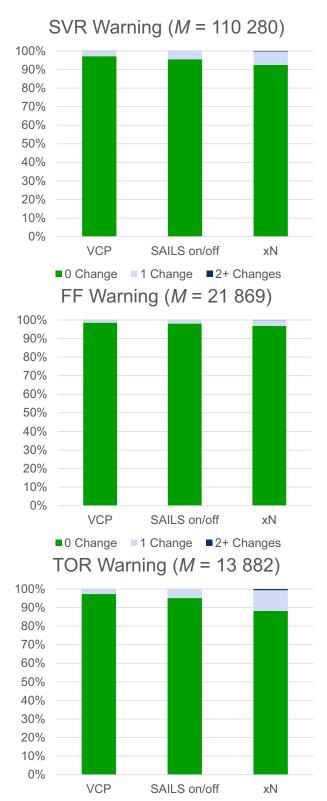


FIG. 1. Percentage of cases during 1 h prior to when a warning decision was made that a scanning mode (VCP number, SAILS on/off, and SAILSxN) was not changed, changed once, or changed two or more times; *M* is the number of samples for each warning type.

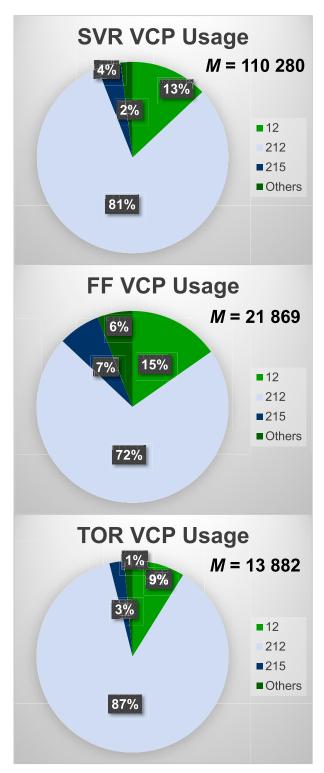


FIG. 2. VCP usage frequency by severe weather warning type.

statistics of warning performance dependence on SAILS usage could be biased by that correlation. Therefore, we computed the FVO distributions for SAILS off versus SAILS

on for each warning type to check for significant statistical differences. As we did not find such notable differences, we felt confident in proceeding with the SAILS impact analysis.

Regarding the warning performance metrics, we defined a detection when a point event was inside the warning polygon or a polygon-delineated event was inside or intersected the warning polygon, and if any portion of the event duration overlapped the warning-valid interval in time; otherwise, the warning was classified as a false alarm. This definition is consistent with NWS severe convective weather verification procedures (NWS 2009). For a detection, the lead time was computed as the event's start time minus the initial warning issuance time. By this definition, negative lead times were included in the POD and MLT calculations. Negative lead times consist of a small fraction of all lead times; see, e.g., Fig. 7 in Cho and Kurdzo (2020b) for a SVR warning lead time distribution. POD was calculated as the number of detections divided by the number of events. FAR was computed as the number of false alarms divided by the number of warnings. Multiple events can occur within the same space and time boundaries of a single warning, which implies that even if there were no missed detections, the number of events can be more than the number of warnings.

3. SAILS impact results

NWS forecasters often initially select VCP modes (including SAILS) based on the expected threat. There may be multiple threat types over disparate geographic areas or over the same area (e.g., coincident severe thunderstorm and flash flood threats), and their relative priorities may evolve with time. The choice between the different levels of SAILS (x1, x2, x3) is a complex decision, because one must trade off base scan update rate with volume update rate. Thus, even though the perceived optimal VCP mode may vary over time, in this complex operational forecast and warning environment, the meteorologists' heavy workload may prevent them from changing VCP modes frequently. Therefore, forecasters may tend to default to SAILS-off or SAILSx1 mode until an initial threat detection is made.

With these potential human factors in mind, we computed results in four categories, wherever possible: 1) All data, 2) data associated with warnings that did not overlap with other warnings in space and time (dubbed "solo" warnings), 3) data associated with lead warnings, and 4) data associated with trailing warnings. We tagged a warning as "trailing" if there was any overlap in its valid period with the valid period of a warning issued earlier, and if there was also any overlap in their geographic polygons. This left "lead" warnings as warnings that had at least one space-time overlap with a trailing warning but were the first ones to be issued. SVR, FF, and TOR warnings were each handled separately. In practice, there is sometimes a delay (~1-4 min) in the issuance of a follow-up warning, especially during periods of high workload, resulting in a short time gap between the expiration of the earlier warning and the beginning valid time of the trailing warning. To account for this, we allowed up to a 4-min gap in the definition of time overlap. Likewise, there may also be a

small spatial gap between an earlier warning polygon and a follow-up polygon. To account for this contingency, we filled out the polygons to their convex hulls and expanded them by 0.05° in latitude–longitude space using the Matlab function "polybuffer." This procedure allowed angular spatial gaps between polygons of up to 0.1° in the definition of spatial overlap, which is on the order of 10 km at midlatitudes. The top plot in Fig. 3 shows the percentages of warnings in each category. Note that a proposed concept, Threats-in-Motion, seeks to mitigate such gaps by continuously updating polygons that move forward with the storm (Stumpf and Gerard 2021).

The FAR for these four categories was calculated directly from the parsed warning data. For MLT, we classified all detected events according to the matching warning. (An event that was successfully predicted by multiple warnings was assigned to the warning with the earliest issuance time.) It was not possible to compute POD separately for solo, lead, and trailing warnings, because its calculation requires knowledge of the number of unwarned events for each category, which cannot be determined.

In operational practice, the event type for which an initial (lead) warning is issued may change for what may be called the trailing warning. Just to give one example, after a TOR warning was issued, a subsequent warning for the same area may be revised to a SVR warning based on a rapidly weakening radar tornado signature. Such "crossover" phenomena are not captured by the event-based warning categories that we used.

Figure 4 displays SAILS usage by warning type. SAILSx1 is the most prevalent scan mode for all warning types, except for TOR trailing warnings where SAILSx3 is most commonly used by a tiny margin. This is partly explained by the fact that MESO-SAILS (SAILSx2 and SAILSx3) did not start being deployed until January 2016. However, year-by-year usage plots (Fig. 5) indicate that the preference for SAILSx1 has continued in more recent years. Overall usage of SAILS for severe weather warning decisions has also declined somewhat after reaching a peak in 2016. It is also clear that forecasters use SAILS (and MESO-SAILS) more aggressively in making TOR warning decisions.

There is a trend of increasing SAILS usage in the order of solo, lead, then trailing warnings, which can be deduced from the fact that SAILS-off usage decreases in this order in Fig. 4. The same order applies to SAILSx3 usage. It must be noted that this paper does not evaluate observed storm mode relative to SAILS usage but rather only warning type and associated SAILS selections. This precludes the ability to directly assess the nature of the convection in terms of mode, areal coverage, and storm motion relative to SAILS trends. Nonetheless, we cautiously surmise that solo warnings tend to be associated with narrower footprint, more transient or faster passing events, e.g., airmass storms, seen as posing lower threat levels. Lead and trailing warnings, by definition part of multiple linked warnings, may tend to be associated with larger, more persistent storms, e.g., mesoscale convective systems (MCSs) and supercells. This proposition is consistent with the median warning area being smallest, and the median warning-valid period

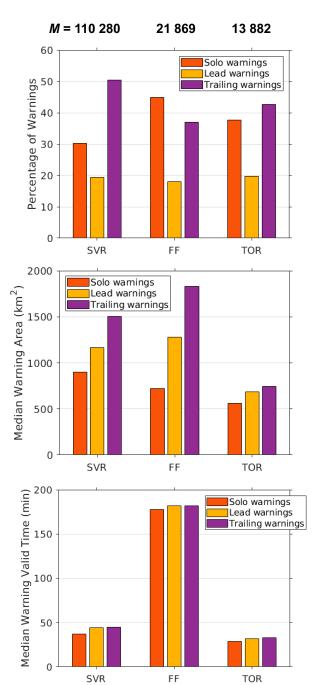


FIG. 3. Warning statistics by category: (top) issued percentage, (middle) median warning area, and (bottom) median warning valid time. (The median was used rather than the mean because it is less sensitive to outliers.)

shortest (to a lesser extent), for solo warnings (middle and bottom plots in Fig. 3).

Because there are many potential factors that influence SAILS usage, we are planning a nationwide survey of WFOs to obtain direct input from forecasters on their thought processes leading to SAILS (and MRLE) usage decisions. Until

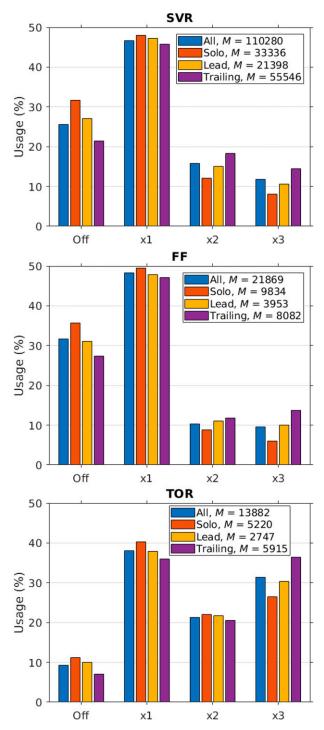


FIG. 4. SAILS usage frequency by severe weather warning type and category.

this survey and associated results are known, we have no direct knowledge of forecaster reasoning for selecting a particular SAILS mode. Given this reality, we urge caution with drawing too many conclusions here. Nonetheless, one of the possibilities is that the continued preferred usage of SAILSx1

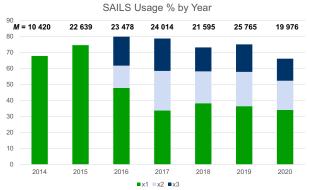


FIG. 5. SAILS usage frequency by year for SVR, FF, and TOR warning decisions.

(versus the MESO-SAILS options) is due to forecasters not wanting to overly increase the length of the volume scan. While this is relevant for all severe weather warning types, it is of particular importance for some wind and most hail events. Given that damaging wind events such as microbursts, as well as large hail events, are often warned by observing descending reflectivity cores and observations of midlevel reflectivity convergence (Roberts and Wilson 1989; Schmocker et al. 1996), it makes intuitive sense that increasing the volume update time too much can be seen as a detriment.

Additionally, automated hail algorithms such as MESH (maximum estimated size of hail; Witt et al. 1998) also utilize full-volume data and are frequently used by WFO forecasters as real-time warning guidance for both hail and summer wind warnings (due to the fact that melting hail with high freezing-level environments can enhance downdraft magnitudes; Straka and Anderson 1993). MESH is generated as a function of the frequency of volume scan updates, meaning that heavier MESO-SAILS usage effectively slows down the update rate of volumetric products such as MESH. Given that MESH is heavily used for many summer hail and wind warnings, especially in the southern and eastern United States, less-frequent selection of MESO-SAILS in these situations would make sense.

a. SVR warnings

Figure 6 shows POD, FAR, and MLT for SVR warnings versus SAILS usage status at the estimated time of warning decision—SAILS off and SAILS on, with the latter further subdivided into SAILSx1, x2, and x3. The short horizontal bars indicate the mean values, and the solid vertical bars are the 95% confidence intervals. For POD and FAR, which are binomial proportion calculations with scores ranging from zero to one, the confidence intervals were estimated by the Wilson score method (Wilson 1927). We chose this formulation due to its conservativeness (on average), coverage probability being consistent and close to the nominal level, and relatively good accuracy even for small sample sizes (Pobocikova 2010). For MLT, the confidence intervals were calculated from $\pm t_{95}\sigma M^{-1/2}$, where M is the number of samples, σ is the

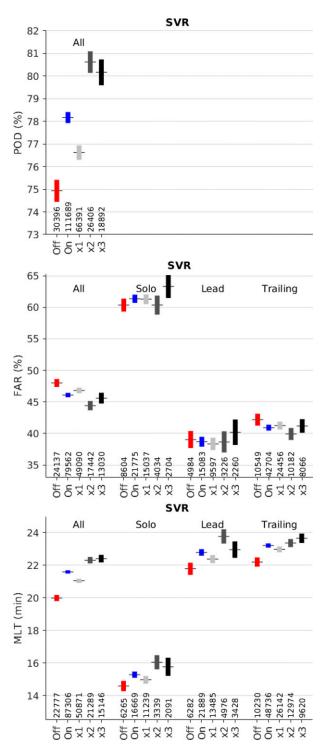


FIG. 6. SVR warning performance metrics vs SAILS usage status. The number of samples per category is displayed at the bottom of each plot. (Colors are used to provide additional visual differentiation between the status categories.) The short horizontal lines denote the mean, and the solid vertical bars are the 95% confidence intervals. (See text for details.)

standard deviation, and t_{95} corresponds to the number of standard deviations from the mean needed to encompass 95% of the Student's t distribution with M-1 degrees of freedom (e.g., Rees 2001). As M grows large, this formula converges to the familiar 95% confidence interval for the mean of a normal distribution with known standard deviation, $\pm 1.96\sigma M^{-1/2}$.

For the all-warning category in Fig. 6, the improvement in SVR warning performance was statistically significant with SAILS on versus SAILS off—POD was higher, FAR was lower, and MLT was longer. (Differences in performance metrics are statistically significant if the confidence intervals of the metrics do not overlap.) Furthermore, the SAILSx2 and x3 results exhibited higher skill than the SAILSx1 results, with no statistically significant difference between x2 and x3. One possible explanation for this result is that, for SVR warning decisions, the benefit of adding extra base scans in the VCP reached saturation with SAILSx2. This is because the volume scan update time also increases with the added base scans (a factor that is analyzed further in the appendix). This suggests that the slower volume updates negated any performance boost provided by the third added base scan.

The parsed FAR results show a complex situation, with only the trailing warning subclass coming close to a statistically significant better performance for SAILS on compared to SAILS off. Perhaps for solo and lead warnings, a 3D volumetric awareness of the evolving storm and/or increased assessment of cutting-edge storm features (e.g., ZDR arc) are more critical than simple quick base scan updates in deciding whether or not to pull the trigger on an initial warning. The timing of the decision, however, appears to benefit from SAILS as seen in the longer lead times.

Because storm morphology and occurrence rate vary by geographic region and season, and warning methodology and culture can differ between WFOs (Andra et al. 2002; Smith 2011), the relationship between SAILS usage and warning performance can vary from office to office. We checked for such an effect by computing the warning performance metrics separately for each of the 115 CONUS WFOs. Only 22 had statistically significant (at the 95% confidence level) differences in POD between the SAILS-off and -on cases. For 19 out of those 22 WFOs, POD was higher with SAILS on. For FAR, the number of WFOs with statistically significant differences was 14, with 11 having lower FAR with SAILS on. For MLT, the number of WFOs with statistically significant differences was 19, with 16 having longer MLT with SAILS on. Thus, even though we are focusing on the overall CONUS results in this paper, it is clear that there is substantial heterogeneity at the WFO level, which we will need to understand better in order to provide useful feedback for operational purposes. We leave this work for the future, with direct input from the WFO meteorologists.

SVR warnings are verified by the presence of a thunderstorm wind or hail event (NWS 2009), which means that POD and MLT (but not FAR) can be computed separately for each event type. Figure 7 shows the consequent POD and MLT results for hail and thunderstorm wind. The overall statistical trends were similar to the aggregate results of Fig. 6; however,

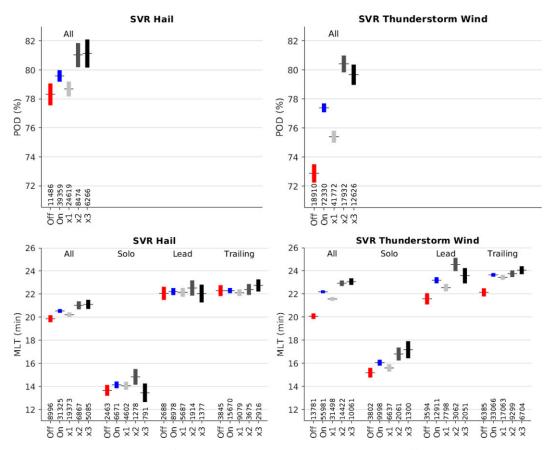


FIG. 7. SVR POD and MLT performance vs SAILS usage status, separated by event type.

the improvements in SVR POD and MLT performance that were associated with SAILS usage were noticeably greater for thunderstorm wind compared to hail. A possible factor is that volumetric update rates are more critical for descending hail cores and automated hail detection algorithms than rapid base scan updates. Large hail is often observable at midlevels, manifesting as a descending reflectivity core before reaching the ground (e.g., Donavon and Jungbluth 2007). When combined with a forecaster's knowledge of the environmental freezing levels, this can become more important during the warning process than rapidly updating base scans. These results agree well with earlier findings from PAR-based severe thunderstorm case studies (Bowden and Heinselman 2016). In short, this is a strong argument for rapid volumetric updates (e.g., via MRLE or a future PAR system) rather than simply rapid base scan updates. However, it should also be noted that warning for hail can be somewhat automated by diagnostic, algorithmic-based guidance such as maximum estimated size of hail (MESH), probability of severe hail (POSH), etc. While these incorporate volumetric data, SAILS data could decrease the effective algorithmic update rate due to slower higher-elevation update rates. It is important to remember, though, that wind hazards do not have the same diagnostic parameters available to forecasters as hail, making the SAILS updates possibly more useful for SVR wind warnings.

Severe winds, on the other hand, can manifest in several different ways. For example, in a pulse storm microburst situation, a descending reflectivity core can be equally useful for a wind threat as it is in other scenarios for hail threats, depending on the thermodynamic environment. However, there are other situations where severe winds can be assessed at the base scan level without as much of a need for rapid volumetric scanning. These scenarios are often prevalent in rapidly moving bow echoes and forward-propagating quasi-linear convective systems (QLCSs) where a rear inflow jet is forcing mesoscale severe winds at the surface, either with the onset of precipitation or as part of a gust front (Markowski and Richardson 2010). In these cases, rapid base scan updates can be particularly useful for a forecaster.

b. FF warnings

Figure 8 displays POD, FAR, and MLT for FF warnings versus SAILS usage status at the estimated time of warning decision. As with SVR warnings, SAILS utilization generally corresponded with improved warning performance, and SAILSx3 was associated with improved warning performance over SAILSx1 in the all-warning category. In addition, SAILSx3 did significantly better than SAILSx2 for POD and MLT, although this gain for MLT comes mainly from the trailing warning category. This implies that a scan strategy

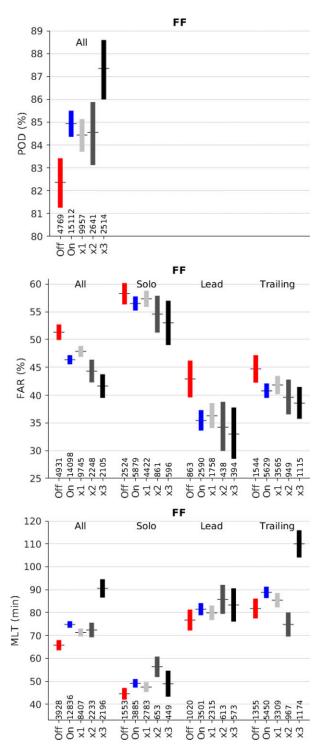


FIG. 8. FF warning performance vs SAILS usage status.

providing faster base scan updates compared to SAILSx3 might lead to even better FF POD and MLT performance (at least for larger systems, e.g., MCSs, that may be more associated with trailing warnings). Note that a recent flash flood model study showed that the peak stream discharge estimate

was 10% lower using 5- versus 1-min radar volume updates (Wen et al. 2021); it would be valuable to deconvolve the effects of faster base scan update from those of faster volume scan updates.

In a parallel study on the effects of SAILS on quantitative precipitation estimation (QPE; Kurdzo et al. 2021), we found that SAILSx3 was a statistically significant indicator for more accurate QPE compared with SAILS-off and other SAILS/MESO-SAILS modes. This was true across several OPE methods, which points to the possibility that faster base scan updates than provided by SAILSx3 might lead to even more accurate QPE. The importance of the base scan cannot be overstated for QPE and FF warnings, since QPE on the WSR-88D is calculated using the hybrid base scan only (Fulton et al. 1998). In a dynamic, rapidly evolving heavy rainfall situation, the faster base scan updates allow for shorter integrations of a given rainfall rate, effectively increasing the spatial resolution of rainfall rate estimation along with the obvious improvement in temporal resolution. Additionally, convective heavy rainfall is often caused by "training" storms over a given area. These cases would also conceivably benefit from faster base scan updates for QPE totals because of the expected faster forecaster recognition of FF potential.

With increased QPE accuracy, we would expect to see better FF warning performance. This argues that faster base scan updates would conceivably further improve FF warning statistics. Additionally, assuming FF warning performance is related to QPE accuracy, the use of vertical profiles of reflectivity (VPR; e.g., Kirstetter et al. 2010) in future operational scenarios would conceivably also increase the performance of both. However, VPR benefits would only be realized in the event of rapid *volumetric* scan updates; again, an argument for the benefits of PAR weather radar architectures in the future.

c. TOR warnings

Figure 9 shows POD, FAR, and MLT for TOR warnings versus SAILS usage status at the estimated time of warning decision. Again, SAILS utilization was generally associated with improved warning performance. For POD and FAR, there were statistically significant differences between the SAILS-off and -on cases. POD significantly increased monotonically with SAILSx1, x2, and x3, while for FAR and MLT, SAILSx3 was associated with statistically better performance compared to SAILSx1 and x2. Nonetheless, the SAILSx1 and x2 cases were statistically indistinguishable for FAR and MLT, which raises the question of whether there is utility in selecting SAILSx2 over x1 in TOR warnings given that skipping over x2 to x3 showed the best statistical performance.

As with SVR and FF, the FAR and MLT performances were significantly worse for solo TOR warnings relative to lead and trailing warnings. As we discussed earlier, this may be due to solo warnings tending to be associated with faster passing, more transient events. This is consistent with TOR warning performance being worse for disorganized storms and OLCSs compared to supercells (Brotzge et al. 2013).

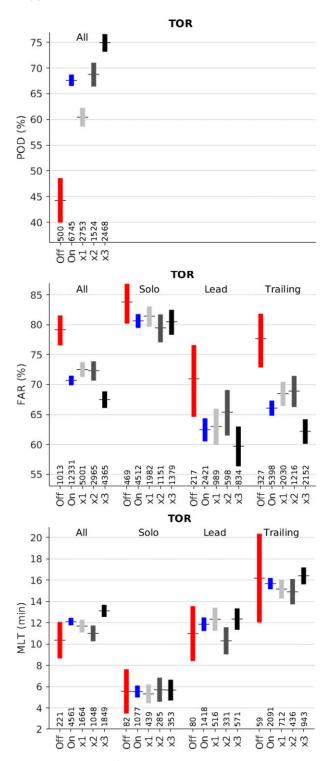


FIG. 9. TOR warning performance vs SAILS usage status.

However, note that the solo warning category could also include situations where a TOR warning transitioned directly to a SVR warning following a spotter report of no low-level rotation and/or a noticeable weakening in the radar velocity signature.

Relative to SVR and FF, TOR events were far fewer in number, so the confidence intervals in Fig. 9 were correspondingly larger. They became even bigger when parsed into subcategories, especially for the sparser SAILS-off cases, with the further division of data points. Even so, the reduction of FAR with SAILS usage was statistically significant for the leading and trailing warning categories.

For MLT, the situation was more ambiguous. There was no clear statistical separation between the SAILS-off and SAILS-on cases. However, under the all-warnings category, the SAILSx3 case yielded a statistically significant increase in MLT compared with all of the other modes.

Taken overall, the Fig. 9 results indicate that, as with FF warning decisions, providing forecasters with more frequent base scan updates than given by SAILSx3 may help improve TOR warning performance further. This agrees, at least conceptually, with the multitude of studies utilizing experimental mobile rapid-scan radar systems for observations of tornadoes and tornadogenesis (e.g., Bluestein et al. 2007; Wurman et al. 2007; Kosiba et al. 2013; Kurdzo et al. 2017). The incredibly rapid changes that can occur in the low-level mesocyclone, the rear-flank downdraft, and the surrounding wind field seen in several rapid-scan radar studies suggests that ~90-s updates of the base scan do not provide enough detail to forecasters issuing warnings (Wilson et al. 2017). These update rates are particularly critical in the most rapidly forming and shorterlived tornadoes, such as those in QLCS or tropical environments. Rapid tornadogenesis can also occur in supercells, especially recently developed supercells in environments primed for tornadoes, as well as in cases of frequent, cyclic tornadogenesis. The precursors to tornadogenesis in these scenarios can be short-lived and subtle (e.g., Sessa and Trapp 2020)—in these situations, the QLCS mesovortex three ingredients method (Schaumann and Przbylinski 2012) can be useful; however, faster scan updates would likely be even more helpful. An important future step for this work will be to isolate these rapidly evolving phenomena from the larger dataset in order to analyze the effects of SAILS and, eventually, MRLE.

We know from past studies that TOR warning performance is strongly related to the tornado's enhanced Fujita scale (EF) damage rating number, with greater warning skill associated with higher EF-scale ratings (e.g., Simmons and Sutter 2005). Therefore, we tried parsing our POD and MLT results in this way. (Such parsing is not possible for FAR, since false alarms do not have associated tornado events.) From past experience (Cho and Kurdzo 2019a,b), we knew that EF0 and EF1 tornadoes tend to have similar POD and MLT statistics, while the occurrence rates for EF3-and-above tornadoes dwindle rapidly. Thus, we divided the results according to three groups—EF0–1, EF2, and EF3–5. The results are shown in Fig. 10.

The EF0-1 results were quite similar to the overall results in Fig. 9, which is to be expected since this group encompassed most (89%) of the tornado events. As for EF2, the only clear difference amongst the various SAILS usage statuses was that the SAILS-on case yielded significantly higher POD than the SAILS-off case, and SAILSx3 had the highest POD. With EF3-5, there were not enough data

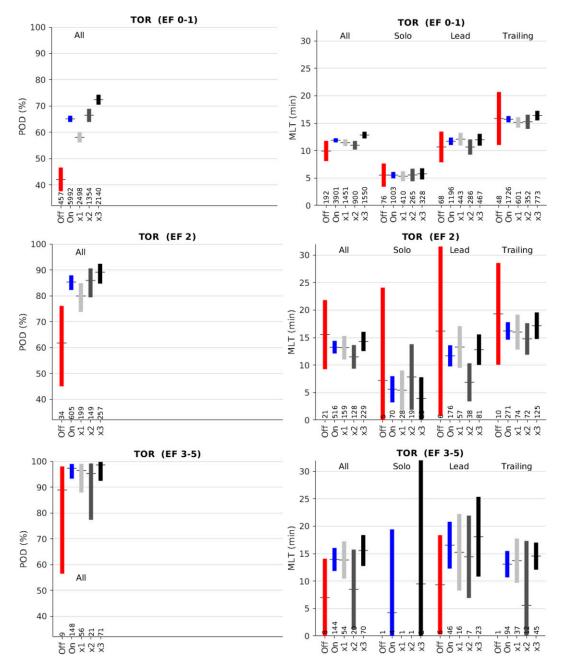


FIG. 10. TOR POD and MLT performance vs SAILS usage status, separated by EF number groups. Results with sample size fewer than two were not plotted.

points to differentiate POD and MLT performance between any of the SAILS status categories.

Since the SAILS-off category contained far fewer events than the SAILS-on group, we tried expanding the database for the former by moving the study period start dates for all sites earlier, i.e., prior to the deployment of SAILS. We set this new start date to be 16 May 2013, which is when the WSR-88D dual-polarization upgrade finished deployment at all CONUS sites, since we did not wish to mix in yet another background variable. This data addition (391 more tornado

events and 843 more warnings in the SAILS-off category) did not, however, create any new statistical distinctions in TOR warning performance among the SAILS usage categories. Another reason why we did not attempt to go back further in time is a sharp decline in POD between 2011 and 2014 that has been attributed to a concerted effort to reduce TOR warning FAR (Brooks and Correia 2018). This significant change in overall TOR warning statistics across different time periods is a reminder that radar data are just one of many variables that impact the ongoing effort to reduce casualties and

damage by providing better severe weather warnings to the public.

4. Case studies of SAILS impacts

Specific case studies can provide practical insights, lend credence to, and offer better understanding of large-data statistical analysis results. At the Norton, Massachusetts (BOX) WFO, SAILS was of great benefit in multiple cases over the last several years. In particular, two cases stood out: a tornadic case on Cape Cod, Massachusetts, on 23 July 2019, and a non-tornadic supercell in Middlesex County, Massachusetts, on 6 July 2020. A discussion of how SAILS was helpful in these cases is provided in the following subsections.

a. Tornadic storm on 23 July 2019—Cape Cod, Massachusetts

The 23 July Cape Cod tornadoes were a particularly major southern New England weather event in 2019, spawning multiple waterspouts and three EF1 tornadoes in Barnstable County, Massachusetts. This was a significant event in part because prior to this date, only three known tornadoes had ever occurred on Cape Cod (an EF0 and two F1 tornadoes). Over the span of 13 min, this historical number was doubled to six with the rapid succession of tornadoes on 23 July. The morning environment (not shown) was characterized by low convective available potential energy (CAPE) and high shear, with surface-based CAPE (SBCAPE) values less than 500 J kg⁻¹, but 0-1-km storm-relative helicity of $100-150 \text{ m}^2 \text{ s}^{-2}$, in a narrow axis between the tip of Long Island, New York, and Cape Cod, Massachusetts. Of note was the strong 0-6-km shear vector from the southwest at 50–60 kt (25.7–30.9 m s⁻¹). The Storm Prediction Center (SPC) mesoanalysis (Bothwell et al. 2002) database (https://www.spc. noaa.gov/exper/mesoanalysis) was used to determine these values.

During the entirety of the event, the NWS forecasters were using SAILSx3 on the KBOX WSR-88D, meaning three out of every four 0.5° elevation scans were SAILS scans. In the figures in this section, "base scan" is the term used for the first 0.5° scan in a volume, while "SAILS scans" are the three added SAILSx3 scans within the volume. This means that the base scans approximately represent the scans that would have taken place had SAILS not been turned on (minus the slower volume update rate from SAILS). As shown in Fig. 11, at 1528:26 UTC, the base scan radial velocity showed mostly convergence in the area that developed into the first tornadic circulation over land (the radar is 65 km to the northwest of the circulation). A tornado warning was ongoing from when the circulation had been farther to the southwest, but a new tornado warning had not yet been issued. The storm had previously tracked over coastal waters for over two hours and there was no confirmation of damage anywhere on land. (Waterspouts had been spawned, but the observation reports had not reached the forecasters at that point.) Although this "primed" the forecasters for severe weather and possible tornadoes, the warning forecaster was strongly incorporating radar signatures to make warning decisions. It should also be

noted that given an ongoing TOR warning was in effect, this would be considered a "trailing" case.

The next base scan, at 1534:28 UTC, showed rotation starting to develop along a kink in the convergence, with a maximum mesocyclone (not gate-to-gate) radial velocity (ν) difference $\Delta \nu$ of 38 m s⁻¹. Over the following three SAILS scans, before the next base scan, the $\Delta \nu$ rapidly increased to 41 m s⁻¹, then 48 m s⁻¹, and finally 49 m s⁻¹, respectively.

During this period of relatively rapid strengthening of the velocity couplet, the NWS forecasters prepared and issued a new tornado warning at approximately the time of the 1539:03 UTC SAILS scan, which displayed the 49 m s⁻¹ Δv . This tornado warning yielded an 18-min lead time for the first tornado (Figs. 12a,b) as well as a 21-min lead time for the second tornado (Figs. 12c,b). A third tornado occurred to the northwest of the second tornado at 1610 UTC. These tornadoes appeared to result from multiple mesocyclones associated with a cyclic supercell. Additionally, making use of the SAILS scans, a special marine warning was issued for waterspouts at 1544:08 UTC, during the second of three SAILS scans for the 1540:42 UTC volume. A waterspout occurred just offshore approximately 6 min later. The forecasters who worked this event believe strongly that SAILS aided them in making decisions about when to issue the tornado and special marine warnings. Seeing the rapid evolution of the mesocyclone, including both the intensification and spatial contraction of the Δv maxima between the 1534:28 and 1539:03 UTC scans, gave the forecasters added confidence in issuing these timely warnings.

b. Non-tornadic supercell on 6 July 2020—Middlesex County, Massachusetts

The 6 July 2020 event was characterized differently than the previous event in that SBCAPE values were larger at 1000–1500 J kg⁻¹ while 0–6-km shear was not as strong from the northwest at 30–35 kt (15.4–18.0 m s $^{-1}$), but this was still marginally favorable for supercells. The 0-1-km stormrelative helicity values were above 100 m² s⁻² and lifting condensation levels (LCLs) were 750-1000 m in southeastern New England. These parameters were taken from the SPC Mesoanalysis page using a regional sector domain. Storms moving south-southeastward from New Hampshire were developing weak rotation at 1-1.5-km beam heights AGL by 0000 UTC. A severe thunderstorm warning, which did not include a "Tornado Possible" tag, was initially issued by WFO BOX at 0037 UTC as a severe thunderstorm was expected to move into Massachusetts from New Hampshire. By 0051:26 UTC, a small hook echo was apparent along the rear flank of a supercell, with a collocated midlevel mesocyclone (Figs. 13a,b). In this case, the radar is approximately 130-160 km south-southeast of the storm of interest. The difficulty in this case for the NWS forecasters was whether or not to issue a tornado warning, especially with an environment marginally favorable for tornadoes and strong rotation rapidly developing in a supercell. Additionally, with the radar beam height above 1 km at 130-km range, lowlevel mesocyclone determination was exceedingly difficult, KBOX - 07/23/2019 - 1528:16 UTC

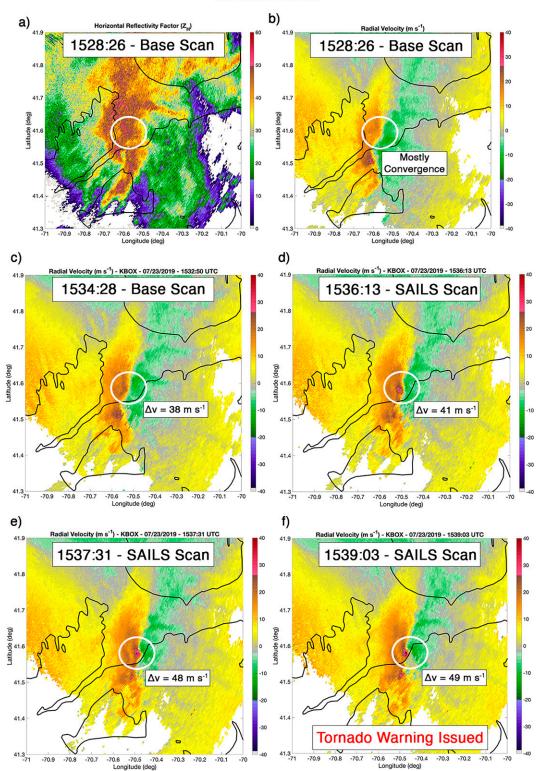


FIG. 11. (a) Horizontal reflectivity Z_H and (b) radial velocity v at the 1528:26 UTC base scan 23 Jul 2019; (c)–(f) v for subsequent base and SAILS scans before the issuance of the new tornado warning.

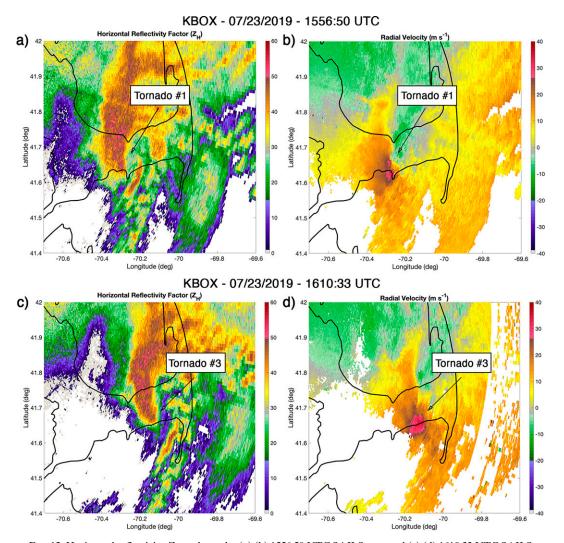


FIG. 12. Horizontal reflectivity Z_H and v at the (a),(b) 1556:50 UTC SAILS scan and (c),(d) 1610:33 UTC SAILS scan, the times of two of the three tornadoes on Cape Cod on 23 Jul 2019.

leading to the need to make assumptions of what may have been happening at low levels based on midlevel sampling (e.g., Togstad et al. 2004).

Two parameters are calculated in order to quantify rotation in this case. First, gate-to-gate shear $(\Delta v_g; \text{m s}^{-1})$ is defined as the maximum difference in radial velocity between two adjacent gates/azimuths. One gate in range in either direction along the radial is considered permissible for qualifying as gate-to-gate, but the radials must be adjacent. Second, a pseudo-vorticity (pseudo- ζ ; s⁻¹) is calculated by finding the maximum difference between outbound and inbound velocities within the mesocyclone, multiplying by two in order to account for the two spatial dimensions, and dividing by the distance in km between the two maxima. This technique has been documented in the literature as a simple proxy for mesocyclone intensity (e.g., Coleman et al. 2018). Past experience has shown that different rotational metrics provide varying levels of diagnostic benefit for different situations. The different metrics calculated and presented for our two case studies were ones that

were most useful for us in making the warning decisions. These calculations were generally performed mentally "on the fly" by forecasters (but not to the precision presented here—the real-time mental computations were more approximate).

Between the 0051:26 UTC base scan (Figs. 13a,b) and the 0055:00 UTC base scan (Fig. 13c), the maximum inbound and outbound velocities become gate-to-gate, maximizing the pseudo- ζ at 29.0 s⁻¹, a near-doubling of the previous value of 16.1 s⁻¹. The Δv_g also rose from 15.0 to 22.0 m s⁻¹. A severe weather statement was issued at 0059 UTC with a "tornado possible" tag, to indicate that tornado potential existed but a tornado was not imminent; close monitoring continued. By the next base scan at 0058:47 UTC (Fig. 13d), the pseudo- ζ had dropped to 17.9 s⁻¹ and the Δv_g had dropped to 18.5 m s⁻¹, and further weakening continued at 0102:46 (Fig. 13e). However, by 0106:28 UTC (Fig. 13f), the pseudo- ζ and Δv_g had increased again to 16.5 s⁻¹ and 19.0 m s⁻¹, respectively.

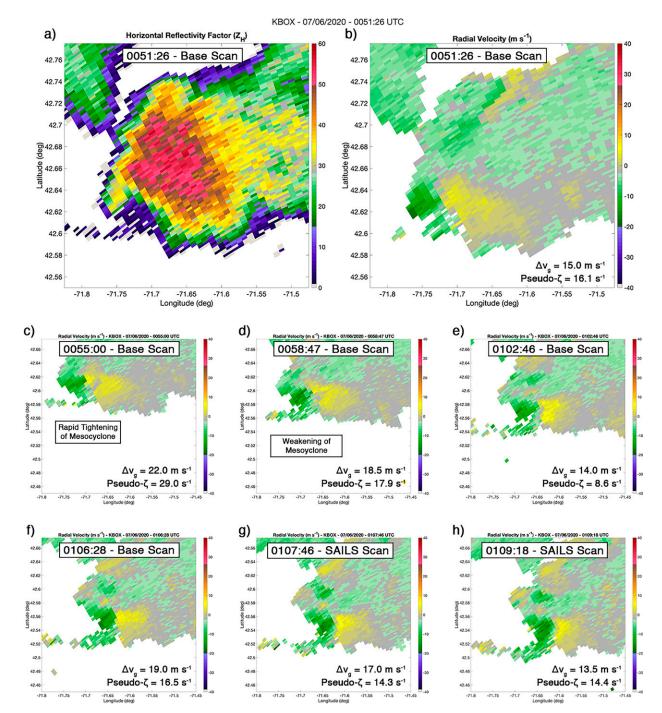


FIG. 13. (a) Z_H and (b) v at the 0051:26 UTC base scan 6 Jul 2020; (c)–(h) v for subsequent base and SAILS scans during the evolution of the mesocyclone.

The radar was not yet in SAILS mode at this time—the forecasters believed that if severe storms developed, the likely threats would be wind and hail such that maintaining a more frequent volume scan update frequency would benefit warning decisions for wind and hail. However, after evidence accumulated that a supercell structure was developing, a decision was made at 0106:30 UTC to change to SAILSx3 in order to

maximize the base scan update rate. The sequence of SAILS and base scans from 0110:38 to 0122:24 UTC are in Fig. 14. By activating SAILSx3 at this point, a closer investigation was possible, as it provided the forecasters more frequent base scan updates in order to more rapidly monitor trends. From the new "peak" in intensities at 0106:28 UTC until the next base scan at 0112:32 UTC (Fig. 14b), the pseudo- ζ and Δv_g

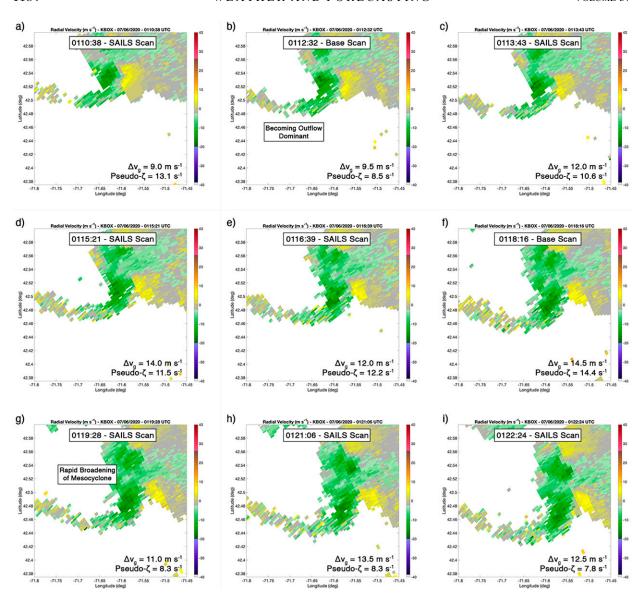


FIG. 14. Radial velocity v for a consecutive series of base and SAILS scans during the evolution of the supercell mesocyclone on 6 Jul 2020.

values dropped by $9.5~{\rm s}^{-1}$ and $8~{\rm m~s}^{-1}$, respectively. Importantly, the SAILS scans allowed the forecasters to trace the steady decline in intensities. A time series of the pseudo- ζ and Δv_g values is shown in Fig. 15. The green shading shows SAILS scans, while the blue shading shows base scans. The first set of SAILS scans corresponded to this time of rapidly decreasing intensity, after the secondary peak at the 0106:28 UTC base scan in Fig. 13f.

Through the use of SAILS, the NWS forecasters were able to more-closely follow the trends in mesocyclone intensity at a time of rapidly changing conditions. It should be emphasized that it is also important to look at the use of SAILS in the sense of *non-tornadic* successes, and not just tornadic successes. Much attention is given in the literature to POD and warning lead times for tornadoes, but the FAR is an

important measure to keep to a minimum. This case is an example of how SAILS can help decrease the FAR by helping monitor rotation trends more accurately and not issuing tornado warnings when they are not warranted. This type of improvement can add to the public's confidence in tornado warnings.

In addition, it is important to note that the storm in question, as well as the adjacent storms, were associated with more than a dozen wind damage reports from southern New Hampshire into eastern Massachusetts. Forecasters issued a number of SVR warnings during this case that arguably allowed for more of a "focus" on the straight-line wind hazards for the public. Not only were TOR warnings rightfully not issued, partially due to the use of SAILS (at least in the opinion of the forecasters), but the appropriate decision between a

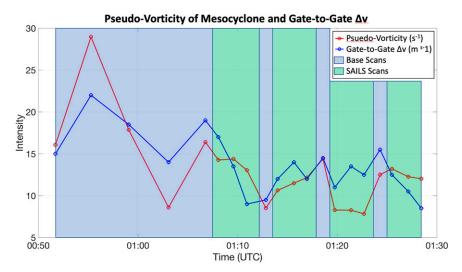


FIG. 15. Pseudo-vorticity (red line) and gate-to-gate Δv (blue line) values from the 6 Jul 2020 supercell mesocyclone. Times of base scans are highlighted by a blue background, while periods of SAILS scans (SAILSx3 in this case) are highlighted by a green background. Note the rapid decrease in intensities during the SAILS scans surrounding 0110 UTC, which aided in the decision not to issue a tornado warning.

TOR warning, a SVR warning, and no warning at all was able to be made. While we have focused on two TOR cases in this section, similar impacts are supported by our evidence for SVR wind warnings specifically. As with TOR warnings, correct decisions on issuing SVR warnings help build public confidence in the NWS warning process.

Finally, with SAILS turned on, forecasters tend to see slight increases and decreases in the trends of rotation more often than they would with regular base scans. We believe this helps tremendously, as velocity signatures associated with Northeast U.S. tornadoes are typically subtle and short-lived. The additional base scans provided by SAILS help fill that void. Prior to SAILS, many of these signatures were not detected, simply because they occurred in between the base scans provided by VCPs 12/212. SAILS helps the warning forecaster maintain higher situational awareness in this regard.

5. Summary discussion

Our statistical analysis on 2014–20 data showed that SAILS usage on WSR-88Ds was associated with statistically significant improvements in SVR, FF, and TOR warning performance. Although direct causality cannot be proven as warning decision-making incorporates many factors, we were able to discount potential "false signal" generators such as the relationship between radar coverage and SAILS usage. Results are generally consistent with contemporaneous studies for tornadoes (Bentley et al. 2021) and severe thunderstorms and tornadoes (Kingfield and French 2022). Two case studies helped demonstrate the statistical results by providing insight into how SAILS helps forecasters make warning decisions. While the warning subcategory (solo, leading, trailing) statistics may be used to infer (by proxy) of storm-type influence on warning performance, we plan to investigate this issue more directly in the future by determining storm

type on an event-by-event basis. We also plan to delve deeper into WFO-specific statistics, in order to provide feedback to regional offices and individual WFOs.

As noted, SAILS increases the base scan update rate at the expense of volume scan update rate. That the relative importance of the two kinds of update rates may vary depending on severe weather type is hinted at by the different effectiveness of MESO-SAILS in improving warning performance for SVR versus FF and TOR. With SVR warnings, SAILSx3 did not provide better warning performance compared to SAILSx2, whereas with FF and TOR, SAILSx3 was apparently more beneficial than SAILSx2. Taken together with the AVSET results (appendix) that showed statistically significant correlation between faster volume update rates and SVR warning performance, it seems that volume update rate may be relatively more important for SVR than for FF and TOR. This situational dependence highlights the benefit for a potential future WSR-88D replacement to have flexible real-time adaptive scanning capability (as would be better accommodated by a PAR) in order to optimize the scan strategy at any given time in any given azimuthal sector. Very rapid updates for both base and volume scans would be ideal, but the radar cost will likely increase accordingly, and there may also be a limit to how much information bandwidth a forecaster is effectively able to handle without further advancements in rapid data processing and analysis tools. Data overload has long been considered a major challenge for operational weather forecasters. This argues for the development of advanced radar data processing techniques, potentially with artificial intelligence/machine learning, to reduce data overload on human forecasters and distill the most pertinent results for forecaster decision-making, including real-time calibrated probabilistic guidance.

Acknowledgments. We would like to thank Dave Smalley (MIT Lincoln Laboratory) for assistance in obtaining supplemental data, as well as Kurt Hondl (NOAA/National Severe Storms Laboratory) for supporting this project. We would also like to thank Katie Wilson (OU/NSSL Cooperative Institute for Severe and High-Impact Weather Research and Operations) and two anonymous reviewers for constructive comments and suggestions. This paper has been approved for public release: distribution unlimited. This material is based upon work supported by the National Oceanic and Atmospheric Administration under Air Force Contract FA8702-15-D-0001. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration.

Data availability statement. Data analyzed in this study are openly available at the following locations: NOAA NCEI storm events data archive, https://www.ncdc.noaa.gov/stormevents/; Iowa Environmental Mesonet NWS Watch/Warnings archive, https://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml; and the NOAA NCEI NEXRAD data archive, https://www.ncdc.noaa.gov/nexradiny/.

APPENDIX

AVSET Effects

As mentioned in the main text, AVSET was available and on by default at all sites during our study period. This means that every SAILS usage category contained a range of volume update times rather than each one having a sharply defined volume update period. These variances create an opportunity for us to examine the effects of volume update rate separately from the SAILS usage (variable base scan update) impacts. Previous experiments with a rapid-scan PAR generally showed performance improvements for non-operational SVR and TOR warnings for volume update periods of ~1 min versus ~2 and ~5 min (Wilson et al. 2017). Our current study, which utilizes large sets of operational data for statistical robustness, provides an excellent complement to those earlier analyses.

Since different VCPs have different baseline volume durations, we decided to analyze AVSET effects on one VCP. The best choice was VCP 212, because it was the dominant mode used for all three warning types (Fig. 2). To minimize the impact of any anomalous values, we took the median of the volume scan durations over the 30-min interval leading up to the warning issuance time. Figure A1 shows the ranges of VCP 212 volume update times for each SAILS usage category. As expected, the volume scan durations increased monotonically with SAILSxN, and there was considerable variance within each category. The improvements in warning performance with SAILS on (and especially the cases with increasing SAILSxN) seen in sections 3a–3c are particularly notable and consistent with prior results given

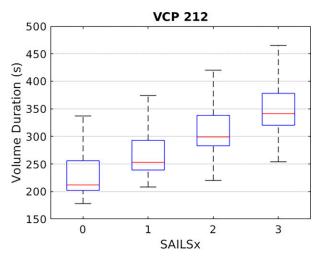


FIG. A1. Volume-scan duration vs SAILS status (SAILSx0 = SAILS off) for VCP 212. The red lines indicate the median values, and the blue boxes mark the extent of the 25th (Q1) to 75th (Q3) percentiles. The maximum extent of the black whiskers corresponds to Q1 - 1.5(Q3 -Q1) (at bottom) and Q3 + 1.5 (Q3 -Q1) (at top), with the plotted whiskers being closest to but inside the maximum extent.

that the volume scan durations are longer with increasing SAILSxN, and speaks to the crucial importance of rapid base scan updates.

To probe the dependence of severe weather warning performance on volume scan duration independent of SAILS usage, we conducted regression analyses-binomial logistic (Berkson 1944) on POD and FAR, as appropriate for events with a binary outcome, and normal linear on MLT. In addition to the volume scan duration (s) as a predictor variable, we also examined event distance from the radar (km), since we knew from past studies that POD is related to event (or source basin for FF) distance from the radar, and that FAR is linked to warning polygon distance from the radar for SVR (Cho and Kurdzo 2020b), FF (Melendez et al. 2018; Cho and Kurdzo 2020a), and TOR (Brotzge and Erickson 2010; Cho and Kurdzo 2019a,b). MLT was only found to be linked to distance from radar for FF (Cho and Kurdzo 2020a), so distance from radar was not included as a predictor variable for the SVR and TOR MLT regression analyses.

The regression analysis results for SVR warnings are listed in Table A1. There are twelve categories for analysis: [POD, MLT, FAR] × SAILS-[Off, x1, x2, x3]. Of the statistically significant results (8 out of 12 categories) for SVR warnings, all but one indicate that faster volume scan updates correlated with higher POD/MLT and lower FAR. These results are independent of the base scan update rates, because they are parsed by SAILS mode. (This observed trend, taken together with the earlier SAILS impact results, implies that there is measurable benefit for SVR warnings to achieving faster volume *and* base scan updates.) The exception was the SAILS-off category for POD. However, the same analysis procedure did not yield any statistically significant results for TOR warnings. For FF warnings, 1 out of 12 categories

TABLE A1. Regression-fit results for SVR warnings. Only results with p values less than 0.05 for all predictor variables (highlighted in bold text) are deemed to be statistically significant. For both the logistic and linear regression fits, positive coefficient estimates for the predictor variables mean that the outcome value rises with increasing predictor variable.

Warning metric	SAILS status	M	Regression output	y intercept	Distance from radar	Volume scan duration
POD	Off	24 711	Coefficient	0.49 ± 0.17	$-8.1 \pm 3.3 \times 10^{-4}$	0.0028 ± 0.0006
		55.047	p value	0.0044	0.016	9.6×10^{-6}
	$\times 1$	55 847	Coefficient	1.3 ± 0.11	$-9.0 \pm 2.2 \times 10^{-4}$	$-7.3 \pm 39 \times 10^{-5}$
	_		p value	1.9×10^{-26}	7.0×10^{-5}	0.85
	$\times 2$	24 509	Coefficient	2.0 ± 0.21	-0.0015 ± 0.0004	-0.0014 ± 0.0006
			p value	6.1×10^{-23}	5.4×10^{-5}	0.016
	$\times 3$	17 669	Coefficient	1.3 ± 0.25	-0.0012 ± 0.0004	$4.2 \pm 6.3 \times 10^{-4}$
			p value	7.7×10^{-8}	0.0087	0.50
FAR	Off	19 538	Coefficient	0.14 ± 0.17	0.0034 ± 0.0003	-0.0020 ± 0.0006
			p value	0.41	7.3×10^{-24}	8.2×10^{-4}
	x1	41 799	Coefficient	-0.60 ± 0.12	0.0022 ± 0.0002	$9.8 \pm 3.8 \times 10^{-4}$
			p value	1.5×10^{-6}	6.3×10^{-20}	0.011
	x2	16 059	Coefficient	-1.2 ± 0.2	0.0027 ± 0.0004	0.0021 ± 0.0006
			p value	2.7×10^{-8}	5.2×10^{-12}	3.7×10^{-4}
	x3	12 199	Coefficient	-0.80 ± 0.24	0.0029 ± 0.0005	$9.6 \pm 6.0 \times 10^{-4}$
			p value	0.0010	5.8×10^{-10}	0.11
MLT	Off	18 324	Coefficient	1900 ± 67	_	-2.5 ± 0.3
			p value	9.8×10^{-175}	_	2.6×10^{-22}
	x1	42 278	Coefficient	1900 ± 46	_	-1.9 ± 0.2
			p value	0	_	2.2×10^{-37}
	x2	19 543	Coefficient	2000 ± 72	_	-1.8 ± 0.2
			p value	1.0×10^{-165}	_	3.4×10^{-18}
	x3	13 948	Coefficient	2100 ± 90	_	-2.0 ± 0.2
			p value	1.0×10^{-120}	_	6.6×10^{-17}

(FAR with SAILS off) had a statistically significant fit, but the sign of the volume duration coefficient was negative, implying that FAR had increased with faster volume updates. The anomalous (or unexpected) results might be explained by circumstances discussed below.

Complicating these regression analyses is the fact that AVSET introduces a certain amount of correlation between storm location (relative to the radar) and volume update time. This occurs because the volume update speed-up is made possible by AVSET skipping the highest elevation scans that have no precipitation signal, a situation that is more likely when a storm is farther away from the radar than close to it. (The correlation is not strict, because the location of highest interest may be farther away from the radar, but there may also be nearby precipitation that forces the radar to scan up to the maximum VCP elevation angle.) For example, a normal linear regression fit of warning polygon distance from radar (km) to volume scan duration (s) in the FF FAR case mentioned in the previous paragraph, yielded a coefficient estimate of $-0.25 \pm$ 0.009 (y intercept = 330 \pm 0.9) and p value of 0, indicating a statistically significant negative correlation between distance to the location of interest and volume-update time. This correlation tends to oppose a correlation between faster scan-volume updates and warning performance, if warning performance degrades with distance from the radar. It may be telling that the most consistent results were for SVR MLT, which has not been statistically linked to distance from the radar.

Because correlation between two predictor variables (distance from radar and volume duration time) could be problematic in interpreting the regression results, we checked for multicollinearity with the variance inflation factor (VIF) in the cases where two predictors were used. VIF never exceeded 1.2, which is far below the commonly used threshold of 5 that indicates a degree of multicollinearity that is problematic for regression analysis (Sheather 2009). Thus, interpretation of the regression results should not be impacted by correlation between distance from radar and volume duration time.

REFERENCES

Andra, D. L., Jr., E. M. Quoetone, and W. F. Bunting, 2002: Warning decision-making: The relative roles of conceptual models, technology, strategy, and forecaster expertise on 3 May 1999. Wea. Forecasting, 17, 559–566, https://doi.org/10.1175/1520-0434(2002)017<0559:WDMTRR>2.0.CO;2.

Bentley, E. S., R. L. Thompson, B. R. Bowers, J. G. Gibbs, and S. E. Nelson, 2021: An analysis of 2016–18 tornadoes and National Weather Service tornado warnings across the contiguous United States. *Wea. Forecasting*, **36**, 1909–1924, https://doi.org/10.1175/WAF-D-20-0241.1.

Berkson, J., 1944: Application of the logistic function to bio-assay.
J. Amer. Stat. Assoc., 39, 357–365, https://doi.org/10.2307/2280041.

Bluestein, H. B., M. M. French, R. L. Tanamachi, S. Frasier, K. Hardwick, F. Junyent, and A. L. Pazmany, 2007: Close-range observations of tornadoes in supercells made with a dual-

- polarization, X-band, mobile Doppler radar. *Mon. Wea. Rev.*, **135**, 1522–1543, https://doi.org/10.1175/MWR3349.1.
- Bothwell, P. D., J. A. Hart, and R. L. Thompson, 2002: An integrated three-dimensional objective analysis scheme in use at the Storm Prediction Center. Preprints, 21st Conf. on Severe Local Storms, San Antonio, TX, Amer. Meteor. Soc., JP3.1, https://ams.confex.com/ams/pdfpapers/47482.pdf.
- Bowden, K. A., and P. L. Heinselman, 2016: A qualitative analysis of NWS forecasters' use of phased-array radar data during severe hail and wind events. Wea. Forecasting, 31, 43–55, https://doi.org/10.1175/WAF-D-15-0089.1.
- Brooks, H. E., and J. Correia Jr., 2018: Long-term performance metrics for National Weather Service tornado warnings. Wea. Forecasting, 33, 1501–1511, https://doi.org/10.1175/WAF-D-18-0120.1.
- Brotzge, J., and S. Erickson, 2010: Tornadoes without NWS warning. Wea. Forecasting, 25, 159–172, https://doi.org/10.1175/2009WAF2222270.1.
- —, S. Nelson, R. Thompson, and B. Smith, 2013: Tornado probability of detection and lead time as a function of convective mode and environmental parameters. *Wea. Forecasting*, 28, 1261–1276, https://doi.org/10.1175/WAF-D-12-00119.1.
- Brown, R. A., J. M. Janish, and V. T. Wood, 2000: Impact of WSR-88D scanning strategies on severe storm algorithms. *Wea. Forecasting*, **15**, 90–102, https://doi.org/10.1175/1520-0434(2000)015<0090:IOWSSO>2.0.CO;2.
- Burgess, D. W., and L. R. Lemon, 1990: Severe thunderstorm detection by radar. *Radar in Meteorology*, R. Atlas, Ed., Amer. Meteor. Soc., 619–647.
- Cho, J. Y. N., and J. M. Kurdzo, 2019a: Monetized weather radar network benefits for tornado cost reduction. Project Rep. NOAA-35, MIT Lincoln Laboratory, 88 pp., https://www.ll. mit.edu/sites/default/files/publication/doc/monetized-weatherradar-network-benefits-cho-noaa-35.pdf.
- —, and —, 2019b: Weather radar network benefit model for tornadoes. *J. Appl. Meteor. Climatol.*, **58**, 971–987, https://doi.org/10.1175/JAMC-D-18-0205.1.
- —, and —, 2020a: Weather radar network benefit model for flash flood casualty reduction. *J. Appl. Meteor. Climatol.*, **59**, 589–604, https://doi.org/10.1175/JAMC-D-19-0176.1.
- —, and —, 2020b: Weather radar network benefit model for nontornadic thunderstorm wind casualty reduction. Wea. Climate Soc., 12, 789–804, https://doi.org/10.1175/WCAS-D-20-0063.1.
- Chrisman, J. N., 2013: Dynamic scanning. NEXRAD Now, No. 22, NOAA/NWS/Radar Operations Center, Norman, Oklahoma, 1–3, https://www.roc.noaa.gov/WSR88D/PublicDocs/ NNOW/NNow22c.pdf.
- —, 2014: The continuing evolution of dynamic scanning. NEXRAD Now, No. 23, NOAA/NWS/Radar Operations Center, Norman, Oklahoma, 8–13, http://www.roc.noaa.gov/ WSR88D/PublicDocs/NNOW/NNow23a.pdf.
- —, 2016: Mid-volume Rescan of Low-level Elevations (MRLE): A new approach to enhance sampling of Quasi-Linear Convective Systems (QLCSs). New Radar Technologies, NOAA/NWS/Radar Operations Center, 21 pp., https:// www.roc.noaa.gov/WSR88D/PublicDocs/NewTechnology/ DQ_QLCS_MRLE_June_2016.pdf.
- Coleman, T. A., A. W. Lyza, K. R. Knupp, K. Laws, and W. Wyatt, 2018: A significant tornado in a heterogeneous environment during VORTEX-SE. Electron. J. Severe Storms Meteor., 13 (2), https://ejssm.org/archives/2018/vol-13-2-2018/.

- Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteor. Soc.*, 74, 1669–1688, https://doi.org/10.1175/1520-0477(1993) 074<1669:TWATWO>2.0.CO;2.
- Donavon, R. A., and K. A. Jungbluth, 2007: Evaluation of a technique for radar identification of large hail across the Upper Midwest and Central Plains of the United States. Wea. Forecasting, 22, 244–254, https://doi.org/10.1175/WAF1008.1.
- Doviak, R. J., and D. S. Zrnić, 1993: Doppler Radar and Weather Observations. Academic Press, 562 pp.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. Wea. Forecasting, 13, 377–395, https://doi.org/10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2.
- Heinselman, P., D. LaDue, D. M. Kingfield, and R. Hoffman, 2015: Tornado warning decisions using phased-array radar data. Wea. Forecasting, 30, 57–78, https://doi.org/10.1175/ WAF-D-14-00042.1.
- Kingfield, D. M., and M. M. French, 2022: The influence of WSR-88D intra-volume scanning strategies on thunderstorm observations and warnings in the dual-polarization radar era: 2011–20. Wea. Forecasting, 37, 283–301, https://doi.org/10.1175/WAF-D-21-0127.1.
- Kirstetter, P.-E., H. Andrieu, G. Delrieu, and B. Boudevillain, 2010: Identification of vertical profiles of reflectivity for correction of volumetric radar data using rainfall classification. *J. Appl. Meteor. Climatol.*, 49, 2167–2180, https://doi.org/10. 1175/2010JAMC2369.1.
- Kosiba, K., J. Wurman, Y. Richardson, P. Markowski, P. Robinson, and J. Marquis, 2013: Genesis of the Goshen County, Wyoming, tornado on 5 June 2009 during VORTEX2. Mon. Wea. Rev., 141, 1157–1181, https://doi.org/10.1175/MWR-D-12-00056.1.
- Kurdzo, J. M., and Coauthors, 2017: Observations of severe local storms and tornadoes with the atmospheric imaging radar. *Bull. Amer. Meteor. Soc.*, 98, 915–935, https://doi.org/10.1175/ BAMS-D-15-00266.1.
- —, J. Y. N. Cho, and B. J. Bennett, 2021: Impact of WSR-88D SAILS usage on quantitative precipitation estimation accuracy. 37th Conf. on Environmental Information Processing Technologies: Radar Technologies and Applications, online, Amer. Meteor. Soc., 10.2, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/378382.
- Markowski, P., and Y. Richardson, 2010: Mesoscale Meteorology in Midlatitudes. John Wiley & Sons, Ltd., 407 pp., https://doi. org/10.1002/9780470682104.
- Melendez, D., K. Abshire, and J. Sokich, 2018: NEXRAD weather radar coverage and National Weather Service warning performance. 2018 Fall Meeting, Washington, DC, Amer. Geophys. Union, Abstract A11K-2394, https://doi.org/10.1002/essoar.10500135.1.
- NOAA, 2020: Weather radar follow-on plan: Research and risk reduction to inform acquisition decisions. Report to Congress, Office of Oceanic and Atmospheric Research, National Oceanic and Atmospheric Administration., 21 pp., https://www.nssl.noaa.gov/publications/mpar_reports/RadarFollowOnPlan_ReporttoCongress_2020June_Final.pdf.
- NWS, 2009: Verification procedures. NWSPD 10-16, Operations and Services, National Weather Service, 83 pp., http://www. nws.noaa.gov/directives/010/archive/pd01016001d.pdf.
- Pobocikova, I., 2010: Better confidence intervals for a binomial proportion. *Communications*, 12, 31–37, 10.26552/com.C.2010. 1.31-37.

- Polger, P. D., B. S. Goldsmith, R. C. Przywarty, and J. S. Bocchieri, 1994: National Weather Service warning performance based on the WSR-88D. *Bull. Amer. Meteor. Soc.*, 75, 203–214, https://doi.org/10.1175/1520-0477(1994)075<0203: NWSWPB>2.0.CO;2.
- Rees, D. G., 2001: Essential Statistics. 4th ed. Chapman and Hall/ CRC, 384 pp.
- Roberts, R. D., and J. W. Wilson, 1989: A proposed microburst nowcasting procedure using single-Doppler radar. *J. Appl. Meteor.*, **28**, 285–303, https://doi.org/10.1175/1520-0450(1989) 028<0285:APMNPU>2.0.CO;2
- Sachidananda, M., and D. S. Zrnić, 1999: Systematic phase codes for resolving range overlaid signals in a Doppler weather radar. J. Atmos. Oceanic Technol., 16, 1351–1363, https://doi. org/10.1175/1520-0426(1999)016<1351:SPCFRR>2.0.CO;2.
- Schaumann, J. S., and R. W. Przbylinski, 2012: Operational application of 0–3 km bulk shear vectors in assessing QLCS mesovortex and tornado potential. 26th Conf. on Severe Local Storms, New Orleans, LA, Amer. Meteor. Soc., P9.10, https://ams.confex.com/ams/26SLS/webprogram/Manuscript/Paper212008/SchaumannSLS2012_P142.pdf.
- Schmocker, G. K., R. W. Przybylinksi, and Y.-J. Lin, 1996: Fore-casting the initial onset of damaging winds associated with a mesoscale convective system (MCS) using the Mid-Altitude Radial Convergence (MARC) signature. Preprints, 15th Conf. on Weather Analysis and Forecasting, Norfolk, VA, Amer. Meteor. Soc., 306–311.
- Sessa, M. F., and R. J. Trapp, 2020: Observed relationship between tornado intensity and pretornadic mesocyclone characteristics. Wea. Forecasting, 35, 1243–1261, https://doi.org/10.1175/WAF-D-19-0099.1.
- Sheather, S., 2009: A Modern Approach to Regression with R. Springer, 392 pp., https://doi.org/10.1007/978-0-387-09608-7.
- Simmons, K. M., and D. Sutter, 2005: WSR-88D radar, tornado warnings, and tornado casualties. Wea. Forecasting, 20, 301– 310, https://doi.org/10.1175/WAF857.1.
- Smith, S. B., 2011: The impact of NWS Weather Forecast Office culture on tornado warning performance. NOAA Office of Science and Technology, Meteorological Development

- Laboratory, 61pp., https://www.nws.noaa.gov/mdl/seminar/ Presentations/November_30_2011.pdf.
- Straka, J. M., and J. R. Anderson, 1993: Numerical simulations of microburst-producing storms: Some results from storms observed during COHMEX. J. Atmos. Sci., 50, 1329–1348, https:// doi.org/10.1175/1520-0469(1993)050<1329:NSOMPS>2.0.CO;2.
- Stumpf, G. J., and A. E. Gerard, 2021: National Weather Service severe weather warnings as threats-in-motion. Wea. Forecasting, 36, 627–643, https://doi.org/10.1175/WAF-D-20-0159.1.
- Togstad, W. E., S. J. Taylor, and J. Peters, 2004: An examination of severe thunderstorm discrimination skills from traditional Doppler radar parameters and Near Storm Environment (NSE) factors at large radar range. Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc., J2.5, https://ams.confex.com/ams/pdfpapers/81462.pdf.
- Weber, M. E., and Coauthors, 2021: Towards the next generation operational meteorological radar. *Bull. Amer. Meteor. Soc.*, 102, E1357–E1383, https://doi.org/10.1175/BAMS-D-20-0067.1.
- Wen, Y., T. J. Schuur, H. Vergara, and C. Kuster, 2021: Effect of precipitation sampling error on flash flood monitoring and prediction: Anticipating operational rapid-update polarimetric weather radars. J. Hydrometeor., 22, 1913–1929, https:// doi.org/10.1175/JHM-D-19-0286.1.
- Wilson, E. B., 1927: Probable inference, the law of succession, and statistical inference. J. Amer. Stat. Assoc., 22, 209–212, https:// doi.org/10.1080/01621459.1927.10502953.
- Wilson, K. A., P. L. Heinselman, C. M. Custer, D. M. Kingfield, and Z. Kang, 2017: Forecaster performance and workload: Does radar update time matter? Wea. Forecasting, 32, 253– 274, https://doi.org/10.1175/WAF-D-16-0157.1.
- Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. Wea. Forecasting, 13, 286–303, https://doi.org/10.1175/1520-0434(1998)013<0286:AEHDAF>2. 0.CO;2.
- Wurman, J., Y. Richardson, C. Alexander, S. Weygandt, and P. F. Zhang, 2007: Dual-Doppler analysis of winds and vorticity budget terms near a tornado. *Mon. Wea. Rev.*, 135, 2392– 2405, https://doi.org/10.1175/MWR3404.1.