Response to Referee #1

We thank Dr. Púčik for his helpful comments. In our reply below, referee comments are shown in black text whereas our corresponding responses are shown in *red italics*.

Summary:

Authors present an interesting research on one of the more challenging tasks for forecasters - how to recognize the potential for severe thunderstorms in weakly sheared / forced environments. They also use a novel approach trying to set thresholds of parameters to differentiate between nonsevere and severe events. Paper is well structured and after some minor revisions to the contents it should be ready for publication.

> *Thank you reviewing our paper. We appreciate your feedback, and a revised manuscript will largely incorporate your comments.*

Introduction

Line 53: I am pretty sure that the criterion for large hail is 2.56 and not 0.56 cm.

> *Thank you for catching this error. It will be corrected in a revision.*
>
> **See line 55 in revision.**

Line 59: Authors use "signal to noise ratio" throughout the paper. Introduction to the term is made here. I am wondering, has anyone used this term in forecasting before? Or are authors introducing it? If not, references should be made.

> *This term is common in climate variability contexts; however, we have not identified any regular forecasting applications in contexts smaller than the seasonal scale. We will add references supporting its use in climate forecasting such as:*
>
> *Sutton, R. T., and D. L. R. Hodson, 2007: Climate response to basin-scale warming and cooling of the North Atlantic Ocean. J. Climate, **20,** 891-907, doi:10.1175/JCLI4038.1.*
> *Hamlington, B. D., R. R. Leben, R. S. Nerem, and K. Y. Kim, 2010: The effect of signal-to-noise ratio on the study of sea level trends. J. Climate, **24,** 1396-1408, doi:10.1175/2010JCLI3531.1.*
> *Trenberth, K. E., 1984: Signal versus noise in the Southern Oscillation. Mon. Wea. Rev., **112,** 326-332, doi:10.1175/1520-0493(1984)112<0326:SVNITS>2.0.CO;2.*
>
> **See lines 62-65 in revision.**

Line 62: Should be represents instead of is represented?

> *This will be corrected in a revision.*
>
> **See line 67 in revision.**

Line 95: SWEAT - Severe WEAther Threat (without and)

> *This will be corrected in a revision.*
>
> **See line 101 in revision.**

Methods

Line 107: I understand that the dataset of weakly forced thunderstorms itself is now published somewhere else, but I strongly suggest that authors at least briefly introduce the definition of weakly forced environment. It would help the reader to better understand the paper.

We will add the table below to help the reader better envision the kinematic and thermodynamic environments that were considered "weakly forced". These convective parameters were computed by Miller and Mote (2017) from composite soundings in Atlanta, GA, on days when each morphological type is dominant. The reader does not necessarily need understand the exact meaning of the "morphological type" to compare the values between the WFT and non-WFT environments.

Table 1. Kinematic and thermodynamic parameters of 1200-UTC composite sounding from Atlanta, Georgia, USA, for each radar-identified morphological type in Miller and Mote (2017). Morphological types classified as WFTs are bolded. All kinematic values are shown in m s^{-1} whereas the units of the thermodynamic parameters are provided in the table.

Type	0–6-km _SHR	0–8-km _SHR	0–12-km Max Wind	0–12-km Mean Wind	ThE _LOW (K)	MLCAPE (J kg^{-1})	Forecast SBCAPE (J kg^{-1})
1	**4.3**	**5.1**	**7.2**	**3.0**	**343.0**	**562**	**1,585**
2	**4.3**	**5.1**	**8.8**	**3.6**	**341.9**	**365**	**1,214**
3	**4.7**	**6.2**	**9.4**	**3.7**	**340.5**	**289**	**1,176**
4	**4.3**	**5.7**	**8.3**	**3.7**	**341.1**	**357**	**1,121**
5	**3.2**	**5.1**	**9.6**	**3.2**	**341.7**	**283**	**1,006**
6	6	7.7	11.6	5.0	339.0	211	973
7	8.2	10.8	16.5	6.1	336.6	66	723
8	4.9	7.7	13.6	3.1	336.0	24	558
9	5.4	8.7	15.4	3.0	330.6	0	32
10	7.9	9.8	13.5	5.8	334.5	0	391

See lines 120-124 and Table 1 in revision.

Line 127: Are you sure you are always creating soundings of pre-storm environments and not soundings that may be contaminated by model simulated convection? While authors subsequently perform a check on the model vs observations performance, has there been any quality control of individual soundings?

It is possible that some soundings experience contamination from convection within the model. However, because the soundings are based on the 0-hr RAP analysis fields any such influence is likely limited. Because the RAP assimilates radar reflectivity and lightning observations from the U.S. (Benjamin et al. 2016), areas of convection in the 0-hr analysis will typically mimic the radar-observed areas of convection. Further, any instances where convective overturning was a source of contamination would be smoothed out by aggregating all storms into SWS, SHS, and nonsevere days. Nonetheless, we will alert the reader to this possibility in Section 2.2 when we discuss the accuracy of the RAP soundings.

See lines 196-206 in revision.

Line 194: How exactly are the results aggregated daily and by the radar sites? Does it mean that you take the average values of parameters for particular day and radar site? This description should be expanded so that reader understands exactly what are the implications of such aggregation.

> *Yes, the convective parameters for all WFTs forming within a radar polygon (Fig. 1) on a single day were averaged (using the mean). This information will be added near line 194.*

> **See lines 223-224 in revision.**

Line 211: Should be precedence, not precedent?

> *Yes, this will be corrected.*

> **See line 241 in revision.**

In general, why are authors even looking at the measures of vertical wind shear when they consider only the weakly sheared environments?

> *We wanted to test whether relative increases/decreases in wind shear (even within "weak shear" environments) may serve to enhanced/diminish severe weather potential. This possibility was not supported by the findings; however, it is still worth including.*

Results

Authors spend a lot of time trying to find the best "threshold" value for each parameter. Have they considered looking at this problem from probabilistic point of view?

> *Though we do spend considerable effort identifying valuable thresholds, the odds ratio (OR), the measure by which the effectiveness of the convective parameters is judged, is itself a probabilistic tool. Rather than solely considering our recommended thresholds, the reader/forecaster is free to directly interpret the relative odds of severe weather at any parameter value from Figures 6 and 8.*

Abbreviations and calculation of different parameters are stated in the Table 1 and then Appendix. However, I still advise authors to at least briefly introduce the mentioned, best discriminating, parameters (beyond their abbreviations) here.

> *We will add brief descriptions of the vertical totals and total totals near line 270 when they are first mentioned as effective differentiating parameters. Meanwhile, MLCAPE and MLLCL are both employed with enough frequency in the atmospheric sciences to be interpreted without further explanation. Because DCAPE, TEI, WNDG, MICROB, PW, PEFF, HGT0, and ApWBZ do not accurately distinguish severe weather days, we are reluctant to dedicate new text to defining all eight in the Results. Interested readers can elect to consult Table 1 or the Appendix.*

> **See lines 303-305 in revision.**

Discussion

Line 359: How would lower freezing level and drier lower troposphere promote more efficient growth of hailstones? The main point here is that melting of hailstones will be less of an issue, which is important particularly for smaller hail sizes.

3

We will augment this text to explain that the freezing level and dry layer are relevant to the melting process.
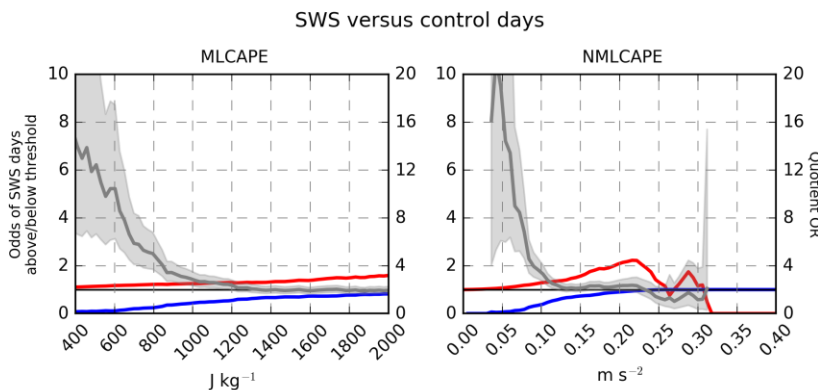
See lines 389-391 in revision.

Same as authors, I was also surprised to see that measures of lapse rates (such as Vertical Totals) perform better than CAPE itself. Apart from possible model errors, I suspect two other reasons for that:

A/ CAPE is spatially variable, more so than the lapse rates. Could it be that the aggregation of soundings and events "smoothed" out CAPE too much?

B/ I presume that in this region of United States, it is easy to get substantial CAPE values owing to the high lower tropospheric moisture content. Then indeed, shape of CAPE (skinny vs fat) that is regulated by lapse rates makes a big difference, with "fatter" CAPE profiles involving stronger updrafts. It would be interesting to see if Normalized CAPE (NCAPE), which is CAPE divided by the depth of convective cloud (EL - LFC) would outperform CAPE by a large margin. I suggest trying out this parameter as well as authors actually have everything they need to calculate it.

We agree with both possibilities and will adapt the manuscript to include them. The possibility that CAPE (and perhaps other variables too) was smoothed out during the daily aggregation process will be mentioned near line 363.

In regards to B, we agree with your logic, and indeed, our dataset already contains the necessary constituent values. We have mimicked Figures 6 and 8 below to compare basic MLCAPE to Normalized MLCAPE. NCAPE is a more effective differentiator at low values (<0.10), meaning WFTs rarely produce severe weather with NCAPEs below these values. Granted, a forecaster would have been unlikely to consider severe weather potential with NCAPEs this low in the first place. Ultimately, it suffers from the same deficiencies as many of the other parameters tested, so we choose not to add the figures below to the paper. Nonetheless, we will revise the manuscript (namely the Tables 1, 4, 5 and the Appendix) to alert the reader than Normalized MLCAPE was also analyzed in this project.

SHS versus control days
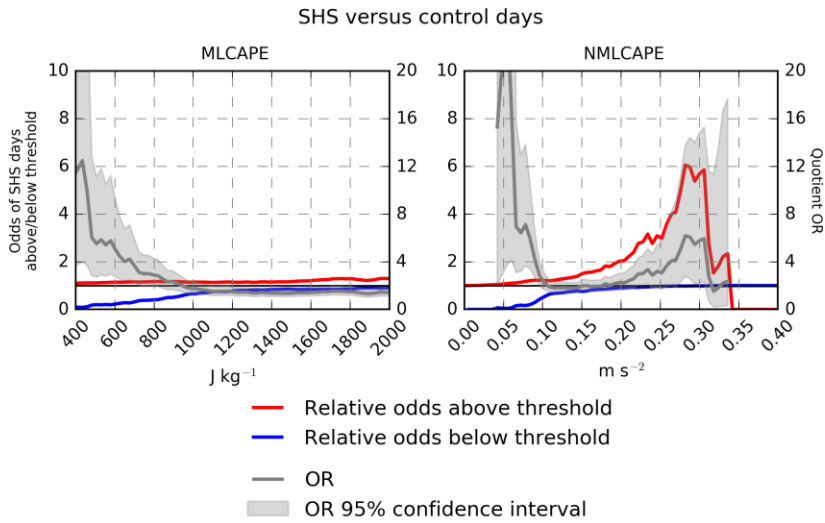
See lines 396-398 as well as Tables 2, 5, 6, and Appendix A in revision.

Changes NOT in response to reviewers:
After submission, we learned it is more appropriate to describe the odds ratio (OR) as indicating that "the odds of an event are X times greater…" rather than "an event is X times more likely…" (e.g., line 223). The revised manuscript will standardize the language describing ORs to reflect the latter.

See lines 30, 36-38, 252-253, 260, 296, 301-306, 320-321, and 446-447 in revision.

Response to Anonymous Referee #2

We thank Referee #2 for their helpful comments. In our reply below, referee comments are reproduced in black text whereas our corresponding responses are shown in *red italics*.

Summary:

In this manuscript the challenge to characterize the severe weather potential in synoptically weakly forced convective environments is tackled. Various convective variables and indices are computed using proxy-soundings of hourly model analyses representing the environment. As observational data base for benign convection the authors' own dataset of radar reflectivities is blended with 'subjective' severe weather reports from Storm Data. The odds ratio is used to measure the conditional likelihood. In essence, the authors find the surprising result (their own wording) that only two convective parameters based on temperatures at two pressure levels (850 and 500 hPa) are able to differentiate between severe and nonsevere convection during such difficult to forecast weather regimes. It is suggested that prominent, heavily moisture-weighted parameters and indices broadly used in severe storm forecasting suffer from inaccuracies in the RAP analysis data.

The manuscript is well-written and merits publication given that the comments below are taken into account.

> *Thank you reviewing our paper, and providing valuable feedback. Please see our responses to your individual comments below.*

Major Points:

1) My main criticism is the use of the RAP analyses providing proxy-soundings to calculate convective variables and indices. Subtle model inaccuracies (or errors) can impede the accurate calculation of convective parameters. The key question is the suitability of the RAP system with dx=13km using presumably a convective parameterization scheme for deep convection for such calculations. I suggest to use proxy-soundings from the convection permitting model HRRR analyses (at least for a subset) and confront your results with those. This extension would considerably broaden the relevance of the manuscript.

> *At the outset of this project, we considered employing the HRRR analysis for the reasons you described. However, NCEI does not maintain a publicly accessible repository of historical HRRR output, likely to due to the enormous storage burden it would pose. The University of Utah does host an independent HRRR archive, but as an ad hoc effort, it does not support FTP downloads, and is designed for smaller-scale projects. Their nascent effort only overlaps with our study period for one year (2015), which is problematic because this may not be an appropriate year for comparison. Ideally, we would perform a subset comparison using a random sample of soundings across the 4-yr period (or at least more than a single year), which unfortunately is not possible given the current state of the HRRR archives.*
>
> *The HRRR is largely a higher-resolution nest of the RAP with the same physical parameterization configuration, except for deep convection (which, as you mentioned, is not parameterized). We would expect the HRRR to resolve similar temperature and moisture fields as the RAP, except near areas of convection (which the HRRR would depict with greater fidelity). Thus, proxy soundings from the HRRR would improve upon the RAP insofar as it reduces convective contamination, which relates to your Major Comment*

#3. We will add text describing the potential of the HRRR to reduce convective contamination in the same addition proposed in our response to Major Comment #3.

UPDATE: *Shortly after offering the reply above during the discussion phase of the review, the University of Utah HRRR archive went offline for much of January. On 29 January 2018, the archive experienced a catastrophic failure, and all archive HRRR files as well as the Google Drive back-up archive were lost. More details can be found at* http://home.chpc.utah.edu/~u0553130/Brian_Blaylock/Pando_archive/Pando_Failure.html. *The revised manuscript maintains its new text describing the advantages that a convection-allowing model, such as the HRRR, can provide to this research. It also notes that at the time of preparation, no such archive is available.*

See lines 196-206 in revision.

2) (line 117ff) More detail is needed describing the data assimilation of RAP. What data is how assimilated in RAP? Any radar reflectivities?

We will supplement the current description of the RAP with more information about its data assimilation from Benjamin et al. (2016), particularly regarding the observational data sources that are assimilated (radar, lightning, aircraft, GOES cloud tops, etc).

See lines 134-139 in revision.

3) (line 128ff) Is the proxy-sounding representative for a storm environment if there is no storm in the model analysis? What happens if there is a shift in space or/and time between observation and model?

The goal of the proxy sounding is to capture a snapshot of the pre-convective environment as close as possible to the time of convective initiation. It is possible that some soundings experience contamination from convection within the model. However, because the soundings are based on the 0-hr RAP analysis fields, any such influence is likely limited. Because the RAP assimilates radar reflectivity and lightning observations (Benjamin et al. 2016), areas of convection in the 0-hr analysis will typically mimic the radar-observed areas of convection. Further, any instances where convective overturning was a source of contamination would be smoothed out by aggregating all storms by day. Nonetheless, we will alert the reader to this possibility in Section 2.2 when we discuss the accuracy of the RAP soundings and the potential of the HRRR to improve upon this limitation.

See lines 196-206 in revision.

Minor Points:

1) (line 59) The term signal-to-noise ratio is awkward in this respect. Is there any reference? Is model error a noise source, too?

This term is common in climate variability contexts; however, we have not identified any regular forecasting applications in contexts smaller than the seasonal scale. Yes, model error does qualify as noise because it is not a process considered by the OR analysis (i.e., convective parameters are taken at face value). We will add references supporting its use in climate forecasting such as:

Sutton, R. T., and D. L. R. Hodson, 2007: Climate response to basin-scale warming and cooling of the North Atlantic Ocean. J. Climate, **20,** *891-907, doi:10.1175/JCLI4038.1.*

Hamlington, B. D., R. R. Leben, R. S. Nerem, and K. Y. Kim, 2010: The effect of signal-to-noise ratio on the study of sea level trends. J. Climate, **24,** *1396-1408, doi:10.1175/2010JCLI3531.1.*

Trenberth, K. E., 1984: Signal versus noise in the Southern Oscillation. Mon. Wea. Rev., **112,** *326-332, doi:10.1175/1520-0493(1984)112<0326:SVNITS>2.0.CO;2.*

See lines 62-65 in revision.

2) (line95) Although a brief explanation of the indices used in this study are given in the appendix, key 'tools' should be explained in the main text. What is SWEAT?

We will add brief descriptions of the vertical totals and total totals near line 270 at first mention as effective differentiating parameters. Meanwhile, MLCAPE and MLLCL are both employed with enough frequency in the atmospheric sciences to be interpreted without further explanation.

See lines 101, 303-305 in revision.

3) (line 108) More detail is needed on the authors' own dataset of radar reflectivities for benign convection that is still under review at the time of this review.

The Miller and Mote (2017) paper describing the development of the WFT dataset is now available in print (https://doi.org/10.1175/JAMC-D-17-0005.1). The current text will be supplemented with more detail describing the radar-based identification of storms; statistical clustering into morphologically similar groups; characterization each group's composite convective environment; and the final designation of each group as WFT/non-WFT. In response to Reviewer #1, we will also be adding a new table better characterizing the convective environments that hosted WFTs.

See lines 114-124 in revision.

Changes NOT in response to reviewers:
After submission, we learned it is more appropriate to describe the odds ratio (OR) as indicating that "the odds of an event are X times greater…" rather than "an event is X times more likely…" (e.g., line 223). The revised manuscript will standardize the language describing ORs to reflect the latter.

See lines 30, 36-38, 252-253, 260, 296, 301-306, 320-321, and 446-447 in revision.

1   Characterizing severe weather potential in synoptically weakly forced

2   thunderstorm environments

3   By

4

5   Paul W. Miller[*]

6

7   and

8

9   Thomas L. Mote

10

11   Department of Geography
12   University of Georgia, Athens, Georgia

13

14   [*]Corresponding author address: Department of Geography, 210 Field Street, Room 204
15   Athens, GA 30602. Email: paul.miller@uga.edu

16

21

Abstract

Weakly forced thunderstorms (WFTs), short-lived convection forming in synoptically quiescent regimes, are a contemporary forecasting challenge. The convective environments that support severe WFTs are often similar to those that yield only nonsevere WFTs, and additionally, only a small proportion of individual WFTs will ultimately produce severe weather. The purpose of this study is to better characterize the relative severe weather potential in these settings as a function of the convective environment. Thirty-one near-storm convective parameters for >200,000 WFTs in the Southeast United States are calculated from a high-resolution numerical forecasting model, the Rapid Refresh (RAP). For each parameter, the relative ~~likelihood~~ odds of WFT days with at least one severe weather event is assessed along a moving threshold. Parameters (and the values of them) that reliably separate severe-weather-supporting from nonsevere WFT days are highlighted.

Only two convective parameters, vertical totals (VT) and total totals (TT), appreciably differentiate severe-wind-supporting and severe-hail-supporting days from nonsevere WFT days. When VTs exceeded values between 24.6–25.1°C or TTs between 46.5–47.3°C, odds of severe-wind days were roughly 5x ~~more likely~~greater. Meanwhile, odds of severe-hail days became roughly 10x ~~more likely~~greater when VTs exceeded 24.4–26.0°C or TTs exceeded 46.3–49.2°C. The stronger performance of VT and TT is partly attributed to the more accurate representation of these parameters in the numerical model. Under-reporting of severe weather and model error are posited to exacerbate the forecasting challenge by obscuring the subtle convective environmental differences enhancing storm severity.

1. Introduction

Weakly forced thunderstorms (WFTs), convection forming in synoptically benign, weakly sheared environments, are a dual forecasting challenge. Not only is the exact location and time of convective initiation difficult to predict, but once present, the successful differentiation of severe WFTs from their benign counterparts is equally demanding. Consequently, severe weather warnings issued on WFTs in the U.S. are less accurate than more organized storm modes, such as squall lines and supercells (Guillot et al., 2008). American operational meteorologists have coined these severe WFTs "pulse thunderstorms" because the surge of the updraft that produces the severe weather occurs in a brief "pulse" (Miller and Mote, 2017). The United States National Weather Service defines "severe weather" as any of the following: winds ≥ 26 m s^{-1}, hail $\geq$ 2~~0~~.5~~46~~ cm in diameter, or a tornado.

Environments thought to support pulse thunderstorms are typically characterized by weak vertical wind shear and strong convective available potential energy (CAPE). However, not all weak-shear, high-CAPE environments facilitate pulse thunderstorms, nor are all pulse thunderstorms confined to environments with the weakest shear and/or strongest instability. The result is a low signal-to-noise ratio (SNR) which obstructs the reliable discernment of pulse-supporting environments. The SNR is common discussion point in climate variability research where it often describes the relative magnitudes of a climate change trend (i.e., the signal) versus inter-annual variability (i.e., the noise) (e.g., Hamlington et al., 2010; Sutton and Hodson, 2007; Trenberth, 1984). In our~~this~~ context, the "signal" refers to the true difference between the large-scale convective environments that support severe weather and those that do not. Meanwhile, the "noise" ~~is~~ represent~~s~~ed the many processes than might cause storms to produce (not produce) severe weather in an environment where it was not expected (expected). Cell interactions,

4

69 stabilization from prior convection, surface convergence, locally enhanced shear, model error, etc,

70 can act as noise in the operational setting.

71   Prior research directed at pulse thunderstorms is limited, and work has not typically

72 included a representative proportion of nonsevere WFTs in their samples (Atkins and Wakimoto,

73 1991; Cerniglia and Snyder, 2002). If the sample contains too many pulse thunderstorms, the SNR

74 may be artificially bolstered, results overstated, and the potential reliability in an operational

75 setting diminished. For instance, in a meta-analysis of studies pertaining to new lightning-based

76 storm warning techniques, Murphy (2017) found that the studies' reported false alarms ratios

77 (FAR)s were directly proportional to the fraction of nonsevere storms contained in the sample.

78 Samples that included a realistic ratio of severe-to-nonsevere storms demonstrated the weakest

79 skill scores.

80   Most research considering pulse thunderstorms in the Southeast U.S. has typically focused

81 on one of its primary severe weather mechanisms: the wet microburst. Severe wet microbursts

82 generally occur in atmospheres characterized by a deep moist layer extending from the surface to

83 4–5 km above ground level (Johns and Doswell, 1992). Above the moist layer lies a mid-level dry

84 layer with lower equivalent potential temperature values (θ_e). In wet microburst environments, the

85 difference between the maximum θ_e observed just above the surface and the minimum θ_e aloft

86 exceeded 20 K, whereas non-microburst-producing thunderstorm days had differences less than

87 13 K (Atkins and Wakimoto, 1991; Roberts and Wilson, 1989; Stewart, 1991; Wheeler and Spratt,

88 1995). However, Atkins and Wakimoto (1991) examined only 14 microburst days versus three

89 non-microburst days. Adding to the uncertainty, James and Markowski (2010) challenged the role

90 of mid-level dry air in severe weather production. The results of their cloud-scale modeling

91 experiment indicated that, for all but the highest instabilities tested, drier mid-level air did not

92 correspond to increased downdraft and cold pool intensity.

93   Building on these findings, several severe weather forecasting parameters have been

94 developed to distill the atmosphere's vertical thermodynamic profile into a single value

95 representing the damaging wind potential. McCann (1994) developed a microburst-predicting

96 "wind index" (WINDEX) to be used in the forecasting of wet downburst potential. However,

97 although WINDEX performed well when tested in known microburst environments, no null cases

98 were presented (McCann, 1994). Additional severe wind potential indices include the wind

99 damage parameter and the microburst index described by the United States Storm Prediction

100 Center (SPC; http://www.spc.noaa.gov/exper/soundings/help/index.html). Tools such as total

101 totals, k-index, the Ssevere WEAweather and Tthreat (SWEAT) index, etc, are also commonly

102 used to forecast convective potential as well as the severity of thunderstorms.

103   However, the comparative utility of these environmental parameters within weakly forced

104 regimes is unclear, particularly when they are tested with a realistic proportion of severe storms.

105 Many of the results above were obtained by analyzing relatively small datasets, and they have not

106 been tested against each other in a weakly forced environment. Therefore, this study seeks to

107 compare the relative skill of convective parameters using a large WFT dataset to determine which

108 are most appropriate for detecting environments supportive of pulse-thunderstorm-related severe

109 weather.

110

111

112

113

2. Data and Methods

2.1 WFT selection and environmental characterization

This study uses the 15-yr WFT dataset developed by Miller and Mote (2017) for the Southeast U.S. (Fig. 1). Their ~~catalogue~~ detection method first identifies thunderstorms as regions of spatiotemporally contiguous composite reflectivities meeting or exceeding 40 dBZ using connected neighborhoods labeling. Each thunderstorm is then assigned five morphological attributes describing it shape, duration, intensity, etc, and all thunderstorms are clustered into ten morphologically similar groups using Ward's clustering (Ward, 1963). The composite convective environments associated with each morphological group were characterized using radiosonde observations from three launch sites in the Southeast U.S. WFTs were designated as ~~with WFTs representing~~ the subset of ~~generally~~ morphological groups with small, short-lived, diurnally driven thunderstorms that also formed in weak-shear, strong-instability composite environments. Table 1 provides the composite kinematic and thermodynamic environmental characteristics for the ten morphological groups from Miller and Mote (2017). The WFTs are spatially referenced according to their first-detection location, the centroid of the composite reflectivities constituting the first appearance on radar. The storms were then paired with severe weather reports from *Storm Data*, a storm event database maintained by the United States National Centers of Environmental Information, to differentiate benign WFTs from pulse thunderstorms. The entire 15-yr dataset contains 885,496 WFTs including 5316 pulse thunderstorms.

Meanwhile, the thermodynamic and kinematic environment of each WFT was characterized using the 0-hr Rapid Refresh (RAP) ~~(RAP; Benjamin et al., 2016)~~ analysis. The RAP, implemented on 9 May 2012, is a 13-km non-hydrostatic weather model initialized hourly for the purpose of near-term mesoscale forecasting which is operated by the United States National

Center for Environmental Prediction. The RAP uses the National Oceanic and Atmospheric Administration (NOAA) Gridpoint Statistical Interpolation (GSI) system to assimilate radar reflectivity, lightning flashes (added in version 3), radiosonde observations, GOES cloud analysis, wind profiler datas, surface station observations, etc. Lateral boundary conditions are provided by the Global Forecast System (GFS). Additional information regarding the RAP assimilation system and model physics can be found in ~~Lateral boundary conditions are provided by the Global Forecast System (GFS). Additional information regarding the RAP assimilation system and model physics can be found in~~ (Benjamin et al., 2016). The model has output available at 37 vertical levels spaced at 25-hPa intervals between 1000 and 100 hPa and 10-hPa intervals above 100 hPa. Several previous studies have relied upon the RAP's predecessor, the Rapid Update Cycle (RUC; Benjamin et al., 2004), to effectively characterize near-storm environments differentiating supercellular versus non-supercellular and tornadic versus non-tornadic thunderstorms (Thompson et al., 2007; Thompson et al., 2014).

For the grid cell containing each WFT's first-detection location, a RAP proxy sounding was created using the SHARPpy software package (Blumberg et al., 2017). Thus, each proxy sounding represents the model-derived storm environment for a point no more than 13 -km and 30 min distant from the WFT first-detection location. The proxy soundings were used to calculate ~~30~~31 near-storm environmental variables and indices, a complete list of which is provided in Table ~~1~~2 with more thorough descriptions in Appendix A. The ~~30~~31 variables were largely selected by virtue of their accessibility in SHARPpy. Four warm seasons of the Miller and Mote (2017) dataset, containing 228,363 WFTs and 1481 pulse thunderstorms, overlapped with the RAP's operational archive period allowing >6 million near-storm parameters to contribute to the analysis.

2.2 RAP error assessment

161   Thompson et al. (2003) demonstrated the suitability of the RUC, version 2 (RUC-2), to

162   represent storm environments as evaluated using co-located radiosonde observations, and the

163   Benjamin et al. (2016) RAP validation statistics show that the RAP is more accurate than its

164   predecessor. Figure 2a shows the results of an error evaluation specific to the purposes of this

165   study. Vertical error profiles were calculated for 3562 co-located RAP predictions and observed

166   radiosonde profiles in the Southeast U.S. The comparisons contain 0000 and 1200 UTC soundings

167   during the warm season (May–September) between 2012 and 2015 at three launch sites along a

168   north-south trajectory through the Miller and Mote (2017) domain: Nashville, Tenn., Peachtree

169   City, Ga., and Tampa, Fla., corresponding to U.S. radar identification codes KOHX, KFFC, and

170   KTBW in Fig. 1. The synoptic station codes for these three sites are the same as their U.S. radar

171   identifications with the exception of Nashville, Tenn., whose synoptic code is KBNA.

172   Similar to the Thompson RUC-2 analysis, the greatest, albeit small, temperature and

173   moisture biases (mean errors) from the RAP reside near the surface and the upper atmosphere (Fig.

174   2a). Aided by the large sample of comparison soundings, the 95% confidence intervals indicate

175   that the true bias of the selected RAP output variables at these sites can be estimated with

176   reasonable confidence. The 95% mixing ratio confidence interval captures zero at all altitudes

177   except 500 hPa, where the RAP predicted drier-than-observed values by 0.08 g kg^{-1}. Temperatures

178   are warmer than observed throughout most of the troposphere with a maximum bias of 0.26°C at

179   850 hPa. In contrast, the RAP underestimated wind speeds on average throughout the depth of the

180   troposphere. The largest bias, 0.46 m s^{-1}, was found at 925 hPa with similar errors above 500 hPa.

181   The 95% confidence interval for wind speed error is largest near the tropopause, and demonstrates

182   larger uncertainty than for temperature and mixing ratio. These results generally agree with the

183   error statistics provided by Benjamin et al. (2016), and the reader should reference that paper for

184   additional information, including validation statistics, about the RAP.

185   Although the RAP appears to resolve temperature, mixing ratios, and wind speeds more

186   accurately than the RUC-2, the transmission of these errors onto the derived convective parameters

187   can be large. ~~Table 2~~Table 3 expresses error measures for surface-based (SBCAPE) and mean-

188   layer CAPE (MLCAPE), 0–3-km and 0–6-km wind shear, total totals, and the theta-e index.

189   Because the focus of this study is surface-based convection, only days when the observed surface-

190   based CAPE was greater than zero were used to calculate the derived quantity error metrics.

191   Similar to previous work (e.g., Lee, 2002), parameters calculated via the vertical integration of a

192   parcel trajectory, such as CAPE, are sensitive to errors in low-level temperature and moisture. The

193   RAP's low-level temperature and moisture biases influence the lifted condensation level (LCL)

194   calculation (negative MLLCL bias; ~~Table 2~~Table 3) yielding a premature transition to the pseudo-

195   adiabatic lapse rate and an overestimate of parcel instability (positive SBCAPE and MLCAPE

196   biases; ~~Table 2~~Table 3)[1]. Thompson et al. (2003) identified smaller CAPE errors generated by the

197   RUC-2; however, the nature of the thermodynamic environments being examined is significantly

198   different in this study. Similar to the RUC-2, the RAP is more adept at representing MLCAPE than

199   SBCAPE with Fig. 2b, and consequently, the mean-layer parcel trajectory will be used for all

200   parcel-related calculations.

201   In some cases, RAP proxy soundings may have been contaminated by premature

202   convective overturning within the model. However, because the RAP assimilates radar reflectivity

203   from the U.S. (Benjamin et al., 2016), the 0-hr RAP analysis fields should generally mirror the

---

[1] The near-surface temperature and moisture errors in Fig. 2a are more pronounced following the upgrade to RAPv2 in February 2014. However, because the RAP is an operational tool and this work has operational relevance, no attempt was made to correct for this change.

204 radar-observed areas of convection. Additionally, any such instances will be dampened by the

205 methodological design decision to aggregate all proxy soundings on a daily level as will be

206 described in Sect. 2.3. The accuracy of the proxy soundings could be improved by employing a

207 convection-permitting numerical model, such as the 3-km High-Resolution Rapid Refresh

208 (HRRR). By explicitly modeling deep convection, the HRRR would limit convective

209 contamination by more closely representing areas of thunderstorm activity. At the time of

210 publication, the absence of a publicly accessible HRRR archive prevented its application in this

211 research.

212   Figures 2b-d demonstrate that although large outliers certainly occur, the majority of RAP-

213 derived thermodynamic and kinematic parameters are concentrated within a narrower range of

214 error. Figure 3 provides an example skewT-logP diagram for a large MLCAPE error shown in

215 Figure 2d. Though the difference in this case exceeded 1000 J kg^{-1}, the discrepancy can largely be

216 attributed to the RAP's minor mischaracterization of low-level moisture. Otherwise, the depiction

217 of the vertical profile is reasonably accurate. The advantage of the RAP to represent the near-storm

218 environment is underscored when compared to results from coarser-scale models. For instance,

219 the coefficients of determination (R^2) for RAP-derived SBCAPE and MLCAPE are appreciably

220 larger than those calculated from the 32-km horizontal and 3-hr temporal resolution North

221 American Regional Reanalysis (NARR; Mesinger et al., 2006) in Gensini et al. (2014).

222

223 *2.3 Assessing convective parameter skill*

224   The quality of severe weather reports is a significant impediment to severe storm research

225 (e.g., Miller et al., 2016; Weiss et al., 2002), particularly regarding the certainty with which

226 nonsevere storms can be declared nonsevere. These storms may only appear benign because their

227 associated severe weather was not reported. Consequently, the results of the proxy soundings are

228 subdivided by nearest radar site (Fig. 1) and aggregated daily (1200–1200 UTC) by computing the

229 mean parameter value associated with all WFTs forming within each polygon on a given day. ~~with~~

230 D~~d~~ays containing at least one severe weather report are considered supportive of severe weather

231 whereas days with no severe weather reports will serve as the control. This approach is similar to

232 the methods the Hurlbut and Cohen (2013) study of severe thunderstorm environments in the

233 Northeast U.S. Severe-wind-supporting (SWS) days and severe-hail-supporting (SHS) days are

234 treated separately because their thermodynamic environments have been shown to contain unique

235 elements related to downdraft and hailstone production (Johns and Doswell, 1992). ~~Table 3~~Table

236 4 provides the specific subdivision details of the frequency of WFT days, SWS days, SHS days,

237 and their respective control days. Figure 4 shows the annual average of WFT days for each radar

238 site within the study area during the 2012–2015 warm seasons. As expected, WFT days are most

239 frequent along coastlines and the Appalachian Mountains (Miller and Mote, 2017).

240   Given the low SNR in WFT environments, t-tests are deceiving. Statistically significant

241 differences in the mean values of parameters on severe versus nonsevere days are routinely

242 reported, but the considerable overlap between the distributions (e.g., Craven and Brooks, 2004;

243 Taszarek et al., 2017) can remove much practical value. This study explores the relationship

244 between convective parameters and pulse thunderstorm environments by means of an odds ratio

245 (OR; e.g., Fleiss et al., 2003). The OR is a common measure of conditional likelihood in human

246 health and risk literature (e.g., Bland and Altman, 2000) with precedenc~~t~~e in the atmospheric

247 sciences (e.g., Black and Mote, 2015; Black et al., 2017). The OR looks past the descriptive

248 statistics of the severe versus nonsevere distributions and more directly compares differences in

249 where the data are~~is~~ concentrated within the distributions.

250   Equation 1 shows the standard definition of the OR, essentially the ratio of two ratios,

$$OR = \frac{A/C}{B/D} \qquad (1)$$

251   where the numerator represents the ratio of events (A) to non-events (C) when a condition is met

252   whereas the denominator is the ratio of events (B) to non-events (D) when the same condition is

253   not satisfied. In this context, "events" are SWS or SHS days whereas "non-events" would be the

254   respective control days. Higher ORs indicate that events are more frequent (relative to non-events)

255   when the condition is met, or conversely, that events are less frequent when the condition is not

256   met. For this study, a condition might be a convective parameter exceeding a specified threshold.

257   For instance, if the SWS OR equals 4 for the condition MLCAPE > 1000 J kg^{-1}, then the odds of

258   an SWS day ~~is~~are 4x ~~more likely~~greater when MLCAPE is greater than 1000 J kg^{-1} than when it

259   is less than 1000 J kg^{-1}.

260   We employ a modified form of the OR in which both the numerator and denominator are

261   standardized by the climatological ratio of events to non-events (Eq. 2), allowing the components

262   of the OR to be separated and interpreted independently by comparison to climatology.

$$OR = \frac{\dfrac{A/C}{(A+B)/(C+D)}}{\dfrac{B/D}{(A+B)/(C+D)}} \qquad (2)$$

263   The modification does not change the value of the quotient OR, but it does improve the

264   interpretability of the numerator and denominator. When the numerator or denominator is near

265   zero (one), then the ~~likelihood~~odds of SWS or SHS days ~~is~~are much lower than (nearly equal to)

266   climatology. The climatological odds ratio was 0.069 for SWS days and 0.025 for SHS days. A

267   95% confidence interval for the OR was calculated using the four-step method presented in Black

268   et al. (2017).

269 **3. Results**

270 *3.1 Convective environments of pulse thunderstorm wind events*

271   During the four-year study period, pulse thunderstorm wind events were documented

272 somewhere in the study area on 49% of WFT days, although the average frequency within any

273 single subdivision was 6.7% (~~Table 3~~Table 4). ~~Table 4~~Table 5 shows the ~~30~~31 convective

274 parameters analyzed from the proxy soundings as well as the number of subdivisions for which

275 each parameter is a statistically significant differentiator of SWS days. A significance threshold of

276 $p < 0.10$ guided the selection of potentially useful parameters which would be examined in more

277 detail. Nine of the ~~30~~31 variables are statistically significant across at least two-thirds of the study

278 area: VT, TT, MLCAPE, MLLCL, MICROB, DCAPE, TEI, RH_LOW, and ThE_LOW.

279   Figure 5a-h depicts the distributions for several parameters from ~~Table 4~~Table 5 for control

280 versus SWS days. These eight parameters are either significant across much of the domain (VT

281 and TT), demonstrate larger relative changes on SWS days (MLCAPE and MLLCL), and/or are

282 traditional operational severe wind forecasting tools (DCAPE, TEI, WNDG, MICROB). However,

283 as the distributions clearly illustrate, any difference in the mean values between the control days

284 and SWS days is small compared to the spread about their means. This results in the

285 characteristically low SNR described in the Sect. 1. Any attempt to establish a forecasting value

286 indicative of pulse-wind potential will yield many missed events occurring beneath the threshold

287 and/or false alarms associated with control days above it.

288   Thus, Fig. 6 employs the OR to characterize the relative skill that some knowledge of the

289 convective environment can contribute to a severe versus nonsevere designation. For each variable

290 in Fig. 5, a progressively larger value is selected, and the OR is calculated at each step. Figure 6

291 displays the OR as well as both the numerator and denominator terms for each iteration. ~~Often~~

14

292 ~~Hh~~High ORs can often result when a near-zero number of severe events exist below the threshold

293 inflating the OR calculation. In these situations, the OR is indicating that severe weather is very

294 unlikely, rather than that the severe weather risk is enhanced. These results are not particularly

295 useful because forecasters would not have needed a decision-support tool in these environments

296 in the first place. Ideally, large ORs will result when the numerator indicates an appreciable

297 increase against the climatology while the denominator simultaneously indicates an appreciable

298 decrease below climatology. Further, these ORs would ideally occur in a range where the severe

299 weather risk may be uncertain. In Fig. 6, the OR is shown in a gray line, but the line is drawn in

300 black whenever the OR results from a numerator ≥ 2 and a denominator ≤ 0.5. ORs resulting from

301 this combination indicate that the threshold yields a simultaneous two-fold increase (decrease) in

302 the ~~likelihood~~ odds of SWS days above (below) the specified value. These ORs will be hereon

303 referenced as "two-fold" ORs, and represent a goal scenario.

304 　　Figures 6a-h show ORs for the same eight parameters in Fig. 5. Of all eight parameters,

305 only VT and TT achieve two-fold ORs for any range of thresholds, as indicated by the black

306 segments in Fig. 6a-b. The maximum two-fold OR for VT is 5.16 at 24.6°C, meaning that the odds

307 of an SWS day~~s~~ are 5.16x ~~more likely~~ greater when this threshold is met. TT offers slightly more

308 skill with a maximum two-fold OR of 5.70 at 46.5°C. As described in Appendix A, VT (vertical

309 totals) and TT (total totals) are relatively primitive indices. VT is purely a temperature lapse rate

310 whereas TT is predominantly a measure of lapse rate with an additional dewpoint term included.

311 Meanwhile, MLCAPE and MLLCL demonstrate consistently lower ORs between 2 and 4.

312 ~~Surprisingly, t~~The four wind-specific variables in Fig. 6e-h are relatively poor differentiators of

313 SWS days in the WFT regime. The maximum OR achieved by any of these parameters is

314 approximately 10 driven by very low values of DCAPE with corresponding wide confidence

315 intervals.

316 Though ORs are greater at lower VT and TT thresholds, these values are also somewhat

317 common. Placing the aforementioned values (24.6°C and 46.5°C, respectively) in the context of

318 the 12,759 WFT environments included in this study, they represent the 58.8[th] and 58.9[th]

319 percentiles of their distributions. Alternatively, the maximum VT threshold that yields a two-fold

320 OR is 25.1°C, which corresponds to the 70.9[th] percentile of all VTs in the dataset; however, the

321 OR for this value is smaller, 4.77. This result illustrates the trade-off involved by seeking

322 climatologically exceptional values to serve as guidance. As greater values are selected as the

323 threshold, meteorologists can focus on a fewer number of days. However, the OR decreases as

324 more severe weather events occur in environments not satisfying the threshold. As for TT, the

325 maximum two-fold OR value is 47.3°C, corresponding to the 70.6[th] percentile, but demonstrates

326 an OR of 5.16. This means that when TT meets or exceeds 47.3°C, the odds of a pulse thunderstorm

327 severe wind event~~s~~ are 5.16x ~~more likely~~greater than when it does not.

328

329 *3.2 Convective environments of pulse thunderstorm hail events*

330 ~~Table 5~~Table 6 replicates ~~Table 4~~Table 5 except for SHS days. Many of the same

331 parameters that are statistically significant differentiators of SWS days also rank high for SHS

332 days. However, fewer parameters in ~~Table 5~~Table 6 are statistically significant over two-thirds of

333 the domain. Whereas 10 parameters in ~~Table 4~~Table 5 showed spatially expansive statistical skill

334 on SWS days, only three quantities do so on SHS days. We attribute this result to the pattern in

335 ~~Table 3~~Table 4 and Fig. 4b-c whereby there are fewer SHS days than SWS days, which increases

336 uncertainty related to the statistical tests and makes it harder to confidently detect differences.

337   Nonetheless, VT and TT are once again skillful differentiators, and are now joined by their

338   related parameter CT. Additionally, several new convective variables demonstrate statistical

339   significance across roughly half of the domain on SHS days that demonstrated little skill on SWS

340   days: PW, PEFF, HGT0, and ApWBZ. For comparison, Fig. 7a-d duplicates Fig. 5a-d now

341   comparing distributions between the control and SHS days while Fig. 7e-h displays boxplots for

342   the SHS-specific convective parameters listed above. The distributions for MLCAPE and MLLCL

343   are similar; however, there is a larger separation between control and SHS days for VT and TT

344   than was apparent on SWS days. This observation is corroborated by the relative changes in VT

345   and TT on SHS days that are several percentage points larger than for SWS days (~~Table 5~~Table

346   6). PW, PEFF, HGT0, and ApWBZ demonstrate smaller differences.

347   Figure 8 replicates Fig. 6 except by representing SHS days and substituting the four wind-

348   specific parameters (DCAPE, TEI, WNDG, MICROB) with the four hail parameters listed above

349   (PW, PEFF, HGT0, ApWBZ). The ORs for VT and TT are large, greater than 10, throughout the

350   entire range of thresholds tested, and contain larger swathes of two-fold ORs. The maximum two-

351   fold OR for VT is 13.1 at 24.4°C, and the maximum ~~two-fold-OR-achieving~~ VT threshold that

352   achieves a two-fold OR is 26.0°C with an OR of 9.61. These values relate to the 53.4^{th} and 86.0^{th}

353   percentiles of the VT distribution. As for TT, the maximum two-fold OR is 14.98 at 46.3°C, and

354   the maximum two-fold-OR threshold is 49.2°C with an OR of 11.79. These two TT cut-offs

355   translate to the 55.7^{th} and 88.4^{th} percentiles. Similar to SWS days, MLCAPE and MLLCL show

356   little skill with ORs generally between 1–2. PW, PEFF, HGT0, and ApWBZ perform more capably

357   than MLCAPE and MLLCL; however, they do not produce any two-fold ORs. Values for these

358   metrics are generally around 4 with several instances of higher ORs driven by a small denominator

359   with wide 95% confidence intervals.

360 *3.3 Separating marginal pulse thunderstorm days*

361   Because the severe weather generated by pulse thunderstorms is often near the lower limit

362 used to define severe weather in the United States, some pulse thunderstorm environments may

363 closely resemble nonsevere regimes. Consequently, the influence of these "marginal" pulse

364 thunderstorm days on the OR analysis is further scrutinized. For this purpose, "marginal" SWS

365 and SHS days are defined as those on which only one severe wind or hail report was received.

366 Marginal days constitute 48.7% of the SWS days and 57.7% of the SHS days in ~~Table 3~~Table 4.

367 Figure 9 replicates the OR analysis for VT and TT, the two most promising environmental

368 parameters from Sects. 3.1 and 3.2, but with only marginal SWS and SHS days being considered.

369 Comparing Figs. 6a-b and 8a-b to Fig. 9, marginal SWS and SHS days resemble the OR patterns

370 of the broader set of SWS (Fig. 6a-b) and SHS (Fig. 8a-b) days. Though the ORs for the marginal

371 subset are slightly smaller than for the broader group, they bear similar OR patterns as the

372 thresholds are increased. Overall, marginal SWS and SHS days are generally characterized by

373 similar VT and TT values as when all SWS and SHS days were aggregated. Corroborating this

374 finding, an OR analysis comparing marginal SWS and SHS days to those with >1 severe event

375 (not shown) revealed that ORs generally remained near 1 regardless of the VT or TT threshold

376 selected. Thus, although marginal pulse thunderstorm days are by no means easily distinguishable

377 from non-severe WFT days, they do not appear to be particularly more challenging to differentiate

378 than active pulse thunderstorm days.

379

380 **4. Discussion**

381   The relative changes in the convective variables in ~~Table 4~~Table 5 on SWS days versus

382 control days correspond well to previous microburst research. Compared to the nonsevere control

383   days, SWS days are characterized by a drier near surface layer (i.e., lower RH, higher LCLs).

384   Simultaneously, steep mid-level lapse rates (i.e., larger VT and TT) aid an increase in CAPE which

385   supports stronger updrafts. As the strong updraft transitions to a downdraft-dominant storm, the

386   drier surface layer supports evaporative cooling, downdraft acceleration, and severe outflow

387   winds. This same conceptual model has been promoted by previous severe convective wind

388   research (e.g., Atkins and Wakimoto, 1991; Kingsmill and Wakimoto, 1991; Wolfson, 1988).

389   The results of SHS days also support previous findings (Johns and Doswell, 1992; Moore

390   and Pino, 1990; Púčik et al., 2015). The distributions in Fig. 7 (and relative changes in ~~Table~~

391   ~~5~~Table 6) indicate that SHS days are characterized by relative decreases in PW, a lower freezing

392   level, a lower wet-bulb freezing level, and dry near-surface air. Smaller PWs result in less

393   waterloading and greater parcel buoyancy (larger VT, TT, and MLCAPE) which maximizes

394   updraft strength. Meanwhile, lower freezing levels and a dry layer between 1000–850 hPa support

395   evaporative cooling which can together yield a lower wet-bulb zero height~~,~~ and ~~more efficient~~

396   ~~growth of hailstones~~limit hail stone melting during its descent to the surface. Interestingly, these

397   two concepts are both represented in the PEFF calculation (Appendix A) which was not developed

398   as a hail indicator. PEFF as defined by Noel and Dobur (2002), equals the product of PW and the

399   mean 1000–700-hPa RH. As both values decrease, PEFF becomes smaller and hail is more likely

400   for the reasons stated above.

401   The poor performance of MLLCLs and MLCAPEs in differentiating SWS and SHS days

402   from their controls is surprising given their prominence in severe storm forecasting. One

403   possibility is that the daily aggregation of MLCAPEs may have smoothed out locally higher values

404   near the WFTs that were responsible for severe weather production. ~~In contrast~~Alternatively, VT

405   and TT were among the strongest indicators of both SWS and SHS days. Recalling from Sect. 2.2,

406   VT and TT are also very well represented by the RAP. TTs were replicated by the model with a

407   <1°C bias and a MAE representing only 3% of the average value (~~Table 2~~Table 3). Additionally,

408   mid-level temperatures, from which VT is computed, also compared very well to the observed

409   soundings (Fig. 2a). ~~Thus, t~~The strong performance of VT and TT compared to other more heavily

410   moisture-weighted metrics may be due to their more accurate representation in the proxy

411   soundings.

412   Regardless, because the severe weather SNR is already low in WFT environments, any

413   systematic error introduced by the data source (in this case the RAP) may significantly dampen,

414   or even remove, whatever environmental differences exist. As Sect. 2.2 indicated and previous

415   work has also concluded, low-level moisture biases can impede the accurate calculation of

416   convective parameters relying on those terms (e.g., Gensini et al., 2014; Thompson et al., 2003).

417   In this study, MLCAPE, MLLCL, PW, PEFF, and others were vulnerable to such errors. The

418   poorer performance of these variables' ORs (relative to the lapse-rate-based parameters) and the

419   sensitivity of PW, PEFF, and ApWBZ to simulated RAP errors suggests that model inaccuracies

420   may be obscuring their potential skill to detect weakly forced severe weather environments. The

421   perception of the WFT environment as a difficult-to-forecast regime may partly be driven by model

422   inconsistency exacerbating an already small SNR.

423   Another confounding factor is the quality of the *Storm Data* severe weather reports. Section

424   3.3 discussed that marginal SWS and SHS days are more similar to days with >1 report than days

425   with no reports. Thus, the basis for the similarity may be that severe weather was simply under-

426   reported on "marginal" days. Extending this logic, the pulse regime's low SNR may also be

427   partially attributed to under-reporting of severe weather on "nonsevere" days. Given that the severe

428   weather generated by pulse convection is often short-lived, isolated, and narrowly exceeds severe

429 criteria, the notion that some pulse-related severe weather events go undetected is likely. If some

430 "nonsevere" days existing above the tested parameter thresholds in Figs. 6 and 8 did in fact host

431 severe weather, then the ORs would have been larger than those found in Sects. 3.1 and 3.2.

432

433 **5. Conclusions**

434 Hazardous weather within WFT environments is characterized by a lower SNR than other

435 severe thunderstorm regimes. Though past research has developed promising tools for forecasting

436 pulse thunderstorm environments, their relatively small samples sizes may have understated the

437 SNR, and by corollary overstated the reliability of their tools. With recent research suggesting that

438 the performance of new severe weather forecasting tools is closely tied to the proportion of

439 nonsevere thunderstorms in the sample (Murphy, 2017), this study sought to test the relative skill

440 of ~~30~~31 convective forecasting parameters using realistic proportions of severe and nonsevere

441 WFT environments (severe: 7.9%; nonsevere: 92.1%). Future research may consider broadening

442 the methods of Murphy (2017) to standardize the skill values across previous studies of severe

443 convective environments.

444 Only 13 (5) of the ~~30~~31 convective parameters tested were statistically significant (p <

445 0.10) differentiators of SWS (SHS) days across at least half of the domain. Though the distinctive

446 variables for SWS and SHS days were consistent with previous theories of severe microburst and

447 hail formation, considerable overlap between the distribution of values on severe and nonsevere

448 days is problematic. Similarities between the SWS, SHS, and their corresponding control

449 distributions inhibit consistent identification of pulse thunderstorm potential based on the value of

450 any individual parameter. Nonetheless, VT and TT did perform more skillfully than the others.

451 When VTs exceed values between 24.6–25.1°C or TTs between 46.5–47.3°C, the relative

452 ~~likelihood~~ odds of a wind event increases roughly 5x. Meanwhile, the odds of a hail event~~s~~ become

453 roughly 10x ~~more likely~~greater when VTs exceed values between 24.4–26.0°C or TTs between

454 46.3–49.2°C.

455   The noteworthy performance of VT and TT, two quantities calculated from the more

456 reliable RAP output fields, is unlikely a coincidence. Our findings suggest that the already weak

457 severe weather SNR in WFT environments is exacerbated by model limitations in the low-level

458 moisture and temperature fields. Meteorologists may perhaps alleviate the challenges of the WFT

459 environment by examining convective parameters that are well-represented by models, such as

460 VT, TT, and other measures of lapse rate. Future research might seek to track the transmission of

461 the model errors through calculation of forecast skill statistics, and more concretely ascertain the

462 contribution of model error to the SNR.

463

467   **7.   References**

468   Atkins, N. T. and Wakimoto, R. M.: Wet microburst activity over the southeastern United States:

469   Implications for forecasting, Wea. Forecasting, 6, 470–482, doi:10.1175/1520-

470   0434(1991)006<0470:WMAOTS>2.0.CO;2, 1991.

471   Benjamin, S. G., Dévényi, D., Weygandt, S. S., Brundage, K. J., Brown, J. M., Grell, G. A., Kim,

472   D., Schwartz, B. E., Smirnova, T. G., Smith, T. L., and Manikin, G. S.: An hourly

473   assimilation–forecast cycle: The RUC, Mon. Wea. Rev., 132, 495–518, doi:10.1175/1520-

474   0493(2004)132<0495:AHACTR>2.0.CO;2, 2004.

475   Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., Olson,

476   J. B., James, E. P., Dowell, D. C., Grell, G. A., Lin, H., Peckham, S. E., Smith, T. L.,

477   Moninger, W. R., Kenyon, J. S., and Manikin, G. S.: A North American hourly assimilation

478   and model forecast cycle: The Rapid Refresh, Mon. Wea. Rev., 144, 1669–1694,

479   doi:10.1175/MWR-D-15-0242.1, 2016.

480   Black, A. W. and Mote, T. L.: Characteristics of Winter-Precipitation-Related Transportation

481   Fatalities in the United States, Wea. Climate Soc., 7, 133–145, doi:10.1175/WCAS-D-14-

482   00011.1, 2015.

483   Black, A. W., Villarini, G., and Mote, T. L.: Effects of Rainfall on Vehicle Crashes in Six U.S.

484   States, Wea. Climate Soc., 9, 53–70, doi:10.1175/WCAS-D-16-0035.1, 2017.

485   Bland, J. M. and Altman, D. G.: The odds ratio, BMJ, 320, 1468, 2000.

486   Blumberg, W. G., Halbert, K. T., Supinie, T. A., Marsh, P. T., Thompson, R. L., and Hart, J. A.:

487   SHARPpy: An open source sounding analysis toolkit for the atmospheric sciences, Bull.

488   Amer. Meteor. Soc., 98, 1625–1636, doi:10.1175/BAMS-D-15-00309.1, 2017.

23

489   Cerniglia, C. S. and Snyder, W. R.: Development of warning criteria for severe pulse

490   thunderstorms in the Northeastern United States using the WSR-88D. National Weather

491   Service, Albany, NY, 2002.

492   Craven, J. P. and Brooks, H. E.: Baseline climatology of sounding derived parameters associated

493   with deep, moist convection, Natl. Wea. Dig, 28, 13-24, 2004.

494   Fleiss, J. L., Levin, B., and Paik, M. C.: Statistical Methods for Rates and Proportions, Wiley,

495   2003.

496   Gensini, V. A., Mote, T. L., and Brooks, H. E.: Severe-thunderstorm reanalysis environments and

497   collocated radiosonde observations, J. Appl. Meteor. Climatol., 53, 742–751,

498   doi:10.1175/JAMC-D-13-0263.1, 2014.

499   Guillot, E. M., Smith, T. M., Lakshmanan, V., Elmore, K. L., Burgess, D. W., and Stumpf, G. J.:

500   Tornado and severe thunderstorm warning forecast skill and its relationship to storm type,

501   New Orleans, LA, 20–24 January 2008 2008.

502   Hamlington, B. D., Leben, R. R., Nerem, R. S., and Kim, K. Y.: The effect of signal-to-noise ratio

503   on the study of sea level trends, J. Climate, 24, 1396-1408, doi:10.1175/2010JCLI3531.1,

504   2010.

505   Hurlbut, M. M. and Cohen, A. E.: Environments of Northeast U.S. severe thunderstorm events

506   from 1999 to 2009, Wea. Forecasting, 29, 3-22, doi:10.1175/WAF-D-12-00042.1, 2013.

507   James, R. P. and Markowski, P. M.: A numerical investigation of the effects of dry air aloft on

508   deep convection, Mon. Wea. Rev., 138, 140–161, doi:10.1175/2009MWR3018.1, 2010.

509   Johns, R. H. and Doswell, C. A. I.: Severe local storms forecasting, Wea. Forecasting, 7, 588–612,

510   doi:10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2, 1992.

511   Kingsmill, D. E. and Wakimoto, R. M.: Kinematic, dynamic, and thermodynamic analysis of a

512   weakly sheared severe thunderstorm over northern Alabama, Mon. Wea. Rev., 119, 262–

513   297, doi:10.1175/1520-0493(1991)119<0262:KDATAO>2.0.CO;2, 1991.

514   Lee, J. W.: Tornado proximity soundings from the NCEP/NCAR reanalysis data, M.S., Dept. of

515   Meteorology, University of Oklahoma, 61 pp., 2002.

516   McCann, D. W.: WINDEX - a new index for forecasting microburst potential, Wea. Forecasting,

517   9, 532–541, doi:10.1175/1520-0434(1994)009<0532:WNIFFM>2.0.CO;2, 1994.

518   Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D.,

519   Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li,

520   H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American regional reanalysis,

521   Bull. Amer. Meteor. Soc., 87, 343–360, doi:10.1175/bams-87-3-343, 2006.

522   Miller, P. and Mote, T.: A climatology of weakly forced and pulse thunderstorms in the Southeast

523   United States, J. Appl. Meteor. Climatol., 56, 3017–3033,

524   doi:https://doi.org/10.1175/JAMC-D-17-0005.1, 2017.

525   Miller, P. W., Black, A. W., Williams, C. A., and Knox, J. A.: Quantitative assessment of human

526   wind speed overestimation, J. Appl. Meteor. Climatol., 55, 1009–1020,

527   doi:10.1175/JAMC-D-15-0259.1, 2016.

528   Moore, J. T. and Pino, J. P.: An interactive method for estimating maximum hailstone size from

529   forecast soundings, Wea. Forecasting, 5, 508–525, doi:10.1175/1520-

530   0434(1990)005<0508:AIMFEM>2.0.CO;2, 1990.

531   Murphy, M.: Preliminary results from the inclusion of lightning type and polarity in the

532   identification of severe storms, 8th Conf. on the Meteor. Appl. of Lightning Data, Seattle,

533   WA, 1–12, 2017.

534 Noel, J. and Dobur, J. C.: A pilot study examining model-derived precipitation efficiency for use

535   in precipitation forecasting in the eastern United States, Natl. Wea. Dig., 26, 3–8, 2002.

536 Púčik, T., Groenemeijer, P., Rýva, D., and Kolář, M.: Proximity soundings of severe and nonsevere

537   thunderstorms in Central Europe, Mon. Wea. Rev., 143, 4805–4821, doi:10.1175/MWR-

538   D-15-0104.1, 2015.

539 Roberts, R. D. and Wilson, J. W.: A proposed microburst nowcasting procedure using single-

540   Doppler radar, J. Appl. Meteor., 28, 285–303, doi:10.1175/1520-

541   0450(1989)028<0285:APMNPU>2.0.CO;2, 1989.

542 Stewart, S. R.: The prediction of pulse-type thunderstorm gusts using vertically integrated liquid

543   water content (VIL) and the cloud top penetrative downdraft mechanism. NOAA, 1991.

544 Sutton, R. T. and Hodson, D. L. R.: Climate response to basin-scale warming and cooling of the

545   North Atlantic Ocean, J. Climate, 20, 891-907, doi:10.1175/JCLI4038.1, 2007.

546 Taszarek, M., Brooks, H. E., and Czernecki, B.: Sounding-derived parameters associated with

547   convective hazards in Europe, Mon. Wea. Rev., 145, 1511-1528, doi:10.1175/MWR-D-

548   16-0384.1, 2017.

549 Thompson, R. L., Edwards, R., Hart, J. A., Elmore, K. L., and Markowski, P.: Close proximity

550   soundings within supercell environments obtained from the Rapid Update Cycle, Wea.

551   Forecasting, 18, 1243–1261, doi:10.1175/1520-

552   0434(2003)018<1243:CPSWSE>2.0.CO;2, 2003.

553 Thompson, R. L., Mead, C. M., and Edwards, R.: Effective storm-relative helicity and bulk shear

554   in supercell thunderstorm environments, Wea. Forecasting, 22, 102–115,

555   doi:10.1175/WAF969.1, 2007.

556 Thompson, R. L., Smith, B. T., Dean, A. R., and Marsh, P. T.: Spatial Distributions of Tornadic

557   Near-Storm Environments by Convective Mode, Electronic J. Severe Storms Meteor., 8,

558   1–22, 2014.

559 Trenberth, K. E.: Signal versus noise in the Southern Oscillation, Mon. Wea. Rev., 112, 326-332,

560   doi:10.1175/1520-0493(1984)112<0326:SVNITS>2.0.CO;2, 1984.

561 Ward, J. H.: Hierarchical grouping to optimize an objective function, Journal of the American

562   Statistical Association, 58, 236–244, doi:10.1080/01621459.1963.10500845, 1963.

563 Weiss, S. J., Hart, J., and Janish, P.: An examination of severe thunderstorm wind report

564   climatology: 1970–1999, 21st Conf. Severe Local Storms, San Antonio, TX, 446–449,

565   2002.

566 Wheeler, M. and Spratt, S. M.: Forecasting the potential for central Florida microbursts. NOAA,

567   1995.

568 Wolfson, M.: Characteristics of microbursts in the continental United States, Lincoln Lab Journal,

569   1, 49–74, 1988.

570

571 **Appendix A.**

572 **Table A1.** Additional detail describing the convective parameters in Table 2~~1~~.

Parameter	Comments	Parameter	Comments
MLCAPE		DCAPE	Downdraft CAPE with respect to parcel with the minimum 100 hPa layer-averaged theta-e found in the lowest 400 hPa of the sounding.
MLCIN	Mean-layer parcel mixed over the lowest 100 hPa		
MLLCL			
MLLFC			
MLEL			
NCAPE	MLCAPE / MLEL	WNDG	(MLCAPE)/2000*(0–3-km lapse rate)/9*(1–3.5-km mean wind)/15*[(MLCIN + 50)/40)]. Values larger than 1 indicate an increased risk for strong outflow gusts.
K_IND	$T_{850} - T_{500} + T_{d850} - (T_{700} - T_{d700})$		
TT	CT + VT		
CT	$T_{d850} - T_{500}$		
VT	$T_{850} - T_{500}$		
PW	Depth of liquid water if all water vapor were condensed from the sounding	TEI	Difference between the surface theta-e and the minimum theta-e value in the lowest 400 hPa AGL
HGT0	Pressure level of the 0°C isotherm	MICROB	Weighted sum of the following individual parameters: surface theta-e, SBCAPE, surface-based lifted index, 0–3-km lapse rate, VT, DCAPE, TEI, and PW. Values exceeding 9 indicate that microbursts are likely.
ApWBZ	Height above ground level of the RAP pressure level with the wet bulb temperature nearest to 0°C		
W_LOW	Mean mixing ratio between 1000–850 hPa		
W_MID	Mean mixing ratio between 850–500 hPa		
RH_LOW	Mean RH between 1000–850 hPa	SWEAT	$12(T_{d850}) + 20(TT - 49) + 2(U_{850}) + (U_{500}) + 125[\sin(U_{dir500} - U_{dir850}) + 0.2]$
RH_MID	Mean RH between 850–500 hPa		
ThE_LOW	Mean theta-e from 1000–850 hPa		
ThE_MID	Mean theta-e from 850–500 hPa		
ML_BRN	Bulk Richard Number of the mean-layer parcel	0-3-km_SHR	Magnitude of vector shear between surface and 3 km AGL
Tc	Temperature of parcel lowered dry adiabatically from the convective condensation level	0-6-km_SHR	Magnitude of vector shear between surface and 6 km AGL
		0-8-km_SHR	Magnitude of vector shear between surface and 8 km AGL
PEFF	As defined by Noel and Dobur (2002). PEFF equals the product of PW and the mean 1000–700-hPa RH.	EBWD	Magnitude of vector shear between effective inflow base and one half of the MU equilibrium level height

573

574
575
576

577   **Tables**

578   **Table 1.** Kinematic and thermodynamic parameters of 1200-UTC composite soundings from

579   Atlanta, Ga., USA, for each radar-identified morphological type in Miller and Mote (2017).

580   Morphological types classified as WFTs are bolded. All kinematic values are shown in m s^{-1}

581   whereas the units of the thermodynamic parameters are provided in the table. Explanations for the

582   variable abbreviations can be found in Table 1 and Appendix A.

Type	0–6-km SHR	0–8-km SHR	0–12-km Max Wind	0–12-km Mean Wind	ThE LOW (K)	MLCAPE (J kg^{-1})	Forecast SBCAPE (J kg^{-1})
1	**4.3**	**5.1**	**7.2**	**3.0**	**343.0**	**562**	**1585**
2	**4.3**	**5.1**	**8.8**	**3.6**	**341.9**	**365**	**1214**
3	**4.7**	**6.2**	**9.4**	**3.7**	**340.5**	**289**	**1176**
4	**4.3**	**5.7**	**8.3**	**3.7**	**341.1**	**357**	**1121**
5	**3.2**	**5.1**	**9.6**	**3.2**	**341.7**	**283**	**1006**
6	6	7.7	11.6	5.0	339.0	211	973
7	8.2	10.8	16.5	6.1	336.6	66	723
8	4.9	7.7	13.6	3.1	336.0	24	558
9	5.4	8.7	15.4	3.0	330.6	0	32
10	7.9	9.8	13.5	5.8	334.5	0	391

583

584   **Table 21.** List of the 3031 convective parameters computed from the proxy soundings where

585   CAPE, CIN, LCL, LFC, and EL and correspond to convective available potential energy,

586   convective inhibition, lifted condensation level, level of free convection, and equilibrium level,

587   respectively.

Abbrev.	Full Name	Units
MLCAPE	Mean-layer CAPE	J kg^{-1}
MLCIN	Mean-layer CIN	J kg^{-1}
MLLCL	Mean-layer LCL	m
MLLFC	Mean-layer LFC	m
MLEL	Mean-layer EL	m
NCAPE	Normalized MLCAPE	m s^{-2}
K_IND	K index	°C
TT	Total totals	°C
CT	Cross totals	°C
VT	Vertical Totals	°C
PW	Precipitable Water	mm
HGT0	Height of 0°C temperature isotherm	hPa
ApWBZ	Approximate height of 0°C wet bulb temperature	m
W_LOW	Mean low-level mixing ratio	g kg^{-1}
W_MID	Mean mid-level mixing ratio	g kg^{-1}
RH_LOW	Mean low-level relative humidity	--
RH_MID	Mean mid-level relative humidity	--
ThE_LOW	Mean low-level theta-e	K
ThE_MID	Mean mid-level theta-e	K
ML_BRN	Mean layer bulk Richardson number	--
Tc	Convective temperature	°C
PEFF	Precipitation efficiency	--
DCAPE	Downdraft CAPE	J kg^{-1}
WNDG	Wind damage parameter	--
TEI	Theta-e index	°C
MICROB	Microburst composite index	--
SWEAT	Severe weather and threat index	--
0-3-km_SHR	0–3-km vertical wind shear	m s^{-1}
0-6-km_SHR	0–6-km vertical wind shear	m s^{-1}
0-8-km_SHR	0–8-km vertical wind shear	m s^{-1}
EBWD	Effective layer vertical wind shear	m s^{-1}

588

589 **Table 32.** RAP error statistics for surface-based CAPE (SBCAPE) and several of the variables

590 listed in Table 21. The statistics are presented similarly to Thompson et al. (2003) by providing

591 the mean RAP-derived value, the mean arithmetic error (bias), and the mean absolute error (MAE).

Parameter	Mean	Bias	MAE	R^2
SBCAPE	1354.3	141.3	530.4	0.59
MLCAPE	943.4	112.6	338.0	0.64
MLLCL	1077.4	-32.9	151.8	0.82
Total Totals	44.8	0.51	1.54	0.74
TEI	21.1	-2.30	3.80	0.69
0–3-km Shear	6.33	-0.48	1.38	0.82
0–6-km Shear	8.39	-0.28	1.40	0.88

592

593   **Table 43.** WFT, SWS, and SHS day frequency by radar site.

Site	WFT Days	Wind Control	SWS Days	% SWS	Hail Control	SHS Days	% SHS
KAKQ	376	351	25	6.6	363	13	3.5
KAMX	581	569	12	2.1	575	6	1.0
KBMX	376	364	12	3.2	372	4	1.1
KCAE	401	339	62	15.5	377	24	6.0
KCLX	450	407	43	9.6	440	10	2.2
KDGX	426	403	23	5.4	416	10	2.3
KEOX	384	366	18	4.7	382	2	0.5
KEVX	467	449	18	3.9	463	4	0.9
KFCX	408	318	90	22.1	370	38	9.3
KFFC	400	358	42	10.5	387	13	3.3
KGSP	417	334	83	19.9	383	34	8.2
KGWX	362	349	13	3.6	354	8	2.2
KHPX	299	282	17	5.7	294	5	1.7
KHTX	373	343	30	8.0	369	4	1.1
KJAX	555	520	35	6.3	546	9	1.6
KJGX	384	356	28	7.3	377	7	1.8
KLIX	504	492	12	2.4	501	3	0.6
KLTX	452	439	13	2.9	444	8	1.8
KMHX	497	496	1	0.2	495	2	0.4
KMLB	540	532	8	1.5	532	8	1.5
KMOB	451	444	7	1.6	446	5	1.1
KMRX	415	349	66	15.9	384	31	7.5
KMXX	357	346	8	2.2	350	4	1.1
KNQA	356	336	20	5.6	345	11	3.1
KOHX	349	336	13	3.7	345	4	1.1
KPAH	330	305	25	7.6	318	12	3.6
KRAX	367	337	30	8.2	355	12	3.3
KTBW	546	525	21	3.8	535	11	2.0
KTLH	482	461	21	4.4	479	3	0.6
KVAX	457	430	27	5.9	452	5	1.1
Mean	425	398	27	6.7	415	10	2.5

594

595   **Table 54.** Summary of convective parameters on SWS days. The "Sites" column indicates the

596   number of spatial subdivisions within which the difference between the SWS mean and the control

597   mean was accompanied by p < 0.10; the "percent change" column shows the relative increase or

598   decrease of the mean on SWS days.

Parameter	Sites	Percent change
VT	28	5.1
TT	27	4.2
MLCAPE	25	31.2
MICROB	23	44.0
DCAPE	22	17.3
TEI	22	13.1
MLLCL	21	12.9
ThE_LOW	21	0.9
RH_LOW	20	-5.5
WNDG	19	41.2
CT	19	3.2
Tc	19	5.8
MLEL	18	8.0
SWEAT	14	7.8
W_LOW	10	3.0
K_IND	8	3.8
RH_MID	7	-3.2
ThE_MID	6	0.1
PEFF	6	-3.8
0-6-km_SHR	6	-4.5
0-8-km_SHR	6	-6.5
ApWBZ	5	-0.5
HGT0	4	0.1
W_MID	3	0.0
MLBRN	3	-0.7
NCAPE	2	23.9
PW	2	0.9
0-3-km_SHR	2	-1.2
MLCIN	0	6.6
MLLFC	0	0.9
EBWD	0	-1.9

599

600 **Table 65.** Same as Table 54, except for SHS days.

Parameter	Sites	Percent change
VT	27	8.0
TT	27	7.5
CT	21	7.1
PEFF	16	-11.0
MLLCL	15	13.2
HGT0	14	2.4
ApWBZ	14	-6.0
RH_LOW	14	-5.3
DCAPE	13	23.3
MLCAPE	12	28.8
PW	12	-6.7
W_MID	11	-9.2
ThE_MID	10	-0.7
WNDG	10	27.4
RH_MID	9	-7.8
TEI	7	10.4
MICROB	7	21.6
SWEAT	7	10.1
W_LOW	6	-2.1
Tc	6	3.2
0-6-km_SHR	6	9.7
0-8-km_SHR	5	6.9
MLEL	4	3.8
K_IND	3	2.7
ThE_LOW	3	-0.1
0-3-km_SHR	3	5.3
NCAPE	1	27.0
MLCIN	1	17.7
MLLFC	1	4.1
MLBRN	1	-15.8
EBWD	1	9.5

601

602   **Figures**

603
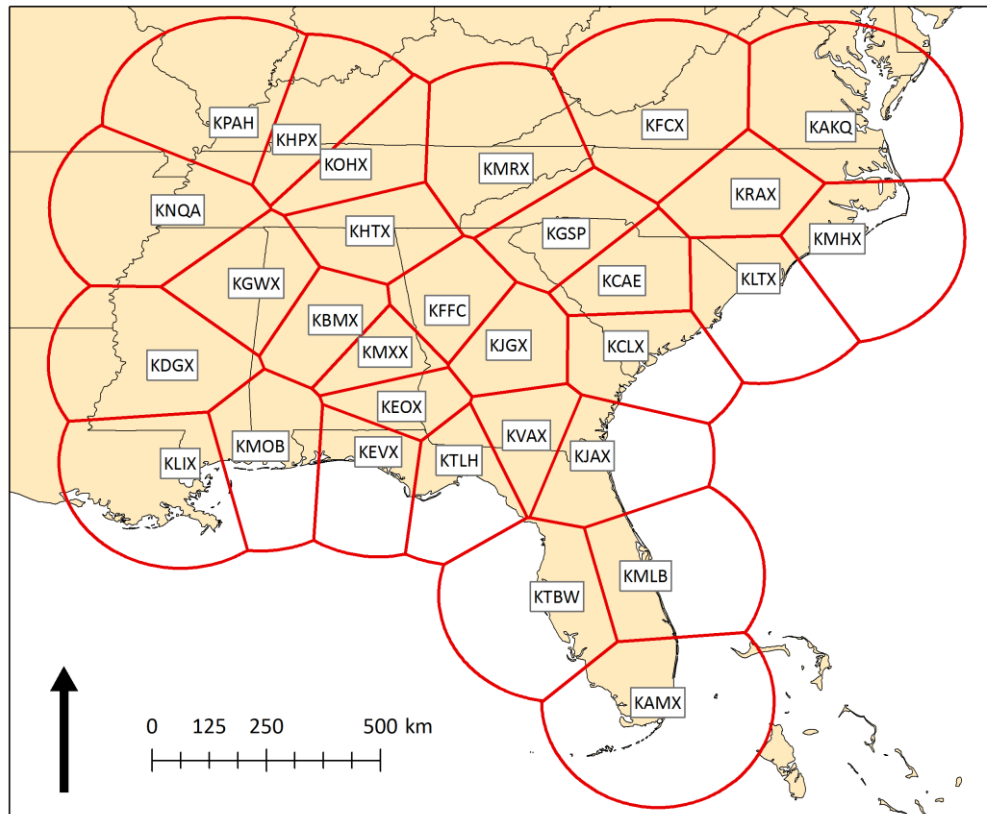
604   **Figure 1**. WSR-88D sites contributing to the Miller and Mote (2017) WFT climatology.
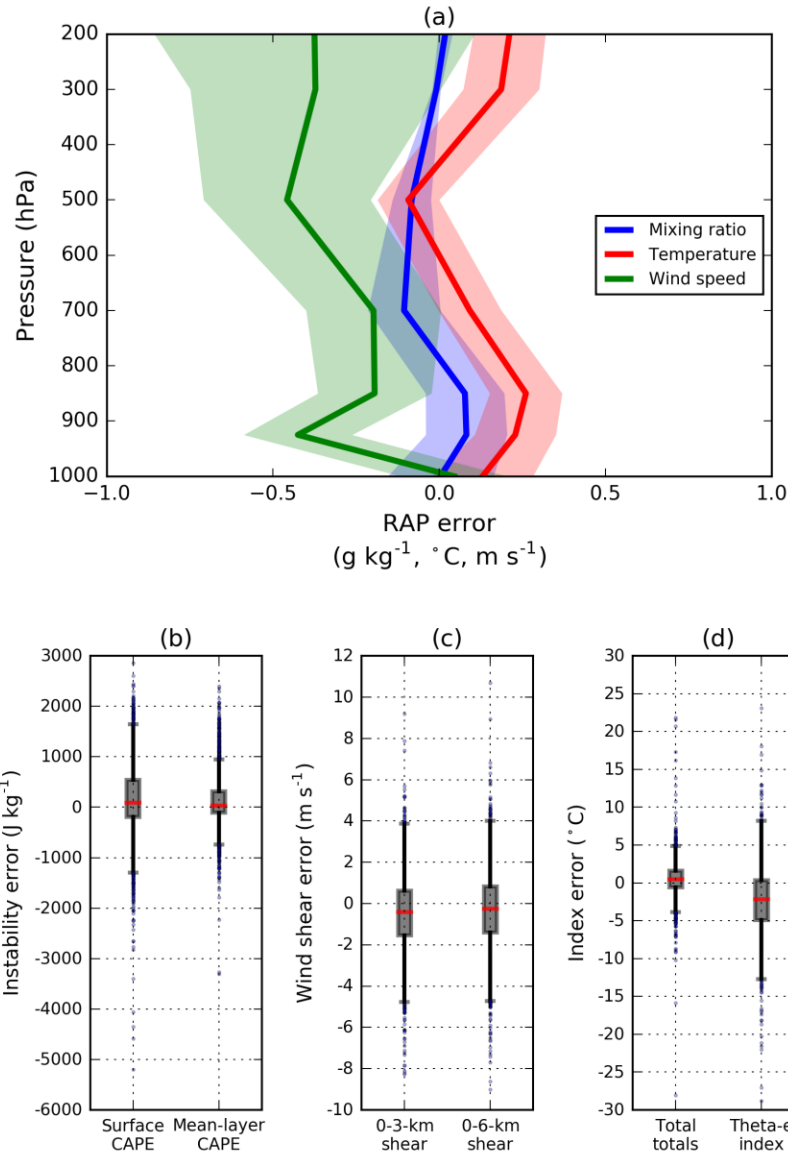
Figure 2. Vertical profiles of RAP output errors measured by co-located radiosonde observations (a). Errors were calculated at 1000, 925, 850, 700, 500, 300, and 200 hPa. The 95% confidence interval for the mean error (solid lines) is shaded. Boxplots of the resulting error for six derived quantities is shown in (b)-(d). The interquartile range (IQR), representing the middle 50% of values, is depicted by the gray box. Values lying more than 1.5*IQR from the median (red line) are marked with dots.
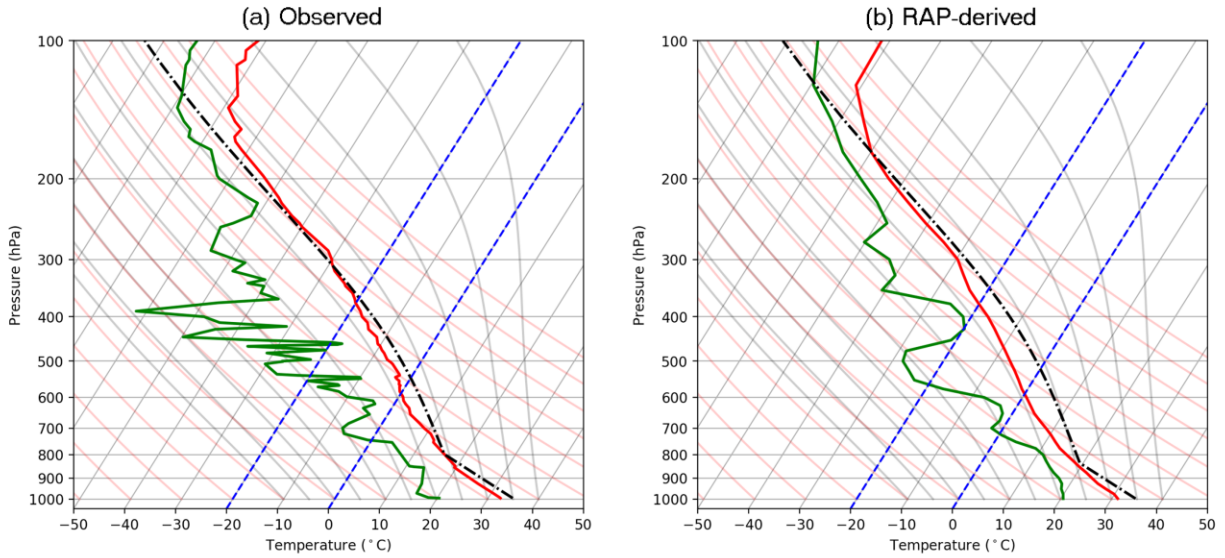
Nashville, Tenn., USA
0000 UTC, 18 June 2015

612

Figure 3. Comparison of observed (a) versus RAP-derived (b) soundings for a case when the MLCAPE discrepancy exceeded 1000 J kg^{-1} (observed: 1028 J kg^{-1}; RAP: 2051 J kg^{-1}). Minor mischaracterizations of low-level moisture contributed to a large response in MLCAPE during the vertical integration of the parcel trajectory.

617
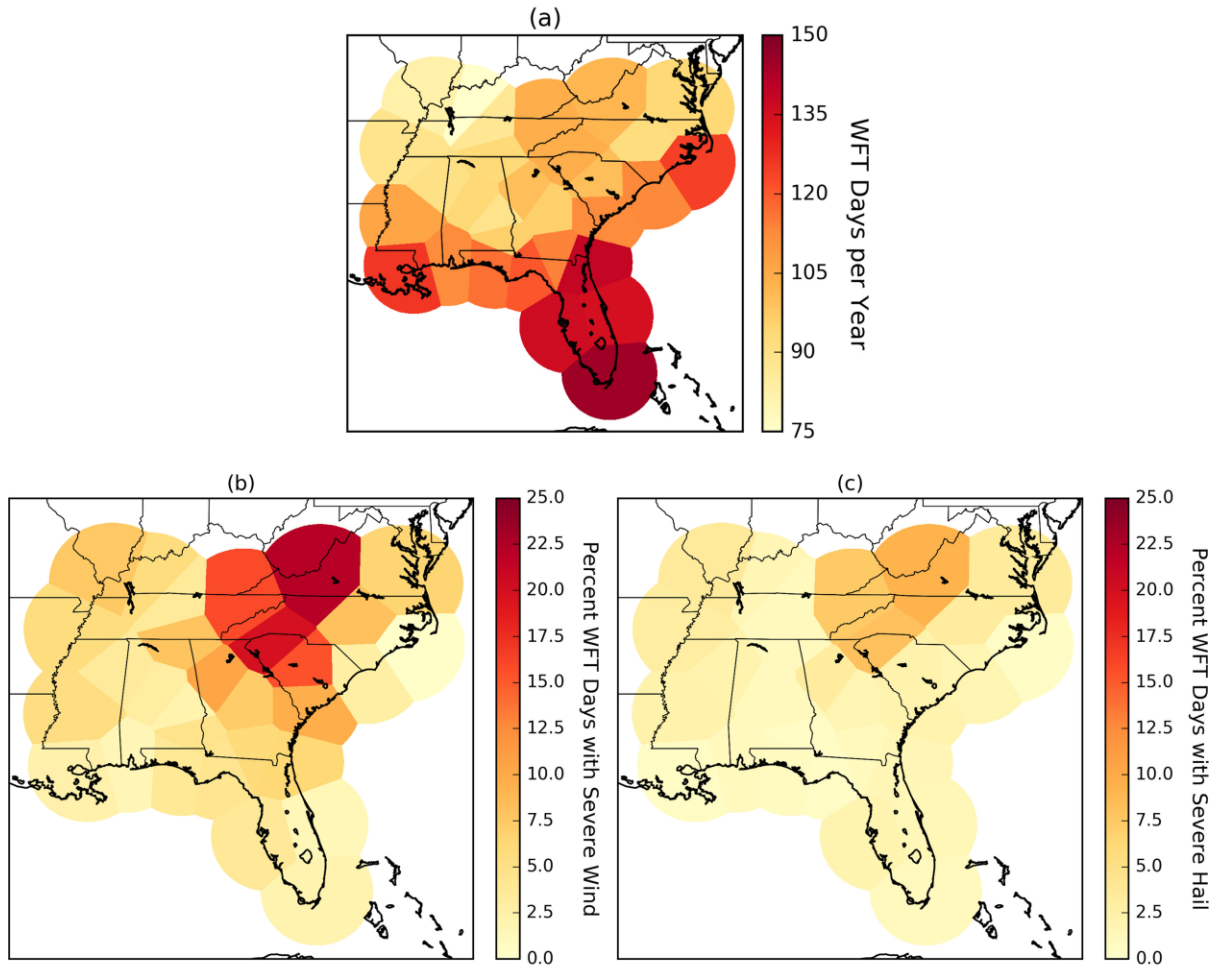
618 **Figure 4.** Average number of WFT days during the four-year study period (a) compared to the

619 proportion of WFT days affiliated with severe wind (b) and severe hail (c) events.

620
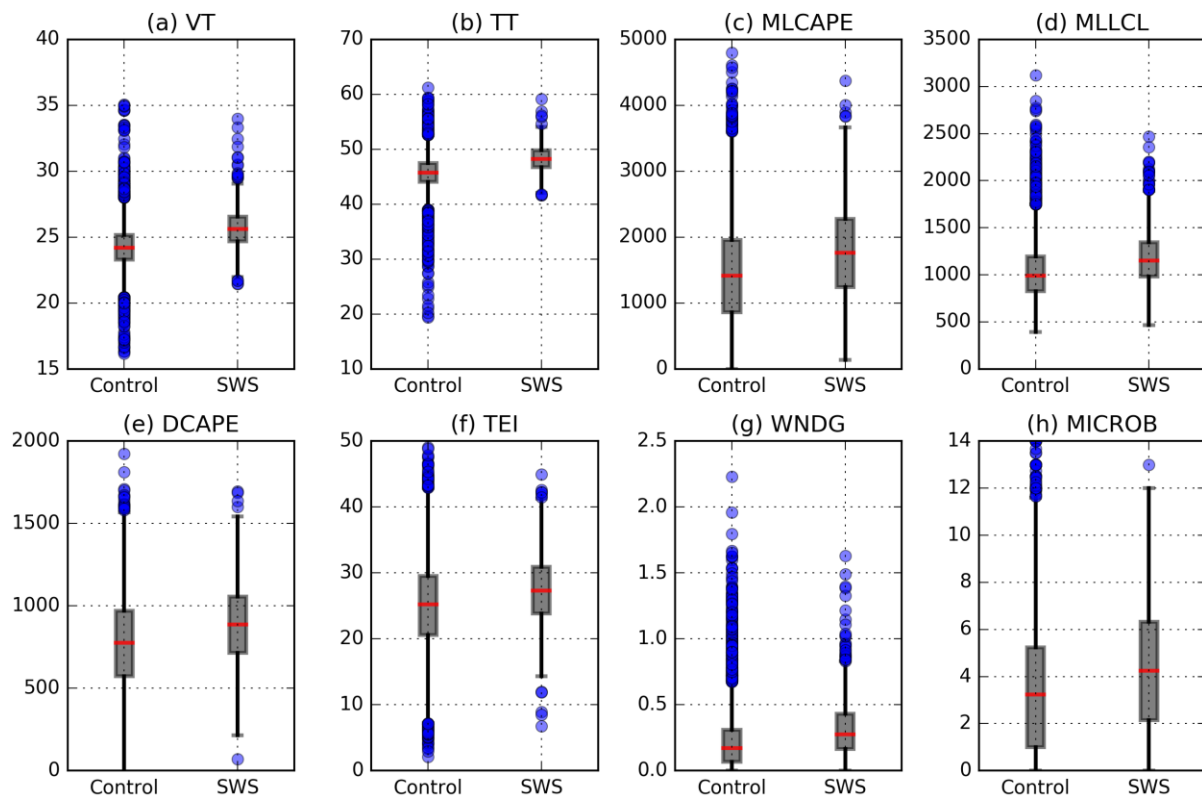
Figure 5. Boxplots of selected convective parameters that demonstrated skill in differentiating
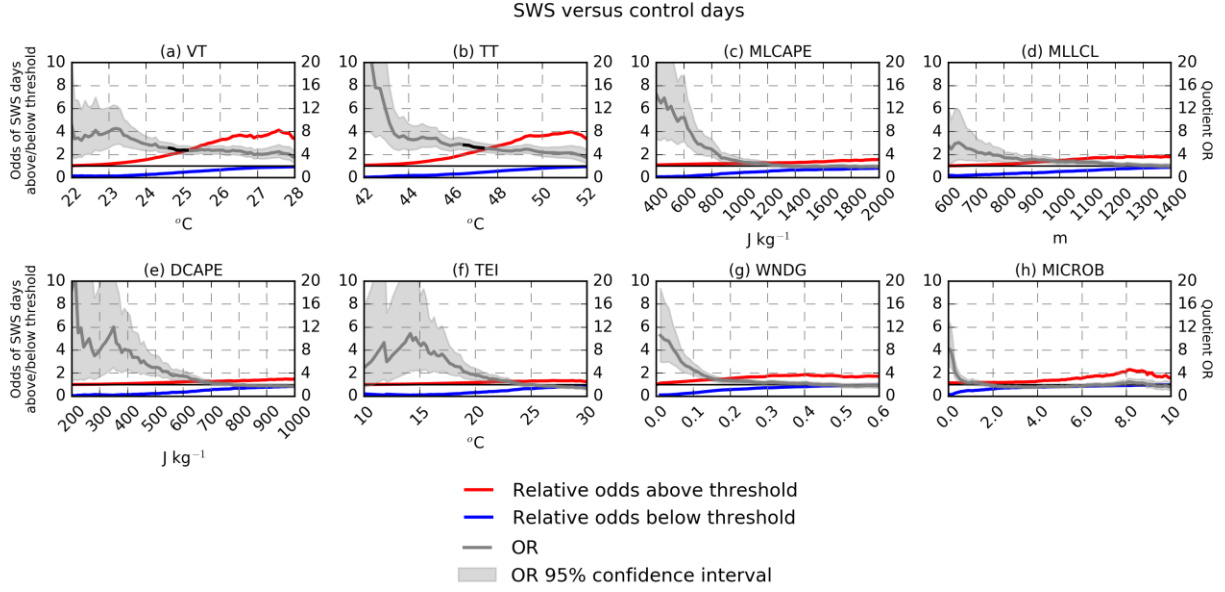
622   between the control days and SWS days.

Figure 6. ORs for the same eight convective parameters shown in Fig. 5. Whenever the OR, defined by Eq. (2), results from a numerator (red) ≥2 and a denominator (blue) ≤0.5, then the OR is drawn in black. The left y-axis expresses values corresponding to the OR's numerator and denominator (red and blue lines), and the right y-axis corresponds to the OR value (gray line). At very low and very high threshold values, the variance of the OR may be undefined, and the 95% OR confidence interval cannot be computed.
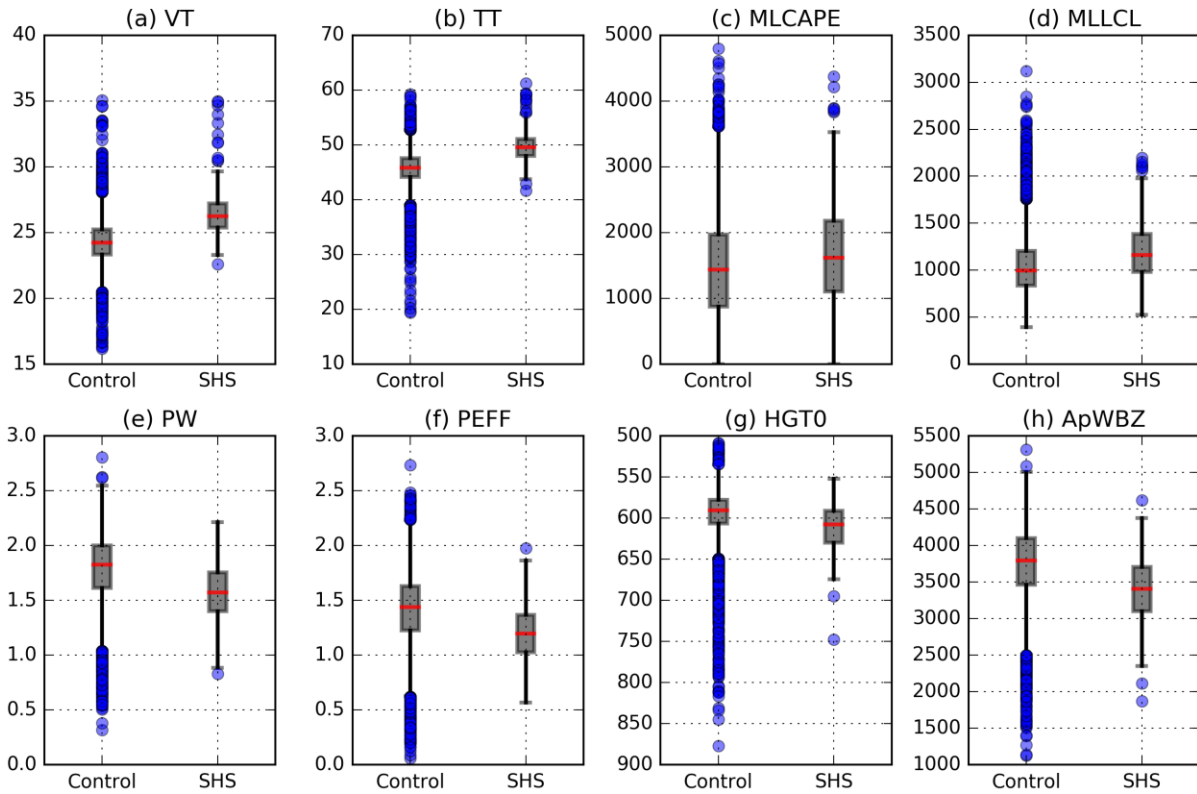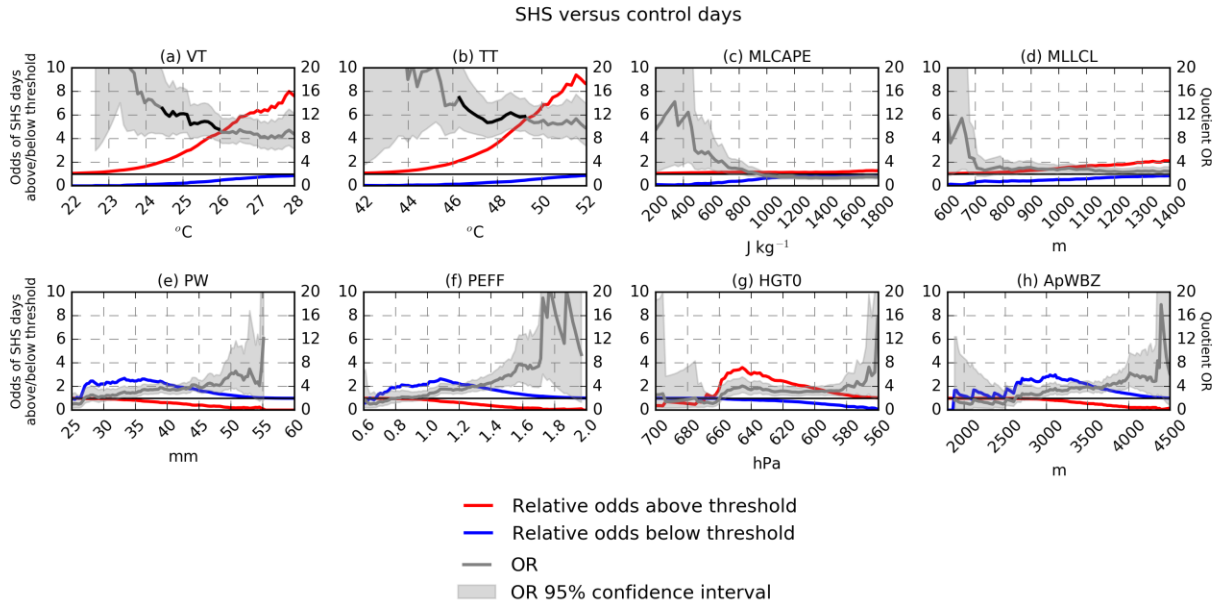
630

Figure 7. Same as Fig. 5 except for SHS days. Panes (a)-(d) replicate the same variables shown in

Fig. 5 whereas (e)-(h) are replaced with four SHS-specific parameters from ~~Table 5~~Table 6.

SHS versus control days

Relative odds above threshold
Relative odds below threshold
OR
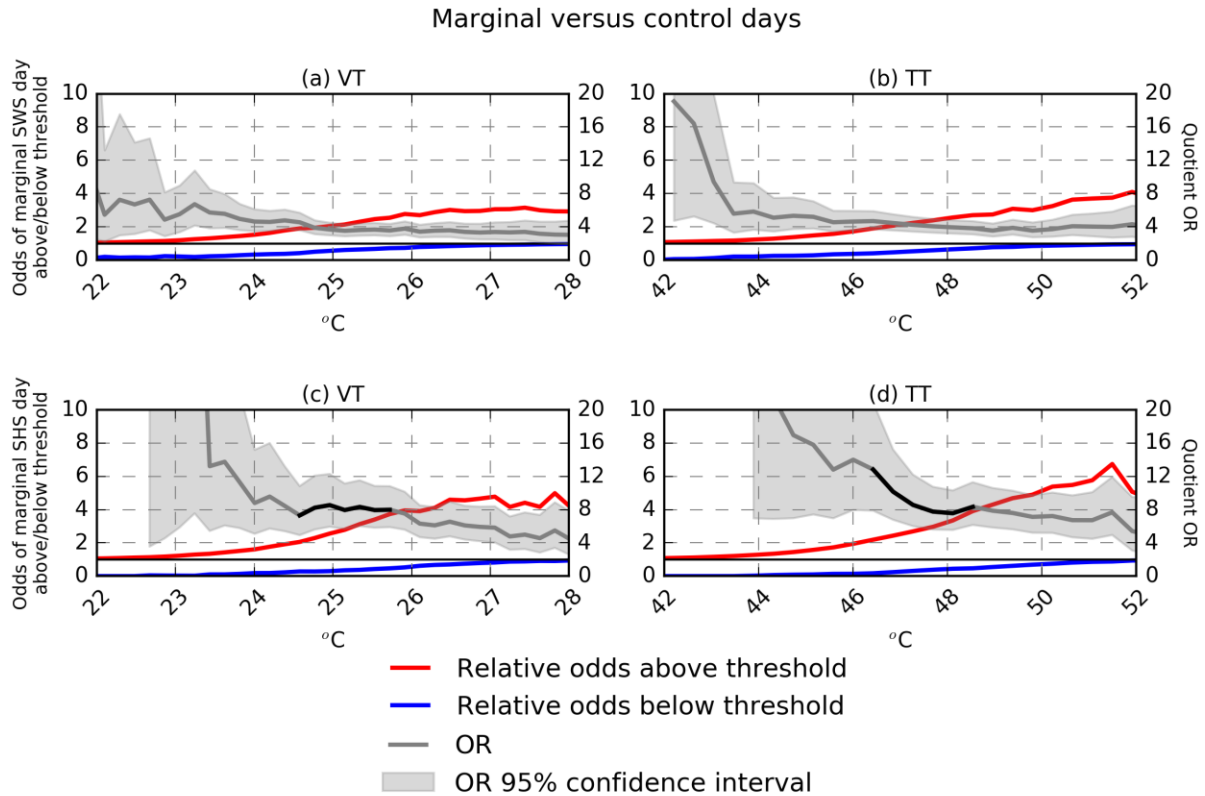OR 95% confidence interval

633

634   **Figure 8.** Same as Fig. 6 except for SHS days. Panes (a)-(d) replicate the same variables shown in

635   Fig. 6 whereas (e)-(h) are replaced with four SHS-specific parameters from ~~Table 5~~Table 6. At

636   very low and very high threshold values, the variance of the OR may be undefined, and the 95%

637   OR confidence interval cannot be computed.

638

Marginal versus control days

Figure 9. Same as Fig. 6a-b (a-b) and Fig. 8a-b (c-d) except that only marginal SWS and SHS days are used to calculate the OR. At very low and very high threshold values, the variance of the OR may be undefined, and the 95% OR confidence interval cannot be computed.

639

640

641

642