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.*

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.*

Line 62: Should be represents instead of is represented?

> *This will be corrected in a revision.*

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

> *This will be corrected in a 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

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.

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.

Line 211: Should be precedence, not precedent?

Yes, this will be corrected.

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.*

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.

> *We will augment this text to explain that the freezing level and dry layer are relevant to the melting process.*

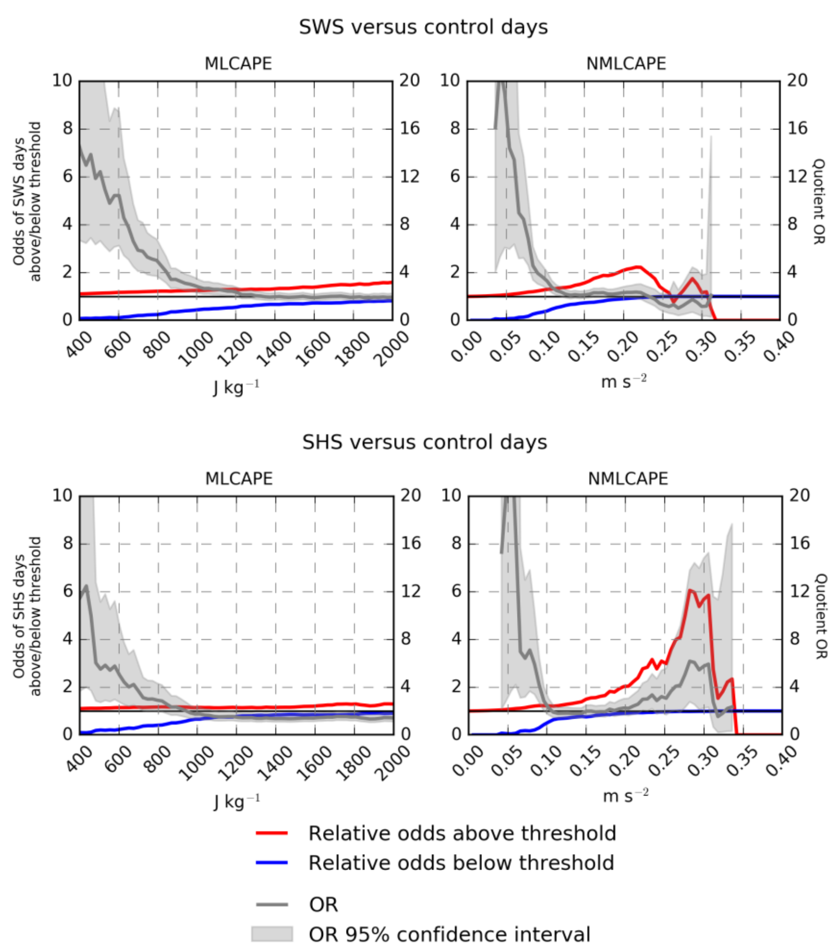
Same as authors, I was also surprised to see that measures of lapse rates (such as Vertical Totals) perform better than CAPE itself. Apart from possible model errors, I suspect two other reasons for that:

> 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.



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.