Suggestions for revision or reasons for rejection

I commend the authors for their efforts responding to the reviewers' comments. These comments have been dealt with in an appropriate manner and have provided sufficient clarification (also in the manuscript) that my recommendation is to accept this for final publication.

I will suggest a set of (related) technical corrections:

1. Report ascent rate in m/s rather than km/min, as the former is more immediately recognisable. This affects Figure 11 and line 402.

This suggestion has been completed and is reflected in the most recent revision of the paper.

2. Regarding Figure 11, I do not see the "dashed line" mentioned in the caption.

The dashed line has been made thicker to be more apparent on the plots in Figure 11.

3. To make the ascent rate discussion slightly more statistically interesting/robust. The authors could consider including a set of quantiles (e.g. 25th, median, and 75th) in Figure 11, to better understand the likelihood of large ascent rates or descent at different stages of the life cycle.

We have added a median line (solid black) and 10th and 90th percentiles lines (dotted black) to Figure 11 to further elucidate the likelihood of larger ascent/descent rates from the methodology presented.

Thank you for your additional feedback to improve the quality and readability of this paper.

Suggestions for revision or reasons for rejection

All line numbers refer to the tracked version of the manuscript.

Major comments:

The authors attempt to show climatology statistics of convective cells near the Houston, TX area using 4 years of radar and satellite data during the summer seasons. A modified MCIT tracking algorithm was used. My main concerns are as follows:

(1) Little radar data quality control was performed, such as attenuation correction and signal-noise-ratio (SNR) within the polar coordinates before regridding the data into Cartesian coordinate. In addition, key variable VIL should limit Z to 56 dBZ as introduced by Greene and Clark (1972), this will also require a rerun for the entire dataset and tracking.

We have re-run the dataset using a Z_H limit of 56 dBZ and remade all of the figures. In short, there was essentially no change to the results from adding this Z_H limit. As for attenuation corrections, since KHGX operates at S-Band, as do all WSR-88D radars, the effects of attenuation are minimal and would only be present with the most extreme convection. As for SNR, the data freely provided from the WSR-88D network does not contain SNR and these corrections cannot be done.

(2) Using AOD as proxy for aerosol condition within clouds is a poor choice. AOD is a column integrated optical product, its own bias as to cloud invigoration cannot be quantified. According to Stier (2016) AOD explained only 25% of the CCN variance, not to mention the underlying key microphysical parameters like droplet concentration. It is suggested to use Rosenfeld et al., 2016 cloud base retrieved CCN concentration as a more direct aerosol signal to cloud invigoration statistics.

The method suggested in Rosenfeld et al. (2016) uses polar orbiting satellites, which would severely limit the analyses presented herein. Using GOES-16 AOD including only "medium" and "high" quality measurements for calculating a 30-minute pre-cell initiation mean AOD value for the location of cell initiation already reduced our shallow convective dataset by 93.4% (from 35,974 to 2,361 cells), reduced our modest deep convective dataset by 96.2% (from 7,930 to 303 cells), and reduced our vigorous deep convective dataset by 95.6% (from 4,869 to 212 cells). Using polar orbiting satellite data would limit our sample size even further and may leave us with only a handful of cases or none at all. These reasons are why we chose to go with GOES-16 AOD as a proxy for aerosol concentration.

(3) The methodology of dividing the cells into shallow, modest deep, and vigorous deep require further justification or maybe simplification here. See detailed comments below.

The reasoning for dividing into three categories rather than just "shallow" and "deep" is because deep convection can vary substantially in intensity. As shown by the analyses presented herein, simply dividing deep convective cases into these two groups shows noticeable differences in cell characteristics. The echo top height is used as the discriminator because the other thresholds were originally designed to do as you suggest and simply divide into shallow and deep convection, but adding the different echo top height thresholds to further divide into "modest" and "vigorous" deep convective cells showed clear differences in behavior that we believe is important in discriminating between weaker and stronger deep convection.

(4) The fonts for most figures are too small to read. Please modify.

Fonts were enlarged on most figures to improve readability.

As my main concerns will require a rerun of the overall dataset, I will stop here and continue review process once the updated dataset is ready.

The concerns mentioned in your major comments have also been further addressed below while responding to your minor comments.

Minor comments:

1. Line 10, "Radars have been traditionally used to provide the convective clouds characteristics." This statement is a bit odd here. Radar was first invented during war time and was used for missile detection instead of weather. In addition, once radar is applied in meteorology studies, both warm and cold season weather phenomenon are studied without priority rankings.

Reworded this sentence to provide clarity of meaning.

2. Line 14, consider changing "warm" to "summer" as Houston is warm from April to almost the end of Nov.

We used "warm" here since summer is typically considered June, July, and August, whereas September begins meteorological autumn. Therefore, we have chosen to leave the wording as is.

3. Line 36-42, it is a bit confusing here what modelers are really missing during the debate of warm and cold phase convective invigoration, can you elaborate on what is being debated here? In addition, doesn't the models intercomparison project from van den Heever's group show the models cannot agree on each other in terms of precipitation around Houston on the same set up and case? This sounds like the modelers are more debating on their model's inconsistence instead of an invigoration theory.

We have added information to further describe warm- and cold-phase invigoration and inconsistency among models. While we agree with your point, this study is not meant to investigate model inconsistencies, rather it is meant to provide climatological analyses of convective case types for easy comparison with model output.

4. Line 49 – 50, TRACER and ESCAPE were choosing Houston mainly for the variation of aerosol conditions here. In addition, it is well equipped with WSR-88D radar coverage, LMA coverage, TCEQ network, etc. Basically, we are treating Houston as a natural laboratory to study aerosol induced microphysical processes here. Please modify and include this in the text.

We have added text to include the importance of the variation in aerosol conditions.

5. Line 91-95, Houston is a natural lab as stated earlier due to its variability in terms of aerosol conditions. Although the authors identified the local pollution source, it is not reasonable to simply state SW is pristine, and NE is polluted. The state of pollution is not only determined by local sources but also synoptic weather conditions. In other words, according to Rosenfeld et al, 2016's satellite retrieval technique, both SW and NE of Houston can be polluted or pristine in a case-by-case scenario.

While we agree with the sentiment here, going through and manually identifying the surface flow for all hours analyzed is not reasonable. The reasoning for simplifying as we have is based on the general flow across the Houston area during the months considered. We have added text that specifies that increasingly large errors may come about as the flow in a given case deviates more from the general flow.

6. Line 99, using AOD as a proxy for aerosol within cloud or related to cloud invigoration can be quite noisy and subject to false conclusions. AOD is a column integrated optical product, its own bias as to cloud invigoration cannot be quantified. In an extreme case, one can expect a high AOD case while the aerosol contributed to is over/under the cloud column entirely. In other words, AOD has little to do with the CCN actually got activated into cloud droplets.

To combat the noisy nature of AOD, we only used AOD data that were denoted as "medium" or "high" quality observations. We then collected these values for the 30 minutes prior to cell initiation and used the mean AOD value as the cell's AOD value for initiation. We never try to separate CCN specifically from the total aerosol population. AOD are also not collected when clouds are present which mitigates the concern of AOD being over/under clouds. We were aiming to maintain the largest number of cases possible by using this method. Using the method presented in Rosenfeld et al. (2016) with a polar orbiting satellite would substantially reduce the number of cases analyzed here and would introduce huge temporal discrepancies between the time of satellite flyover and cell initiation.

7. Line 113, GOES BT13 is an IR product and has a resolution of 2 km and has a 5 min resolution. So what procedure has been done to match the radar and satellite data both in space and time? Simply the nearest neighbor perhaps? Then how much bias does this procedure will introduce to the overall dataset?

The GOES data were regridded to the same grid as ZH from KHGX. We used nearest neighbor temporally. While this may introduce some error, since the scan time of KHGX and GOES-16 are approximately the same, the longest possible difference between these two datasets is 2.5 minutes. We have added text to mention this.

8. Line 120-125, For VIL calculation, Z should be capped at 56 dBZ, as introduced by Greene and Clark 1972 to avoid possible ice phase hydrometeor contamination. Please rerun the cases with this threshold.

We have rerun the cases with the capped threshold and the results have not changed. Text has been added to mention the 56 dBZ cap.

9. Line 147, ETH definition should not be a fixed Z>-10 dBZ as the SNR degrades the signal with distance. Please use an SNR threshold instead here, say SNR > 10 or 15 dB. This step should be done in the polar coordinate before regrid the dataset into Cartesian coordinate.

Level-2 WSR-88D data does not contain SNR. Therefore, we cannot complete this request.

10. Line 149-150, according to the authors description, Hcell is not corrected with increasing range, then if the authors use cells 100 km from the KHGX, the base scan tilt is already approaching melting layer height (4 km), this is quite strong simplification suggested in the manuscript and subject to underestimation of Hcell. In addition, Hcell depth should not use detectible signal without quality control, but like ETH, use SNR masked signal.

We agree that H_{cell} is subject to underestimation with increasing distance. However, this underestimation would affect distant shallow cells primarily, not deep convective cells. Given the number of shallow cells we are analyzing herein, losing some shallow cells toward the edges of our domain is acceptable. As mentioned previously, level-2 WSR-88D data does not contain SNR, and as such, we cannot apply an SNR mask.

11. Table 1., it seems the only difference between Modest Deep vs Vigorous Deep is the lifetime max ETH difference. Why is it? What's the author's justification and objectives here? Would it be simpler if you combine the two deep scenarios? In addition, why does shallow convection cells are limited to 30 km2? Any justification? Sounds a bit random.

The empirically-derived thresholds in Table 1 are there only to ensure that we are looking at shallow or deep cells. The reason for the difference in ETH only for deep convection is as follows: the cell has already reached "deep convection" status if all other thresholds are satisfied. However, not all deep convection is created equally. Modest deep convective cells are meant to capture convective cells which have reached a deep convective state, but are not intense enough to grow into the most intense convective storms that reach the tropopause. As shown in our analysis, while modest deep convective cells share many similarities with vigorous deep convective cells, there are also stark differences between them. Shallow cells are limited to an area of 30 km2 to prevent tracking large shields of stratiform precipitation, which shares many of the existing thresholds we identified to isolate shallow cells. The area threshold for shallow cells ensures that we are looking at a small, discrete cell, not a large cohesive blob of stratiform rain.

Thank you for your in-depth comments to improve the quality of this manuscript.