Overall, it looks great. I think there is some content describing statistical methods that can be moved to the appendices and also I'm a bit confused why you're so insistent that the 2-size model is the best. I think there is most likely a continuum of microburst sizes and the 1-size model is best to probe that continuum. I'll describe that further below.

Line#: Comment

4: S. G. Claudepierre -> add Seth's middle initial

9: "microbursts" -> plural

11: remove "on the order of a few 10s of km" and replace with "ranging from ~40 to ~130 km"

17: add "the" here: "...orbit and the upper atmosphere"

21: "is" should be "are" on this line following "microbursts"

22-23: "contribution" to what? Suggestion: "...contribution as a global loss process."

23: remove "further" here

36: remove "the dynamics" here and replace with "various"

36: reword the sentence starting "One of the acceleration and loss mechanisms extensively studied is..." to "One of the extensively studied mechanisms responsible for both acceleration and loss is..."

59-60: add parentheses around the "i.e." statement.

83: after "latitude" here put ", i.e. along-track in orbit."

93: correction: "The Aerospace Corporation" with a capital T

101: AC6 dos1 is >35 keV electrons: see Blake and O'Brien 2016 and be consistent with that

110-111: Should this read "temporally correlated" OR "observed simultaneously"?

116: See Lorentzen et al. (2001a; b); O'Brien et al. (2004); Alex Crew's PhD thesis

131: "was" should be "were" here

143, 145, 147: Just spell out "cross-correlate" here; you go a bit overboard with acronyms which can confuse/frustrate readers

159: Define explicitly what you mean by "spatially aligned"

166-167: revise to "...above any AC6 separation, s."

172-194, 238-268, and 270-286: These sections should maybe be refined into either a section on Statistical Methods or just summarized briefly in the main text with the full details put into an Appendix.

Also, in those sections, you should really demonstrate how equation 1 is equal to equation 8 so long as B is completely independent of A and P(A). Both equations are forms of Bayes Theorem, with Equation 8 being one of the most widely accepted forms despite you never referring to equation 8 as Bayes Theorem... I think as written, this is all a bit misleading and the theory is all so spread out that it might confuse/mislead/bore readers. As written, it is a slog to get through these sections since each one starts with methods then goes into results on their own; that resulted in me really losing the main points to be compared from each set of results with the different models.

186: revise to "...in the ith bin"

Figure 3: Define "PD" in the Figure caption or directly in the axis label so that readers don't have to dig for it in the text

234: add "The" here: "The microburst footprint is..."

238: "MC", again... going a bit overboard with the use of acronyms. I suggest just spelling out to avoid confusing/frustrating readers

245: add "the" here: "in the ith bin"

253: add "it" here: "If it exists,..."

Figure 5: Does the analytical model assume the microburst needs to be centered over one of the two spacecraft? The figure kind of indicates that, but if that isn't true, then you should specify so explicitly. Also, be sure to stress this is not true for the Monte Carlo model either.

Figure 5: Label A(r,s) as the shaded region on the figure.

274: remove the comma after "x" and replace the word "are" with "as" here: "...as the P(r) model parameters.."

314: add a comma here: "...of microbursts, have a size..."

320: "microbursts" should be plural here

330: replace "these" with "our new" here: "Lastly, our new results are..."

344-345: I don't think you're "naively" assuming microbursts are only one scale size. By using the single-size microburst model over a range of scales, you show that the observed distribution can be constrained by microbursts scaled between ~40 and ~130 km in size. That model doesn't at all mean that the microbursts need to be just one size; what you've shown is the observations are fit by a range of microburst sizes and you've quantified the limits of the range. Considering that, Figure 6 is really the crux of your results...

348-349: What makes you think the 2-size model is any better? I think the discussion and details on that model can actually be removed entirely and you'll have a much clearer set of conclusions. I think you should stick with the range from the 1-size model (Figure 6) and just lump up all of the 2-size model with the line stating: "A variety of continuous P(r) such as..." This additional set of results from the 2-size model really confuses things when I think you nailed it that it fits the observations a bit better because it has more free parameters. All around, the results from it are consistent with that range of sizes from the 1-size model, so why does it deserve so much more discussion/credit? I think it is very, very unlikely physically that there are two distinct sizes of all microbursts. Why not three? Why not four? I think a distribution of sizes is the best we can estimate here from the AC6 results, and you've done that already with Figure 6.

365: You should also cite Turner et al. (2017) who used MMS to show that individual chorus elements are phase coherent up to at least 70 km scales at L ~ 8 but that the coherency starts to break down around that scale too (not all of the elements analyzed were phase coherent).

386-387: Again... the 2-scale model should not be held above all others here. I think what you've done with all the different models is shown there is a range of microburst scale sizes that can fit the observations and that range is between ~40 and ~130 km with a median size ~70 km.

391: "affect" should not have an "s" here

391: revise here: "an X-ray imager"

397: you might want to note here that those "relatively small regions" are on the order of thermal electron scale sizes. (10s of km)

Nice job, Mike, and thanks for including me on this

Drew