Response to Reviewer #3

Thank you for your careful and thorough reading of the manuscript and your thoughtful comments and suggestions. We apologize for the delay in revising the manuscript, as we spent a significant amount of time on the manual validation and inspection of the identified derechos. According to all three reviewers' comments and suggestions, we have made several significant improvements in our manuscript, which we want to highlight before the point-by-point response.

Firstly, we have changed the wording of some terms so that they are more distinguishable and accurate, and we have also improved the language of many sentences so that the manuscript is more understandable. A derecho is a windstorm, while an MCS is a convective system. They are different concepts. Therefore, we change "derecho" to "derecho-producing MCS" when we refer to the MCS associated with a derecho, and "derecho feature" has been changed to "derecho." The dataset developed in this study includes tracking of both derechos and corresponding derecho-producing MCSs.

Secondly, we have incorporated a "forward propagating" criterion in our derecho detection algorithm. Our initial understanding of the term "forward propagating" was incomplete, and we failed to recognize its critical role in defining a derecho proposed by Corfidi et al. (2016). After careful consideration and evaluation, we have adopted and modified two criteria from Corfidi et al. (2016) to establish the definition of "forward propagating." One is that the acute angle between the averaged bow echo orientation and the bow echo series' propagation direction is larger than 45°, and the other is that the bow echo series' propagation speed is at least 30% faster than the 500-hPa background wind speed. Implementing the "forward propagating" criterion removes many windstorms externally forced by extratropical cyclones, aligning with the purpose of Corfidi et al. (2016), which intends to define derechos as internally driven windstorms. "Externally forced" and "internally driven" reflect distinct physical formation mechanisms of those windstorms, which is why Corfidi et al. (2016) proposed a physically based derecho definition. With the updated detection algorithm, the derecho number between 2004 and 2021 has been substantially reduced from 556 to 274 (for ISD) and 220 (for SED). In addition, due to the inclusion of the "forward propagating" constraint in our derecho algorithm, we have decided not to change the name of our dataset to "high wind-producing bow echo." We have updated all the results in Section 6 based on the improved dataset.

Thirdly, we have developed another parallel dataset using gust speeds from the Storm Events Database (SED). Now, our derecho dataset consists of two subsets: one based on gust measurements from the global hourly Integrated Surface Database (ISD) and the other based on SED gust speeds. Although there are some discrepancies between the two subsets, their agreement is much larger than their difference (Figures 9-12 in the revised main manuscript; or Figure R1 below). Moreover, both ISD and SED gust speeds have limitations and uncertainties, hence differences between the ISD-based and the SED-based datasets are expected and understandable. In addition, it also indicates that our usage of lower gust speed criteria for ISD measurements is reasonable and does not change the derecho number much. We must emphasize that using lower gust speed criteria for ISD measurements than SED reports does not mean that the ISD-identified derechos are weaker than the SED-identified ones (or even not derechos). This

is a compromise, considering that ISD stations are limited and may miss many damaging gusts, as we highlighted in Lines 460-469 in the revised main manuscript (as below).

"We emphasize that, in Criterion 4, our ISD gust speed criteria are weaker than the SED gust speed criteria as well as those of previous studies (Squitieri et al., 2023; Bentley and Mote, 1998; Johns and Hirt, 1987), which also estimated the gust swath based on SED damaging gusts. As mentioned in Section 2.2.2, most SED gust reports are estimates, while ISD provides gust measurements from weather stations. SED estimates can capture potential damaging gust occurrences over a much larger area, although with large uncertainties. In contrast, due to the limited coverage of observational sites, real-time ISD measurements may miss substantial damaging gust occurrences in nearby regions. Therefore, we lower the gust speed criteria to capture potential derechos when using ISD measurements. It does not mean that the ISD-based derechos are weaker than the SED-based ones or even not derechos, as elaborated in Section 5."

Fourthly, due to the incorporation of "forward propagating" in the detection algorithm and the development of the SED-based derecho dataset, we have updated some sensitivity tests, evaluations, and the comparisons of our datasets with the NOAA SPC data in 2004 and 2005 and previous studies. Please see Section 5 for further details.

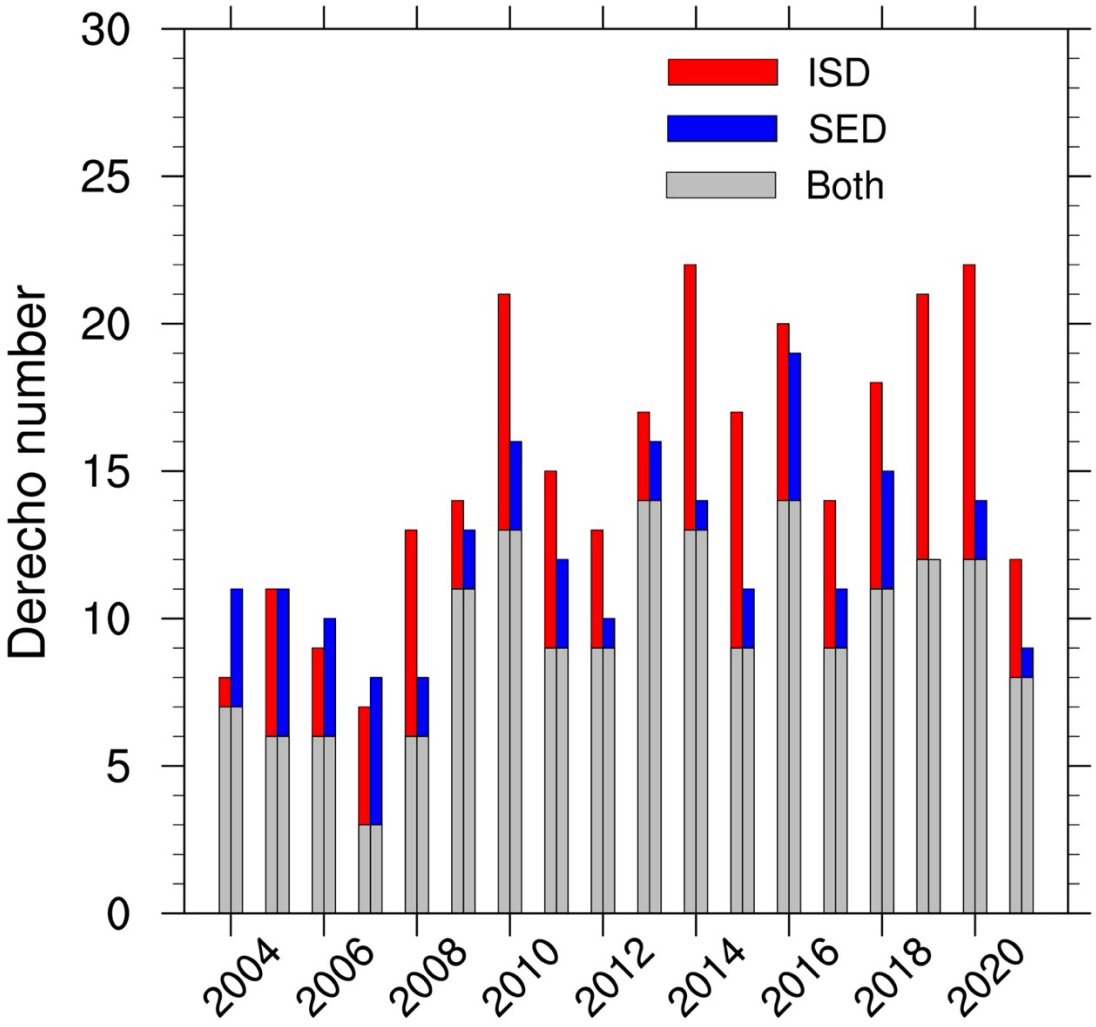

Figure R1. Bar chart of the annual derecho numbers from the ISD-based and the SED-based datasets from 2004 to 2021. Gray shading denotes derechos captured by both datasets, red shading refers to derechos only identified when using ISD gust observations, and blue shading represents SED-only derechos. The figure is the same as Figure 9 in the revised main manuscript.

Major Comments

This paper presents a novel, machine-learning scheme to objectively identify derecho-producing convective systems. The paper reads well and, not withstanding the minor comments indicated below, to this reviewer seems both well-organized and well-presented. I have only two main comments.

First, being familiar with the vagaries of the severe-weather report database, I support the authors' use of the Integrated Surface Database (ISD). While the ISD arguably is subject to its own limitations, the quality-control algorithms used offer a higher and more universal level of uniformity than that associated with the severe-weather report database. The large number of derecho-producing convective systems identified with the current approach of the authors compared with those of previous studies largely reflects the rather low wind threshold employed; use of a somewhat higher threshold (and/or duration threshold) would lower the number of

events identified. Obviously the true frequency of derecho-producing events remains unknown; the lower frequencies suggested by previous studies may be somewhat low.

Reply:

We sincerely appreciate your support for our usage of the ISD. To make the study more reliable and robust, and in response to the other reviewers, we have developed an additional dataset based on SED. Comparison between the two datasets justifies the usage of lower wind thresholds for the ISD. The overestimation in our initial manuscript is mainly due to incomplete implementation of the "forward propagating" criterion from Corfidi et al. (2016). Revising the "forward propagating" in our detection algorithm removes many externally forced systems. For more details, please see the highlights of our responses above.

The "true frequency" point brings into mind the main purpose of the present study --- objective identification of derecho-producing MCSs. The number of systems identified is sensitive to the underlying definition used in the scheme. This is where difficulty has arisen in the past and to some extent continues with the present paper. The omission of "forward propagating" from the current definition of the authors (page 516 ff) is problematic. Sustained forward-propagation is a fundamental aspect of derecho-producing convective systems. In absence of such a criterion one could argue that the approach described in the present paper is closer to that of a bow-echo detection scheme. Derecho-producing convective systems could be described as arising from bow-echo producing processes --- including rapid, sustained forward propagation --- that remain active for extended periods of time. I suggest re-visiting the abortive attempts made (lines 517-519) to identify the presence of forward propagation and refine the ML approach used here.

Reply:

Thank you so much for the suggestions. As mentioned above, our initial understanding of "forward propagating" is incomplete. We thought it was just a direction metric. After careful reading of Corfidi et al. (2016), we realize the importance of this criterion and its relationship with physical processes. After revising the "forward propagating" criterion (Lines 381-388 in the revised main manuscript), the dataset quality has been much improved.

> "The derecho bow echo series must exhibit forward propagation, based on two modified criteria from Corfidi et al. (2016):
>
> - The acute angle between the averaged bow echo orientation and the bow echo series' propagation direction must exceed 45° (Figure 6).
> - The propagation speed of the bow echo series must be at least 30% greater than the background mean wind speed at 500 hPa, derived from ERA5 data. The methodology

for calculating the averaged bow echo orientation, bow echo series' propagation direction and speed, and the background mean wind speed is detailed in Appendix A."

Minor comments (numbers refer to line numbers in version of 24 June 2024)
40. Change "magnitude" to "impact"
Reply*:*

Corrected. Thank you!

190. Consider adding a parenthetical description of "skip connections"
Reply*:*

Added. Thank you!

"Dense Nets are notable for their large number of skip connections (which create multiple paths for data to flow through the network without passing through every layer), and they can achieve comparable performance to very large classifier CNNs with only a fraction of the trainable parameters."

201. Clarify what is meant by "more distinct"
Reply*:*

The initial 556 positive samples are from 54 named derechos. The new negated 500 positive samples may be from 100 convective systems. Here, "more distinct" just refers to the samples from diverse "sources" to improve its representation.

267. Define or reference "binary cross entropy loss"
Reply*:*

A reference added. Thanks.

"It is trained using binary cross entropy loss (Bishop, 2006) on masks generated from its 384, 192, 96, 48, 24, and 12-pixel resolution feature representations (Huang et al., 2020), though only the full-resolution (384×384-pixel) output mask is used at inference time."

307-310. Well-stated
Reply*:*

Thank you!

312. Not completely sure what is meant by the "upper" and "lower" parts of the table
Reply*:*

Corrected to "upper triangular" and "lower triangular". Thank you!

366. Add parenthetically, "Derecho feature" after "DF"

Reply*:*

Thank you for your suggestion. We change the wording of “derecho feature” to “derecho” throughout the manuscript.

371-372. Not sure that this criterion would always be helpful...
Reply*:*

It is not always helpful. But it does help remove some extratropical cyclones. We manually check every identified derecho and high-wind-producing convective system to ensure the criterion does not produce any adverse impact.

385. Makes sense!
Reply*:*

Thank you! By comparing our SED-based and ISD-based subsets, we find lowering the gust speed criterion is reasonable for the ISD.

433-434. Agree with this focus; consider also maintaining the forward-propagating aspect.
Reply*:*

Thank you for your suggestion. The sentence has been deleted since we rewrite the evaluation section (Section 5) due to the incorporation of forward-propagating and the inclusion of SED.

455. Should "as" be "than"?
Reply*:*

The sentence has been deleted. Thank you.

480. Consider parenthetically adding "cold-cloud shield" given that acronym has not been used since line 129.
Reply*:*

Corrected. Thank you!

“The 50% areal overlap threshold in PyFLEXTRKR, which links consecutive cold cloud shields (CCSs), may fail to capture very fast-moving convective systems using hourly satellite and NEXRAD data.”

520. The addition of a simple schematic to illustrate the angles mentioned would be helpful
Reply*:*

Thank you for the suggestion. We have updated the algorithm schematic (as Figure R2 below) to illustrate the angle. Besides, in Appendix A, we provide more details on how we calculate them.

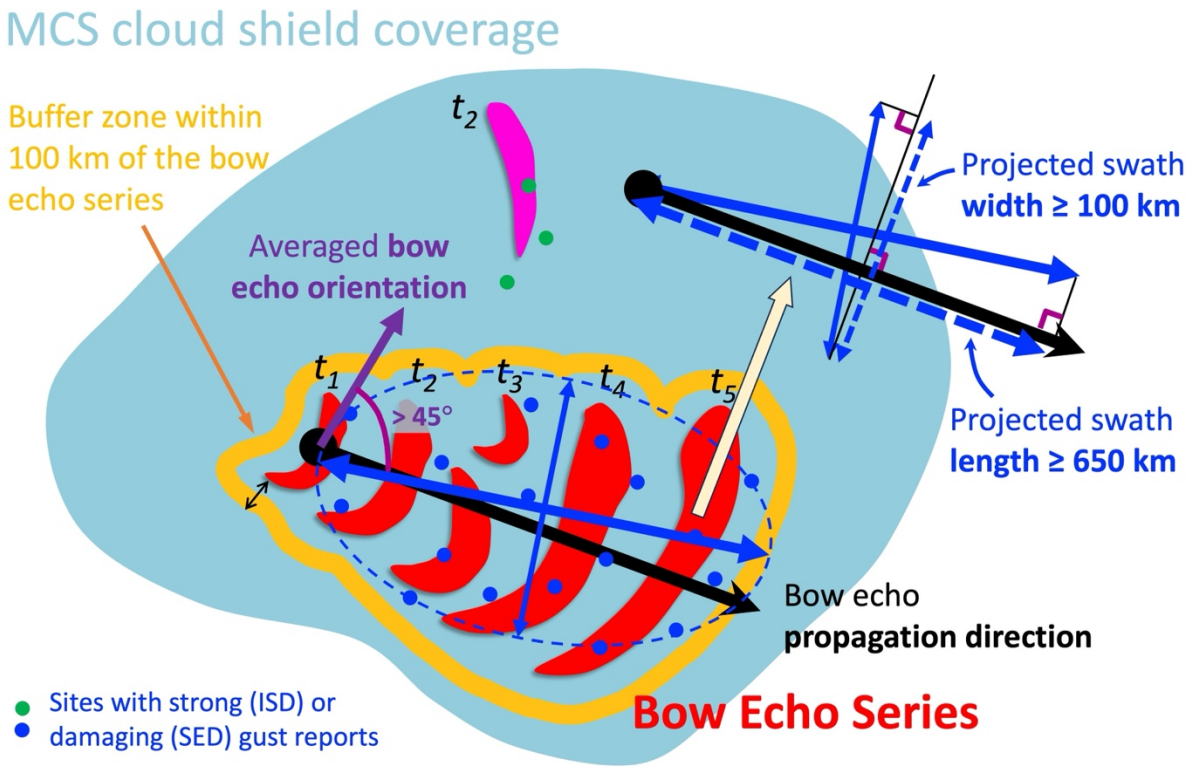

Figure R2. Schematic of the automated detection algorithm.

525-526. Good to see this explicitly stated
Reply*:*

Thank you for positive comment. Unfortunately, our understanding of "forward propagating" was incomplete. We remove the sentence and incorporate "forward propagating" in the derecho detection algorithm.

556. Should "west-east" be "northwest-southeast"?
Reply*:*

The sentence is deleted. We remove this sensitivity test, which is unnecessary after the development of the SED-based dataset.

588. Should "2014" in Figure 12 caption be "2004"?
Reply*:*

Corrected. Thank you for pointing out the error.

676. Capitalize "Weather" and "Review"
Reply*:*

Corrected. Thank you!

712. Add publication in which this manuscript appeared
Reply*:*

This is a report, and we add the website of the report.

731. Add year of publication (2020 (?), per line 234)
Reply*:*

Corrected. Thank you!

738. Capitalize "Python"
Reply*:*

Corrected. Thank you!

791. Capitalize "Atmospheric Sciences"
Reply*:*

Corrected. Thank you!

References

Corfidi, S. F., Coniglio, M. C., Cohen, A. E., & Mead, C. M. (2016). A proposed revision to the definition of “derecho”. Bulletin of the American Meteorological Society, 97(6), 935-949. https://doi.org/https://doi.org/10.1175/BAMS-D-14-00254.1

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., & Wu, J. (2020). Unet 3+: A full-scale connected unet for medical image segmentation. ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP),