

This manuscript combines LOFAR high-resolution VHF lightning imaging with dual-pol C-band weather radar to investigate “sparkles”, which are small intermittent discharges near storm tops. A two-stage DBSCAN clustering approach is used to separate sparkles from other lightning structures, followed by quantitative and qualitative comparisons by using radar variables (Z_h , V_{rad} , W_{rad}) and a fuzzy-logic hydrometeor classification (HMC). The key findings are that sparkles preferentially occur at high altitudes with relatively high Z_h , increased graupel occurrence, enhanced W_{rad} , and heterogeneous V_{rad} patterns, interpreted as enhanced turbulence. The results are broadly consistent with existing hypotheses involving strong updrafts lofting large charged hydrometeors toward a charged screening layer, with turbulence contributing to small spatial extent and intermittency. Overall, the topic is timely and the dataset combination is valuable; the study is promising for publication. I recommend Major Revision. The main issues are (i) potential methodological biases and incomplete reproducibility details in sparkle classification and radar–lightning matching, (ii) insufficient uncertainty quantification and statistical significance assessment, and (iii) the interpretation of enhanced W_{rad} as turbulence requires stronger controls of confounders. Addressing these will substantially strengthen the manuscript.

1. Beyond normalization constants and minimum cluster size, key DBSCAN settings should be explicitly reported (preferably in a parameter table).
2. The sensitivity tests must be accompanied and show whether sparkle counts, spatial patterns, and radar-based conclusions remain stable.
3. Please provide at least a limited validation using manual labeling for a subset of LOFAR images or a weak ground truth based on Scholten et al. (2023), reporting approximate false-positive/false-negative rates or overlap metrics.
4. The method likely excludes sparkle-like activity linked to positive leader breakdown (needles). This scope limitation should be stated more explicitly in the abstract and methods to prevent over-generalization of “sparkles”.
5. ERA5 may deviate from storm-scale winds, and within-sweep timing differences exist (Appendix B1). Please quantify the uncertainty of advection by comparing alignment metrics before/after correction, and report typical horizontal uncertainties.
6. Radius(2 km) is large relative to sparkle scales; include sensitivity tests (1/1.5/3 km) to show the $Z_h/W_{rad}/HMC$ contrasts are robust.
7. Even with an 8-km threshold, the detailed height distributions may differ; provide height histograms and stratified comparisons or explicitly control for height in a statistical model.
8. Figures 6–7 are primarily descriptive. Add at least one significance test/effect size (KS,

Mann-Whitney U, Cliff's delta, etc.), while noting limitations due to correlation.

9. The interpretation of positive Vrad patches as signatures of strong updrafts plus rotation (Figs. 10–11) may be correct, but needs clearer geometry and quantification.
10. Add a schematic conceptual model showing the hypothesized geometry of updrafts, graupel/hail charge regions, screening layer, turbulence, and sparkles.
11. The manuscript repeatedly uses “lighting” where “lightning” is intended (including the abstract). Please perform a systematic language/typo pass (also fix minor misspellings such as “system”, etc.).
12. For Fig. 6, consider reporting medians and IQRs (or 90% ranges) in addition to means. For Fig. 7 difference plots, mask or flag low-count bins as low confidence.
13. Add an objective echo-top/overshooting-top metric (radar or satellite) to directly test the proposed relationship.