<u>Review</u>: A Universal Multifractals perspective into the link between rainfall variability and temperature (egusphere-2025-3571)

Summary

The Authors apply the Universal Multifractals (UM) framework to rainfall and temperature data from three high-resolution disdrometer campaigns in northern France, and investigate how UM parameters (α, C_1, β, and H) and the scale-invariant maximum observable singularity γ_s vary with sample-mean surface temperature across two scale ranges: (a) event scale (i.e., 30 sec to 1 h) and larger scale (i.e., up to 11 days). They report a general increase of γ_s with temperature across both ranges, interpret this as a scale-invariant framework for understanding the commonly reported scale-dependent rate of increase of rainfall extremes with warming, and relate their findings to the Clausius-Clapeyron (CC) literature.

The underlying idea of using γ_s as a unifying descriptor that folds α and C_1 into a single extremes-relevant quantity is appealing, and the high-resolution disdrometer data are valuable. However, in its current form the manuscript substantially overstates its statistical rigor and novelty, neglects a directly relevant parallel body of literature, and raises several methodological concerns. I, therefore, recommend **major revision** prior to proceeding with publication in HESS.

Overall Contributions

- The UM framework is summarized clearly for a non-specialist audience, with appropriate pointers to foundational literature.
- The use of γ_s to combine α and C_1 trends into a single interpretable, extremes-relevant quantity is generally useful framing.
- Data and code availability via Gires (2025) is good practice and well-executed.
- The high temporal resolution (i.e., 30 sec) of the disdrometer data distinguishes this work from most similar existing studies.

Major Concerns

1. **The statistical evidence currently provided is weak and is not acknowledged as such**:
   - The headline claim that γ_s increases with temperature rests, at the individual-sample level, on Pearson correlations mostly in the range $| r | = [0.1, 0.5]$ (see e.g., Tables 3 and 5). A correlation of $r = 0.29$ explains roughly 8% of the variance.
   - With n in the hundreds, statistical significance at $p < 0.05$ is nearly automatic and says little about effect size.
   - The abstract statement that "*the scale invariant maximum observable singularity increases on average with greater temperature*" implies considerably less uncertainty than the data in Tables 3, 5, and Figure 9 support.
   - The Authors should report effect sizes with bootstrap confidence intervals (i.e., not just p-values), add Spearman rank correlations as a robustness check given visible heteroscedasticity in Figures 5 and 8, and substantially moderate the tone of the the abstract and conclusions.

2. **The binned analysis (see Figure 6) is mostly rhetorical work, and it is the weakest methodological point in the manuscript:**

   - The high correlations (i.e., r = 0.7 to 0.9) reported in Figure 6 arise from aggregating individual samples into 11 temperature bins and regressing across bin averages. This is a typical fallacy, as correlations across bin averages are almost always much higher than underlying sample-level correlations, and they blend "between-bin" and "within-bin" variations.
   - Bins are also unequally populated. The Authors discard bins with n = 1 and n = 4, but retain bins with modest n, thus the regression is unweighted.
   - This analysis should be treated as illustrative only, not as confirmation of the individual-event analysis. As currently presented, it invites the readership to conclude the trends are more robust than the underlying data actually support.

3. **Temperature is poorly justified as the predictor at the scales analyzed:**

   - The Authors acknowledge in the introduction (see lines 25 – 27) that surface temperature is a poor proxy for column water content at time scales beyond a day. They then use a single sample-mean temperature as the regressor on samples that extend up to 11 days. This tension is never resolved.
   - At the 11-day scale, $< T >$ plausibly reflects the seasonal cycle and prevailing circulation regime as much as any thermodynamic state variable.
   - The seasonal analysis (see Figure 9) in fact shows that γ_s at larger scales behaves differently in winter than in other seasons (i.e., the sign of the trend can reverse), which hints that the temperature signal across all seasons is partly a seasonal/regime-confounding artefact.
   - This possibility is mentioned but not properly reckoned with. The Authors need either to show the trends survive within-season at each site, or to reframe the claim as "*γ_s covaries with temperature-associated regime shifts*" rather than a temperature dependence per se.

4. **The snowfall filter is crude and asymmetric:**

   - Dropping large-scale samples where any single timestep shows rainfall with T < 2°C, versus dropping events with mean T < 2°C at event scale, are two very different filters, and neither is physically defended.
   - The 2°C threshold is explicitly chosen to avoid "*discarding too much data*" (i.e., a convenience threshold, not a physical one) and the Authors note the more defensible 4°C threshold is avoided for this reason.
   - Mixed-phase precipitation between 0 and 4°C is well-known. Given the claims regarding low-temperature behaviour at larger scales, a sensitivity analysis using a stricter threshold (e.g., 4°C) and using a symmetric definition between the two scale ranges is essential.

5. **Instrumental uncertainty is dismissed too quickly:**

   - Parsivel^2 disdrometers are subject to 10% biases and known wind-induced effects (the Authors cite this themselves; see lines 83 – 87), then declare the issue "*outside the scope*" of this work. For a paper whose central claim concerns how higher-order moments and extreme singularities shift with temperature, and where temperature covaries with wind regime, convective vs. stratiform rainfall type, and drop-size distribution, this is not defensible.

- Wind-induced undercatch is itself temperature-correlated. At minimum, a sensitivity analysis or a qualitative discussion of the direction of potential bias is required; the correction scheme of Chinchella *et al*. (2024), which the Authors currently cite, should at least be considered.

6. **Sample sizes are marginal at the large scale:**

- For the ENPC Campus, 8 years of data split into approximately 11-day samples gives about 250 samples before snowfall filtering; SIRTA yields only 22 large-scale samples covering less than a year.

- Drawing conclusions from 22 samples over a narrow temperature range and then pooling them into in Tables 3 and 5 is not rigorous. The weakness of the SIRTA subset should be stated plainly and reflected in the interpretation.

7. **Critical omission of a parallel and directly relevant literature on multifractal rainfall extremes:**

The bibliography reflects almost entirely the Schertzer/Lovejoy UM lineage and omits a parallel body of work that has been addressing essentially the same questions for nearly two decades. The Authors should engage with, at minimum:

- Veneziano & Furcolo (2002) and Langousis & Veneziano (2007), which established the derivation of Intensity-Duration-Frequency (IDF) curves from multifractal representations of rainfall;

- Langousis, Veneziano, Furcolo, & Lepore (2009), which developed practical estimation of IDF curves from multifractal models under pre-asymptotic conditions;

- Langousis, Carsteanu, & Deidda (2013), which linked multifractal maxima to GEV approximations;

- Emmanouil, Langousis, Nikolopoulos, & Anagnostou (2020, 2022, 2023), which use multifractal scaling arguments to estimate IDF curves from short records, assess the CONUS-wide spatiotemporal evolution of rainfall extremes under a changing climate, and to project future extremes under high-emission scenarios.

Emmanouil *et al*. (2022) in particular address directly the question this manuscript claims to open: how multifractal scaling of rainfall changes under a warming climate and what this implies for extremes across scales. They document that shorter-duration extremes intensify faster, and interpret this through multifractal scaling arguments. Emmanouil *et al*. (2020) is also methodologically important here, as it benchmarks multifractal-parametric IDF estimation against annual-maxima and peaks-over-threshold methods on short records, which is directly relevant to the sample-size limitations of the present work. These papers must be cited and discussed substantively, not merely added to the reference list.

8. **Novelty is overstated and should be re-scoped:**

Once the Veneziano/Langousis/Emmanouil line of work is acknowledged, several novelty claims in the manuscript also become difficult to support:

- Using multifractal parameters to characterize rainfall across sub-hourly to synoptic scales is well-established (see also Tessier *et al*., 1993; de Lima & Grasman, 1999, and many others the paper already cites).

- Deriving extremes from multifractal parameters, including γ_s specifically, was done by Douglas & Barros (2003), Royer *et al*. (2008), and, more formally, by Veneziano, Langousis, & colleagues.

- Linking changes in multifractal parameters to changes in extremes under climate forcing was done by Royer *et al*. (2008) on climate-model outputs, and by Emmanouil *et al*. (2022, 2023) on observations and downscaled data.

- The qualitative conclusion of Section 4.4 that multifractal scaling explains why extremes at shorter durations intensify faster under warming is anticipated empirically and interpretively by Emmanouil *et al*. (2022). Equation (9) is furthermore an algebraic consequence of assuming γ_s depends on T and that scaling holds; it is a framework-consistent identity, not an empirical test of the framework.

What is new, in my reading, is narrower than the current framing suggests:

(a) the use of surface temperature as a direct stratifying variable for UM parameters estimated sample-by-sample, rather than on time-segmented climate periods;
(b) the application to three 30-sec resolution disdrometer campaigns, finer than most prior UM studies; and
(c) the contrast between event-scale and large-scale regimes within the same framework, including the observation that H behaves oppositely across the two ranges.

These are legitimate contributions and the paper should be re-scoped explicitly around them.

9. ***H* behaves inconsistently across scale regimes, which should be highlighted and analyzed further:**

   - H decreases with T at large scales but increases with T at event scale (see lines 282 – 284). This opposite-sign behavior is flagged in a single sentence and then dropped.
   - Shifting non-conservation in opposite directions across scale regimes is physically significant and deserves interpretation, not just reporting.

Moderate Concerns

- Figure 6 contains a broken cross-reference (see line 247: "Fig. ??").

- The DTM parameter uncertainties reported as ± values in Tables 2 and 4 reflect the DTM fit for a single ensemble, not uncertainty arising from sample variability. Reporting them without qualification may mislead the readership; this should be stated explicitly, or sample-based confidence intervals should be provided.

- Linear regression of γ_s vs. T is used throughout without justification. CC relations are exponential; a log-linear or nonlinear specification should at least be tested.

- The sensitivity analyses on the upper 10% rain-rate percentile and on $T > 12$°C versus $T < 12$°C (see lines 303–306) report "*no significant differences*". Given the statistical significance demonstrated, absence of a detected difference is not evidence of absence. Using these null results to push back against prior literature reports of temperature-range dependence is not defensible.

- "*Excellent scaling*" is asserted repeatedly based on $r^2 > 0.99$ across roughly two decades of scale. High r^2 on log-log plots across two decades is nearly guaranteed for any monotonic process and is weak evidence of true scale invariance. More conservative language is warranted.

- The honest discussion of the stationarity assumption in the conclusions (see lines 369 – 371) is welcome, but belongs earlier in the paper, not in the final paragraph.

Minor Issues

- The abstract states that data are "*from three high resolution measurement campaigns that took place in Northern France between 2018 and 2025*". To the best of my knowledge (from literature), the SIRTA campaign was conducted in 2016 – 2017 and the ENPC Campus 1 campaign started in January 2018; if true, the date range in the abstract should be properly adjusted.
- Line 159: the distinction between H and the Hurst exponent is stated but offloaded to the appendix of a cited paper; a brief clarification in the present manuscript would greatly help the readership.
- Line 206: the DTM η range of [1.1, 1.5] is used without justification; this choice affects parameter estimates and should be motivated.
- Line 271: "*Very similar qualitative and quantitative*" is missing a noun.
- Tables 2 and 4 mix ± uncertainties of very different magnitudes (e.g., 0.0022 vs. 0.013) without comment.
- The manuscript would benefit from a careful language edit throughout.

Overall Recommendation

Major revision: The framework, data, and targeted contributions are worth publishing, but the current manuscript overclaims relative to the provided evidence and its position in the literature. In summary, the Authors should:

1. Rewrite the abstract and conclusions to reflect the actual statistical strength of the individual-sample evidence; re-scope claims around the binned analysis accordingly. Potentially adjust the date range of the campaigns in the abstract.
2. Add Spearman rank correlations, bootstrap confidence intervals on slopes, and weighted regressions for the binned analysis.
3. Properly address the seasonality/circulation confound, either by showing that trends survive within-season at each site, or by reframing the "*temperature dependence*" claim accordingly.
4. Add sensitivity analyses for the snowfall threshold (e.g., 4°C) and use a symmetric filter definition across scale ranges.
5. Discuss the direction and potential magnitude of instrumental biases, particularly wind-induced undercatch, given its covariance with temperature-linked rainfall type.
6. Substantively engage with the multifractal-extremes literature, acknowledging the relevant work of Langousis & Veneziano (2007), Langousis *et al*. (2009, 2013), Veneziano & Furcolo (2002), and Emmanouil *et al*. (2020, 2022, 2023) in the introduction, Section 4.4, and the conclusions.
7. Re-scope the novelty claim explicitly around the three defensible contributions identified above (i.e., temperature-stratified UM analysis; high-resolution disdrometer application; event-vs-synoptic contrast, including the opposite-sign behaviour of H).

8. Fix the broken figure reference, clarify the meaning of DTM parameter uncertainties, and perform a thorough language pass.

9. A more modest version of the paper's claim "*within a UM framework applied to three northern French high-resolution disdrometer datasets, sample-mean γs is positively associated with sample-mean surface temperature, with weak individual-sample correlations and with seasonality that cannot be fully disentangled from temperature; this provides a scale-invariant observational complement to the existing multifractal-extremes literature on climate-forced rainfall changes*" would be defensible. The current framing is not.

References

Borga, M., Stoffel, M., Marchi, L., Marra, F., & Jakob, M. (2014). Hydrogeomorphic response to extreme rainfall in headwater systems: Flash floods and debris flows. *Journal of Hydrology*, ***518***, 194–205. https://doi.org/10.1016/j.jhydrol.2014.05.022.

Chinchella, E., Cauteruccio, A., & Lanza, L. G. (2024). Quantifying the wind-induced bias of rainfall measurements for the Thies CLIMA optical disdrometer. *Water Resources Research*, ***60***, e2024WR037366. https://doi.org/10.1029/2024WR037366.

de Lima, M. I. P., & Grasman, J. (1999). Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *Journal of Hydrology*, ***220***(1–2), 1–11. https://doi.org/10.1016/S0022-1694(99)00053-0.

Douglas, E. M., & Barros, A. P. (2003). Probable maximum precipitation estimation using multifractals: Application in the eastern United States. *Journal of Hydrometeorology*, ***4***(6), 1012–1024. https://doi.org/10.1175/1525-7541(2003)004<1012:PMPEUM>2.0.CO;2.

Emmanouil, S., Langousis, A., Nikolopoulos, E. I., & Anagnostou, E. N. (2020). Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records. *Journal of Hydrology*, ***589***, 125151. https://doi.org/10.1016/j.jhydrol.2020.125151.

Emmanouil, S., Langousis, A., Nikolopoulos, E. I., & Anagnostou, E. N. (2022). The spatiotemporal evolution of rainfall extremes in a changing climate: A CONUS-wide assessment based on multifractal scaling arguments. *Earth's Future*, ***10***(3), e2021EF002539. https://doi.org/10.1029/2021EF002539.

Emmanouil, S., Langousis, A., Nikolopoulos, E. I., & Anagnostou, E. N. (2023). Exploring the future of rainfall extremes over CONUS: The effects of high emission climate change trajectories on the intensity and frequency of rare precipitation events. *Earth's Future*, ***11***, e2022EF003039. https://doi.org/10.1029/2022EF003039.

Gires, A. (2025). *Data for "A Universal Multifractals perspective into the link between rainfall extremes and temperature"* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.16406221.

Langousis, A., Carsteanu, A. A., & Deidda, R. (2013). A simple approximation to multifractal rainfall maxima using a generalized extreme value distribution model. *Stochastic Environmental Research and Risk Assessment*, ***27***(6), 1525–1531. https://doi.org/10.1007/s00477-013-0687-0.

Langousis, A., & Veneziano, D. (2007). Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Research*, ***43***(2), W02422. https://doi.org/10.1029/2006WR005245.

Langousis, A., Veneziano, D., Furcolo, P., & Lepore, C. (2009). Multifractal rainfall extremes: Theoretical analysis and practical estimation. *Chaos, Solitons & Fractals*, ***39***(3), 1182–1194. https://doi.org/10.1016/j.chaos.2007.06.004.

Royer, J.-F., Biaou, A., Chauvin, F., Schertzer, D., & Lovejoy, S. (2008). Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. *Comptes Rendus Geoscience*, ***340***(7), 431–440. https://doi.org/10.1016/j.crte.2008.05.002.

Tessier, Y., Lovejoy, S., & Schertzer, D. (1993). Universal multifractals: Theory and observations for rain and clouds. *Journal of Applied Meteorology*, ***32***(2), 223–250. https://doi.org/10.1175/1520-0450(1993)032<0223:UMTAOF>2.0.CO;2.

Veneziano, D., & Furcolo, P. (2002). Multifractality of rainfall and scaling of intensity-duration-frequency curves. *Water Resources Research*, ***38***(12), 1306. https://doi.org/10.1029/2001WR000372.