Response to Referee #2

In this response, the review comments are stated first (in black), followed by our "Replies" (in blue).

Reviewer #2 (Stefan Hergarten):

This manuscript is mainly about the scaling between event sizes and duration of a precursory phase. It is based on LPPLS model for the precursory phase.

First, I have to mention that I like the modeling approach and that I am still impressed how well the log-periodic behavior fits at least in some cases. This was, however, already the case before reading the recent manuscript.

> **Reply:** We thank the referee for the positive assessment of the LPPLS modelling approach and for recognising its ability to capture log-periodic behaviour in precursory geohazard dynamics. As the referee notes, the focus of the present study is the scaling between event size and the duration of the precursory phase, building on our previous work on LPPLS modelling.

Before going into scientific aspects, let me briefly comment on the potential impact of the paper in NHESS. I found the 2025 Science Advances and GRL papers by the same authors more intuitive. From my experience with own papers in NHESS, only a very limited part of the readership will really appreciate this work. If the manuscript was about how well the model can predict the event time t_c from an incomplete record (stopping earlier than t_c), it would probably be much easier to get the interest of the NHESS readers (which was one of the points raised by the first reviewers). However, this concern does not affect the suitability of the topic for NHESS, which is clearly given.

> **Reply:** We thank the referee for this helpful perspective and for confirming the suitability of the topic for NHESS. We agree that prospective prediction of t_c from incomplete records is of direct practical interest to the NHESS readership. However, the present manuscript is intended to address a different and, in our view, more fundamental question.
>
> The central object of this study is not primarily the predictability of failure time once a precursor has been detected. Rather, we investigate the characteristic duration of the dynamical regime that develops before failure and is associated with the nonlinear positive feedbacks driving the system towards a finite-time singularity. In this sense, the term "precursory duration" may perhaps suggest too strongly a forecasting-oriented interpretation, as if our main concern were the duration of an observable warning signal. What we are actually quantifying is the lifetime of the positive feedback regime itself: the interval over which the system departs from its previous dynamics and progressively organises towards terminal failure.
>
> This distinction is important. The question addressed here is therefore not simply "how much warning time is available?" but rather: how long does the instability spend in the positive-feedback regime that ultimately leads to failure, and how does that duration depend on the size of the system? Within the LPPLS framework, the onset time t_1 marks the beginning of this dynamical regime, while t_c marks the finite-time singularity associated with terminal failure. The duration $T = t_c – t_1$) can thus be interpreted as the characteristic timescale of the failure development process itself.
>
> Viewed in this way, the scaling of T with failure volume or characteristic system size is not merely an early-warning result. It is a statement about the temporal organisation of instability across systems of very different scales. The manuscript asks whether the duration of the nonlinear positive feedback regime is systematically controlled by the size of the failing system, and our results indicate that this is indeed the case for mechanically driven geohazards.
>
> This perspective is complementary to prospective forecasting. Prediction asks how accurately t_c can be

inferred from an incomplete record before failure occurs. The present study instead asks what sets the full duration of the dynamical regime leading to that failure. The first is primarily a forecasting problem; the second is a question about the physics and scaling of the failure process itself. We believe that distinguishing these two questions more explicitly in the manuscript will help clarify its scope and significance for the NHESS readership.

Prospective LPPLS-based forecasting of t_c is addressed separately in another recent NHESS preprint by the authors (https://doi.org/10.5194/egusphere-2026-2565), where the analysis is performed pseudo-prospectively using only the data available up to each analysis time. We will revise the present manuscript to make clearer that its objective is complementary: to characterise the global duration of the positive feedback regime preceding terminal failure and to determine how this characteristic timescale scales with event volume and system size.

My thoughts about the scientific aspect are similar to those already raised by the first reviewer. As a central point, the data shown here do not convince me that the Lagrange regularization approach for determining the starting time tau is good. For the examples in Fig. 1(a) and (b), it makes sense to me. Just visually in the sense that the precursory activity increases and exceeds the background. In (c) and (d), however, I cannot interpret it accordingly. In (c), I would even assume that the blue curve was higher than the data if we continued it for times < tau.

Reply: We thank the referee for highlighting this important point. We agree that the onset is visually more evident in Fig. 1a and b than in Fig. 1c and d. However, the purpose of the Lagrange regularisation approach is precisely to avoid relying on visual identification, which becomes subjective when the transition into the accelerating regime is gradual or noisy.

In our framework, t_1 is defined as the beginning of the interval over which the LPPLS dynamics become supported by the data. The model is calibrated over many candidate starting times, and the regularisation balances goodness of fit against window length to avoid the systematic preference for very short fitting windows.

The referee's observation regarding Fig. 1c is particularly useful. The LPPLS curve fitted after t_1 is not expected to provide a valid backward extrapolation to times earlier than t_1, because those earlier observations may belong to a different, pre-precursory dynamical regime. In this sense, a systematic mismatch in the backward extrapolation is consistent with, rather than contradictory to, the interpretation of t_1 as the onset of a distinct precursory regime.

"Do not convince me" does not mean that I assume that it was wrong. It is just that the data and the analysis shown here are not enough for me. Overall, the scaling relation developed here would be interesting enough for being published, even if we assume that it will not turn out to be useful for a prediction. However, I need to become more sure that it is not an artifact.

Reply: We appreciate the referee's clarification and agree that the central issue is whether the reported scaling could be an artefact of the analysis. We would, however, like to clarify an important point concerning how we think this question should be tested. We do not believe that repeating the scaling analysis with an onset time defined by an unrelated method would constitute the appropriate validation of the present result. This is because t_1 is not intended to represent an arbitrary phenomenological change point in the data. It has a specific physical meaning within the framework of the paper: it is the inferred onset of the dynamical regime in which nonlinear positive feedbacks begin to organise the system towards terminal failure.

The purpose of the LPPLS-Lagrange procedure is precisely to identify this regime endogenously from the dynamics. If the onset of the positive feedback regime is a genuine physical transition, then its beginning should be constrained by the subsequent evolution of the system towards failure. In other words, t_1 and the dynamics following t_1 should not be estimated as two independent objects. The value of a physical model is that it connects them. LPPLS provides such a connection by identifying the interval over which the observed dynamics are consistently described by the nonlinear feedback process leading towards the finite-time singularity, while the Lagrange regularisation prevents the trivial preference for increasingly short calibration windows.

For this reason, replacing t_1 by, for example, a displacement threshold, an acceleration threshold, a generic change-point algorithm, or a visually selected onset would not provide a model-independent benchmark. Each of these definitions introduces its own assumptions and, more importantly, generally identifies a different physical quantity. Demonstrating that the scaling survives after changing the definition of the quantity being scaled is therefore not, in our view, the decisive test of whether the LPPLS result is an artefact. It could instead obscure the physical question posed by the manuscript.

The appropriate robustness question is whether the LPPLS-Lagrange inference itself behaves as a genuine physical estimator rather than as a flexible fitting device capable of generating arbitrary onset times. Several features of the analysis directly address this issue. First, the onset is obtained from a well-defined minimum of a regularised cost function rather than selected by eye. Second, the stability analysis illustrated in Supplementary Fig. S2 shows that, as progressively more of the precursory record becomes available, the inferred onset converges toward a stable plateau rather than wandering arbitrarily. This is an important property: once sufficient information about the positive feedback regime is present, the model recovers a persistent estimate of its beginning. We will strengthen the presentation of this analysis in the revised manuscript and make its significance much more explicit.

More fundamentally, the existence of the scaling is not imposed by the LPPLS procedure. Applying the same model to all events ensures a consistent definition of the duration of the inferred positive feedback regime; it does not require that this duration scales with event size, that the exponent be close to 1/3 in volume, or that different mechanically driven hazard classes collapse onto a common relation. The analysis could equally have produced no correlation, a strongly scattered relation, or systematically different dependences among hazard classes. Indeed, volcanic eruptions provide an instructive counterexample in the present dataset: they do not exhibit a statistically significant duration-volume scaling, despite being analysed using the same LPPLS framework. The scaling observed for mechanically driven systems is therefore an empirical outcome of the analysis, not a mathematical consequence of the method.

We therefore propose to revise the manuscript not by substituting alternative definitions of t_1, but by making the physical meaning of t_1 and the robustness of its inference within the LPPLS framework more explicit. In our view, this addresses the referee's concern more directly. The scientific question of the paper is whether the characteristic lifetime of the positive feedback regime leading to terminal failure exhibits a systematic dependence on system size. Answering that question requires a method capable of identifying the beginning of that regime from the dynamics themselves. The endogenisation of t_1 is therefore not a source of circularity to be removed, but a central feature of the physical approach that we seek to test.

Finally, it is difficult for me to give a formal recommendation to the editor. But to the authors: Convince me that this interesting result is true!
Best regards,

Stefan Hergarten

Reply: We thank the referee for this direct and constructive challenge. We agree that the central task in the revision is to demonstrate convincingly that the observed scaling is a genuine physical result rather than an artefact of the LPPLS-Lagrange procedure.

Our revision will therefore focus on the robustness of the inference within the physical framework being tested. We will clarify that t_1 is not an externally imposed change point, but the endogenously inferred onset of the positive feedback regime leading towards terminal failure, and that determining t_1 from the subsequent dynamics is a central feature of the approach rather than a methodological weakness. We will strengthen the presentation of the onset robustness analysis and assess the scaling relation and its residuals across hazard classes.

Importantly, the LPPLS procedure itself does not impose any relation between duration and event size: the analysis could have produced no scaling, different exponents for different systems, or large scatter. The fact that mechanically driven geohazards exhibit a coherent scaling, whereas volcanic eruptions do not show the same significant dependence despite being analysed with the same framework, is therefore an important internal test against simple model imprinting.

We will also sharpen the distinction between the empirical result and its physical interpretation. Our primary claim is that the characteristic duration of the positive feedback regime scales systematically with event size; the near-linear relation with characteristic system size then provides quantitative support for a finite-size criticality interpretation. We believe that presenting the evidence in this way, together with a clearer demonstration of the robustness of the inferred onset, will provide a substantially firmer basis for judging whether this intriguing scaling is indeed a robust property of mechanically driven instabilities.