Referee report

egusphere-2026-877

Global controls on aftershock productivity: roles of magnitude and tectonic setting

A. D'Alessandro — `https://doi.org/10.5194/egusphere-2026-877`

Referee: Francesco Serafini
Affiliation: Global Earthquake Model (GEM) Foundation
Date: August 6, 2026
Recommendation: **Reject**

Summary of the manuscript

The manuscript compiles $M_w \geq 7.0$ earthquakes from the USGS/NEIC ComCat catalogue for 1976–2024, counts $M_w \geq 4.5$ events within 200 km and 30 days of each, classifies each mainshock as subduction, other plate-boundary or intraplate/crustal using distance to PB2002 boundaries, and regresses $\log_{10}(N+1)$ on M_w separately per class. It reports a near-common slope ($b \approx 0.76$–0.81) with class-dependent intercepts, and argues that depth and distance to plate boundary act as weaker secondary controls.

The underlying question — whether aftershock productivity varies systematically with tectonic setting once magnitude is controlled for — is a good and genuinely useful one. Productivity is important for PSHA and is often poorly represented in operational models, so a clean global answer would be valuable. The choice to adopt a single, explicit, reproducible counting window is also defensible and I appreciated that it is stated plainly.

Unfortunately I cannot recommend the manuscript in anything close to its present form. My reasons fall into four groups: (1) the reference list; (2) a mismatch between what the results show and what the text claims they show; (3) the central claim not being tested; and (4) a definition-of-aftershock problem that is acknowledged but never addressed quantitatively.

Major concerns

M1. The reference list

I attempted to retrieve every entry in the reference list, checking each against Crossref and OpenAlex both by author/title/year and by querying what publication actually occupies the cited journal/volume/page slot. **Nineteen of the 37 entries do not correspond to any publication I can find.** A further four are real papers cited with the wrong journal, title or volume; thirteen in-text citations have no reference-list entry; and six reference-list entries are never cited. The 19 fabricated entries are listed in full at the end of this report.

I want to be direct about the consequence, because it is not a formatting matter. Several of the manuscript's physical interpretations rest entirely on sources that do not exist. The clearest case is *Scholz (2019), "Mechanisms of intraplate earthquakes," Nature Communications 10, 1–10*, which is cited four times (L74, L328, L399, L417) to support statements about intraplate stress transfer, rupture fragmentation and fault-zone rheology. There is no such paper. Similarly, the assertion at L80–81 that stress drop and rupture complexity control productivity is supported by three references (Uchide et al. 2013; Nishimura et al. 2020; van der Elst 2021) of which none exists as cited.

I would ask the author to reconstruct the bibliography from scratch, verifying every entry against a DOI, and to re-examine which statements in the text survive that process. I would also point to the relevant recent literature that is currently missing and that bears directly on the manuscript's central claim:

- Dascher-Cousineau, K., Brodsky, E.E., Lay, T., Goebel, T.H.W. (2020), *What Controls Variations in Aftershock Productivity?*, J. Geophys. Res. Solid Earth 125, e2019JB018111.
- Chu, S.X., Beroza, G.C. (2022), *Aftershock productivity of intermediate-depth earthquakes in Japan*, Geophys. J. Int. 230, 448–463.
- Hainzl, S., Sippl, C., Schurr, B. (2019), *Linear Relationship between Aftershock Productivity and Seismic Coupling in the Andean Subduction Zone*, J. Geophys. Res. Solid Earth 124, 8726–8738.
- Page, M.T., van der Elst, N.J., Hardebeck, J.L., Felzer, K.R., Michael, A.J. (2016), *Three Ingredients for Improved Global Aftershock Forecasts*, Bull. Seismol. Soc. Am. 106, 2290–2301.

These make the claim of a "first unified" global framework (L26–28, L456) untenable as written, and they should be engaged with rather than replaced.

M2. The stated conclusions are stronger than the results support

- **Variance explained.** L370 states that magnitude explains "most of the variance in $\log_{10}(N+1)$" and L454–455 that magnitude and tectonic domain explain "the vast majority of global variance." Table 1 gives $R^2 = 0.2308, 0.2354, 0.2572$, and Fig. 12 gives $r = 0.48$ for M_w versus $\log_{10}(N+1)$. Magnitude explains about a quarter of the variance. Roughly 75% of the variability in productivity is *not* explained by any variable considered here — which is itself an interesting and reportable result, but it is the opposite of what the text says.
- **"Significance" of depth and distance.** L377–380 states that depth and distance become "significant mainly for deeper ($\geq$40–60 km) or intraplate events." No test of either variable is reported anywhere in the manuscript — no p-value, no regression coefficient, no model comparison. Fig. 12 gives $|r| \leq 0.13$ for both. Either add the test or remove the word.
- **Two tests that are true by construction.** Table 2 reports an ANOVA on residuals with $F = 0.000058$, $p = 0.99994$, discussed at L339–341 and L351–352 as evidence that tectonic differences vanish after magnitude normalisation. Because the residuals are taken from *class-specific* OLS fits, each class's residuals have mean exactly zero by construction; the test cannot return anything else and carries no information. The same applies to the $r = 0.86$ correlation between residuals and $\log_{10}(N+1)$ noted at L355–356.
- **Slopes described as "nearly identical".** With $SE(b)$ of 0.087–0.125 (Table 1), the three slopes are statistically indistinguishable. That is a legitimate finding, but it should be stated as "consistent within uncertainty" rather than as a demonstrated universality. I would also suggest renaming the slope: using b for a productivity–magnitude slope invites confusion with the Gutenberg–Richter b-value throughout.

M3. The central claim is never tested

The headline result is that tectonic setting produces systematic offsets in baseline productivity. I could not find anywhere in the manuscript where this is tested. What would be needed is a single model with class indicators — for example

$$\log_{10}(N+1) = a + b\,M_w + \sum_k \gamma_k\,\mathbb{I}[\text{class} = k] + \varepsilon,$$

with confidence intervals on the γ_k, or an equivalent ANCOVA, plus a test of whether allowing class-specific slopes improves the fit. Instead the manuscript fits three separate regressions, observes that the fitted lines are ordered in Fig. 3, and treats the ordering as established.

This matters because Table 1 shows that the intercepts alone do not order the way the text claims: $a_{\text{SUB}} = -4.05$, $a_{\text{BOUND}} = -4.54$, $a_{\text{INTRA}} = -4.35$, so intraplate sits *above* other plate-boundary in intercept. The claimed ordering only emerges once slope differences are folded in, and no uncertainty is propagated. Given $SE(a)$ of 0.64–0.92, I would expect the class separation to be marginal at best.

The ANOVA and Kruskal–Wallis tests at L269–270 do not fill this gap, because they are applied to raw $\log_{10}(N+1)$ with magnitude ignored entirely. A significant result there could simply reflect different magnitude distributions across the three classes.

I would also echo the question I raise at L265: given that the goal is to separate magnitude, depth, distance and class, why fit sequentially on residuals rather than fitting one multiple regression? The two-stage approach discards the covariance structure that the analysis is ostensibly about.

M4. What is being measured is not really aftershock productivity

This is the concern I would most like the author to engage with, because I think the paper could be salvaged along these lines even if the rest could not.

The manuscript defines productivity as the count of $M_w \geq 4.5$ events within 200 km and 30 days. It acknowledges (L154–171) that this is a proxy. But the consequences are not evaluated anywhere, and they are not symmetric across the dataset:

- An M_w 9 rupture is several hundred kilometres long, so a 200 km radius measured from the epicentre truncates a substantial part of the sequence; an M_w 7 sequence fits comfortably inside the window and additionally accumulates background events. If large and small events are not evenly distributed across tectonic classes — and they are not — then the measurement bias is *correlated with the very grouping variable under study.* Subduction hosts the largest events, so subduction productivity is preferentially truncated; this biases against the paper's own conclusion, but it needs to be quantified rather than assumed away.

- There is no comparison against any established declustering or window-selection method. A standard declustering algorithm would assign a meaningfully different set of events to each sequence, and some of what is counted here would be classified as background. A side-by-side comparison for even a subset of sequences would show how much the conclusions depend on this choice.

- There is a genuine identifiability problem that is not mentioned. With a background rate that varies in time, an excess of events after a mainshock can reflect either elevated triggering or elevated background; these two are strongly correlated and cannot be separated by counting alone.

A defensible alternative would be to narrow the claim: state explicitly that the paper characterises *seismicity in the 30 days and 200 km following $M_w \geq 7$ earthquakes*, and drop the framing in terms of aftershock productivity in the broader physical sense, along with the ETAS-parameterisation implications (L29–30, L428–437) that the measurement does not support.

Related, at L96–99 and L154–160 the fixed window is defended as providing a consistent *relative* measure. I would like to see the author state how they expect this to alter the results, or argue why it is not problematic — at present the reader is asked to take it on trust.

M5. Analyses described but not shown

Three separate robustness analyses are described in the Methods and their outcomes asserted, with no numbers, tables or figures anywhere in the manuscript or supplement:

- L143–153: the M_c assessment and the three-period temporal stability test. (Also: the three periods 1976–1990, 1991–2005 and 2006–2024 are described as "partially overlapping" — they are disjoint.)
- L172–183: the full sensitivity analysis over $r = 100, 200, 300$ km and $t = 7, 30, 90$ days, whose conclusion ("the key results of this study remain stable") is stated without a single supporting number.
- L289: binned statistics in distance classes (0–50, 50–150, >150 km) that appear nowhere.

These are the analyses that would answer my concern M4, so I would very much like to see them. As it stands they cannot be assessed.

M6. Circularity in the treatment of distance to plate boundary

The three tectonic classes are *defined* by thresholding distance to the nearest PB2002 boundary at 200 km (L239–241). Distance to plate boundary is then used as an independent secondary predictor (L283–285) and shown in Figs. 5, 7, 8 and 12. This is circular, and the manuscript does not acknowledge it. Fig. 8, described at L343–345 as illustrating "the expected separation among tectonic classes," is a direct consequence of the class definition and cannot be evidence for anything.

If distance is to be retained as a predictor, it should be analysed independently of the classification — e.g. as a continuous covariate within a single model, without the thresholded class variable, or with classes defined by an independent criterion (PB2002 provides boundary-type codes: SUB, OTF, OSR, OCB, CTF, CRB, CCB, which would be a more natural basis than a distance threshold, and would avoid sweeping outer-rise, upper-plate crustal and intraslab events into the "subduction" class).

M7. Text and figures disagree

- **Class ordering.** The abstract, Discussion (L404–406) and Conclusions all state subduction > other plate-boundary > intraplate. The caption of Fig. 9 states "subduction . . . followed by intraplate and other plate-boundary events," and the caption of Fig. 10 states "subduction . . . highest, intraplate intermediate, and other plate-boundary the lowest." These are inconsistent with the headline claim and with each other's framing in the text at L346–348.
- **Distance trend, intraplate.** L325–326 reports "a weak but systematic decrease in productivity at intermediate distances (200–1,200 km)." The intraplate LOWESS curve in Fig. 5 is flat to slightly rising over that range.
- **Residual minimum.** L334–335 places a shallow minimum at "∼200–300 km." In Fig. 7 the intraplate LOWESS minimum sits near 700–900 km.
- **Fig. 7 caption** states "a mild decrease in residuals with increasing distance . . . in the intraplate class"; the plotted curve *increases* beyond roughly 800 km, reaching about +0.3 at 1800 km.
- **Depth range.** L184 states that mainshocks deeper than 100 km were retained and tracked; Figs. 4, 6 and 14 are all truncated at 100 km, so either they were in fact excluded or the figures are incomplete.

M8. Sample size and catalogue construction

- **561 versus 532.** L305 reports a catalogue of 561 $M_w \geq 7$ earthquakes; Table 1 sums to $260 + 160 + 112 = 532$. Presumably the difference is the dependent events excluded by the filtering at L199–203, but this is never stated. Please report both numbers explicitly, with the breakdown per tectonic class, and do so in the Data section rather than in Results.

- **Magnitude homogeneity.** L131–132 claims "uniform moment-magnitude estimates worldwide" from 1976, and L145–147 claims M_c for moment magnitude has been below M_w 4.5 since the mid-1970s. Neither is supported by a citation and I do not think either is correct. NEIC ComCat does not report M_w for the majority of $M \approx 4.5$ events in the 1976–1995 period; many are m_b or M_s. Please state exactly which magnitude types were accepted, how mixed types were handled, and support the M_c claim with an actual completeness analysis on the catalogue used.

- **Total count.** A global $M \geq 7.0$ count for 1976–2024 is appreciably larger than 561. Please clarify what was excluded (depth cut? magnitude-type filter?) and report it in the Data section.

M9. Structure and repetition

The operational definition of productivity and its caveats are restated four times in near-identical terms (L92–110, L154–171, L168–171, L186–191), and several conclusions appear in the Methods (L150–153, L178–183) before any result has been presented. Sections 3 and 5 restate the same three findings twice, expanded to five in the Conclusions. Consolidating this would shorten the manuscript considerably and make the actual contribution easier to see.

I also note that the "Highlights" section (L39–47) is an Elsevier convention and is not part of the EGU manuscript structure.

Line-by-line comments

Abstract and Highlights

- **L12** — "all $M_w \geq 7.0$ mainshocks worldwide since 1976" — it would be useful to state how many in total, to give the reader a sense of how robust the results are right away.
- **L18–19** — "nearly identical slopes ($b \approx 0.76$–0.81)" — see M2. Given the standard errors these are indistinguishable rather than "nearly identical"; and b collides with the Gutenberg–Richter b-value.
- **L26–28** — "the first unified, physically interpretable global framework" — see M1; this claim is not tenable given the existing literature.
- **L39–47** — The Highlights section is an Elsevier convention, not EGU.

1. Introduction

- **L77** — "globally homogeneous aftershock datasets" — the difficulty is more fundamental than data homogeneity: the definition of an aftershock is itself ambiguous, and there is no clear observable that tells us whether a given earthquake is an aftershock or not. This should be said explicitly here.
- **L82** — "Yet these effects are often difficult to isolate due to catalog incompleteness, regional variability, and the strong primary control of magnitude itself" — in some cases parameter identifiability (and/or correlation among parameters) is also an issue. With a background rate that can vary in time, is a surge in earthquakes due to above-average aftershock productivity or to increased background seismicity? These two are clearly correlated. Please add this.
- **L87** — "$M_w \geq 7$ mainshocks and their aftershocks ($M_w \geq 4.5$)" — again, how many in total? How many per tectonic class?
- **L91** — "using distances to PB2002 plate boundaries" — have you tried a different methodology to classify? See M6: PB2002's own boundary-type codes would be a more natural basis than a distance threshold.
- **L92** — "defined in an operational and observation-based sense" — it is not clear what is meant by this. Are you referring to a known definition? Is there a reference? If not, please reformulate to explain it more directly.
- **L100** — "While this approach may truncate late or far-field aftershocks of the largest earthquakes. . . " — it may also discard clusters of events occurring at certain distances, especially for the large events. I would like to see a comparison with another method here, to show the difference and the effects of this choice. Alternatively, limit the conclusions to the first month of activity after an M 7 rather than talking about aftershocks in the broader sense. See M4.
- **L103–105** — "the analysis focuses on comparative productivity trends and regression-based offsets among tectonic settings" — so essentially this is about the slope of the frequency–magnitude relation for different tectonic settings, for which we already know there are more homogeneous trends. Productivity, by contrast, is very important for PSHA and is often misrepresented by models (e.g. ESHM20, UCERF3). It would strengthen the paper considerably to make that connection explicit.
- **L105** — "productivity is quantified as the number of earthquakes with $M_w \geq 4.5$ occurring within a fixed spatiotemporal window" — I like having a clear definition. However, what is

being described here is seismic activity in the month after an M 7, rather than aftershocks as such. Any declustering method would identify some of these as background.

- **L109** — "both absolute aftershock counts and the logarithmic scaling $\log_{10}(N+1)$" — why? Please justify the choice rather than only asserting it (the justification at L219–227 arrives much later).

2. Data and Methods

- **L129** — "USGS/NEIC ComCat database" — reference needed.
- **L130** — "between 1 January 1976 and 31 December 2024" — as above, please give the number of earthquakes.
- **L131** — "provides uniform moment-magnitude estimates worldwide" — a time versus magnitude plot would help in visualising this. See also M8: I do not think this claim is correct as stated.
- **L134** — "lies safely above the global magnitude of completeness ... throughout the study period" — reference?
- **L134** — "Previous assessments of global catalogs indicate... " — which ones? Please cite them.
- **L140** — "These thresholds reflect a balance between catalog completeness, consistency with prior global analyses, and robust detectability" — again, a standard declustering algorithm would probably identify different aftershocks. Has any comparison been done prior to the analysis?
- **L144** — "Previous global analyses... " — again, which ones?
- **L145–147** — "M_c for moment magnitude has been consistently below M_w 4.5 since the mid-1970s" — this repeats what is said a few lines above, and is unsupported. See M8.
- **L149** — "three partially overlapping time periods (1976–1990, 1991–2005, and 2006–2024)" — these periods are disjoint, not overlapping. Also, the number of mainshocks and aftershocks per period would be informative.
- **L150–153** — "Across all time windows, the M_w–aftershock productivity scaling ... remain stable" — this reads as an anticipation of the results, and no supporting numbers are given anywhere. See M5.
- **L160** — "because the same spatiotemporal criteria are applied uniformly to all events, this approach provides a consistent, reproducible measure of *relative* aftershock productivity" — I like this, but I would also like to know how the author expects it to alter the results, or why it is not problematic.
- **L163** — "comparative productivity scaling and systematic offsets among tectonic domains under a common observational framework" — this productivity is itself altered by the methodology. Assuming large earthquakes cluster in certain tectonic regions and smaller ones in others, then in regions of large events all sequences are truncated, and in regions of small events all are augmented. See M4.
- **L165–168** — "potential long-term improvements in network sensitivity primarily affect background detection levels ... only second-order noise" — this assumes the variation comes only from network capability and not from the process changing with time.

- **L172–177** — The sensitivity analysis over alternative radii and time windows is exactly the right thing to do — but the results are not shown anywhere. Please add them. See M5.
- **L178–183** — "The sensitivity tests show that while absolute productivity values vary. . . " — this may belong in the Results section rather than here; the conclusions have already been repeated quite a bit at this point.
- **L184** — "Mainshocks occurring deeper than 100 km were retained" — why 100 km? And note that Figs. 4, 6 and 14 are truncated at 100 km, so it is unclear whether these events were in fact retained.
- **L185–186** — "Aftershocks were not declustered, as the objective of this study is to characterize the full cascading productivity" — this should have come earlier, while discussing the definition of aftershocks. It also directly contradicts L93–95, which states that productivity is *not* defined as the total number of triggered events in a fully cascading ETAS sense.
- **L186–191** — "the term aftershock productivity is used in an operational and observational sense. . . " — repeated from earlier.
- **L197–199** — "However, for all analyses that assume independent triggering. . . " — I find this sentence confusing rather than explanatory; what follows is much better.
- **L201–203** — "These events are retained in the catalogue for descriptive analyses, but are excluded from all regression, correlation, and hypothesis-testing procedures" — this is not clear. Suppose I have two mainshocks at t_1 and t_2: are the aftershocks of the first only the events between t_1 and t_2, with the second taking all events after t_2? Or are both excluded? Please give the rule explicitly.
- **L209–213** — The definition of N has already been stated a number of times by this point.
- **L239–241** — The tectonic classes are defined by thresholding distance to plate boundary, which is then used as an independent predictor. See M6.
- **L255–258** — "To further verify temporal robustness, we repeated the magnitude–productivity analysis for different sub-periods. . . " — repeated again, and again without supporting numbers.
- **L259–261** — "In this sense, our results provide global empirical constraints . . . that can inform, but do not replace, the parameterization of physics-based or ETAS-type seismicity models" — also repeated; and I do not see why this needs specifying here.
- **L265–266** — "[residuals] were then used to examine secondary controls from depth and distance to plate boundaries after removing the primary M_w dependence" — why not go to multiple regression directly? See M3.
- **L269–270** — One-way ANOVA and Kruskal–Wallis — references needed for both. Note also that applying them to raw $\log_{10}(N + 1)$ ignores magnitude entirely (M3).
- **L272** — "Given the limited sample size. . . " — which we still do not know at this point in the manuscript.
- **L273–275** — "p-values are interpreted cautiously and in conjunction with effect sizes, confidence intervals, and physical consistency" — this also seems more appropriate for a Results or Discussion section. Note too that no confidence intervals or effect sizes are actually reported.
- **L279** — Missing space: "both to $\log_{10}(N + 1)$and to regression residuals".

- **L289** — Binned statistics in distance classes (0–50, 50–150, >150 km) are described but never shown. Note also that these bins are inconsistent with the 200 km class threshold.
- **L290** — "ε residuals" appears as "e residuals".

3. Results

- **L305** — "The global catalogue of 561 $M_w \geq 7$ earthquakes" — finally the number of events! This should really go in the Data section. Note also that Table 1 sums to 532, not 561 (M8).
- **L312–313** — "Scatterplots with OLS regressions (Figure 2) show robust positive slopes in each tectonic class, with limited scatter at the highest magnitudes ($M_w \geq 8.5$)" — I wonder what would happen to these regression lines if the top 1% or 5% in magnitude were excluded, as they are undersampled. Relatedly, the "limited scatter" at high magnitude is a consequence of small n, not of genuinely lower variability.
- **L316** — "Table 1 reports regression coefficients and R^2 values" — which are very low. This should be stated and discussed openly rather than passed over. See M2.
- **L322** — "Depth explains only modest variance in productivity relative to magnitude" — why not also show a LOWESS trend for magnitude? Then we would be comparing like with like, and could compute R^2 and the variance explained by each variable.
- **L323–325** — "Subduction and other boundary events cluster tightly within <150 km . . . while intraplate events span distances >1,000 km" — looking at this, I would really consider using the log of distance, to reduce the difference in domain.
- **L325–326** — "Only the intraplate class shows a weak but systematic decrease in productivity at intermediate distances (200–1,200 km)" — the LOWESS curve in Fig. 5 is flat to slightly rising over that range. See M7.
- **L330–332** — "Residuals vs depth (Figure 6). . . " — Fig. 6 looks very similar to Fig. 4; it seems that considering the residuals did not make a lot of difference. Please comment on this.
- **L333–335** — "Residuals vs distance to plate boundaries (Figure 7)" — I am curious to see the same results on a log scale. Also, L334–335 places the intraplate minimum at ∼200–300 km, whereas in Fig. 7 the minimum is near 700–900 km (M7).
- **L337–341** — The ANOVA on residuals (Table 2, $F = 0.000058$, $p = 0.99994$) is zero by construction and carries no information. See M2.
- **L343–345** — "Violin plots of distance to PB2002 boundaries (Figure 8) illustrate the expected separation among tectonic classes" — this belongs in the Data section, and it is a direct consequence of how the classes were defined (M6).
- **L346–348** — "Raw aftershock counts (Figure 9) differ substantially across classes. . . " — this also belongs in the Data section, along with the discussion of $\log(N + 1)$. Note also that the ordering in the captions of Figs. 9 and 10 (intraplate above other plate-boundary) contradicts the ordering claimed in the abstract and Discussion (M7).
- **L351–352** — "differences among tectonic classes are strongly reduced, consistent with statistical tests (Tables 2–3)" — they were not very different to start with. The variances in Fig. 11 do look quite different, though, and that is not commented on.
- **L353–356** — What is `depth_main_km`? It is not defined anywhere in the caption or text. Please label figure variables using the notation defined in the text. Note also that the $r = 0.86$ correlation between residuals and $\log_{10}(N + 1)$ is an artefact of construction (M2).

- **L361–363** — "shallow events (0–40 km) are slightly more productive, whereas deep events (>60 km) produce fewer aftershocks" — true, but the greater depths are severely undersampled; we essentially have good coverage of about a quarter of the domain.
- **L365–367** — "the robustness of the results is established through the convergence of multiple lines of evidence" — this reads more like a Methods statement than a result.
- **L377–380** — "becoming significant mainly for deeper (≥40–60 km) or intraplate events" — this may simply be undersampling: looking at Fig. 14, for depth > 60 km the data are sparse and all below M 8, so combined with the magnitude scaling, depth becomes relevant to predicting productivity essentially *through* magnitude. Note also that no significance test for depth is reported anywhere (M2).

4. Discussion

- **L389** — "provide a unified physical framework" — I would not call this physical, as it was found through statistical modelling; and I would not call it a framework, as this is simply an analysis. Please moderate.
- **L392–396** — The slope range is compared with rate-and-state predictions citing Helmstetter (2003), Nandan et al. (2017) and Page et al. (2016); Nandan et al. is not in the reference list, and Page et al. is cited with fabricated bibliographic details.
- **L399–400** — "This invariance also explains why ETAS productivity parameters tend to cluster globally" — reference? And note that Kanamori & Anderson (1975) is not in the reference list, while Scholz (2019) does not exist.
- **L407–412** — The physical mechanisms attributed to subduction megathrusts cite Lay & Kanamori (1981), which is absent from the reference list, and Hayes et al. (2012), which is Slab1.0 — a subduction *geometry* model, which does not support statements about pore-fluid pressure or fault-zone width.
- **L416–417** — "This reconciles contrasting findings in earlier studies . . . (e.g., Das & Henry 2003; Scholz 2019) or magnitude-controlled universality (e.g., Felzer et al., 2004)" — none of these three appear in the reference list under those years, and Scholz (2019) does not exist.
- **L419–423** — "Depth-dependent productivity decay . . . corroborates earlier findings" citing Sibson (1982) and Houston (2015) — neither is in the reference list.
- **L428–437** — "These results have direct implications for global hazard modeling. . . " — already repeated in the Conclusions a few lines below. I would also note that items 3 and 4 in this list assert depth-dependent and distance-dependent effects that were never tested.

5. Conclusions

- **L454–455** — "M_w and tectonic domain explain the vast majority of global variance" — contradicted by Table 1 ($R^2 \approx 0.23$–0.26). See M2.
- **L456–459** — "These findings establish a coherent, physically interpretable global framework" — see the comment at L389.

Data statement, acknowledgements, references

- **L465–476** — The data statement gives only a generic ComCat API endpoint and a plate-boundary URL. EGU journals require an assets availability statement with a persistent identifier. Please deposit the derived catalogue (mainshock list with class assignment, counts, depth, d_{PB}) and the analysis code in a repository with a DOI. Given the discrepancies noted above, this is necessary for the results to be checkable at all. Note also that "last accessed in 2025" should be a precise date.

- **L473–476** — "custom Python scripts developed by the authors" and L479 "The authors declare no competing interests" — the manuscript has a single author.
- **L489–594** — See the list of fabricated references at the end of this report. Nineteen entries do not exist, four more are incorrect, thirteen in-text citations are unmatched, and six entries are never cited.
- **L493** — Craig et al. (2017) — this does not seem to exist. *Confirmed*: GJI volume 208 ends at page 1592, so pages 1642–1664 do not exist. The likely intended work is Craig, Copley & Middleton (2014), EPSL 392, 94–99.
- **L496** — Das & Scholz (1981), "Theory of large intraplate earthquakes" — I could not find this either. Do you mean Scholz (1982), "Scaling laws for large earthquakes: consequences for physical models"? *Confirmed*: the real Das & Scholz (1981) is "Theory of time-dependent rupture in the Earth," JGR 86, 6039–6051.
- **L501** — Felzer et al. (2002) — does not exist. The real 2002 paper by that group is "Triggering of the 1999 M_w 7.1 Hector Mine earthquake...", JGR 107(B9).
- **L507** — Hainzl et al. (2016) — does not exist. Those exact pages (JGR 121, 2575–2590) belong to Hainzl, Fischer, Čermáková, Bachura & Vlček (2016) on fluid-triggered aftershocks.

List of fabricated references

The following 19 entries in the reference list do not correspond to any publication traceable in Crossref or OpenAlex. Each was checked both by author/title/year and by querying what publication actually occupies the cited journal/volume/page slot; where that slot is occupied by a different paper, or where a plausibly intended real work exists, this is noted.

1. **Craig, T.J., Copley, A.C., Middleton, T., 2017.** Constraining fault friction in intraplate earthquakes. *GJI* 208, 1642–1664.
   GJI vol. 208 ends at p. 1592; those pages do not exist. Closest real work: Craig, Copley & Middleton (2014), "Constraining fault friction in oceanic lithosphere...", EPSL 392, 94–99.

2. **Das, S., Scholz, C.H., 1981.** Theory of large intraplate earthquakes. *Science* 213, 1153–1160.
   Slot unoccupied. The real Das & Scholz (1981) is "Theory of time-dependent rupture in the Earth," JGR 86, 6039–6051.

3. **Felzer, K.R., Becker, T.W., Abercrombie, R.E., 2002.** Correlation between aftershock productivity and earthquake stress drops. *GRL* 29, 2286.
   No such paper. The 2002 paper by this group is "Triggering of the 1999 M_w 7.1 Hector Mine earthquake...", JGR 107(B9).

4. **Hainzl, S., Moradpour, J., Davidsen, J., 2016.** Aftershock generation as dependent on mainshock rupture properties. *JGR* 121, 2575–2590.
   Those exact pages belong to Hainzl, Fischer, Čermáková, Bachura & Vlček (2016), "Aftershocks triggered by fluid intrusion..." — a fabricated title and author list on a real page range.

5. **Hayes, G.P., et al., 2020.** The global development of NEIC moment tensor solutions. *SRL* 91, 2230–2243.
   Slot unoccupied; no such paper. Never cited in the text.

6. **Lima, R.G., Courboulex, F., Cotton, F., 2021.** Aftershock productivity across global regions. *BSSA* 111, 1531–1547.
   Slot unoccupied; no such author combination or paper. Never cited in the text.

7. **Marzocchi, W., Lombardi, A.M., 2009.** Real-time aftershock forecasting using ETAS. *BSSA* 99, 172–189.
   BSSA 99:172–193 is Bondár & McLaughlin on seismic location bias. The real paper is "Real-time forecasting following a damaging earthquake," GRL 36, L21302.

8. **Marsan, D., Lengliné, O., 2008.** Extending aftershock triggering models with near-field interactions. *JGR* 113, B09304.
   No such paper. The real Marsan & Lengliné (2008) is "Extending Earthquakes' Reach Through Cascading," Science 319, 1076–1079.

9. **Nishimura, T., Uchide, T., Sekiguchi, H., 2020.** Rupture complexity and aftershock productivity in Japan. *EPS* 72, 1–15.
   No such paper; EPS 72 pp. 1–10 is a geomagnetics article. Cited at L63 and L80.

10. **Omi, T., Ogata, Y., Hirata, Y., Aihara, K., 2013.** Forecasting large aftershocks with ETAS models. *JGR* 118, 3642–3653.
    That slot is Fang & Dunham on fault roughness. The real paper by this team is "Forecasting large aftershocks within one day after the main shock," Sci. Rep. 3, 2218.

11. **Page, M.T., van der Elst, N.J., Hardebeck, J.L., Felzer, K.R., 2016.** Optimal forecasting of aftershock activity. *JGR* 121, 5088–5108.
Slot unoccupied. The real paper is Page, van der Elst, Hardebeck, Felzer & Michael (2016), "Three Ingredients for Improved Global Aftershock Forecasts," BSSA 106, 2290–2301 — note the additional author.

12. **Reasenberg, P.A., Jones, L.M., 1994.** Earthquake aftershock occurrence: Relations to mainshock magnitude. *JGR* 99, 6527–6540.
That slot is a paper on the ionosphere of Titan. Reasenberg & Jones (1994) is a two-page Science update (265, 1251–1252); the model paper is Reasenberg & Jones (1989), Science 243, 1173–1176.

13. **Scholz, C.H., 2019.** Mechanisms of intraplate earthquakes. *Nature Communications* 10, 1–10.
No such paper exists. Scholz's 2019 output is the 3rd edition of "The Mechanics of Earthquakes and Faulting." ***Cited four times*** *— L74, L328, L399, L417 — in support of the manuscript's physical interpretation.*

14. **Shcherbakov, R., Turcotte, D.L., Rundle, J.B., 2004.** A generalized Omori law. *PRL* 93, 178501.
PRL 93, 178501 is Marsan et al. on Arctic sea-ice deformation. The real paper is "A generalized Omori's law for earthquake aftershock decay," GRL 31, L11613.

15. **Stein, S., Okal, E.A., 2007.** Ultralong aftershock sequences of great earthquakes. *Science* 316, 1737.
Slot unoccupied. Stein & Okal (2007) is in BSSA 97, S279–S295 (an ultralong-period *study). The concept belongs to Stein & Liu (2009), Nature 462, 87–89. Never cited in the text.*

16. **Uchide, T., Imanishi, K., Nakai, M., et al., 2013.** Aftershock productivity and mainshock complexity in Japan. *EPS* 65, 1607–1612.
Slot unoccupied; no such paper. Cited at L80.

17. **van der Elst, N.J., 2021.** Aftershock productivity linked to stress drop variability. *JGR* 126, e2020JB021145.
van der Elst's 2021 JGR paper is e2020JB021027, "B-Positive: A Robust Estimator of Aftershock Magnitude Distribution. . . ". Cited at L65–66 and L81.

18. **van der Elst, N.J., Page, M.T., 2014.** Aftershock productivity and background seismicity rate. *GRL* 41, 8460–8467.
That GRL slot is an oceanography paper on mineral ballast and particulate organic carbon export. Cited at L54 and L81–82.

19. **Wilcock, W.S.D., Tolstoy, M., Waldhauser, F., et al., 2009.** Seismicity of mid-ocean ridges. *Nature Geoscience* 2, 646–652.
Slot unoccupied; no such paper. Never cited in the text.

Verification performed against the Crossref REST API (`https://api.crossref.org`) and OpenAlex (`https://api.openalex.org`), querying both by author/title/year and by journal ISSN + volume + first page. Queries and raw outputs available on request.