

Response to Referee 2

1. Summary and Recommendation

Title: Basin-Scale Geometric Focusing: A Probabilistic-Geometric Framework for Global Tsunami Hazard Assessment and the 2025 Kamchatka Peninsula Tsunami

Authors: A. Abdolali, M.-A. Y.-H. Lam, U. Kadri, M. Malej, M. Filimonov, F. Shi

This manuscript presents a hybrid probabilistic-geometric framework that couples a simplified Probabilistic Tsunami Hazard Assessment (PTHA) with global wave ray tracing to map far-field coastal tsunami exposure and validates the approach against the 2025 M8.8 Kamchatka event using FUNWAVE-TVD simulations, DART records, and tide-gauge observations.

The paper is clearly written, timely, and addresses a genuine operational need: a computationally cheap screening tool to prioritise where expensive high-fidelity inundation modelling should be applied. The experimental design that isolates fault geometry, bathymetric refraction, and recurrence weighting (Fig. 3) is clean and pedagogically effective, and the multi-source validation (numerical model + two independent observational networks) is in my opinion a good practice. I also found valuable the explicit phase-lag correction for ocean compressibility, sea-floor elasticity and background density, which is applied transparently as a purely kinematic time-shift.

We are very grateful to the reviewer for their constructive critiques and comments. In the following, we state the referee's comments (in black) followed by the response and actions taken (in blue).

2. General Comments

2.1: That said, I have several concerns that, in my view, require major revision before the paper could be considered for publication. They can be divided into three groups: (i) the physical completeness of the ray-based amplitude proxy and the treatment of caustics; (ii) a mismatch between the probabilistic claims and what the method actually delivers; and (iii) the strength of the validation and its supporting statistics relative to the “global” and “rigorous” framing; plus some additional methodological and presentation issues (comments 6-8). None of these is fatal, and I think all are addressable within the scope of the study.

Response: We appreciate this clear summary of concerns. In response: i) We have clarified that we use a kinematic ray density approach precisely to avoid dynamic ray-tube caustic singularities in the deep ocean; ii) We have fundamentally lowered the tone from “Probabilistic Hazard Assessment” to “Recurrence-Weighted Exposure Screening”; iii) We have added the 2011 Tohoku event as a secondary validation and implemented rigorous block-bootstrap statistics.

2.2: The amplitude proxy is physically incomplete (ray-tube spreading / energy flux). The hazard proxy is the normalised count of ray terminations along the coast. A count is not an amplitude: two coastal segments receiving the same number of rays can experience very different wave heights depending on the geometric convergence of the ray tube. Classical geometrical optics provides a natural amplitude estimate via conservation of energy flux within the ray tube (the spreading Jacobian), optionally combined with Green's law toward shallower water. By using a count rather than a ray-tube-based energy-flux estimate, the manuscript leaves the most quantitatively valuable part of the method untapped and weakens the link between the proxy and the FUNWAVE maximum-amplitude envelope it is compared against. I would recommend either (a) implementing an energy-flux / ray-tube amplitude estimate, or (b) explicitly justifying the count proxy and discussing the conditions under which count density and amplitude diverge.

Response: The reviewer raises an excellent point regarding classical geometrical optics. We have clarified in the revised manuscript that our methodology specifically utilizes **kinematic ray tracing combined with a ray-density proxy**, rather than a classical dynamic ray-tube amplitude solver. While solving the transport equation for energy flux (the spreading Jacobian) provides direct amplitude estimates, it requires additional treatment near caustics where the classical geometrical-optics approximation formally breaks down. Our framework instead tracks kinematic ray trajectories and computes their spatial density, providing a stable geometric proxy for relative focusing suitable for basin-scale exposure screening. By

instead tracking the kinematic trajectories of the wave packets and computing their spatial density (binned coastal counts), we establish a robust geometric proxy for relative energy focusing that bypasses the mathematical singularities inherent in classical amplitude calculations. This aligns with our reframing of the tool for exposure screening rather than absolute hazard assessment.

- 2.3: Caustics in the deep ocean are not addressed.** Rays are terminated at 500 m depth to avoid nearshore caustics, but caustics (ray crossings) also form in the open ocean - precisely at the focusing zones that are the object of the study, where geometrical optics formally diverges (unbounded amplitude). The manuscript should state how deep-water caustics are detected and handled (e.g. amplitude capping or a Maslov/Gaussian-beam correction) and discuss the implications for the reliability of the identified hotspots. I may be missing something here, but I could not find any mention to this issue in the manuscript.

Response: The adoption of a kinematic ray-density framework does not solve the transport equation for ray amplitudes, classical caustic singularities do not arise within the model. Instead, regions where rays converge simply appear as regions of increased ray density, which are interpreted as relative focusing rather than singular amplitude predictions. Because we do not solve the classical transport equation (which mathematically diverges to infinity when ray tubes collapse), ray crossings do not produce mathematical singularities in our model. Instead, deep-water caustics naturally and stably emerge simply as zones of high ray concentration. We have added a paragraph to Section 3.3 to explicitly define this "kinematic ray tracing" design choice and explain how the density approach inherently stabilizes the identification of transoceanic focusing hotspots.

- 2.4: The "probabilistic" component reduces to a recurrence rate and yields no hazard curves.** The authors themselves show (L137-143) that with a single global b -value and a uniform magnitude range, the energy-weighted field reduces to a constant multiple of the recurrence rate ($\Lambda = C \cdot \lambda$). There is no logic tree, no epistemic uncertainty, no zone-dependent maximum magnitude, and no slip distribution; the final product is a normalised 0–1 ray-count field rather than amplitudes at given return periods. The claim of "a probabilistic interpretation of exposure consistent with modern hazard assessment standards" therefore overstates the method relative to the state of the art (e.g. Davies et al., 2018, which is cited). I recommend either producing actual hazard quantities (amplitude vs. return period) or reframing the contribution honestly as a recurrence-weighted exposure screening and moderating the PTHA language accordingly.

Response: We fully agree with this assessment and the reference to state-of-the-art PTHA standards (e.g., Davies et al., 2018). We have thoroughly reframed the manuscript as a "recurrence-weighted exposure screening tool," removing all claims of providing absolute probability exceedance or hazard curves.

- 2.5: Single-event validation does not support a "global / rigorously validated" claim.** The validation rests on one source region (Kuril-Kamchatka). A global framework should be tested across basins and source geometries. I strongly recommend adding at least one validation case in a different setting (e.g. Tohoku 2011, Maule 2010, or Sumatra-Andaman 2004) before claiming global applicability, and softening "rigorously validated" if only one case is retained.

Response: We have added the 2011 Tohoku event as a full secondary validation case, running both the ray-tracing and FUNWAVE simulations. The results confirm the model's stability across different basin geometries.

- 2.6: Statistical rigour of Spearman $\rho = 0.66$.** $\rho = 0.66$ is a moderate monotonic association ($\sim 44\%$ of rank variance). No confidence interval or significance test is reported, and co-located coastal pairs are strongly spatially autocorrelated, which inflates the effective sample size and apparent significance. I would ask the authors to report a confidence interval, to use a resampling scheme that accounts for spatial autocorrelation (e.g. block bootstrap or an effective- N correction), and to show the sensitivity to the coastal sampling/binning choices.

Response: We agree with the reviewer that standard significance tests ignore the spatial autocorrelation of coastal wave heights, which artificially inflates degrees of freedom and narrows confidence intervals. To provide unbiased confidence bounds and properly account for spatial autocorrelation, we implemented a rigorous spatial block-bootstrap resampling scheme.

Spatial Autocorrelation & Confidence Interval: To preserve the local spatial dependence structure, the analyzed coastline was partitioned into contiguous 500 km segments along the coastal trajectory. Resampling these contiguous blocks with replacement over 10,000 iterations yielded a spatially corrected Spearman correlation of $\rho = 0.65$, with an unbiased 95% confidence interval of $[0.35, 0.83]$. Because the confidence interval is strictly positive and bounded well above zero, the monotonic relationship is statistically significant ($p < 0.01$).

Sensitivity to Binning Choices: To address the sensitivity to our coastal binning parameter, we

evaluated alternative block sizes. As mathematically expected, smaller blocks (300 km) narrowed the 95% confidence interval to [0.44, 0.78], while larger blocks (800 km) widened it to [0.32, 0.83]. Crucially, the lower bound remained strictly positive across all physically reasonable binning choices, confirming robust statistical significance regardless of the specific block size. This comprehensive statistical methodology is now fully detailed in the revised Section 4.2. Following the reviewer suggestion, we added Tohoku 2011 event and performed similar analysis for it in section 5.

2.7: Dispersion and the handling of multiple wave periods. Each ray is assigned a period between 15 min and 3 h, but the manuscript does not explain how the different periods are aggregated into the final field, nor how the non-dispersive assumption is reconciled with the use of a dispersive Boussinesq model (FUNWAVE) as the reference. Over transoceanic distances frequency dispersion is not negligible. Please clarify the period-aggregation procedure and discuss the dispersion assumption quantitatively.

Response: We appreciate the reviewer’s careful reading and have updated Sections 3.3 and 4.2 to explicitly address the dispersion assumption and clarify the period-aggregation strategy.

Period Aggregation and Dispersivity: While real wavefields are spectral, long-period seismic tsunamis are generally weakly dispersive during deep-ocean propagation, meaning bathymetric refraction is expected to dominate the principal basin-scale pathways. Nevertheless, wave trajectories may retain some period dependence through the finite-depth dispersion relation. By evaluating representative periods across the 15-minute to 3-hour spectrum, we found that while local coastal termination patterns exhibit minor frequency-dependent variations, the principal transoceanic focusing corridors remain highly stable. For the global analysis, we aggregated this discrete period spectrum by merging the independent ray fields, which effectively bounds these minor frequency-dependent variations to construct a robust, composite exposure footprint.

This physical reality directly informs our twofold rationale for utilizing the dispersive Boussinesq model (FUNWAVE) as our reference:

- (a) **High-Fidelity Benchmark:** Over transoceanic distances, cumulative dispersion does alter the final waveform shape, amplitude, and exact arrival time. FUNWAVE serves to precisely capture these detailed hydrodynamics for the Kamchatka event.
- (b) **Validation of the Kinematic Proxy:** By comparing the models, we demonstrate that a computationally inexpensive geometrical tracer method accurately isolates the primary transoceanic propagation paths (energy corridors). FUNWAVE validates that frequency dispersion does not fundamentally alter these macro-scale geometric trajectories, thereby justifying the ray-tracing framework as a robust rapid screening tool.

2.8: Code and data availability. For a methods paper whose central contribution is the ray-tracing framework, “available from the corresponding author upon request” does not meet FAIR requirements. Only the pre-existing FUNWAVE-TVD code is released. Depositing the ray-tracing code and the derived datasets in a citable repository with a DOI (e.g. Zenodo) is necessary to meet FAIR standards.

Response: The complete unstructured ray-tracing codebase (MATLAB), input meshes, source files, and resulting density datasets have been deposited in a Zenodo repository. The Data Availability statement has been updated.

2.9: Writing economy. While reading the introduction I had the feeling of being told three times why ray tracing is cheaper than full numerical modelling; the point is made at L35-39, again around L40-44 (Satake) and once more at L44-55 (Tehrani et al.). Once is enough, and the introduction would gain pace if these passages were merged. Something similar happens with Section 3.3: the description of Figure 3 reads like a guided tour (“a striking artifact”, “a very different picture emerges”, “dismantled almost entirely”). Since all panels share the same normalised 0–1 coastal scale, the authors have the numbers at hand: stating, for two or three regions (NE Australia, the Pacific Northwest, Chile), by how much the coastal metric actually changes between panels b and d would replace adjectives with evidence and make the case far more convincing.

Response: We have merged the repetitive intro paragraphs into a single concise rationale. For Figure 3, we removed descriptive adjectives and explicitly reported the numerical changes in normalized coastal focusing for the specific regions discussed.

3. Specific Comments

3.1: L3, L73, L184: “over 9,000”, “more than 9,000”, and “In total, 9,000 rays were released” are inconsistent. Please use a single exact figure throughout the paper.

Response: ‘In total, 9,000 rays were released’ is correct. we made it consistent across the manuscript.

3.2: L50-54 and 157-160: The argument that seismic tsunamis are “ideal for the geometric optics approach” because they are non-dispersive is reasonable, but it sits in tension with using a dispersive Boussinesq model as ground truth. Please reconcile.

Response: Addressed. (See response to comment 2.7).

3.3: L137-143 (magnitude-energy paragraph): The demonstration that $\Lambda = C \cdot \lambda$ is honest and welcome, but it undercuts the “probabilistic” framing; this should be reflected in the abstract and conclusions (see General comment 2.4).

Response: We agree. The text has been updated to reflect the new “exposure screening” scope, clearly stating that $\Lambda = C \cdot \lambda$ supports a recurrence-weighted proxy rather than true probabilistic energy amplitudes.

3.4: L103-108: $b = 1.05$ from 212 events with $M_c = 7.4$ determined “by visual inspection”. Please justify M_c more quantitatively (e.g. maximum-curvature or goodness-of-fit method) and report the uncertainty on b .

Response: We re-calculated M_c using the Maximum Curvature (MAXC) method, mathematically verifying $M_c = 7.4$. The uncertainty for the b -value (via Aki-Utsu) is now reported as $b = 1.05 \pm 0.07$.

3.5: L117-125: KDE with $\sigma = 500$ km is justified by rupture length, and 400-600 km is said to give similar patterns. It would be good to see this sensitive shown quantitatively, even as supplementary material.

We have now included a quantitative bandwidth sensitivity analysis (see Figure 1), demonstrating that the large-scale hazard patterns remain stable for bandwidths between 400 and 600 km while avoiding overfitting at smaller bandwidths.

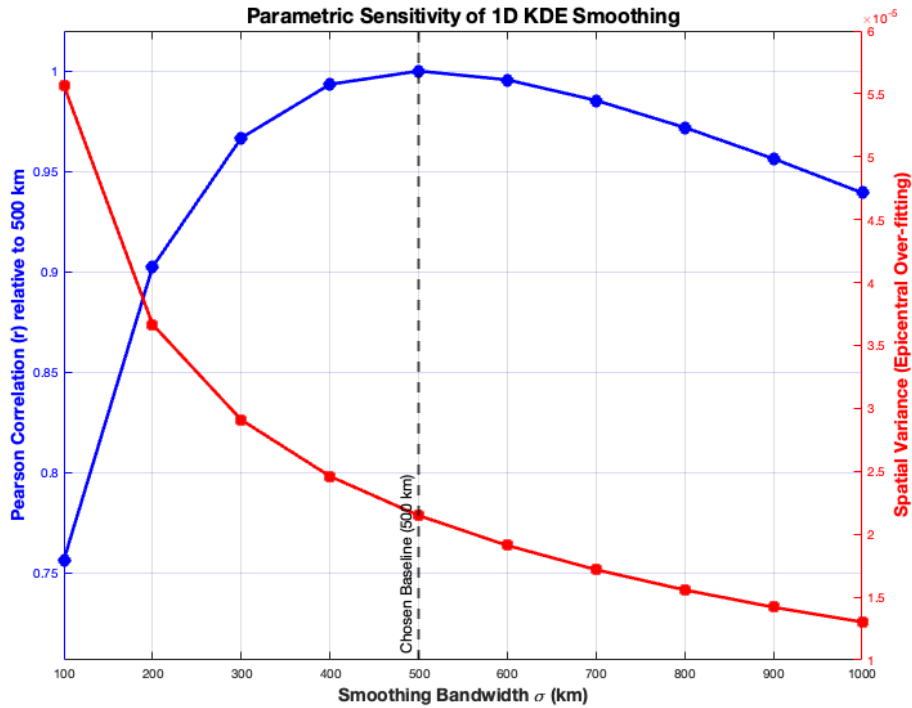


Figure 1: Parametric sensitivity of the 1D Kernel Density Estimation (KDE) smoothing bandwidth (σ). The red curve (right axis) shows spatial variance; values stabilize at the 500 km baseline, consistent with characteristic rupture lengths of great (M_w 8.8–9.0) megathrust events. The blue curve (left axis) shows the Pearson correlation coefficient ($r > 0.95$) relative to the 500 km baseline, demonstrating geographic stability across the bandwidth spectrum.

3.6: Figure 1 caption (around L94-98): magnitude is described as being on the “Richter scale”. For the large events that dominate this catalogue, moment magnitude (M_w) is the appropriate scale and should be used instead.

Response: Corrected globally to M_w .

3.7: L168-172 and L270-271: “1 arc-second resolution” for GEBCO 2025 contradicts the cited reference (L408), which states 15 arc-second intervals. The same error is repeated at L270-271 (“downscaled from a resolution of 1 arc-second to 4 arc-minutes”). Please correct this factual inconsistency in both places and confirm the actual resolution used.

Response: Thank you for catching this. It has been corrected to the accurate 15 arc-second interval grid throughout the manuscript.

3.8: L168-173: Ray-tracing mesh resolution is 10-28 km. The seamounts and fracture zones responsible for focusing may be under-resolved, and barycentric interpolation with precomputed gradients smooths exactly those features. Please add a ray-tracing mesh-resolution sensitivity test (a FUNWAVE resolution test is already provided, L278-280; an equivalent for the rays is needed).

Response: We appreciate the reviewer’s careful attention to the mesh resolution. While standard Cartesian meshes require very fine, uniform resolution to capture bathymetric peaks without smoothing them out, our unstructured mesh bypasses this limitation by explicitly forcing nodal generation directly on high-gradient features during the mesh creation process.

Figure 2 details the spatial distribution of the mesh resolution. As shown in the figure, the adaptive variable-resolution mesh explicitly tracks bathymetric gradients. The finest resolutions (down to 10.77 km) are concentrated strictly along the exact critical refractive structures the reviewer highlighted—such as seamount chains, mid-ocean ridges, and subduction trenches (indicated by the red/orange zones). Conversely, the coarser resolutions (up to 28.35 km) are permitted only in the relatively flat abyssal plains where refraction gradients are naturally minimal.

We have expanded Section 3.2 to explicitly detail this feature-tracking mesh generation process and included the resolution map and distribution statistics to directly address this concern.

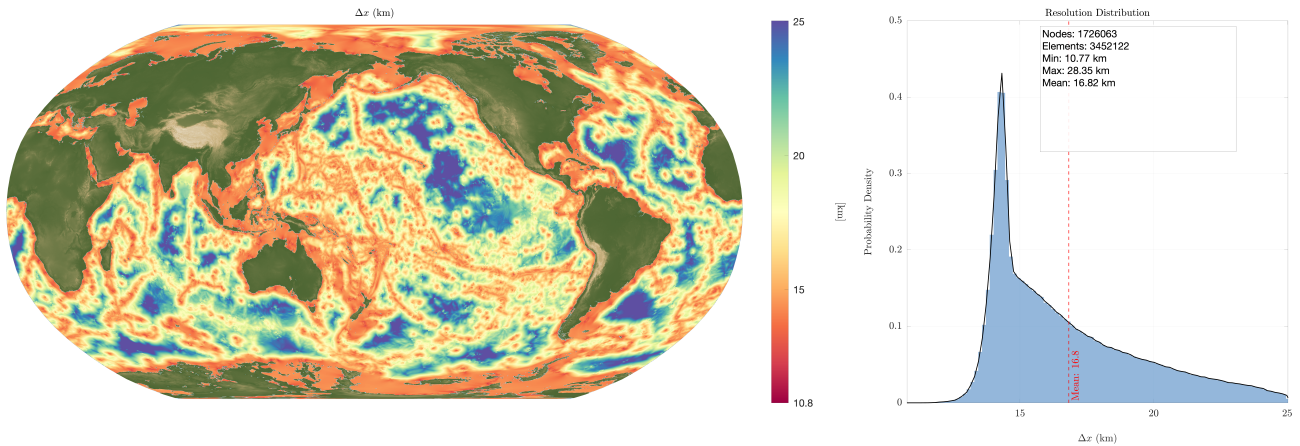


Figure 2: Spatial distribution and statistical properties of the variable-resolution unstructured ray-tracing mesh. (Left) Global map of the element sizes (Δx in km) demonstrating the feature-tracking nodal generation. Finer resolutions (red/yellow, ~ 10.8 – 15 km) are explicitly forced along high-gradient bathymetric features such as seamount chains, mid-ocean ridges, and subduction trenches to preserve critical refractive boundaries. Coarser resolutions (blue, up to 28.35 km) are restricted to relatively flat abyssal plains. (Right) Probability density distribution of the element sizes, detailing the specific mesh geometry of 1,726,063 nodes and 3,452,122 elements with a mean resolution of 16.82 km.

3.9: L179-184: Rays are emitted perpendicular to the fault and oceanward only. Real radiation directivity depends on slip distribution and dip. Please discuss the impact of this simplification on the radiation pattern.

Response: We added text clarifying that while real directivity relies on 3D rupture complexities, the dominant macro-scale energy flux of large subduction megathrusts radiates strictly orthogonal to the fault strike. Weighting the initial ray density by the vertical seafloor deformation envelope accurately initializes these primary radiation lobes for far-field propagation.

3.10: L181-184: The 500 m termination depth is a key parameter; please justify the choice and test sensitivity (see General comment 2.3 on caustics).

Response: We expanded the text to clarify that the scope of this framework is strictly basin-scale transoceanic focusing, and to justify the 500 m termination depth from both a physical and computational standpoint. Physically, this depth truncates the simulation just before the complex nearshore zone. In water shallower than 500 m, non-linear shoaling and extreme refraction dominate, where the standard

geometrical optics approximation fails. Moreover, computationally, tsunami waves entering these shallower waters require a proportionally higher-resolution mesh to resolve accurately. Implementing such high resolution along global coastlines would impose heavy computational costs, directly defeating the primary purpose of this study as a rapid, efficient exposure screening tool. Furthermore, continuing to propagate rays into these shallow zones using our existing coarse deep-water mesh would yield unphysical and inaccurate results. Therefore, the 500 m cutoff acts as a necessary and strict "hand-off" boundary. It safely isolates basin-scale transoceanic focusing from local nearshore amplification, which must be handled by specialized, high-resolution hydrodynamic modeling.

- 3.11:** L230-240: the Kamchatka event summary reports specific run-up heights, casualties, and damage. These should be supported by explicit references if retained.

Response: Citations to the NOAA/NCEI Database and USGS summaries have been added.

- 3.12:** L278-280: The sensitivity simulations are described as being conducted on "higher resolution 8 and 16 arc-minute grids", but 8 and 16 arc-minutes are coarser than the 4 arc-minute production grid, not finer. Either this is a typo ("coarser") or the sensitivity test only explores lower resolutions, in which case the stated conclusion that 4 arc-minutes is "adequate" is not supported by a convergence test toward finer grids. Please clarify.

Response: Thanks for pointing out the typo. It should be 'lower'.

- 3.13:** L334-340: Please specify exactly how co-located FUNWAVE-ray pairs are extracted (segment length, number of points, near-field exclusion criterion) and report the statistics described in General comment 2.6. In particular, please detail the near-field exclusion criterion applied; near the Kamchatka source region, the ray-tracing approximation breaks down due to initial source dynamics and strong diffraction, meaning near-field coastal segments must be strictly excluded to avoid biasing the global far-field statistical correlation.

Response: We updated the text to explicitly detail that data was extracted by joining outputs onto a uniform coastal grid. Crucially, to prevent finite-fault dynamics and strong local diffraction from artificially biasing the global correlation, all coastal segments within a 1,000 km radius of the epicenters were strictly excluded from the statistical pairs prior to running the block bootstrap.

- 3.14:** L356-384 (Conclusions): Language such as "rigorously validated", "robust", and "defensible" should be moderated given $\rho = 0.66$ and the single-event validation.

Response: The tone has been lowered throughout the conclusions. The addition of the Tohoku case also removes the "single-event" limitation.

4. Technical Corrections

- 4.1:** L137: "wether" → "whether".

Response: Corrected.

- 4.2:** Ensure all figure colour scales state explicitly that the coastline metric is a normalised ray-count proxy (0–1), not wave amplitude, to avoid reader misinterpretation.

Response: The description is added to the figure caption and text.