Response to reviewers

We thank the two reviewers for their time spent reviewing our manuscript, “The Role of El Niño Southern Oscillation in driving Coastal Hazards in the U.S. Pacific Northwest,” and their clear feedback on how to improve the communication of our results and the organization of the text. Below, we have included the reviewers’ comments **in bold**, our initial response during the interactive discussion period in *blue italics*, and our new comments in **blue, bold text**. We have incorporated the majority of the reviewers’ feedback into our revision and feel the new version of the manuscript has improved clarity and is less lengthy and dense.

Reviewer 1:

This manuscript investigates the relationship between the El Niño Southern Oscillation (ENSO) and coastal hazards in the U.S. Pacific Northwest (PNW) using a stochastic climate emulator. By generating probabilistic simulations, the authors assess the role of ENSO phase, intensity, and synoptic weather in driving flooding and erosion proxies. A central finding is that ENSO, while linked to hydrodynamic anomalies, is not as reliable a predictor of coastal hazard impacts in the PNW as often inferred from the observational record. This represents a valuable contribution to the literature and has implications for climate-informed coastal hazard management.

We thank the reviewer for their supportive remarks.

Currently, Section 4 merges discussion and conclusion, which makes the text somewhat lengthy and diffuse. I recommend splitting into two sections:

Discussion: further organized into subsections (e.g., ENSO mechanisms, interaction with other climate modes, methodological limitations, comparison with prior studies).

Conclusion: concise summary of the main findings and their implications.

Future research directions. Although the manuscript briefly mentions PDO and other climate modes, the outlook for future work could be strengthened. A dedicated subsection at the end of the discussion would improve clarity. Possible directions include:

Incorporating PDO and other large-scale climate modes into stochastic simulations to assess multi-mode interactions.

Coupling the emulator with shoreline change models to capture longshore sediment transport and ENSO-related shoreline rotation.

Exploring the implications of enhanced ENSO variability under anthropogenic climate change, as recent literature suggests.

Some paragraphs in the discussion are dense and could benefit from restructuring or condensation. For example, the comparison between simulated and historical anomalies could be streamlined with summary tables or schematic figures to highlight the key differences.

Thank you for this suggestion. We have restructured the discussion into sections following your recommendations (see below) and have made several paragraphs less dense by rewording comparisons to other studies.

Recommendation: Moderate revision. The study is timely and well executed, but the manuscript would benefit from clearer structuring of the discussion, a more concise conclusion, and a stronger forward-looking perspective.

Interactive Discussion Response: Thank you for your supportive comments and clear direction on how to improve the manuscript. Your feedback on separating the discussion and conclusion is much appreciated, particularly your suggestions for subsections to improve clarity. We will re-organize our text into the described sections in our revision and include a concrete and expanded discussion of avenues for future work.

We have restructured the Discussion and Conclusion following your recommendation. New Discussion subsections include:

1. **ENSO mechanisms affecting PNW coastal hazards**
2. **ENSO-Coastal Hazard Teleconnections interactions with Other Climate Modes**
3. **Study Limitations and Future Research Directions**
    a. **By restructuring our discussion, we more explicitly call out future research directions including opportunities to include other climate modes (Line 690), shoreline change modeling (Line 695), and future projections of ENSO (Line 685, as suggested by the reviewer, into similar analyses.**
4. **Implications for Future Coastal Hazard Impacts**

Reviewer 2:

This study investigates the relationship between ENSO and coastal hazard risk in the U.S. Pacific Northwest using a stochastic climate emulator. The results show that the ENSO may not be a necessarily have a strong relationship with the coastal hazards in this region. Overall, this study is comprehensive and well designed. However, I still have several questions and suggestions for improving the current work.

Thank you for your comments, both general and specific in how to improve this manuscript. I will address them below, with your comments in bold.

1) In the abstract, only the term “coastal hazards” was mentioned. However, this study only investigated the cross-shore hazard without including the shoreline hazard. It is suggested to make it more specific to avoid misleading conclusions.

- We agree we should be more specific here. We will specify that our study is focused on cross-shore coastal hazards teleconnections to ENSO if we are invited to submit a revised manuscript.

We have now updated this phrase in the abstract and in the re-organized discussion and conclusion.

2) The Introduction Section is too long. It would be better to make it more concise, only highlight the existing limitation and research gaps, and remove some unnecessary details in previous studies. Also, in the Discussion & Conclusion, several previous studies were mentioned. It is suggested to incorporate some into the Introduction Section, and only highlight the new findings and new questions raised from this study.
Thank you for this feedback. During a previous review process, we received comments requesting that we include more context surrounding ENSO teleconnections and dynamics, leading to a longer introduction section. We will consider how to make this section more concise per your recommendation while retaining important background, we agree it is a bit lengthy.

We have removed several sections from the introduction that focus on ENSO background and processes, streamlining our text to focus on research gaps and providing context for our specific analysis. The introduction was initially 2667 words, now it is 2123 words long.

Similarly, following your advice and the advice of reviewer 1, we will reorganize our Discussion & Conclusion to improve clarity and focus on the findings and questions raised from this study. However, we do think that some of references to other studies in the discussion are necessary and better situated at the end of the paper, particularly the discussion of the influence of other climate modes and changing frequency and intensity of ENSO. So, we will likely retain references to some previous studies in the Discussion.

We have also reduced our synthesis of other studies in the Discussion. While we have not removed new mentions of previous literature in this section, we hope the reviewer will be satisfied that the changes made improve the readability & concision of the discussion.

3) In Fig. 1, what does the number in the top left of the sub-figures of AWT and DWT mean?
These numbers refer to the AWT and DWT labels. So, in this figure we are showing AWT 1 (linked to El Niño) and 6 (linked to La Niña). Likewise, we are showing DWTs 1, 2, 7, and 8. We can specify what these labels mean in the caption.

We have updated Figure 1 to include a brief explanation.

4) In the k-means clustering, how to determine the value of k?
The value of k (which determines the number of weather types) was inherited from previous published applications of the stochastic climate emulator, which were more focused on model development (e.g., Anderson et al., 2019 https://doi.org/10.1029/2019JC015312; Cagigal et al.,

2021 https://doi.org/10.1029/2020JC016919 ; Cagigal et al., 2020 https://doi.org/10.1016/j.ocemod.2020.101695; Camus et al., 2014 https://doi.org/10.1002/2014JC010141 ; Camus et al., 2011 https://doi.org/10.1016/j.coastaleng.2011.02.003). During these studies extensive sensitivity testing occurred to balance the number of clusters (k) while having enough observations within each cluster to perform robust statistical analyses on them. We can specify in the revision process that while we did not perform sensitivity testing for our application, there are several applications we are drawing from where this tuning has been performed.

We have added text to line 326, highlighting previous studies that have focused on choosing the number of weather types.

5) Could you justify the usage of Gaussian copulas in the stochastic climate emulator?

Within the current text we justify the usage of gaussian copulas on line 369: "Employing gaussian copulas maintains the historical dependence structures between sea-state parameters and weather types while allowing for extrapolation from historical observations (SI.1b; Cagigal et al., 2020)." We also expand on this justification in the supplemental section.

However, understanding the pros and cons of the choice of gaussian copulas as compared to other copula types is an open area of research. Anderson et al. (2019), in their original description of the emulator methodology, highlight this: "The application of Gaussian copulas for joint probabilities was another user decision applied after initial random sampling from each univariate distribution resulted in underestimation of the extremes. This element of the methodology could be improved in the future with an automated goodness-of-fit decision choosing which of the many copula formulations best defines the joint probabilities."

We have added a line to the supplemental section that references the above quote from Anderson et al. (2019) and describes that copula choice an open area of research.

6) Line 310, one of the empirical wave runup formulas was used to calculate the TWL, but it is expected that different empirical formulas may lead to different estimates. Why this particular one was employed? How would you quantify the uncertainty propagation through the simulation process?

On line 311 (now 395), we mention that we are using two different runup formulas for two different situations.

"On dune-backed beach types, the Stockdon (Stockdon et al., 2006) formula was used, while on cliff or riprap backed beach types, a modified TAW (Technical Advisory Committee for Water Retaining Structures) barrier runup method was applied when TWL elevation exceeded the barrier toe (Allan et al., 2015; Leung, et al., 2024b; Pullen et al., 2007; van der Meer, 2002)."

The Stockdon formula is globally the most popular wave runup model as it was developed based on field experiments at 10 different beaches, 2 of which were in the Oregon/ Cascadia *region. So, it is as well-tuned as any formula for the majority of beaches assessed in this study. The combination of Stockdon and the modified TAW formulas has been employed in FEMA flood hazard analysis reports in the region (e.g., Allan et al., 2015 https://doi.org/10.13140/RG.2.1.1656.5608) so employing this methodology facilitates comparisons to existing hazard studies in Cascadia and familiarity for the coastal practitioners in the region.*

We have slightly reworded this section to emphasize that we are using two different runup formulas and our justification for the formulas (Line 397).

The uncertainty of the Stockdon formulation is well known (~0.2 m RMSE) and carries through to all of our total water level calculations.

7) Line 316: How did you define an "unsafe beach"?

We specify in line 403: "These proxies state that unsafe beach conditions are met when the total water level (TWL) is high enough that the beach width is less than 10 m (this threshold width is designed to be flexible based on the unique geomorphic setting or community needs).."

This definition is expanded in Leung et al. 2024 -which we reference- where it was originally presented:

"The unsafe beach hours metric was developed through extensive engagement with community partners in the Cascadia region. Partners highlighted that solely assessing flooding and erosion proxies tends to center hazard impact discussions on (typically wealthy) coastal property owners only. Interested parties expressed interest in a proxy that can communicate how the general population may also be impacted by chronic coastal hazards. On their suggestion, we co-created a beach safety proxy that underscores how visitors to beaches (either for work or leisure) may feel unable to utilize the beach for their preferred activities based on its time-dependent width. To quantify beach safety, we track the number of daylight hours during which the beach is 'unsafe', or too narrow to comfortably recreate without safety concerns. The definition of 'too narrow' should be determined based on the unique conditions of a particular beach and how visitors use it. In this study a threshold width of 10 m was applied regionally for simplicity after testing varying thresholds of 10, 15, and 20 m based on average beach widths in the region and input from our partners. The 10 m threshold was chosen for presentation here because it is the least conservative in terms of how wide a beach needs to be in order to be "useable", while still capturing changes in dry beach width on seasonal to decadal scales."

8) Line 348, what methods were used to construct the probability density functions of the hydrodynamic variables? Some inherent parameters of these methods may have a significant effect on the shape of the probability density curves.

We used the seaborn python package (https://seaborn.pydata.org/generated/seaborn.violinplot.html) to create the violin and split violin plots, using their default settings to create the probability density functions. In this case, the binwidth method used is 'scott', and the density_norm method is 'area'. We tested multiple settings but did not find perceptible differences in the shape of the *probability density functions. We have included these tests in the attached PDF file.*

9) Lines 373-374: How did you define the "anomalies" of those variables?

To assess anomalies, we isolate the wave energy flux, water level, and wave direction during winter months (December, January, and February). We take the average over the entire timeseries (so we have a winter average for the historical and a winter average for the simulated timeseries). An average is then calculated for each winter "year" and then the winter average from the full timeseries is subtracted.

Our approach for computing anomalies is now detailed in lines 482-484.

10) Line 401: Why could the EP El Nino exhibit negative winter wave energy anomalies? Are there any evidence observed from any historical events?

Thank you for the opportunity to expand on how EP El Niño could exhibit negative winter wave energy anomalies. Here, we will discuss this from a statistical perspective in the model, as well as a dynamical perspective.

Regarding the statistical techniques used in the model, negative winter wave energy could occur during an EP El Niño year due to sampling statistically feasible, but different than observed, daily weather types (DWTs) and associated hydrodynamic variables (e.g., wave height, period).

As shown in figure 4, some DWTs are more likely to occur during winter months (4b) and EP El Niño years (4c). For a given EP El Niño year, new DWT timeseries are generated based on these probabilities (and the probability to transition from one DWT to another) using the auto-logistic regression (ALR) model. The inherit stochasticity of the ALR model means that in some EP El Niño winters, more stormy DWTs may be sampled than in other EP El Niño winters (contributing to higher or lower wave energy flux anomalies respectively).

Further, the synthetic wave characteristic timeseries are populated based on the simulated DWT timeseries using gaussian copulas of the relevant hydrodynamic variables. As such, there will be simulations where randomly sampling wave characteristics from the copulas will result in more or less extreme cases by design. By stochastically simulating different ENSO years and the associated hydrodynamic variables, the model attempts to represent the full range of probability based on observations. For the synthetic EP El Niño years with negative wave energy flux anomalies, the roll of the dice for those simulations resulted in fewer stormy daily weather types and lower wave heights/ smaller wave periods.

We calculated in our simulations of EP El Niño (1283 events), 360 have negative winter wave energy flux anomalies (28%). For simulated CP El Niño events, 646 events or 49% of events have negative winter wave energy flux anomalies.

Dynamically speaking, as large wave events are driven by more frequent high-energy storms that have different directionality (compared to non-El Niño years) the wave impacts can be highly dependent on the unique storm directionalities of a particular El Niño year. This can be *seen in our supplemental figure 1a and b, which compares winter wave anomalies in Northern Washington to Northern California. In particular, the 1991-92 EP El Niño the spatial variability in winter wave energy anomalies: in Northern Washington wave energy flux is ~20 W/m, while it gets progressively lower in Northern Oregon (~10W/m; figure 5a), and Northern California (~2W/m). While this Northern California wave energy flux is still positive, it highlights that directionality of storminess can lead to a wide range of differences in local wave impacts across the Cascadia region.*

To address your question regarding historical events, in our analysis of observed wave energy, no EP El Niño events have exhibited negative winter anomalies in Cascadia (here we assess only 4 observed different events). However, the 2023-2024 (mixed EP-CP) event further highlights, (as discussed in our introduction) that waves impacting Cascadia are not always extreme during even extreme El Niños (only 3% higher than average).

11) Lines 420-421: Fig. 5(d) actually shows that the probability density curves for the historical and simulated variables are different to some extent. Also, The legend of Fig. 5(d) is not consistent with the color in the figure.

Thank you very much for bringing our attention to this issue. After comparing figure versions, we realized we accidentally included an old version of this figure that did not include the bias correction in Hs

(This is discussed in Line 510 of the revised submission).

We have attached a pdf that compares the un-bias corrected version (in our original submission) and the bias corrected version of the figure. The bias correction in Hs only improves the simulated distributions of wave energy flux. This bias correction also minorly affects wave height in fig 5a,b,c which is also included in the pdf, though it does not change the overall findings of this paper. We have confirmed that the hazard proxy analysis included in the paper did use the bias corrected wave heights. Should the manuscript move forward to the revision process, we will ensure all reported numbers reflect the bias corrected wave energy flux.

We have updated the figures and reported numbers in the text (both main manuscript and SI figure 1 in the supplemental section.

There are still slight differences in the distributions of all three of the variables explored in this plot (wave energy flux, water level, direction). These differences are more visible in the anomaly plots than in the full distributions, as we are intentionally highlighting the extremes. Naturally, the tails of the distributions have more uncertainty, in both the observed and simulated data, as there are fewer observations of extreme events and our simulations rely on some statistical assumptions we make about the shape of the observed distributions. We see more dramatic differences in the DWT compared to AWT since we are parsing the data at a more granular level. For example, DWT 6 has a slight offset in wave direction (figure 5d). This DWT is one of the lowest probability DWTs to occur, and has only 72 days of hourly data, or ~1500 data *points, in the historical period compared to 24,005 days, or ~580,000 data points, in the simulated dataset. By contrast, in the AWT anomaly plots we compare the historical period, with 4 winter seasons (4 x 3 months x ~30 days x 24hrs) →~8,600 data points, to the simulated period with (1295 winter seasons→~ 2.7million data points.*

Furthermore, the simulated wave height evolves hourly based on parameterized hydrographs, while simulated wave period, direction, and storm surge is constant throughout the simulated hydrograph. This contributes to more differences in the historical distributions to the synthetic distributions. Comparing the historical vs simulated hydrographs rather than hourly data does improve the comparisons of the two distributions across DWTs, though as we use the hourly data in the hazard proxy impact analysis, we felt using the hourly distributions made more sense in the progression from figure 5 to figure 6.

We can add more information about these factors / decisions affecting the historical vs simulated distributions in our analysis.

We have added text in the interpretation of this plot (line 531) that expands on why there are some differences between the historical and simulated distribution for figure 5d.

The legend uses the colors from WT1 to indicate that the historical data is full opacity (darker), while the simulated data is half opacity (lighter), though the colors change for the different weather types, the darker vs lighter trend is what we were trying to highlight.

The figure caption describes the saturated/ opaque label for the observed distributions compared to the lighter/ desaturated colors for the simulated distributions.

12) Line 524: Could you elaborate what the model limitations are?

We can add to the line as follows to highlight a few model limitations:

The finding that El Niño is not necessarily a strong predictor of intensified coastal hazards could be attributable to model limitations. For example, short or imperfect observations used as input data, statistical assumptions used to derive joint probabilities, and parameterizations of wave characteristics all contribute to model uncertainty and limitations in skill.

We have adjusted lines 665-668 as described above

Minor Issues

13) The label of vertical axis of Fig. 2(b) was covered by Fig. 2(c).

Thank you for catching this, we can adjust this figure.

The axis has been modified

14) In Figs. 5 and 6, the numbers in the violin plots are not clear and some overlapped with each other.

Thank you for catching this, we can adjust this figure.

We have manually adjusted text placement.