# Response to Anonymous Referee 1

Dear Referee,

We sincerely thank you for your time and effort in reviewing our manuscript: *SPIN (v1.0): A Spontaneous Synthetic Tropical Cyclone Model Empowered by NeuralGCM for Hazard Assessment*. We are grateful for your supportive assessment of our work and for the helpful comments you provided. We have carefully addressed each of your points and revised the manuscript where appropriate. Please find our detailed responses below.

**Major comments:**

1. First, the authors argue that SPIN enables two-way interaction between TC and environment, which in my opinion is only partially true. Storm intensity in SPIN is post-processed using FAST, so it does not really 'feedback' to NeuralGCM's environment. As a result, MTCs in SPIN do not really reflect true storm-to-storm interaction. Please add a couple sentences of this limitation.

***Response:***

Thank you very much for pointing this out. We agree with the reviewer that the original statement overstated the degree of two-way coupling between TCs and the environment in SPIN. Storm intensity is diagnosed offline using FAST, and NeuralGCM is an atmosphere only model. Therefore, air–sea interactions related to storm intensity, such as storm induced cold-wake cooling, do not feed back onto the simulated environment or influence other TCs. We have therefore removed the statement that SPIN enables two-way TC–environment interaction.

The revised manuscript now clarifies that SPIN's key advantage lies in diagnosing synthetic storms from hourly evolving environmental fields, thereby representing intra-month variability relevant to TC genesis and providing synthetic storms together with their ambient environments.

***Changes in the manuscript:***

We have added this limitation to the manuscript as follows (Lines 556–558):

> "Moreover, TC intensity is diagnosed offline using FAST, and NeuralGCM is an atmosphere only model. Therefore, air–sea interactions related to storm intensity, such as storm induced cold-wake cooling, do not feed back onto the simulated environment or influence other TCs."

2. Could you elaborate on how the simulations were conducted? Are the 14 ensemble members' simulations initialized 6 hours apart from each other? Can these simulations be considered SST-forced runs, meaning that after the initial time, the only input from ERA5 is the monthly SST and SIC? And there is a 2.5-month spin-up period, am I correct and is this necessary?

***Response:***

Thank you for this comment. We agree with your interpretation. The 14 ensemble members were initialized at 6-hour intervals in mid-October of each year. After initialization, the model was run freely, with SST and SIC as the only inputs. The only inputs after the initial condition were SST and SIC, which were linearly interpolated from monthly values to 12-hourly intervals. No ERA5 atmospheric fields were used after initialization.

The approximately 2.5-month spin-up period follows the experimental setup of Kochkov et al. (2024) and Cheng et al. (2022). It is used because our evaluation covers the full January–December period, initializing later would reduce the spin-up before January and could affect early-year TC activity, particularly in Southern Hemisphere basins where the TC season is active during January–March.

***Changes in the manuscript:***

We have therefore revised the manuscript as follows to clarify the simulation process (Lines 115–123):

> "For each target year from 1980 to 2022, we initialize a 14-member NeuralGCM ensemble simulation. The ensemble is generated using lagged initial conditions, with members initialized at 6-hour intervals in mid-October of the preceding year and integrated continuously through December of the target year. For example, the simulations used for the 1980 analysis are initialized in mid-October 1979 and integrated through December 1980. The preceding October–December period is treated as spin-up, and the January–December period of the target year is used for TC tracking and subsequent analysis. All initial and lower-boundary fields are derived from ERA5 (see Section 2.1). The initial atmospheric states include wind, geopotential, temperature, specific humidity, and specific cloud ice and liquid water content on pressure levels. Monthly mean SST and SIC fields are linearly interpolated to 12-hourly resolution and prescribed as lower-boundary conditions throughout the simulations."

And Lines 276–280:

> "We use January–December simulations here to maintain comparability with previous hazard-model evaluations. For application-oriented experiments, however, initializing closer to the target TC season can shorten the forecast lead time and reduces the accumulation of model drift. Consistent with this interpretation, April initialization simulations followed by evaluation over the Northern Hemisphere hurricane season yields higher basin-scale TC-frequency correlations (Figure 4)."

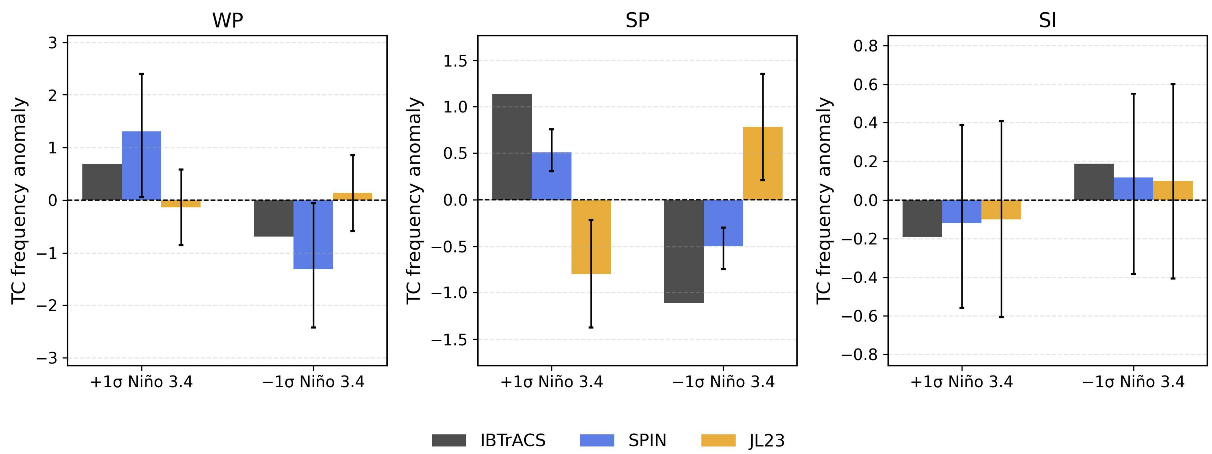
3. I am not really following the argument here — it basically says that NeuralGCM better captures the ENSO modulation of TCs, but both the JL models and other models (Lin et al. 2024; Lee et al. 2025) show that they can simulate ENSO modulation of TCs as well. A clearer demonstration would be a direct comparison of interannual TC frequency or intensity anomalies conditioned on ENSO phase across models.

***Response:***

Thank you for this helpful comment. We agree that the previous discussion did not sufficiently demonstrate the difference between SPIN and existing downscaling models in terms of ENSO-related TC variability. According to JL23's own manuscript, they denoted that their model's difficulty in simulating annual TC frequency in these basins may be due to:

> "El Niño-Southern Oscillation, the main source of tropical interannual variability, has single-signed signals in TC genesis in the Atlantic and Eastern Pacific basins, but mixed-signed signals in the Western Pacific, South Pacific, and South Indian basins. It is likely that the interannual signal in genesis gets averaged out in basins such as the Western Pacific."

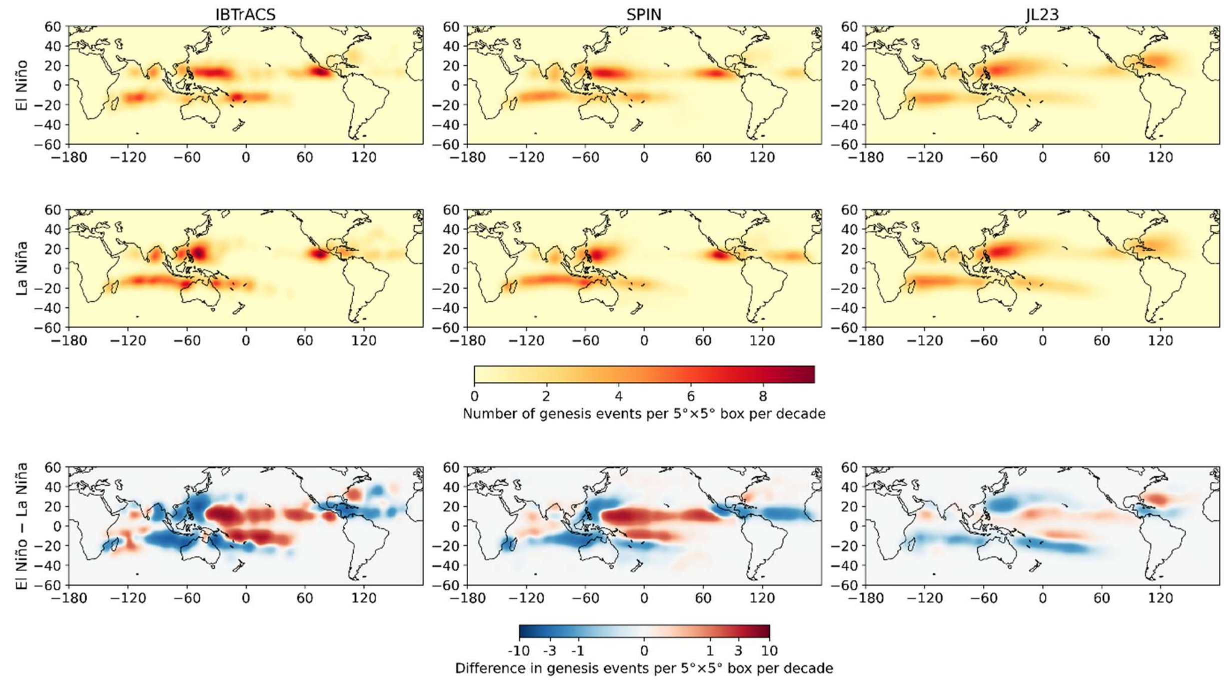
We have therefore examined ENSO conditioned TC frequency anomalies in regions with mixed-sign ENSO responses (Response Fig. R1). SPIN shows closer agreement with IBTrACS than JL23. In the WP, SPIN and IBTrACS both show positive anomalies under El Niño-like conditions and negative anomalies under La Niña-like conditions, whereas JL23 remains close to zero. In the SP, SPIN is also consistent with IBTrACS in sign, while JL23 shows the opposite tendency. In the SI, the ENSO-related anomalies are weaker and more uncertain; nevertheless, SPIN remains closer to the weak IBTrACS response than JL23. These results suggest that SPIN better captures ENSO modulation of TC frequency in these regions with mixed-sign ENSO responses.




**Response Fig. R1.** ENSO modulation of TC frequency in the WP, SP and SI basins. Bars show the TC frequency anomalies expected under +1σ (El Niño-like) and −1σ (La Niña-like) Niño 3.4 conditions, estimated from the linear relationship between TC frequency anomalies and the Niño 3.4 index over 1980–2022. TC frequency and Niño 3.4 index are calculated over May–December for the WP basin and December–April for the SP and SI basins. TC anomalies are expressed relative to each dataset's own climatology. Error bars show the 10th–90th percentile range across SPIN ensemble members and the 95% OLS confidence interval for JL23. JL23 counts are scaled separately within each basin so that the total JL23 TC count in each basin matches that of IBTrACS.

In addition, we show the maps of the El Niño minus La Niña difference in TC genesis density. Both SPIN and JL23 reproduce the sign of the multi-year mean ENSO-related genesis signal. Notably, in the WP, SP, and SI basins, positive and negative genesis anomalies coexist within the same basin. However, SPIN shows the magnitude of genesis density difference between El Niño and La Niña conditions that are closer to IBTrACS, whereas JL23 produces weaker and smoother anomalies (Response Fig. R2).

A possible reason is the difference in the treatment of seed disturbances. Previous diagnostic frameworks express TC genesis rates as the product of seed frequency and a transition probability controlled by the large-scale environment (Emanuel, 2022; Hsieh et al., 2020). In JL23, the seed supply is prescribed randomly and does not vary dynamically with ENSO related circulation changes. In contrast, SPIN identifies candidate storms directly from the evolving circulation in NeuralGCM, allowing both environmental favourability and precursor disturbance supply to vary from year to year. This may help SPIN retain regional ENSO responses that are partly smoothed in JL23. In other words, random seeding in JL23 may introduce disturbances in regions where activity should be suppressed, or miss regions where activity should be enhanced. We therefore present this mechanism as a plausible explanation, rather than a definitive attribution.




**Response Fig. R2.** TC genesis counts in 5° × 5° grid boxes per decade from IBTrACS, SPIN and JL23 during 1980-2022. JL23 genesis densities are shown after applying a constant global scaling factor so that the total JL23 genesis count over 1980–2022 matches that of IBTrACS. Top and middle rows show genesis counts during El Niño-like and La Niña-like months, respectively, defined using the Niño 3.4 SST index with thresholds of +1σ and −1σ. The bottom row shows the El Niño minus La Niña difference in genesis counts per decade. The colour scale is linear between −1 and 1 and logarithmic for absolute differences greater than 1.

### *Changes in the manuscript:*

To make this point clearer in the manuscript, we have revised the text in Lines 292–305 as follows:

> "SPIN also shows markedly improved agreement with observations relative to JL23 in the WP, SP, and SI basins. Lin et al. (2023) suggested that lower predictability in these basins may reflect the mixed-sign influence of El Niño–Southern Oscillation (ENSO) on TC genesis across subregions, (Figure S6; Camargo et al., 2007; Lin et al., 2023), which can cause the basin-integrated interannual genesis signal to be averaged out. In contrast, SPIN better retains ENSO modulation in these basins, with ENSO conditioned TC frequency anomalies closer to IBTrACS than JL23 (Figure S7). A possible reason is the difference in the treatment of seed disturbances. Previous diagnostic frameworks express TC genesis rates as the product of seed frequency and a transition probability controlled by the large-scale environment (Emanuel, 2022; Hsieh et al., 2020). JL23 uses random seeding (Emanuel et al., 2008; Lin et al., 2023), in which seeds are placed randomly in space and time and then allowed to evolve within the large-scale environment. Because this seed supply is not dynamically coupled to ENSO related circulation changes, it may artificially increase genesis probability in low-activity regions while decreasing it in high-activity regions. In contrast, SPIN diagnoses candidate storms from the evolving circulation fields generated by NeuralGCM. This allows SPIN to retain year-to-year changes in both environmental favorability and precursor disturbance supply. We regard this as a plausible mechanism rather than a

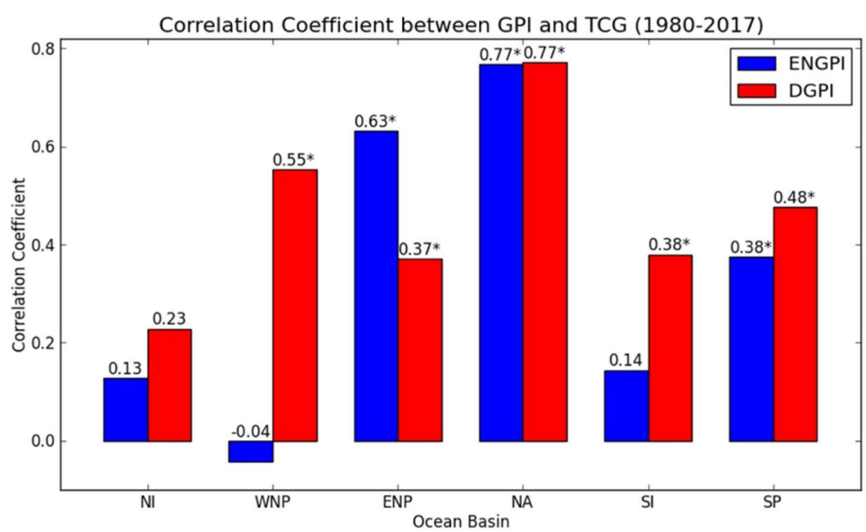
definitive attribution. Overall, these results highlight the potential of SPIN to capture interannual variability in TC activities across most basins."

4. Also, regarding Line 190, one assumption of the random-seeding approach in the JL model is that the genesis process is simply part of intensification. However, your argument seems to suggest that this assumption may not hold for interannual variability. Could you further discuss this, and whether approaches that use genesis indices would be a better way to handle interannual variability?

***Response:***

Thank you for raising this important point. As mentioned in our response to Comment 3, previous diagnostic frameworks express TC genesis rates as the product of seed frequency and a transition probability controlled by the large scale environment (Emanuel, 2022; Hsieh et al., 2020). JL23 uses random seeding, in which seeds are placed randomly in space and time and then allowed to evolve within the large-scale environment. Because this seed supply is not dynamically coupled to ENSO related circulation changes, it may artificially increase genesis probability in low-activity regions while decreasing it in high-activity regions. In contrast, SPIN identifies candidate storms directly from the evolving NeuralGCM circulation. This allows SPIN to better retain year to year changes in both environmental favourability and precursor disturbance supply.

Regarding the use of genesis indices, we agree that they provide a useful complementary diagnostic for evaluating interannual variability in TCG. Previous studies have shown that ENGPI and DGPI can capture aspects of annual TCG variability, although their skill varies substantially by basin (Response Fig. R3, adapted from Wang and Murakami (2020). In particular, DGPI shows improved performance over ENGPI in several basins, including the WP, SI, and SP. However, for hazard modelling, genesis potential indices alone are not sufficient. They provide diagnostic estimates of environmental favourability, but they do not generate individual TC tracks, including storm location, intensity evolution, lifetime, and landfall characteristics.




**Response Fig. R3.** Adapted from Fig. 5 of Wang and Murakami (2020). Correlation skills of the GPI in diagnosing the interannual variation of the basin total TCG. Shown are the correlation coefficients between the observed and diagnosed interannual variation of the basin total TCG number. Blue bars denote ENGPI, whereas red bars denote DGPI. The numbers over the bar indicate correlation coefficients. The asterisk indicates statistically significant at the 95% significance level using student t-test. The data period is 1980–2017.

5. Figure 8. The area definition is not precise. It is not just the eastern US — you also include the Gulf of Mexico, which includes the southern US.

***Response:***

Thank you for pointing this out. We have revised the figure label to “Gulf–Atlantic Region” to better reflect the geographic coverage.

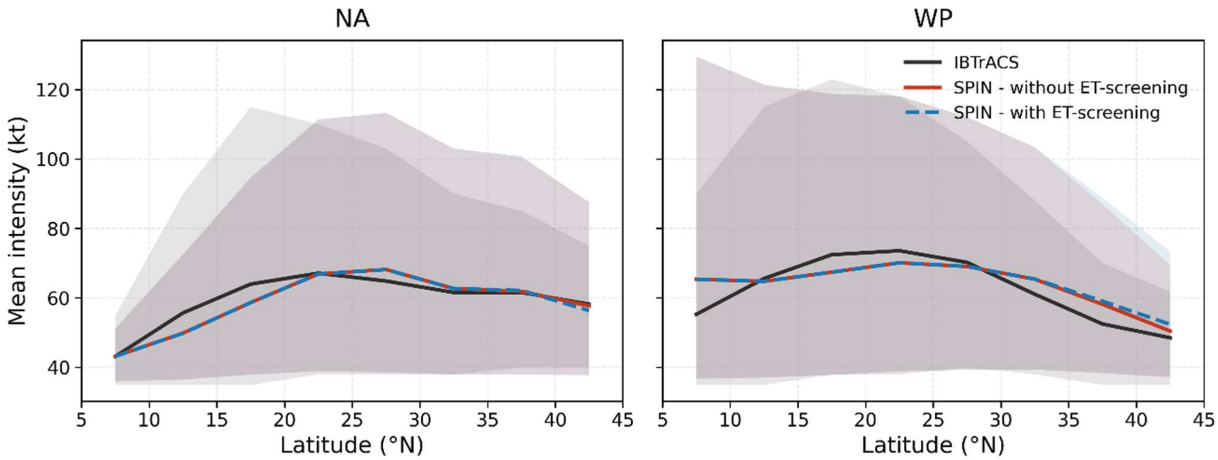
6. How did you handle extratropical transition (ET) storms? If you simply run FAST all the way to the mid-latitudes, you are likely to overestimate storm intensity and introduce a positive bias in the number of MTCs.

***Response:***

Thank you very much for raising this important point. In SPIN v1.0, extratropical transition is not explicitly represented in the intensity model, consistent with many synthetic tropical cyclone hazard models that focus on genesis, track, intensity, and dissipation rather than on post-tropical structural evolution (Jing et al., 2024; Jing and Lin, 2020; Lee et al., 2018; Lin et al., 2023). However, we agree that this issue might be relevant for the present study, and we have therefore included an optional ET-screening module in the updated model. This module allows users to account for ET in applications involving mid-latitude TC events or MTCEs, while keeping the framework flexible for other applications.

ET involves structural changes in the storm core and would ideally be diagnosed using three-dimensional thermodynamic metrics such as cyclone phase space diagnostics (Hart, 2003). Applying such diagnostics to NeuralGCM-1.4° output, however, would require a dedicated validation of the model’s three-dimensional thermal phase of the storm. Moreover, CPS values are sensitive to the resolution of the input data, and no universal thresholds exist for determining when a cyclone has transitioned from one phase to another (Wood et al., 2023). Here, our objective is not to classify ET explicitly, but to limit retaining TC intensity after storms enter ET-favourable environments. We therefore use SST and latitude as simple environmental indicators. This choice is consistent with the physical understanding that ET is often associated with poleward motion into cooler SSTs (Evans et al., 2017), and with the statistical results of Bieli et al. (2020), which identified SST and latitude as important predictors of ET occurrence. Specifically, when both conditions are satisfied for two consecutive time steps, we truncate the storm starting from the first of those two time steps: (1) the area-averaged SST within a 2° radius of the storm center is ≤ 25.5°C, and (2) the storm’s latitude exceeds the basin-specific median ET latitude derived from IBTrACS. Sensitivity tests using SST thresholds of 25.0 and 26.0 °C and latitude thresholds shifted by ±2° show little influence on the reported results.

We updated the results by applying the ET-screening module. Applying this module has a limited effect on the analyses presented in this study. For the annual MTCE frequency, the correlation coefficient and significance level remain almost unchanged. We also find that the simulated mean intensity as a function of latitude generally agrees with IBTrACS, both with and without ET truncation (Response Fig. R4). We also performed sensitivity tests using nearby thresholds, including SST thresholds of 25.0°C and 26.0°C and latitude thresholds shifted by ±2° from the IBTrACS-derived median ET latitude. The main results remain robust under these alternative settings.




**Response Fig. R4.** Mean TC intensity as a function of latitude in the NA and WP basins during 1980–2022. IBTrACS are shown in black. SPIN simulations without ET screening are shown in red, and SPIN simulations with the optional ET-screening module are shown in blue dashed lines. Shaded bands indicate the 10th–90th percentile range.

***Changes in the manuscript:***

We have added a new section, Sect. 2.2.4 "Extratropical transition screening module", and included the following text (Lines 164–182):

> "2.2.4. Extratropical transition screening module
>
> For storms extending into the mid-latitudes, the FAST intensity model may continue to estimate intensity under a TC framework even after the storm has entered environments favourable for extratropical transition (ET) (Jones et al., 2003). We therefore implemented an optional ET-screening module that masks storm records after the onset of ET-like environmental conditions in analyses that depend on TC intensity.
>
> ET involves structural changes in the storm core and would ideally be diagnosed using three-dimensional thermodynamic metrics such as cyclone phase space (CPS) diagnostics (Hart, 2003). However, applying CPS-based metrics to NeuralGCM-1.4° would require dedicated validation of the model's three-dimensional thermal phase of the storm. In particular, these metrics are sensitive to the resolution of the input data and lack universal thresholds for determining when ET occurs (Wood et al., 2023). Here, our objective is not to classify ET explicitly, but to limit retaining TC intensity after storms enter ET-favourable environments. We therefore use SST and latitude as simple environmental indicators. This choice is consistent with the physical understanding that ET is often associated with poleward motion into cooler SSTs (Evans et al., 2017), and with the statistical results of Bieli et al. (2020), which identified SST and latitude as important predictors of ET occurrence. Specifically, when both conditions are satisfied for two consecutive time steps, we truncate the storm starting from the first of those two time steps: (1) the area-averaged SST within a 2° radius of the storm center is ≤ 25.5°C, and (2) the storm's latitude exceeds the basin-specific median ET latitude derived from IBTrACS. Sensitivity tests using SST thresholds of 25.0 and 26.0 °C and latitude thresholds shifted by ±2° show little influence on the reported results."

The screening is intended to limit the inclusion of records that are likely affected by extratropical transition, although the conclusions of this study are not sensitive to its application.

7. Line 300. You may need to check with JL23 for details — I think most existing statistical-dynamical downscaling models can provide date information. It may stop at monthly resolution because the input is monthly data, and thus 'daily' information is simply artificially generated due to the seeding rate. However, they do have date information, and MTCs will exist when the monthly environmental conditions are more favorable than in other months. So in a way, is this not similar to your approach of using monthly SST input.

***Response:***

Thank you very much for this helpful clarification. Our point is that JL23's output provides storm year, month and lifetime rather than a dynamically simulated genesis date. Therefore, daily genesis dates are typically assigned in post-processing, for example by uniform sampling within each month.

SPIN differs because NeuralGCM continuously evolves the atmospheric state at synoptic-to-submonthly timescales. Although SST and sea ice are prescribed from monthly means, they act as slowly varying lower-boundary forcing and do not constrain the atmosphere to monthly variability. This allows TC genesis to emerge from the evolving synoptic-scale atmospheric disturbances. Importantly, NeuralGCM has shown skill in simulating MJO teleconnections (Peings et al., 2026), which are relevant to MTCE occurrence (Gao and Li, 2011; Ito and Yamauchi, 2026).

This distinction matters because random date assignment can bias inter-genesis intervals. In our analysis, uniform assignment overestimates the probability of 0–5-day inter-genesis intervals, especially in the WP basin (Figure 11), which can artificially increase the number of storms grouped into the same MTCE cluster (Figure 13).

***Changes in the manuscript:***

To avoid ambiguity, we have revised the original text as follows (Lines 402–406):

> "Traditional statistical downscaling models, such as JL23, provide the month of TC genesis and the corresponding storm lifetime, but do not resolve intra-month variability needed to determine daily genesis time."

And Lines 414–418:

> "Although SPIN uses monthly SST and sea-ice boundary forcing, it resolves submonthly atmospheric variability. Its improved representation of inter-genesis intervals may therefore reflect the ability of the NeuralGCM component to capture intramonth environmental evolution, including MJO-related teleconnections (Peings et al., 2026). By allowing TC to emerge from the evolving atmospheric state, rather than imposing independence among storms, SPIN provides a more physically consistent basis for MTCE estimates."

8. L330. In SPIN, when you apply NeuralGCM output to FAST, do you use instantaneous output or monthly averaged fields? Also, do you have any idea why JL23-ERA5 performs better than JL23-NeuralGCM?

***Response:***

Thank you for this comment. In SPIN, when applying NeuralGCM output to FAST, we use the same type of monthly environmental inputs as in JL23-ERA5, namely monthly statistics of daily environmental winds at 250 and 850 hPa, together with monthly mean temperature, relative humidity, and sea surface temperature.

Regarding why JL23-ERA5 performs better than JL23-NeuralGCM, we speculate that this difference may mainly arise for two reasons.

First, TC genesis is highly sensitive to high-frequency environmental variability, whereas monthly averaging substantially smooths the signals most relevant to genesis, such as short-lived periods of reduced vertical wind shear, enhanced humidity, and increased low-level vorticity. For a data-driven model such as NeuralGCM, the amplitude of environmental anomalies may already be somewhat underestimated relative to ERA5, and monthly averaging may further weaken these signals. As a result, when JL23 is driven by monthly mean large-scale fields from NeuralGCM, part of the high-frequency and extreme information critical for TC genesis may already have been lost.

Second, the parameters in JL23 are tuned based on ERA5. Applying JL23 to NeuralGCM output introduces a mismatch in the input environmental fields, including differences in grid resolution and systematic biases, which likely contributes to the degraded performance of JL23.

By contrast, directly extracting TCs from hourly NeuralGCM output is likely less sensitive to these issues, because TempestExtremes identifies TCs primarily from local anomalies relative to the surrounding environment, gradient-based features, and vortex structure, rather than relying solely on the absolute magnitude of monthly mean environmental fields.

***Changes in the manuscript:***

We have clarified the FAST input fields in the revised manuscript (Lines 186–188):

> “The JL23 model is driven by monthly statistics of daily environmental winds at 250 and 850 hPa, together with monthly-mean temperature, relative humidity, and sea surface temperature.”

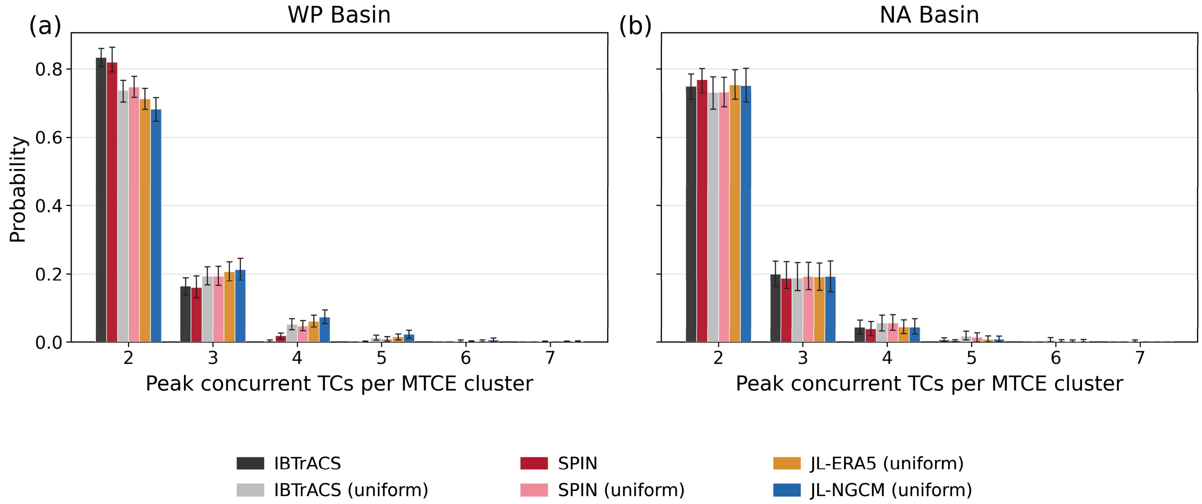
We have also added a brief discussion in the revised manuscript (Lines 442–446):

> “The poorer performance of JL23-uniform when driven by NeuralGCM (Figure 12 e,f), relative to ERA5 (Figure 12 c,d), may arise for two reasons. First, environmental anomalies may be weaker in NeuralGCM than in ERA5 (Duan et al., 2025), and monthly averaging may further damp the signals relevant to TC genesis. Second, because the parameters in JL23 are tuned based on ERA5, applying it to NeuralGCM output introduces differences in the input environmental fields, including grid resolution and systematic biases, which may further reduce its performance."

9. Can you show me the sample errors of your MTCE analysis, like those in Figure 10?

***Response:***

Thank you for this helpful suggestion. We have added error bars to this figure to quantify the sampling uncertainty in the MTCE cluster-size distributions. Error bars indicate the 10th–90th percentile range. For IBTrACS, the range is estimated by year bootstrapping over 1980–2022. For SPIN, it represents the spread across ensemble members. For the "uniform" experiments, the range is estimated across 1000 Monte Carlo realizations.




**Response Fig. R5.** Histogram of MTCE cluster size for the (a) WP and (b) NA basins during 1980–2022. Cluster size is defined as the peak number of concurrent TCs within each MTCE cluster. The x-axis shows the peak concurrent TC count, and the y-axis shows the probability. Results are shown for IBTrACS (black), IBTrACS with uniformly sampled genesis dates (grey), the SPIN ensemble mean (red), SPIN with uniformly sampled genesis dates (pink), JL23 downscaling driven by ERA5 environmental fields with uniformly sampled genesis dates (yellow), and JL23 downscaling driven by NeuralGCM environmental fields with uniformly sampled genesis dates (blue). Error bars indicate the 10th–90th percentile range. For IBTrACS, the range is estimated by year bootstrapping over 1980–2022. For SPIN, it represents the spread across ensemble members. For the "uniform" experiments, the range is estimated across 1000 Monte Carlo realizations.

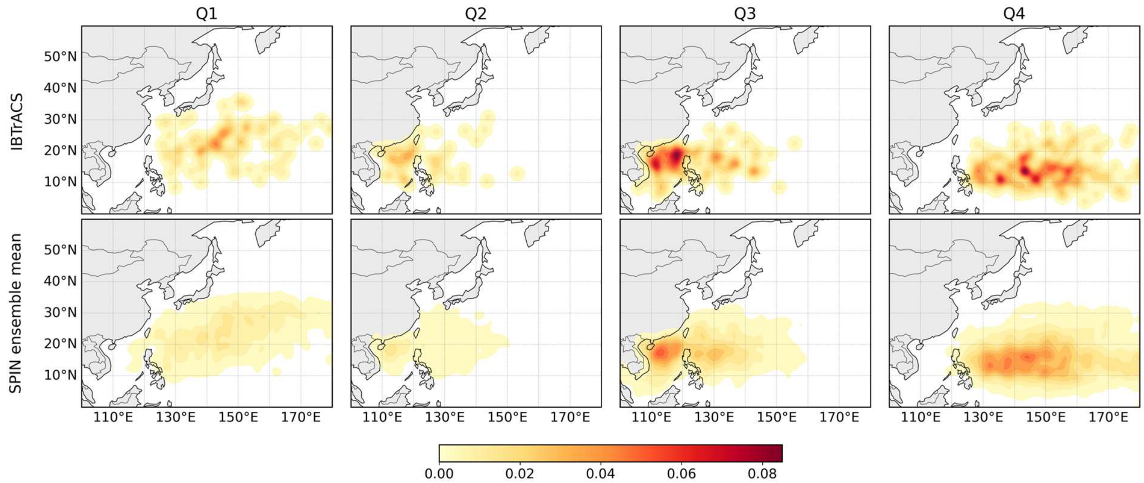
10. Figure 11: Do STCs forming in these quadrants have a geographic or seasonal preference?

***Response:***

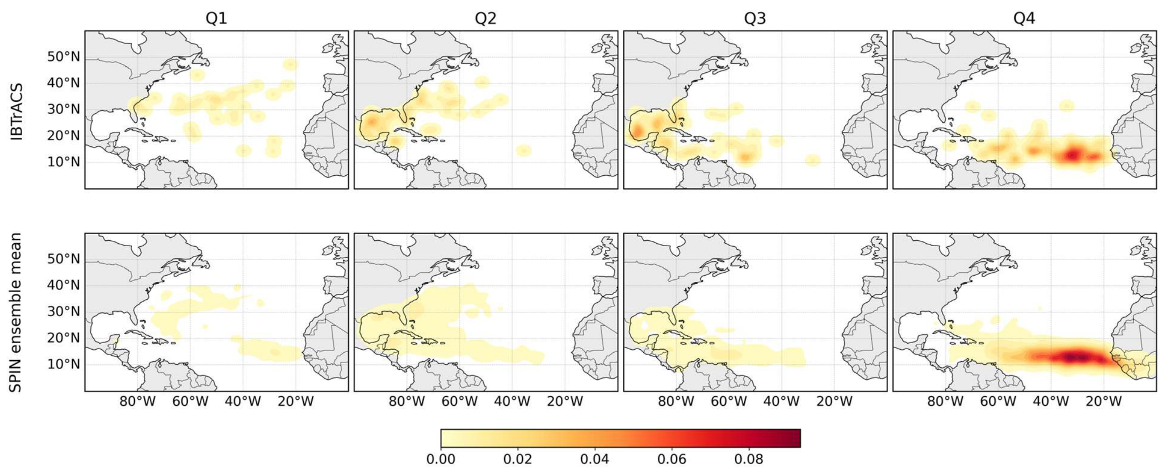
Thank you for your questions. Yes, subsequent TCs forming in different quadrants exhibit distinct geographic preferences (Response Figs. R6 and R7). In both the WP and NA basins, Q1 and especially Q4 STCs are distributed more broadly along the main genesis belt, whereas Q2 and Q3 STCs are more confined to the western basin. SPIN captures this observed quadrant-dependent geographic organization of STC genesis.

The pattern likely reflects a combination of TC propagation (Fu et al., 2025), TC-induced Rossby-wave dispersion (Ritchie and Holland, 1999), and basin-scale environmental controls (Ito and Yamauchi, 2026). In the WP and NA basins, TCs typically move westward, north-westward, or poleward under the influence of the subtropical high and the beta effect. If subsequent TCs still preferentially form within the low-latitude main genesis belt, they naturally tend to appear to the east or southeast of the parent TC in a TC-relative framework. This may explain the broader geographic coverage of Q1 and Q4 STCs. TC-induced Rossby-wave dispersion may further contribute to the Q4 preference, whereas the more confined Q2 and Q3 distributions may be influenced by environmental controls. These interpretations remain hypotheses. A robust attribution is beyond the scope of this study but may be worth investigating in future work.

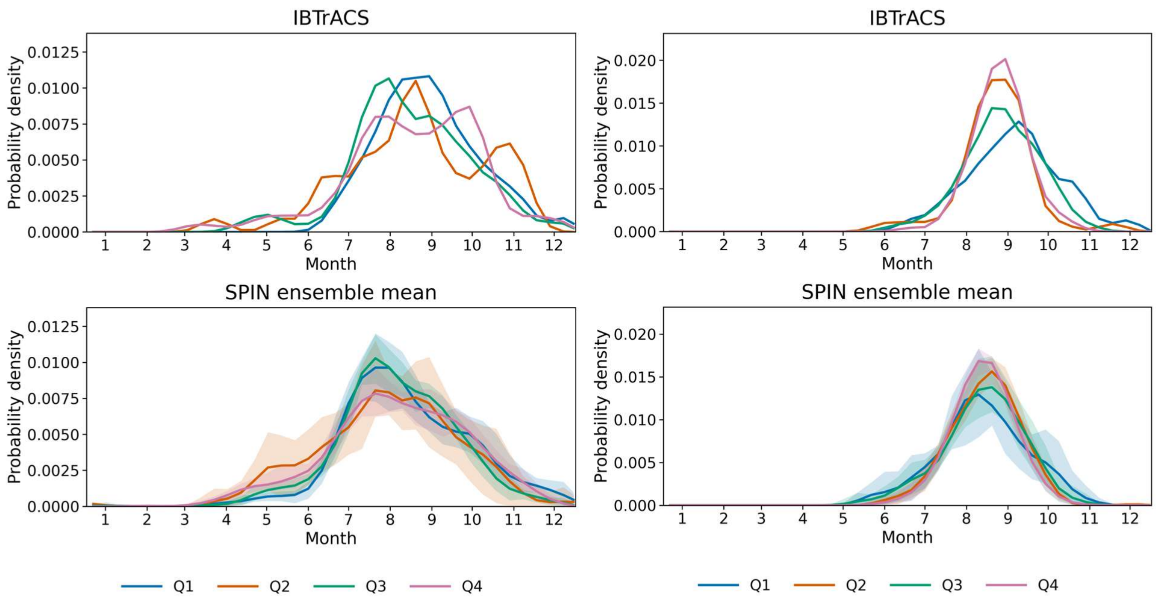
Seasonal differences among quadrants are present but modest (Response Figure R8). In both basins, STCs in all four quadrants peak within the hurricane season. In the WP, STC genesis peaks from late summer to early autumn across all quadrants. In the NA, Q2, Q3, and Q4 peak earlier, from August to September, whereas Q1 peak later and has a broader distribution. SPIN captures these basin-dependent seasonal features, although its ensemble mean is smoother than IBTrACS.




**Response Fig. R6.** Genesis density maps of subsequent TCs in MTCEs over the WP basin during 1980–2022. The upper row shows IBTrACS, and the lower row shows the SPIN ensemble mean. Q1–Q4 denote the quadrants in which subsequent TCs form relative to their parent TCs.




**Response Fig. R7.** Genesis density maps of subsequent TCs in MTCEs over the NA basin during 1980–2022. The upper row shows IBTrACS, and the lower row shows the SPIN ensemble mean. Q1–Q4 denote the quadrants in which subsequent TCs form relative to their parent TCs.




**Response Fig. R8.** Seasonal distribution of subsequent TC genesis in MTCEs over the WP (left) and NA (right) basin during 1980–2022. Curves show the probability density of genesis date for STCs in each quadrant, with each quadrant normalized independently. The upper panel shows IBTrACS, and the lower panel shows the SPIN ensemble mean. Shading denotes the 10–90% ensemble range for SPIN.

**References:**

Cheng, K.-Y., Harris, L., Bretherton, C., Merlis, T. M., Bolot, M., Zhou, L., Kaltenbaugh, A., Clark, S., and Fueglistaler, S.: Impact of Warmer Sea Surface Temperature on the Global Pattern of Intense Convection: Insights From a Global Storm Resolving Model, Geophysical Research Letters, 49, e2022GL099796, https://doi.org/10.1029/2022GL099796, 2022.

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., Klöwer, M., Lottes, J., Rasp, S., Düben, P., Hatfield, S., Battaglia, P., Sanchez-Gonzalez, A., Willson, M., Brenner, M. P., and Hoyer, S.: Neural General Circulation Models for Weather and Climate, Nature, 632, 1060–1066, https://doi.org/10.1038/s41586-024-07744-y, 2024.

Jonathan Lin, Chia-Ying Lee, Suzana Camargo et al. The Response of Tropical Cyclone Hazard to Natural and Forced Warming Patterns, 21 October 2024, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-5248169/v1]

Camargo, S. J., Emanuel, K. A., and Sobel, A. H.: Use of a Genesis Potential Index to Diagnose ENSO Effects on Tropical Cyclone Genesis, https://doi.org/10.1175/JCLI4282.1, 2007.

Hsieh, T.-L., Vecchi, G. A., Yang, W., Held, I. M., and Garner, S. T.: Large-scale control on the frequency of tropical cyclones and seeds: a consistent relationship across a hierarchy of global atmospheric models, Clim. Dyn., 55, 3177–3196, https://doi.org/10.1007/s00382-020-05446-5, 2020.

Emanuel, K.: Tropical Cyclone Seeds, Transition Probabilities, and Genesis, https://doi.org/10.1175/JCLI-D-21-0922.1, 2022.

Wang, B. and Murakami, H.: Dynamic genesis potential index for diagnosing present-day and future global tropical cyclone genesis, Environ. Res. Lett., 15, 114008, https://doi.org/10.1088/1748-9326/abbb01, 2020.

Bieli, M., Sobel, A. H., Camargo, S. J., and Tippett, M. K.: A Statistical Model to Predict the Extratropical Transition of Tropical Cyclones, Weather and Forecasting, 35, 451–466, https://doi.org/10.1175/WAF-D-19-0045.1, 2020.

Evans, C., Wood, K. M., Aberson, S. D., Archambault, H. M., Milrad, S. M., Bosart, L. F., Corbosiero, K. L., Davis, C. A., Pinto, J. R. D., Doyle, J., Fogarty, C., Galarneau, T. J., Grams, C. M., Griffin, K. S., Gyakum, J., Hart, R. E., Kitabatake, N., Lentink, H. S., McTaggart-Cowan, R., Perrie, W., Quinting, J. F. D., Reynolds, C. A., Riemer, M., Ritchie, E. A., Sun, Y., and Zhang, F.: The Extratropical Transition of Tropical Cyclones. Part I: Cyclone Evolution and Direct Impacts, Monthly Weather Review, 145, 4317–4344, https://doi.org/10.1175/MWR-D-17-0027.1, 2017.

Hart, R. E.: A Cyclone Phase Space Derived from Thermal Wind and Thermal Asymmetry, Monthly Weather Review, 131, 585–616, https://doi.org/10.1175/1520-0493(2003)131%3C0585:ACPSDF%3E2.0.CO;2, 2003.

Gao, J. and Li, T.: Factors Controlling Multiple Tropical Cyclone Events in the Western North Pacific*, Mon. Weather Rev., 139, 885–894, https://doi.org/10.1175/2010MWR3340.1, 2011.

Ito, K. and Yamauchi, K.: Limited Influence of Pre-Existing Tropical Cyclones on Subsequent Cyclogenesis in the Western North Pacific, J. Geophys. Res. Atmospheres, 131, e2025JD044793, https://doi.org/10.1029/2025JD044793, 2026.

Ritchie, E. A. and Holland, G. J.: Large-Scale Patterns Associated with Tropical Cyclogenesis in the Western Pacific, 1999.

Yoshida, R. and Ishikawa, H.: Environmental Factors Contributing to Tropical Cyclone Genesis over the Western North Pacific, https://doi.org/10.1175/MWR-D-11-00309.1, 2013.

Duan, S., Zhang, J., Bonfils, C., and Pallotta, G.: Testing NeuralGCM's capability to simulate future heatwaves based on the 2021 Pacific Northwest heatwave event, Npj Clim. Atmospheric Sci., 8, 251, https://doi.org/10.1038/s41612-025-01137-2, 2025.