Dear Editor and Anonymous Reviewer

We use text **in blue** to respond to the review comments and text **in green** to show revised text. Please review our response below:

Reviewer#1

General comments

This manuscript investigates the potential of applying Seasonal Climate Outlooks to localized precipitation forecasting by compiling 0.5-month lead time NOAA CPC seasonal precipitation tercile probability forecasts for two study basins. Specifically, these forecasts are incorporated as conditioning information in a local-scale stochastic weather generator to produce precipitation ensembles. Through the evaluation, two non-parametric ensemble generation methods are found to be suitable for different seasonal regimes, and the predictive skill is shown to be associated with large-scale climate signals (i.e., ENSO).

Overall, the manuscript is of good quality, with a clear and easy-to-follow structure, and it addresses a scientifically promising question using an innovative approach. What is more important, the proposed framework and the associated findings are generalizable, and it provides a guidance for further exploiting seasonal precipitation forecasts at the local scale. In my opinion, the manuscript is suitable for publication following minor revisions. My concerns mainly relate to the clarity of the presentation and the way the results are presented, as detailed in the specific comments below.

Specific comments

1. The authors appear to overemphasize hydrological modeling, as references to it appear in the abstract, introduction, and conclusion. While precipitation forecasts are indeed important for hydrological modeling, the methods and results presented in this study do not directly involve any hydrological model. I therefore suggest either removing these references or adding a dedicated discussion section to assess the potential impact of such precipitation forecasts on hydrological modeling, which would help improve the overall logical flow of the manuscript. If such a discussion is added, I also believe it would be highly beneficial to include some additional commentary on the predictability of precipitation.

   Response:
   We thank the reviewer for this important comment. We agree that the primary contributions of this study are the evaluation of localized seasonal precipitation forecasts and the development of forecast-conditioned rainfall ensemble generation methods, rather than the direct application of a hydrologic model. In response, we revised the manuscript to remove the reference that stated the hydrologic-modeling component, particularly in the introduction and conclusion section.

   Specifically, the Introduction (original Lines 68–70) previously included the statement: If forecast skill improves, probabilistic information could be integrated into existing hydrological models to anticipate anomalously wet or dry periods and adjust operations proactively." Similarly, the Conclusion (original Lines 388–389) stated: "This adaptive approach is expected to improve the accuracy of hydrologic simulation, e.g., streamflow forecasts, and subsequently benefit water resources management, which will be the focus of an extended study." Both statements have been removed from the revised manuscript. The framing now correctly positions hydrological application as a logical future extension of this work, rather than an implicit component of the present study

2. The figures in this manuscript warrant further revision, particularly regarding font size and color differentiation. For example, in Figure 1, the font size for the latitude/longitude labels and basin names is

too small; in Figure 2, the text is so small that it is difficult to read; and in Figures 3, 5, and 6, the colors and marker styles lack sufficient contrast, making it hard to distinguish different elements.

Response:

We thank the reviewer for this helpful suggestion. In response, we revised the figures throughout the manuscript to improve readability and overall presentation quality. Specifically, font sizes were increased where necessary, labels were made clearer, and the color and marker schemes in the relevant figures were adjusted.

In the revised manuscript, Figure 1 was updated by increasing the font size of the latitude/longitude labels and basin names. The overall clarity of the study area map was also improved.

In Figure 2, the flowchart was revised to improve readability and streamline the presentation of the methodology. Specifically, the text size was increased, wording was simplified, the previous CDF sorting and tercile-partitioning steps were consolidated into a single component, and the three sampling approaches were reformatted and color-coded to provide a clearer representation of the workflow.

In Figure 3, the font size of all text was increased. Lighter colors were used to represent above-normal and below-normal forecast events, and hatched patterns were added to distinguish the counts of above-normal and below-normal hits.

In Figure 4, the font size was increased, and the color scheme for above-normal and below-normal categories was made consistent with Figure 3.

In Figure 5, the text size was increased, and the boxplot format was revised. Solid black lines were used to represent above-normal forecast events, and dashed black lines were used to represent below-normal forecast events. The marker size of the ENSO points was also increased. In addition, the overall size of the text was improved, and the panel arrangement was changed from horizontal to vertical to improve readability.

In Figure 6, the text size was increased, the markers representing the three methods were made solid for better visibility, and the statistics previously shown in the figure were removed to produce a cleaner presentation

In Figure 7, the marker styles were also revised to make them consistent with Figure 6.

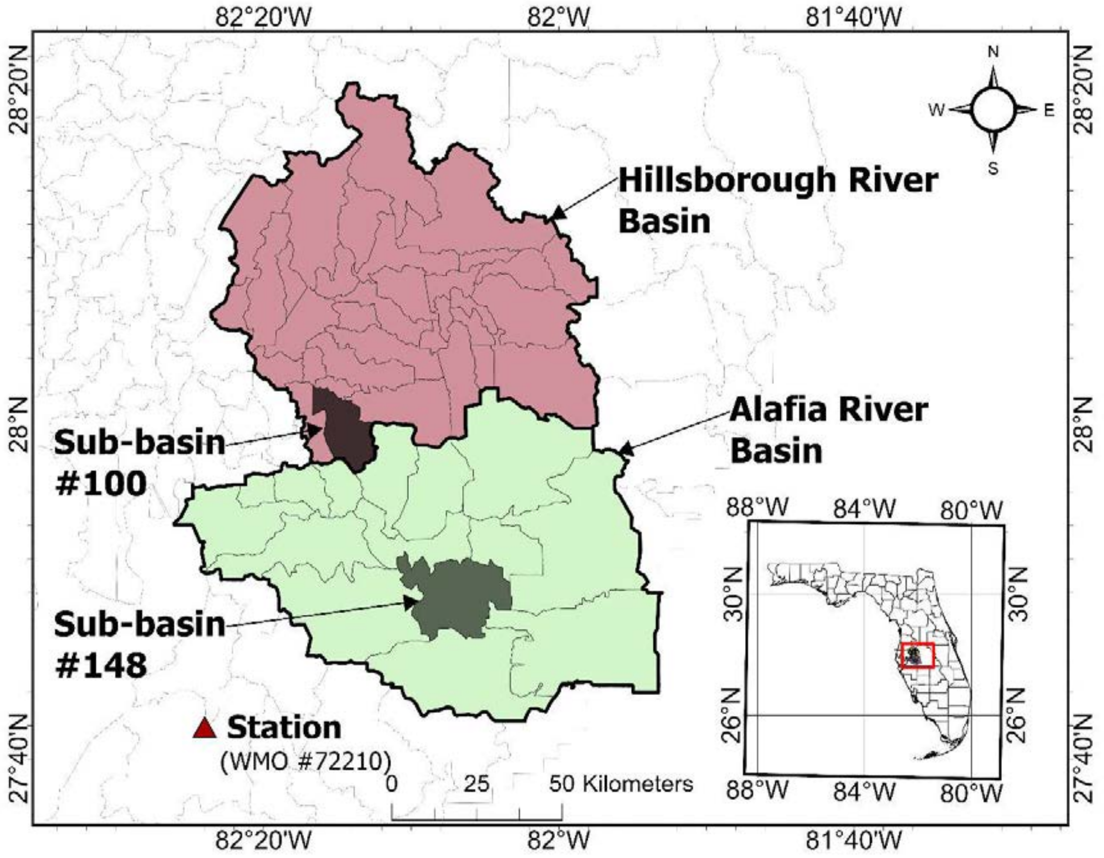

Figure 1: Hillsborough River Basin and Alafia River Basin in west-central Florida, with sub-basins #100, #148 and forecast station (WMO #72210) highlighted.

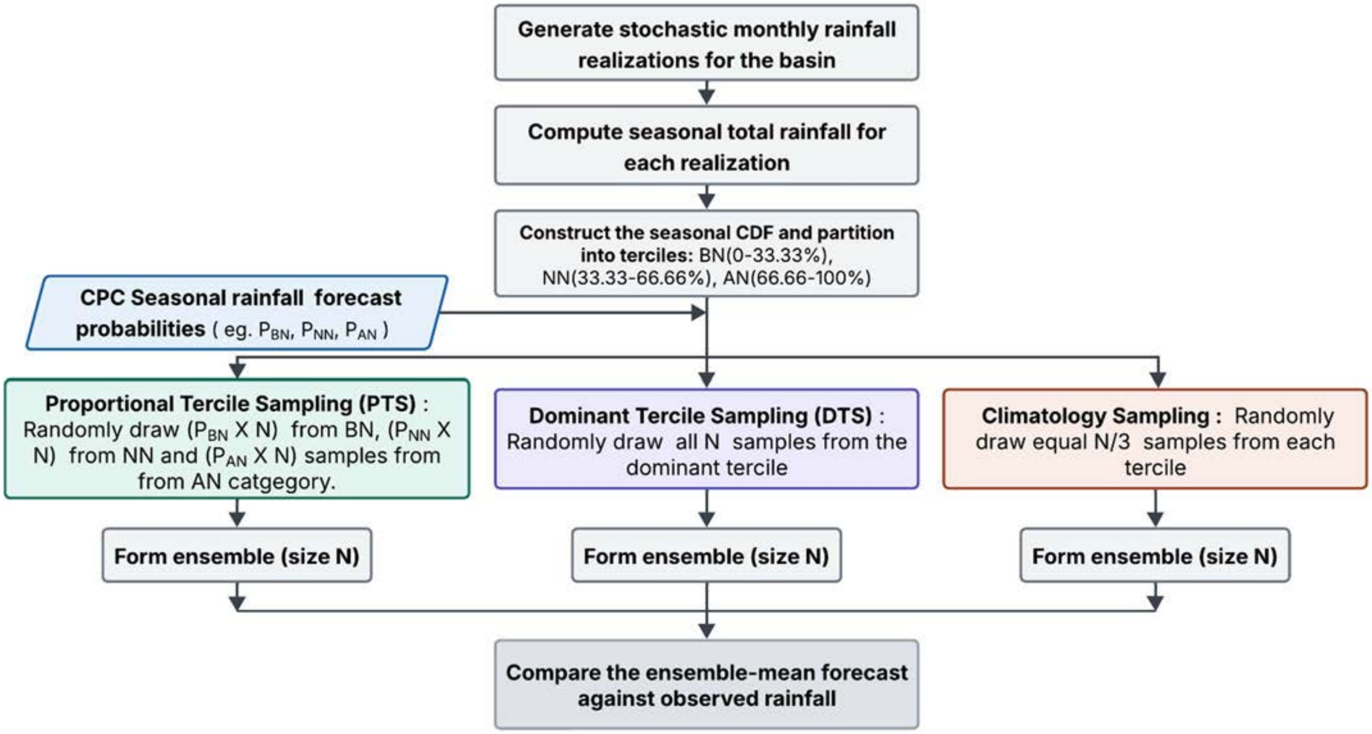

Figure 2: Schematic diagram illustrating the process of generating ensemble forecasts informed by seasonal climate outlooks.

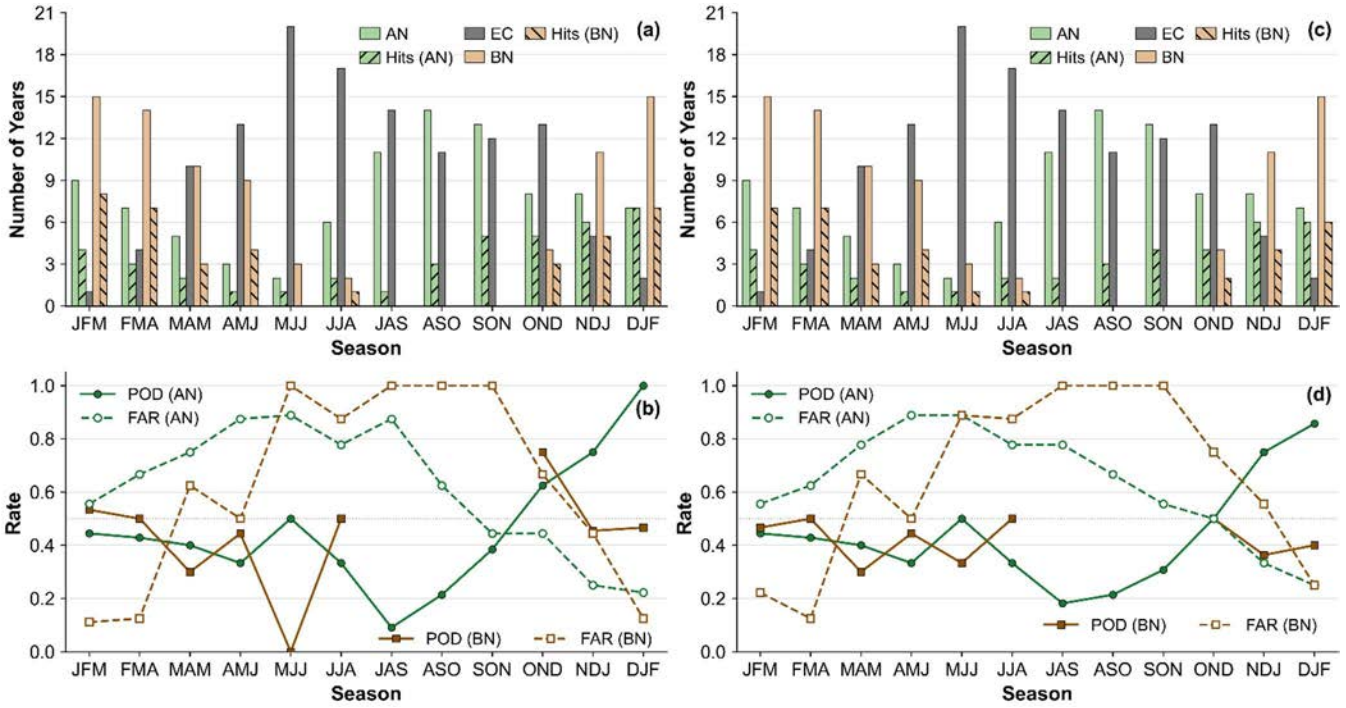

Figure 3: Seasonal distribution of categorical precipitation forecasts and associated performance metrics for the Alafia River Basin (a, b) and Hillsborough River Basin (c, d). Panels (a) and (c) display the number of years forecasted as above-normal (AN), below-normal (BN), and equal-chance (EC), along with the number of years correctly predicted as above-normal [Hits (AN)] and below-normal [Hits (BN)]. Panels (b) and (d) present the probability of detection (POD) and false alarm rate (FAR) for AN and BN categories across seasons from 1995-2019.

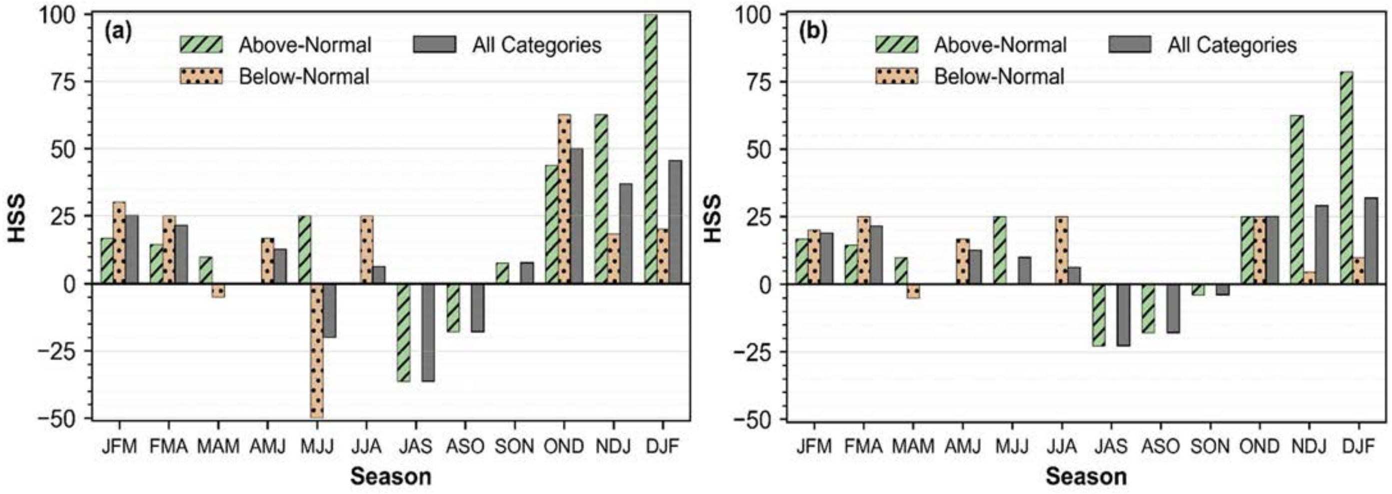

Figure 4: Heidke Skill Score (HSS) for CPC 0.5-month-lead seasonal precipitation forecasts in the Alafia River Basin (a) and Hillsborough River Basin (b), shown separately for above-normal, below-normal, and both forecast categories combined.

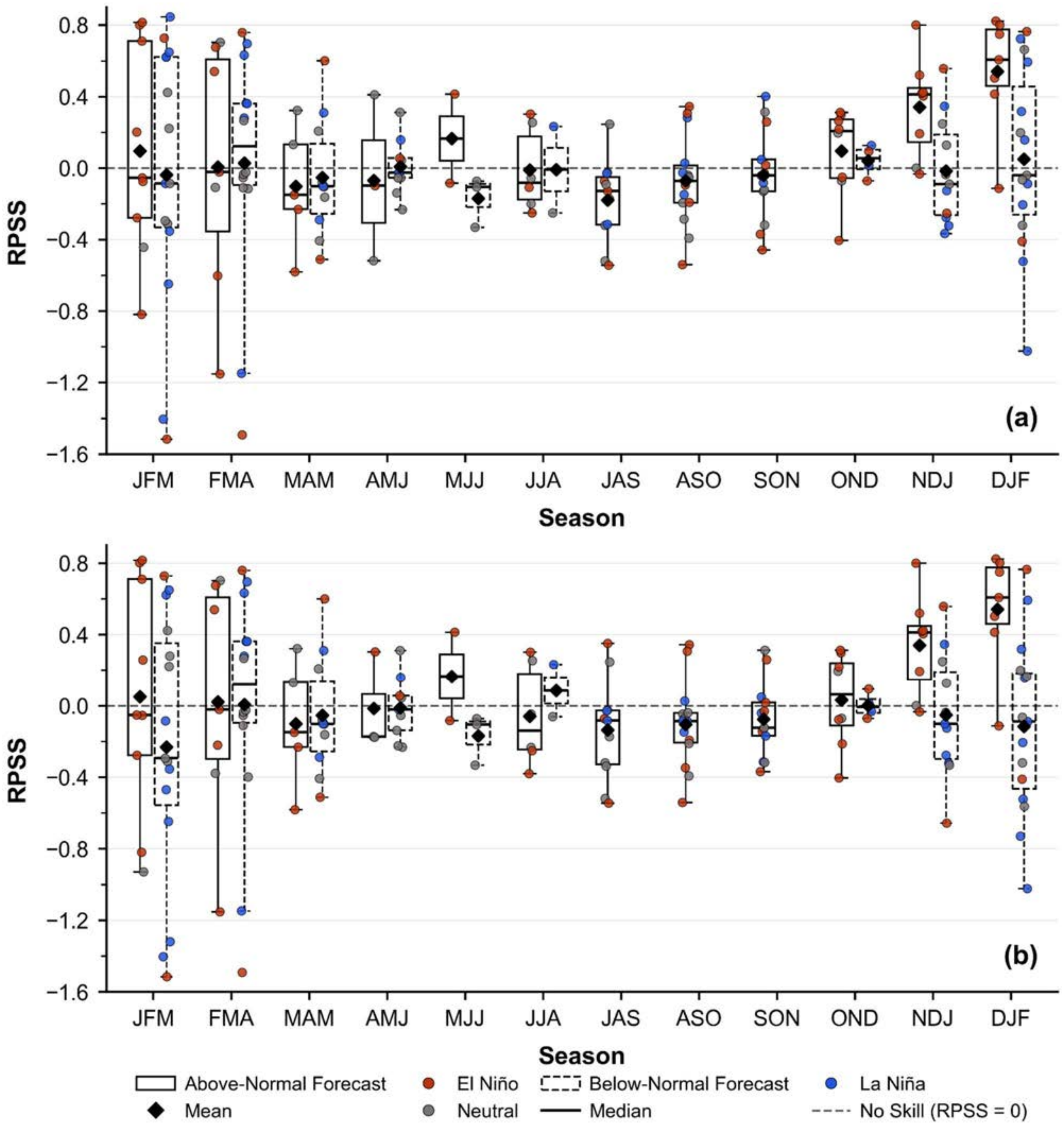

Figure 5: Ranked probability skill score (RPSS) for CPC 0.5-month-lead seasonal precipitation forecasts in the Alafia River Basin (a) and Hillsborough River Basin (b). Boxplots depict the seasonal distribution of RPSS for above-normal (AN) and below-normal (BN) forecast categories from 1995 to 2019. Individual points represent seasonal cases, color-coded by ENSO phase (El Niño, La Niña, Neutral). The dashed horizontal line represents the no-skill reference (RPSS = 0), indicating performance equivalent to climatology.

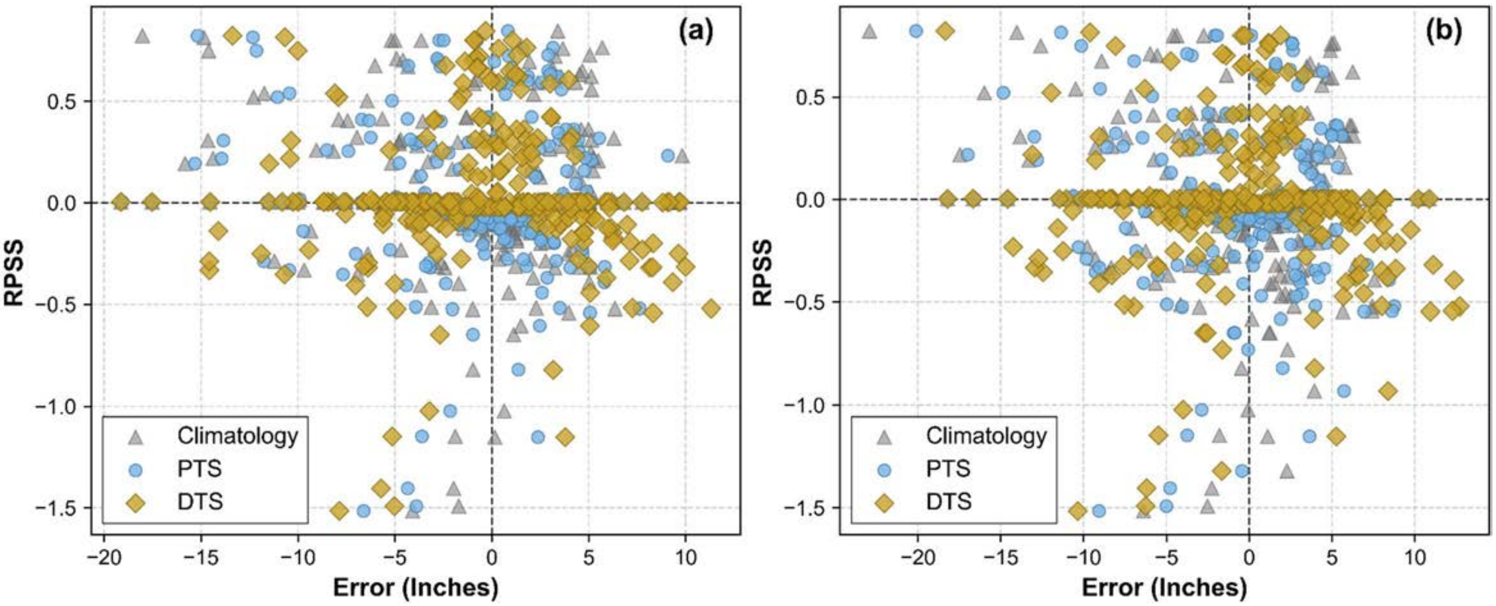

Figure 6: Scatterplot of forecast error (mean ensemble forecast minus observed precipitation, in inches) for each season from 1995 to 2019, corresponding to proportional tercile sampling (PTS), dominant tercile sampling (DTS), and climatology sampling. Errors are plotted against the ranked probability skill score (RPSS) of that season for sub-basin #148 (a) and sub-basin #100 (b).

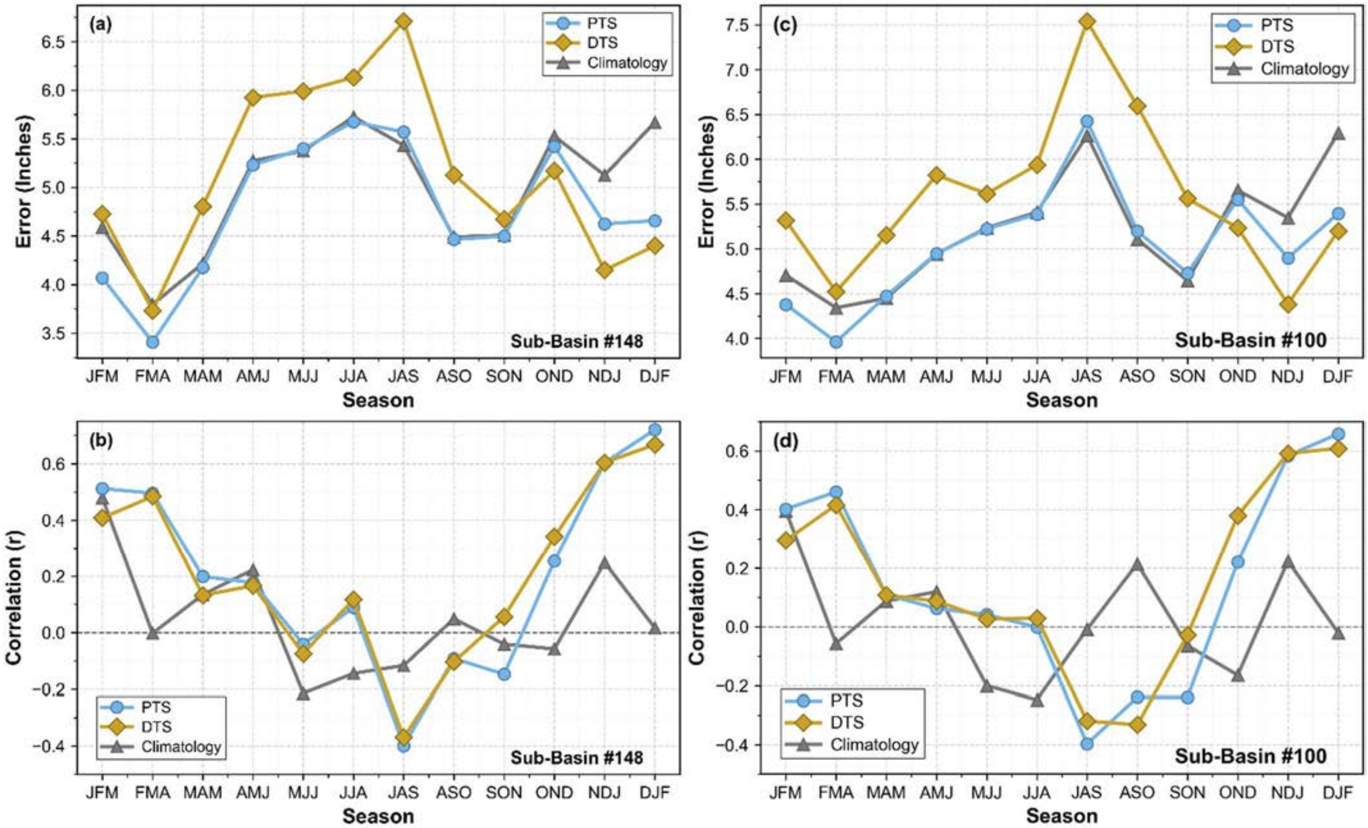

Figure 7: Seasonal variation in forecast performance for proportional tercile sampling (PTS), dominant tercile sampling (DTS), and climatology sampling for sub-basin #148 (a, b) and sub-basin #100 (c, d). Panels (a) and (c) display the root mean square error (RMSE, in inches) for each season, while panels (b) and (d) show the Pearson correlation coefficient (r) between ensemble forecasts and observed precipitation from 1995 to 2019.

3. The manuscript contains an excessive amount of numerical comparisons and descriptions, which would be more clearly presented in a table. This primarily concerns the section from Lines 346–354, as the current format makes it difficult to follow.

Response:

We thank the reviewer for this helpful recommendation. To improve clarity, we reorganized the quantitative comparisons of the sampling methods by summarizing the key results in Table 2. This table makes it easier to compare the performance of PTS, DTS, and climatology across the two sub-basins and under different forecast-skill conditions.

In the revised manuscript, we added Table 2:

Table 2. Summary of PTS, DTS, and climatology performance in sub-basins #148 and #100

Sub-basin / season set	**PTS**	**PTS**	**DTS**	**DTS**	**Climatology**	**Climatology**
	RMSE (in.)	r	RMSE (in.)	r	RMSE (in.)	r
#148 – All seasons	**4.81**	**0.81**	5.21	0.78	5.01	0.79
#148 – Skillful only (RPSS > 0)	5.83	0.78	**5.15**	**0.83**	6.25	0.73
#100 – All seasons	**5.08**	**0.81**	5.64	0.77	5.23	0.79
#100 – Skillful only (RPSS > 0)	6.11	0.79	**5.39**	**0.84**	6.53	0.75

We deleted the earlier lines explaining these statistics from Lines 370–379. In their place, we added the following interpretive text to summarize Table 2:

Across both sub-basins, a consistent performance pattern emerged for RMSE and correlation. When all seasons were included, the PTS approach produced only modest improvements, typically reducing RMSE by about 3–4% relative to climatology, while DTS actually performed worse than both methods. This indicates that PTS provides a relatively stable advantage under mixed-skill conditions. In contrast, when the analysis focused solely on skillful seasons (RPSS > 0), the DTS approach became clearly superior. DTS became clearly superior. DTS achieved RMSE reductions of approximately 12% relative to PTS and 17–18% relative to climatology, while PTS still outperformed climatology by about 7%. Correlation patterns were consistent with these findings, with DTS yielding the highest r values during skillful seasons, compared to PTS and climatology.

Taken together, these results highlight that the relative value of the two sampling strategies is strongly dependent on forecast quality: PTS offers greater stability when all seasons are considered, whereas DTS is markedly more effective during seasons with demonstrable predictive skill, improving both error magnitude and correspondence with observed variability.

4. The description of the observed precipitation data is not sufficiently clear. I suggest that the authors provide additional details in Section 2.2, specifically indicating which stations' precipitation observations were used to validate the forecasts. Additionally, the relative locations of the stations within each basin should be marked in a figure.

Response:

We thank the reviewer for highlighting this point. The manuscript has been revised to clearly specify the source of the observed precipitation data used for verification. We now explicitly describe the basin-scale observational dataset, its derivation, and provide the link to access it in the Data Availability section. In addition, Figure 1 has been updated to show the location of the forecast station used in the analysis.

In the revised Section 2.2, we added the following description of the precipitation data used in the study:

“Observed basin precipitation for the study basins was obtained from Tampa Bay Water’s Bayesian rainfall dataset for the Integrated Northern Tampa Bay (INTB) domain, which integrates rain-gauge observations and Doppler radar measurements within a Bayesian statistical framework and interpolates the resulting daily rainfall estimates to the basins (GSI Environmental Inc., 2021a,b,c). The resulting basin-scale daily precipitation record was then aggregated to seasonal totals for use in forecast verification.”

References:
GSI Environmental Inc. (2021a). Evaluation of Rain Gauge and Doppler Radar Data in the Integrated Northern Tampa Bay Hydrologic Model Domain. Prepared for Tampa Bay Water, Clearwater, Florida.
GSI Environmental Inc. (2021b). Integration of Rainfall Data From Rain Gauges and Doppler Radar in the Integrated Northern Tampa Bay Hydrologic Model Domain. Prepared for Tampa Bay Water, Clearwater, Florida.
GSI Environmental Inc. (2021c). Supplemental Report on Integration of Rainfall Data From Rain Gauges and Doppler Radar in the Integrated Northern Tampa Bay Hydrologic Model Domain for 2017 Through 2019. Prepared for Tampa Bay Water, Clearwater, Florida.

We also added the link to access this dataset in the Data Availability section: “https://tampabaywater.sharefile.com/share/view/f3376a6f07534256/focac193-6df5-48bf-b3c0-6f8e7614a15f”.

In the revised manuscript, Figure 1 was also updated to show the location of the forecast station used in the analysis, namely World Meteorological Organization Station #72210.

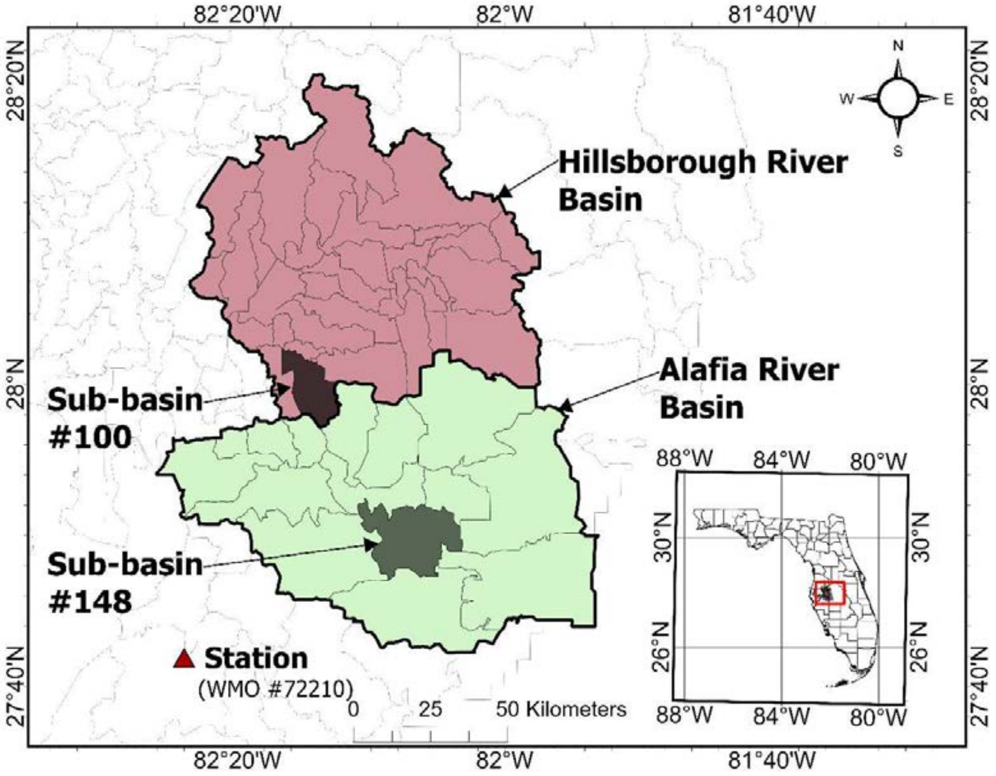

Figure 1: Hillsborough River Basin and Alafia River Basin in west-central Florida, with sub-basins #100, #148 and forecast station (WMO #72210) highlighted.

5. It is also recommended that the abstract includes a brief description of the observational data used to assess the NOAA precipitation outlooks.

   Response:
   We thank the reviewer for this helpful comment. The Abstract has been revised to include a concise description of the observational dataset used in the analysis, consistent with the clarification added in Section 2.2.

   The revised sentence in the Abstract (Lines 14–15) is as follows:
   Forecast performance is assessed seasonally using categorical and probabilistic metrics against basin-scale observed precipitation from a Bayesian gauge-radar dataset.

technical corrections

1. Line 245. There is a mistake in the equation

We thank the reviewer for identifying this error. The equation on Line 245 has been corrected in the revised manuscript. The corrected equation is updated in equation 17.

Dear Editor and Anonymous Reviewer

We use text **in blue** to respond to the review comments and text **in green** to show revised text. Please review our response below:

Reviewer#2

The manuscripts evaluates NOAA CPC 3-month precipitation outlooks skills at 0.5-month lead for two basin in west-central Florida basins over 25 years. It then proposes two simple non-parametric methods, Proportional Tercile Sampling (PTS) and Dominant Tercile Sampling (DTS), to translate CPC tercile probabilities into basin-scale ensemble rainfall realizations using an existing stochastic rainfall simulation dataset. The results show the strongest categorical and probabilistic skill during late fall and winter, particularly for above-normal precipitation forecast and in El Nino Years. A hybrid recommendation is suggested (DTS during high-skill seasons and PTS otherwise). Overall, the study offers a coherent and potentially transferable approach, but a few edits to terminology and methodology are needed.

Below are several specific comments

1. Line 65: The basin location mentioned as 'Southwest Florida' conflicts with the title. Please use one term consistently.

Response: We thank the reviewer for identifying this inconsistency. The terminology has been revised in the manuscript to ensure consistency with the title.

In the revised manuscript, "Southwest Florida" in the Introduction (Line 65) was changed to "west-central Florida.

2. The Equal Chances (EC) category is defined clearly; however, it is unclear how EC cases are handled in the categorical verification framework when building 2×2 tables (e.g., whether they are excluded or pooled into the "non-event" category). This should be clarified in the methodology.

Response: We thank the reviewer for this important comment. We agree that the treatment of Equal Chances (EC) cases in the categorical verification framework should be stated more explicitly. The methodology has been revised to clarify how EC forecasts are handled when constructing the 2×2 contingency tables for categorical verification and mentioned that EC cases were grouped into the non-event category.

In the revised manuscript, Section 2.5.1 was updated as follows: Separate contingency tables were constructed for wet events (above-normal) and dry events (below-normal) for each basin to enable a more detailed evaluation, with Equal Chances (EC) forecasts classified as non-events.

3. FAR is labeled as "False Alarm Rate" in the figure caption, but Equation (10) and the following text define FAR as "False Alarm Ratio." Consistent terminology for FAR is needed throughout.

Response: We thank the reviewer for noting this inconsistency. The terminology for FAR has been revised throughout the manuscript to ensure consistency. It is now referred to consistently as False Alarm Rate in the text, equations, and figure captions.

In the revised manuscript, Section 2.5.1 was revised accordingly, and the Figure 3 caption was also updated to use "false alarm rate (FAR)" consistently.

4. Equation (15) requires correction.

Response: We thank the reviewer for identifying this error. The typo mistake in Equation (15) has been corrected in the revised manuscript.

5. Section 2.4 describes how ensemble rainfall forecasts are generated using tercile probabilities, but it is unclear how the method is applied when the CPC outlook is classified as Equal Chances. The handling of EC cases in the ensemble-generation step should be stated explicitly.

Response: We thank the reviewer for this helpful comment. We agree that the treatment of Equal Chances (EC) cases in the ensemble-generation step should be stated explicitly. Section 2.4 has been revised to clarify how EC forecasts are handled in the generation of rainfall ensembles. For EC forecasts, realizations were sampled without tercile preference to reflect climatological conditions.

In the revised manuscript, Section 2.4.2 was updated by adding the sentence: For EC forecasts, realizations were sampled equally from the below-normal, near-normal, and above-normal tercile pools for each method.

6. The manuscript references the wrong equation: it states that the "above-normal" pool is defined by Equation (6), but Equation (6) defines the below-normal (BN) category. Please correct this citation.

Response: We thank the reviewer for identifying this incorrect equation reference. The citation has been corrected in the revised manuscript so that the above-normal pool now refers to the appropriate equation.

In the revised manuscript, Section 2.4.1 now states: "…50% of the ensemble members are randomly drawn from the pool defined by Equation (8), 30% from Equation (7), and 20% from Equation (6)."

7. Please correct the panel label in the Figure 5 caption: the Alafia River Basin panel should be labeled "(a)" (not "(b)").

Response: We thank the reviewer for noting this labeling error. The Figure 5 caption has been corrected in the revised manuscript.

In the revised manuscript, the Figure 5 caption still reads: "...in the Alafia River Basin (a) and Hillsborough River Basin (b)."

8. In the section on performance of sampling schemes, qualitative comparisons should be complemented (or replaced) with quantitative improvements (e.g., the decrease in RMSE of one sampling method relative to another).

Response: We thank the reviewer for this suggestion. We agree that quantitative comparisons provide a clearer basis for evaluating the relative performance of the sampling schemes. In response, we revised the corresponding section to include more explicit quantitative comparisons between PTS and DTS, including relative improvement in RMSE and correlation.

In the revised manuscript, Section 3.3 was expanded to include the following interpretation: Across both sub-basins, a consistent performance pattern emerged for RMSE and correlation. When all seasons were included, the PTS approach produced only modest improvements, typically reducing RMSE by approximately 3–4% relative to climatology, whereas DTS performed worse than both comparison methods. This finding suggests that PTS provides a relatively stable advantage under mixed-skill conditions. In contrast, when the analysis was restricted to skillful seasons (RPSS > 0), DTS became clearly superior. Under these conditions, DTS reduced RMSE by approximately 12% relative to PTS and by 17–18% relative to climatology, while PTS still outperformed climatology by about 7%. Correlation patterns were consistent with these results, with DTS yielding the highest correlation coefficients during skillful seasons, followed by PTS and climatology.

9. The typo in the data availability statement ("sis") should be corrected.

Response: We thank the reviewer for identifying this typographical error. The typo in the data availability statement has been corrected in the revised manuscript.

Dear Editor and Anonymous Reviewer

We use text **in blue** to respond to the review comments and text **in green** to show revised text. Please review our response below:

Climate forecasts generated by global climate models (GCMs) are valuable for hydrological modelling. In this paper, attention is paid to NOAA's seasonal precipitation forecasts. The skill at different lead times is evaluated. In general, the analysis is useful for West-Central Florida.

There are four comments for further improvements in the paper.

1. Besides the NOAA's forecasts, there are other sets of forecasts available in the NMME (). The authors are suggested to add more GCM forecasts and conduct a more comprehensive evaluation of GCM forecast skill. In the meantime, the multiple GCM forecasts under investigation can be summarized by a table.

   Response:
   We thank the reviewer for this valuable suggestion. We agree that the North American Multi-Model Ensemble (NMME) provides important context for seasonal hydroclimate forecasting and that individual GCM performance is highly relevant when working directly with raw model forecasts. However, the objective of the present study is to evaluate the skill of the NOAA CPC operational probabilistic tercile outlooks at the local basin scale and to develop a framework for translating those outlooks into basin-scale rainfall ensembles. Because CPC outlooks already capture information from multiple dynamical and statistical guidance sources, including NMME-based products, accessing the full intercomparison of the GCM members would therefore require a fundamentally different study design, including raw model output processing, bias correction, and a separate verification framework which lies beyond the scope of the present work.

   But we have meaningfully expanded our manuscript with the NMME literature in direct response to this comment. The literature consistently demonstrates that the NMME multimodel ensemble outperforms any individual constituent model, including in the southeastern United States where the study basins are located. Even among the stronger-performing individual models in this region, namely CCSM4 and CFSv2, the multimodel ensemble still exhibits superior skill (Wang, 2014; Slater et al., 2019). For probabilistic 3-month forecasts specifically, Becker and van den Dool (2016) demonstrated that the NMME generally provides more skillful and reliable probabilistic forecasts than the CFSv2 single-model baseline. So, our approach of working directly with CPC outlooks, which already synthesize NMME-based guidance and operationally practical starting point. A direct comparison of individual NMME member skill against CPC operational outlooks at the basin scale is identified as a valuable direction for future research.

   The following expanded discussion has been added to the Introduction of the revised manuscript:
   The NOAA CPC operationally maintains the NMME, a dynamic multimodel ensemble system comprising NCEP-CFSv2, ECCC-CanESM5, ECCC-GEM5.2-NEMO, NCAR-CESM1, NCAR-CCSM4, and NASA-GEOS-S2S-2 (NOAA CPC, n.d.), which provides more skillful probabilistic seasonal forecasts than single-model systems (Kirtman et al., 2014). Evaluation studies have shown that NMME precipitation skill varies substantially by season, region, and lead time, with the multimodel ensemble generally outperforming individual members, and models such as CCSM4 and CFSv2 showing stronger skill in the southeastern USA (Wang, 2014; Slater et al., 2019; Becker & van den Dool, 2016). When raw GCM precipitation fields are used directly rather than probabilistic outlooks, postprocessing such as quantile mapping is often applied to correct systematic biases, though its effectiveness depends on the forecast objective (Zhao et al., 2017).

In addition, the following sentence has been added to the Conclusion to contextualize the ENSO-driven skill patterns observed in this study within the broader NMME literature:
Peak forecast skill is closely aligned with El Niño winters, during which above-normal precipitation is more likely, a finding consistent with Infanti and Kirtman (2014), who documented elevated NMME skill during El Niño winters in the southeastern United States.

2. In the evaluation of GCM forecasts, the authors are suggested to refer to Slater et al. (2019) that presented a comprehensive of the NMME forecasts across the continental USA.

   Response:
   We thank the reviewer for this helpful recommendation. We have added Slater et al. (2019) to the revised Introduction section, where it appears alongside Wang (2014) and Becker and van den Dool (2016) in support of the expanded NMME discussion described in our response to Comment 1.

   In the revised manuscript, the following text was added in the Introduction:
   Evaluation studies have shown that NMME precipitation skill varies substantially by season, region, and lead time, with the multimodel ensemble generally outperforming individual members, and models such as CCSM4 and CFSv2 showing stronger skill in the southeastern USA (Wang, 2014; Slater et al., 2019; Becker & van den Dool, 2016).

3. Given the existence of spatiotemporal biases in GCM forecasts, forecast post-processing plays a key part in exploiting the potential raw GCM forecasts. The quantile mapping is probably the simplest yet effective method to use.

   Response:
   We agree that post-processing is often important when working with raw quantitative GCM forecasts and appreciate this suggestion. The present study, however, does not work directly with raw GCM precipitation fields. Instead, we use NOAA CPC operational tercile probability outlooks, which are categorical probabilistic products that summarize expected deviations from climatological norms. Because these outlooks have already been subject to multi-model synthesis and expert post-processing within the CPC framework, applying quantile mapping, which requires a continuous quantitative output is not directly applicable in our workflow.

   We have nevertheless incorporated this important methodological context into the revised manuscript, acknowledging quantile mapping as a relevant technique when working with raw GCM outputs, and identifying it explicitly as a direction for future work.

   In the revised manuscript, the following sentence was added in the Introduction:
   When raw GCM precipitation fields are used directly rather than probabilistic outlooks, postprocessing such as quantile mapping is often applied to correct systematic biases, though its effectiveness depends on the forecast objective (Zhao et al., 2017).

4. The Oceanic Niño Index (ONI) is mentioned in the paper. There are statistical forecasts produced from hydroclimatic teleconnections such as ONI. Is it possible to perform some comparisons of GCM forecasts with statistical forecasts?

   Response:
   We thank the reviewer for this insightful suggestion. While a full ONI-based statistical forecast model is beyond the scope of the current study, we note that our existing ENSO-phase stratification in Section 3.2 provides a meaningful partial comparison. The strong alignment between El Niño conditions and above-normal CPC forecasts in NDJ and DJF suggests that CPC's GCM-based system successfully captures the dominant ONI teleconnection signal. In this sense, CPC's forecasts implicitly encode the ENSO-

precipitation relationship that an ONI-based statistical model would seek to exploit. This interpretation is consistent with Infanti and Kirtman (2014), who demonstrated that the NMME effectively leverages ENSO-driven teleconnections to improve skill in the southeastern United States during El Niño winters.

References

Becker, E., & Van Den Dool, H. (2016). Probabilistic Seasonal Forecasts in the North American Multimodel Ensemble: A Baseline Skill Assessment. Journal of Climate, 29(8), 3015–3026. https://doi.org/10.1175/JCLI-D-14-00862.1

NOAA Climate Prediction Center. (n.d.). North American Multi-Model Ensemble (NMME). National Oceanic and Atmospheric Administration. Retrieved March 16, 2026, from https://www.cpc.ncep.noaa.gov/products/NMME/NMME_description.html

Infanti JM, Kirtman BP (2014) Southeastern U.S. rainfall prediction in the North American multi-model ensemble. J Hydrometeorol 15(2):529–550. doi:10.1175/JHM-D-13-072.1

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., Van Den Dool, H., Saha, S., Mendez, M. P., Becker, E., Peng, P., Tripp, P., Huang, J., DeWitt, D. G., Tippett, M. K., Barnston, A. G., Li, S., Rosati, A., Schubert, S. D., … Wood, E. F. (2014). The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. Bulletin of the American Meteorological Society, 95(4), 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1

Slater, L. J., Villarini, G., & Bradley, A. A. (2019). Evaluation of the skill of North-American Multi-Model Ensemble (NMME) Global Climate Models in predicting average and extreme precipitation and temperature over the continental USA. Climate Dynamics, 53(12), 7381–7396. https://doi.org/10.1007/s00382-016-3286-1

Wang, H. (2014). Evaluation of monthly precipitation forecasting skill of the National Multi-model Ensemble in the summer season: FORECASTING SKILL OF THE NATIONAL MULTI-MODEL ENSEMBLE. Hydrological Processes, 28(15), 4472–4486. https://doi.org/10.1002/hyp.9957