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.

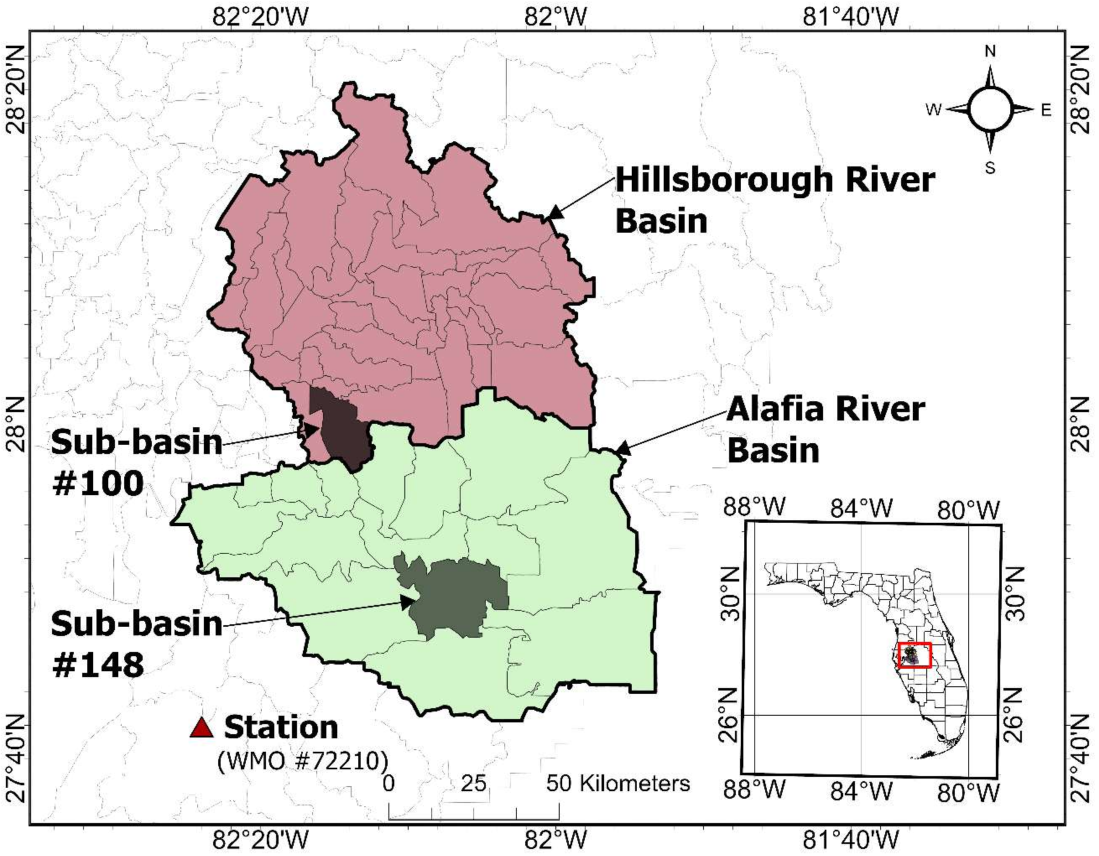

Figure 1: Hillsborough River Basin and Alafia River Basin in west-central Florida, with sub-basins #100, #148 and forecast station (WMO #72210) highlighted.

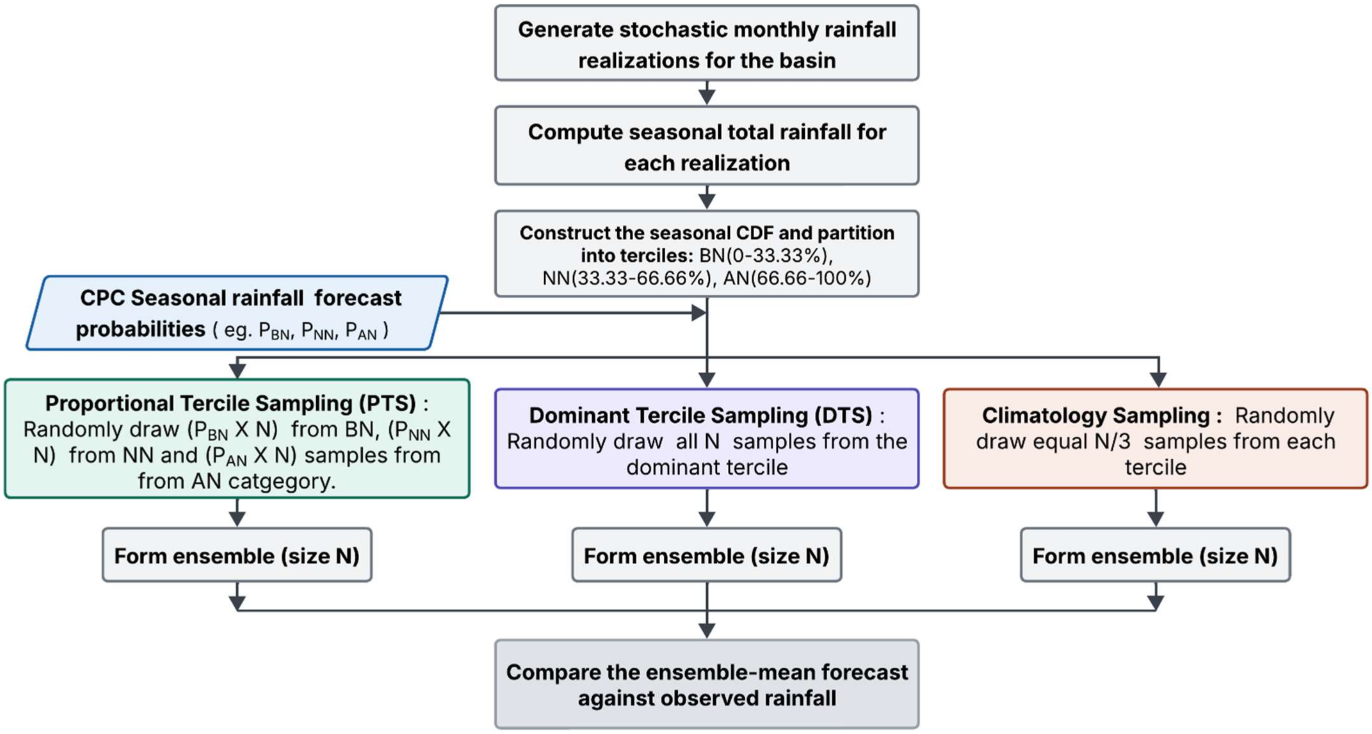

Figure 2: Schematic diagram illustrating the process of generating ensemble forecasts informed by seasonal climate outlooks.

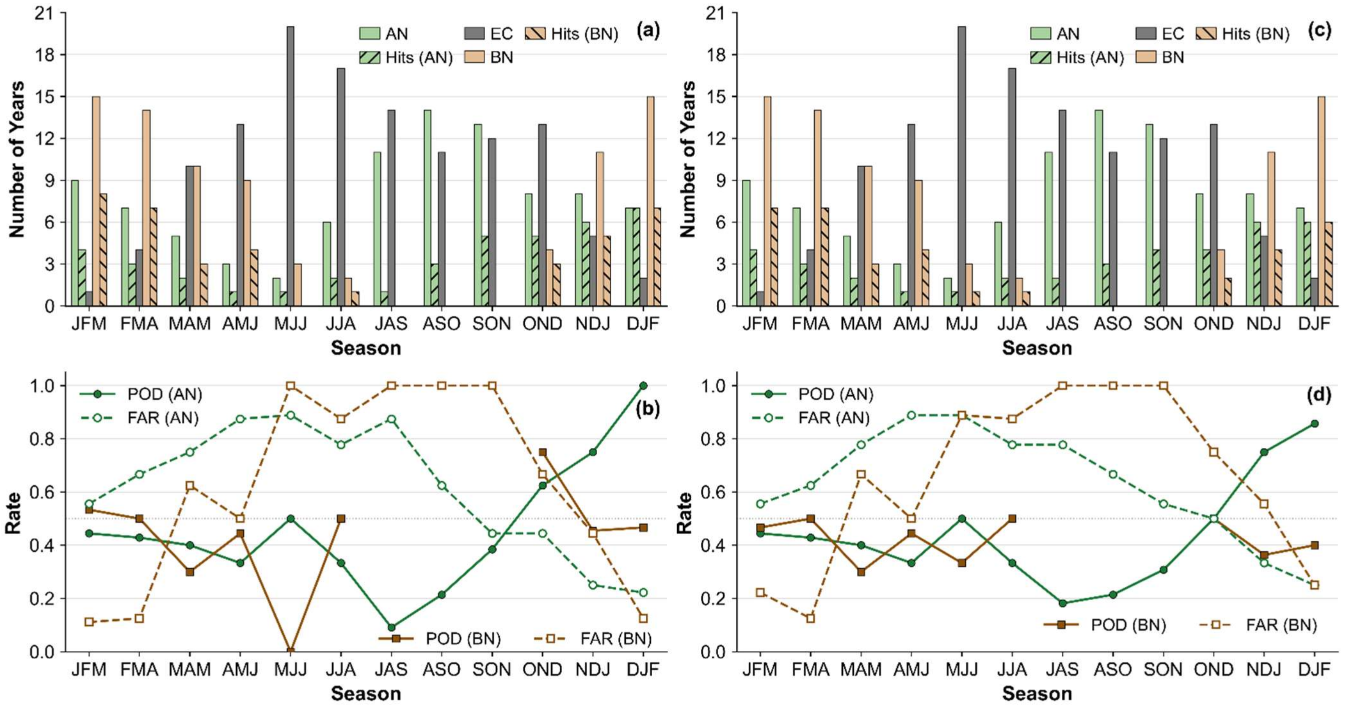

Figure 3: Seasonal distribution of categorical precipitation forecasts and associated performance metrics for the Alafia River Basin (a, b) and Hillsborough River Basin (c, d). Panels (a) and (c) display the number of years forecasted as above-normal (AN), below-normal (BN), and equal-chance (EC), along 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.

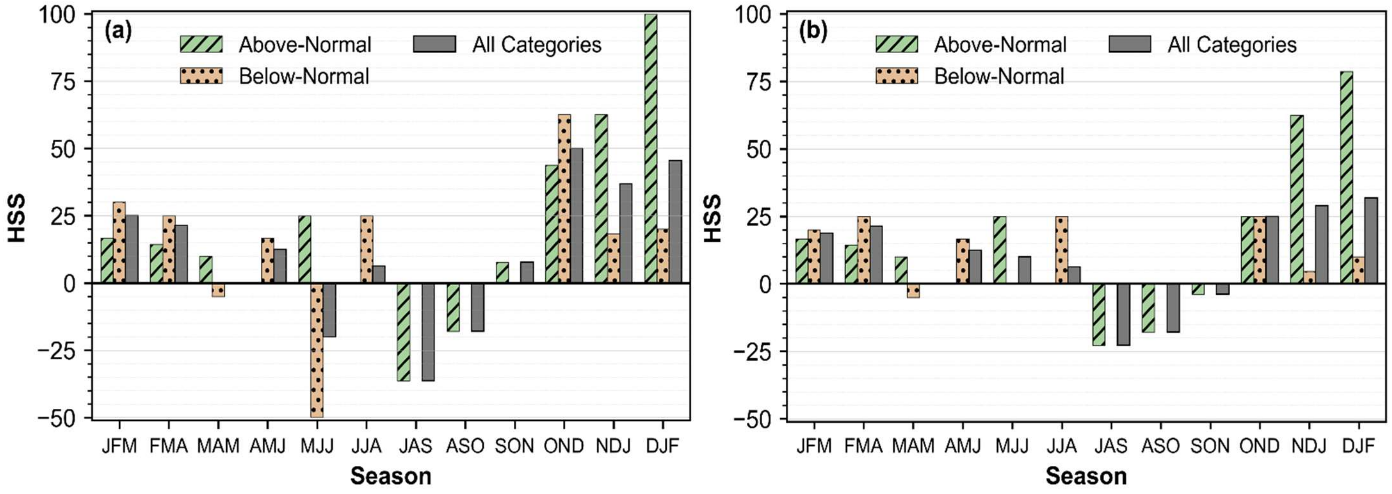

Figure 4: Heidke Skill Score (HSS) for CPC 0.5-month-lead seasonal precipitation forecasts in the Alafia River Basin (a) and Hillsborough River Basin (b), shown separately for above-normal, below-normal, and both forecast categories combined.

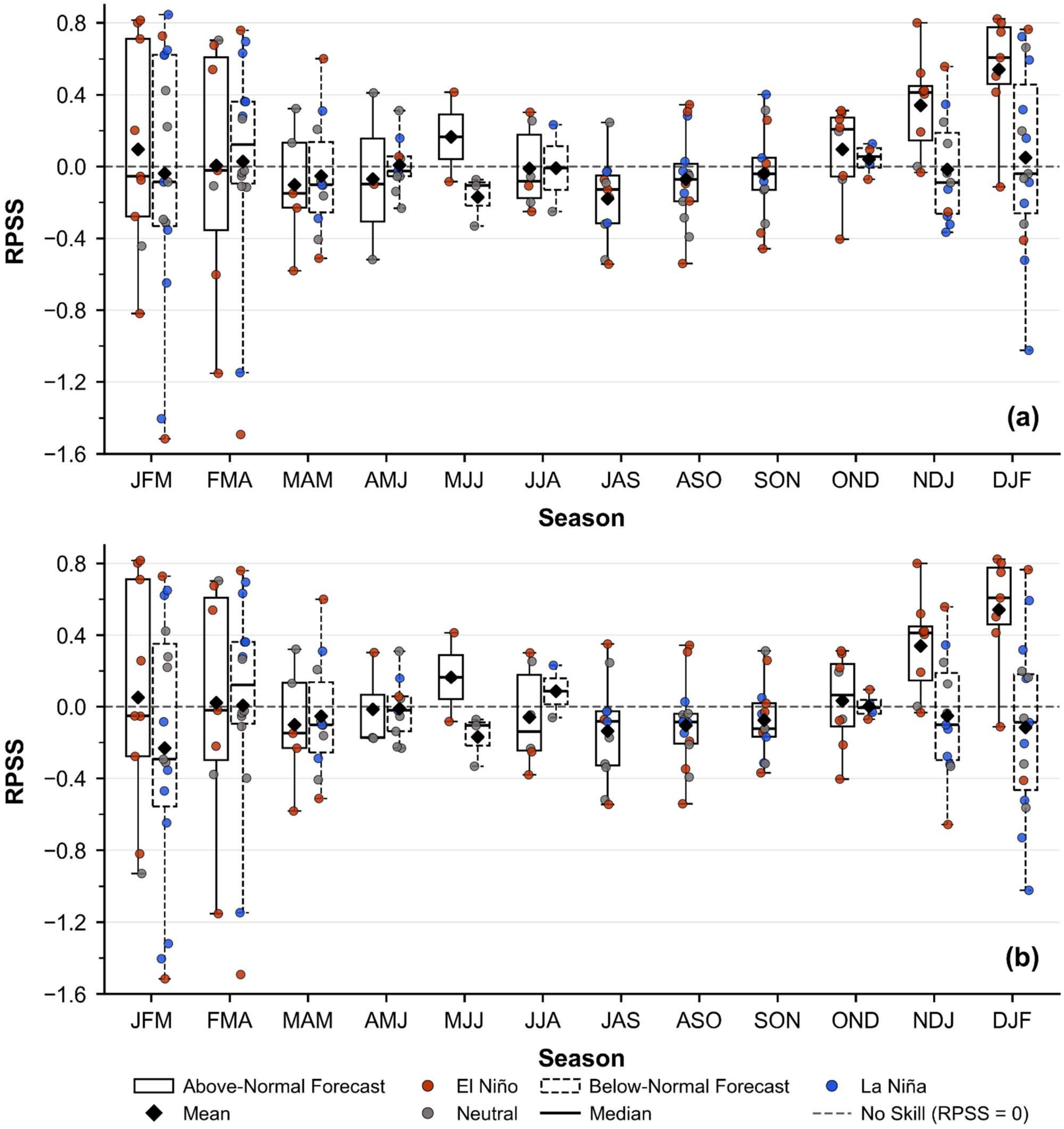

Figure 5: Ranked probability skill score (RPSS) for CPC 0.5-month-lead seasonal precipitation forecasts in the Alafia River Basin (a) and Hillsborough River Basin (b). Boxplots depict the seasonal distribution of RPSS for above-normal (AN) and below-normal (BN) forecast categories from 1995 to 2019. 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.

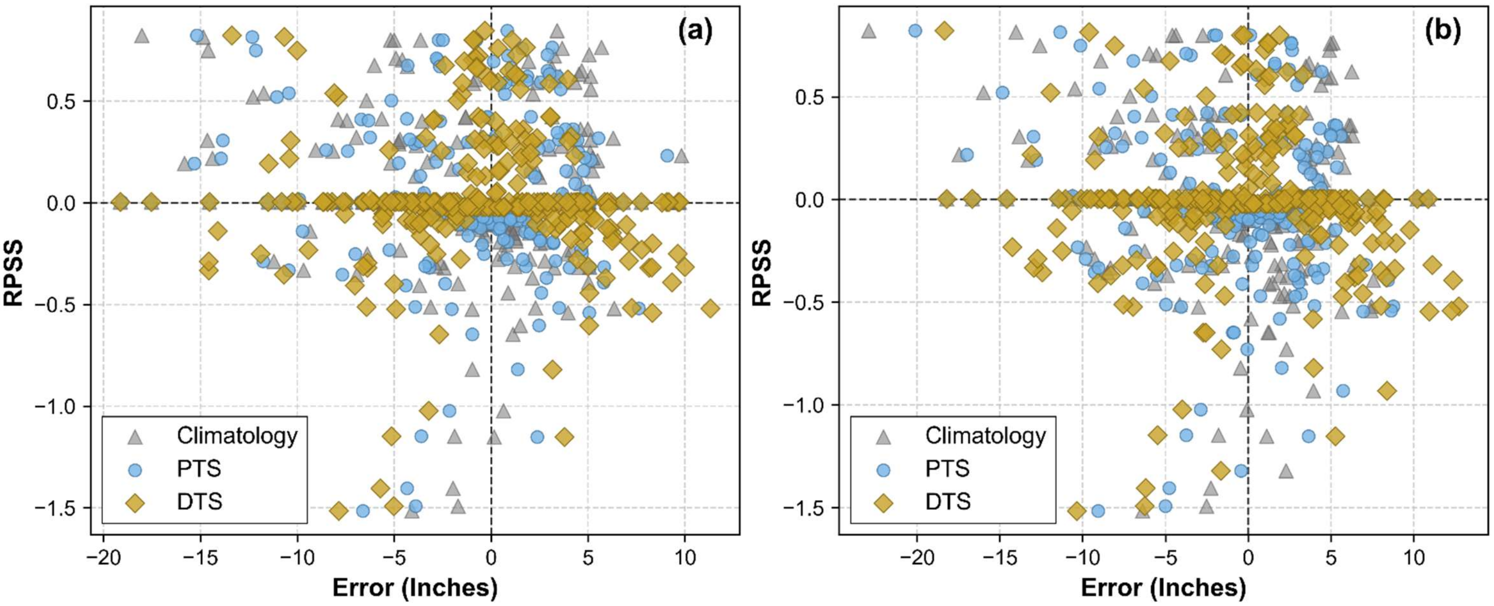

Figure 6: Scatterplot of forecast error (mean ensemble forecast minus observed precipitation, in inches) for each season from 1995 to 2019, corresponding to proportional tercile sampling (PTS), dominant tercile sampling (DTS), and climatology sampling. Errors are plotted against the ranked probability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

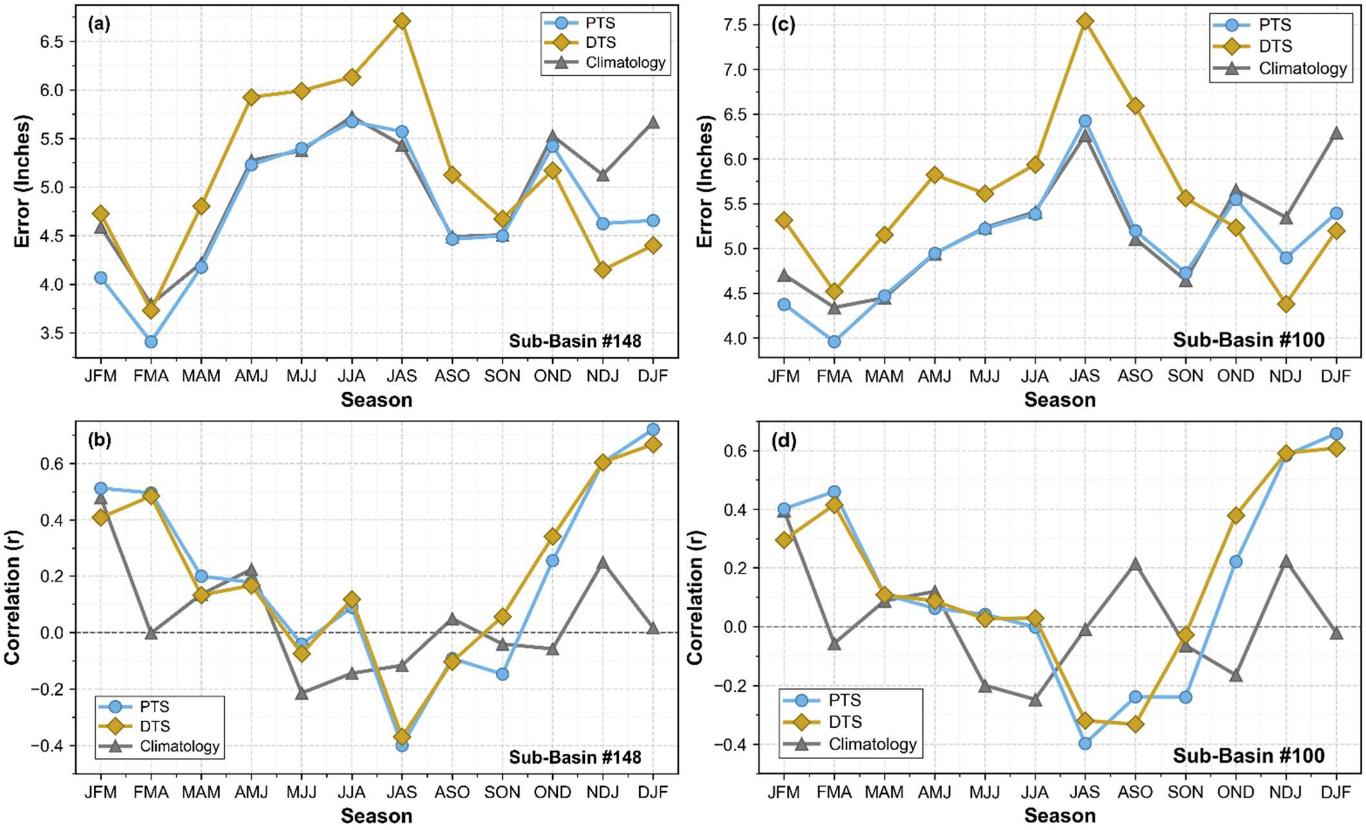

Figure 7: Seasonal variation in forecast performance for proportional tercile sampling (PTS), dominant tercile sampling (DTS), and climatology sampling for sub-basin #148 (a, b) and sub-basin #100 (c, d). Panels (a) and (c) display the root mean square error (RMSE, in inches) for each season, while panels (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.

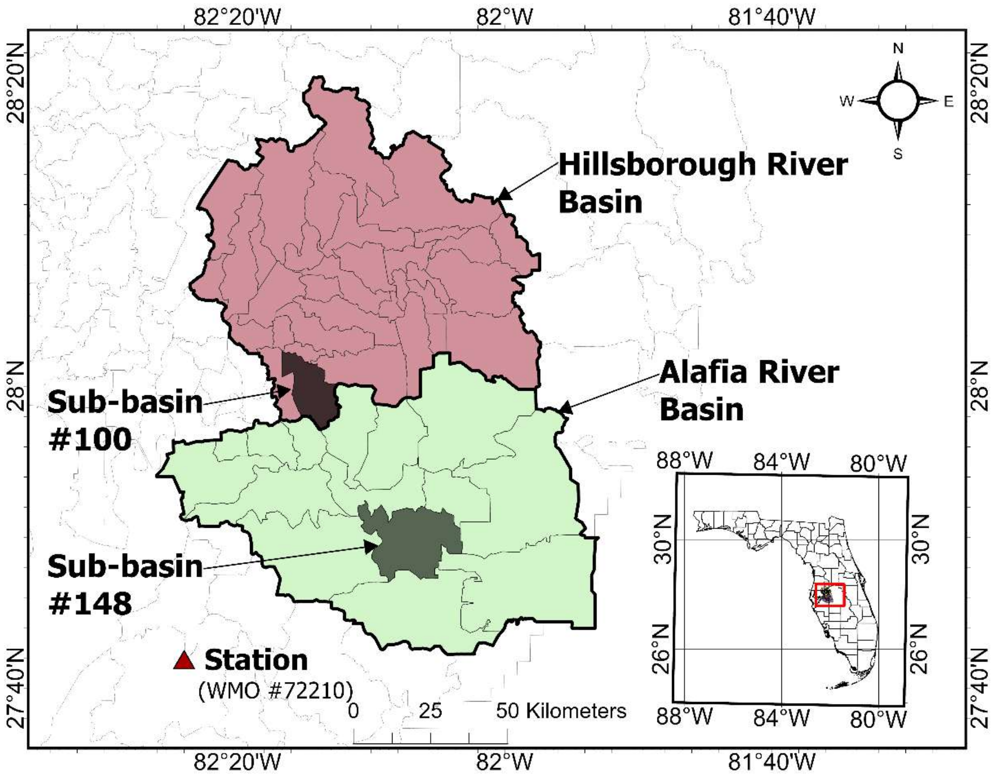

Figure 1: Hillsborough River Basin and Alafia River Basin in west-central Florida, with sub-basins #100, #148 and forecast station (WMO #72210) highlighted.

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.