

Validation of ERA5-Land and NCEP/NCAR Reanalysis-Based Precipitation Data over the Sichuan-Yunnan Region Using Satellite Observations

Zongze Li^{1,a,*}

*¹School of Atmospheric Sciences, Nanjing University of Information Science & Technology, No. 219
Ningliu Road, Pukou District, Nanjing, Jiangsu, China
a. 2352010763@qq.com*

**corresponding author*

Abstract: This study evaluates the applicability of ERA5-Land and NCEP/NCAR reanalysis precipitation data in the Sichuan-Yunnan region by comparing them with GPM-IMERG satellite observations from 1998 to 2023. The analysis focuses on seasonal variations in precipitation across Sichuan-Yunnan Region and its surrounding areas. Results indicate that both datasets effectively capture the seasonal precipitation patterns observed in IMERG, with precipitation peaking in summer, followed by autumn and spring, and reaching its lowest levels in winter. On average, summer precipitation is approximately seven times higher than in winter. Over the full study period (1998–2023), both datasets tend to overestimate precipitation compared to satellite observations. However, from 2010 to 2023, ERA5-Land exhibited smaller biases than NCEP/NCAR, whereas from 1998 to 2010, their biases were relatively similar. In terms of spatial distribution, ERA5-Land generally overestimates precipitation, particularly in the western Hengduan Mountains and in parts of Sichuan and Yunnan, where summer and autumn precipitation exceed 1000 mm. Meanwhile, NCEP/NCAR consistently identifies a high-precipitation center (peaking at around 500 mm) in Sichuan Province, with significant precipitation extending eastward into Sichuan and western Yunnan. During winter and spring, this high-precipitation center shifts further east, reaching southern Yunnan. Notably, in summer and autumn, NCEP/NCAR underestimates precipitation by approximately 650 mm in areas west of Sichuan and east of Yunnan.

Keywords: GPM-IMERG, NCEP/NCAR Reanalysis, ERA5-Land Reanalysis, Precipitation

1. Introduction

Precipitation is a crucial component of the global hydrological cycle, significantly impacting both natural ecosystems and human societies [1,2]. However, traditional ground-based precipitation data often suffer from uneven spatial distribution [3], limited long-term continuity [4], and discrepancies introduced by variations in in-situ gauge measurements across different stations [5]. While satellite observations provide broader coverage, they typically have lower temporal resolution. Consequently, for weather and climate analyses, reanalysis datasets are essential, as they offer greater homogeneity and longer time series [6].

Given the inherently non-linear characteristics of precipitation, it is crucial to evaluate whether reanalysis data can accurately capture its spatial and temporal distribution [7–10]. Among commonly used reanalysis datasets, ERA5-Land, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), and NCEP/NCAR, jointly developed by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR), have been widely assessed in various regions. Guo et al. [11] demonstrated that ERA5-Land effectively represents precipitation trends and spatial distributions in the Yellow River Basin, including extreme precipitation events. Li et al. [12] found that ERA5-Land accurately reproduces the temporal and spatial variations in monthly average precipitation across the Qilian Mountains and the Tibetan Plateau. Similarly, Fu et al. [13] showed that NCEP/NCAR reanalysis data can reliably capture regional precipitation characteristics in Australia.

The Sichuan-Yunnan region of China, encompassing the Sichuan Basin, the Yunnan-Guizhou Plateau, and the Hengduan Mountains, experiences significant temporal and spatial variability in precipitation. Huang et al. [14] analyzed ERA5-Land precipitation data from 2018 to 2020 and concluded that it aligns well with station observations, effectively capturing overall precipitation characteristics across five southwestern Chinese provinces. Gao et al. [15] examined ERA5 precipitation data from 1979 to 2018 in Anhui Province and found that it accurately reflected seasonal precipitation patterns, with summer experiencing peak rainfall, spring receiving more precipitation than autumn, and winter having the least. Their study also confirmed that seasonal averages of precipitation rates can reliably reflect regional climate characteristics.

This study uses GPM-IMERG satellite observation data as a benchmark to analyze the temporal and spatial variability of precipitation in the Sichuan-Yunnan region, evaluating the applicability of ERA5-Land and NCEP/NCAR reanalysis datasets in this area. The structure of the paper is organized as follows: Section 2 describes the data and methodology, Section 3 presents and analyzes the results, and Section 4 offers a summary of the key findings.

2. Data and Method

2.1. Data Sources

This study utilizes three main datasets: GPM-IMERG satellite observations, ERA5-Land reanalysis data, and NCEP/NCAR reanalysis data.

- 1) Produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), ERA5-Land (ERA5L) is a global land surface reanalysis dataset spanning over four decades from 1981 to the present. It has a spatial resolution of 9 km and a temporal resolution of 1 hour. In this study, monthly average ERA5L reanalysis data from January 1998 to December 2023 were used, resampled to a spatial resolution of $0.1^\circ \times 0.1^\circ$.
- 2) Developed by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) in the United States, NCEP/NCAR Reanalysis Dataset is a comprehensive reanalysis dataset. This study utilizes monthly average NCEP/NCAR reanalysis precipitation data from January 1998 to December 2023, with a spatial resolution of $2.5^\circ \times 2.5^\circ$.
- 3) GPM-IMERG (IMERG) Data is a satellite-based precipitation product developed by NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency) for global precipitation monitoring and estimation. It has a temporal resolution of 30 minutes and a spatial resolution of 10 km. For this study, the Level-3 product GPM_3IMERGM (GPM IMERG Final Precipitation L3, 1-month, $0.1^\circ \times 0.1^\circ$, V07) was selected, with a final temporal resolution of 1 month and a spatial resolution of $0.1^\circ \times 0.1^\circ$.

2.2. Data Processing and Methodology

The study period spans from 1998 to 2023. Seasons are categorized as follows: spring encompasses March, April, and May; summer includes June, July, and August; autumn covers September, October, and November; and winter consists of December, January, and February. The region of interest (ROI) includes Sichuan and Yunnan provinces and their surrounding areas, covering a latitude range of 20°N–35°N and a longitude range of 95°E–110°E. The precipitation values from all three datasets are standardized to millimeters per day (mm/day).

To ensure comparability, both the ERA5L and IMERG datasets were interpolated to match the grid points of the NCEP/NCAR dataset using the nearest-neighbor method. Subsequently, various analytical techniques, including synthetic analysis, correlation coefficient calculations, and significance testing, were applied to systematically compare the temporal and spatial distribution characteristics of precipitation across ERA5L, NCEP/NCAR reanalysis data, and satellite observations.

3. Results and Discussion

3.1. Precipitation Variation Characteristics in the Study Area (1998–2023)

Figure 1 presents the annual and seasonal multi-year average precipitation values derived from reanalysis and observational datasets in the study area from 1998 to 2023, along with the relative biases of the two reanalysis datasets compared to the observational data.

As shown in Figure 1(a), the multi-year average precipitation recorded by the IMERG dataset is lowest in winter, at 64 mm. Spring precipitation is slightly higher, at 135 mm, followed by a significant increase in summer, which records the highest precipitation at 580 mm. Autumn precipitation is 447 mm, slightly lower than in summer. The total annual precipitation amounts to 1,225 mm.

In comparison, the multi-year average precipitation recorded by the ERA5-Land reanalysis dataset is 116 mm in winter, 208 mm in spring, 731 mm in summer, and 555 mm in autumn, with a total annual precipitation of 1,610 mm. Meanwhile, the NCEP/NCAR reanalysis dataset records 139 mm in winter, 221 mm in spring, 724 mm in summer, and 598 mm in autumn, with a total annual precipitation of 1,683 mm.

Overall, both reanalysis datasets overestimate precipitation across all seasons compared to IMERG observations. However, both datasets adequately capture the general seasonal precipitation patterns in the study area, characterized by the following: lowest precipitation in winter, slightly higher precipitation in spring, peak precipitation in summer, and a slight decrease in autumn compared to summer. These findings are consistent with those of Huang et al. [14], who reported that ERA5-Land effectively represents the overall precipitation characteristics of Sichuan and Yunnan provinces.

The relative bias analysis in Figure 1(b) reveals that the ERA5-Land dataset exhibits relative biases of 45.20% in winter, 35.17% in spring, and 19.39% in autumn, all of which are lower than the corresponding biases in NCEP/NCAR reanalysis data (54.40%, 39.03%, and 25.28%). However, in summer, the relative bias of ERA5-Land is 20.68%, which is slightly higher than that of NCEP/NCAR (19.92%). For total annual precipitation, the relative bias of ERA5-Land is 23.87%, which is lower than NCEP/NCAR's 27.19%.

These results suggest that, overall, ERA5-Land provides a more accurate estimation of precipitation in the study area compared to NCEP/NCAR, particularly in non-summer seasons.

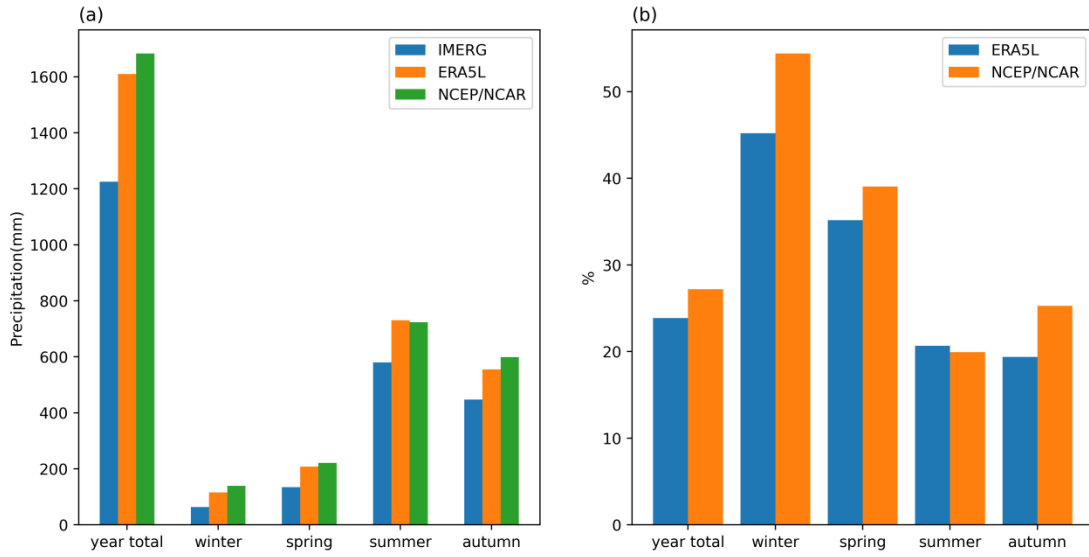


Figure 1: Multi-year seasonal average precipitation (mm) from 1998-2023. (a) annual and seasonal average from both reanalysis data and observations, (b) relative bias of ERA5-Land and NCEP/NCAR reanalysis to observations.

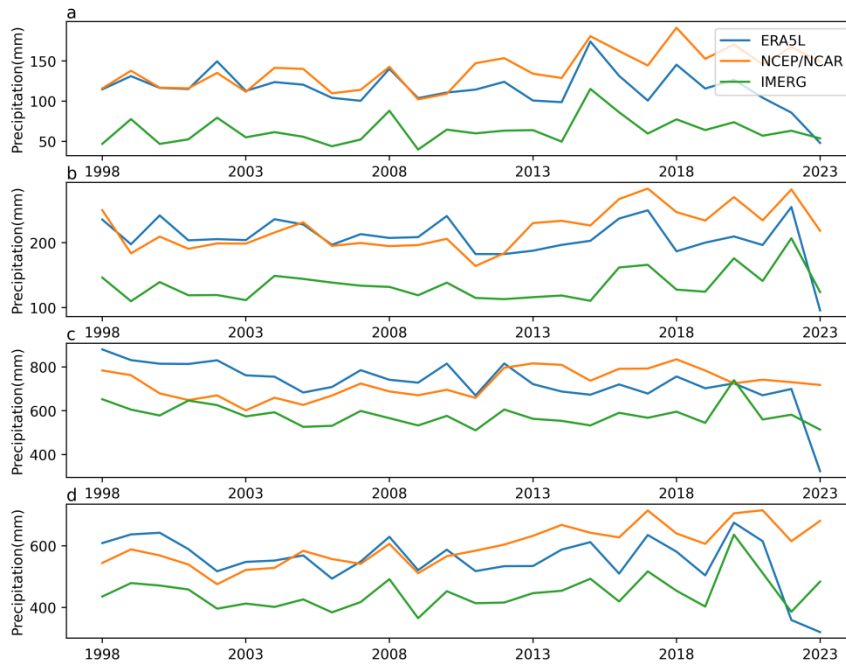


Figure 2: Long-term trends of regional seasonal average precipitation for both reanalysis datasets and observations. (a) winter, (b) spring, (c) summer, and (d) autumn.

The interannual variation trends of regional average precipitation for both reanalysis datasets and observational data across winter, spring, summer, and autumn in the study area are presented in Figure 2.

From Figures 2(a) and 2(b), it is evident that in winter and spring, the average precipitation from both reanalysis datasets between 1998 and 2023 is consistently higher than that recorded in the observational data—except for ERA5-Land, which slightly underestimates precipitation in winter 2023. Between 1998 and 2010, the differences in average precipitation between the two reanalysis

datasets were relatively small, remaining below 20 mm. However, from 2010 to 2023, NCEP/NCAR reanalysis data consistently exceeded ERA5-Land in terms of average precipitation.

From Figure 2(c), in summer, both reanalysis datasets also tend to overestimate precipitation compared to observations throughout 1998–2023—with the exception of ERA5-Land, which slightly underestimates precipitation in 2020 and 2023, and NCEP/NCAR, which slightly underestimates in 2020. Between 1998 and 2010, the average precipitation of NCEP/NCAR was consistently lower than that of ERA5-Land; however, from 2010 to 2023, NCEP/NCAR's precipitation values consistently exceeded those of ERA5-Land.

From Figure 2(d), in autumn, both reanalysis datasets generally report higher precipitation levels than the observational data from 1998 to 2023—with ERA5-Land slightly underestimating in 2022 and 2023. Between 1999 and 2005, NCEP/NCAR's average precipitation was consistently lower than ERA5-Land's, but from 2010 to 2023, NCEP/NCAR consistently reported higher precipitation values than ERA5-Land. In particular during 1998 to 2010 winter and spring, the biases of both reanalysis datasets were relatively close to the observational data. Yet in summer of 1998 to 2010, NCEP/NCAR exhibited smaller biases than ERA5-Land. While in autumn of 1998 to 2010, NCEP/NCAR showed smaller biases from 1998 to 2005, while from 2005 to 2010, both datasets had similar biases. However, during 2010–2023, ERA5-Land exhibited smaller biases than NCEP/NCAR across all seasons when compared to observational data.

3.2. Regional Distribution Characteristics of Precipitation in the Study Area

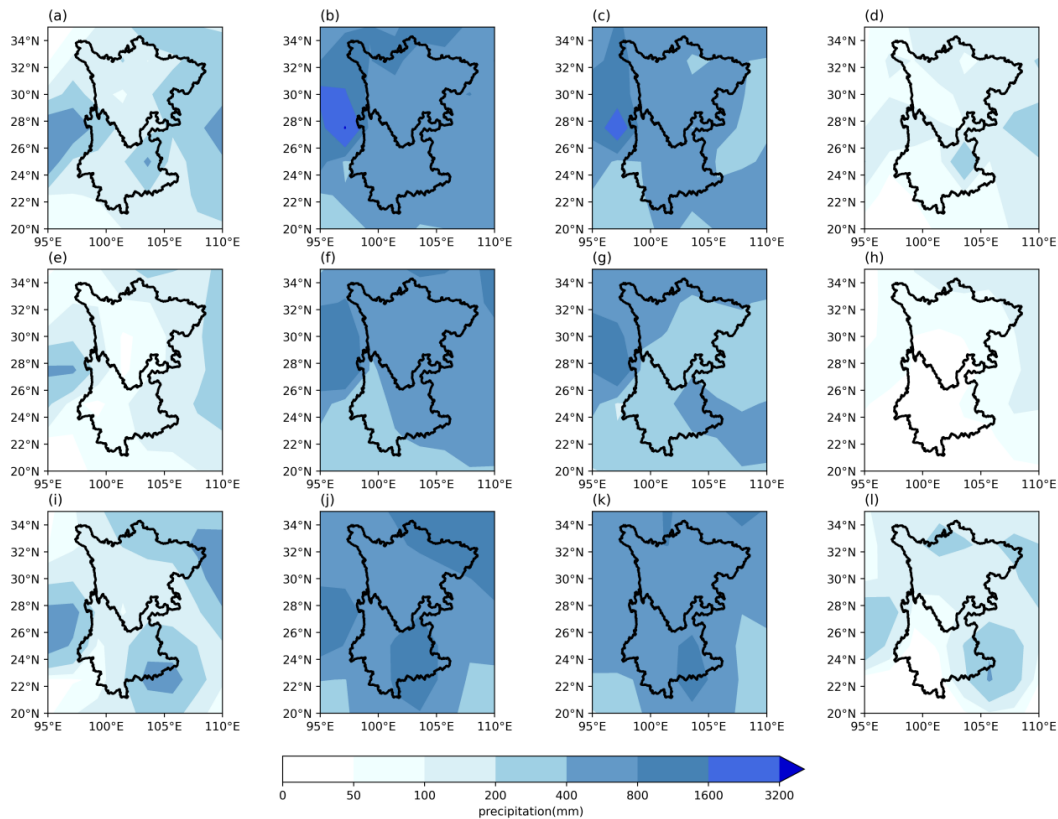


Figure 3: Spatial distribution of multi-year average precipitation from 1998 to 2023 for the two reanalysis datasets and observation data in the study area, for spring (a, e, i), summer (b, f, j), autumn (c, g, k), and winter (d, h, l). The first row shows ERA5L, the second row shows IMERG, and the third row shows NCEP/NCAR.

From Figures 3(a, e, i), it can be observed that in spring, both reanalysis datasets identify a precipitation zone exceeding 400 mm west of the Hengduan Mountains (N27.5, E97.5), which aligns well with the observational data (IMERG). However, both datasets also depict precipitation centers exceeding 400 mm in eastern Yunnan Province and the region east of Sichuan Province, which are not present in the observational data. From Figures 3(b, f, j), in summer, both reanalysis datasets successfully capture a precipitation center exceeding 800 mm west of the Hengduan Mountains. However, ERA5L estimates a maximum precipitation value exceeding 3200 mm in this region, representing a significant overestimation compared to the observational data. This suggests that ERA5L may have substantial biases when estimating extreme precipitation events in summer.

From Figures 3(c, g, k), in autumn, ERA5L continues to depict a precipitation center exceeding 800 mm west of the Hengduan Mountains, though this value remains higher than the observational data suggests. In contrast, NCEP/NCAR fails to capture any precipitation center exceeding 800 mm, indicating its limited ability to reflect the spatial characteristics of autumn precipitation in this region. From Figures 3(d, h, l), in winter, both reanalysis datasets overestimate precipitation by depicting precipitation centers exceeding 200 mm in eastern Yunnan Province, the western Hengduan Mountains, and eastern Sichuan Province. However, the observational data does not exhibit such high-value zones, suggesting that both reanalysis datasets significantly overestimate winter precipitation, particularly in complex terrain regions.

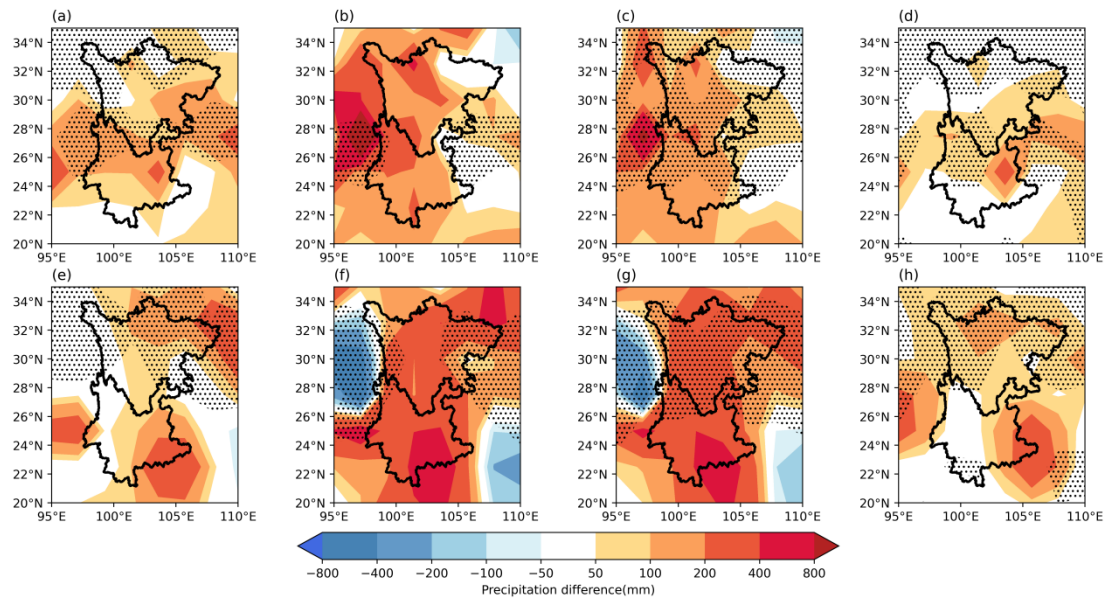


Figure 4: Spatial distribution of absolute bias in multi-year average precipitation from 1998 to 2023 for the two reanalysis datasets and observation data, for spring (a, e), summer (b, f), autumn (c, g), and winter (d, h). The first row shows ERA5L, and the second row shows NCEP/NCAR. The dotted areas indicate regions where the bias passed the 95% confidence significance test ($p = 0.05$).

From Figures 4(a, d), it is evident that ERA5L precipitation reanalysis data significantly overestimates precipitation in spring and winter in the regions west of the Hengduan Mountains, western Sichuan Province, and the area east of the Sichuan-Yunnan border. These overestimation zones pass the 95% significance test. However, north of these overestimated regions (north of N30°), the bias in ERA5L is relatively small and also passes the 95% significance test. From Figures 4(b, c), in summer and autumn, ERA5L exhibits substantial overestimation in the western Hengduan Mountains, with maximum overestimations exceeding 800 mm, which pass the 95% significance test. Additionally, the overestimation of ERA5L precipitation in southwestern China is notably greater in

the western regions than in the eastern regions during these seasons. This indicates a systematic overestimation of ERA5L precipitation in the study area, consistent with the findings of Huang et al. [14], who also reported significant overestimation of ERA5L compared to station observations in southwestern China.

From Figure 4(e), NCEP/NCAR precipitation reanalysis data significantly overestimates precipitation in spring across eastern Sichuan Province and the area to its east, with maximum overestimations exceeding 400 mm, passing the 95% significance test. Overestimation is also observed in the western and southeastern parts of Yunnan Province, with a maximum bias exceeding 400 mm, though these overestimations do not pass the 95% significance test. From Figures 4(f, g), in summer and autumn, NCEP/NCAR significantly underestimates precipitation in the western Hengduan Mountains and western Yunnan Province, with maximum underestimations exceeding 400 mm, which pass the 95% significance test. In contrast, significant overestimation is observed in eastern Yunnan Province and its surrounding areas, also passing the 95% significance test. Furthermore, in the southeastern part of Sichuan Province, a significant overestimation exceeding 400 mm is observed, decreasing eastward. Near E110°, a notable underestimation exceeding 200 mm emerges, though this region does not pass the 95% significance test.

From Figure 4(h), in winter, NCEP/NCAR precipitation reanalysis data shows significant overestimation across Yunnan Province, including both western and eastern regions, with maximum overestimations exceeding 100 mm, passing the significance test. In contrast, significant overestimation is also detected in western and eastern Sichuan, though these overestimations do not pass the significance test. Cheng et al. [5] previously noted that NCEP/NCAR reanalysis data tends to exhibit a positive bias in the Sichuan Basin and the Yunnan-Guizhou Plateau, which aligns with the findings of this study, further confirming the overestimation of precipitation by NCEP/NCAR in these regions.

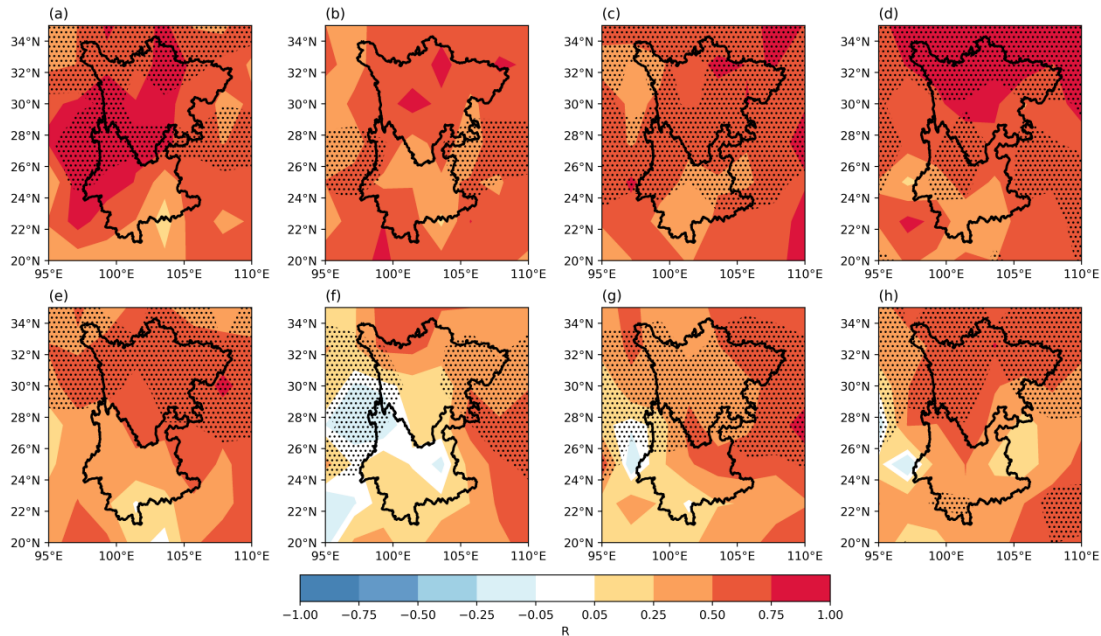


Figure 5: Spatial distribution of correlation coefficients between the two precipitation reanalysis datasets and observation data in the study area from 1998 to 2023 for spring (a, e), summer (b, f), autumn (c, g), and winter (d, h). The first row shows ERA5L, and the second row shows NCEP/NCAR. Dotted areas indicate regions where the correlation passed the 95% significance test ($p = 0.05$).

The spatial distribution of correlation coefficients between the multi-year average precipitation from the two reanalysis datasets and the observational data for each season (spring, summer, autumn, and winter) in the study area from 1998 to 2023 is shown in Figure 5.

From Figure 5(a), in spring, the correlation coefficient between ERA5L precipitation reanalysis data and observational data exceeds 0.75 in western Yunnan Province, southern Sichuan Province, and regions west of the Hengduan Mountains, all of which pass the 95% significance test. However, in southeastern Yunnan Province, the correlation coefficient is below 0.25 and does not pass the significance test. From Figure 5(b), in summer, the correlation coefficient between ERA5L and the observational data exceeds 0.5 in most of Sichuan Province. However, in most of Yunnan Province, the correlation coefficient is below 0.25, and these regions do not pass the 95% significance test. From Figure 5(c), in autumn, the correlation coefficient between ERA5L and the observational data exceeds 0.5 in most of Sichuan Province, while in most of Yunnan Province, it remains below 0.25. The correlation coefficients in Sichuan Province and northern Yunnan pass the 95% significance test.

From Figure 5(d), in winter, the correlation coefficient between ERA5L and the observational data is greater than 0.75 in northern Sichuan Province and further north, passing the 95% significance test. However, the correlation coefficient decreases toward the south, dropping below 0.25 in southern Yunnan Province. From Figures 5(e, h), for NCEP/NCAR, the correlation coefficient with observational data is greater than 0.5 in most of Sichuan Province and eastern Sichuan in both winter and spring, passing the 95% significance test. However, in Yunnan Province, the correlation coefficients are generally below 0.5 and do not pass the significance test. From Figure 5(f), in summer, the correlation coefficient between NCEP/NCAR and the observational data is below 0.25 across most of Sichuan Province, Yunnan Province, and regions west of the Hengduan Mountains, indicating poor consistency with observational data. From Figure 5(g), in autumn, the correlation coefficient between NCEP/NCAR and the observational data is greater than 0.5 in eastern Sichuan Province and northern Yunnan Province. However, the correlation coefficient decreases from northeast to southwest, dropping below 0.05 in regions west of the Hengduan Mountains.

4. Conclusion

This study evaluates the variation trends and spatial distribution of absolute bias, relative bias, and correlation coefficients in precipitation estimates from two reanalysis datasets—ERA5-Land (ERA5L) and NCEP/NCAR—by comparing them with satellite-based IMERG observational data across different seasons in Sichuan, Yunnan, and surrounding areas. The key findings are as follows:

- 1) General Overestimation and Seasonal Patterns. Both ERA5L and NCEP/NCAR overestimate precipitation in all seasons compared to IMERG observations. However, both datasets effectively capture the overall seasonal precipitation patterns in the study area. And from 2010 to 2023, ERA5L exhibited smaller biases compared to NCEP/NCAR, whereas from 1998 to 2010, both datasets had similar biases in winter and spring. However, in summer and autumn, NCEP/NCAR showed smaller biases compared to the observational data.
- 2) ERA5L precipitation data exhibits a systematic overestimation across the study area. In winter and spring, overestimation is mainly concentrated west of the Hengduan Mountains, western Sichuan Province, and east of the Sichuan-Yunnan border. In summer and autumn, overestimation is more pronounced in the western regions. NCEP/NCAR precipitation data also overestimates precipitation across all four seasons in Sichuan Province, eastern Sichuan, and western Yunnan. In winter and spring, the overestimation center in southern Yunnan shifts eastward. In summer and autumn, NCEP/NCAR significantly underestimates precipitation in areas west of Sichuan and east of Yunnan. These findings are consistent with Xiao-Yu Cheng et al. [6], who reported that NCEP/NCAR exhibits a positive bias in the Sichuan Basin and the Yunnan-Guizhou Plateau.

- 3) ERA5L shows higher correlation coefficients in winter and spring compared to summer and autumn, with correlations in Sichuan being generally higher than in Yunnan. Similarly, NCEP/NCAR exhibits higher correlation coefficients in winter and spring in Sichuan than in Yunnan. In summer, NCEP/NCAR correlation coefficients decrease from east to west, while in autumn, they decrease from northeast to southwest.
- 4) Both ERA5L and NCEP/NCAR perform better in Sichuan Province than in Yunnan Province and more reliably in winter and spring than in summer and autumn.

ERA5L outperforms NCEP/NCAR in representing precipitation patterns in both Sichuan and Yunnan in summer and autumn, while both datasets exhibit similar performance in winter and spring. These findings highlight the seasonal and regional differences in the applicability of reanalysis precipitation datasets, particularly in complex terrains such as western Yunnan, where both datasets demonstrate weaker performance.

References

- [1] Skofronick-Jackson, G., Kirschbaum, D., Petersen, W., et al. The global precipitation measurement (GPM) mission's scientific achievements and societal contributions: Reviewing four years of advanced rain and snow observations. *Quarterly Journal of the Royal Meteorological Society*, 2018, 144(S1): 27-48.
- [2] Liu, Tingting, Zhu, Xiufang, Guo, Rui, et al. Analysis of the applicability of ERA5 reanalysis precipitation data in China. *Arid Zone Geography*, 2022, 45(1): 66-79.
- [3] Zhang, Ruoyu. The applicability of PREC/L reanalysis precipitation data in the Shiyang River Basin and its application study. [D] Wuhan: Huazhong University of Science and Technology, 2019: 1-8.
- [4] Lan, Yufeng, Hou, Junxing, Huang, Jiahong. Evaluation of three reanalysis precipitation datasets in southern China. *Meteorological Research and Applications*, 2020, 41(02): 14-20. DOI: 10.19849/j.cnki.CN45-1356/P.2020.2.03.
- [5] Cheng, Xiaoyu, Wang, Yanhua, Li, Guochun, et al. Comparative evaluation of three reanalysis precipitation datasets in China. *Progress in Climate Change Research*, 2013, 9(4): 258-265.
- [6] Ye, M. S., Yao, X. P., Zhang, T., et al. Intercomparison of CRA-Interim precipitation products with ERA5 and JRA-55. *Journal of Tropical Meteorology*, 2021, 27(2): 136-147.
- [7] Pendergrass, A. G., et al. "Nonlinear response of extreme precipitation to warming in CESM1." *Geophysical Research Letters*, 46.17-18 (2019): 10551-10560.
- [8] Jamet, Cédric, and William W. Hsieh. "Nonlinear atmospheric variability in the winter northeast Pacific associated with the Madden-Julian oscillation." *Geophysical Research Letters*, 32.13 (2005).
- [9] Qin, Yanhua, et al. "Spatio-temporal variations of nonlinear trends of precipitation over an arid region of northwest China according to the extreme-point symmetric mode decomposition method." *International Journal of Climatology*, 38.5 (2018): 2239-2249.
- [10] Bengtsson, Lisa, et al. "Impact of flow-dependent horizontal diffusion on resolved convection in AROME." *Journal of Applied Meteorology and Climatology*, 51.1 (2012): 54-67.
- [11] Guo, Chunrui, et al. "Does ERA5-Land Effectively Capture Extreme Precipitation in the Yellow River Basin?" *Atmosphere*, 15.10 (2024): 1254.
- [12] Li, Yanzhao, et al. "Evaluation of long-term and high-resolution gridded precipitation and temperature products in the Qilian Mountains, Qinghai-Tibet Plateau." *Frontiers in Environmental Science*, 10 (2022): 906821.
- [13] Fu, Guobin, et al. "Comparison of NCEP-NCAR and ERA-Interim over Australia." *International Journal of Climatology*, 36.5 (2016).
- [14] Huang, Xiaolong, Wu, Wei, Xu, Jianhui, et al. Evaluation of ERA5-Land precipitation reanalysis data in southwestern China. *Plateau Meteorology*, 2023, 42(06): 1562-1575.
- [15] Gao, Jinlan, Wen, Huayang, Zheng, Xiaoyi, et al. Evaluation of three reanalysis precipitation datasets in Anhui Province. *Journal of Meteorology and Environment*, 2023, 39(4): 47-56.