frontiers in Climate

TYPE Original Research
PUBLISHED 18 February 2026
DOI 10.3389/fclim.2026.1731069

Check for updates

OPEN ACCESS

EDITED BY
Xiang Gao,
Massachusetts Institute of Technology,
United States

REVIEWED BY
Gloria Chinwendu Okafor,
Nigeria Maritime University, Nigeria
Popat Salunke,
Massachusetts Institute of Technology,
United States

*CORRESPONDENCE
Temesgen Gashaw Tarkegn
gtemesgen114@gmail.com

RECEIVED 23 October 2025
REVISED 21 January 2026
ACCEPTED 26 January 2026
PUBLISHED 18 February 2026

CITATION
Tarkegn TG, Ray RL, Tefera GW and Tsegaye Wubetu S (2026) Performance of gauge-based and reanalysis gridded temperature datasets in representing means and extremes across different climate zones of the Brazos River Basin, United States.
Front. Clim. 8:1731069.
doi: 10.3389/fclim.2026.1731069

Performance of gauge-based and reanalysis gridded temperature datasets in representing means and extremes across different climate zones of the Brazos River Basin, United States

Temesgen Gashaw Tarkegn[1]*, Ram Lakhan Ray[1], Gebrekidan Worku Tefera[1] and Selemon Tsegaye Wubetu[2]

[1]College of Agriculture, Food and Natural Resources, Prairie View A&M University, Prairie View, TX, United States, [2]Department of Natural Resource Management, College of Agriculture and Environmental Science, Bahir Dar University, Bahir Dar, Ethiopia

Introduction: Identifying the most reliable gridded temperature datasets for each region is crucial for supporting evidence-based decision-making.

Methods: This study evaluates the performance of gauge-based and reanalysis gridded temperature datasets, Daymet, PRISM, MERRA, and ERA5, in estimating means and extremes across the diverse climate zones of the Brazos River Basin in the United States. The evaluation spans 1998–2020 and examines the datasets' ability to estimate maximum (Tmax) and minimum (Tmin) temperatures across daily, monthly, and annual scales. Additionally, the datasets' ability to estimate temperature extremes are assessed using 12 indices recommended by the Expert Team on Climate Change Detection and Indices (ETCCDI). The evaluations were conducted using the Global Historical Climatology Network (GHCN) as the reference dataset. Four continuous statistical metrics were used to assess performance, and the overall ranking was determined using the Comprehensive Rating Index (CRI).

Results and discussion: The results indicate that dataset performance varies across climate zones, temporal scales, and specific temperature extremes. PRISM and Daymet demonstrated the highest accuracy in estimating daily Tmax and Tmin, respectively in most climate zones. For monthly and annual time scales, Daymet was most effective for estimating Tmax, and PRISM for Tmin. Conversely, MERRA showed the weakest performance for Tmax, and ERA5 for Tmin, across all temporal scales. In certain climate zones, reanalysis datasets are better than gauge-based datasets. For temperature extremes, PRISM outperformed the other datasets across most indices, while ERA5 showed the poorest performance for most of the indices. Thus, the study recommends selecting the highest-performing dataset for Tmax and Tmin separately, tailored to the temporal scale relevant to the study's objectives.

KEYWORDS
climate zones, Daymet, ERA5, extremes, MERRA, PRISM, reanalysis

1 Introduction

Due to the logistical and financial challenges associated with installing a sufficient number of meteorological stations, particularly across both developed and developing regions, gridded temperature datasets have become indispensable tools for evidence-based decision-making (Al-Sakkaf et al., 2024; Bodjrenou et al., 2023; Peng et al., 2023; Lan et al., 2022). These datasets are extensively utilized in a wide range of applications, including hydrological modeling, climate projections, crop simulation, and various other environmental and agricultural analyses (Smith et al., 2025; Tefera et al., 2025; Wang et al., 2025; Gashaw et al., 2024; Araghi et al., 2022; Mihalevich et al., 2022). However, the performance of gridded temperature datasets can vary significantly across different geographic regions and temporal scales (Smith et al., 2025; Tadesse et al., 2025). Therefore, a comprehensive understanding of their spatiotemporal accuracy and reliability is essential to select the most appropriate dataset for specific decision-making contexts.

Nevertheless, comprehensive evaluations of gridded temperature datasets remain limited both in the United States and globally. Most existing evaluations have primarily focused on gauge-based gridded temperature datasets, often using them as reference products in climate studies. Daymet (Tefera et al., 2025; Tefera et al., 2024) and PRISM are among the widely used temperature datasets in the United States (Newman et al., 2024). However, because these datasets are derived from observational station data through interpolation or statistical algorithms, their accuracy can be influenced by factors such as station density, topographic complexity, and the specific methodologies employed in their development (Mankin et al., 2025). Consequently, the performance of gauge-based gridded datasets is not uniform and may differ significantly by region and timescale (Tadesse et al., 2025; Tarek et al., 2020). Therefore, it is essential to rigorously evaluate these datasets themselves before applying them in research, modeling, and decision-making across diverse spatial and temporal contexts.

In several studies conducted globally, reanalysis datasets have been also employed either as benchmarks for evaluating gridded temperature datasets, as inputs for modeling and too many other applications (Wang et al., 2025; Al-Sakkaf et al., 2024). For instance, ERA5 was used to assess specific temperature datasets, temperature extremes, and as input for modeling studies worldwide (Al-Sakkaf et al., 2024; Almeida and Coelho, 2023; Alaminie et al., 2021). Additionally, MERRA have been widely applied for climate model evaluation, temperature extremes, and hydrological and climate modeling across numerous watersheds worldwide (Gashaw et al., 2024; Valappil et al., 2023). However, as reanalysis datasets are not derived from direct observations but rather from a combination of model outputs and assimilated data, their accuracy and applicability can vary significantly across spatial and temporal scales. Thus, comprehensive validation of reanalysis datasets is essential before their use in scientific and operational applications. Nevertheless, despite some evaluation efforts in other parts of the world (Tadesse et al., 2025; McNicholl et al., 2022), studies assessing the performance of reanalysis temperature datasets within the United States remain notably limited.

Furthermore, previous evaluations of gridded temperature datasets have often relied on areal averages over entire study domains, without disaggregating performance by distinct climate regions. As a result, these studies have not adequately captured spatial variability in dataset performance. However, research conducted in North America (Smith et al., 2025; Tarek et al., 2020) and other global regions (Tadesse et al., 2025; Gashaw et al., 2023; McNicholl et al., 2022) has demonstrated that the accuracy of gridded temperature datasets can differ significantly across climate zones. Additionally, many earlier evaluations have focused on a single temporal scale, such as daily, monthly, or annual periods, without conducting a comprehensive assessment across multiple timeframes. Evidence from studies in diverse geographic settings indicates that the performance of temperature datasets can vary considerably depending on the temporal resolution considered (Peng et al., 2023). Therefore, a multi-scale and climate-region-specific evaluation is essential for accurately assessing the utility of gridded temperature datasets.

Studies have largely overlooked the evaluation of gridded temperature datasets in relation to temperature extremes (Bhattacharyya et al., 2025; Smith et al., 2025). The limited existing studies that have addressed extremes often failed to incorporate the full set of temperature extreme indices recommended by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Tan et al., 2023; Lan et al., 2022). Research has shown that the performance of gridded temperature datasets can vary significantly depending on the specific temperature extreme being considered. In addition, many prior evaluations have not assessed the datasets separately for maximum (Tmax) and minimum (Tmin) temperatures (Almeida and Coelho, 2023; Birkel et al., 2022). However, existing evidence highlights that the accuracy of gridded temperature products often differs between Tmax and Tmin (Tadesse et al., 2025), underscoring the importance of evaluating these variables independently to ensure robust and reliable analysis.

The Brazos River Basin is one of the largest and developed river systems (Lee and Singh, 2020; Damavandi et al., 2019; Vogl and Lopes, 2009) in Texas and New Mexico. Studies conducted in parts of the river basin indicated that this area is experiencing hydrological changes such as a decrease in annual streamflow volume, declining peak flow during summer, and declining groundwater levels (Wolaver et al., 2024). Projection study also indicates an expected increase in temperature and temperature extreme indices across climate change scenarios (Tefera et al., 2024). Thus, reliable temperature data that estimates Tmax and Tmin across its diverse climate zones is vital for supporting evidence-based decision making. This study, therefore, evaluates the performance of both gauge-based and reanalysis gridded temperature products in estimating Tmax and Tmin at daily, monthly, and annual scales, as well as in capturing temperature extremes across the basin's diverse climate zones. To the best of our knowledge, no previous study has comprehensively evaluated these datasets for estimating Tmax and Tmin across multiple temporal scales and capturing temperature extremes across climate zones in this area. By identifying the top-performing products for each climate zone, our findings furnish a crucial foundation for data-driven decision-making across the Brazos River Basin.

2 Study area and datasets

2.1 Study area

The Brazos River Basin covers approximately 117,519 km^2 across Texas and New Mexico (Figure 1). Elevation ranges from 4 meters below sea level to 1,457 meters above sea level (m a.s.l), with a mean elevation of 524 m a.s.l. The basin encompasses nine climate

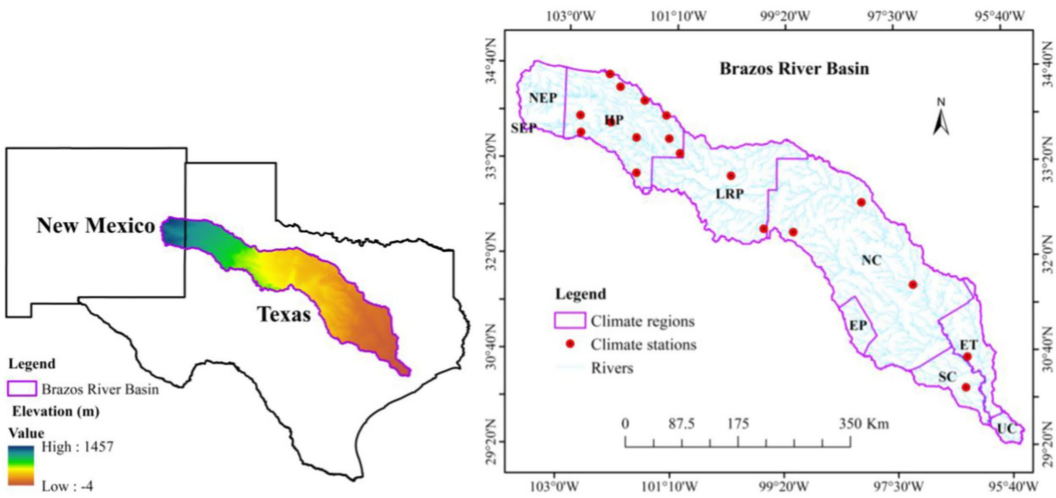

FIGURE 1
Location of the study area and its corresponding climate zones. The abbreviations used for the climate regions are as follows: ET, East Texas; HP, High Plains; LRP, Low Rolling Plains; NC, North Central; SC, South Central; EP, Edwards Plateau; NEP, Northeastern Plains; SEP, Southeastern Plains; UC, Upper Coast.

zones-Edwards Plateau (EP), High Plains (HP), Northeastern Plains (NEP), Southeastern Plains (SEP), Low Rolling Plains (LRP), East Texas (ET), South Central (SC), North Central (NC), and Upper Coast (UC) (Figure 1). The long-term mean annual (1998–2020) precipitation values computed for each climate zone using GHCN data show substantial spatial variability, ranging from 496 mm in the HP climate zone to 1,206 mm in the SC climate zone. Over the same period, mean annual Tmax ranged from 23.6 °C in the HP climate zone to 26.7 °C in the ET climate zone, while mean annual Tmin varied from 7.5 °C in the HP climate zone to 15.1 °C in the ET climate zone, underscoring pronounced climatic variability across the basin.

2.2 Datasets

This study utilized five datasets: GHCN, Daymet, PRISM, ERA5, and MERRA. GHCN provides station-based observations, whereas Daymet and PRISM are gauge-based gridded datasets. In contrast, ERA5 and MERRA-2 are reanalysis products. The primary reason for evaluating the aforementioned gauge-based and reanalysis gridded datasets in this study is that these products are widely used as benchmarks in climate modeling and gridded temperature evaluation studies, as well as serving as inputs for numerous modeling applications (Wang et al., 2025; Al-Sakkaf et al., 2024; Newman et al., 2024; Valappil et al., 2023).

The GHCN dataset provides long-term, ground-based meteorological observations collected from thousands of weather stations distributed across the globe. Despite its extensive spatial reach, the number of stations and the length of available records differ substantially from one region to another, reflecting variations in monitoring infrastructure and historical data availability (Menne et al., 2012). Because the GHCN archive is composed solely of in-situ measurements taken directly at meteorological stations, the dataset is considered to represent purely observed, rather than modeled or remotely sensed, climate information. As a result, the GHCN dataset is used as a reference for evaluating various climate datasets and other applications in North America (Wang et al., 2024; Lee and Singh, 2020; Walton and Hall, 2018) and globally (Lan et al., 2022; McNicholl et al., 2022). In this study, we similarly employed version 3 of the GHCN's Tmax and Tmin data as reference for assessing the performance of Daymet, PRISM, ERA5, and MERRA temperature datasets.

Daymet is a station-based gridded product for North America with 1 km spatial and daily temporal resolution (Thornton et al., 2021), available for 1980–2023. It is widely used in recent studies (Tefera et al., 2025; Newman et al., 2024; Tefera and Ray, 2024; Tefera et al., 2024). PRISM, likewise derived from meteorological stations across the conterminous United States, employs an elevation-aware interpolation scheme (Buban et al., 2020). It provides 4 km spatial and daily temporal resolution for 1981 to present. Since both products originate from observed data, PRISM, like Daymet, is frequently adopted as a reference dataset in diverse research applications (Newman et al., 2024). In this study, Daymet version 4 R1 was used.

MERRA provides global coverage at 0.5 ° × 0.625 ° (~50 km) spatial resolution (Gelaro et al., 2017). Daily and hourly fields are available from 1981 to the present and from 2001 to the present, respectively. Owing to its broad temporal span and physically consistent reanalysis framework, MERRA is widely employed both as model input and as a benchmark dataset in the United States and elsewhere (Gashaw et al., 2024; Valappil et al., 2023). Like MERRA, ERA5 supplies global reanalysis data (Hersbach et al., 2020). ERA5 data are available from 1950 onward at a daily temporal resolution, with a spatial resolution of 0.25° (~30 km). This dataset is widely used both as input for models as a reference for evaluating other climate datasets, and many other applications across the United States (Ahn et al., 2024) and globally (Al-Sakkaf et al., 2024; Almeida and Coelho, 2023; Alaminie et al., 2021). The version of MERRA evaluated in this study is version 2.

3 Methodology

The screening of GHCN meteorological stations in this study was conducted in two stages. In the first stage, all stations located within the study area were downloaded, resulting in a total of 304 stations. The second stage involved screening these stations for the study period

(1998–2020), during which stations lacking data at either the beginning or the end of the period were excluded. Applying these criteria, 18 meteorological stations that span five climate zones were identified and contained less than 1.5% missing data. These climate zones include ET, HP, LRP, NC, and SC. Consequently, four climate zones-EP, NEP, SEP, and UC-were excluded from this study due to insufficient data. Among the 18 selected stations, 11 are located in HP, 3 in NC, and 2 in LRP. The ET and SC zones are each represented by a single station, as no additional stations met the screening requirements. Consequently, analyses for the ET and SC zones were conducted using the available single stations.

Missing Tmax and Tmin data were imputed using the Multivariate Imputation by Chained Equations (MICE) package in R software (Buuren et al., 2025). The MICE algorithm fills missing values at an individual station by using the complete observations from all other stations as predictors. It generates multiple estimates for each gap, incorporates the associated uncertainties, and reports the corresponding standard errors. Consequently, the MICE package is a widely accepted method for handling missing climate data (Tadesse et al., 2025; Lebeza et al., 2024; Mohammed et al., 2022).

Since this study utilized station-based Tmax and Tmin data from the GHCN, while the evaluated datasets are gridded, a point-to-pixel evaluation approach was employed. This method aligns with previous studies that have used station observations to assess gridded climate datasets (Tadesse et al., 2025; Bodjrenou et al., 2023; Tan et al., 2023). Specifically, values from each gridded product were extracted at the geographic coordinates corresponding to the meteorological stations. To facilitate zonal analysis across the study River Basin, the selected stations were grouped according to their respective climate zones. For each zone, the station values were averaged to evaluate the performance of the gridded datasets in estimating Tmax, Tmin, and temperature extremes over the 1998–2020 period. Thus, no interpolation was applied. The performances of the studied gridded datasets were evaluated in the original datasets and therefore did not apply any bias correction in the assessment, which is aligned with these kinds of studies (Tadesse et al., 2025; Gashaw et al., 2023).

The performance of the studied datasets in estimating daily, monthly, and annual Tmax and Tmin across different climate zones was evaluated using four widely adopted statistical metrics: the correlation coefficient (R), Root Mean Square Error (RMSE), Percent Bias (PBIAS), and the Kling-Gupta Efficiency (KGE) (Aniley et al., 2025; Andrade et al., 2024; Deng et al., 2024; Luo et al., 2024; Almeida and Coelho, 2023). The correlation coefficient (R) quantifies the strength and direction of the linear relationship between the observed and estimated values, ranging from -1 to 1. An R value of ± 1 indicates a perfect positive or negative correlation. RMSE assesses the amount of estimation errors relative to the reference data, with values ranging from 0 to ∞; a value of 0 indicates perfect accuracy. PBIAS evaluates the average tendency of the estimated values to be larger or smaller than their observed counterparts. Positive PBIAS indicates underestimation, while negative values indicate overestimation. The optimal value is 0, representing no bias. KGE integrates correlation, variability, and bias into a single metric, with values ranging from $-\infty$ to 1; 1 represents perfect agreement with the observed data.

To assess the capability of the studied temperature datasets in capturing temperature extremes, this study employed all 12 indices (Table 1) recommended by the ETCCDI, which were computed with the RClimDex (1.0) software (Zhang and Yang, 2004) based on the daily Tmax and Tmin records. Unlike many previous evaluations that focus on a limited subset of extreme temperature indices (Tan et al., 2023; Lan et al., 2022), this study incorporates the complete set to

TABLE 1 Temperature extremes considered in this study.

Indices	Index name	Definition of the index	Unit
TXx	Max Tmax/ warmest day	Maximum value of daily maximum temperature	°C
TXn	Min Tmax/ coldest day	Minimum value of daily maximum temperature	°C
TNx	Max Tmin/ warmest night	Maximum value of daily minimum temperature	°C
TNn	Min Tmin/ coldest night	Minimum value of daily minimum temperature	°C
TN10p	Cool nights	Percentage of days when TN < 10th percentile	%
TN90p	Warm nights	Percentage of days when TN > 90th percentile	%
TX10p	Cool days	Percentage of days when TX < 10th percentile	%
TX90p	Warm days	Percentage of days when TX > 90th percentile	%
WSDI	Warm spell duration indicator	Annual count of days with at least 6 consecutive days when TX > 90th percentile	Days
CSDI	Cold spell duration indicator	Annual count of days with at least 6 consecutive days when TN < 10th percentile	Days
SU25	Summer days	Annual count when TX (daily maximum) > 25 °C	Days
FD0	Frost days	Annual count when TN (daily minimum) < 0 °C	Days

provide a comprehensive assessment. The inclusion of all temperature extremes ensures a more robust evaluation, recognizing that different datasets may perform better for specific types of extremes.

The overall performance rankings of the datasets for estimating Tmax and Tmin across various temporal scales and capturing temperature extremes were determined using the Comprehensive Rating Index (CRI), consistent with approaches adopted in previous climate model evaluation studies (Drisya and Al-Zubari, 2025; Lebeza et al., 2024; Du et al., 2022; Rivera and Arnould, 2020). The CRI integrates the number of performance metrics employed (four in this study), the total number of datasets evaluated (also four), and the rank assigned to each dataset for each metric. For a detailed description of the CRI methodology, readers are referred to the cited scientific publications.

4 Results

4.1 Performances for simulating Tmax

Along with Tmin, the performance of the studied gridded temperature datasets in estimating daily, monthly, and annual Tmax is presented in Figures 2–6. For daily Tmax estimation, PRISM performs best in the HP, LRP, and SC climate zones, but ranks lowest in ET and NC (Figure 2). Daymet demonstrates the highest accuracy in ET and NC, and ranks second in LRP and third in the SC region. Together with MERRA, Daymet is also ranked third in the HP climate zone. ERA5 outperforms MERRA in HP, LRP, and SC. In contrast, MERRA performs better than ERA5 in ET and NC (Figure 2). Overall, PRISM is identified as the most accurate dataset for daily Tmax estimation. At the same time, MERRA shows the weakest performance in most climate zones.

Regarding estimation bias, most datasets underestimated daily Tmax across the climate zones, with the exception of MERRA in HP; MERRA and ERA5 in LRP; and Daymet and MERRA in SC (Figure 3). MERRA exhibited the highest estimation bias in HP and LRP, while ERA5 showed the highest bias in ET and NC. In SC, all datasets except PRISM displayed comparable levels of bias. Conversely, the datasets with the lowest estimation bias were PRISM in SC, Daymet in ET and LRP, ERA5 and Daymet in HP, and MERRA in NC (Figure 3).

At the monthly temporal scale, except in SC zone, Daymet is the best-performing dataset for estimating monthly Tmax across the remaining four climate zones, while PRISM ranks second in these climate zones (Figure 2). In the SC zone, PRISM is the leading dataset, while Daymet is the second-best dataset. The results indicate the performance differences of the datasets across climate zones. ERA5 outperforms MERRA in the HP and SC zones, while MERRA performs better than ERA5 in NC. In the remaining two zones, both MERRA and ERA5 exhibit comparably low performance. Overall, Daymet and MERRA are identified as the best and poorest datasets, respectively, for monthly Tmax estimation in most climate zones.

With respect to estimation bias, most datasets underestimated observed monthly Tmax, except for MERRA and ERA5 in LRP, and Daymet and MERRA in SC (Figure 4). MERRA shows the highest estimation bias in SC, LRP, and HP. At the same time, ERA5 exhibits the greatest bias in NC and ET. In contrast, the lowest estimation bias is observed for Daymet in ET and LRP, ERA5 in HP, MERRA in NC, and PRISM in SC.

Similar to the performance difference of the studied gridded temperature datasets for estimating monthly Tmax across climate regions, the performance variations of the studied datasets for estimating Tmax across climate zones is also exhibited in the long-term (1998–2020) monthly mean Tmax values (Figure 5). Majority of the higher biases are demonstrated from May to October (Figure 5). In addition to these months, in LRP climate zone, ERA5 has also showed higher biases from January to March.

At the annual temporal scale, Daymet is the best-performing dataset in the ET, HP, and NC climate zones, while PRISM ranked second in these areas (Figure 2). Conversely, in the LRP and SC zones, PRISM leads in performance, followed by Daymet. MERRA is the poorest-performing dataset in HP, LRP, and SC, while ERA5 shows the weakest performance in ET and NC. Overall, Daymet consistently ranks as the most accurate dataset, whereas MERRA performs the poorest across most of the studied climate zones.

With regard to estimation bias, most datasets underestimated annual Tmax, except for MERRA in HP; MERRA and ERA5 in LRP; and Daymet and MERRA in SC. The highest estimation biases are observed for ERA5 in ET and NC and MERRA in HP, LRP, and SC. In contrast, the lowest estimation biases are recorded for Daymet in ET and LRP; Daymet and ERA5 in HP; MERRA in NC; and PRISM in SC (Figure 6).

4.2 Performances for simulating Tmin

Daymet is the best-performing dataset for estimating Tmin in the HP, LRP, and NC climate zones at the daily temporal scale, with PRISM ranking second in these areas. In SC, PRISM ranks highest, followed by Daymet. However, both Daymet and PRISM perform poorly in the ET zone, where they are the lowest-ranking datasets (Figure 2). In contrast, ERA5 is the poorest-performing dataset across the remaining climate zones. Overall, Daymet demonstrates the best performance, while ERA5 shows the weakest performance for estimating daily Tmin in the majority of the studied climate zones.

For monthly Tmin estimation, PRISM is the best-performing dataset in all climate zones except HP, where Daymet ranks first (Figure 2). Daymet is the second-best performer in LRP, NC, and SC. The poorest performance is observed for ERA5 in HP, LRP, NC, and SC, and Daymet in the ET zone. These findings indicate that the performance of the datasets for estimating monthly Tmin varies across climate zones. Thus, PRISM is identified as the most reliable dataset for estimating monthly Tmin, while ERA5 is the least reliable in most climate zones. The long-term (1998–2020) monthly mean Tmin values of the datasets also indicate the performances variations of the studied datasets across climate zones (Figure 5). For estimating the long-term monthly average Tmin, ERA5 and MERRA in HP, LRP and SC, ERA5 in NC, and Daymet and ERA5 in ET showed higher biases compared to the observed dataset (Figure 5).

At the annual temporal scale, PRISM is the best-performing dataset for estimating Tmin in the ET, LRP, and NC climate zones. It also ranks second in HP and, along with MERRA in SC (Figure 2). Daymet ranks first in HP and SC, and second in LRP and NC. Among the studied datasets, ERA5 in HP, LRP, NC and SC, and Daymet in ET exhibits the weakest performance. Overall, PRISM is the most reliable dataset for estimating annual Tmin, while ERA5 is the least reliable across most climate zones.

Regarding estimation bias, except ERA5 in ET, and MERRA and ERA5 in HP and LRP climate zones, the remaining datasets underestimated Tmin across daily, monthly, and annual temporal scales in ET, HP, and LRP climate zones. Conversely, in the SC zone, except Daymet, all datasets overestimated Tmin at the studied temporal scales. Likewise, except Dayemt for daily and monthly periods, and Daymet and ERA5 for annual time scales, the remaining datasets have overestimation bias in the NC climate zone (Figure 6). ERA5 exhibits the highest estimation bias for Tmin across daily, monthly, and annual scales in HP, LRP, NC, and SC. In contrast, Daymet shows the greatest bias in ET. In contrast, PRISM demonstrates the lowest estimation bias across these time steps in all climate zones except HP (Figure 3).

4.3 Performances for simulating temperature extremes

The performance of the evaluated datasets in estimating various temperature extremes is detailed in Figure 7 and Tables 2–7. Daymet consistently demonstrates superior accuracy for estimating TXx in the HP and LRP climate zones and ranks second in the NC and SC zones (Figure 7). In the ET zone, Daymet and PRISM are better than other gridded data. In contrast, ERA5 exhibits the poorest performance in the NC and ET zones, while MERRA is the lowest-performing dataset in the HP zone. Both ERA5 and MERRA show comparably weak performance in the LRP and SC zones. Overall, Daymet stands out as the

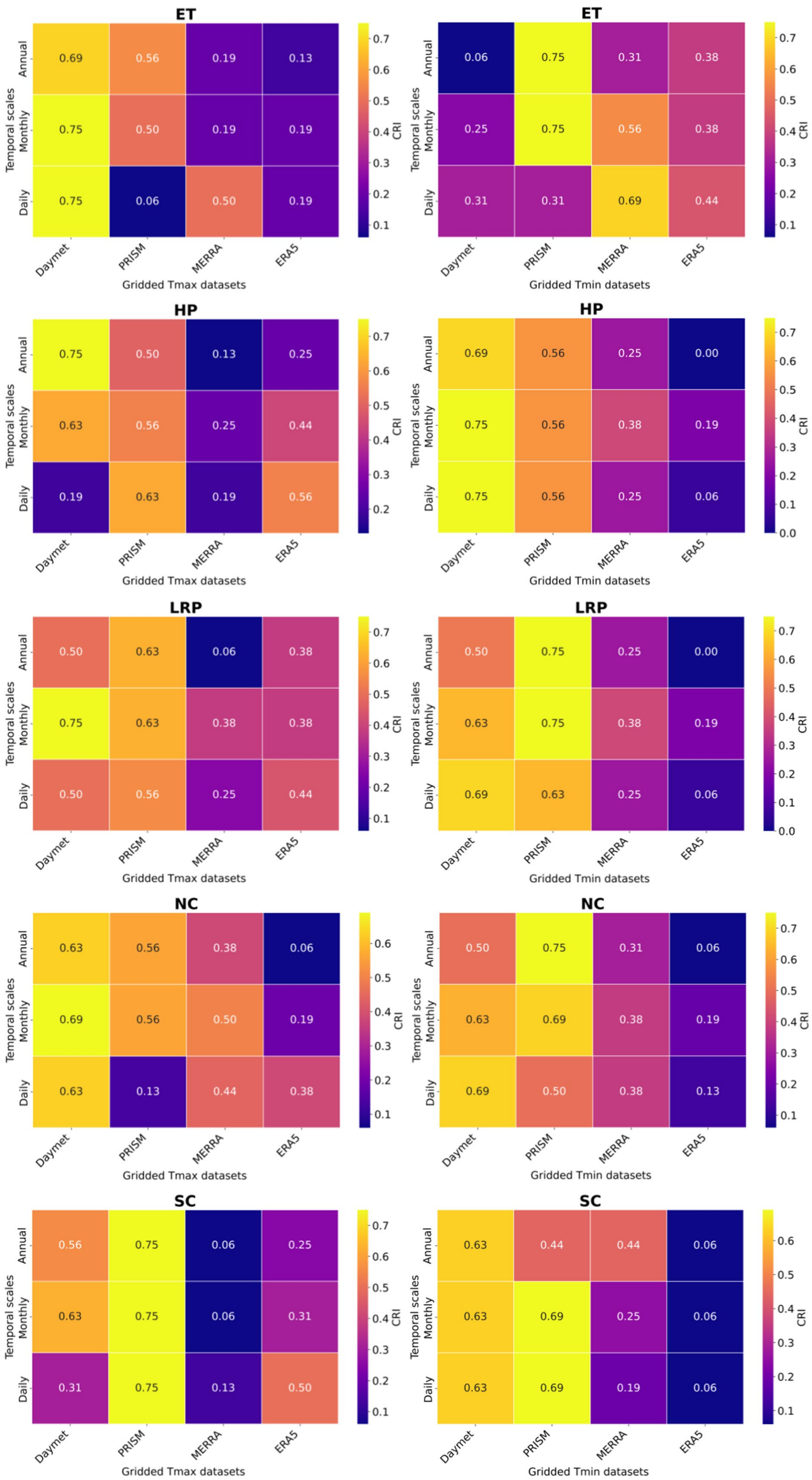

FIGURE 2
Overall ranking of gridded temperature products for simulating Tmax (left) and Tmin (right) across daily, monthly, and annual temporal scales in different climate zones. The abbreviations used for the climate regions are given in Figure 1.

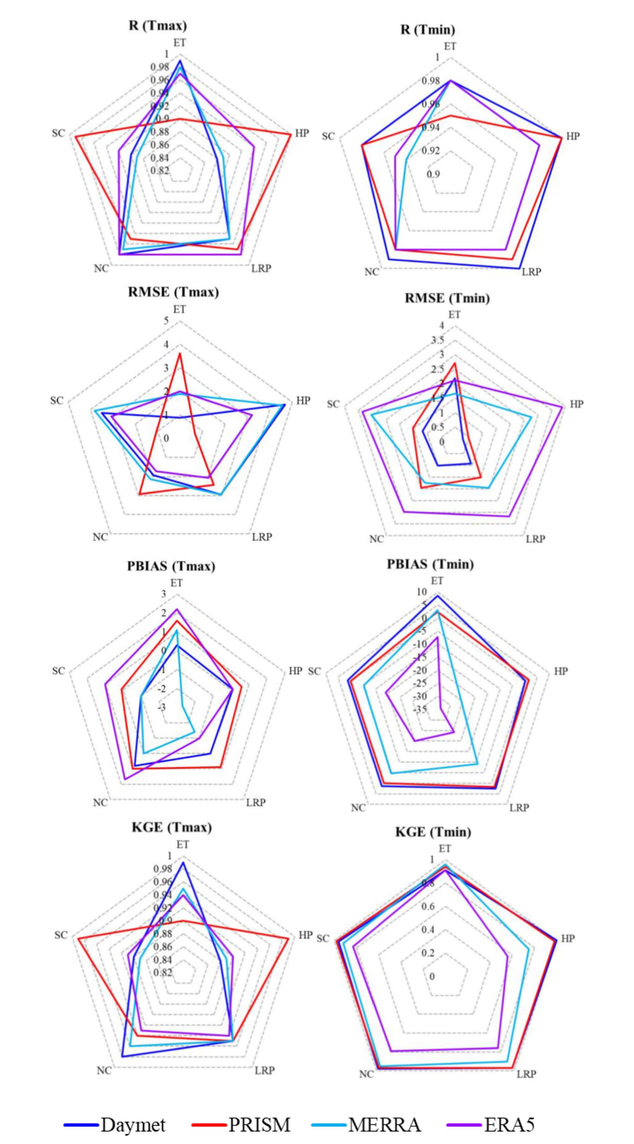

FIGURE 3
Performance of gridded temperature datasets in simulating daily Tmax (left) and Tmin (right) across different climate zones of the Brazos River Basin. The abbreviations used for the climate regions are given in Figure 1.

best-performing dataset, whereas ERA5 performs the poorest in estimating TXx.

Regarding estimation bias, except ERA5, all datasets tend to overestimate TXx in the SC, HP and LRP climate zones. Conversely, in the ET and NC zones, all datasets except MERRA demonstrate underestimation bias (Table 2). While ERA5 displays the highest estimation bias in the ET and NC climate zones, MERRA exhibits the greatest bias for TXx across the other three climate zones. The lowest estimation biases are found in Daymet (ET, HP, and LRP), PRISM (NC), and both Daymet and PRISM in the SC region (Table 2).

For estimating TXn, Daymet demonstrated leading performance in the ET, HP, LRP, and NC climate zones, while PRISM was the top-performing dataset in the SC area (Figure 7). PRISM ranked as the second-best dataset in the ET and HP climate zones. Alongside ERA5 in the LRP and NC regions, PRISM consistently held the second position in terms of performance. Conversely, the poorest-performing datasets for estimating TXn were ERA5 (ET, HP, and SC) and MERRA (LRP and NC). Overall, Daymet and ERA5 represent the best and poorest performing datasets, respectively, across most climate zones of the Brazos River Basin for estimating TXn (Figure 7).

Regarding estimation bias for TXn, all evaluated datasets in the SC zone exhibit an overestimation bias (Table 2). Except PRISM, the remaining datasets show overestimation bias in the ET zone. Conversely, in the HP and NC zones, Daymet, MERRA, and ERA5 tend to underestimate TXn, while PRISM shows overestimation bias. In the LRP zone, both Daymet and PRISM overestimate observed TXn, whereas MERRA and ERA5 underestimate it. The largest estimation biases are observed in ERA5 for the ET, HP, and SC zones, and in PRISM for the LRP and NC zones. Conversely, the lowest estimation biases are associated with Daymet in the ET, HP, LRP, and NC zones, while PRISM exhibits the minimum estimation bias in the SC area (Table 2).

For TNx estimation, PRISM leads in the ET, LRP, and NC climate zones, while Daymet is the top-performing dataset in the HP zone (Figure 7). Both Daymet and PRISM perform equally well as the best datasets for estimating TNx in the SC area. In contrast, ERA5 ranks lowest in the NC and SC climate zones; ERA5 and Daymet show the poorest performance in the ET zone; and both MERRA and ERA5 exhibit the lowest performance in the HP and LRP zones. Overall, PRISM and ERA5 represent the best and poorest performing datasets for estimating TNx (Figure 7).

The estimation bias and direction of the studied datasets for TNx vary across climate zones. In the HP, LRP, NC, and SC zones, MERRA and ERA5 show an overestimation bias, whereas Daymet and PRISM tend to underestimate TNx (Table 3). In the ET zone, all datasets except ERA5 display an underestimation bias. The greatest estimation bias is observed in ERA5 (HP, LRP, and NC), MERRA (SC), and Daymet (ET). Conversely, the lowest estimation bias is found in PRISM (LRP, NC, and SC), MERRA (ET), and Daymet (HP).

For estimating TNn, PRISM is the top-performing dataset in the ET, LRP, and SC climate zones (Table 3). PRISM ranks as the second-best dataset in the NC zone. On the other hand, Daymet leads in the NC zone and ranks second in the SC, LRP, and ET climate zones. Both Daymet and PRISM are the best-performing datasets in the HP climate zone. ERA5 exhibits the poorest performance in the ET and SC zones, while MERRA is the lowest-performing dataset in the HP and LRP zones. In the NC zone, ERA5 and MERRA show comparable low performance. PRISM is the best dataset for estimating TNn, whereas ERA5 and MERRA are the poorest performers (Table 3).

Concerning estimation bias, in the ET, LRP, and SC climate zones, all datasets except Daymet underestimate the observed TNn, while all datasets exhibit underestimation bias in the NC climate zone (Table 3). In the HP zone, Daymet and PRISM display overestimation bias, whereas ERA5 and MERRA tend to underestimate TNn. With respect to the magnitude of estimation bias, the highest biases are observed in ERA5 (ET, NC, and SC) and MERRA (HP and LRP). Conversely, the lowest estimation biases are found in Daymet (HP, LRP, and NC), PRISM (SC), and MERRA (ET).

For TN10p estimation, PRISM is the best-performing dataset in the ET, LRP, and NC climate zones, while ranking second in the HP zone (Figure 7). Conversely, Daymet leads in the HP and ranks second in ET, LRP, and NC climate zones. In SC, both Daymet and MERRA have comparable best performance. ERA5 is the poorest-performing dataset in the ET, HP, and NC zones. In the SC zone, both ERA5 and PRISM demonstrate the lowest performance. Additionally, MERRA exhibits the poorest performance in the LRP zone. Overall, PRISM

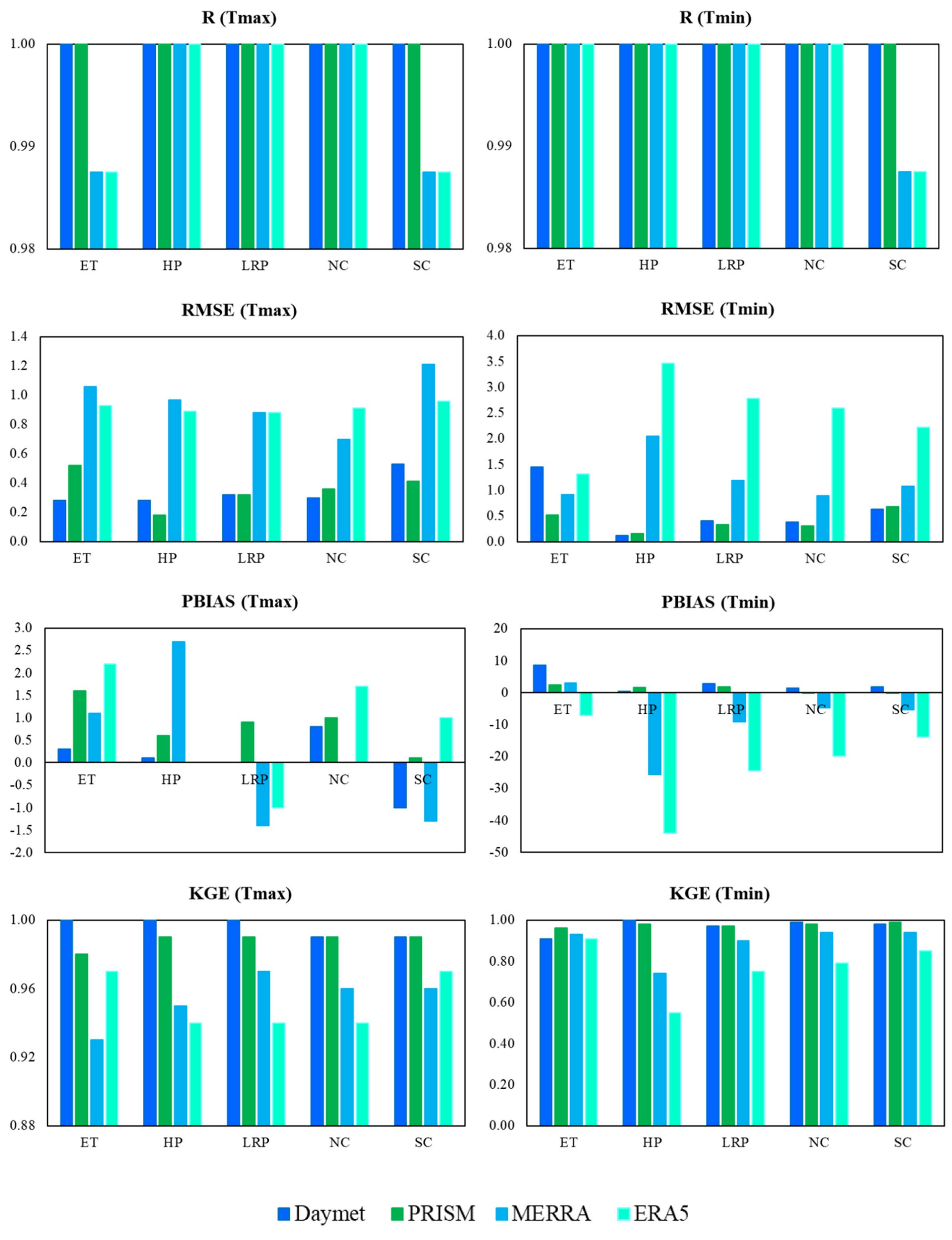

FIGURE 4
Monthly performance of gridded temperature datasets for Tmax (left) and Tmin (right). The abbreviations used for the climate regions are given in Figure 1.

and ERA5 represent the best and poorest datasets in most climate zones, respectively, for estimating TN10p.

Regarding estimation bias, all datasets exhibit overestimation bias for TN10p in the ET, LRP, and NC zones (Table 4). In the SC zone, all datasets except PRISM show overestimation. In the HP zone, Daymet and ERA5 exhibit overestimation bias, whereas PRISM and MERRA show underestimation issues. The maximum estimation bias is observed in ERA5 (ET and SC), PRISM (HP), Daymet and MERRA (LRP), and Daymet (NC) (Table 4). Conversely, with the exception of the HP climate zone, PRISM demonstrates the lowest estimation bias.

The best dataset for estimating TN90p is PRISM in the ET, LRP, and NC zones, while Daymet leads in the SC zone (Figure 7). Both PRISM and Daymet rank first in the HP zone. Except for the ET climate zone, where MERRA is the lowest-performing dataset, ERA5 is the poorest performer across the remaining climate regions. Overall, PRISM is the best-performing dataset in most climate zones, whereas

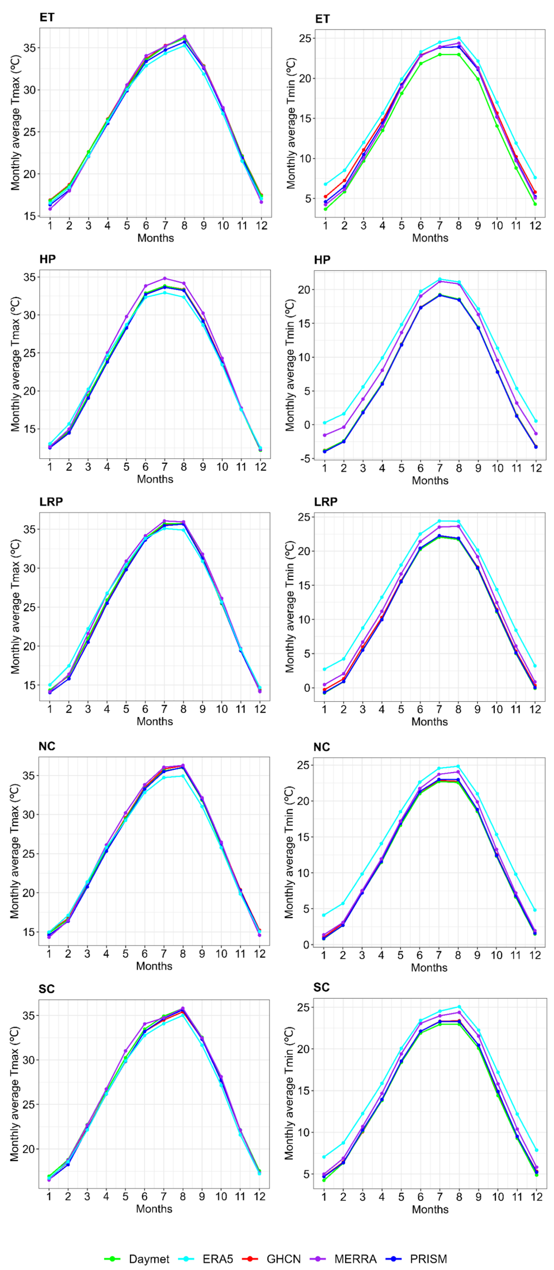

FIGURE 5
Long-term monthly average Tmax (left) and Tmin (right) of GHCN (observed data) and other gridded precipitation products across different climate zones. The abbreviations used for the climate regions are given in Figure 1.

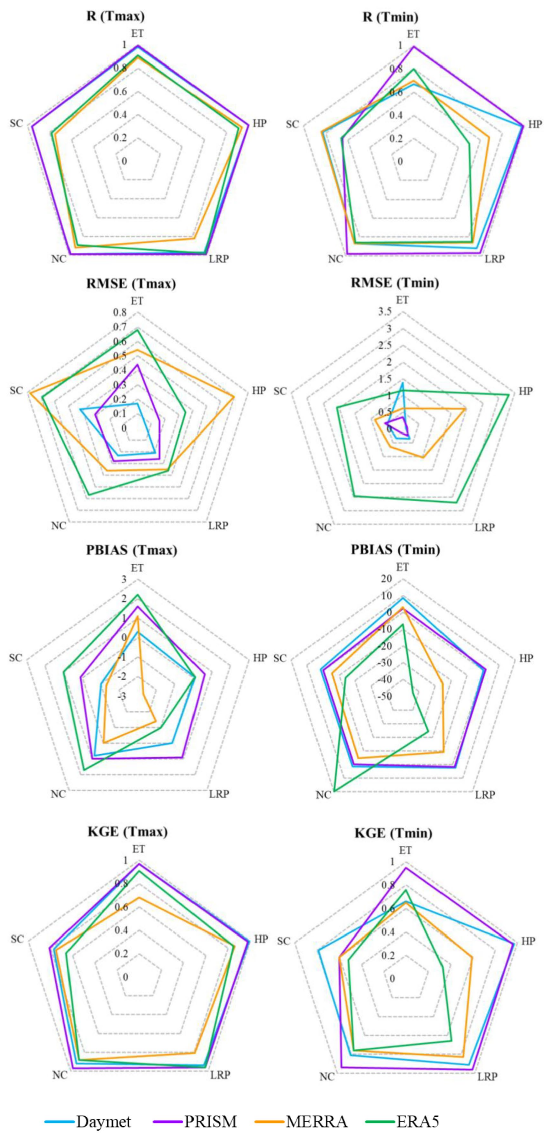

FIGURE 6
Performance of the temperature products in simulating annual Tmax (left) and Tmin (right). The abbreviations used for the climate regions are given in Figure 1.

ERA5 is the least performing dataset for estimating TN90p in most climate zones (Figure 7).

Regarding estimation bias for TN90p, all datasets exhibit overestimation bias in the ET, LRP, NC, and SC zones, whereas they show underestimation bias in the HP zone (Table 4). The highest estimation bias is observed in Daymet for the ET, NC, and SC zones, while ERA5 in HP and PRISM in LRP also exhibit significant estimation bias. The lowest estimation biases are found in PRISM (ET and HP), MERRA (LRP and NC), and ERA5 (SC).

PRISM is the best-performing dataset for estimating TX10p across all studied climate zones (Figure 7). Daymet ranks as the second-best dataset in the ET, HP, NC, and SC regions and holds the third position in the LRP area. MERRA is the poorest-performing dataset in the ET, LRP, and NC climate zones. In the HP and SC zones, both MERRA and ERA5 exhibit comparably low performance. Overall, PRISM emerges as the best dataset, while MERRA is the poorest performer for estimating TX10p in most climate zones (Figure 7).

Regarding estimation bias for TX10p, all datasets exhibit overestimation bias in the ET, HP, LRP, and SC climate zones (Table 5). Conversely, in the NC zone, all datasets except ERA5 show underestimation bias. Except for the HP climate zone, where MERRA displays the greatest estimation bias, the highest biases in the remaining four climate zones are observed in ERA5. Conversely, the lowest estimation biases are found in PRISM (ET, LRP, and NC), ERA5 (HP), and MERRA (SC) (Table 5).

Concerning the performance of the datasets for estimating TX90p, PRISM is the best-performing dataset in the ET, LRP, and NC zones, while ranking second in the HP and SC climate zones (Figure 7). Conversely, Daymet leads in the HP and SC zones and

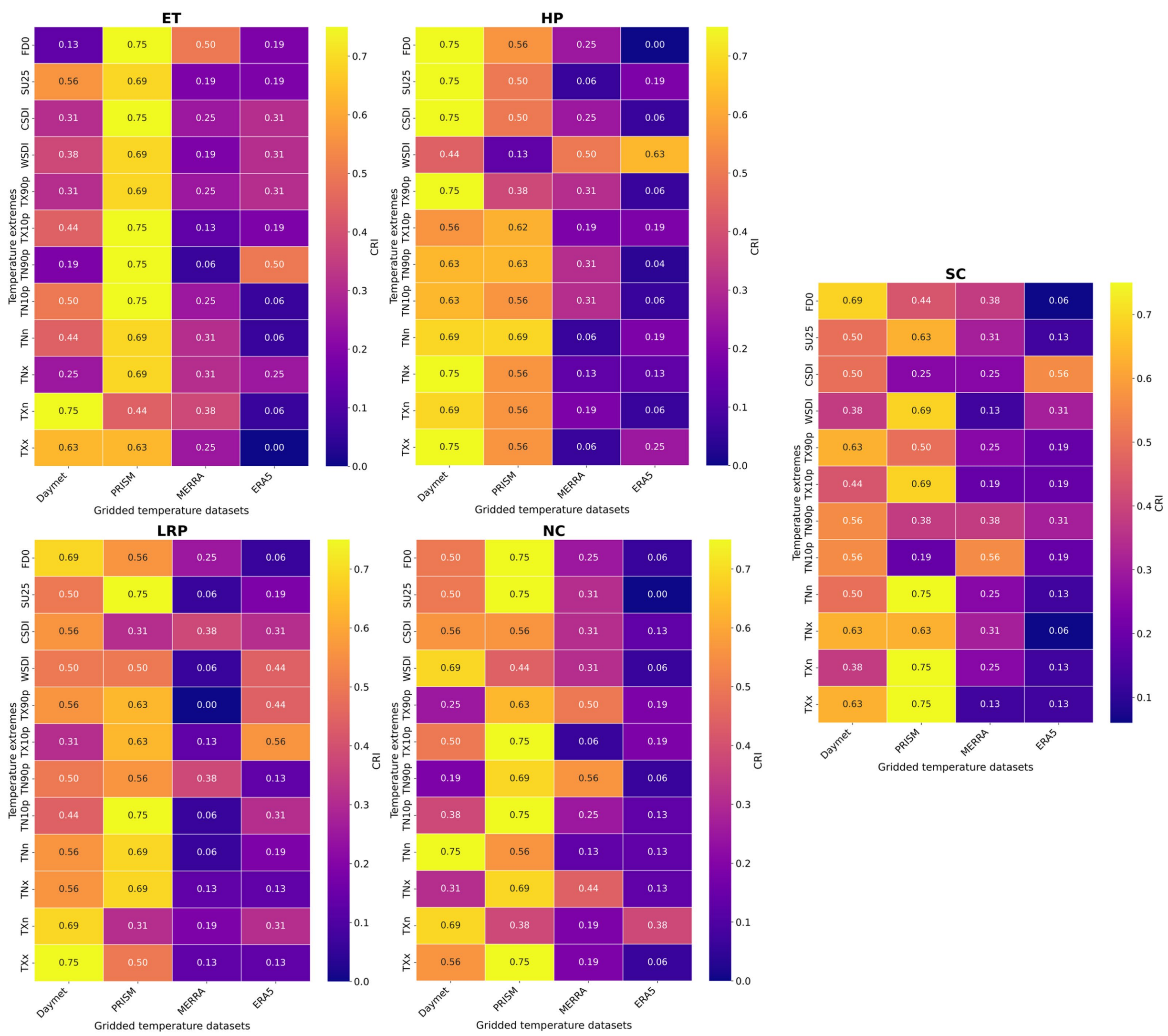

FIGURE 7
Overall performance of gridded temperature products in simulating temperature extremes across different climate zones. The abbreviations used for the climate regions are given in Figure 1.

ranks second in the LRP zone. In the ET zone, Daymet and ERA5 share a comparable second-place performance. In the NC zone, Daymet's performance is lower than that of MERRA. The poorest-performing datasets include MERRA (ET and LRP) and ERA5 (HP, NC, and SC). Overall, PRISM and ERA5 represent the best and poorest datasets, respectively, for estimating TX90p in most climate zones.

Regarding estimation bias for TX90p, all datasets exhibit overestimation bias in the ET, LRP, and SC climate zones (Table 5). In the HP zone, PRISM and MERRA have overestimation bias whereas the remaining two datasets contain underestimation bias. In the NC zone, Daymet and PRISM demonstrate overestimation bias, whereas MERRA and ERA5 exhibit underestimation bias. The highest estimation biases are observed in Daymet and MERRA (ET), ERA5 (HP), MERRA (LRP), Daymet (NC), and both PRISM and MERRA (SC). Conversely, the lowest estimation biases are found in ERA5 (ET and SC), Daymet (HP), Daymet and ERA5 (LRP), and MERRA and ERA5 (NC).

PRISM outperforms other datasets for estimating WSDI in the ET and SC climate zones, whereas Daymet is the leading dataset in the NC zone (Figure 7). Both Daymet and PRISM share the top position in the LRP zone. In the HP zone, although both Daymet and PRISM perform below ERA5 and MERRA, Daymet's performance surpasses that of PRISM. Thus, for WSDI estimation, PRISM and Daymet demonstrate comparable, leading performance. The lowest-performing datasets include MERRA (ET, LRP, and SC), PRISM (HP), and ERA5 (NC). Notably, in the HP zone, both ERA5 and MERRA outperform relative to the gauge-based datasets Daymet and PRISM. Overall, PRISM and Daymet exhibit similarly strong performance, while MERRA is the poorest-performing dataset for estimating WSDI across most climate zones.

Regarding estimation bias, all datasets in the ET climate zone, and Daymet, PRISM, and MERRA in the SC climate zone exhibit overestimation of WSDI, whereas all studied datasets show underestimation bias in the remaining three climate zones (Table 6). The highest estimation biases are observed in MERRA (ET, LRP, and SC) and PRISM (HP and NC). Conversely, the lowest estimation biases are exhibited by ERA5 (ET, HP, LRP, and SC) and Daymet (NC) (Table 6).

TABLE 2 Performance of the gridded datasets in estimating TXx and TXn across the climate zones of the study area.

Climate regions	Datasets	TXx				TXn			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.96	0.65	1.1	0.96	0.97	0.82	−23.2	0.76
	PRISM	0.97	0.74	1.6	0.97	0.91	1.15	27	0.7
	MERRA	0.85	1.34	−2.2	0.83	0.85	1.44	−25.3	0.7
	ERA5	0.64	1.71	2.6	0.51	0.71	1.44	−38.3	0.62
HP	Daymet	0.97	0.4	−0.3	0.97	0.96	0.86	1.2	0.93
	PRISM	0.97	0.45	−0.4	0.92	0.98	0.86	−13.56	0.84
	MERRA	0.95	1.66	−4	0.91	0.84	0.86	48.2	0.48
	ERA5	0.95	1.19	2.7	0.93	0.71	0.86	49	0.49
LRP	Daymet	0.98	0.32	−0.2	0.96	0.87	1.47	−14.9	0.8
	PRISM	0.95	0.59	−0.75	0.95	0.9	1.88	−55.8	0.42
	MERRA	0.89	1.48	−3.1	0.87	0.75	2	29.6	0.52
	ERA5	0.87	0.95	1.2	0.85	0.84	1.96	39.8	0.56
NC	Daymet	0.96	0.63	1.2	0.96	0.91	1.18	4.6	0.89
	PRISM	0.97	0.38	0.3	0.96	0.91	1.22	−62.2	0.64
	MERRA	0.85	1.41	−2.6	0.77	0.74	1.91	24	0.43
	ERA5	0.84	1.71	3.7	0.78	0.95	1.22	26.9	0.22
SC	Daymet	0.89	0.91	−0.2	0.81	0.95	1.49	−85.1	0.44
	PRISM	0.93	0.78	−0.2	0.81	0.99	0.44	−11.1	0.88
	MERRA	0.77	1.95	−3.5	0.76	0.79	2.52	−96.4	0.46
	ERA5	0.72	1.54	1.4	0.45	0.72	2.85	−99.2	0.55

TABLE 3 Performance metrics of gridded temperature datasets for estimating TNx and TNn.

Climate regions	Datasets	TNx				TNn			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.62	1.46	4.9	0.62	0.95	1.09	−15.09	0.82
	PRISM	0.7	0.66	1.4	0.67	0.98	0.48	5.5	0.94
	MERRA	0.41	0.83	0.3	0.41	0.84	1.18	1.8	0.61
	ERA5	0.55	0.86	−2.1	0.5	0.82	2.01	31	0.62
HP	Daymet	0.97	0.23	0.2	0.96	1	0.24	−0.4	0.99
	PRISM	0.97	0.81	3	0.9	1	0.22	−0.7	0.99
	MERRA	0.75	2.53	−10.8	0.55	0.74	4.27	29	0.48
	ERA5	0.82	2.92	−12.8	0.59	0.59	3.52	20.6	0.51
LRP	Daymet	0.97	0.81	3	0.9	0.98	0.44	−0.6	0.97
	PRISM	0.94	0.42	0.8	0.93	0.99	0.31	0.7	0.99
	MERRA	0.72	1.52	−4.9	0.66	0.76	2.74	22.5	0.51
	ERA5	0.91	2.38	−9.1	0.76	0.75	2.47	18.4	0.68
NC	Daymet	0.63	1.05	2.2	0.53	0.98	0.55	2.8	0.92
	PRISM	0.76	0.72	0.1	0.69	0.98	0.78	6.3	0.91
	MERRA	0.73	1.27	−3.8	0.73	0.5	2.83	21.5	0.26
	ERA5	0.63	1.99	−6.8	0.61	0.75	3.28	31	0.59
SC	Daymet	0.71	0.85	2.3	0.71	0.92	1.05	−6.8	0.83
	PRISM	0.61	0.69	0.4	0.6	0.92	0.98	0.6	0.85
	MERRA	0.61	0.98	−2.7	0.59	0.79	1.58	4.7	0.59
	ERA5	0.41	1.65	−2.6	0.33	0.82	2.13	31.8	0.59

TABLE 4 Performance metrics of the datasets for estimating TN10p and TN90p across different climate zones.

Climate regions	Datasets	TN10p				TN90p			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.66	3.22	−11.3	0.59	0.69	3.75	−21.3	0.6
	PRISM	0.97	1.27	−9.8	0.9	0.95	1.69	−12.1	0.87
	MERRA	0.49	3.9	−11.8	0.43	0.64	4.17	−18.5	0.5
	ERA5	0.44	3.9	−12.6	0.37	0.81	2.95	−14.9	0.69
HP	Daymet	0.96	0.97	−0.1	0.91	0.98	0.73	0.7	0.95
	PRISM	0.96	0.94	1.3	0.94	0.97	0.85	0.3	0.95
	MERRA	0.71	2.52	0.3	0.71	0.9	2.27	0.5	0.65
	ERA5	0.59	2.89	−1	0.59	0.9	3	1.3	0.47
LRP	Daymet	0.86	1.82	−3.4	0.85	0.97	1.4	−4	0.88
	PRISM	0.92	1.34	−1.4	0.91	0.98	1.14	−4.9	0.93
	MERRA	0.73	2.37	−3.4	0.7	0.89	2.31	−3	0.88
	ERA5	0.8	2.08	−1.8	0.79	0.92	2.46	−4.5	0.75
NC	Daymet	0.94	1.28	−1.7	0.84	0.88	2.74	−2.5	0.83
	PRISM	0.96	0.96	−0.4	0.93	0.97	1.34	−1.5	0.95
	MERRA	0.74	2.49	−1.5	0.72	0.93	1.88	−0.7	0.93
	ERA5	0.51	3.39	−0.9	0.49	0.81	3.28	−1.8	0.8
SC	Daymet	0.75	2.82	−16	0.68	0.76	3.61	−19.1	0.69
	PRISM	0.57	3.33	12.1	0.54	0.7	3.99	−13.3	0.6
	MERRA	0.68	2.75	−15.9	0.6	0.65	3.94	−15.4	0.66
	ERA5	0.63	2.97	−16.4	0.57	0.64	3.99	−13	0.6

TABLE 5 Statistical metrics indicating the performance of the studied datasets in estimating TX10p and TX90p across climate zones.

Climate regions	Datasets	TX10p				TX90p			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.94	1.39	−9.2	0.87	0.98	2.54	−18.8	0.75
	PRISM	0.98	0.98	−7.6	0.91	0.99	1.91	−16.7	0.82
	MERRA	0.71	2.96	−7.9	0.5	0.94	3.19	−18.8	0.77
	ERA5	0.78	2.26	−9.3	0.73	0.87	3.88	−16.3	0.79
HP	Daymet	0.97	0.66	−1	0.96	0.99	0.7	0.3	0.97
	PRISM	0.98	0.56	−1.3	0.96	0.98	1.05	−0.46	0.92
	MERRA	0.94	1.24	−1.7	0.75	0.95	1.73	−0.4	0.93
	ERA5	0.82	1.81	−0.5	0.74	0.97	1.75	0.5	0.8
LRP	Daymet	0.84	1.8	−1.6	0.83	1.98	1.41	−0.9	0.95
	PRISM	0.9	1.34	−1.1	0.88	0.99	0.89	−1	0.99
	MERRA	0.69	2.52	−1.6	0.68	0.89	2.87	−2	0.89
	ERA5	0.94	1.06	−1.8	0.93	0.96	1.79	−0.9	0.95
NC	Daymet	0.93	1.32	0.7	0.93	0.95	2.22	−1.9	0.95
	PRISM	0.96	1.05	0.4	0.95	0.99	0.99	−1.2	0.96
	MERRA	0.85	1.98	0.8	0.84	0.96	2.05	0.3	0.94
	ERA5	0.91	1.49	−0.9	0.89	0.84	3.95	0.3	0.81
SC	Daymet	0.88	1.96	−7.66	0.85	0.92	3.79	−12.3	0.85
	PRISM	0.94	1.42	−6.8	0.91	0.94	3.37	−12.8	0.84
	MERRA	0.56	3.82	−4.6	0.52	0.8	5.57	−12.8	0.75
	ERA5	0.69	2.86	−8.1	0.67	0.64	6.98	−10.8	0.56

TABLE 6 Performance of the datasets in estimating WSDI and CSDI across climate zones.

Climate regions	Datasets	WSDI				CSDI			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.98	7.65	−69.7	0.13	0.57	7.52	−85	0.24
	PRISM	0.98	4.25	−41.2	0.54	0.63	2.19	−40	0.42
	MERRA	0.93	9.87	−70.14	0.35	0.46	3.88	−90	0.24
	ERA5	0.77	9.97	−32.8	0.49	0.34	3.41	−70	0.23
HP	Daymet	0.94	3.4	25.2	0.71	0.73	2.22	−3.2	0.68
	PRISM	0.93	3.4	26.1	0.7	0.59	2.73	−6.5	0.56
	MERRA	0.96	3.04	25.2	0.68	0.26	4.87	−32.3	0.21
	ERA5	0.94	3.26	20.2	0.78	0.26	5.5	−51.6	0.18
LRP	Daymet	0.92	4.6	12	0.79	0.73	3.31	−99	0.42
	PRISM	0.95	3.84	23.9	0.75	0.41	2.8	−111	0.3
	MERRA	0.87	6.57	29.1	0.5	0.36	2.8	−89	0.25
	ERA5	0.86	5.83	9.4	0.82	0.43	6.03	101.3	0.33
NC	Daymet	0.96	5	29.6	0.62	0.67	2.48	−32	−0.4
	PRISM	0.97	5.25	44.2	0.52	0.67	2.29	76.5	0.44
	MERRA	0.92	7.05	32.7	0.51	0.48	3.39	28	0.22
	ERA5	0.83	8.78	43.1	0.39	0.43	5.93	−76	0.33
SC	Daymet	0.89	10.07	−54.8	0.39	0.52	2.91	−10.7	0.47
	PRISM	0.93	6.49	−20.6	0.78	0.63	3.33	−39.3	0.46
	MERRA	0.74	12.99	−61.9	0.32	0.54	3	−17.9	0.43
	ERA5	0.6	13.19	9.7	0.48	0.55	2.42	25	0.47

TABLE 7 Performance of the datasets in estimating SU25 and FD0 across the climate zones.

Climate regions	Datasets	SU25				FD0			
		R	RMSE	PBIAS	KGE	R	RMSE	PBIAS	KGE
ET	Daymet	0.95	4.36	−0.8	0.87	0.71	9.98	−62	0.26
	PRISM	0.98	3.15	1	9.2	0.95	3.71	−24.7	0.75
	MERRA	0.83	10.23	2.9	0.72	0.67	5.92	−24.7	0.59
	ERA5	0.83	10.23	3	0.77	0.64	7.82	46.8	0.32
HP	Daymet	0.98	2.23	−0.3	0.93	0.98	2.44	−0.5	0.98
	PRISM	0.97	2.6	0.5	0.88	0.98	3.25	−2	0.95
	MERRA	0.89	7.05	−2.7	0.79	0.86	7	24.46	0.68
	ERA5	0.9	6.42	1.6	0.68	0.64	9.61	49	0.51
LRP	Daymet	0.93	6.57	−2.6	0.91	0.96	6.23	−9.5	0.68
	PRISM	0.95	4.15	−1.1	0.94	0.95	6.7	−11	0.74
	MERRA	0.7	11.88	−4.4	0.69	0.85	8.64	12	0.63
	ERA5	0.86	10.78	−4.6	0.85	0.85	23.68	44.9	0.61
NC	Daymet	0.91	5.12	0.6	0.91	0.91	4.21	2.6	0.88
	PRISM	0.94	4.05	0	0.94	0.93	3.74	−1.6	0.92
	MERRA	0.81	7.46	0.6	0.81	0.79	8.83	13.8	0.75
	ERA5	0.75	8.91	1.7	0.75	0.79	9.18	21.3	0.74
SC	Daymet	0.92	8.9	−2.9	0.84	0.86	4.33	−14.5	0.77
	PRISM	0.92	7.24	−1.7	0.82	0.81	5.02	−14.8	0.76
	MERRA	0.69	12.44	−1.6	0.68	0.67	5.8	13.3	0.59
	ERA5	0.67	12.34	1.9	0.65	0.67	10.94	50.96	0.54

For estimating CSDI, Daymet is the leading dataset in the HP and LRP climate zones. It ranks as the second-best performing dataset in the SC zone (Figure 7). Alongside ERA5, Daymet is also the second-best performing dataset in the ET zone. Together with PRISM, Daymet ranks as the top dataset in the NC zone. The lowest-performing datasets are MERRA (ET), ERA5 (HP and NC), PRISM and ERA5 (LRP), and PRISM and MERRA (SC) (Figure 7). In the SC zone, ERA5 is the leading dataset and holds a comparable second-place ranking with Daymet in the ET zone. Similarly, MERRA outperforms PRISM in the LRP zone. These results underscore that, in certain climate zones, reanalysis datasets demonstrate better or comparable performance relative to gauge-based datasets. Overall, Daymet emerges as the best-performing dataset, while ERA5 is the poorest performer across most climate zones (Figure 7).

Regarding estimation bias for CSDI, all datasets exhibit overestimation bias in the ET and HP climate zones (Table 6). Similarly, in the LRP and SC zones, all datasets except ERA5 show overestimation bias. In the NC zone, Daymet and ERA5 display overestimation bias, whereas the other two datasets exhibit underestimation bias. The highest estimation biases are observed in PRISM (LRP, NC, and SC), MERRA (ET), and ERA5 (HP). In contrast, the lowest estimation biases are found in Daymet (HP and SC), MERRA (LRP and NC), and PRISM (ET) (Table 6).

For SU25 estimation, except for the HP zone, PRISM is the leading dataset in four climate zones, whereas Daymet is the second-performing dataset in these zones (Figure 7). In the HP zone, Daymet is the top-performing dataset, whereas PRISM is situated second. The poorest-performing datasets include MERRA in the HP and LRP zones, and ERA5 in the NC and SC zones. Both MERRA and ERA5 exhibit similarly low performance in the ET zone. Overall, PRISM emerges as the best dataset across most studied climate zones, while MERRA and ERA5 are equally the poorest performers (Figure 7).

Regarding estimation bias for SU25, all studied datasets display overestimation bias in the LRP zone (Table 7). Additionally, except for ERA5, the other datasets exhibit overestimation bias in the SC zone. In the NC zone, all datasets except PRISM, which shows no bias, have underestimation bias. Similarly, in the ET zone, all datasets except Daymet tend to underestimate SU25. In the HP zone, Daymet and MERRA exhibit overestimation bias, while the remaining datasets show underestimation bias (Table 7). The highest estimation biases are observed in ERA5 (ET, LRP, and NC), MERRA (HP), and Daymet (SC). In contrast, the lowest estimation biases are found in Daymet (ET and HP), PRISM (LRP and NC), and MERRA (SC).

The performance of datasets for estimating FD0 varies notably across the Brazos River Basin climate zones (Figure 7). Daymet demonstrates superior performance in the HP, LRP, and SC zones, yet ranks lowest in the ET zone, highlighting zonal differences in its effectiveness. In contrast, reanalysis datasets such as MERRA and ERA5 outperform Daymet-a gauge-based dataset-in the ET zone. PRISM consistently ranks as the best dataset in the ET and NC zones and holds second place in the remaining three climate zones. ERA5 performs poorly in the HP, LRP, NC, and SC zones. Overall, Daymet emerges as the most reliable dataset for FD0 estimation, while ERA5 generally underperforms across most zones (Figure 7).

Regarding estimation bias, Daymet and PRISM tend to overestimate FD0 in the HP, LRP, and SC zones, whereas MERRA and ERA5 exhibit underestimation bias in these areas (Table 7). In the NC zone, all datasets except PRISM underestimate FD0. Conversely, in the ET zone, all datasets except ERA5 tend to overestimate. This variation underscores that the direction and magnitude of estimation bias are region-dependent. The highest bias occurs with Daymet in the ET zone, while ERA5 exhibits the greatest bias in the other zones. Conversely, the lowest estimation biases are associated with PRISM in the NC zone, Daymet in HP and LRP, MERRA in SC, and both PRISM and MERRA in ET.

5 Discussion

The performance of the evaluated temperature datasets varies notably across distinct climate zones. For instance, PRISM demonstrated the highest accuracy in estimating daily Tmax in the HP, LRP, and SC climate zones; however, it exhibited the poorest performance in the ET and NC zones. Similarly, ERA5 was the most accurate dataset for estimating WSDI in the HP climate zone, but it is the poorest dataset in the NC climate zone. These spatial inconsistencies in dataset performance across climate zones observed in our study align with findings reported in other studies conducted in North America (Smith et al., 2025; Tarek et al., 2020) and globally (Gashaw et al., 2023; McNicholl et al., 2022). For example, Smith et al. (2025) reported that the performance of the dataset varies across different regions in the United States.

The variations in dataset performance across climate regions and temperature indices have important implications for climate studies, impact assessments, and decision-making processes (Bhattacharyya et al., 2025; Smith et al., 2025; Schreiner-McGraw and Ajami, 2022). First, the findings underscore the need for zonal-specific validation before selecting a dataset for applications such as hydrological modeling, agricultural planning, or extreme weather analysis (Tadesse et al., 2025; Gashaw et al., 2023). Relying on a single dataset without considering its spatial performance may lead to biased or unreliable outcomes, particularly in areas where the dataset underperforms (Xie et al., 2025). Additionally, the observed inconsistencies highlight the importance of multi-dataset approaches or ensemble methods to improve reliability and reduce uncertainty in climate-related analyses.

The findings of our study also reveal that the most suitable dataset for estimating temperature varies across daily, monthly, and annual temporal scales. For example, PRISM best estimated daily Tmax, while Daymet outperformed at monthly and annual scales. Conversely, Daymet was most accurate for daily Tmin, while PRISM performed better at monthly and annual temporal scales across most climate zones. The results also indicated that the widely used reanalysis datasets, such as ERA5 and MERRA are neither good nor poor for both Tmax and Tmin. In our study, MERRA exhibits the poorest performance for Tmax, while ERA5 performs the worst for Tmin. These findings are consistent with Tadesse et al. (2025)'s evaluations of ERA5 and MERRA, who reported that the most effective datasets for estimating Tmax and Tmin at the daily scale in Ethiopia are not identical. Given the widespread use of ERA5 and MERRA globally, applying the same dataset for both Tmax and Tmin may introduce greater uncertainty compared to using the most suitable dataset for each variable separately.

Our findings also demonstrate that although PRISM is better than Daymet for estimating most temperature extreme indices, Daymet outperforms PRISM in computing certain temperature extremes.

These results highlight the importance of selecting datasets based on their intended application. Consistent with our findings, other studies have also reported similar conclusions (Smith et al., 2025; Walton and Hall, 2018). For instance, Smith et al. (2025) evaluated PRISM and MERRA with other gridded datasets in the United States, indicating the superior performance of some datasets for specific hazards. In addition, a study in India also indicated that no single gridded temperature dataset is better for estimating all temperature extremes (Bhattacharyya et al., 2025).

The results of this study also reveal that reanalysis datasets generally outperform gauge-based gridded datasets in specific contexts. For instance, both MERRA and ERA5 demonstrate superior performance compared to Daymet and PRISM in estimating daily Tmin in the ET zone. Similarly, for estimating WSDI in the HP zone, ERA5 and MERRA also outperformed the gauge-based counterparts. The finding suggests that reanalysis products may offer a more reliable alternative in regions with sparse or unevenly distributed ground observations (Yilmaz, 2023; Zandler et al., 2020). The findings also highlight the importance of evaluating dataset performance regionally rather than assuming universal applicability, thereby supporting more informed and context-specific dataset selection in climate-related studies.

Comparisons of the evaluated reanalysis datasets revealed that ERA5 for Tmax and MERRA for Tmin at daily, monthly, and annual temporal scales are performing better in most climate zones. Aligned to this finding, ERA5 is better than MERRA for estimating monthly Tmax in the Upper Blue Nile Basin of Ethiopia (Alaminie et al., 2021). Similar to our finding, MERRA is better than ERA5 for estimating Tmin at daily, monthly, and annual temporal scales in most agro-ecological zones (AEZs) of Ethiopia (Tadesse et al., 2025). However, another finding in the Bale eco-regions of Ethiopia indicated that MERRA is better than ERA5 for both Tmax and Tmin at two AEZs at daily, monthly, and annual temporal scales, but in the other AEZ, ERA5 has better performance (Gashaw et al., 2023). In addition, unlike our finding, ERA5 is better than MERRA for estimating temperature both at daily and monthly scales in Turkey (Hasan Karaman and Akyürek, 2023). The results from our study and others indicated that either ERA5 or MERRA may not show consistently better performance in simulating both Tmax and Tmin at all temporal scales and regions.

In our study, MERRA is also the best dataset for estimating most of the temperature extreme indices in most climate zones of the study area. In addition, similar to the better performance of MERRA compared to ERA5 for estimating most of the temperature extremes in our study, other studies also reported the better performance of MERRA compared to ERA5 for estimating temperature extremes (Ullah et al., 2024). Unlike our findings, ERA5 is better than MERRA for estimating most of the temperature extreme indices in India (Bhattacharyya et al., 2025). The finding also indicates that the performance of the reanalysis products for estimating temperature extremes is region-specific.

The findings also indicate that the direction and magnitude of estimation bias vary across climate zones. Most of the studied datasets for Tmax in the five climate zones at daily, monthly, and annual temporal scales, and for Tmin in ET, HP and LRP climate zones demonstrate underestimation bias. Conversely, most of the datasets in the SC zone demonstrate overestimation bias for daily, monthly, and annual Tmin. Consistent with these observations, several studies have reported that the estimation bias of datasets differs by geographic region (Smith et al., 2025; McNicholl et al., 2022; Tarek et al., 2020). These variations may stem from limitations in the algorithms used to generate the datasets, which may perform differently depending on the regional characteristics. Additionally, discrepancies in estimation bias across climate zones are likely influenced by the density and distribution of meteorological stations (Mankin et al., 2025; Yin et al., 2014).

6 Conclusion

This study evaluated the performances of four gridded temperature datasets in simulating daily, monthly, and annual Tmax and Tmin values as well as temperature extremes across the climate zones of the Brazos River Basin. Performance, benchmarked against GHCN observations, varied across climate zones, temporal scales, and extreme indices. For Tmax, PRISM delivered the highest accuracy at the daily scale, while Daymet showed marginally better skill at monthly and annual scales across most climate zones. For Tmin, Daymet was superior at the daily scale, whereas PRISM dominated at monthly and annual scales. MERRA for Tmax and ERA5 for Tmin consistently ranked lowest across most climate zones and all studied temporal scales. In estimating temperature extremes, PRISM outperformed the other datasets for estimating most indices, whereas ERA5 generally proved least reliable in most of them. However, reanalysis products occasionally surpassed gauge-based datasets in certain zones.

Because PRISM performs best for daily Tmax and Daymet for daily Tmin, using a single dataset for both variables in daily climate and meteorological studies may introduce uncertainty relative to applying the optimal dataset for each variable separately. Similar uncertainties arise in monthly and annual analyses when either PRISM or Daymet is used exclusively for both Tmax and Tmin, rather than selecting the most suitable dataset for each variable. Moreover, reliance on Daymet for analyses of all temperature extremes may increase uncertainty compared to PRISM. These uncertainties have important policy implications, as they can affect the reliability of climate risk assessments, water resource planning, infrastructure design standards, and adaptation strategies that depend on accurate characterization of temperature variability and extremes. Accordingly, the study recommends selecting the top-performing dataset for Tmax and Tmin separately, tailored to the study's temporal focus. As neither Daymet nor PRISM consistently demonstrates superior performance across all temporal scales, it is recommended to revisit and enhance their underlying algorithms for Tmax and/or Tmin to ensure more consistent accuracy from daily to annual resolutions.

This study can play an important role in climate and meteorological research, particularly for regions with sparse observational networks. However, the study contains certain limitations. First, the number of meteorological stations used to represent each climate zone is relatively small, as few stations meet the selection criteria outlined in the methodology. In addition, the distribution of stations across climate zones is not uniform. As a result, limited gauge availability (station density) in certain zones may influence the results, particularly bias. Therefore, future research is recommended to evaluate these datasets using a more balanced and comparable number of meteorological stations within each climate zone. Moreover, although understanding the reasons why certain datasets perform better or worse, such as the influence of topography, climate zone characteristics, and data assimilation methods, is important, such analyses are beyond the scope of this study. Future research addressing these factors is therefore strongly encouraged.

Data availability statement

The employed GHCN (https://www.ncei.noaa.gov/maps/daily/), Daymet (https://daymet.ornl.gov/getdata), PRISM (https://prism.oregonstate.edu/), MERRA (https://power.larc.nasa.gov/data-access-viewer/), and ERA5 (https://climexp.knmi.nl/start.cgi) maximum and minimum temperature datasets for this study are freely accessible in the given websites.

Author contributions

TGT: Conceptualization, Visualization, Writing – original draft, Formal analysis, Investigation, Writing – review & editing, Data curation, Methodology, Software. RLR: Investigation, Conceptualization, Writing – review & editing, Methodology, Validation, Supervision. GWT: Supervision, Methodology, Writing – review & editing, Investigation, Conceptualization, Validation. STW: Data curation, Software, Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Acknowledgments

This study used maximum and minimum temperature data from GHCN, Daymet, PRISM, MERRA, and ERA5, and hence the authors greatly acknowledge the developers of these datasets.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Ahn, Y., Tuholske, C., and Parks, R. M. (2024). Comparing approximated heat stress measures across the United States. *Geohealth* 8:e2023GH000923. doi: 10.1029/2023GH000923

Alaminie, A. A., Tilahun, S. A., Legesse, S. A., Zimale, F. A., and Jury, M. R. (2021). Evaluation of past and future climate trends under CMIP6 scenarios for the UBNB (Abay), Ethiopia. *Water* 13:2110. doi: 10.3390/w13152110

Almeida, M., and Coelho, P. (2023). A first assessment of ERA5 and ERA5-land reanalysis air temperature in Portugal. *Int. J. Climatol.* 43, 6643–6663. doi: 10.1002/joc.8225

Al-Sakkaf, A. S., Zhang, J., Yao, F., Hamed, M. M., Simbi, C. H., Ahmed, A., et al. (2024). Assessing exposure to climate extremes over the Arabian peninsula using ERA5 reanalysis data: spatial distribution and temporal trends. *Atmos. Res.* 300:107224. doi: 10.1016/j.atmosres.2024.107224

Andrade, J. M., Ribeiro Neto, A., Nóbrega, R. L. B., Rico-Ramirez, M. A., and Montenegro, S. M. G. L. (2024). Efficiency of global precipitation datasets in tropical and subtropical catchments revealed by large sampling hydrological modelling. *J. Hydrol.* 633:131016. doi: 10.1016/j.jhydrol.2024.131016

Aniley, E., Zemen, A., Belay, M., Birku, B., and Taye, H. (2025). Evaluating rainfall forecasting capability of multi-source weather (MSWX) product over rainfall regimes of Ethiopia. *Geocarto Int.* 40:2480298. doi: 10.1080/10106049.2025.2480298

Araghi, A., Martinez, C. J., Olesen, J. E., and Hoogenboom, G. (2022). Assessment of nine gridded temperature data for modeling of wheat production systems. *Comput. Electron. Agric.* 199:107189. doi: 10.1016/j.compag.2022.107189

Bhattacharyya, S., Hassan, M. A., Sreekesh, S., and Choudhary, V. (2025). How well do the reanalysis datasets capture hot and cold extremes and their trends in India? *Atmos. Res.* 321:108073. doi: 10.1016/j.atmosres.2025.108073

Birkel, S. D., Mayewski, P. A., Perry, L. B., Seimon, A., and Andrade-Flores, M. (2022). Evaluation of reanalysis temperature and precipitation for the Andean altiplano and adjacent cordilleras. *Earth Space Sci.* 9:e2021EA001934. doi: 10.1029/2021ea001934

Bodjrenou, R., Cohard, J.-M., Hector, B., Lawin, E. A., Chagnaud, G., Danso, D. K., et al. (2023). Evaluation of reanalysis estimates of precipitation, radiation, and temperature over Benin (West Africa). *J. Appl. Meteorol. Climatol.* 62, 1005–1022. doi: 10.1175/jamc-d-21-0222.1

Buban, M. S., Lee, T. R., and Baker, C. B. (2020). A comparison of the U.S. climate reference network precipitation data to the parameter-elevation regressions on independent slopes model (PRISM). *J. Hydrometeorol.* 21, 2391–2400. doi: 10.1175/jhm-d-19-0232.1

Buuren, S. v., Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., and Rockenschaub, P. (2025). Multivariate imputation by chained equations. Version 3.18.0. *J. Stat. Software* 45:i03. doi: 10.18637/jss.v045.i03

Damavandi, H. G., Shah, R., Stampoulis, D., Wei, Y., Boscovic, D., and Sabo, J. (2019). Accurate prediction of streamflow using long short-term memory network: a case study in the Brazos River basin in Texas. *Int. J. Environ. Sci. Dev.* 10, 294–300. doi: 10.18178/ijesd.2019.10.10.1190

Deng, Y., Wang, X., Ruan, H., Lin, J., Chen, X., Chen, Y., et al. (2024). The magnitude and frequency of detected precipitation determine the accuracy performance of precipitation data sets in the high mountains of Asia. *Sci. Rep.* 14:17251. doi: 10.1038/s41598-024-67665-8

Drisya, J., and Al-Zubari, W. K. (2025). Investigating the future precipitation changes over the Kingdom of Bahrain using CMIP6 projections. *Earth Syst. Environ.* 10, 275–298. doi: 10.1007/s41748-025-00578-2

Du, Y., Wang, D., Zhu, J., Wang, D., Qi, X., and Cai, J. (2022). Comprehensive assessment of CMIP5 and CMIP6 models in simulating and projecting precipitation over the global land. *Int. J. Climatol.* 42, 6859–6875. doi: 10.1002/joc.7616

Gashaw, T., Worqlul, A., Lakew, H., Taye, M., Seid, A., and Haileslassie, A. (2023). Evaluations of satellite/reanalysis rainfall and temperature products in the bale eco-region (southern Ethiopia) to enhance the quality of input data for hydro-climate studies. *Remote Sens. Appl. Soc. Environ.* 31:100994. doi: 10.1016/j.rsase.2023.100994

Gashaw, T., Worqlul, A. W., Taye, M. T., Lakew, H. B., Seid, A., Ayele, G., et al. (2024). Performance evaluations of CMIP6 model simulations and future projections of rainfall and temperature in the bale eco-region, southern Ethiopia. *Theor. Appl. Climatol.* 155, 5069–5092. doi: 10.1007/s00704-024-04904-y

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* 30, 5419–5454. doi: 10.1175/JCLI-D-16-0758.1

Hasan Karaman, Ç., and Akyürek, Z. (2023). Evaluation of near-surface air temperature reanalysis datasets and downscaling with machine learning based random forest method for complex terrain of Turkey. *Adv. Space Res.* 71, 5256–5281. doi: 10.1016/j.asr.2023.02.006

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146, 1999–2049. doi: 10.1002/qj.3803

Lan, H., Guo, D., Hua, W., Pepin, N., and Sun, J. (2022). Evaluation of reanalysis air temperature and precipitation in high-latitude Asia using ground-based observations. *Int. J. Climatol.* 43, 1621–1638. doi: 10.1002/joc.7937

Lebeza, T. M., Gashaw, T., Bayabil, H. K., van Oel, P. R., Worqlul, A. W., Dile, Y. T., et al. (2024). Performance of specific CMIP6 GCMs for simulating the historical rainfall and temperature climatology of Lake Tana sub-basin, Ethiopia. *Sci. Afr.* 26:e02387. doi: 10.1016/j.sciaf.2024.e02387

Lee, K., and Singh, V. P. (2020). Analysis of uncertainty and non-stationarity in probable maximum precipitation in Brazos River basin. *J. Hydrol.* 590:125526. doi: 10.1016/j.jhydrol.2020.125526

Luo, Z., Zhang, S., Shao, Q., Wang, L., Wang, S., and Wang, L. (2024). A new method to improve precipitation estimates by blending multiple satellite/reanalysis-based precipitation products and considering observations and terrestrial water budget balance. *J. Hydrol.* 635:131188. doi: 10.1016/j.jhydrol.2024.131188

Mankin, K. R., Mehan, S., Green, T. R., and Barnard, D. M. (2025). Review of gridded climate products and their use in hydrological analyses reveals overlaps, gaps, and the need for a more objective approach to selecting model forcing datasets. *Hydrol. Earth Syst. Sci.* 29, 85–108. doi: 10.5194/hess-29-85-2025

McNicholl, B., Lee, Y. H., Campbell, A. G., and Dev, S. (2022). Evaluating the reliability of air temperature from ERA5 reanalysis data. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/lgrs.2021.3137643

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *J. Atmos. Ocean. Technol.* 29, 897–910. doi: 10.1175/jtech-d-11-00103.1

Mihalevich, B. A., Neilson, B. T., and Buahin, C. A. (2022). Evaluation of the ERA5-land reanalysis data set for process-based river temperature modeling over data sparse and topographically complex regions. *Water Resour. Res.* 58:e2021WR031294. doi: 10.1029/2021wr031294

Mohammed, J. A., Gashaw, T., Worku Tefera, G., Dile, Y. T., Worqlul, A. W., and Addisu, S. (2022). Changes in observed rainfall and temperature extremes in the upper Blue Nile Basin of Ethiopia. *Weather Clim. Extremes* 37:100468. doi: 10.1016/j.wace.2022.100468

Newman, A. J., Kalb, C., Chakraborty, T. C., Fitch, A., Darrow, L. A., Warren, J. L., et al. (2024). The high-resolution urban meteorology for impacts dataset (HUMID) daily for the conterminous United States. *Sci Data* 11:1321. doi: 10.1038/s41597-024-04086-2

Peng, Y., Duan, A., Zhang, C., Tang, B., and Zhao, X. (2023). Evaluation of the surface air temperature over the Tibetan plateau among different reanalysis datasets. *Front. Environ. Sci.* 11:1152129. doi: 10.3389/fenvs.2023.1152129

Rivera, J. A., and Arnould, G. (2020). Evaluation of the ability of CMIP6 models to simulate precipitation over southwestern South America: climatic features and long-term trends (1901–2014). *Atmos. Res.* 241. doi: 10.1016/j.atmosres.2020.104953

Schreiner-McGraw, A. P., and Ajami, H. (2022). Combined impacts of uncertainty in precipitation and air temperature on simulated mountain system recharge from an integrated hydrologic model. *Hydrol. Earth Syst. Sci.* 26, 1145–1164. doi: 10.5194/hess-26-1145-2022

Smith, E., Grant, N., Luo, X., and Diaz, D. (2025). Evaluating the ability of gridded climate datasets to capture temperature and precipitation trends and extremes. *Sci. Rep.* 15:12607. doi: 10.1038/s41598-025-97570-7

Tadesse, T. A., Tarkegn, T. G., Ray, R. L., Tefera, G. W., Demessie, S. F., Shawul, A. A., et al. (2025). Performances of reanalysis products in representing the temperature climatology of Ethiopia. *Theor. Appl. Climatol.* 156:70. doi: 10.1007/s00704-024-05332-8

Tan, M. L., Armanuos, A. M., Ahmadianfar, I., Demir, V., Heddam, S., Al-Areeq, A. M., et al. (2023). Evaluation of NASA power and ERA5-land for estimating tropical precipitation and temperature extremes. *J. Hydrol.* 624:129940. doi: 10.1016/j.jhydrol.2023.129940

Tarek, M., Brissette, F. P., and Arsenault, R. (2020). Large-scale analysis of global gridded precipitation and temperature datasets for climate change impact studies. *J. Hydrometeorol.* 21, 2623–2640. doi: 10.1175/jhm-d-20-0100.1

Tefera, G. W., and Ray, R. L. (2024). Hydrology and hydrological extremes under climate change scenarios in the bosque watershed, north-Central Texas, USA. *Environ. Sci. Pollut. Res. Int.* 31, 40636–40654. doi: 10.1007/s11356-023-27477-1

Tefera, G. W., Ray, R. L., and Wootten, A. M. (2024). Evaluation of statistical downscaling techniques and projection of climate extremes in Central Texas, USA. *Weather Clim. Extrem.* 43:100637. doi: 10.1016/j.wace.2023.100637

Tefera, G. W., Ray, R. L., and Wootten, A. M. (2025). Changes in plant hardiness zones under climate change scenarios in the south-Central United States. *J. Appl. Meteorol. Climatol.* 64, 339–351. doi: 10.1175/jamc-d-24-0080.1

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S. C., Wei, Y., and Wilson, B. E. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Sci Data* 8:190. doi: 10.1038/s41597-021-00973-0

Ullah, W., Alabduoli, K., Ullah, S., Al-Ghamdi, S. G., Alhebsi, K., Almazroui, M., et al. (2024). Comparison of 2-m surface temperature data between reanalysis and observations over the Arabian peninsula. *Atmos. Res.* 311:107725. doi: 10.1016/j.atmosres.2024.107725

Valappil, N. K. M., Hamza, V., and de Oliveira Junior, J. F. (2023). Evaluation of daily average temperature trends in Kerala, India, using MERRA-2 reanalysis data: a climate change perspective. *Environ. Sci. Pollut. Res. Int.* 30, 26663–26686. doi: 10.1007/s11356-022-23895-9

Vogl, A. L., and Lopes, V. L. (2009). Impacts of water resources development on flow regimes in the Brazos River. *Environ. Monit. Assess.* 157, 331–345. doi: 10.1007/s10661-008-0538-5

Walton, D., and Hall, A. (2018). An assessment of high-resolution gridded temperature datasets over California. *J. Clim.* 31, 3789–3810. doi: 10.1175/jcli-d-17-0410.1

Wang, T., Hamann, A., and Sang, Z. (2024). Monthly high-resolution historical climate data for North America since 1901. *Int. J. Climatol.* 45:e8726. doi: 10.1002/joc.8726

Wang, S., Li, G.-C., Zhang, Z.-H., Zhang, W.-Q., Wang, X., Chen, D., et al. (2025). Recent warming trends in Antarctica revealed by multiple reanalysis. *Adv. Clim. Chang. Res.* 16, 447–459. doi: 10.1016/j.accre.2025.03.003

Wolaver, B. D., Reynolds, L. V., Caldwell, T. G., Bongiovanni, T., Paul Pierre, J., Breton, C., et al. (2024). Hydrologic changes in the Brazos River basin and implications for Great Plains fishes. *J. Hydrol.* 629:130351. doi: 10.1016/j.jhydrol.2023.130351

Xie, Y., Kong, D., Zhang, Y., Zhong, Y., Ma, N., Gong, R., et al. (2025). Evaluation of the performance of multiple reanalysis forcing data in potential evapotranspiration estimation and its implication for actual evapotranspiration modeling. *J. Hydrol.* 660:133472. doi: 10.1016/j.jhydrol.2025.133472

Yilmaz, M. (2023). Accuracy assessment of temperature trends from ERA5 and ERA5-land. *Sci. Total Environ.* 856:159182. doi: 10.1016/j.scitotenv.2022.159182

Yin, H., Donat, M. G., Alexander, L. V., and Sun, Y. (2014). Multi-dataset comparison of gridded observed temperature and precipitation extremes over China. *Int. J. Climatol.* 35, 2809–2827. doi: 10.1002/joc.4174

Zandler, H., Senftl, T., and Vanselow, K. A. (2020). Reanalysis datasets outperform other gridded climate products in vegetation change analysis in peripheral conservation areas of Central Asia. *Sci. Rep.* 10:22446. doi: 10.1038/s41598-020-79480-y

Zhang, X., and Yang, F. (2004). "RClimDex (1.0) user manual" in Climate research branch Environment Canada. eds. X. Zhang and F. Yang (Downsview, ON: Environment Canada).