

Evaluation of the WRF Model's Performance at Gray-Zone Resolution in Simulating Climate over the Tibetan Plateau

YIJIA LI,^a XUEJIA WANG^{①, a, b}, TINGHAI OU,^c JIAYU WANG,^a YUE YUAN,^a XIAOHUA GOU,^{a, b} GUOJIN PANG,^d MEIXUE YANG,^c HANS W. LINDERHOLM,^c DELIANG CHEN,^f AND MENGQIAN LU^g

^a Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, China

^b Gansu Liancheng Forest Ecosystem Field Observation and Research Station, Lanzhou University, Lanzhou, China

^c Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden

^d Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China

^e Key Laboratory of Cryospheric Science and Frozen Soil Engineering, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China

^f Department of Earth System Science, Tsinghua University, Beijing, China

^g Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Hong Kong, China

(Manuscript received 25 July 2025, in final form 15 January 2026, accepted 25 February 2026)

ABSTRACT: The Tibetan Plateau (TP), also known as the “Water Tower of Asia,” profoundly impacts regional and global climates. Existing climate models show substantial biases over the TP, primarily due to low spatial resolution, deficient driving data, and inadequate model domains. Leveraging meteorological station data, the CN05.1 meteorological dataset, and reanalysis datasets, we comprehensively evaluate the performance of the Weather Research and Forecasting Model at gray-zone resolution (9 km) (WRF9km) in simulating TP air temperature and precipitation during 1980–2019 and identify bias causes. WRF9km effectively captures observed spatial air temperature patterns but shows a cold bias. This is mainly due to overestimated surface albedo (accounting for 64%), along with underestimated downward radiation and ground heat fluxes. WRF9km also simulates observed spatial precipitation patterns well, with significant seasonal correlations; however, precipitation biases exhibit pronounced spatial heterogeneity. Compared to CN05.1, precipitation is overestimated over the southern and eastern slopes of the TP, while it is underestimated over the interior TP, especially in the western part. These biases primarily arise from inadequate characterization of wind field dynamics and moisture transport within the model framework. Meanwhile, WRF9km captures annual precipitation variation with minor deviations, although it does not fully reproduce the temporal trends over the TP. Overall, compared to driving ERA5 data, WRF9km yields only marginal improvement for the cold bias but reduces the regional mean precipitation wet bias by 79%, particularly over the southern and eastern slopes. This evaluation provides critical insights for dynamic downscaling in complex terrain, highlighting the need for improved surface albedo parameterizations and higher-quality input data.

KEYWORDS: Tibetan Plateau; Dynamics; Climate; Precipitation; Temperature; Climate models

1. Introduction

Climate warming is one of the focal points of the international research community. Global temperatures have risen by approximately 1.1°C since the Industrial Revolution (IPCC 2023), where the increase varies with latitude and altitude (Pepin et al. 2015). Regarding precipitation, in general wet regions are experiencing increased precipitation, while decreasing trends are observed in dry regions (Polson et al. 2013), although studies have identified more complex regional variations (e.g., Ghosh 2018; Wang et al. 2018). Along with these changes, increases in the frequency and intensity of extreme weather events such as heatwaves, floods,

droughts, and storms have been observed, resulting in significant impacts (IPCC 2023; Wang et al. 2022). With the aim of better coping with such challenges, regional studies provide crucial insights into local climate change and its impacts (Giorgi 2019).

The Tibetan Plateau (TP), also known as the “Water Tower of Asia,” is the source of several large Asian rivers, providing water resources for billions of people living downstream, and plays a crucial role in regulating Asian and even global hydrological cycles (Immerzeel et al. 2010; Yao et al. 2022). The TP is a typical alpine region, exhibiting substantial ecological vulnerability and sensitivity to climate change, while holding great ecological and environmental significance for China and the larger Asian region (Latif et al. 2019; X. Wang et al. 2025). The TP has exhibited a marked warming trend over recent decades, with a warming rate twice the global average (Yao et al. 2019, 2022), resulting in glacier retreat, permafrost degradation, and intensified hydrological extremes. This threatens regional water security and places biodiversity and ecosystems under increasing pressure (Wang et al. 2022; Xu et al. 2024). Therefore, studying TP climate change is of great importance to better understand regional climate dynamics and their associated impacts.

Climate models simulate the climate system based on the physical, chemical, and biological properties of its components, along

 Denotes content that is immediately available upon publication as open access.

 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-25-0419.s1>.

Corresponding author: Xuejia Wang, wangxuejia@lzu.edu.cn

DOI: 10.1175/JCLI-D-25-0419.1

© 2026 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

with their recognized interactions and feedback processes. Climate models are essential tools when attempting to understand historical climate dynamics and predict future changes (Jakob et al. 2023). Global climate models (GCMs) are widely used for climate projections; however, their low spatial resolution limits their effectiveness in capturing climate variations that are generally influenced by local topography (Karki et al. 2020; Yimer et al. 2022), particularly over the TP. Previous studies have consistently shown that although the widely used climate models from the Coupled Model Intercomparison Project phase 6 (CMIP6) demonstrate improved simulation accuracy compared to CMIP5 models, they still exhibit persistent cold biases over the TP (Wei et al. 2025). Furthermore, raw GCM outputs tend to underestimate the magnitude of warming in China. In contrast, bias corrections markedly enhance projected warming amplitudes, particularly for high-altitude regions and long-term projections for China (Wei et al. 2025). Nevertheless, existing correction methods fail to fully rectify errors connected to physical processes, such as local valley circulations and land–atmosphere interactions, caused by the low resolution of GCMs. This limitation underscores the necessity for synergistic optimization between dynamic downscaling and statistical corrections (Chen et al. 2022; Wei et al. 2025). Achieving a realistic modeling of the regional characteristics of climate change therefore requires more refined research.

Regional climate models (RCMs), having high horizontal resolution (i.e., a spatial resolution of several kilometers), can effectively reproduce regional and local features and have been widely used in simulating and projecting climate change over the TP (e.g., Giorgi 2019; Niu et al. 2021; Wang et al. 2021a). The high-resolution gridded meteorological data generated by such models represent surface characteristics and explicitly simulate certain cloud and convection processes better than traditional coarser-resolution models. International initiatives, such as the Coordinated Regional Climate Downscaling Experiment, have further advanced these capabilities, with its flagship project—the Convective-Permitting Third Pole—employing ensemble-based kilometer-scale modeling to address challenges specific to the TP, including glacier–atmosphere interactions and extreme snowfall events (Prein et al. 2023). This has markedly enhanced climate simulations over the TP, especially in terms of precipitation, contributing to a better understanding of climate change and the regional water cycle across the TP (Prein et al. 2023).

RCMs effectively reproduce the spatial distribution of temperature and humidity over the TP and substantially improve spatiotemporal precipitation simulation accuracy by explicitly resolving terrain effects, with annual precipitation simulations outperforming CMIP6 in matching observations. Compared with GCMs, RCMs reduce precipitation simulation errors by 30%–50% over the TP (Singh et al. 2019). Despite these advances, multimodel evaluations reveal persistent cold–wet biases in RCMs, with annual-mean temperature underestimated by 1.2°–2.5°C and annual precipitation overestimated by 20%–40% on the central TP. These errors primarily stem from misrepresented convective parameterizations and biased downward shortwave radiation (SR) associated with complex terrains (L. Wang et al. 2025). A critical unresolved challenge lies in the gray-zone resolution range of 4–15 km, where

conventional cumulus schemes can be disabled but cloud microphysics remain underresolved, leading to distorted representations of valley circulations and energy exchanges, an issue that limits the reliability of TP climate projections (Ou et al. 2020).

The Weather Research and Forecasting (WRF) Model, which is one of the most widely used meteorological models globally, has been increasingly used in high-resolution (approaching a few hundred meters) simulations. A recent study employed both the Regional Climate Model, version 4 (RegCM4), and the WRF Model with a horizontal resolution of 30 km to carry out long-term regional climate simulations in China; this demonstrated that the WRF Model excels in simulating annual average temperature and precipitation trends, showcased by a particularly high spatial pattern correlation (Gao et al. 2022; Yang et al. 2025). A second study assessed the performance of the same models at the same horizontal resolution in representing temperature extremes in China; this indicated that the RegCM4 performs better in reproducing average temperature extremes, especially over the TP, while the WRF excels in capturing interannual variations in extreme indices (Kong et al. 2019). Overall, the WRF Model brings notable benefits in simulating climate variables in regions of complex terrain in China, providing essential technical support in the study of regional climate change.

In regions with steep terrain, the choice of horizontal resolution is particularly crucial in regional climate simulations, especially when simulating precipitation and water vapor transport (Lin et al. 2018; Xu et al. 2018). Recent studies have suggested that a spatial resolution of 9 km, within the so-called gray-zone range, strikes a balance between sufficiently detailed spatial simulation and computational cost (Shin and Hong 2013; Lin et al. 2018). Ou et al. (2020) simulated the diurnal cycle of precipitation on the TP using the WRF Model with a spatial resolution of 9 km and found that simulations without the cumulus parameterization scheme performed better than those with the scheme. Kilometer-scale modeling over the TP has been shown to faithfully reproduce diurnal and seasonal cycles, lapse rates, spatial distribution of climate variables, precipitation frequency and intensity, and water vapor transport patterns (Karki et al. 2020; Sugimoto et al. 2021). Convection-permitting simulations at ~3-km grid spacing (Zou and Zhou 2024; Zou et al. 2024) enable an even more faithful representation of the TP climate system. Compared to gray-zone resolutions, these simulations yield more accurate representations of the spring water cycle, seasonal precipitation, and temperature patterns over complex terrain. This provides significant added value for both current climate evaluation and future projections, albeit at a higher computational cost.

In addition to horizontal resolution, the domain configuration is also crucial. In this respect, when compared with the High Asia Refined (HAR) analysis, the updated HAR v2, with an extended domain, provides a better representation of air temperature and the spatial distribution of seasonal-mean precipitation but overestimates the amount of precipitation (Wang et al. 2021c). Most recently, Ou et al. (2023) conducted a regional downscaling study over a broader domain extending to 50°N using the WRF Model at a resolution of 9 km (WRF9km) driven by the fifth generation ECMWF atmospheric reanalysis (ERA5) reanalysis, which indicated that the WRF9km

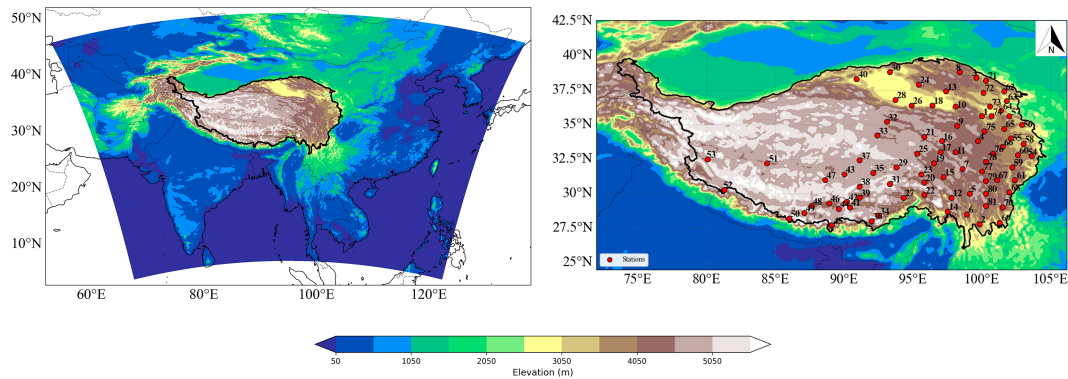


FIG. 1. (left) Model domain and (right) topography and distribution of meteorological stations used in this study. Detailed information on the meteorological stations can be found in Table S1.

substantially reduced the moisture bias in summer precipitation over the northwestern TP.

Previous work applying the WRF at gray-zone resolutions over the TP (e.g., [Ou et al. 2020, 2023](#)) has shown promising results. However, a systematic and quantitative evaluation of model performance is a prerequisite for resolving more complex physical process biases. The aforementioned studies, along with multimodel evaluations (e.g., [L. Wang et al. 2025](#)), have highlighted persistent cold–wet biases and uncertainties in representing precipitation processes. Thus, a comprehensive understanding of how a specific, widely used WRF configuration at 9-km resolution performs across the entire TP—capturing both its strengths and systematic errors—is still needed. It is evident that high-resolution modeling and appropriate domain setting are vital to climate research, especially over the TP. Building upon this previous work, this study aims to 1) quantify the performance of a WRF 9-km simulation in reproducing temperature and precipitation over the TP and 2) identify the causes of its biases. A pivotal finding we present is a substantial 79% reduction in the domain-averaged annual precipitation wet bias compared to the driving ERA5 reanalysis. Furthermore, by decomposing the surface energy balance, we trace the dominant cause of the pervasive cold bias to an overestimated surface albedo. This quantitative evaluation and mechanistic diagnosis provide a reliable benchmark and clear pathways for future model development, ultimately contributing to more trustworthy climate projections for the TP.

2. Methods and data

a. Simulation experiment using the WRF9km model

In this study, we used WRF9km developed by [Ou et al. \(2023\)](#) from the Regional Climate Group, University of Gothenburg, Sweden, to conduct model evaluations over the TP. The selected WRF Model, version 3.7.1, used a nonhydrostatic balance configuration ([Skamarock et al. 2008](#)) to dynamically downscale global reanalysis data, with a model domain centered over the TP. The northern boundary of the model domain (8° – 50° N, 65° – 125° E) ([Fig. 1](#)) was chosen to cover the core of the summer subtropical westerly jet stream in the upper

troposphere. The initial and boundary conditions were provided by the ERA5 reanalysis, updated every 3 h.

Our WRF9km experiment used the New Goddard shortwave radiation scheme, the Rapid Radiative Transfer Model for GCMs ([Iacono et al. 2008](#)), the WRF double-moment 6-class microphysics scheme, the unified Noah land surface model, and the Yonsei University planetary boundary layer scheme ([Hong et al. 2006](#)) as a free run without spectral nudging ([Ou et al. 2025](#)). To optimize computational efficiency, parallel continuous simulations were conducted for 1979–98 and 1998–2019. Comparative analysis of the overlapping year (1998) revealed negligible differences between simulations in regional averages of precipitation, relative humidity, 2-m air temperature, skin temperature, and surface net radiation. Therefore, this 1-yr overlap strategy introduced minimal disturbance to the simulations, allowing the two simulations to be merged into a continuous simulation covering the period 1979–2019 ([Ou et al. 2023](#)).

b. Validation data

To assess the performance of the WRF9km in simulating near-surface air temperature and precipitation, we used the CN05.1 high-resolution gridded daily air temperature and precipitation dataset at a resolution of 0.25° , released by the China National Meteorological Information Center (<https://data.cma.cn/>), as our observational data. The CN05.1 dataset is generated through an anomaly approach, which includes three sequential stages. First, a smoothed $0.25^{\circ} \times 0.25^{\circ}$ gridded climatology (1971–2000 baseline) is created from 2416 Chinese meteorological stations using elevation-corrected thin-plate spline interpolation. Second, daily station anomalies (deviations from climatological norms) are interpolated to the same grid via angular distance weighting, capitalizing on their inherent spatial smoothness. Finally, high-resolution gridded daily meteorological fields are reconstructed by superimposing the gridded anomalies onto the climatological background field ([Wu and Gao 2013](#)). This dataset has been widely used in climate model validation and climate change studies ([Miao and Wang 2020](#); [Wang et al. 2021b, 2025](#)). We also utilized in situ observational data from the existing 81 meteorological stations on the TP provided by the National Climate Center of the China Meteorological Administration (<https://ncc-cma.net/cn>)

to validate our model simulations at a station scale (Fig. 1b and Table S1 in the online supplemental material).

To evaluate potential improvements in the WRF9km simulations compared to other reanalysis data, we compared air temperature, precipitation, specific humidity, and wind components derived and other variables from the ERA5 reanalysis dataset (<https://cds.climate.copernicus.eu/>), at a spatial resolution of $0.25^\circ \times 0.25^\circ$ and covering the period from 1950 onward (Muñoz-Sabater et al. 2021). This dataset undergoes daily updates and monthly quality control procedures, during which preliminary estimates are replaced with validated products (Hersbach et al. 2020).

To examine the causes of temperature biases in the WRF9km simulations, we used the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2), with a spatial resolution of $0.5^\circ \times 0.625^\circ$, developed by the National Aeronautics and Space Administration (<https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>); MERRA2 replaced the original MERRA reanalysis system, integrating the upgraded Goddard Earth Observing System Model, version 5, data assimilation system and updating both the model and global statistical interpolation schemes. It contains meteorological elements such as radiation flux, surface albedo, and water vapor transport, making it an important dataset for understanding global climate change and evaluating climate model performance (Gelaro et al. 2017).

To comprehensively evaluate the performance of the WRF9km over the TP, it was compared with relevant reference datasets, ERA5-Land, and the meteorological forcing dataset for the Third Pole (TPMFD). The ERA5-Land dataset is based on a variety of data sources, including remote sensing and meteorological station observations (<https://cds.climate.copernicus.eu/>), which are obtained from the European Centre for Medium-Range Weather Forecasts. It has a spatial resolution of $0.1^\circ \times 0.1^\circ$ and provides continuous temporal coverage from 1950 to the present (Muñoz-Sabater et al. 2021). The TPMFD is a long-term (1979–2020) and high-resolution ($1/30^\circ$) near-surface meteorological forcing dataset for the Third Pole region (<https://data.tpdc.ac.cn/zh-hans/data/>). In this dataset, 2-m air temperature, 2-m specific humidity, 10-m wind speed, and surface pressure are obtained by integrating ERA5 reanalysis data, short-term high-resolution atmospheric simulations, and ground-station observations, while downward longwave radiation (DLR) is calculated using a semiphysical semiempirical method (Jiang et al. 2025). This dataset serves as a counterpart to the widely used China Meteorological Forcing Dataset.

To evaluate the causes of the cold bias in the WRF9km simulation, we also incorporated satellite-derived snow cover and albedo data from the Moderate Resolution Imaging Spectroradiometer (MODIS). Specifically, we used the MODIS albedo product (MCD43C3), which provides 0.05° data on black-sky and white-sky albedo, and the MOD10A2 V5 snow-cover product (Hall et al. 2007). The MOD10A2 product provides a spatial resolution of 500 m in 8-day composites, representing the maximum snow extent within each period. The MODIS products have been widely validated in numerous studies, showing consistently good agreement with the ground observation data across various regions, including high mountain areas like

the TP (Klein and Barnett 2003; Parajka and Bloeschl 2006; Simic et al. 2004; Pang et al. 2025).

c. Evaluation strategy

The model simulations were collectively assessed in terms of annual averages and seasonal data, with the four seasons being winter [December–February (DJF)], spring [March–May (MAM)], summer [June–August (JJA)], and autumn [September–November (SON)]. To quantitatively evaluate model performance, we compared the bias, root-mean-square error (RMSE), correlation coefficient, and linear trend in the WRF9km simulations with observations (CN05.1) and reanalysis datasets (ERA5, ERA5-Land, and MERRA2). To ensure consistency, we interpolated the WRF9km simulations, ERA5-Land, and MERRA2 reanalysis data onto a common grid resolution of $0.25^\circ \times 0.25^\circ$ to match that of the CN05.1 and ERA5 data. We selected the common time frame of 1980–2019 in all datasets when conducting our analyses. Additionally, we compared the WRF9km simulations and ERA5 data with observational data from 81 stations to avoid potential biases resulting from the relatively low number of stations used in existing gridded datasets.

To better understand the accuracy of WRF9km simulations relative to the commonly used datasets on the TP, we performed quantitative comparisons of the WRF9km with ERA5, MERRA2, ERA5-Land, and TPMFD; this was based on several indicators, including degree of deviation, temporal correlation, and spatial correlation.

The deviation indicators included bias (BIAS) and RMSE, calculated as follows:

$$\text{BIAS} = \frac{1}{N} \sum_{i=1}^N (M_i - O_i), \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2}, \quad (2)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |M_i - O_i|, \quad (3)$$

where M_i indicates the value of the dataset under evaluation at grid point i , O_i indicates the value of the observational dataset CN05.1 at grid point i , and N indicates the total number of grid points.

The correlation indicators included the temporal correlation coefficient (TCC) and spatial correlation coefficient (SCC), describing the linear correlation between two random variables [see Eqs. (4) and (5), respectively]:

$$\text{TCC} = \frac{\sum_{t=1}^T (M_t - \bar{M}_t)(O_t - \bar{O}_t)}{\sqrt{\sum_{t=1}^T (M_t - \bar{M}_t)^2} \sqrt{\sum_{t=1}^T (O_t - \bar{O}_t)^2}}, \quad (4)$$

where M_t indicates the value of the dataset under evaluation at time step t , O_t indicates the value of the observational

dataset CN05.1 at time step t , $\overline{M}$ indicates the temporal mean of the dataset under evaluation, $\overline{O}$ indicates the temporal mean of the observational dataset CN05.1, and T indicates the total length of the time series:

$$SCC = \frac{\sum_{k=1}^K (M_k - \overline{M})(O_k - \overline{O})}{\sqrt{\sum_{k=1}^K (M_k - \overline{M})^2} \sqrt{\sum_{k=1}^K (O_k - \overline{O})^2}}, \quad (5)$$

where M_k indicates the value of the dataset under evaluation at grid point k , O_k indicates the value of the observational dataset CN05.1 at grid point k , $\overline{M}$ indicates the spatial mean of the dataset under evaluation, $\overline{O}$ indicates the spatial mean of the observational dataset CN05.1, and K indicates the total number of grid points.

After all individual indicators were calculated, a comprehensive scoring system was employed to quantitatively evaluate the overall performance of each dataset. The processes included the following steps: 1) Indicator scoring based on rank: For each of the five evaluation indicators, the performance of all datasets was ranked. The dataset with the best performance for a given indicator (e.g., the smallest MAE or the largest TCC) was assigned the highest score of 5. The dataset with the worst performance was assigned the lowest score of 1. The remaining datasets received scores of 2, 3, or 4 based on their percentile rank within the distribution. This nonparametric scoring approach ensures that the scores reflect the relative performance among the evaluated datasets, independent of the absolute magnitude of the indicator values. 2) Score aggregation: The comprehensive score for a dataset was computed as the sum of the standardized scores from all five individual indicators. 3) Interpretation: Since each of the five indicators could contribute a maximum of five points, the maximum attainable comprehensive score was 25. A higher total score indicates that the dataset is, on average, closer to the observations, exhibiting smaller deviations, higher correlation, and more realistic variability. This comprehensive evaluation framework enables both a visual and quantitative assessment of the relative strengths and weaknesses of each dataset, as well as their overall performance.

Furthermore, to quantify the relative dynamic contributions of surface albedo, downward shortwave radiation, downward longwave radiation, sensible heat flux (SHF), latent heat flux (LHF), and ground heat flux to temperature biases, we applied the decomposed temperature metric developed by Juang et al. (2007). The surface energy balance equation is

$$\varepsilon\sigma LST^4 = (1 - \alpha_s)SW_{in} + LW_{in} - SH - LH - G, \quad (6)$$

where LST is the land surface temperature, ε is the effective emissivity, and σ is the Stefan-Boltzmann constant ($5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$). Following established methodologies for the decomposed temperature metric (Juang et al. 2007; Li et al. 2024), a constant value of 1.0 was assumed for ε . This common simplification facilitates a first-order approximation and enables direct comparison with prior studies, as obtaining

accurate, spatiotemporally continuous ε fields for both models and reanalysis remains challenging. Nevertheless, it is acknowledged that this assumption introduces a source of uncertainty by neglecting the natural variations of emissivity with land cover, soil moisture (SM), and snow cover. The potential implications of this simplification are further discussed below. The term α_s is the surface albedo, SW_{in} is the downward shortwave radiation, LW_{in} is the downward longwave radiation, SH is the sensible heat flux, LH is the latent heat flux, and G is the ground heat flux.

Next, we considered the temperature difference Δ between the reference dataset (MERRA2) and the WRF9km simulations. The rearranged Eq. (7) gives

$$\Delta LST = \frac{1}{4\sigma LST^3} [-SW_{in}\Delta\alpha_s + (1 - \alpha_s)\Delta SW_{in} + \Delta LW_{in} - \Delta LH - \Delta SH - \Delta G + \Delta I], \quad (7)$$

where LST , SW_{in} , and α_s are the reference mean climate state surface temperature, downward shortwave radiation, and surface albedo, respectively; ΔI represents the residual resulting from the high-order derivative, including the covariance between attributing variables and the variables that could not be quantified in the decomposition metric. As such, the contribution of energy budget changes to LST biases could be evaluated based on Eq. (7).

3. Results

a. Evaluation of WRF9km performance in depicting air temperature

We analyzed the spatial distributions of annual- and seasonal-mean temperatures over the TP during the period 1980–2019 based on TPMFD, WRF9km, ERA5, MERRA2, ERA5-Land, and CN05.1 data (Fig. S1). The WRF9km captured the observed spatial patterns of temperature, with a spatial correlation exceeding 0.8 across all seasons. However, the WRF9km exhibited cold biases in most regions, including the Qiangtang Plateau and Kunlun Mountains, whereas a warm bias was observed in the northern Qaidam Basin (Fig. 2). In terms of regional averages, the WRF9km exhibited cold biases relative to CN05.1 for the annual mean and all four seasons, being largest in March (-4.0°C) and smallest in August (-1.8°C). The ERA5 also effectively captured the spatial patterns of observed temperature but exhibited a notable warm bias across the north and east of the TP and a cold bias in the southeast, except in JJA. Compared with ERA5, WRF9km displayed higher temperatures across the southeastern TP, specifically in the Hengduan Mountains, and lower temperatures across the northwestern Kunlun Mountains (Fig. 2c), indicating that the WRF9km mitigated the cold bias in the southeastern region and warm bias in the northwestern region inherent in ERA5. Comparing the WRF9km with MERRA2, cold biases prevailed for the WRF9km across nearly the entire TP, except in the Qaidam Basin. We further compared both the WRF9km and ERA5 with in situ station observations. The results revealed that

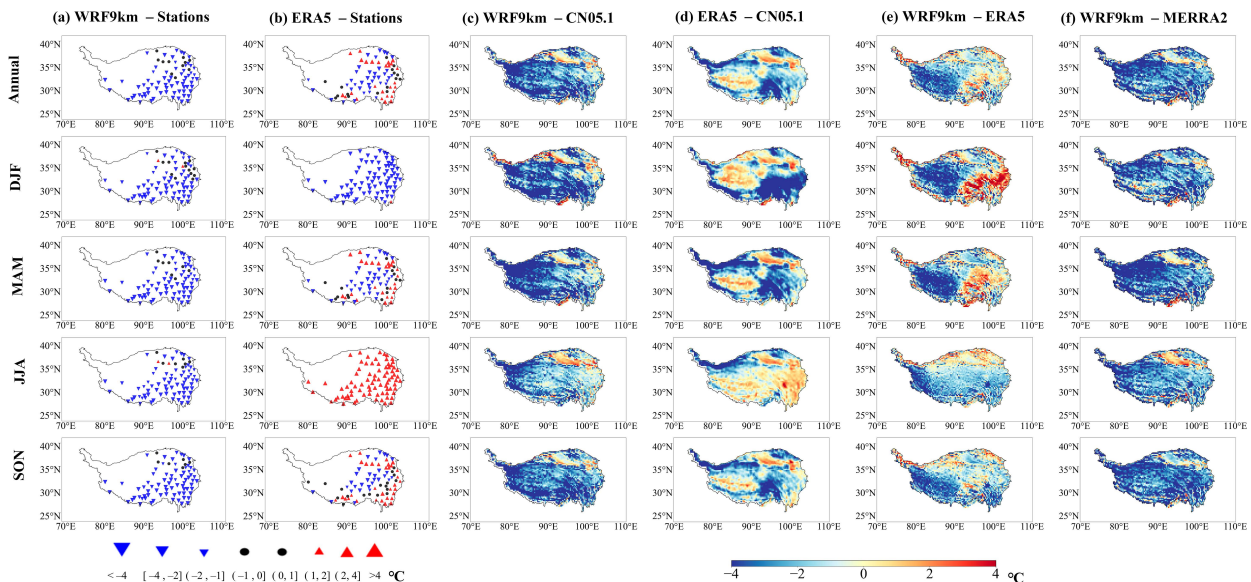


FIG. 2. Spatial distributions of air temperature obtained from (a) WRF9km–meteorological station data, (b) ERA5–meteorological station data, (c) WRF9km–CN05.1, (d) ERA5–CN05.1, (e) WRF9km–ERA5, and (f) WRF9km–MERRA2 for the period 1980–2019.

WRF9km displayed cold biases at most stations in all seasons, while ERA5 displayed warm biases at most stations in JJA.

In terms of the annual mean, the ERA5 data displayed a cold bias (Fig. 3b). The WRF9km simulation also exhibited a consistent cold bias throughout the year when compared with CN05.1 and MERRA2 data. However, the WRF9km displayed a larger cold bias throughout the year compared to ERA5, indicating that the WRF9km did not markedly reduce the regional-averaged cold bias inherent in its driving data. The application of a $0.65^{\circ}\text{C} (100 \text{ m})^{-1}$ temperature lapse rate during model-to-station interpolation substantially mitigated systematic cold biases associated with terrain discrepancies between modeled and observed locations, consistent with a previous study (Li et al. 2025). The evaluation against observational data from 81 meteorological stations revealed that although both the WRF9km and ERA5 exhibited reduced cold biases when using this elevation-correction approach, persistent cold biases remained (Fig. 3b).

b. Evaluation of WRF9km performance in depicting precipitation

The WRF9km effectively captured the spatial distributions of annual- and seasonal-mean precipitation across the TP (Fig. S2), with a small deviation relative to the CN05.1 observational dataset (Fig. 3c). Compared with CN05.1, ERA5 overestimated precipitation throughout the year, with the largest wet bias of 2.58 mm day^{-1} occurring in July. Although the WRF9km also overestimated precipitation year-round, its overestimation was lower than that of ERA5, with the largest bias of 0.60 mm day^{-1} in May and almost no deviation in August. This indicates that although both the ERA5 and WRF9km datasets exhibited a tendency to overestimate precipitation, the WRF9km model performed better, with especially notable improvements in simulating summer precipitation. The WRF9km reduced 79% of the wet bias inherent in ERA5 throughout the year.

We compared the spatial distributions of annual- and seasonal-mean precipitation over the TP during 1980–2019 among CN05.1, WRF9km, ERA5, and ERA5-Land (Fig. 4). Compared to CN05.1, the precipitation bias in WRF9km exhibited pronounced spatial heterogeneity rather than a uniform regional bias. Specifically, WRF9km tended to overestimate precipitation along the slope regions of the southern and eastern TP, while underestimating it over the interior TP, particularly the western TP, with the dry bias being most evident during JJA. As a result, the localized dry bias over the interior western TP partially offset the wet bias over the slope regions, leading to a relatively smaller bias in regional-mean summer precipitation. Comparison with station observations further indicated that WRF9km exhibited relatively small biases at most stations, whereas ERA5 showed a pronounced wet bias over the southeastern TP, especially during JJA. Overall, WRF9km represented a noticeable improvement over ERA5 by reducing excessive precipitation across much of the TP, particularly over the eastern and southeastern slope regions, where ERA5 substantially overestimated precipitation.

Figure 5 shows the spatial distributions of peak precipitation months based on the CN05.1, ERA5, ERA5-Land, and WRF9km datasets. Precipitation peaks on the TP generally exhibit a homogeneous summer dominance that aligns with Indian monsoon moisture intrusion through the Brahmaputra Valley (Yang et al. 2022). Across all four datasets, peak precipitation occurred later on the southwestern TP than in other regions. The month of peak precipitation was well captured by ERA5 and ERA5-Land, showing high consistency with CN05.1. While WRF9km effectively simulated the summer precipitation peak, it also revealed substantial spatial heterogeneity across the northwestern TP, with maximum precipitation occurring in MAM. The MAM peak simulated by WRF9km could be attributed to the convergence effect of topographic lifting, driven by

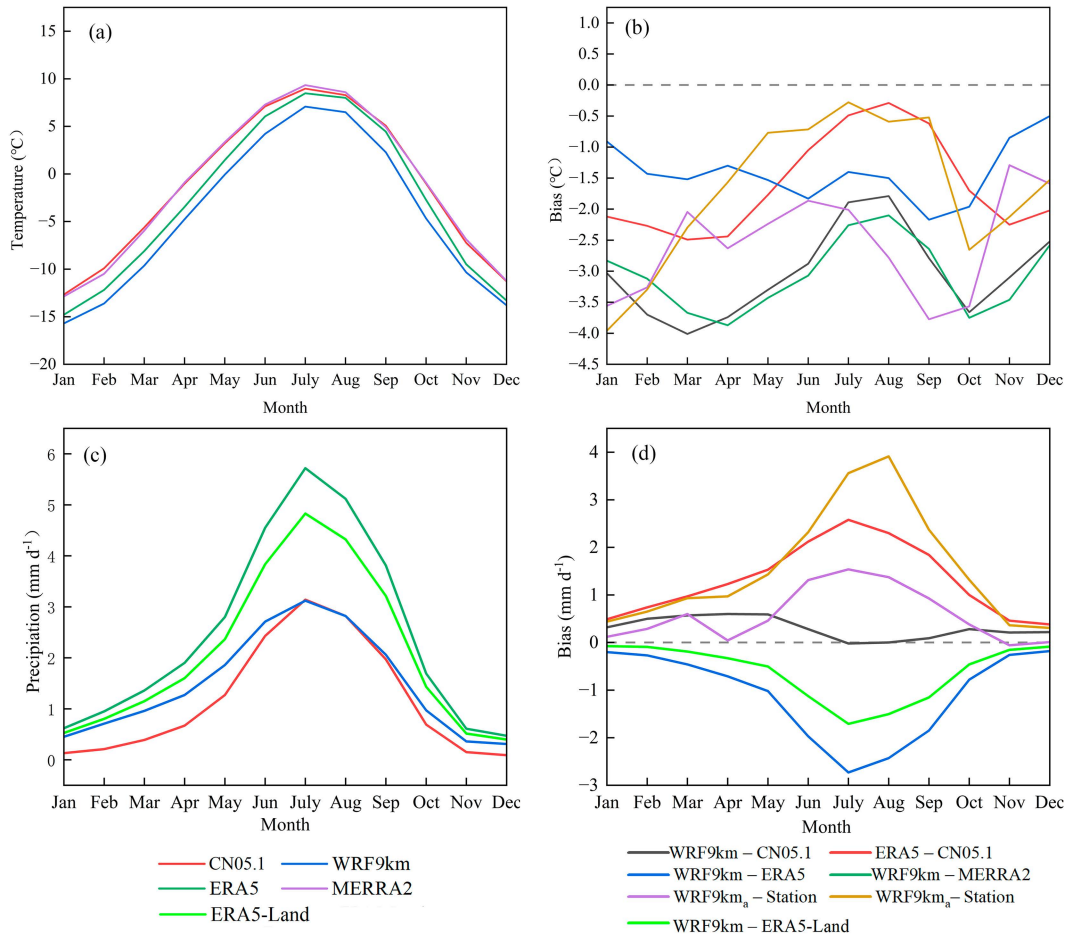


FIG. 3. Annual cycles of (a) regional average air temperature and (c) precipitation obtained from CN05.1, ERA5, MERRA2, ERA5-Land, and WRF9km. Differences in (b) air temperature and (d) precipitation between WRF9km and reference data for the period 1980–2019. WRF9km_a and ERA5_a represent the WRF9km and ERA5 results after lapse rate correction, respectively. The evaluation against station data is based on a point-to-grid comparison, using the model/reanalysis averages at the station locations.

the dominant westerly belt over the northwestern TP during DJF and MAM. However, this “bias” in WRF9km may not simply reflect an error but was more likely attributable to its high spatial resolution and specific physical parameterizations. The high resolution of WRF9km enabled it to resolve fine-scale topographic features. The southeastern TP is a region with extremely steep topographic gradients, where the high-resolution model can simulate stronger and more frequent topographically forced convection. This type of convection may be particularly pronounced in spring—when moisture availability increases but the large-scale summer monsoon circulation is not yet fully established—potentially causing the model to transition into the peak precipitation phase earlier than observed (MauSSION et al. 2014; Bieniek et al. 2016).

Previous studies suggested that the WRF Model tends to overestimate the intensity of daily heavy precipitation and simulates an early precipitation peak over the complex terrain of the TP. These biases may be related to the relatively coarse horizontal resolution used in the current convection-permitting

model (CPM), as well as observational uncertainties (Frei and Schär 1998; Isotta et al. 2014). Furthermore, microphysics parameterization schemes may not be fully optimized for CPMs. Although some microphysics schemes can achieve more realistic hydrometeor distributions, they often involve substantial tuning, and many key processes—such as droplet–droplet interactions, ice particle formation, and particle size distribution shapes—remain poorly understood. Fully realizing the potential of these schemes will require further in-depth research to improve the numerical representation of these processes (Prein et al. 2015).

c. Spatial distribution of trend in air temperature

We compared the spatial distributions of annual- and seasonal-mean temperature trends over the TP during the period 1980–2019 in CN05.1, WRF9km, ERA5, and MERRA2 data (Fig. 6). The CN05.1 dataset exhibited significant warming trends in annual- and seasonal-mean temperatures across the entire TP, with a stronger warming trend in the northern

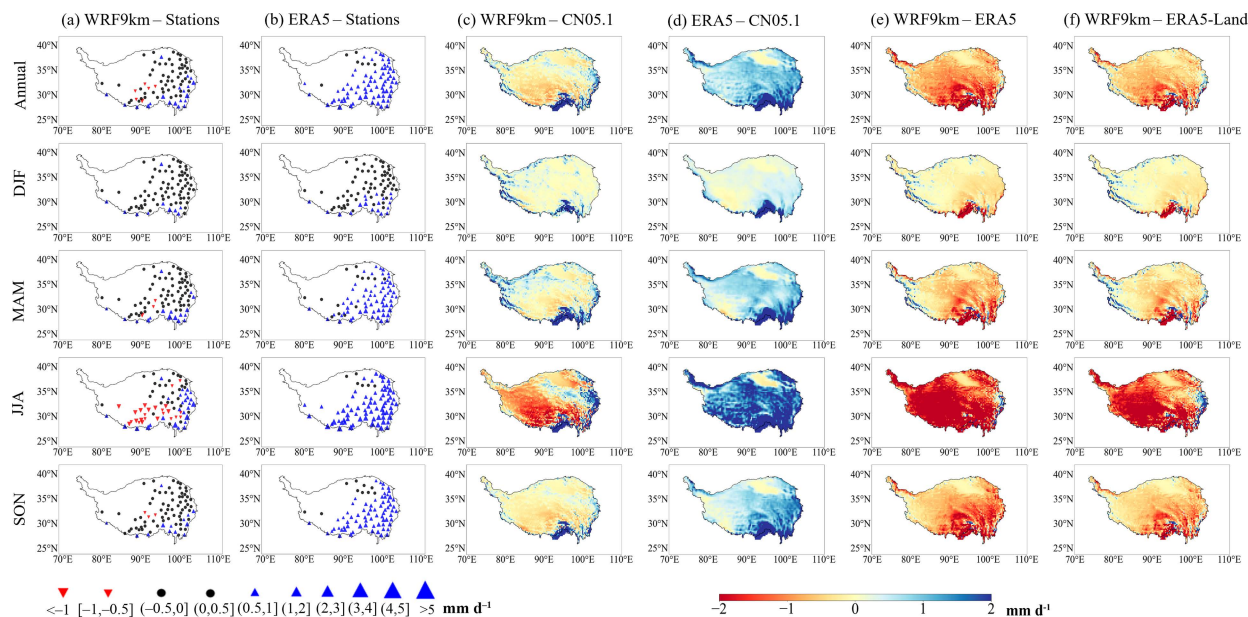


FIG. 4. Spatial distributions of precipitation obtained from (a) WRF9km–meteorological station data, (b) ERA5–meteorological station data, (c) WRF9km–CN05.1, (d) ERA5–CN05.1, (e) WRF9km–ERA5, and (f) WRF9km–ERA5-Land for the period 1980–2019.

region compared with that in the southern region. The WRF9km broadly reproduced the warming pattern observed in CN05.1, maintaining good agreement in terms of spatial distribution patterns and statistical significance across the TP. However, there were seasonal and regional discrepancies. Specifically, in DJF, the WRF9km underestimated the warming amplitude over the central TP. On the other hand, the WRF9km spatial pattern and warming trend in JJA aligned closely with that of CN05.1. The WRF9km and ERA5 were highly consistent in their temperature change signals, reflecting the inheritance by

WRF9km of large-scale trends from ERA5. Nonetheless, the WRF9km reduced the overwarming trend of ERA5, especially across the southern TP in DJF and the northeastern TP in MAM, yielding simulations that were closer to CN05.1 data. Overall, the WRF9km simulations offered a more balanced and regionally consistent picture of temperature trends. In contrast, the MERRA2 spatial pattern markedly diverged from those of other datasets, particularly failing to capture warming trends on the western TP. Moreover, MERRA2 struggled to accurately simulate spatial temperature variation trends because

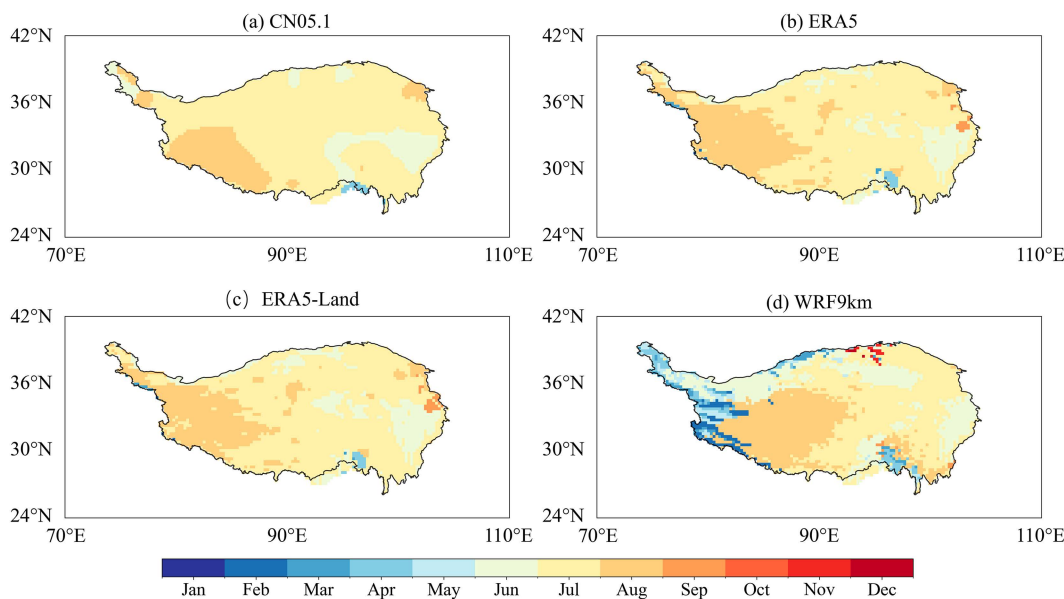


FIG. 5. Spatial distributions of peak precipitation months in (a) CN05.1, (b) ERA5, (c) ERA5-Land, and (d) WRF9km for the period 1980–2019.

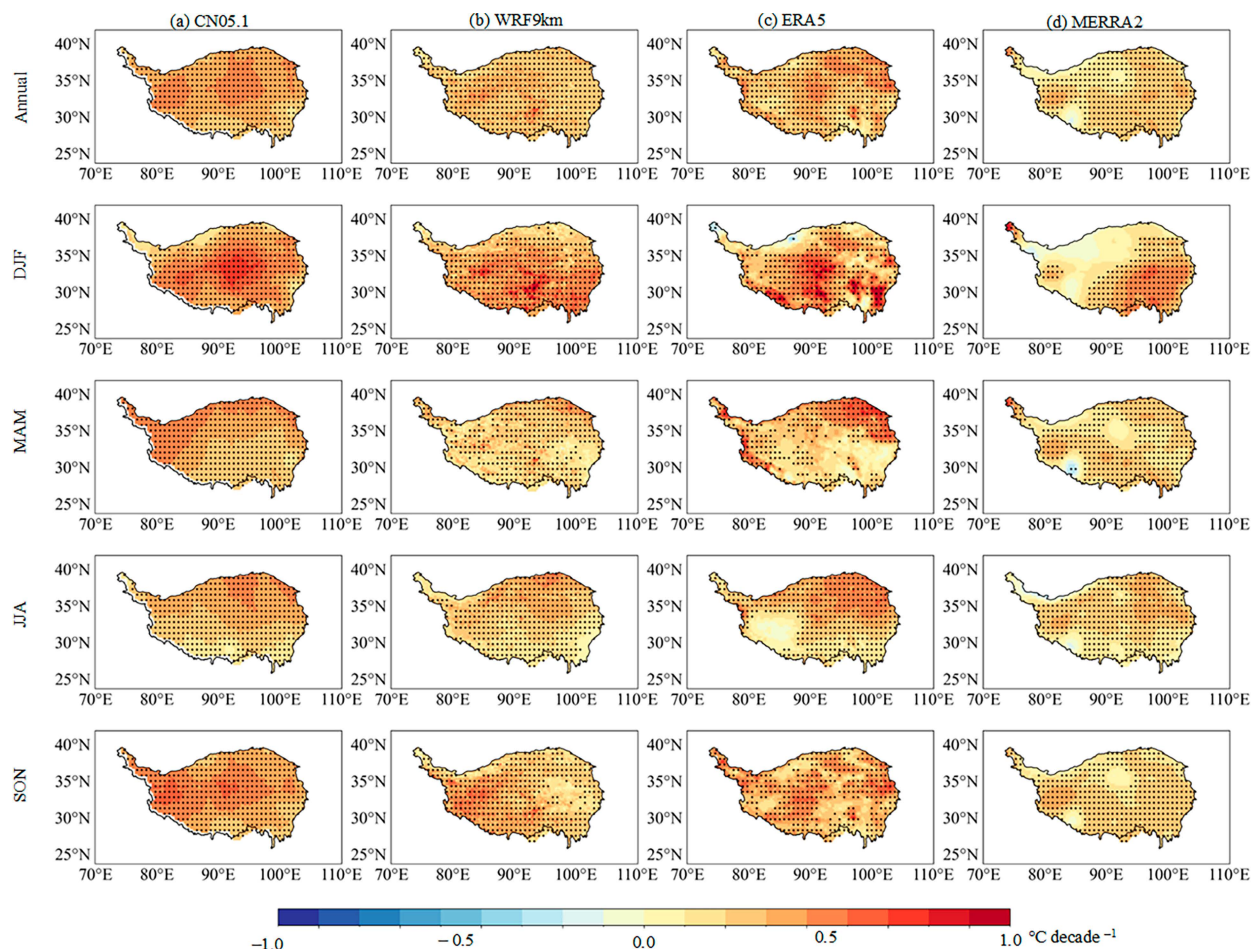


FIG. 6. Spatial distributions of linear trends in annual- and seasonal-mean air temperatures over the TP from (a) CN05.1, (b) WRF9km, (c) ERA5, and (d) MERRA2 for the period 1980–2019. Black dots indicate trends exceeding the 95% confidence level.

its horizontal resolution is insufficient to resolve small-scale temperature differences caused by the complex terrain of the TP (Wang et al. 2018). In general, the WRF9km simulated the spatial variation trend of temperature on the TP, exhibiting high consistency with CN05.1.

d. Spatial distribution of trend in precipitation

A comparison of the spatial distributions of annual and seasonal precipitation trends over the TP for the period 1980–2019 in CN05.1, WRF9km, ERA5, and ERA5-Land is shown in Fig. 7. On the basis of CN05.1 data, annual precipitation changes over the TP exhibited marked spatial heterogeneity. The westerlies-dominated region, i.e., western and northern TP, exhibited a trend of significantly increasing precipitation, while the monsoon-dominated region, i.e., southeastern TP, displayed a decreasing trend, and the pattern was particularly pronounced in JJA. Other seasons exhibited diverging trends. Specifically, the overall precipitation trend in DJF was weak, MAM showed a slight drying in the west and significant wetting in the southeast, and SON exhibited localized wetting in the northeast but widespread drying elsewhere. Although both ERA5 and

ERA5-Land largely captured the broad spatial patterns of CN05.1, they exhibited systematic biases: a consistent overestimation of drying intensity in the monsoon-dominated region across all seasons and an exaggerated wetting trend in the westerlies-dominated region during JJA. The general spatial patterns of the WRF9km simulation were inherited from ERA5, but with marked improvements in trend magnitudes in key regions, particularly a reduction in overestimated summer wetting in the westerlies-dominated region and drying trends in the monsoon-dominated region, ultimately achieving precipitation change patterns more consistent with those of CN05.1.

e. Quantitative evaluation of WRF9km performance in temperature and precipitation simulations

To evaluate the overall performance of the WRF9km over the TP against released reference datasets, we conducted systematic comparisons with ERA5, ERA5-Land, MERRA2, and TPMFD through comparative analysis of individual indicator scores and the composite assessment scores (Fig. 8). Figure 8a shows the detailed metric-level performance of different datasets in depicting air temperature. MERRA2 exhibited consistently

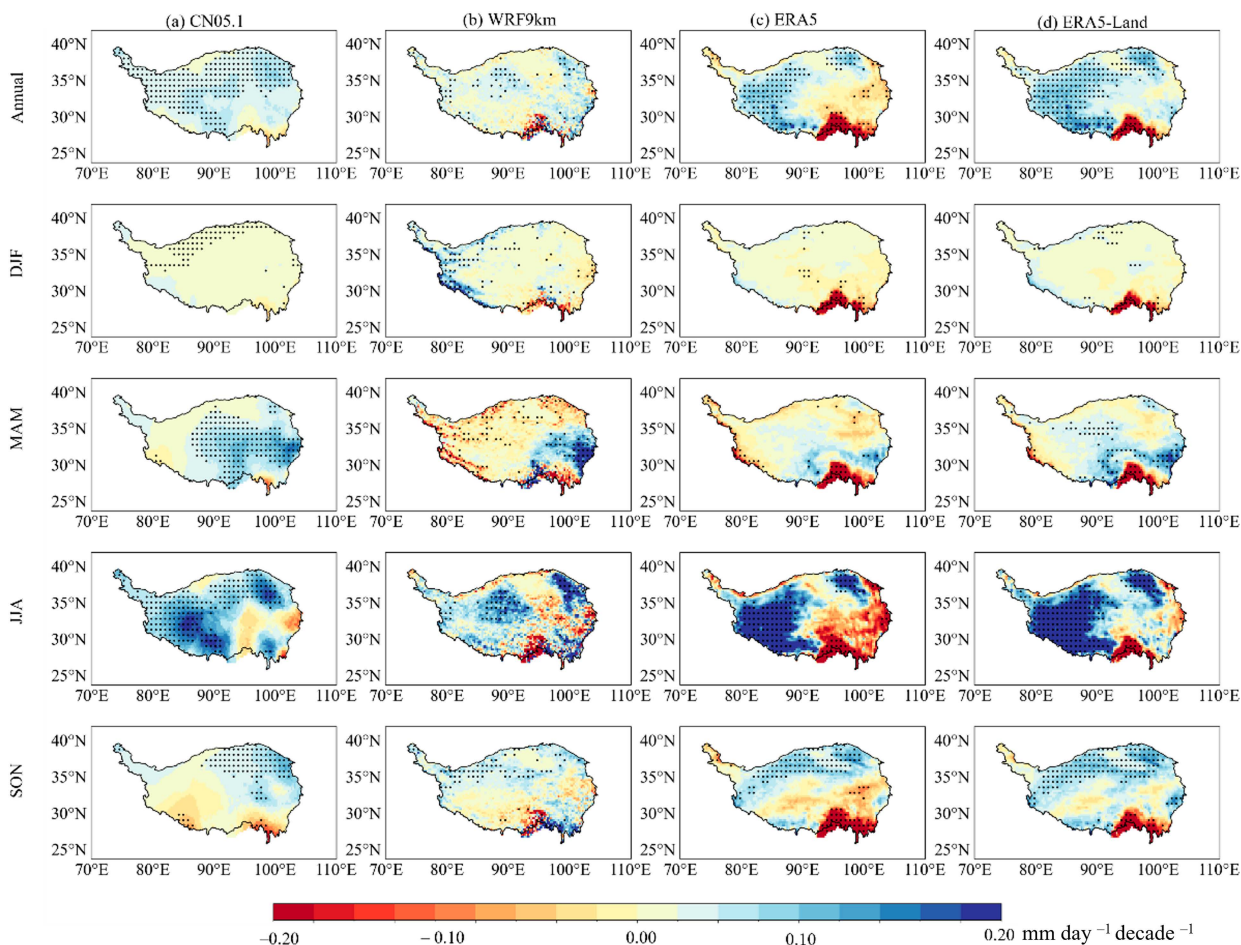


FIG. 7. Spatial distributions of linear trends in annual- and seasonal-mean precipitation over the TP in (a) CN05.1, (b) WRF9km, (c) ERA5, and (d) ERA5-Land for the period 1980–2019. Black dots indicate trends exceeding the 95% confidence level.

outstanding performance across all evaluation indicators, faithfully capturing the temperature characteristics recorded in the CN05.1 observational data. It achieved high scores in bias, temporal variation, and spatial distribution, thus making it well suited for subsequent temperature bias analysis. The WRF9km demonstrated comparatively robust performance in terms of spatial and temporal correlations, while its error magnitude metrics attained lower scores than the other datasets. As shown in Fig. 8b, the comprehensive evaluation scores ranked as follows: MERRA2 was the highest, followed by TPMFD, ERA5, WRF9km, and ERA5-Land. This ranking indicates that WRF9km did not yield a marked improvement in overall temperature representation relative to ERA5, with the exception of its enhanced performance in capturing spatial trends (Fig. 6).

Figures 8c and 8d depict the comprehensive evaluation of performance in precipitation across the five datasets, revealing that TPMFD obtained the highest overall performance, with a total score of 24. In notable contrast to the temperature evaluation, the WRF9km outperformed ERA5 in representing precipitation, particularly in the error magnitude. The advantage of the WRF9km over ERA5 in capturing the spatial trends of

precipitation can be observed in Fig. 7. These were key factors driving its higher composite score, again indicating pronounced improvement in precipitation simulation relative to its driving data. However, improvements in WRF9km simulations of temporal changes and spatial distribution patterns remain necessary.

f. Potential causes of air temperature biases

The WRF9km simulations exhibited certain temperature biases across the TP. To investigate the possible causes of these biases, we compared relevant variables, such as energy fluxes, surface albedo, snow cover, and cloud cover, in WRF9km and MERRA2/MODIS (Figs. 9–13). The WRF9km underestimated both downward longwave radiation and net SR (NSR) across most of the TP. The spatial differences in downward radiation forcing, primarily driven by net shortwave and downward longwave radiation, closely resembled those in temperature (Fig. 2). Therefore, these underestimations of downward longwave radiation and net downward shortwave radiation were inferred to be important factors in the detected cold bias. Ground heat flux directly affects surface temperature fluctuations by controlling the transfer, storage, and release of surface energy. Compared to MERRA2, the

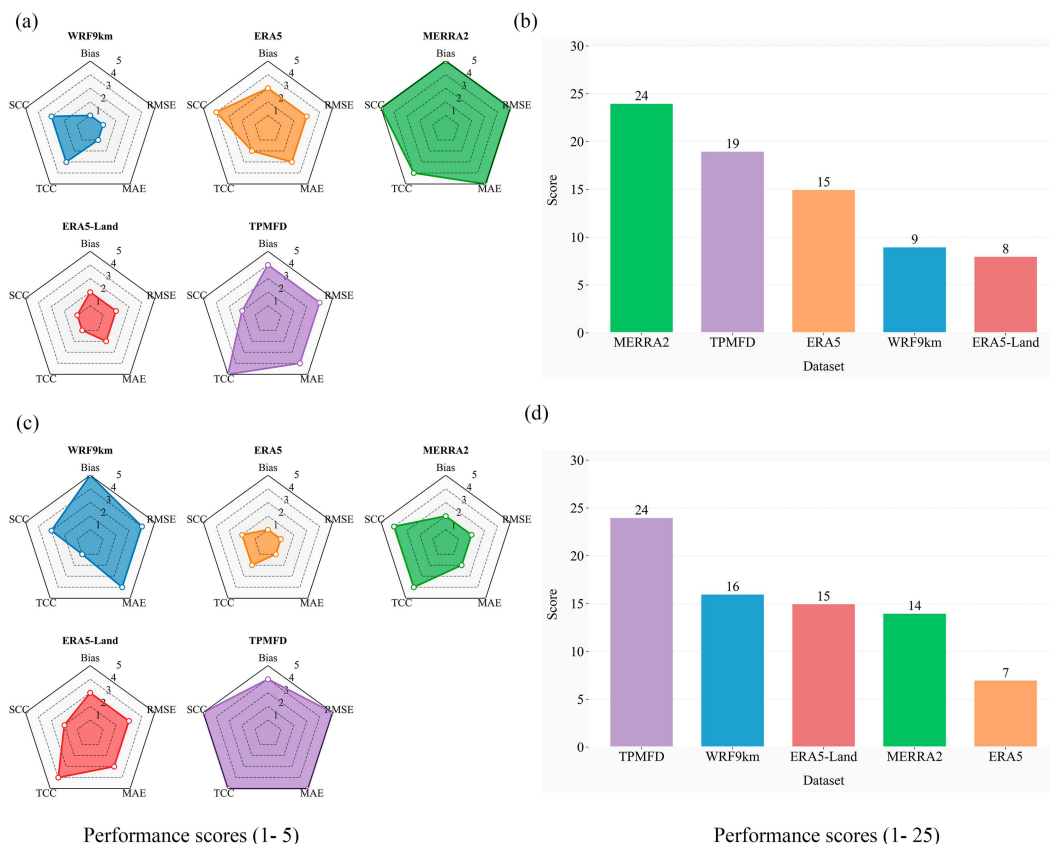


FIG. 8. Comprehensive evaluation scores of the WRF9km, ERA5, MERRA2, ERA5-Land, and TPMFD datasets in relation to (a),(b) air temperature and (c),(d) precipitation on the TP for the period 1980–2019 according to CN05.1 observational data. See [section 2c](#) for definitions of evaluation metrics.

WRF9km substantially underestimated ground heat flux, partially contributing to the cold bias.

To diagnose the cause of the shortwave radiation bias, surface albedo from WRF9km was compared against MERRA2 and MODIS over the TP (Fig. 10). The analysis revealed a widespread albedo overestimation in WRF9km, which directly explained its simulated shortwave radiation deficit. This elevated albedo was likely attributable to an excessive snow-cover fraction in the model, as WRF9km exhibited greater snow cover than both MERRA2 and MODIS across the region (Fig. 11). Furthermore, clouds played a critical role in the surface radiation budget by absorbing and re-emitting longwave radiation. The WRF9km simulation produced less high-cloud cover than MERRA2 (Fig. 12), leading to a consequent reduction in downward longwave radiation.

Figure 13 shows the quantitative contributions of various factors to the detected temperature biases. Surface albedo had the highest contribution to the cold bias, accounting for 64% of the total. Ground heat flux ranked second followed by sensible heat and latent heat fluxes. However, the warm bias induced by downward shortwave radiation may stem from the misrepresentation of cloud–radiation processes over the TP in the model; the WRF9km simulated reduced high-cloud cover, which allowed

excessive shortwave radiation to reach the land surface, leading directly to the warm bias.

In addition, some model bias may be related to the distribution of meteorological stations, especially in high mountain areas. Meteorological stations are typically located in valleys at lower altitudes, where temperatures tend to be higher (Wang et al. 2016). Therefore, the high-altitude temperatures interpolated from these stations often include high temperature values, resulting in a cold bias in the simulated values.

g. Potential causes of precipitation biases

The WRF9km simulations indicated overall wet precipitation biases over the TP across both annual and seasonal scales compared to the CN05.1 observational dataset. However, these wet biases were significantly smaller compared to both ERA5 and ERA5-Land. To investigate the causes of these biases, we conducted a systematic comparative analysis focusing on both near-surface land–atmosphere interactions and midtropospheric moisture transport processes.

At the near-surface layer, we compared the annual- and seasonal-mean wind speed at 10-m height, specific humidity, and Bowen ratio between WRF9km and ERA5-Land. Compared with ERA5, WRF9km exhibited relatively higher 10-m wind speeds, which were accompanied by enhanced

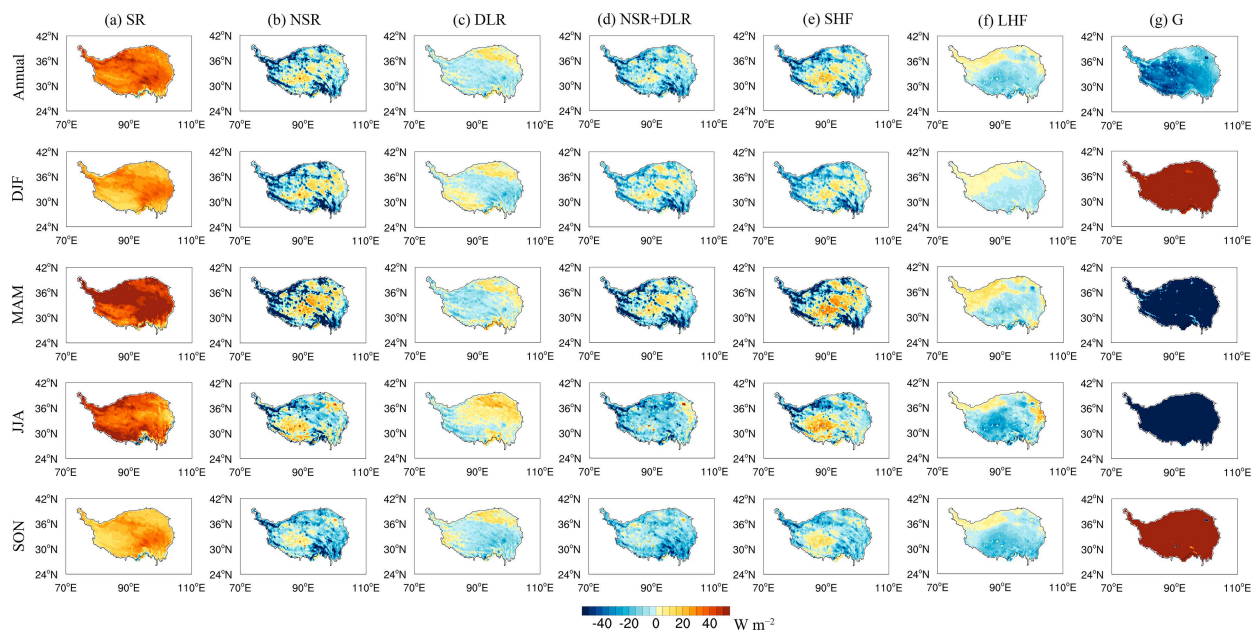


FIG. 9. Differences between WRF9km and MERRA2 ($\text{WRF9km} - \text{MERRA2}$) across the TP in (a) SR, (b) NSR, (c) DLR, (d) net radiation forcing (NSR + DLR), (e) SHF, (f) LHF, and (g) ground heat flux (G) for the period 1980–2019.

surface turbulent exchange and sensible heat flux, leading to a larger Bowen ratio and reduced near-surface specific humidity (Fig. 14). These near-surface thermodynamic differences were closely associated with reduced atmospheric moisture availability and suppressed cloud formation, although they may also partly reflect the consequences of reduced precipitation. Additional diagnostics further indicated that WRF9km simulated consistently lower soil moisture as well as reduced sensible heat flux and latent heat flux compared to ERA5 and ERA5-Land (Figs. 16a–c). The reduced latent heat flux directly weakened the supply of moisture and latent heat to the atmosphere, while the concurrent reductions in sensible and latent heat fluxes promoted near-surface cooling and enhanced atmospheric stability, thereby inhibiting convective initiation and vertical moisture transport.

Thermodynamically, as noted above, the overestimated surface albedo in WRF9km reduced surface absorption of solar radiation. This resulting near-surface cooling increased atmospheric stability, which, in turn, suppressed turbulent mixing and limited the vertical transport of moisture to cloud-forming levels aloft (Zhang et al. 2018; Santanello et al. 2018). Together, the enhanced near-surface drying and constrained vertical moisture transport provided a physically consistent explanation for the reduced precipitation simulated by WRF9km relative to the reanalyses.

In the midtroposphere, we analyzed the annual- and seasonal-mean wind speed, specific humidity, and integrated water vapor transport at the 500-hPa level between WRF9km and ERA5 (Fig. 15). While both datasets agreed on the dominance of westerly moisture transport outside the summer monsoon season, notable discrepancies emerged in certain key regions. In JJA, the summer monsoon brings warm and moist air from the Indian Ocean into the southern TP, causing a weakening of the westerly transport. However, a pronounced

discrepancy appeared over the southwestern TP, where the WRF9km simulation produced weaker southerly and southeasterly winds than ERA5. This resulted in reduced northward moisture transport and contributed to a distinct precipitation reduction in the model for this region. Conversely, over the Hengduan Mountains, WRF9km simulated stronger moisture convergence, thereby enhancing precipitation in the southeastern TP. During SON, as the monsoon retreats, westerly moisture transport strengthens and reestablishes dominance. The WRF9km simulation accurately captured this seasonal transition in the southern TP, shifting from a summertime southeasterly regime to a westerly one. In contrast, ERA5 maintained an anomalous southeasterly wind pattern over the southeastern TP. This systematic deviation in the ERA5 circulation could explain its pronounced wet bias relative to the WRF9km simulation during this season.

The circulation differences described above can be attributed to the combined effects of surface thermal forcing, boundary layer processes, and complex topography over the TP. Reduced sensible and latent heat fluxes in WRF9km weakened the spatial gradients of surface heating, thereby altering the intensity and pattern of thermal forcing that drive regional circulation (Avissar and Schmidt 1998; Wang et al. 2022). Through boundary layer mixing and vertical heat redistribution (Liu et al. 2018), these surface anomalies modulated the lower-tropospheric thermal structure and subsequently influenced large-scale atmospheric circulation. In regions of complex terrain, interactions between high-resolution topography and boundary layer parameterizations further reshaped local circulations (Xu et al. 2018), affecting the efficiency of orographic lifting and the location and strength of moisture convergence.

In addition, the precipitation differences between WRF9km and ERA5 reflected the combined effects of physical

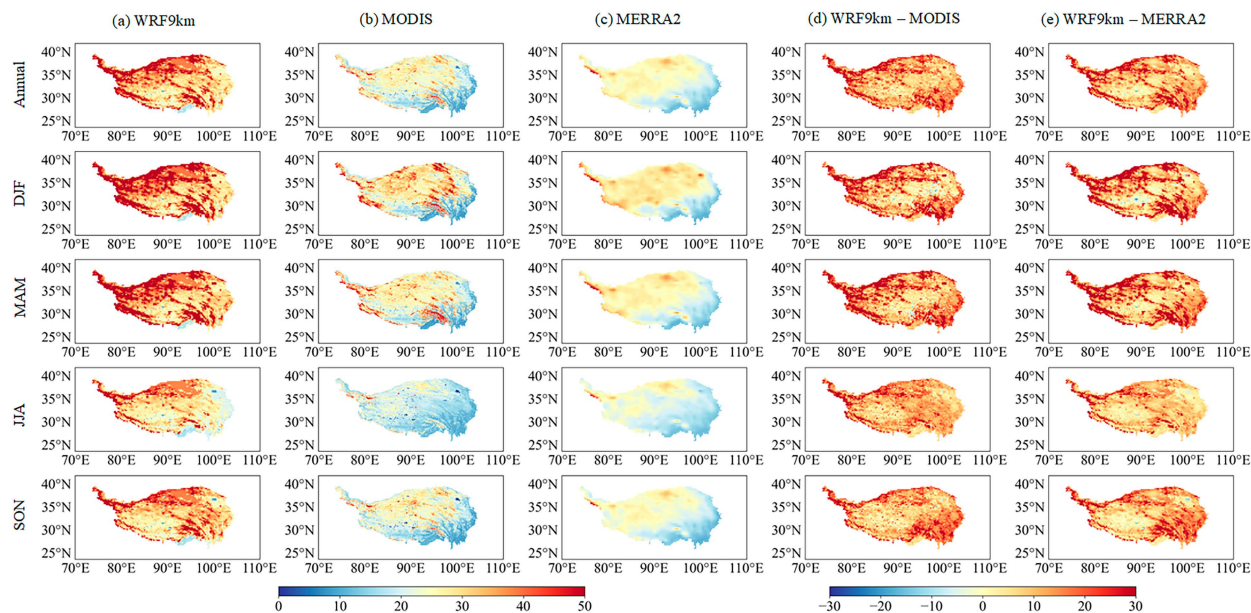


FIG. 10. Spatial distributions of annual- and seasonal-mean albedo obtained from (a) WRF9km, (b) MODIS, and (c) MERRA2, as well as their differences from (d) WRF9km – MODIS and (e) WRF9km – MERRA2 for the period 1980–2019.

parameterizations and reanalysis uncertainties. Inconsistencies in the representation of cumulus convection and boundary layer processes altered latent heat release and the associated atmospheric heat source, which plays a critical role in shaping circulation over the TP (Dee et al. 2011; Zhang et al. 2023). Moreover, inherent uncertainties in ERA5, particularly in data-sparse regions such as the TP, could be further amplified during regional dynamical downscaling (Li et al. 2024). Overall, the reduced precipitation bias in WRF9km relative to ERA5 and

ERA5-Land can be attributed to a coherent mechanism initiated by a drier and colder land surface, leading to weakened surface moisture and energy fluxes, enhanced atmospheric stability, altered thermal forcing, and ultimately suppressed moisture convergence and uplift.

In addition to mean-state precipitation biases, the differences in precipitation trends between WRF9km and the reanalyses were also linked to long-term changes in land–atmosphere coupling. The spatial trends in soil moisture,

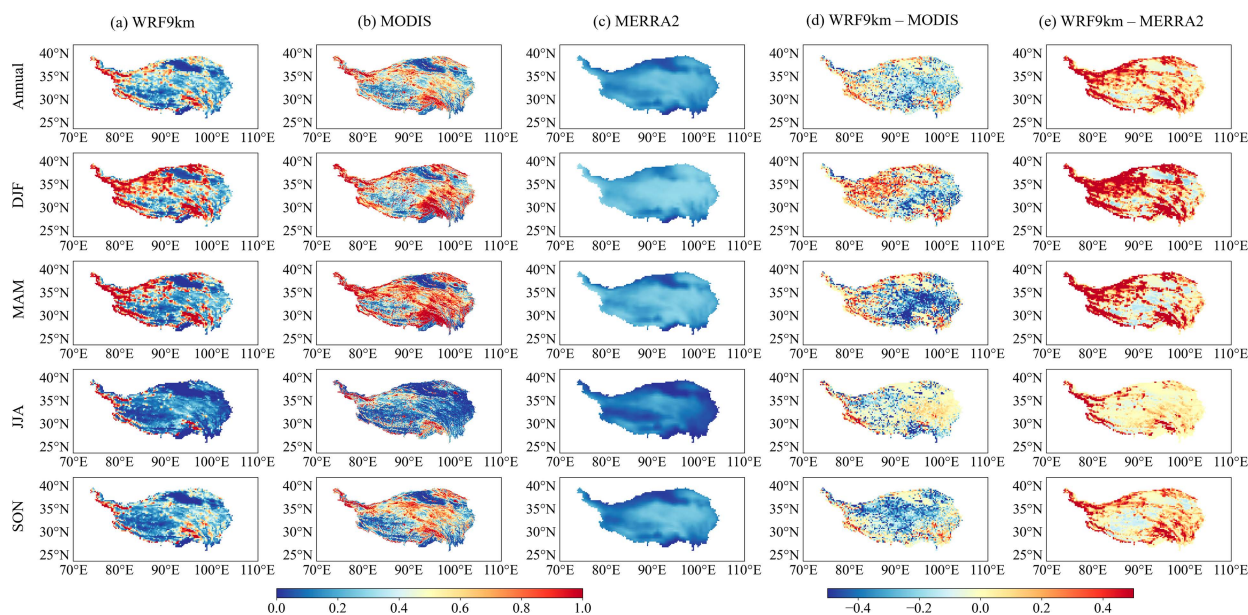


FIG. 11. Spatial distributions of annual- and seasonal-mean snow cover obtained from (a) WRF9km, (b) MODIS, and (c) MERRA2, as well as their differences from (d) WRF9km – MODIS and (e) WRF9km – MERRA2 for the period 1980–2019.

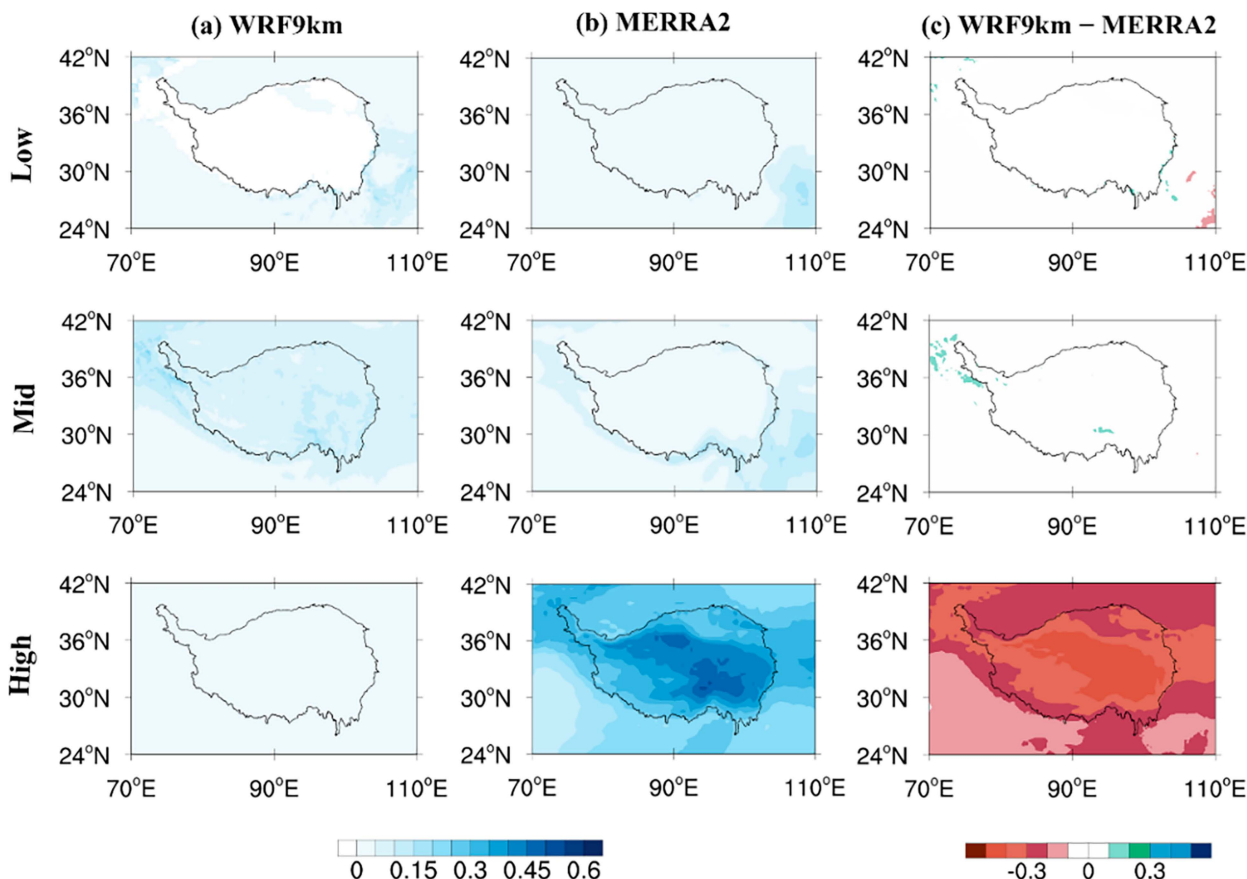


FIG. 12. Spatial distributions of the depth of cloud at different heights obtained from (a) WRF9km and (b) MERRA2, as well as (c) their differences (WRF9km – MERRA2) for the period 1980–2019.

sensible heat flux, and latent heat flux showed pronounced regional contrasts over the western TP, where WRF9km exhibited a markedly weaker wetting trend in soil moisture, a stronger increase in sensible heat flux, and a weaker

rise in latent heat flux compared to ERA5 and ERA5-Land (Figs. 16d–f).

The weaker soil moisture trend in WRF9km limited the long-term increase in surface moisture availability, thereby constraining the increase in latent heat flux. As a result, more available surface energy was partitioned into sensible heat flux, strengthening its increasing trend. This enhanced sensible heating promoted a warmer and drier boundary layer, suppressing moistening of the lower troposphere—a key process for sustaining precipitation increases. In contrast, the stronger soil moisture and latent heat flux trends in ERA5 and ERA5-Land favored more efficient moisture and latent heat transfer to the atmosphere, supporting a more wetting trend in precipitation.

From a dynamical perspective, divergent trends in surface energy partitioning also imply different evolutions of surface thermal forcing. In WRF9km, the intensified sensible heating altered the long-term thermal contrast between the western TP and surrounding regions, which could weaken the trend in moisture convergence and orographic uplift. This limited the increase in precipitation over time, particularly in the western TP where precipitation is highly sensitive to land–atmosphere feedbacks. Conversely, the reanalyses, with stronger latent heat flux and soil moisture trends, maintained more favorable conditions for moisture convergence and uplift, yielding a stronger wetting trend.

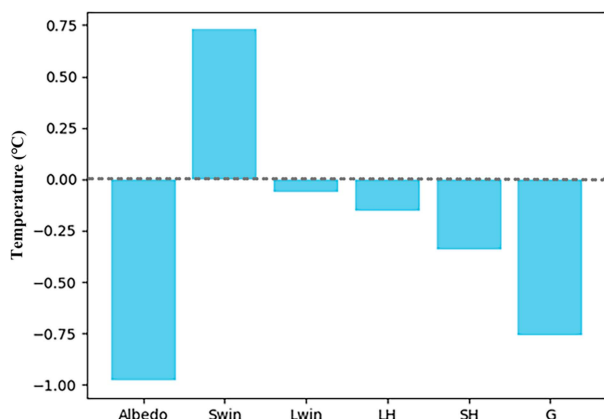


FIG. 13. Quantitative contributions of different factors to air temperature biases during the period 1980–2019. Note: SW_{in} = downward shortwave radiation; LW_{in} = downward longwave radiation; G = ground heat flux.

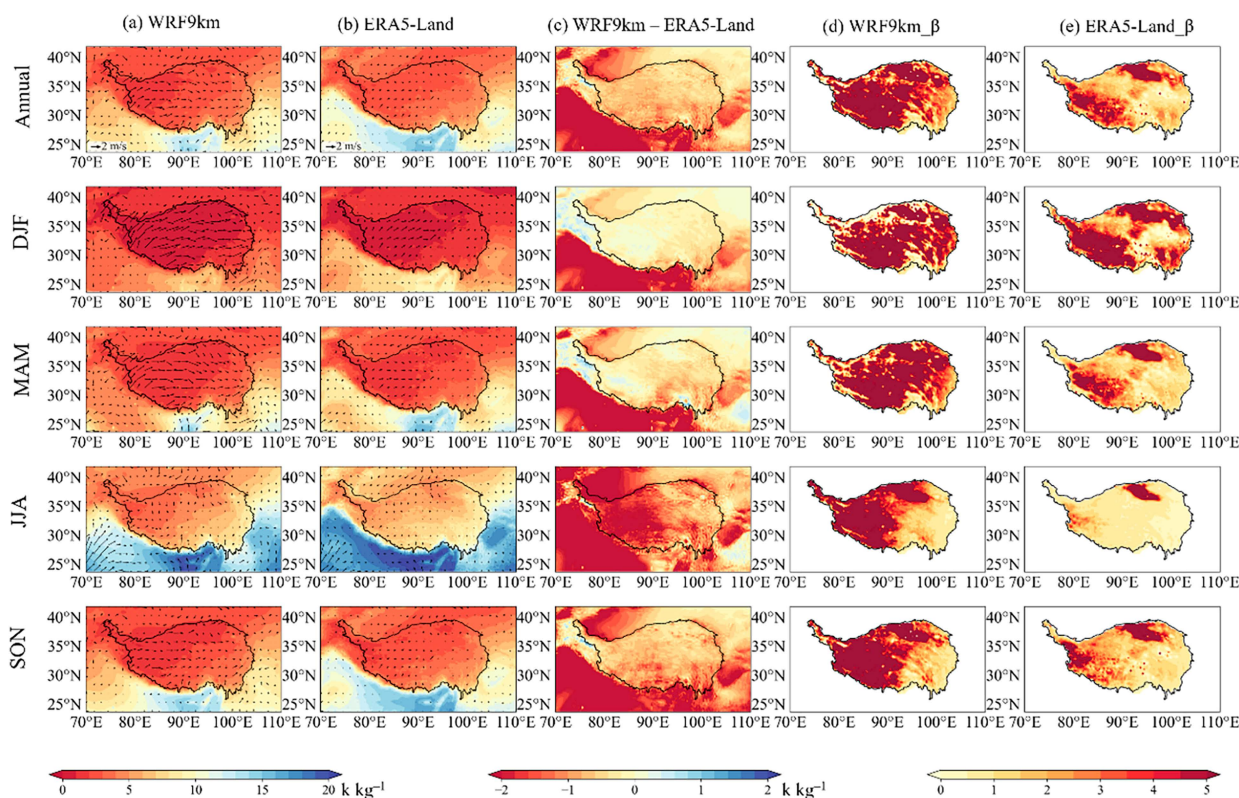


FIG. 14. Spatial distributions of annual- and seasonal-mean water vapor transport [vectors; $\text{kg} (\text{m s}^{-1})^{-1}$] and specific humidity (g kg^{-1}) at 10 m obtained from (a) WRF9km and (b) ERA5-Land, and Bowen ratio β from (d) WRF9km and (e) ERA5-Land, as well as (c) the differences in specific humidity (g kg^{-1}) between WRF9km and ERA5-Land during the period 1980–2019.

Overall, the weaker precipitation wetting trend in WRF9km over the western TP can be attributed to a self-consistent land–atmosphere feedback mechanism, characterized by muted soil moisture increase, reduced latent heat flux trend, and enhanced sensible heating trend. These processes collectively suppressed the long-term intensification of atmospheric moistening and moisture convergence, providing a physically consistent explanation for the precipitation trend differences between WRF9km and ERA5/ERA5-Land.

4. Discussion

Our study reveals that the WRF9km exhibits cold and wet biases in simulating temperature and precipitation across the TP; these are closely related to the model resolution, physical processes, external moisture transport simulations, and treatment of topography. Although the WRF9km performs well in simulating temperature and precipitation over the TP relative to commonly used datasets, notable differences persist in magnitude, trend patterns, and spatial variability relative to observations, highlighting the need for substantial improvements in the simulation of both variables.

The quantitative evaluation unequivocally demonstrates the superior performance of the WRF9km configuration in simulating JJA precipitation over the TP compared to ERA5. A pivotal finding of this study is the substantial reduction (79%)

in the domain-averaged wet bias compared to the driving ERA5 reanalysis data. Jiang et al. (2024) evaluated the daily average precipitation over the TP using kilometer-scale simulations and found that such high-resolution simulations effectively reduced the summer wet bias but even led to a slight underestimation. Ding et al. (2023) conducted simulations for East Asia with the Icosahedral Nonhydrostatic (ICON)-CPM model, comparing CPM against traditional parameterized convection across various resolutions. Their results confirmed CPM's superior performance in simulating precipitation characteristics, with further improvements in precipitation rate, spatial extent, and flux metrics when the grid spacing was refined to approximately 4 km. Similarly, Li et al. (2021) conducted a TP-wide CPM experiment at a grid spacing of 4 km and contrasted its summer performance with two coarser mesoscale models (13 and 35 km). The CPM eliminated spurious afternoon rainfall, thereby significantly reducing the wet bias present in the coarser simulations. The better-resolved orography in high-resolution models is a key factor in improving precipitation simulation over the complex terrain of the TP. In summary, our results underscore the efficacy of CPM in significantly reducing precipitation wet bias over the TP. The specific quantitative improvements and mechanistic analysis provided herein pave the way for more refined and accurate precipitation simulations.

MERRA2 performs poorly in terms of temperature and precipitation trend simulations; this may be related to the

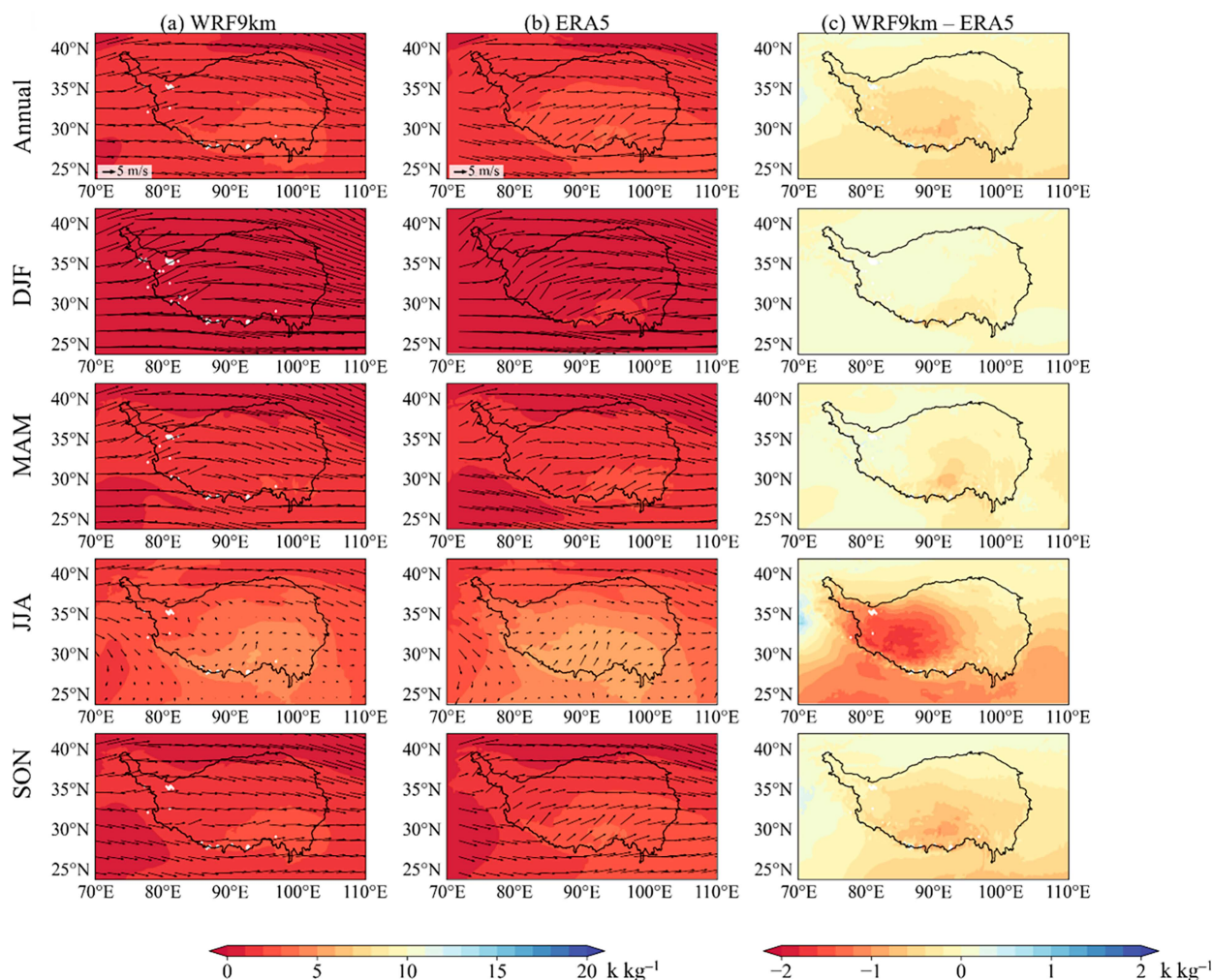


FIG. 15. Spatial distributions of annual- and seasonal-mean water vapor transport [vectors; $\text{kg (m s}^{-1})^{-1}$] and specific humidity (g kg^{-1}) at 500 hPa obtained from (a) WRF9km and (b) ERA5, and (c) the differences in specific humidity (g kg^{-1}) between WRF9km and ERA5 during the period 1980–2019.

data generation process. The data assimilation system and land surface model employed in MERRA2 oversimplify critical factors such as land cover, local circulation, and aerosol–cloud interactions; this leads to systematic biases in simulating the spatial heterogeneity of energy budgets, ultimately undermining the reliability of regional temperature and precipitation gradients and variation trends (Wen et al. 2022).

While CPM at kilometer-scale resolutions have been demonstrated to significantly improve summer precipitation simulations over the TP—effectively reducing wet bias by better representing moisture transport and uplift processes (Jiang et al. 2024; Li et al. 2021) and even outperforming traditional parameterized convection in multiple aspects (Ou et al. 2023)—our WRF9km simulations reveal a more complex reality: Improved precipitation performance does not translate to better snow simulation as expected. The primary reason lies in a strong counteracting mechanism driven by a pronounced cold bias originating from excessively high surface albedo. This overestimated albedo significantly limits shortwave radiation absorption

at the surface, thereby strongly suppressing snow ablation. As a result, although improved precipitation reduces wet bias, it simultaneously diminishes snow accumulation. These two opposing effects ultimately offset each other, resulting in no net improvement in simulated snow depth and snow cover. This indicates that the main shortcoming of the current model lies in the snow–ice–albedo feedback mechanism and that optimizing the snow albedo parameterization scheme will be essential for achieving future breakthroughs in snow simulation (Liu et al. 2022).

In the analysis of the month of peak precipitation, peak precipitation occurred later on the southwestern TP than in other regions for four datasets. This may be because during June–July, the South Asian monsoon makes landfall over the Indian Peninsula; thus, moisture can be transported to the southeastern TP, leading to strong moisture flux convergence and a precipitation peak across this region. However, owing to the topographic obstruction of the Himalayas, the South Asian monsoon must accumulate kinetic energy to rise over the high

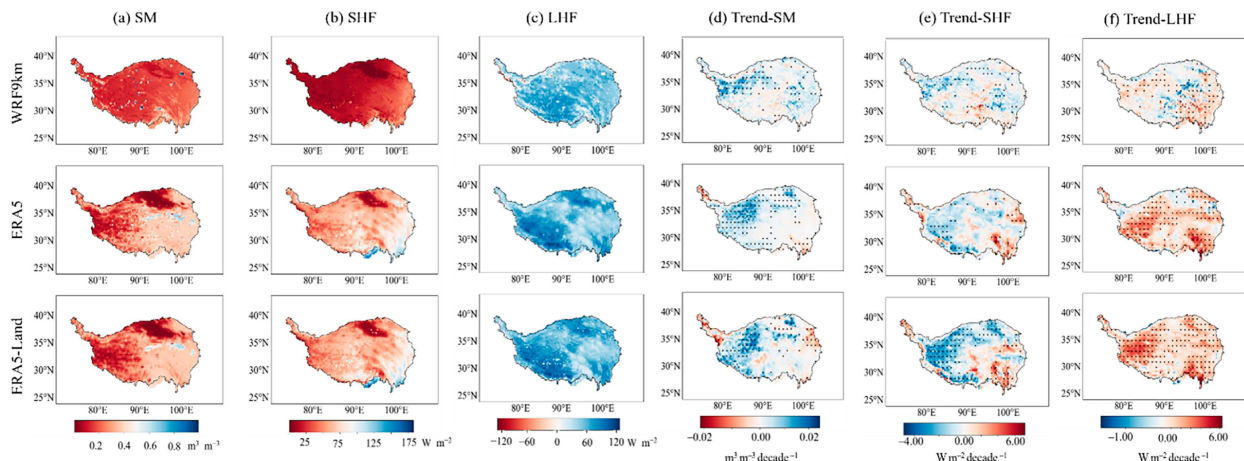


FIG. 16. (a)–(c) Spatial distributions and (d)–(f) linear trends of SM, SHF, and LHF over the TP for the period 1980–2019 from WRF9km, ERA5, and ERA5-Land. Black dots indicate that trends exceeded the 95% confidence level.

mountains. Only in August, when the monsoon circulation attains peak intensity, can moist air penetrate the orographic barrier and form strong moisture convergence across the leeward southwestern TP region. WRF9km shows obvious spatial heterogeneity in the northwestern TP, and the peak precipitation months occur during MAM. This simulated peak month may not be spurious but rather an “added value,” because the high-resolution WRF9km simulation can resolve terrain-driven micrometeorological processes (Ou et al. 2023). Domain size is also a key factor for RCMs (Prein et al. 2023); the domain of the WRF9km experiment covers major atmospheric dynamic systems, including the westerlies and Asian monsoons, enabling better capturing of low-level winds and summer precipitation over the TP (Ou et al. 2023). Therefore, the peak precipitation month distribution in the WRF9km simulation reflects the enhanced capacity of the model to resolve topographic features and appropriate domain configuration in capturing regional water vapor transport variability. In the northwestern TP, the WRF9km model accurately represents the spring precipitation peak documented in CN05.1. Despite this agreement, the simulated peak occurs earlier than observed, highlighting a need for more detailed research into the underlying model physics.

For the land surface temperature decomposition, we applied the first-order approximation, assuming a constant land surface emissivity ($\varepsilon = 1.0$) for consistency with prior studies (e.g., Juang et al. 2007; Li et al. 2024). Acknowledging that this neglects actual ε variations and associated uncertainties, we argue that the approach remains valid for achieving the study’s primary objective: quantifying the relative contributions of major energy balance components. Future research that incorporates more accurate and spatiotemporally higher-resolution emissivity data would further enhance the accuracy of the related estimates. Additionally, regarding the model configuration, it should be clarified that the WRF9km simulation, driven by 3-hourly ERA5 data, deviated from the initial experimental design, and spectral nudging was not active due to a technical configuration issue. A subsequent investigation (Ou et al. 2025) revealed that while both the simulations with and without spectral

nudging reproduce broadly consistent spatial patterns of wind vectors at 500 hPa and precipitation in summer, spectral nudging performs better in mitigating the existing wind speed and wet biases. Consequently, the WRF9km results presented herein are based on a free simulation without large-scale nudging, which should be considered when interpreting the model’s internal variability and its departure from the driving fields.

We also find substantial dry biases across the southwestern interior of the TP, which may be related to the resolution of the WRF9km model. The 9-km horizontal resolution, while informative, somewhat limits the accurate representation of the steep terrain on the southern edge of the TP. Consequently, some complex moisture features may not be fully captured. With increasing resolution, the characterization of moisture transport and precipitation formation should become more realistic (Zhou et al. 2023). Looking ahead, a more comprehensive evaluation that includes additional variables—such as soil temperature, soil moisture, and land–atmosphere interaction processes—is essential. Such efforts, especially through the integration of in situ and remote sensing data within CPM frameworks, hold strong potential to yield more reliable and physically coherent precipitation estimates over the TP (Gao et al. 2022; Lin et al. 2018).

5. Conclusions

In this study, we used multisource data to synthetically and quantitatively evaluate the performance of WRF9km climate simulations across the TP during the period 1980–2019. The WRF9km model performed reasonably well in simulating the spatial distribution of temperature, with a spatial correlation coefficient exceeding 0.86, although there was a cold bias of approximately -2.06°C . This cold bias may be related to downward radiation bias and underestimated ground heat flux; the greatest contribution to the cold bias (64%) came from the overestimation of albedo. We attributed the underestimation of downward longwave radiation to the underestimation of high-level cloud cover, while the overestimation of surface albedo was associated with an overestimation of snow cover.

Both CN05.1 and WRF9km data exhibited a significant warming trend across all seasons, but the spatial distribution of this trend differed between datasets, with the simulated warming trend generally being of a smaller magnitude than that of the CN05.1 observations in most areas. Compared with ERA5, the WRF9km substantially reduced the bias in the warming trend across the northern and southwestern TP.

The WRF9km also effectively captured the observed spatial distribution patterns of precipitation, albeit with pronounced spatially heterogeneous biases. The bias pattern was characterized by overestimation on the southern and eastern TP slopes and underestimation across the interior, particularly in the western TP, with the largest biases in JJA. These precipitation biases were associated not only with differences in wind fields and moisture transport but also with variations in near-surface land–atmosphere interactions. By increasing resolution and turning off the cumulus parameterization scheme, the WRF9km better resolved convective processes and land surface–atmosphere interactions over the TP than the driving ERA5 data, leading to a mean decrease of 79% in the wet bias. Although both the WRF9km and ERA5 exhibited spatially opposing trends relative to CN05.1 observations, the WRF9km simulation demonstrated improved performance relative to the ERA5 driving data. This improvement was particularly pronounced on the southern TP across all seasons, where the overestimated drying trend in ERA5 was notably moderated through dynamic downscaling.

Compared with other datasets commonly used across the TP, the WRF9km performed well in simulating the spatial distribution of temperature and precipitation; however, it exhibited poor accuracy against CN05.1 observations in terms of magnitude (bias), spatial variability, and temporal variability. For comprehensive performance, WRF9km ranked fourth in temperature (outperforming ERA5-Land) and second in precipitation (surpassing ERA5, ERA5-Land, and MERRA2). To further improve the simulation accuracy of the WRF9km model, future research directions should focus on refining the snow-cover albedo scheme and the quality of the driving data.

Acknowledgments. This study was jointly supported by the National Natural Science Foundation of China (42477483), the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025-XDXM910), and the International Science and Technology Cooperation Program under the 2024 Shanghai Action Plan for Science, Technology and Innovation (24230780200). The computations/data handling of WRF9km were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through Grant Agreement 2022-06725. This is a contribution to the Swedish national strategic research program MERGE and 24 to CORDEX-FPS-CPTP. This work was also supported by Tsinghua University (100008001) and the Supercomputing Center of Lanzhou University. M.L. acknowledges the support of the Hong Kong Research Grants Council's General Research Fund (16300424) and the support by the Otto Poon Centre for Climate Resilience and Sustainability at HKUST.

This work is part of the United Nations Educational Scientific and Cultural Organization (UNESCO) International Decade of Sciences for Sustainable Development (2024–33) and contributes to the Seamless Prediction and Services for Sustainable Natural and Built Environment (SEPRESS) Program (2025–32), an initiative endorsed under this global framework.

Data availability statement. All the data in this paper are publicly available from the following websites: CN05.1 is from <https://data.cma.cn/>; ERA5 is from <https://cds.climate.copernicus.eu/>; MERRA2 is from <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>; ERA5-Land is from <https://cds.climate.copernicus.eu/>; TPMFD is from <https://data.tpc.ac.cn/zh-hans/data/>; and WRF9km is from <https://biggeo.gvc.gu.se/TPReanalysis/>. MODIS data are downloaded from <https://modis.gsfc.nasa.gov/>.

REFERENCES

- Avissar, R., and T. Schmidt, 1998: An evaluation of the scale at which ground-surface heat flux patchiness affects the convective boundary layer using large-eddy simulations. *J. Atmos. Sci.*, **55**, 2666–2689, [https://doi.org/10.1175/1520-0469\(1998\)055<2666:AEOTSA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<2666:AEOTSA>2.0.CO;2).
- Bieniek, P. A., U. S. Bhatt, J. E. Walsh, T. S. Rupp, J. Zhang, J. R. Krieger, and R. Lader, 2016: Dynamical downscaling of ERA-Interim temperature and precipitation for Alaska. *J. Appl. Meteor. Climatol.*, **55**, 635–654, <https://doi.org/10.1175/JAMC-D-15-0153.1>.
- Chen, R., H. Li, X. Wang, X. Gou, M. Yang, and G. Wan, 2022: Surface air temperature changes over the Tibetan Plateau: Historical evaluation and future projection based on CMIP6 models. *Geosci. Front.*, **13**, 101452, <https://doi.org/10.1016/j.gsf.2022.101452>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Ding, T., Z. Guo, L. Zou, and T. Zhou, 2023: Impact of convection-permitting and model resolution on the simulation of meso-scale convective system properties over East Asia. *J. Geophys. Res. Atmos.*, **128**, e2023JD039395, <https://doi.org/10.1029/2023JD039395>.
- Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.*, **18**, 873–900, [https://doi.org/10.1002/\(SICI\)1097-0088\(19980630\)18:8<873::AID-JOC255>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9).
- Gao, Y., J. Xu, M. Zhang, Z. Liu, and J. Dan, 2022: Regional climate dynamical downscaling over the Tibetan Plateau—From quarter-degree to kilometer-scale. *Sci. China Earth Sci.*, **65**, 2237–2247, <https://doi.org/10.1007/s11430-022-9968-4>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Ghosh, K. G., 2018: Analysis of rainfall trends and its spatial patterns during the last century over the Gangetic West Bengal, eastern India. *J. Geovisualization Spat. Anal.*, **2**, 15, <https://doi.org/10.1007/s41651-018-0022-x>.
- Giorgi, F., 2019: Thirty years of regional climate modeling: Where are we and where are we going next? *J. Geophys. Res. Atmos.*, **124**, 5696–5723, <https://doi.org/10.1029/2018JD030094>.

- Hall, D. K., G. A. Riggs, and V. V. Salomonson, 2007: MODIS/Terra Snow Cover 8-Day L3 Global 500m SIN Grid, version 5 (updated weekly). National Snow and Ice Data Centre, accessed 18 September 2025, <https://doi.org/10.5067/C574UGKQQUIT>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, <https://doi.org/10.1029/2008JD009944>.
- Immerzeel, W. W., L. P. Van Beek, and M. F. Bierkens, 2010: Climate change will affect the Asian water towers. *Science*, **328**, 1382–1385, <https://doi.org/10.1126/science.1183188>.
- IPCC, 2023: Summary for policymakers. *Climate Change 2023: Synthesis Report*, Core Writing Team, H. Lee, J. Romero, Eds., IPCC, 35–115, <https://doi.org/10.59327/IPCC/AR6-9789291691647>.
- Isotta, F. A., and Coauthors, 2014: The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int. J. Climatol.*, **34**, 1657–1675, <https://doi.org/10.1002/joc.3794>.
- Jakob, C., A. Gettelman, and A. Pitman, 2023: The need to operationalize climate modelling. *Nat. Climate Change*, **13**, 1158–1160, <https://doi.org/10.1038/s41558-023-01849-4>.
- Jiang, H., Y. Gao, and G. Wang, 2024: Discrepancies of kilometer-scale dynamic downscaling over the Tibetan Plateau: Underestimation of nocturnal precipitation in summer. *Climate Dyn.*, **62**, 5909–5925, <https://doi.org/10.1007/s00382-024-07183-5>.
- Jiang, Y., and Coauthors, 2025: Development of a high-resolution near-surface meteorological forcing dataset for the Third Pole region. *Sci. China Earth Sci.*, **68**, 1274–1290, <https://doi.org/10.1007/s11430-024-1507-6>.
- Juang, J., G. Katul, M. Siqueira, P. Stoy, and K. Novick, 2007: Separating the effects of albedo from eco-physiological changes on surface temperature along a successional chronosequence in the southeastern United States. *Geophys. Res. Lett.*, **34**, L21408, <https://doi.org/10.1029/2007GL031296>.
- Karki, R., S. Hasson, U. Schickhoff, T. Scholten, J. Böhner, and L. Gerlitz, 2020: Near surface air temperature lapse rates over complex terrain: A WRF based analysis of controlling factors and processes for the central Himalayas. *Climate Dyn.*, **54**, 329–349, <https://doi.org/10.1007/s00382-019-05003-9>.
- Klein, A. G., and A. C. Bennett, 2003: Validation of daily MODIS snow cover maps of the Upper Rio Grande River basin for the 2000–2001 snow year. *Remote Sens. Environ.*, **86**, 162–176, [https://doi.org/10.1016/S0034-4257\(03\)00097-X](https://doi.org/10.1016/S0034-4257(03)00097-X).
- Kong, X., A. Wang, X. Bi, and D. Wang, 2019: Assessment of temperature extremes in China using RegCM4 and WRF. *Adv. Atmos. Sci.*, **36**, 363–377, <https://doi.org/10.1007/s00376-018-8144-0>.
- Latif, A., S. Ilyas, Y. Zhang, Y. Xin, L. Zhou, and Q. Zhou, 2019: Review on global change status and its impacts on the Tibetan Plateau environment. *J. Plant Ecol.*, **12**, 917–930, <https://doi.org/10.1093/jpe/rtz038>.
- Li, P., K. Furtado, T. Zhou, H. Chen, and J. Li, 2021: Convection-permitting modelling improves simulated precipitation over the Central and Eastern Tibetan Plateau. *Quart. J. Roy. Meteor. Soc.*, **147**, 341–362, <https://doi.org/10.1002/qj.3921>.
- Li, X., and Coauthors, 2024: Land data assimilation: Harmonizing theory and data in land surface process studies. *Rev. Geophys.*, **62**, e2022RG000801, <https://doi.org/10.1029/2022RG000801>.
- Li, Y.-J., and Coauthors, 2025: Evaluation of the performance of WRF9km in simulating climate over the upper Yellow River basin. *Adv. Climate Change Res.*, **16**, 58–72, <https://doi.org/10.1016/j.accre.2024.12.003>.
- Lin, C., D. Chen, K. Yang, and T. Ou, 2018: Impact of model resolution on simulating the water vapor transport through the central Himalayas: Implication for models' wet bias over the Tibetan Plateau. *Climate Dyn.*, **51**, 3195–3207, <https://doi.org/10.1007/s00382-018-4074-x>.
- Liu, L., M. Menenti, Y. Ma, and W. Ma, 2022: Improved parameterization of snow albedo in WRF + NOAA: Methodology based on a severe snow event on the Tibetan Plateau. *Adv. Atmos. Sci.*, **39**, 1079–1102, <https://doi.org/10.1007/s00376-022-1232-1>.
- Liu, M., X. Zhang, and B. Chen, 2018: Assessment of the suitability of planetary boundary layer schemes at “grey zone” resolutions. *Chin. J. Atmos. Sci.*, **42**, 52–69, <https://doi.org/10.3878/j.issn.1006-9895.1704.16269>.
- Maussion, F., D. Scherer, T. Mölg, E. Collier, J. Curio, and R. Finkelnburg, 2014: Precipitation seasonality and variability over the Tibetan Plateau as resolved by the High Asia Reanalysis. *J. Climate*, **27**, 1910–1927, <https://doi.org/10.1175/JCLI-D-13-00282.1>.
- Miao, Y., and A. Wang, 2020: A daily $0.25^\circ \times 0.25^\circ$ hydrologically based land surface flux dataset for conterminous China, 1961–2017. *J. Hydrol.*, **590**, 125413, <https://doi.org/10.1016/j.jhydrol.2020.125413>.
- Muñoz-Sabater, J., and Coauthors, 2021: ERA5-land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data*, **13**, 4349–4383, <https://doi.org/10.5194/essd-13-4349-2021>.
- Niu, X., J. Tang, D. Chen, S. Wang, and T. Ou, 2021: Elevation-dependent warming over the Tibetan Plateau from an ensemble of CORDEX-EA regional climate simulations. *J. Geophys. Res. Atmos.*, **126**, e2020JD033997, <https://doi.org/10.1029/2020JD033997>.
- Ou, T., D. Chen, X. Chen, C. Lin, K. Yang, H.-W. Lai, and F. Zhang, 2020: Simulation of summer precipitation diurnal cycles over the Tibetan Plateau at the gray-zone grid spacing for cumulus parameterization. *Climate Dyn.*, **54**, 3525–3539, <https://doi.org/10.1007/s00382-020-05181-x>.
- , —, J. Tang, C. Lin, X. Wang, J. Kukulies, and H.-W. Lai, 2023: Wet bias of summer precipitation in the northwestern Tibetan Plateau in ERA5 is linked to overestimated lower-level southerly wind over the Plateau. *Climate Dyn.*, **61**, 2139–2153, <https://doi.org/10.1007/s00382-023-06672-3>.
- , —, —, —, —, —, and —, 2025: Erratum: Wet bias of summer precipitation in the northwestern Tibetan Plateau in ERA5 is linked to overestimated lower-level southerly wind over the plateau. *Climate Dyn.*, **63**, 148, <https://doi.org/10.1007/s00382-025-07600-3>.
- Pang, G., L. Meng, X. Wang, F. Liu, M. Shu, C. Ma, and T. Che, 2025: Evaluation of surface albedo products from ERA5, JRA-55, MERRA-2, and CERES-EBAF on the Tibetan Plateau: Insights from in-situ observations and MODIS data. *Int. J. Digit. Earth*, **18**, 2547288, <https://doi.org/10.1080/17538947.2025.2547288>.

- Parajka, J., and G. Bloeschl, 2006: Validation of MODIS snow cover images over Austria. *Hydrol. Earth Syst. Sci.*, **10**, 679–689, <https://doi.org/10.5194/hess-10-679-2006>.
- Pepin, N., R. S. Bradley, and H. F. Diaz, 2015: Elevation-dependent warming in mountain regions of the world. *Nat. Climate Change*, **5**, 424–430, <https://doi.org/10.1038/nclimate2563>.
- Polson, D., G. C. Hegerl, R. P. Allan, and B. B. Sarojini, 2013: Have greenhouse gases intensified the contrast between wet and dry regions? *Geophys. Res. Lett.*, **40**, 4783–4787, <https://doi.org/10.1002/grl.50923>.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, <https://doi.org/10.1002/2014RG000475>.
- Prein, A. F., and Coauthors, 2023: Towards ensemble-based kilometer-scale climate simulations over the Third Pole region. *Climate Dyn.*, **60**, 4055–4081, <https://doi.org/10.1007/s00382-022-06543-3>.
- Santanello, J. A., and Coauthors, 2018: Land–atmosphere interactions: The LoCo perspective. *Bull. Amer. Meteor. Soc.*, **99**, 1253–1272, <https://doi.org/10.1175/BAMS-D-17-0001.1>.
- Shin, H. H., and S.-Y. Hong, 2013: Analysis of resolved and parameterized vertical transport in convective boundary layers at gray-zone resolutions. *J. Atmos. Sci.*, **70**, 327–341, <https://doi.org/10.1175/JAS-D-12-0146.1>.
- Simic, A., R. Fernandes, R. Brown, P. Romanov, and W. Park, 2004: Validation of VEGETATION, MODIS, and GOES + SSM/I snow-cover products over Canada based on surface snow depth observations. *Hydrol. Processes*, **18**, 1089–1104, <https://doi.org/10.1002/hyp.5509>.
- Singh, V., S. K. Jain, and P. K. Singh, 2019: Inter-comparisons and applicability of CMIP5 GCMs, RCMs and statistically downscaled NEX-GDDP based precipitation in India. *Sci. Total Environ.*, **697**, 134163, <https://doi.org/10.1016/j.scitotenv.2019.134163>.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.13140/RG.2.1.2310.6645>.
- Sugimoto, S., K. Ueno, H. Fujinami, T. Nasuno, T. Sato, and H. G. Takahashi, 2021: Cloud-resolving-model simulations of nocturnal precipitation over the Himalayan slopes and foothills. *J. Hydrometeorol.*, **22**, 3171–3188, <https://doi.org/10.1175/JHM-D-21-0103.1>.
- Wang, L., J. Zhu, and D. Wang, 2025: Comparative analysis of high-resolution CMIP6 GCM and CMIP5 RCM: Unveiling biases and advancements in simulating compound extreme events in China. *Climate Dyn.*, **63** (2), 1–22, <https://doi.org/10.1007/s00382-024-07571-x>.
- Wang, X., G. Pang, M. Yang, and G. Wan, 2016: Effects of modified soil water-heat physics on RegCM4 simulations of climate over the Tibetan Plateau. *J. Geophys. Res. Atmos.*, **121**, 6692–6712, <https://doi.org/10.1002/2015JD024407>.
- , —, and —, 2018: Precipitation over the Tibetan Plateau during recent decades: A review based on observations and simulations. *Int. J. Climatol.*, **38**, 1116–1131, <https://doi.org/10.1002/joc.5246>.
- , D. Chen, G. Pang, X. Gou, and M. Yang, 2021a: Historical and future climates over the upper and middle reaches of the Yellow River Basin simulated by a regional climate model in CORDEX. *Climate Dyn.*, **56**, 2749–2771, <https://doi.org/10.1007/s00382-020-05617-4>.
- , —, —, S. A. Anwar, T. Ou, and M. Yang, 2021b: Effects of cumulus parameterization and land-surface hydrology schemes on Tibetan Plateau climate simulation during the wet season: Insights from the RegCM4 model. *Climate Dyn.*, **57**, 1853–1879, <https://doi.org/10.1007/s00382-021-05781-1>.
- , V. Tolksdorf, M. Otto, and D. Scherer, 2021c: WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis. *Int. J. Climatol.*, **41**, 743–762, <https://doi.org/10.1002/joc.6686>.
- , and Coauthors, 2022: Contrasting characteristics, changes, and linkages of permafrost between the Arctic and the Third Pole. *Earth-Sci. Rev.*, **230**, 104042, <https://doi.org/10.1016/j.earscirev.2022.104042>.
- , and Coauthors, 2025: Accelerated China's cold regions contraction under climate warming. *Gondwana Res.*, **149**, 229–245, <https://doi.org/10.1016/j.gr.2025.08.004>.
- Wei, S., and Coauthors, 2025: Bias correction of CMIP6 GCMs for historical and future air temperatures across China. *Atmos. Res.*, **323**, 108193, <https://doi.org/10.1016/j.atmosres.2025.108193>.
- Wen, A., and Coauthors, 2022: Evaluation of MERRA-2 land surface temperature dataset and its application in permafrost mapping over China. *Atmos. Res.*, **279**, 106373, <https://doi.org/10.1016/j.atmosres.2022.106373>.
- Wu, J., and X. Gao, 2013: A gridded daily observation dataset over China region and comparison with the other datasets (in Chinese). *Chin. J. Geophys.*, **56**, 1102–1111, <https://doi.org/10.6038/cjg20130406>.
- Xu, J., and Coauthors, 2018: On the role of horizontal resolution over the Tibetan Plateau in the REMO regional climate model. *Climate Dyn.*, **51**, 4525–4542, <https://doi.org/10.1007/s00382-018-4085-7>.
- Xu, X., H. Zhang, Y. Sun, B. Xu, J. Li, G. Dong, and B. Pan, 2024: The effects of climatic change and inter-annual variability on glacier retreat from ~1850s AD moraines in the Kuoqiong-gangri peak region, southern Tibetan Plateau. *Climate Dyn.*, **62**, 2941–2951, <https://doi.org/10.1007/s00382-023-07041-w>.
- Yang, P., J. Tang, Y. Lu, Y. Xu, and S. Wang, 2025: Evaluation of very high-resolution regional climate downscaling over CORDEX East Asia: Climatology and interannual variability. *J. Meteor. Res.*, **39**, 1228–1246, <https://doi.org/10.1007/s13351-025-4249-4>.
- Yang, Y.-X., and Coauthors, 2022: Progress of recent 60 years' climate change and its environmental impacts on the Qinghai-Xizang Plateau. *Plateau Meteorol.*, **41** (1), 1–10, <https://doi.org/10.7522/j.issn.1000-0534.2021.00117>.
- Yao, T., and Coauthors, 2019: Recent Third Pole's rapid warming accompanies cryospheric melt and water cycle intensification and interactions between monsoon and environment: Multi-disciplinary approach with observations, modeling, and analysis. *Bull. Amer. Meteor. Soc.*, **100**, 423–444, <https://doi.org/10.1175/BAMS-D-17-0057.1>.
- , and Coauthors, 2022: The imbalance of the Asian water tower. *Nat. Rev. Earth Environ.*, **3**, 618–632, <https://doi.org/10.1038/s43017-022-00299-4>.
- Yimer, S. M., A. Bouanani, N. Kumar, B. Tischbein, and C. Borgemeister, 2022: Assessment of climate models performance and associated uncertainties in rainfall projection from CORDEX over the Eastern Nile basin, Ethiopia. *Climate*, **10**, 95, <https://doi.org/10.3390/cli10070095>.
- Zhang, C., S. Xie, S. A. Klein, H. Ma, S. Tang, K. Van Weverberg, C. J. Morcrette, and J. Petch, 2018: CAUSES: Diagnosis of the summertime warm bias in CMIP5 climate models at the ARM Southern Great Plains site. *J. Geophys. Res. Atmos.*, **123**, 2968–2992, <https://doi.org/10.1002/2017JD027200>.

- Zhang, S., M. Wang, L. Wang, X.-Z. Liang, C. Sun, and Q. Li, 2023: Sensitivity of the simulation of extreme precipitation events in China to different cumulus parameterization schemes and the underlying mechanisms. *Atmos. Res.*, **285**, 106636, <https://doi.org/10.1016/j.atmosres.2023.106636>.
- Zhou, P., M. Shao, M. Ma, T. Ou, and J. Tang, 2023: WRF gray-zone dynamical downscaling over the Tibetan Plateau during 1999–2019: Model performance and added value. *Climate Dyn.*, **61**, 1371–1390, <https://doi.org/10.1007/s00382-022-06631-4>.
- Zou, L., and T. J. Zhou, 2024: Convection-permitting simulations of current and future climates over the Tibetan Plateau. *Adv. Atmos. Sci.*, **41**, 1901–1916, <https://doi.org/10.1007/s00376-024-3277-9>.
- , T. Zhou, and Y. Zhao, 2024: How does regional convection-permitting modeling improve the simulation of the atmospheric water cycle in spring over the Tibetan Plateau? *J. Geophys. Res. Atmos.*, **129**, e2024JD040964, <https://doi.org/10.1029/2024JD040964>.