npj | climate and atmospheric science

Article

Published in partnership with CECCR at King Abdulaziz University

https://doi.org/10.1038/s41612-025-00916-1

Excessive equatorial light rain causes modeling dry bias of Indian summer monsoon rainfall

Gudongze Li[1], Chun Zhao[1,2,3] ✉, Jun Gu[1], Jiawang Feng[1], Mingyue Xu[1], Xiaoyu Hao[4], Junshi Chen[2,4], Hong An[2,4] ✉, Wenju Cai[2] & Tao Geng[2]

Simulating accurately the South Asian summer monsoon is crucial for food security of several South Asian countries yet challenging for global climate models (GCMs). The GCMs suffer from some systematic biases including dry bias in mean monsoon rainfall over the India subcontinent and excessive equatorial light rain between which the relationship was rarely discussed. Numerical experiments are conducted for one month during active monsoon with global quasi-uniform resolution of 60 km (U60 km) and 3 km (U3 km) separately. Evaluation with observations shows that U3 km reduces the dry bias over northern India and excessive light rain over the equatorial Indian Ocean (EIO) that are both prominent in U60 km. Excessive light rain in U60 km contributes critically to stronger rainfall and latent heating over the EIO. A Hadley-type anomalous circulation is thus induced, whose subsidence branch suppresses updrafts and reduces moisture transport into northern India, contributing to the dry bias. The findings highlight the importance of constraining excessive light rain for regional climate projection in GCMs.

South Asian summer monsoon rainfall is one of the most important climatic phenomena, crucial for the food and economic security of several South Asian nations. It profoundly impacts the lives of over 1 billion inhabitants[1–3]. Global climate models (GCMs) are essential tools for advancing the understanding and projection of the monsoon system. The community have witnessed the success of GCMs in capturing the overall spatial and temporal variation of South Asian monsoon precipitation[4–7]. However, the scientific community has evaluated GCMs from the CMIP3 to CMIP6 (the third, fifth and sixth coupled model intercomparison projects) and has identified systematic shortcomings that are common and persistent in most GCMs, including the dry bias in mean summer monsoon rainfall over the Indian subcontinent[4,5,8–11]. This systematic bias may lead to uncertainties in predicting summer floods over India. It also affects the ability to capture the future evolution of monsoon rainfall under rapid climate change[12–14].

This dry bias over the Indian subcontinent is associated with both air-sea coupling and atmospheric processes. Cold sea surface temperature (SST) biases over the northern Indian Ocean can diminish the moisture flux from the Arabian Sea and reduce precipitation frequency and intensity over India, contributing to the reduction of monsoon rainfall[9,10,15]. Within the Indian longitudes, there is a competition between two convergence zones over the equator and subcontinent[16–18], respectively, so that the weak Indian summer precipitation is usually accompanied by a positive rainfall bias over the equatorial Indian Ocean (EIO)[5,19]. Thus, warm SST biases and distorted SST-precipitation interaction over the EIO induce an anomalous Hadley-type circulation whose descending branch suppresses precipitation over the Indian subcontinent[20]. On the other hand, Hagos et al. [14] suggested that the mean dry bias in summer monsoon rainfall in CMIP5 models may result from excessive mean precipitation over the EIO during spring. This excessive precipitation is a consequence of erroneous profiles of water vapor and divergence over the EIO, which may stem from the convection schemes employed in CMIP5 models. Previous studies have also demonstrated that atmosphere-only GCMs forced by observed SST can reproduce similar biases[19,21,22] to those in coupled models. Moreover, individual atmospheric-only models exhibit even stronger biases[23]. Analysis of these models revealed that convections over the South Asia were suppressed by inaccurate sensible heat flux and released latent heat in pre-monsoon through weakening the meridional heating gradient[21]. An atmospheric-only model, MetUMGA3.0, could achieve significant improvements in the summer precipitation bias

[1]Deep Space Exploration Laboratory/School of Earth and Space Sciences, CMA-USTC Laboratory of Fengyun Remote Sensing, State Key Laboratory of Fire Science, Institute of Advanced Interdisciplinary Research on High-Performance Computing Systems and Software, University of Science and Technology of China, Hefei, China. [2]Laoshan Laboratory, Qingdao, China. [3]CAS Center for Excellence in Comparative Planetology, University of Science and Technology of China, Hefei, China. [4]School of Computer Science and Technology, University of Science and Technology of China, Hefei, Anhui, China.
✉e-mail: chunzhao@ustc.edu.cn; han@ustc.edu.cn

over the western EIO and India subcontinent by solely increasing the entrainment/detrainment rates in the western EIO[19], which are tunable parameters of the entrainment process in convection scheme. Process-based diagnostics applied to the NCAR CAM4, also indicated that the entrainment process was a potential source of biases in monsoon rainfall[22]. Furthermore, sensitive experiments using NCAR CAM4 demonstrated that strengthening environmental mixing globally improved moist convection over the Arabian Sea, mitigating the wet bias over the Arabian Sea and dry bias over the South Asia[22].

Although there has been disagreement on the critical source of the dry bias over the Indian subcontinent[9,14,19,20,24], the key process underlying this bias is moist convection through which both air-sea coupling and atmospheric processes modulate the monsoon summer precipitation. However, the convection schemes widely used in GCMs often exhibit common biases[25]. For example, the convection schemes tend to generate excessive convections with weakened intensity, primarily contributing to the widespread bias "too much drizzle" which is still persistent in CMIP6 GCMs[26–28]. Even if the global precipitation amount is comparable with observations, precipitation with different intensity could result in various vertical profiles of moisture and diabatic heating, which may affect the meridional heating gradient and thus the summer monsoon rainfall[21]. Moreover, a recent study[29] demonstrated that improvement of precipitation frequency can contribute to better intraseasonal evolution of tropical precipitation. Therefore, an inaccurate precipitation intensity spectrum is likely to be one important source of the dry bias over the Indian subcontinent, which is discussed in few studies[10,24]. The convection schemes are necessary for most GCMs at a resolution of tens to hundreds of kilometers, due to associated small-scale processes such as updrafts and downdrafts of convections[30]. Rather, global convection-permitting models (GCPMs) enable deep convections to be simulated explicitly, avoiding the usage of convection parameterization at resolutions of kilometers[31,32]. Simulations using GCPMs can thus serve as validations with realistic climate and weather patterns for data assimilation systems, parameterizations and current GCMs etc.[27,33,34].

Recent advancements in exascale computing have enabled the implementation of a nonhydrostatic atmospheric model with a global uniform resolution of convection-permitting scale (e.g., 3 km), the integrated Atmospheric Model Across Scales (iAMAS) on the heterogeneous Sunway supercomputer[35–37]. In this study, we demonstrate that continuous simulation during active monsoon using iAMAS at global convection-permitting scale can well capture the monsoon rainfall and circulation to reduce the dry bias over the Indian subcontinent that is present in the simulation at a global uniform resolution of 60 km. Further analysis reveals that there is an overestimation of the light-rain frequency in convection parameterized simulation compared to the convection-permitting simulation. This overestimation is particularly significant over the EIO, primarily contributing to stronger latent heating vertically. Consequently, stronger latent heating induces a Hadley-type anomalous meridional circulation in convection parameterized simulation, with its subsident branch weakening precipitation over northern India. Therefore, the "too much drizzle" bias associated with convection schemes is one important scientific issue that needs to be addressed to improve the simulation of South Asian summer monsoon rainfall and its projection under climate change.

Results

Validation of simulated monsoon rainfall and circulation

Before conducting a detailed validation of the simulated Indian monsoon, an overall analysis of global rainfall in models can provide a fundamental reference. Both global convection-parameterized (U60 km in Methods) and convection-permitting (U3 km in Methods) simulations show general agreement with observations GPCP and GPM, capturing major spatial distribution of mean rainfall averaged from 11th June to 9th July in 2020 (Fig. S1). Quantitatively, U3 km shows a slight improvement over U60 km when compared to GPM: the root-mean-square error (RMSE) decreases from 2.45 mm day^{-1} in U60 km to 2.40 mm day^{-1} in U3 km, while the correlation coefficient increases from 0.74 to 0.76 (Fig. S1c, d). These metrics suggest that U3 km achieves a modest enhancement in global rainfall and furthermore provides more refined patterns. The spatial distribution of rainfall bias in U3 km (Fig. S2b) also closely resembles that of U60 km (Fig. S2a). However, U3 km generally shows reduced biases compared to U60 km, with one particularly pronounced improvement located over the Indian monsoon rainfall region (Fig. S2c). The spatial distribution of mean rainfall during the same period is generally reproduced by three numerical experiments (Methods) across the study area, with notable discrepancies over certain rainbelts (Fig. S3). During this period, the South Asian monsoon has reached the India subcontinent and brought abundant moisture, leading to heavy rainfall concentrated over the western coast of Indian continent, northern India, Bay of Bengal and the southern foothills of the Himalayas (Fig. 1a, b). U60 km forced by observed SST overestimates precipitation over the southern foothills of the Himalayas but fails to capture the heavy precipitation over the India subcontinent (Fig. 1c), particularly over northern India (72.5–92.5°E, 15–25°N; marked in Fig. 1). This is consistent with the systematic dry bias in Indian summer rainfall that is commonly observed in many GCMs[10,24]. In contrast, the rainfall over the southern slope of the Himalayas is improved in the simulation with a regional refinement at convection-permitting scale over the India subcontinent (V4 km in Methods) relative to U60 km (Fig. 1d), likely due to the higher-resolution that better resolves complex topography[38]. But the dry bias still occurs over northern India in V4 km (Fig. 1d), indicating that locally refined convections and topography, as discussed in previous studies[8,21,39], are not sufficient to overcome the dry bias in our experiments. U3 km not only reduces overestimation of mean rainfall over the south foothills of the Himalayas as V4 km but also improves the dry bias over northern India, even leading to a slight overestimation of mean rainfall in this region (Fig. 1e; Fig. S3). Additionally, the mean rainfall pattern of South Asian summer monsoon includes a secondary rainbelt over the EIO (50–100°E, 0–10°S; marked in Fig. 1e). In U60 km and V4 km, the precipitation over this secondary rainbelt is significantly stronger than that in observations and covers larger area (Fig. 1c, d). In contrast, the precipitation over this secondary rainbelt in U3 km is improved relative to U60 km, approaching the observed precipitation levels with minimal spatial discrepancies (Fig. 1e; Fig. S3). The regional average precipitation over the EIO in U60 km and V4 km is 11.66 and 13.67 mm day^{-1}, respectively, which are much higher than observations (8.17 mm day^{-1}). But the precipitation in U3 km is only 8.47 mm day^{-1}. The relative magnitudes of precipitation averaged over northern India and EIO suggest a competitive relationship between these two rainbelts[16,17,20] so that dry biases over northern India in U60 km and V4 km are accompanied with the wet biases over the EIO but both rainfall biases are improved simultaneously in U3 km (Fig. 1f). Therefore, U60 km reproduces the systematic biases of summer monsoon rainfall common in many GCMs as expected, while U3 km captures a more realistic mean pattern. Moreover, the systematic biases occurring in V4 km shows that regional refinement limited to the India subcontinent is not sufficient to simulate the Indian summer monsoon rainfall realistically. Consequently, the results from V4 km will not be further discussed.

Figure 2a depicts the precipitation intensity spectrums of observations and experiments counted across the focus area (35°E-115°E, 20°S-40°N). Owing to the constrained duration of the simulation, the precipitation intensity classification employs three-hour precipitation in lieu of the commonly used daily precipitation (Methods). The precipitation frequency decreases sharply with stronger intensity. The frequency of light rain (0.1–3 mm 3 hr^{-1}) is 14.1% in CMORPH and 15.9% in GPM, while the frequency of intense rain (> 10 mm 3 hr^{-1}) is 1.0% or less. In U60 km, the light rain occurs with a frequency of 42.5%, nearly three times that in observations, yet the frequency of intense rain is significantly underestimated. This is the well-known bias "too much drizzle", which may be attributed to the convection parameterization triggering excessive convections and subsequently inhibiting the accumulation of convective available potential energy (CAPE) favorable to intense rain[40]. In U3 km, where convection is resolved explicitly avoiding the use of a convection scheme, the frequency of light rain is reduced to 31.5%, and the frequency of intense rain

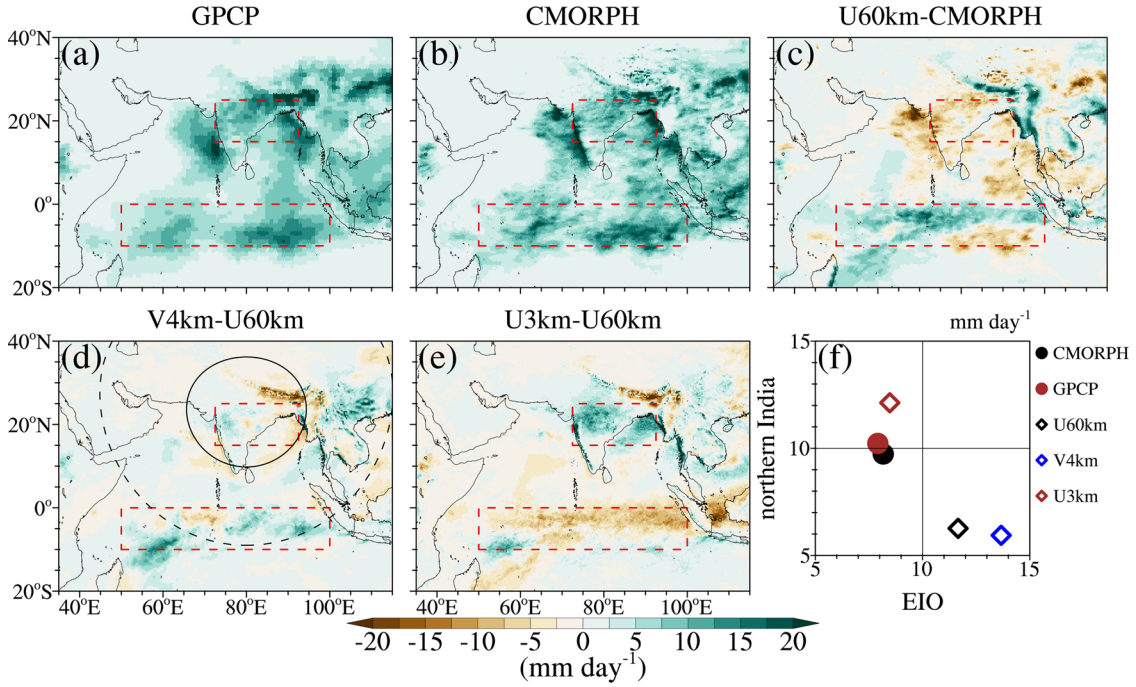

Fig. 1 | South Asian summer monsoon rainfall during active monsoon in 2020. The spatial distributions of mean rainfall (mm day^{-1}) in (**a**) the observation datasets GPCP and (**b**) CMORPH, averaged from 11th June to 9th July in 2020. The spatial distributions of mean rainfall difference (mm day^{-1}) between (**c**) U60 km and CMORPH; (**d**) V4 km and U60 km; (**e**) U3 km and U60 km. The two regions surrounded by red dotted line in **a**-**e** indicate the primary rainbelt over the northern India (72.5–92.5°E, 15–25°N) and the secondary rainbelt over the equatorial Indian Ocean (EIO; 50–100°E, 0–10°S) respectively. **f** The mean rainfall averaged over two rainbelts in observations and experiments above. The area enclosed by the black solid line in **d** is in a resolution of ~4 km, the area outside the black dotted line is in a resolution of ~60 km, and the area between two lines is the transition zone.

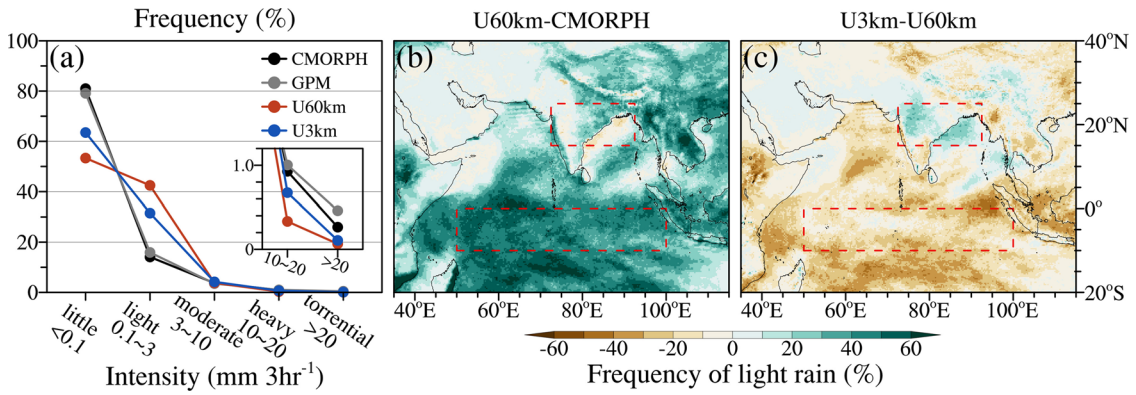

Fig. 2 | Precipitation intensity spectrum during active monsoon in 2020. a the frequency of five rainfall intensity levels in CMORPH, GPM, U60 km and U3 km, counted from 11th June to 9th July in 2020. The spatial distribution of difference in light rainfall frequency between (**b**) U60 km and CMORPH; (**c**) U3 km and U60 km during the same period. The regions surrounded by red dotted line are same as that in Fig. 1.

is approximately doubled compared with that in U60 km. Considering that light rain occurs most frequently and has the largest difference of frequency except little rain, it is necessary to analyze spatial distributions of light-rain frequency in U60 km and U3 km. Over the Indian subcontinent, excessive light rain in U60 km covers the regions including the western coast, northern India, and the south foothills of the Himalayas where summer monsoon rainfall concentrates (Fig. 2b). Light rain in U60 km also occurs frequently over most of the tropical Indian Ocean where there is little rain covering in observation instead (Fig. 2b; Fig. S4e, f). But the spatial correlation coefficient of light-rain frequency between U3 km and CMORPH is 0.80, a significant improvement compared to the correlation between U60 km and CMORPH (0.72). Although the frequency of light rain is still overestimated over the entire focus area in U3 km compared to observations, the light rain in U3 km is less excessive than that in U60 km, particularly over the Indian Ocean (Fig. 2c; Fig. S4e–h), which is consistent with the difference in regional average frequency. It is worth noting that in U60 km the overestimated frequency of light rain coincides with the wet bias in mean monsoon rainfall over the EIO. Given that the observed bias in precipitation frequency is likely attributable to the convection scheme, one fundamental component of GCMs, it can be hypothesized that the inaccurate precipitation intensity spectrum over the focus area contributes to wet bias over the EIO. This, in turn, leads to the dry bias in mean monsoon rainfall over northern India in U60 km due to the competition between rainbelts over northern India and EIO.

One advantage of the GCMs is the ability to reproduce atmospheric circulations continuously in both vertical and horizontal directions. Before

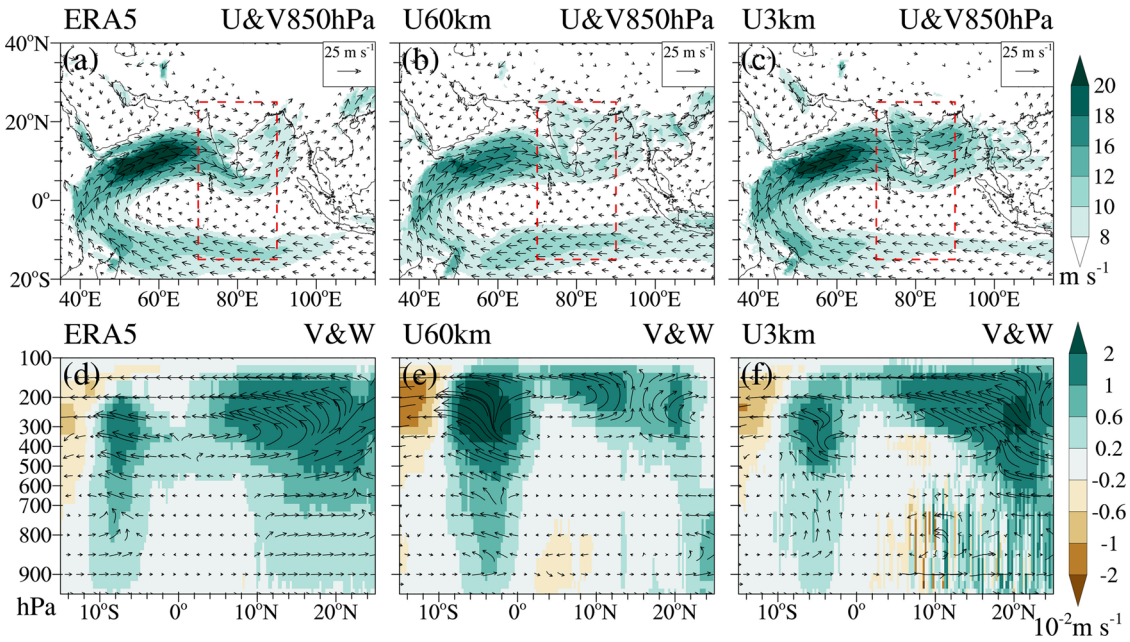

Fig. 3 | South Asian summer monsoon circulation during active monsoon in 2020. The horizontal wind field (vectors) and speed (color shading) at 850 hPa in (**a**) the reanalysis ERA5, the experiment (**b**) U60 km and (**c**) U3 km averaged from 11th June to 9th July in 2020. The areas surrounded by the red dotted lines in a–c indicate the region (70–90°E, 15°S-25°N) where the zonal-mean vertical circulation is analyzed. The zonal-mean vertical circulation (with vertical velocity amplified 1000 times, vectors) and vertical velocity (without amplified, color shading) in (**d**) the reanalysis ERA5, the experiment (**e**) U60 km and (**f**) U3 km, averaged from 11th June to 9th July in 2020.

applying these simulations to analyze how differences in rain intensity spectra affect monsoon circulation and precipitation, the monsoon circulations in two experiments need to be evaluated against the reanalysis ERA5. During the South Asian summer monsoon, monsoon winds blow northward across the equator from the southern hemisphere, swing clockwise into the Somali jet, which reaches its maximum strength (23.8 m s^{-1}) at the latitude of 11.5°N, before turning into westerlies over the Arabian sea (Fig. 3a). The Somali jet in the low level transports moisture from the equator, across the Arabian sea and into the west coast of India subcontinent, playing a crucial role in summer monsoon rainfall over the India subcontinent[19]. The maximum wind speed of the Somali jet is 19.4 m s^{-1} in U60 km and 21.9 m s^{-1} in U3 km (Fig. 3b, c). Although the horizontal wind fields in low level are generally similar across two experiments and ERA5, the Somali jet is significantly weakened in U60 km but in U3 km is comparable compared with that in ERA5. Meanwhile, the competition between two rainbelts over northern India and EIO is linked to two convergence zones over the same region. During this simulated period, the convergence zone over northern India features organized convections with strong convergence and updraft while the convergence zone over the EIO appears intermittently and then moves northward, contributing to the persistence of the convergence zone over northern India[16]. As a result, the mean pattern during this period reveals that the continental convergence zone has stronger updrafts than the maritime convergence zone, playing a dominant role (Fig. 3d). In ERA5, updrafts over the primary convergence zone cover most of the India subcontinent (north of 10°N) and are strengthened from low to high level, then turns southward at upper troposphere. The ascending flow over the secondary convergence zone has weaker vertical velocity and is primarily located from 4°S to 11°S. In U60 km, the maximum vertical velocity at high level over the EIO is significantly increased and shifts northward, along with a general decrease in the vertical velocity at high level over northern India relative to that in ERA5 (Fig. 3d, e). This indicates that updrafts are strengthened over the EIO but weakened over northern India in U60 km. Thus, the maritime convergence zone dominates during the simulated period instead of the continental convergence zone in U60 km, which is opposite to the pattern in ERA5. The magnitudes of vertical velocity in U3 km are comparable to those in ERA5 over both two convergence zones with a slight overestimation of vertical velocity at high level (Fig. 3d, f). Therefore, U3 km successfully reproduces the strength of updrafts over the continental and maritime convergence zones and their competitive relationship. In addition, there are some locally meso-scale fluctuations with upward and downward flow occurring widely over the Indian subcontinent below 600 hPa, rather than the consistently upward flow observed in ERA5.

The South Asian summer monsoon rainfall and circulations in two experiments are found to agree with observations and the reanalysis ERA5. Simulation of U60 km exhibits a systematic dry bias over northern India and generates excessive light rain but lacks intense rain similar to many GCMs. The frequency of light rain in U60 km is particularly high over the tropical Indian Ocean, including the EIO where there is the positive bias in mean monsoon rainfall simultaneously. Further evaluations of monsoon circulation reveal that updrafts over the EIO are strengthened in U60 km, even surpassing those over northern India. Therefore, the maritime convergence zone assumes a dominant role, supplanting the continental convergence zone. These findings support the hypothesis that the inaccurate precipitation intensity spectrum in U60 km leads to the dry bias over northern India by influencing the competition between continental and maritime convergence zones. In contrast, U3 km not only improves the biases in mean summer monsoon rainfall but also accurately reproduces the horizontal and vertical monsoon circulations during the simulated period. This suggests that the simulation of U3 km is closer to the realistic South Asian summer monsoon. The differences in precipitation amount, frequency and monsoon circulation between U60 km and U3 km can be used to validate the aforementioned hypothesis and explain how the excessive light rain in U60 km affects the monsoon circulation and then induces biases over the two rainbelts.

Impacts of excessive light rain over the EIO

There are significant discrepancies in precipitation amount and frequency between U60 km and U3 km over the EIO (Fig. 1e; Fig. 2c) so that Fig. 4a, b further illustrate the regional average precipitation frequency and accumulated amounts, respectively, over the EIO. In CMORPH (GPM), the light-rain frequency increases to 23.4% (25.5%) over the EIO while the heavy-rain frequency increases to 1.9% (2.0%) relative to that over the whole

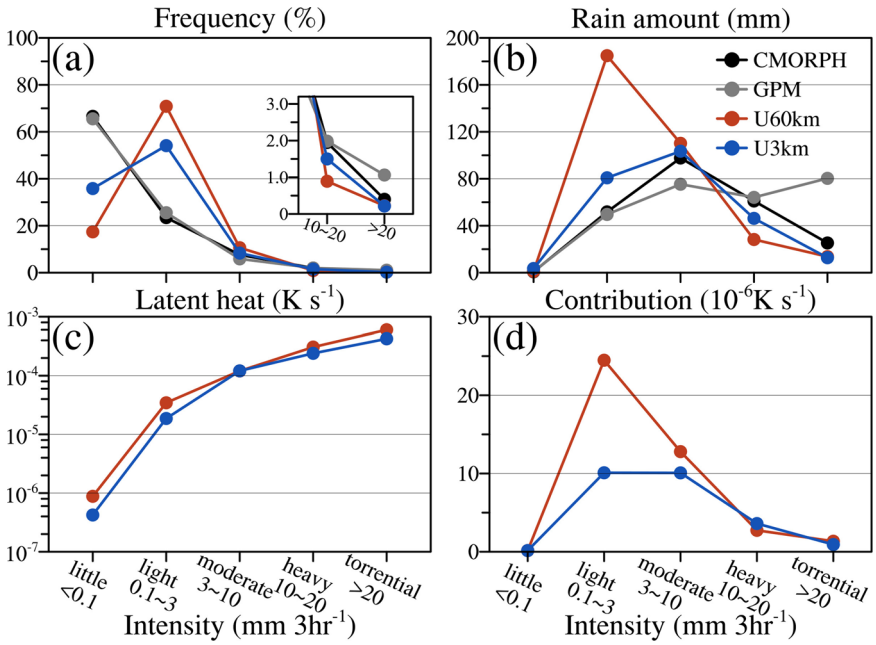

Fig. 4 | Frequency, accumulated amount and latent heat of precipitation with different intensity levels over the EIO during active monsoon. **a** the frequency of five rainfall intensity levels in CMORPH, GPM, U60 km and U3 km, counted from 11th June to 9th July in 2020 over the EIO (50–100°E, 0–10°S). **b** The accumulated rainfall of rainfall with different intensities. **c** The column-averaged latent heat released by rainfall with different intensities in U60 km and U3 km. **d** The contribution (frequency times column-averaged latent heat with the same level) to averaged latent heat by rainfall with different intensities in U60 km and U3 km.

focus area (Fig. 2a; Fig. 4a). Meanwhile, the cumulative precipitation in light-rain level is 51.8 mm (49.6 mm) in CMORPH (GPM), contributing to a relatively small portion of the mean monsoon rainfall (Fig. 4b). In contrast, the U60 km generates excessive light rain over the EIO with a frequency of 70.9% and, still lacks heavy rain though its frequency increases to 0.89% (Fig. 4a). The discrepancy in precipitation frequency between U60 km and observations is further exacerbated over the EIO, suggesting that the inaccurate representation of the rain intensity spectrum in the U60 km is particularly pronounced in this region. The excessive light rain over the EIO in U60 km contributes significantly to the total cumulated precipitation, with a total of 184.9 mm, far exceeding the contributions of other precipitation levels (Fig. 4b). Compared with the observations, U60 km greatly overestimates the accumulated precipitation amount of light rain while the accumulated precipitation contributed by other intensity levels exhibits relatively little discrepancy. The differences in the accumulated precipitation suggest that the excessive light rain in U60 km play a dominant role in summer monsoon rainfall over the EIO and competition between continental and maritime rainbelt. In U3 km, the light-rain frequency over the EIO is 54.1% with the heavy-rain frequency increasing to 1.5%, which represents an improvement relative to that in U60 km (Fig. 4a). The light rain over the EIO has an accumulated precipitation of 80 mm in U3 km, indicating that the light rain does not dominate as it does in U60 km (Fig. 4b) though the light-rain frequency is also overestimated. The other four rainfall levels in U3 km contribute to a similar accumulated precipitation amount as in CMORPH over the EIO (Fig. 4b). The rainbelt over the EIO is actually part of the intertropical convergence zone (ITCZ) over the Indian longitudes[16] where precipitation is usually associated with active convections. The vertical latent heat released by convections plays a critical role in the interaction between precipitation and circulation over the EIO. The column-averaged latent heat over the EIO during the simulated period is 41.5 (10^{-6}K s^{-1}) in U60 km and is 24.8 (10^{-6}K s^{-1}) in U3 km. Figure 4c further illustrates the column-averaged latent heat for each rainfall level. In U60 km, the released latent heat of all levels is also stronger than that in U3 km. The rainfall intensity is classified discretely rather than continuously, which means that the average intensity of the same rain level can vary in U60 km and U3 km. This may result in different average latent heat released by rainfall with the same level in U3 km and U60 km. Variation in the share of grid-scale and convective precipitation may also be a source of discrepancies in latent heat released by each precipitation level, which is beyond the scope of discussion in this study[41]. To assess the contribution of different rainfall intensities to vertical latent heating, we calculate the frequency-weighted column-averaged latent heat for each level (Fig. 4d). In U3 km, the contribution of light rain to vertical heating over the EIO is approximately equal to that of moderate rain, with other three levels contributing relatively little. In contrast, light rain in U60 km dominates the total contribution, with a magnitude significantly greater than that in U3 km. Therefore, the excessive light rain in U60 km leads to the positive bias in mean summer monsoon rainfall over the EIO and plays a dominant role in enhancing vertical latent heating compared to U3 km.

Differences of vertical and horizontal wind fields between U60 km and U3 km are representative of how discrepancy in released latent heating by precipitation affects the associated monsoon circulation. The air temperature at the mid-upper troposphere over the EIO is heated by stronger latent heating in U60 km (Fig. S5). This leads to an increase in the upward wind velocity, resulting in an anomalous upward air flow whose velocity achieves maximum at high altitudes (Fig. 5a). Meanwhile, the upward air flow diverges at the upper troposphere over the EIO and moves towards the poles. A compensatory downward motion is then induced over northern India (Fig. 5a), which can suppress moist convection and the release of latent heat, accompanied by a slight anomalous cooling (Fig. S5c) at the mid-upper troposphere. These differences in vertical circulation between U60 km and U3 km form a Hadley-type anomalous circulation, similar to previous studies[13,20]. There is a significantly high-pressure deviation at the low troposphere over northern India, which favors the additional divergence of airflow (Fig. 5b). An anticyclonic anomaly is hence triggered with the easterlies over its southern region decelerating the westerlies from the Arabian sea to the Indian subcontinent (Fig. 5b; Fig. 3a). This relationship between the monsoon trough and cross-equatorial flow is also supported by numerical experiments in a previous study[42]. In conclusion, the descending branch of this Hadley-type not only suppresses upward motion but also weakens the moisture transport over northern India (Fig. S6), consequently contributing to the dry bias of this continental rainbelt.

However, the south Asian summer monsoon is also under the control of other large-scale circulations. Comparison between these two experiments alone cannot fully isolate the sensitivity of monsoon rainfall and circulation from discrepancies in other regions (Fig. S2c). To address this limitation, we conducted an additional experiment (V4 kmTropics in Methods) with regional refinement at a convection-permitting scale (4 km) within the equatorial region (40°E-100°E, 10°S-0°; Fig. S7). In the EIO, V4 kmTropics substantially reduced the average precipitation to

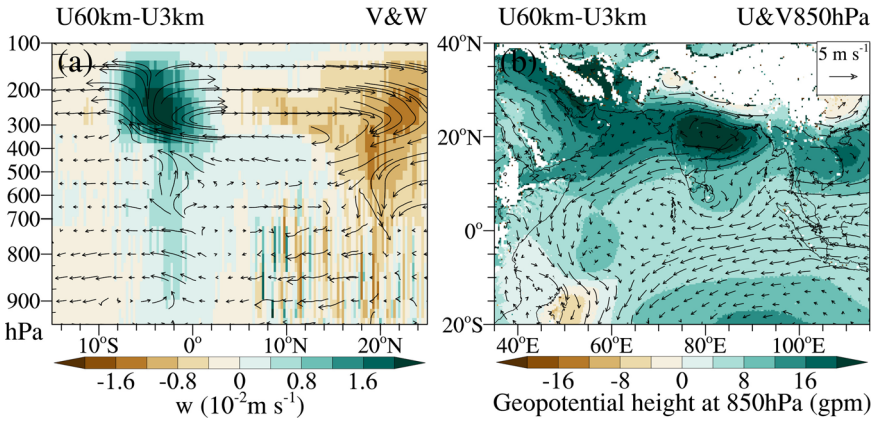

Fig. 5 | Impacts of excessive light rain over EIO on south Asian summer monsoon circulation. a The difference in zonal-mean vertical circulation (with vertical velocity amplified 1000 times, vectors) and vertical velocity (without amplified, color shading) between U60 km and U3 km, averaged from 11th June to 9th July in 2020. **b** The difference in horizontal wind field (vectors) and geopotential height (color shading) at 850 hPa, averaged from 11th June to 9th July in 2020.

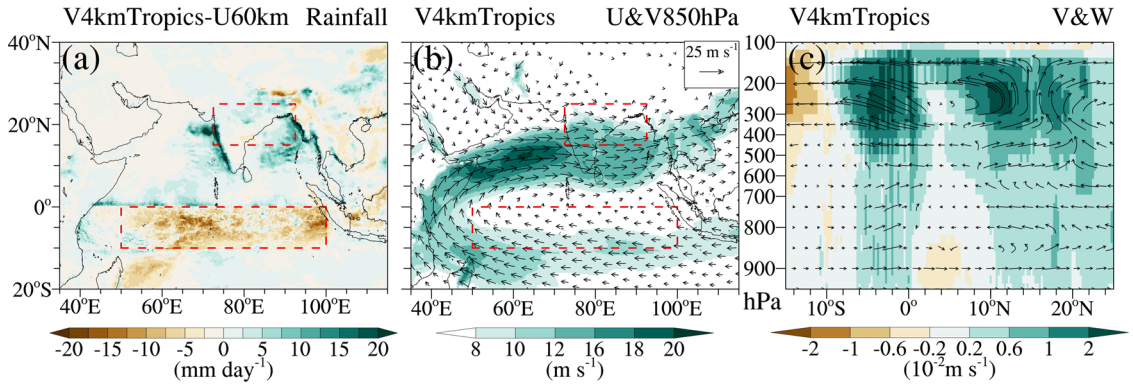

Fig. 6 | South Asian summer monsoon rainfall and circulation during active monsoon in 2020. a The difference in mean rainfall between V4 kmTropics and U60 km, averaged from 11th June to 9th July in 2020. **b** The horizontal wind field (vectors) and speed (color shading) at 850 hPa in V4 kmTropics, averaged from 11th June to 9th July in 2020. **c** The zonal-mean vertical circulation (with vertical velocity amplified 1000 times, vectors) and vertical velocity (without amplified, color shading) in V4 kmTropics, averaged from 11th June to 9th July in 2020.

6.88 mm day^{-1}, compared to 11.66 mm day^{-1} in U60km (Fig. 6a; Fig. S8a; Fig. S3d). This reduction effectively mitigates the wet bias observed in U60 km over the EIO (Fig. 1c). Concurrently, the Somali Jet intensity in V4 kmTropics more closely aligns with ERA5 relative to U60 km (Fig. 6b; Fig. 3a, b), and the ascending flow over the EIO in V4 kmTropics is also in better agreement with ERA5 than that in U60 km (Fig. 6c; Fig. 3d, e). Therefore, the vertical air motion over northern India is enhanced (Fig. 6c; Fig. 3e) under the control of competition relationship between two rainbelts, alleviating the dry bias over the northern India observed in U60 km (Fig. 6a; Fig. S8c). In summary, the comparison between V4 kmTropics and U60 km further proves the conclusion that suppressed rainfall over EIO at higher resolution can improve the simulated monsoon circulation and monsoon rainfall over northern India. Moreover, V4 kmTropics demonstrates a stronger decrease in precipitation over the EIO but a less pronounced improvement in dry bias over northern India compared with U3 km (Figs. 6a, 1e; Fig. S8c). This discrepancy can be attributed to two key factors: the important role of explicitly simulated local convection in modeling monsoon rainfall, and the impact of uncertainties potentially brought by the resolution transition zone on monsoon circulation, particularly the Somali Jet. Resolution transitions can compromise numerical algorithm precision[43] and significantly impact key meteorological fields such as cloud fraction, integrated water vapor and meridional wave propagation[44]. A narrow transition zone with abrupt variations in grid spacing may generate spurious numerical interactions, for example, nonphysical gradients of integrated water vapor. In V4 kmTropics, these transition zone effects could be evidenced by a high-precipitation band over the resolution transition zone (Fig. S7; Fig. S8a) and a notably weakened Somali Jet (Figs. 3c, 6b).

Discussion

Our evaluation using observations and reanalysis showed that the U3 km reduced the dry bias of the South Asia summer monsoon rainfall over northern India, the inaccurate rainfall intensity spectrum, and the monsoon circulation compared with the results of U60 km. Meanwhile, the frequency of light rain was particularly overestimated over the EIO in U60 km relative to the observations, coexisting with the wet bias of rainfall. We thus hypothesized that the excessive light rain in U60 km contributed to stronger rainfall over the EIO and then weakened the continental rainfall through competition between the maritime and continental convergence zone[13,20]. The precipitation amount and latent heat of two experiments had significant discrepancies over the EIO with the differences of light rain playing a dominant role, which supported our hypothesis. Differences in monsoon circulation between U60 km and U3 km could illustrate how excessive light rain and associated stronger latent heat modulating the competition between the continental and maritime convergence zone. As showed in Fig. 7, the upward flow was strengthened by stronger latent heating over the maritime convergence zone and resulted in a compensatory subsidence over the continental convergence zone, forming a Hadley-type anomalous circulation. The subsidence branch of this anomalous circulation suppressed upward motion and weakened moisture transport over northern India, thus contributing to dry bias of the South Asian summer monsoon rainfall (Fig. 7). The additional experiment with regional refinement over the EIO also demonstrates the interplay between excessive equatorial light rain and dry bias over northern India. Furthermore, it highlights the added

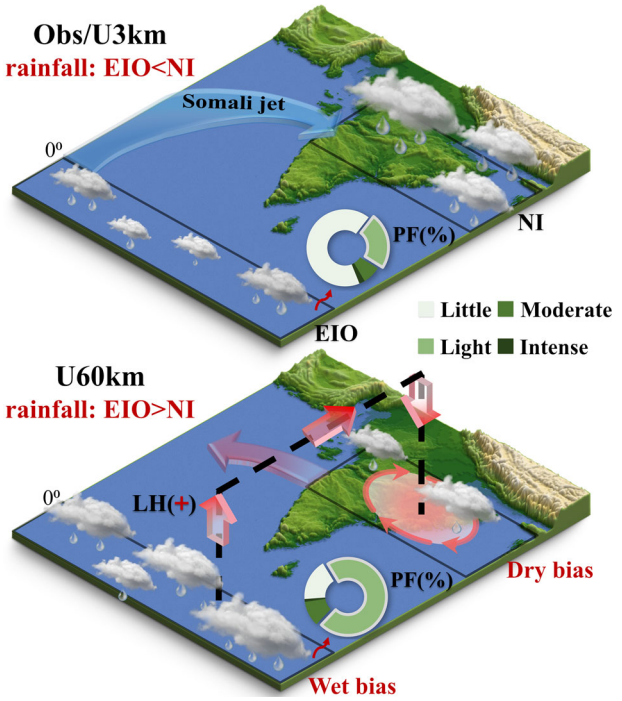

Fig. 7 | Schematic of the linkage between equatorial excessive light rain and dry bias over northern India. The regions outlined by black solid line are the equatorial India Ocean (EIO) and northern India (NI). The arrows in blue denote the Somali jet with abundant moisture. The arrows in red and black dotted line represent the anomalous wind induced by excessive equatorial light rain in U60 km relative to U3 km. The red '+' also represents stronger latent heating (LH) in U60 km.The doughnut charts show the precipitation frequency (PF, %) counted over the EIO in CMORPH (upper) and U60 km (bottom), respectively.

values of convection-permitting simulation in capturing local convection processes.

The Hadley-type anomalous circulation could be viewed as the up-scale effect of uncertainties from the convection parameterization. But the light rain is still overestimated in U3 km with explicitly resolved convection compared with observations. This residual issue may stem from multiple sources: (1) subgrid-scale variability of convective processes, particularly unresolved vertical mass flux[45]; (2) boundary layer processes including turbulence and boundary convergence; (3) limitations in microphysics parameterization. Consequently, the influence of excessive equatorial light rain in U60 km is actually stronger than that revealed by discrepancies between U60 km and U3 km. Meanwhile, even at a high resolution of 3 km, the iAMAS exhibits a persistent wet bias over the western EIO (Fig. S3a; Fig. S3d–f), which may impact the simulated monsoon circulation pattern. Encouragingly, the differences between the U3 km and U60 km experiments could mitigate the impact of this systematic bias in iAMAS on our findings. Moreover, experiment groups for different resolutions only cover one month and consist of five ensemble members due to the expensive computation cost of U3 km. Thus, our simulated datasets are not able to further reveal the up-scale effect of excessive light rain frequency on monsoon rainfall at intraseasonal and seasonal scale, which deserves being studied in the future. Meanwhile, analysis on fourteen atmospheric models from CMIP6 reveals that models with greater light rain biases over the EIO during summer (June, July and August) tend to also show reduced monsoon rainfall over northern India (Fig. S9). This provides evidence that our conclusions may be valid in the summer monsoon season.

The rainfall intensity spectrum is usually used to evaluate the performance of the GCMs and the "too much drizzle" bias common in current GCMs is usually analyzed in a global view[26–28,46]. However, few studies discussed its impacts on regional circulation, such as the South Asian summer monsoon. Our results suggest that the heterogeneous distribution of such bias does modulate the monsoon circulation and then introduces bias of mean summer rainfall. Precipitation frequency hence can be seen as an important indicator to check the simulating summer monsoon besides rainfall amount and temporal evolution. This relationship between inaccurate precipitation frequency and summer monsoon rainfall may amplify the drying trend of Indian subcontinent induced by the warming Indian Ocean draw from the CMIP GCMs[13]. Furthermore, the role of light rain is emphasized because in U60 km and some atmospheric models (Fig. S10), the overestimation of light rain significantly outweighs the underestimation of moderate to heavy rain, primarily causing the wet bias over the EIO and then anomalous Hadley-type circulation (Fig. 4b, d). However, some models overestimate light rain while severely underestimating heavier precipitation, leading to a counterbalance that make total precipitation match observations over the EIO. In such models, the impact of excessive light rain may be obscured[29,46,47]. Nevertheless, improving rainfall intensity spectrum through high-resolution modeling or parameterization optimization[29,48,49] remains crucial for alleviating the systematic bias of the South Asian summer monsoon rainfall and mitigation of extreme events.

Methods

Model and experimental design

The model used in this study is the integrated Atmospheric Model Across Scales (iAMAS) that was developed based on the atmospheric solver of the Model for Prediction Across Scales (MPAS)[50] and ported to the Sunway supercomputer heterogeneous architecture to enable accelerated computation and implementation of advanced physics and chemistry processes[35,37]. Three sets of ensemble experiments are conducted continuously for one month during the active monsoon. Two sets employ meshes at global quasi-uniform resolution of 60 km (U60 km) and 3 km (U3 km), respectively, while the third set employs a mesh at global variable resolution with regional refinement at 4 km (V4 km) over the Indian subcontinent. The amount of mesh cells in U60 km, V4 km and U3 km is 163842, 785410 and 65536002, respectively, with corresponding time-steps of 300, 20 and 15 s. The U3 km simulations, with 65536002 mesh cells, were performed utilizing 3,900,000 processor cores for approximately 5 days per simulation run[35]. These three sets of ensemble experiments are conducted with identical experimental settings except for the mesh resolution and time-step. Each set of ensemble experiment consists of five ensemble members, initialized using the reanalysis ERA5 at 18:00 UTC 9th, 21:00 UTC 9th, 00:00 UTC 10th, 03:00 UTC 10th, 06:00 UTC 10th of June in 2020 separately and finalized at 00:00 UTC 10th July in 2020. The SST, sea ice and soil state in ERA5 are used as surface forcing. More details of physical parameterization configuration can be found in our previous studies[35]. Ensemble mean for each set is analyzed to mitigate uncertainties inherent in individual members and obtain the most probable scenario. Meanwhile, only results spanning the period from June 11th to July 10th are considered, excluding results during the spin-up phase. Furthermore, an additional experiment with regional refinement at a resolution of 4 km (V4 kmTropics) within the equatorial region (40°E-100°E, 10°S-0°; Fig. S7) is conducted using identical experimental settings to validate the link between modeling biases over two rainbelts.

Latent heat and 300–500hPa mean temperature calculation

In addition, the column-averaged latent heat of each experiment is calculated as:

$$\overline{LH} = \frac{\int_{z_0}^{z_t} \rho H dz}{\int_{z_0}^{z_t} \rho dz} \quad (1)$$

The mean air temperature at mid-upper level is calculated as:

$$\bar{T} = \frac{\int_{500}^{300} T dp}{\int_{500}^{300} dp} \quad (2)$$

Observation and reanalysis datasets

Three precipitation datasets are used to evaluate simulations. The Global Precipitation Climatology Project (GPCP) dataset is obtained from surface and satellite measurements at a global spatial resolution of 1degree and temporal resolution of 1 day[51]. The CMORPH dataset is produced based on satellite precipitation estimation with Climate Prediction Center morphing Technique, at a scale of 0.25 degree and 30 min[52]. The IMERG final run of GPM is a product retrieved from multi-satellite precipitation products corrected by gauge data, at a resolution of 0.1 degree and 30 min[53]. Because analysis on U3 km whose number of meshes is particularly large, rainfall intensity is only graded into five levels: little rain (< 0.1 mm $3\ hr^{-1}$), light rain ($0.1 \sim 3$ mm $3\ hr^{-1}$), moderate rain ($3 \sim 10$ mm $3\ hr^{-1}$), heavy rain ($10 \sim 20$ mm $3\ hr^{-1}$) and torrential rain (> 20 mm $3\ hr^{-1}$). The precipitation amount and frequency in experiments are interpolated into the same grid as the CMORPH in order to calculate the difference in spatial distribution. The simulated monsoon circulation is validated using the reanalysis ERA5, which is provided hourly at a scale of 0.25 degree[54]. The vertical velocity in ERA5 is vertical pressure velocity ($Pa\ s^{-1}$) and turned into the same unit in the iAMAS ($m\ s^{-1}$) using the function 'omega_to_w' in the NCAR Command Language (NCL). The vertical and horizontal wind fields in experiments are interpolated into the same grid as ERA5 for comparison of monsoon circulation.

Data availability

The meshes used in this study can be download from http://aemol.ustc.edu.cn/product/list/ or contact chunzhao@ustc.edu.cn. The GPCP dataset is available at https://climatedataguide.ucar.edu/climate-data/gpcp-monthly-global-precipitation-climatology-project. The CMORPH dataset can be obtained from https://www.ncei.noaa.gov/products/climate-data-records/precipitation-cmorph. The IMERG final run of GPM dataset is available at https://gpm.nasa.gov/data/directory. The ERA5 data on pressure levels can be downloaded from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=form and the ERA5 data on single levels can be downloaded from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form. The code analyzing the precipitation and wind fields can be found in this GitHub repository (https://github.com/zeger314/U3km_IndianMonsoon.git).

Code availability

The code analyzing the precipitation and wind fields can be found in this GitHub repository (https://github.com/zeger314/U3km_IndianMonsoon.git).

Received: 25 July 2024; Accepted: 8 January 2025;
Published online: 18 January 2025

References

1. Van Bronkhorst & Bernice K, Disaster risk management in South Asia: regional overview (English). http://documents.worldbank.org/curated/en/648281468170977802/Disaster-risk-management-in-South-Asia-regional-overview (2012).
2. Turner, A. G. & Annamalai, H. Climate change and the South Asian summer monsoon. *Nat. Clim. Chang.* **2**, 587–595 (2012).
3. Huang, X. et al. South Asian summer monsoon projections constrained by the interdecadal Pacific oscillation. *Sci. Adv.* **6**, eaay6546 (2020).
4. Annamalai, H., Hamilton, K. & Sperber, K. R. The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations. *J. Clim.* **20**, 1071–1092 (2007).
5. Sperber, K. R. et al. The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. *Clim. Dyn.* **41**, 2711–2744 (2013).
6. Mitra, A. A Comparative Study on the Skill of CMIP6 Models to Preserve Daily Spatial Patterns of Monsoon Rainfall Over India. *Front. Clim*. **3**, (2021).
7. Mahendra, N. et al. Interdecadal modulation of interannual ENSO-Indian summer monsoon rainfall teleconnections in observations and CMIP6 models: Regional patterns. *Int. J. Climatol.* **41**, 2528–2552 (2021).
8. Boos, W. R. & Hurley, J. V. Thermodynamic Bias in the Multimodel Mean Boreal Summer Monsoon. *J. Clim.* **26**, 2279–2287 (2013).
9. Levine, R. C., Turner, A. G., Marathayil, D. & Martin, G. M. The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall. *Clim. Dyn.* **41**, 155–172 (2013).
10. He, L., Zhou, T. & Chen, X. South Asian summer rainfall from CMIP3 to CMIP6 models: biases and improvements. *Clim. Dyn.* **61**, 1049–1061 (2023).
11. Li, G., Xie, S.-P. & Du, Y. Monsoon-Induced Biases of Climate Models over the Tropical Indian Ocean. *J. Clim.* **28**, 3058–3072 (2015).
12. Ramesh, K. V. & Goswami, P. Assessing reliability of regional climate projections: the case of Indian monsoon. *Sci. Rep.* **4**, 4071 (2014).
13. Roxy, M. K. et al. Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient. *Nat. Commun.* **6**, 7423 (2015).
14. Hagos, S., Leung, L. R., Ashfaq, M. & Balaguru, K. South Asian monsoon precipitation in CMIP5: a link between inter-model spread and the representations of tropical convection. *Clim. Dyn.* **52**, 1049–1061 (2019).
15. Levine, R. C. & Turner, A. G. Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. *Clim. Dyn.* **38**, 2167–2190 (2012).
16. Sikka, D. R. & Gadgil, S. On the Maximum Cloud Zone and the ITCZ over Indian, Longitudes during the Southwest Monsoon. *Mon. Wea. Rev.* **108**, 1840–1853 (1980).
17. Gadgil, S. & Sajani, S. Monsoon precipitation in the AMIP runs. *Clim. Dyn.* **14**, 659–689 (1998).
18. Bollasina, M. & Nigam, S. Indian Ocean SST, evaporation, and precipitation during the South Asian summer monsoon in IPCC-AR4 coupled simulations. *Clim. Dyn.* **33**, 1017–1032 (2009).
19. Bush, S. J., Turner, A. G., Woolnough, S. J., Martin, G. M. & Klingaman, N. P. The effect of increased convective entrainment on Asian monsoon biases in the MetUM general circulation model. *Q.J.R. Meteorol. Soc.* **141**, 311–326 (2015).
20. Bollasina, M. A. & Ming, Y. The general circulation model precipitation bias over the southwestern equatorial Indian Ocean and its implications for simulating the South Asian monsoon. *Clim. Dyn.* **40**, 823–838 (2013).
21. Ashfaq, M., Rastogi, D., Mei, R., Touma, D. & Leung, L. R. Sources of errors in the simulation of south Asian summer monsoon in the CMIP5 GCMs. *Clim. Dyn.* **49**, 193–223 (2017).
22. Hanf, F. S. & Annamalai, H. Systematic Errors in South Asian Monsoon Precipitation: Process-Based Diagnostics and Sensitivity to Entrainment in NCAR Models. *J. Clim.* **33**, 2817–2840 (2020).
23. Meehl, G. A. et al. Monsoon Regimes in the CCSM3. *J. Clim.* **19**, 2482–2495 (2006).
24. Wang, Z., Li, G. & Yang, S. Origin of Indian summer monsoon rainfall biases in CMIP5 multimodel ensemble. *Clim. Dyn.* **51**, 755–768 (2018).
25. Rio, C., Del Genio, A. D. & Hourdin, F. Ongoing Breakthroughs in Convective Parameterization. *Curr. Clim. Chang. Rep.* **5**, 95–111 (2019).
26. Dai, A. Precipitation Characteristics in Eighteen Coupled Climate Models. *J. Clim.* **19**, 4605–4630 (2006).

27. Na, Y., Fu, Q. & Kodama, C. Precipitation Probability and Its Future Changes From a Global Cloud-Resolving Model and CMIP6 Simulations. *J. Geophys. Res. Atmos.* **125**, e2019JD031926 (2020).
28. Chen, D., Dai, A. & Hall, A. The Convective-To-Total Precipitation Ratio and the "Drizzling" Bias in Climate Models. *J. Geophys. Res. Atmos.* **126**, e2020JD034198 (2021).
29. Wang, Y., Zhang, G. J. & Craig, G. C. Stochastic convective parameterization improving the simulation of tropical precipitation variability in the NCAR CAM5. *Geophys. Res. Lett.* **43**, 6612–6619 (2016).
30. Arakawa, A. The Cumulus Parameterization Problem: Past, Present, and Future. *J. Clim.* **17**, 2493–2525 (2004).
31. Satoh, M. et al. Global Cloud-Resolving Models. *Curr. Clim. Chang. Rep.* **5**, 172–184 (2019).
32. Caldwell, P. M. et al. Convection-Permitting Simulations With the E3SM Global Atmosphere Model. *J. Adv. Model Earth Syst.* **13**, e2021MS002544 (2021).
33. McCarty, W., Errico, R. M. & Gelaro, R. Cloud Coverage in the Joint OSSE Nature Run. *Mon. Wea. Rev.* **140**, 1863–1871 (2012).
34. Miyakawa, T. et al. Convective Momentum Transport by Rainbands within a Madden–Julian Oscillation in a Global Nonhydrostatic Model with Explicit Deep Convective Processes. Part I: Methodology and General Results. *J. Adv. Sci.* **69**, 1317–1338 (2012).
35. Gu, J. et al. Establishing a non-hydrostatic global atmospheric modeling system at 3-km horizontal resolution with aerosol feedbacks on the Sunway supercomputer of China. *Science Bulletin* **67**, 1170–1181 (2022).
36. Hao, X. et al. swMPAS-A: Scaling MPAS-A to 39 Million Heterogeneous Cores on the New Generation Sunway Supercomputer. *IEEE Transactions on Parallel and Distributed Systems* **34**, 141–153 (2023).
37. Feng, J. et al. Simulating Atmospheric Dust With a Global Variable-Resolution Model: Model Description and Impacts of Mesh Refinement. *J. Adv. Model Earth Syst.* **15**, e2023MS003636 (2023).
38. Li, G. et al. Impacts of Topographic Complexity on Modeling Moisture Transport and Precipitation over the Tibetan Plateau in Summer. *Adv. Atmos. Sci.* **39**, 1151–1166 (2022).
39. Xie, S.-P., Xu, H., Saji, N. H., Wang, Y. & Liu, W. T. Role of Narrow Mountains in Large-Scale Organization of Asian Monsoon Convection. *J. Clim.* **19**, 3420–3429 (2006).
40. Chen, D. & Dai, A. Precipitation Characteristics in the Community Atmosphere Model and Their Dependence on Model Physics and Resolution. *J. Adv. Model Earth Syst.* **11**, 2352–2374 (2019).
41. Huang, Y., Wang, Y. & Cui, X. Differences between Convective and Stratiform Precipitation Budgets in a Torrential Rainfall Event. *Adv. Atmos. Sci.* **36**, 495–509 (2019).
42. Annamalai, H., Hafner, J., Sooraj, K. P. & Pillai, P. Global Warming Shifts the Monsoon Circulation, Drying South Asia. *J. Clim.* **26**, 2701–2718 (2013).
43. Skamarock, W. C. & Gassmann, A. Conservative transport schemes for spherical geodesic grids: high-order flux operators for ODE-based time integration. *Mon. Wea. Rev.* **139**, 2962–2975 (2011).
44. Zhao, C. et al. Exploring the impacts of physics and resolution on aqua-planet simulations from a nonhydrostatic global variable-resolution modeling framework. *J. Adv., Model Earth Syst.* **8**, 1751–1768 (2016).
45. Rauscher, S. A. et al. A multimodel intercomparison of resolution effects on precipitation: simulations and theory. *Clim. Dyn.* **47**, 2205–2218 (2016).
46. Wang, Y. et al. Disproportionate control on aerosol burden by light rain. *Nat. Geosci.* **14**, 72–76 (2021).
47. Wang, Y., Zhang, G. & Jiang, Y. Linking stochasticity of convection to large-scale vertical velocity to improve Indian summer monsoon simulation in the NCAR CAM5. *J. Clim.* **31**, 6985–7002 (2018).
48. Kang, I.-S., Yang, Y.-M. & Tao, W.-K. GCMs with implicit and explicit representation of cloud microphysics for simulation of extreme precipitation frequency. *Clim. Dyn.* **45**, 325–335 (2015).
49. Zhang, G. J. & Mu, M. Effects of modifications to the Zhang-McFarlane convection parameterization on the simulation of the tropical precipitation in the National Center for Atmospheric Research Community Climate Model, version 3. *J. Geophys. Res. Atmos*. **110** (2005).
50. Skamarock, W. C. et al. A Multiscale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering. *Mon. Wea. Rev.* **140**, 3090–3105 (2012).
51. Adler, R. F. et al. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeorol.* **4**, 1147–1167 (2003).
52. Joyce, R. J., Janowiak, J. E., Arkin, P. A. & Xie, P. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeorol.* **5**, 487–503 (2004).
53. Jin, D., Oreopoulos, L., Lee, D., Tan, J. & Cho, N. Cloud–Precipitation Hybrid Regimes and Their Projection onto IMERG Precipitation Data. *J. Appl. Meteor. Clim.* **60**, 733–748 (2021).
54. Hersbach, H. et al. The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.* **146**, 1999–2049 (2020).

Acknowledgements

This research was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB0500303), the Natural Science Foundation of Anhui (2208085UQ09, 2208085UQ02), the USTC Research Funds of the Double First-Class Initiative (YD2080002007, KY2080000114), the Science and Technology Innovation Project of Laoshan Laboratory (LSKJ202300305), and the National Key Scientific and Technological Infrastructure project "Earth System Numerical Simulation Facility" (EarthLab). The study used the computing resources from the Supercomputing Center of the University of Science and Technology of China (USTC) and the Qingdao Supercomputing and Big Data Center. The authors also thank the GPCP, CMORPH, GPM and ERA5 projects for providing these atmospheric data publicly, and the NCL developing group for data analysis and visualizations. We are grateful to anonymous reviewers for their valuable comments and constructive suggestions that helped improve the manuscript significantly.

Author contributions

C.Z. designed the experiments and study. J.G., J.F. and G.L. conducted the experiments. C.Z. and G.L. analyzed the results and wrote the draft. All of the co-authors helped revise and improve the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-025-00916-1.

Correspondence and requests for materials should be addressed to Chun Zhao or Hong An.