Supplement for

Design and trial implementation of a continental-scale, kilometre-resolution hourly precipitation analysis for Australia using satellite, radar, and gauges

Yuhang Zhang[1], Quan J. Wang[1, 4], Andrew J. Frost[2], Jayaram Pudashine[3], Blair Trewin[3], Carlos Velasco-Forero[3], Chun-Hsu Su[3], Vincent Villani[3]

[1]Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
[2]Bureau of Meteorology, Sydney, NSW 1300, Australia
[3]Bureau of Meteorology, Melbourne, VIC 3001, Australia
[4]College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

Correspondence to: Quan J. Wang (quan.wang@unimelb.edu.au) and Andrew J. Frost (andrew.frost@bom.gov.au)

Text S1. Bias correction effect in different subregions

Figure S1 presents the distribution of mean bias error (mm/h) for satellite and radar precipitation estimates across different climate zones: tropical, subtropical, and temperate. The results highlight the improvements achieved through bias correction for satellite precipitation and the alternative radar calibration approach used in this study. Before correction, satellite precipitation (S0) exhibits a broader bias distribution, particularly in the tropical region, indicating significant overestimations. After bias correction (S1), the distribution becomes more centred around zero, reflecting reduced bias and improved accuracy. Similarly, radar precipitation calibrated using the operational Rainfields system (R0) shows a systematic bias, largely due to the limited availability of gauges in the operational system, which is affected by data transfer constraints and the use of a smaller spatial and temporal window (150-km radius and a 15-minute time window in the operational system, compared to 300-km and 1-hour in this study). In contrast, the calibration approach used in this study (R1) yields a sharper distribution, suggesting better agreement with observed values.

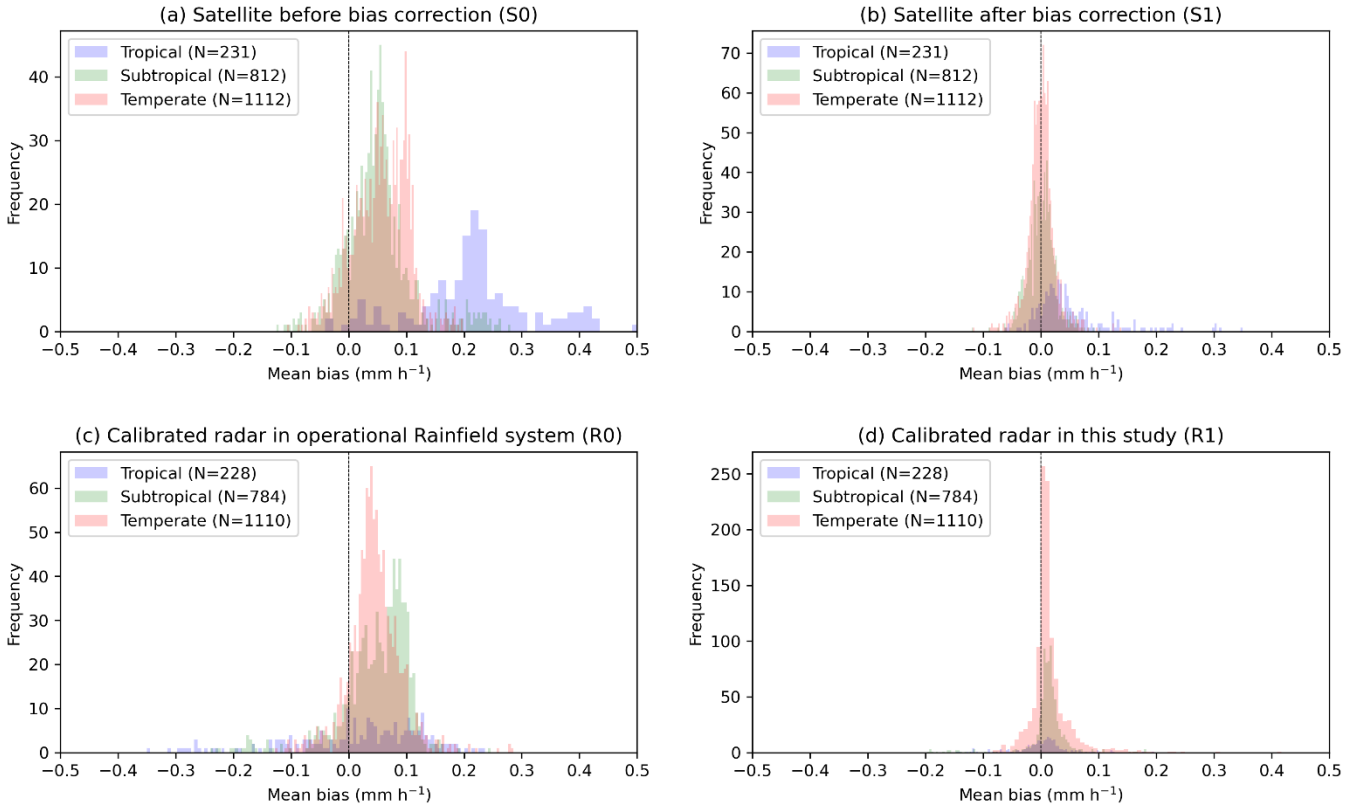

Figure S1. Histograms of mean bias error (mm/h) for different climate zones (tropical, subtropical, and temperate) across hourly gauges located in each subregion. Panels show (a) satellite precipitation before bias correction (S0), (b) radar precipitation calibrated using the operational Rainfields system (R0), (c) satellite precipitation after bias correction (S1), and (d) radar precipitation calibrated using the method in this study (R1). The Rainfields system calibration (R0) utilizes sub-daily gauges within a 150-km radius and a 15-minute time window, whereas the calibration in this study (R1) expands to a 300-km radius and a 1-hour time window. The analysis includes all available hourly gauges from 2022 to 2023. The vertical dashed line represents zero bias. Differences in sample sizes across climate zones are indicated in the legends. Note that the number of gauges in satellite plots differs from radar plots because gauges outside the radar coverage are excluded from the radar analysis.

Text S2. Further investigation of BR-SRG performance

S2.1. Stratification by subregions

To further examine performance differences across latitudes, we compared bias-corrected satellite data (S1), BR-SRG, and IDW-based interpolation for daily gauges across subregions located outside the radar coverage area, as illustrated in Fig. S2. This subregional analysis specifically focuses on areas beyond the radar range.

Although the bias-corrected satellite data (S1) shows some improvement, it remains the least accurate among the three methods. This is largely due to limitations in the satellite retrieval algorithms and the bias correction process, which relies on only a limited number of sub-daily gauges.

BR-SRG, which blends satellite data with gauge observations and radar information near the radar boundary, outperforms both S1 and the two-source blending product (BR-SG), as shown in Fig. 6. However, in regions beyond radar coverage, BR-SRG performs worse than IDW-based interpolation. This is illustrated in the BR-SRG vs. IDW comparison in Fig. 6, and the reasons for this outcome are discussed in the main text. The subregional results in Fig. S2 further reinforce these findings, revealing a consistent pattern across different climate zones.

In tropical regions, where gauge density is low, IDW-based interpolation exhibits greater variability but still achieves a lower median RMSE than BR-SRG. In subtropical regions, as gauge density increases, IDW becomes more stable and continues to outperform BR-SRG. This indicates that IDW benefits significantly from higher gauge density. In contrast, for BRAIN blending beyond the radar range, satellite data becomes the dominant input. Consequently, systematic or persistent biases in satellite estimates can lead to higher daily RMSE, despite relatively strong correlation values, as shown in Fig. 6b.

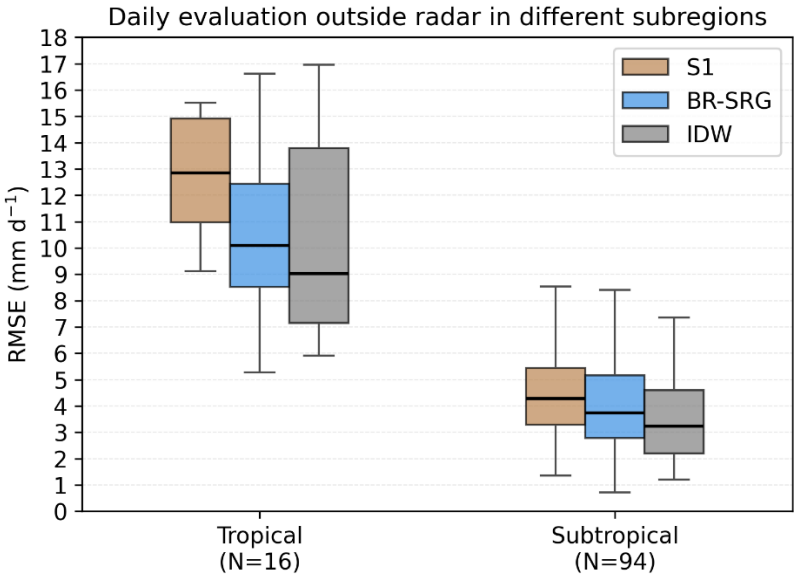

Figure S2. Boxplots of root mean square error (RMSE, mm/d) for three precipitation products—S1, BR-SRG, and IDW—outside the radar range (See Fig. 1b) in different climate zones. The comparison is shown for the tropical (N=16) and subtropical (N=94) regions. As there is only one gauge located in the temperate region outside the radar range, it is excluded from this analysis.

S2.2. Investigation of performance according to distance

To better understand how performance varies with input gauge sparsity, we further analysed the performance of BR-SRG, IDW-based interpolation, and individual data sources by grouping the daily validation results based on the average distance to the five nearest sub-daily gauges, as shown in Fig. S3. For each daily validation gauge, the mean distance to its five closest hourly gauges was calculated. This approach aligns with the configuration used in operational flood modelling, where IDW-based interpolation relies on the five nearest gauges as input sources.

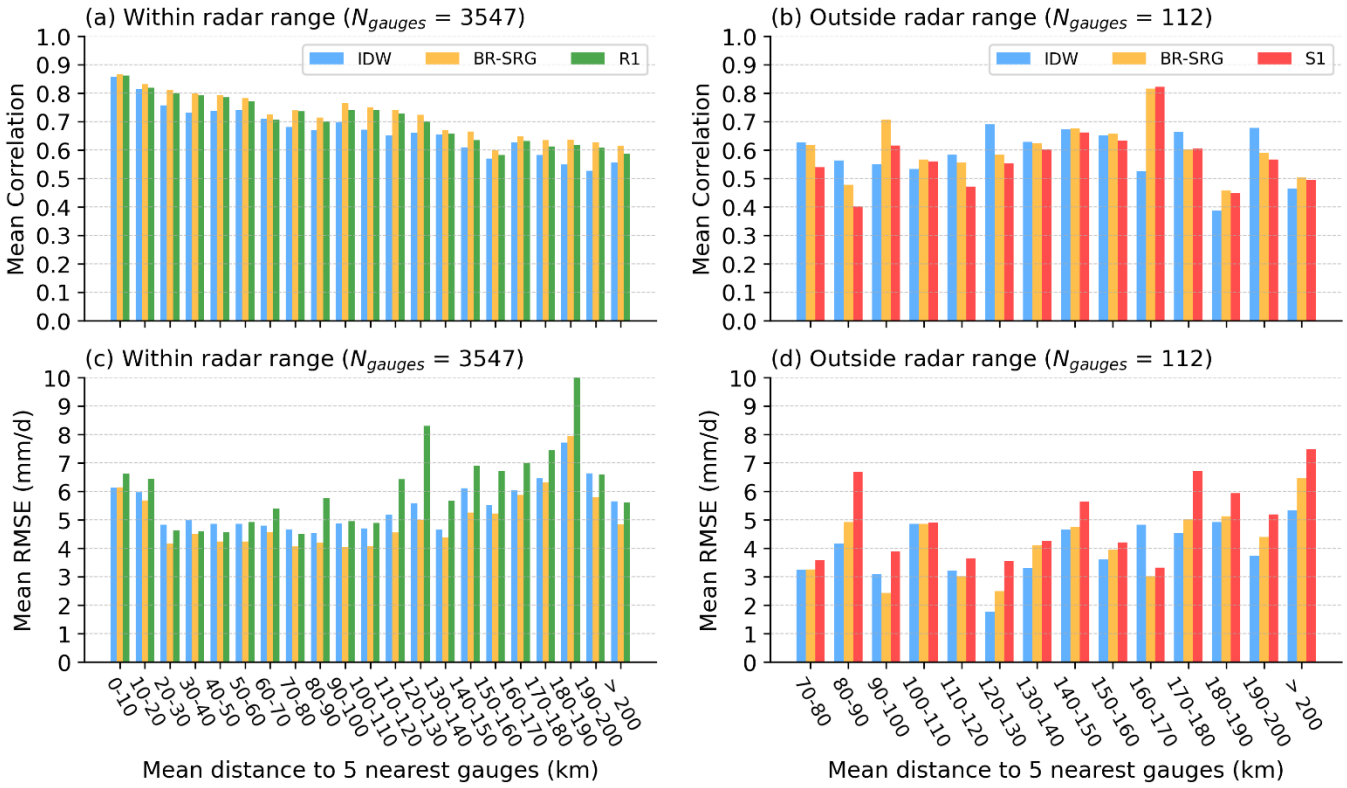

Figure S3. Validation of BR-SRG, IDW-based interpolation, and individual data sources within and outside radar range, categorized by mean distance to the five nearest hourly gauges. (a, c) Correlation and RMSE for daily gauges within radar range (see Fig. 1b), comparing bias-corrected radar (R1), IDW, and BR-SRG. (b, d) Correlation and RMSE for daily gauges outside radar range (see Fig. 1b), comparing bias-corrected satellite (S1), IDW, and BR-SRG.

The results are categorised into gauges located within and outside the radar range, reflecting the primary blending data sources. Radar data are used within the radar range due to their higher accuracy and proximity to the target location. In contrast, satellite data dominate beyond the radar range, as radar observations are distant from the target location and typically receive very little or no weight in the blending process. This distinction enables a clearer assessment of how gauge

sparsity influences the performance of different methods in areas where either radar or satellite data serve as the primary source of precipitation information.

For gauges within the radar range, correlation (Fig. S3a) decreases, while RMSE (Fig. S3c) increases with increasing distance to the nearest gauges. This pattern reflects the impact of distance-based weighting methods, where the influence of gauge data weakens with increasing distance, leading to higher errors. In terms of correlation, BR-SRG consistently outperforms both IDW and R1 (bias-corrected radar), highlighting the added value of integrating satellite data in the blending process. For RMSE, BR-SRG also demonstrates the best performance, whereas R1 performs the worst. These results indicate that the weighting scheme in BR-SRG effectively integrates satellite information alongside radar and gauge data while accounting for their respective error variances, leading to superior performance compared to IDW, which relies solely on gauge data without incorporating satellite or radar information.

For gauges outside the radar range, correlation (Fig. S3b) and RMSE (Fig. S3d) exhibit less variability across distance bins, likely due to the very limited number of gauges in these regions, as well as the uneven distribution of gauges within each bin. In terms of correlation (Fig. S3b), the performance of BR-SRG is highly dependent on the bias-corrected satellite (S1). Where S1 performs better than IDW, BR-SRG also outperforms IDW by benefiting from the inclusion of S1 in the blending process, along with gauge information and edge radar data. However, it is hypothesized that in cases where satellite bias persists, this may degrade the final BR-SRG product.

For RMSE (Fig. S3d), BR-SRG performs comparably to or slightly better than IDW when the mean gauge distance is less than 100 km, except in the 80–90 km bin, where satellite errors are notably high. Beyond 100 km, IDW consistently outperforms BR-SRG in most distance bins, except for the 160–170 km range, where variations in the number of gauges within specific bins may influence the results.

The performance differences between BR-SRG and IDW across distance bins arise from their underlying methodologies. IDW is a purely distance-based interpolation technique, which can incorporate the influence of gauges at distances greater than 200 km in this study. In contrast, BR-SRG employs a more complex weighting scheme that accounts for errors and spatial correlation alongside distance, with correlations set to zero beyond 200 km in this study. Additionally, at larger distances (e.g., beyond 100 km), the fitted error correlations tend to stabilize at a near-constant value (e.g., ~0.1 in Fig. 3b). Consequently, BR-SRG increasingly relies on satellite data as background information, with very limited contributions from gauge observations, as increments at gauge locations are considered weakly correlated with the target grid cell. This highlights the need to further improve the quality of satellite data to enhance background information within the BR-SRG framework.

Overall, these results demonstrate that BR-SRG consistently outperforms IDW within the radar range and in regions with shorter distances to gauges outside the radar range. IDW only shows a slight advantage at larger distances (e.g., beyond 100 km) outside the radar range.

Text S3. Extreme rainfall records in two cases

In the main text, we compared the blended product (BR-SRG) and the daily operational product (AGCD) in terms of their spatial patterns and their ability to represent extreme rainfall events (Fig. 7). Here, we further explore their differences by mapping gauge observation records (Fig. S4) and analysing the top 20 rainfall records along with their corresponding gauge information for both cases (Tables S1 and S2).

For the first case on December 24, 2022, the highest recorded rainfall occurred in the Northern Territory. Both sub-daily and daily gauges captured this extreme rainfall event. However, due to the low density of sub-daily gauges, only one station recorded a high rainfall value of 168.2 mm, while nearby stations, being farther from the rainfall centre, failed to capture similarly high values (Fig. S4 and Table S1). The second-highest record among sub-daily gauges was only 51.4 mm. This limitation reduces the ability of the BR-SRG blended product, which is based on sub-daily gauges, to accurately depict the highest rainfall centre. In contrast, the higher density of daily gauges allowed multiple stations to record significant rainfall values (208.0 mm, 184.6 mm, 141.4 mm, etc.), enhancing the AGCD product's ability to represent extreme rainfall through daily gauge interpolation.

For the second case on December 14, 2023, the highest recorded rainfall occurred in the Cape Tribulation area of Queensland. Both sub-daily and daily gauges captured this extreme event. Unlike the first case, multiple hourly gauges recorded high rainfall values (Fig. S4 and Table S2), helping the BR-SRG product better capture the maximum rainfall. However, a similar issue remains—due to the sparser distribution of hourly gauges, the spatial extent of the rainfall maximum in BR-SRG is smaller compared to that derived from daily gauge interpolation in the AGCD product (Fig. 7).

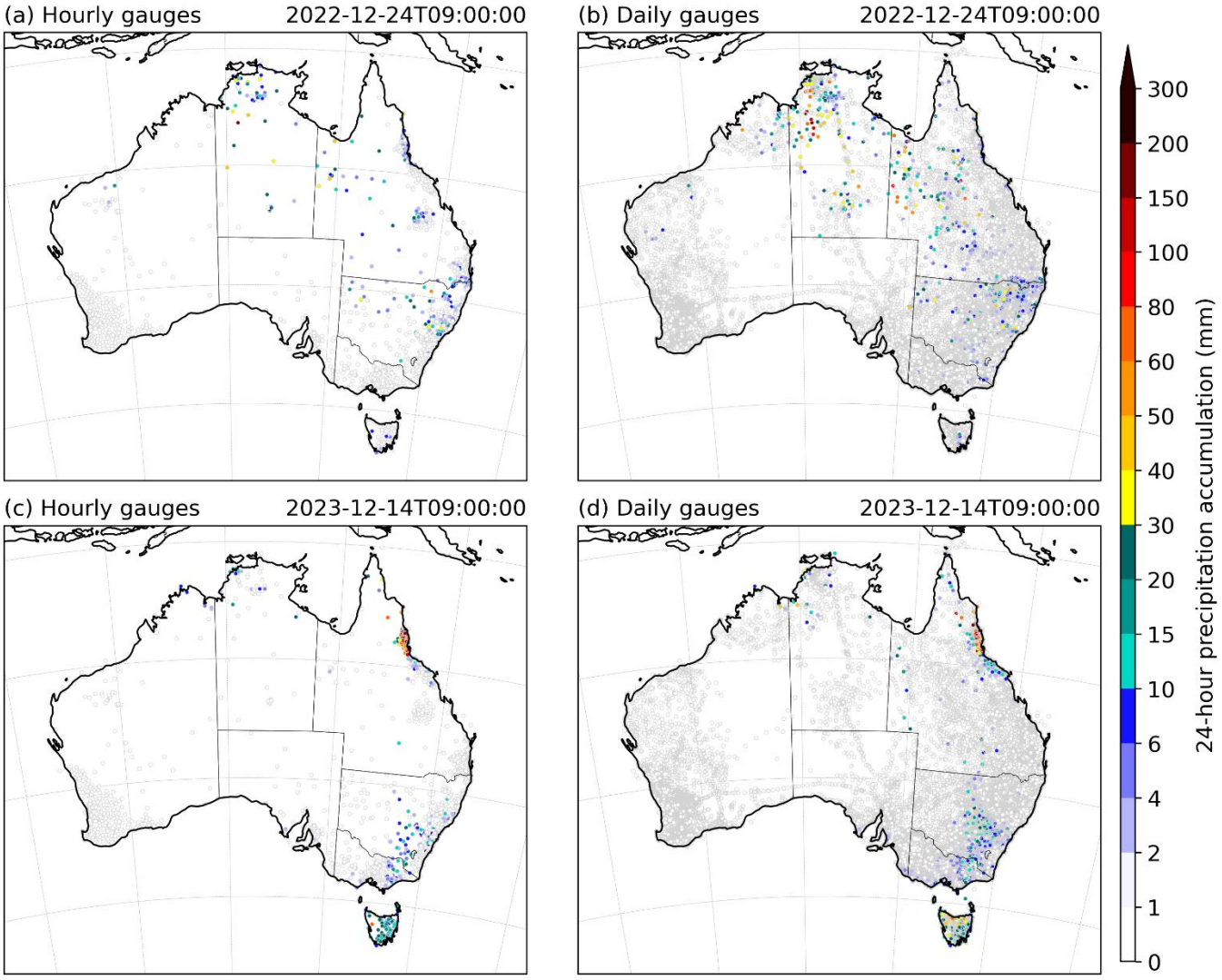

Figure S4. Spatial distribution of 24-hour precipitation accumulation (mm) from gauge observations for two extreme rainfall events: (a, b) 24 December 2022 and (c, d) 14 December 2023. Panels (a) and (c) show hourly gauge observations, while panels (b) and (d) display daily gauge observations. The colour scale represents the recorded precipitation amount, highlighting differences in spatial coverage and station density between hourly and daily gauge networks.

Table S1. Top 20 recorded rainfall values from sub-daily and daily gauges on 24 December 2022. For sub-daily gauges, the daily values are aggregated from hourly values.

Rank	Daily gauge value	Daily longitude	Daily latitude	Sub-daily gauge value	Sub-daily longitude	Sub-daily latitude

	(mm/d)			(mm/d)		
1	208.0	130.454	-15.599	168.2	131.015	-16.403
2	184.6	131.015	-16.403	51.4	149.847	-29.490
3	141.4	130.573	-15.440	49.8	130.015	-20.182
4	137.0	131.160	-16.853	48.2	139.536	-17.748
5	114.5	130.956	-16.118	48.2	139.296	-20.936
6	93.0	131.117	-17.387	47.0	133.881	-23.636
7	92.4	138.367	-21.598	46.0	150.082	-32.520
8	92.0	144.485	-20.739	44.6	133.878	-23.579
9	90.0	144.154	-19.254	44.2	131.918	-15.744
10	78.0	139.283	-18.295	42.5	133.753	-23.575
11	71.0	138.251	-17.338	38.0	138.367	-21.598
12	70.0	139.536	-17.748	38.0	132.894	-12.659
13	69.8	145.008	-22.207	36.8	143.074	-30.852
14	69.6	145.380	-16.459	35.8	130.854	-12.793
15	68.2	130.879	-13.668	35.2	134.183	-19.642
16	68.0	132.526	-12.904	34.0	130.573	-15.440
17	66.8	151.337	-32.791	33.4	145.311	-14.967
18	65.0	130.882	-14.189	33.3	133.863	-23.650
19	64.0	138.835	-22.520	32.1	133.945	-23.543
20	62.0	130.493	-13.744	32.0	131.866	-12.914

135

Table S2. Top 20 recorded rainfall values from sub-daily and daily gauges on 14 December 2023. For sub-daily gauges, the daily values are aggregated from hourly values.

Rank	Daily gauge value (mm/d)	Daily longitude	Daily latitude	Sub-daily gauge value (mm/d)	Sub-daily longitude	Sub-daily latitude
1	572.4	145.458	-16.097	479.2	145.612	-16.800
2	461.8	145.351	-16.391	456.5	145.511	-16.732
3	450.0	145.638	-16.818	378.4	145.675	-16.979
4	429.0	145.375	-16.473	284.5	145.546	-16.891
5	340.6	145.568	-16.658	280.1	145.764	-17.134
6	326.0	145.675	-16.979	275.5	145.686	-16.875
7	226.0	145.922	-17.344	266.8	145.669	-17.197
8	220.0	145.731	-17.469	221.1	145.742	-16.862
9	207.0	145.716	-17.545	207.7	145.716	-17.545
10	186.0	145.791	-17.093	203.8	145.747	-16.946
11	178.6	145.746	-16.874	193.4	145.746	-16.874

12	176.0	145.133	-16.532	179.8	145.789	-17.101
13	171.0	145.731	-17.469	178.7	145.724	-17.185
14	165.8	145.747	-16.946	178.3	145.597	-17.625
15	162.0	145.600	-17.588	177.5	145.601	-17.833
16	162.0	145.878	-17.164	174.7	145.705	-17.831
17	160.0	145.600	-17.588	167.4	145.600	-17.588
18	142.0	145.548	-17.169	166.6	145.443	-16.842
19	139.0	145.529	-17.591	161.9	145.529	-17.591
20	133.8	146.073	-18.586	146.4	145.565	-17.777

Text S4. Spatial performance of individual sources and blending

Here, we extend our analysis from Fig. 5 and Fig. 6 by presenting spatial performance maps of the individual data sources (bias corrected satellite, bias corrected radar, gauge-based IDW interpolation) and the three-source (BR-SRG) blended product. These maps provide a more detailed view of how each method performs across the continent.

S4.1. Spatial maps for hourly gauges

Figures S5 to S7 show the cross-validation evaluation results based on hourly gauges. Specifically, Fig. S5 presents the spatial distribution of correlation, Fig. S6 shows RMSE, and Fig. S7 displays bias.

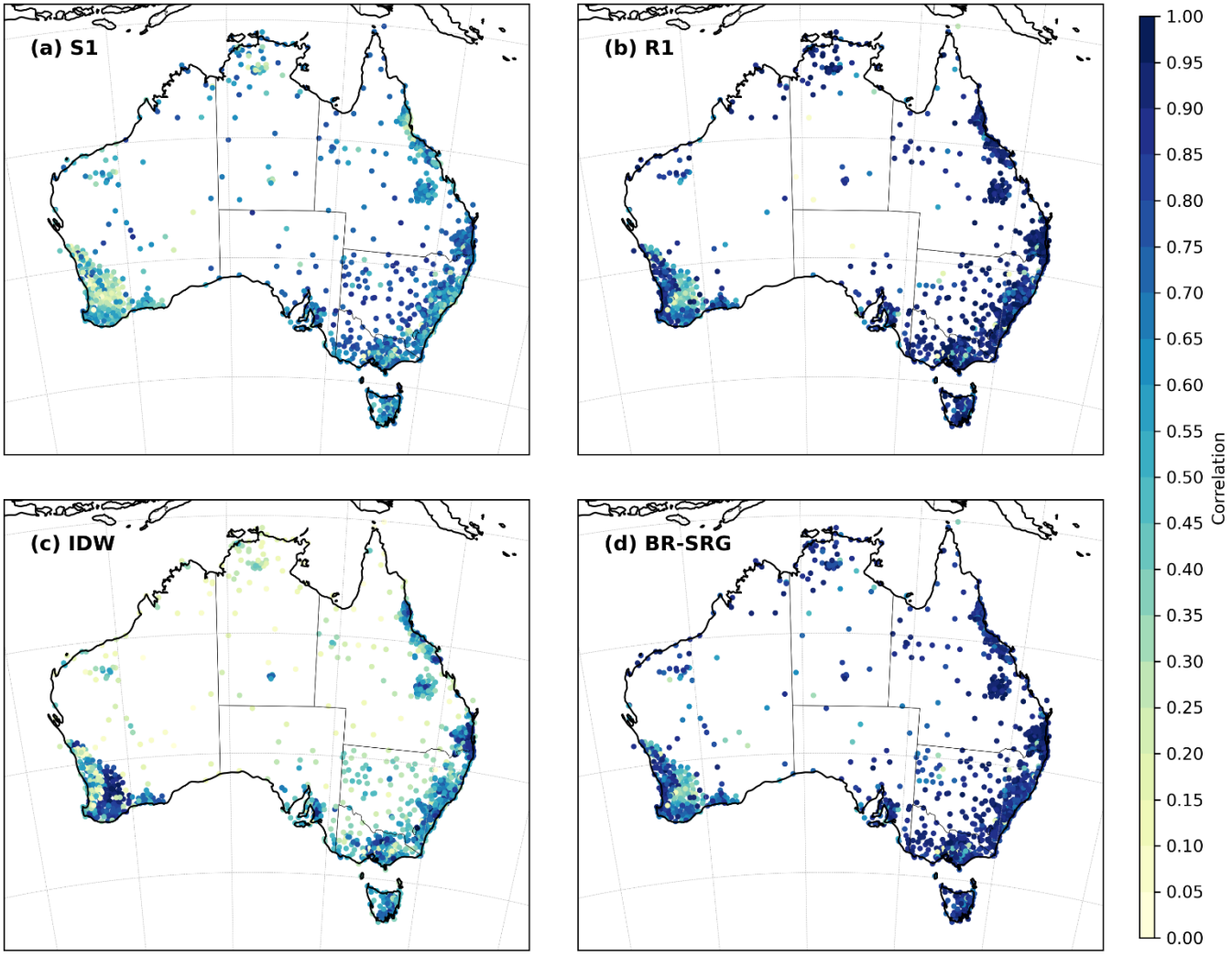

Figure S5. Spatial distribution of correlation between estimated and observed hourly precipitation at gauge locations, based on 20-fold cross-validation using hourly gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that S1 and R1 use all available hourly gauges for bias correction without cross-validation, while IDW and BR-SRG are evaluated using cross-validation. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded.

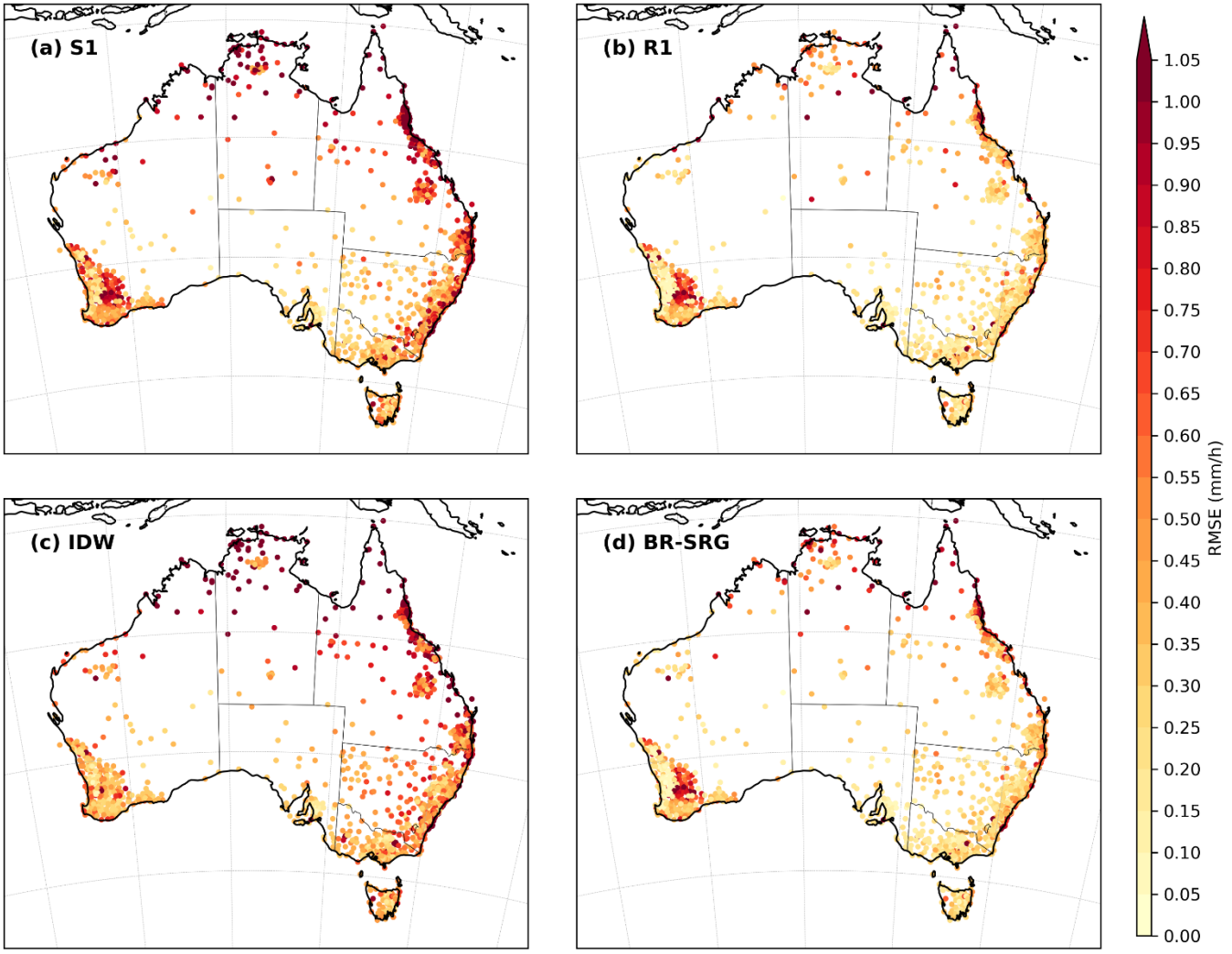

Figure S6. Spatial distribution of RMSE between estimated and observed hourly precipitation at gauge locations, based on 20-fold cross-validation using hourly gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that S1 and R1 use all available hourly gauges for bias correction without cross-validation, while IDW and BR-SRG are evaluated using cross-validation. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded.

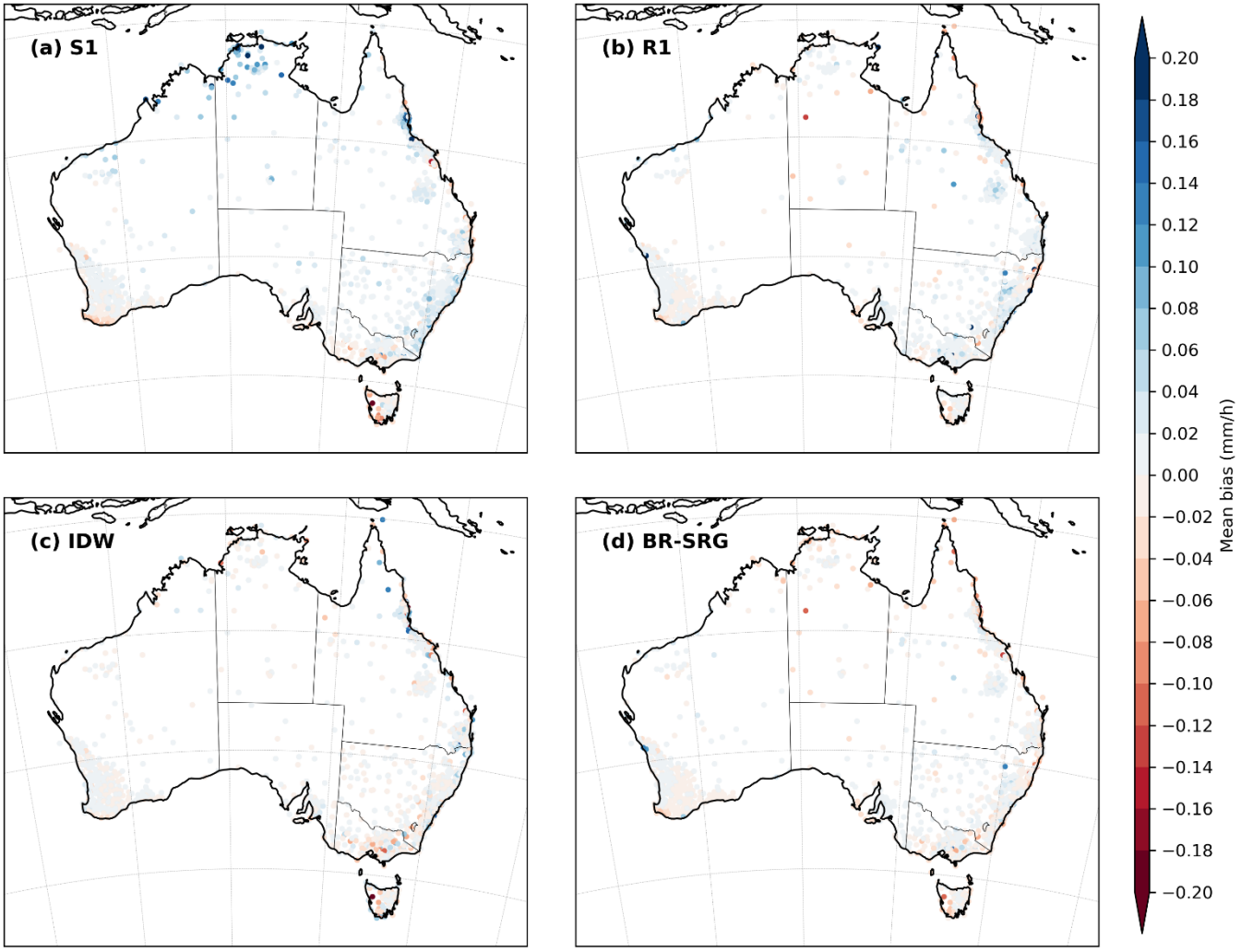

Figure S7. Spatial distribution of mean bias between estimated and observed hourly precipitation at gauge locations, based on 20-fold cross-validation using hourly gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that S1 and R1 use all available hourly gauges for bias correction without cross-validation, while IDW and BR-SRG are evaluated using cross-validation. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded.

S4.2. Spatial maps for daily gauges

Figures S8 to S10 provide corresponding evaluations using independent daily gauges. Fig. S8 presents correlation, Fig. S9 shows RMSE, and Fig. S10 displays bias. This set of figures offers additional insight into performance at daily resolution using an independent dataset across Australia.

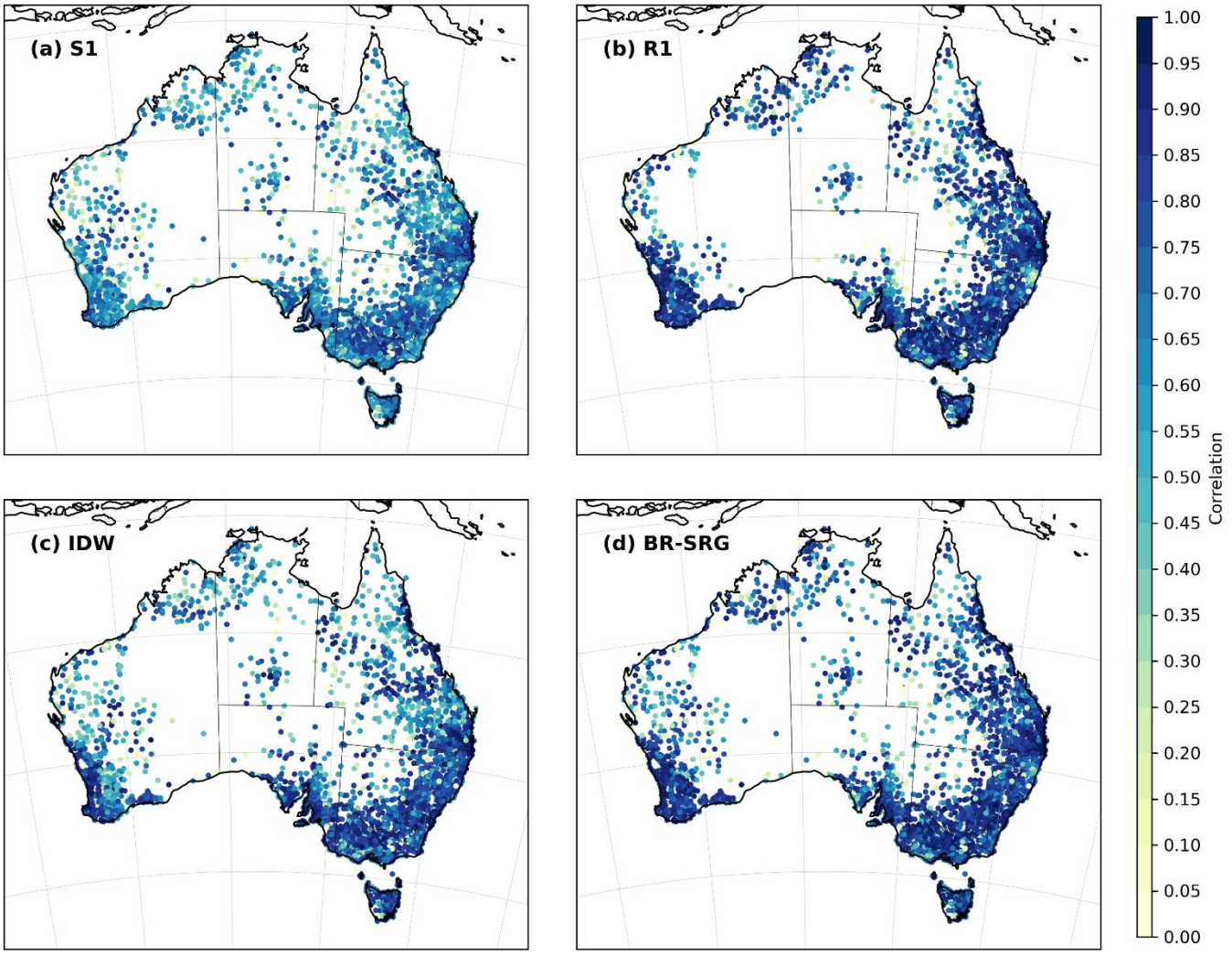

Figure S8. Spatial distribution of correlation between estimated and observed daily precipitation at gauge locations, evaluated using independent daily gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that all daily gauges are fully independent and have no overlap with the hourly gauges used in previous evaluations. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded.

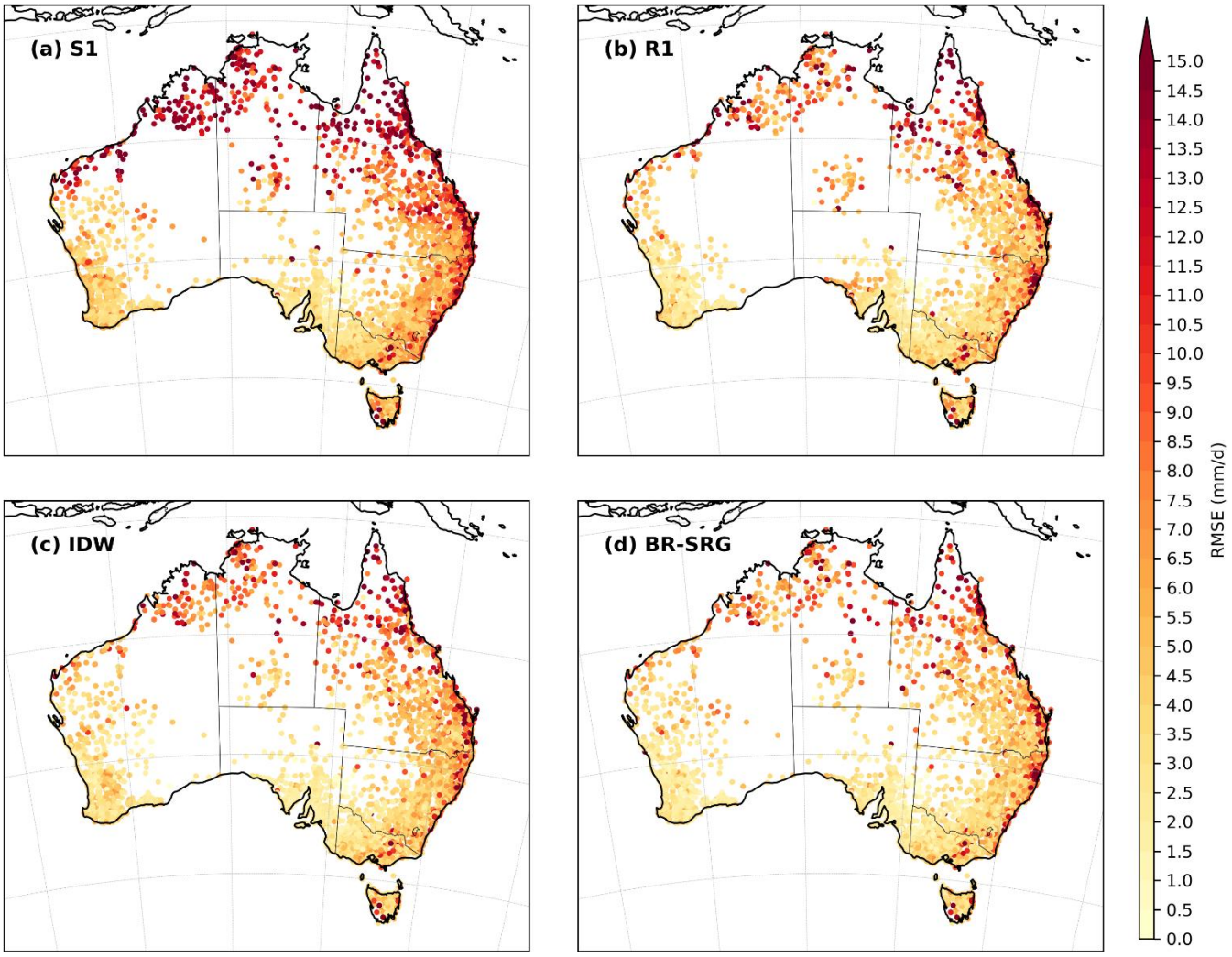

Figure S9. Spatial distribution of RMSE between estimated and observed daily precipitation at gauge locations, evaluated using independent daily gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that all daily gauges are fully independent and have no overlap with the hourly gauges used in previous evaluations. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded.

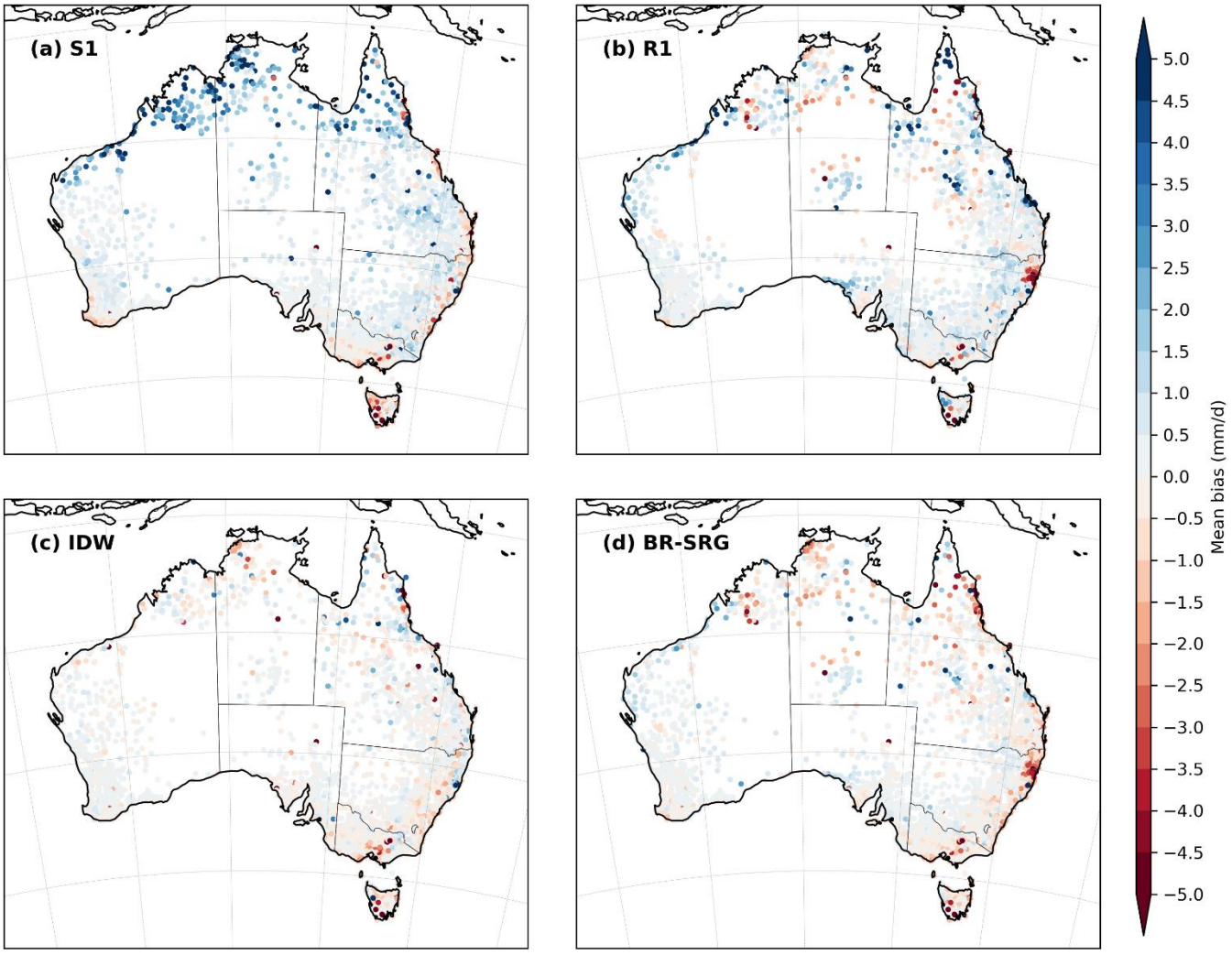

Figure S10. Spatial distribution of mean bias between estimated and observed daily precipitation at gauge locations, evaluated using independent daily gauges. Results are shown for (a) bias-corrected satellite (S1), (b) bias-corrected radar (R1), (c) gauge-based IDW interpolation, and (d) three-source blended product (BR-SRG). Note that all daily gauges are fully independent and have no overlap with the hourly gauges used in previous evaluations. The number of gauges shown in the radar plot differs from the others because gauges located outside the radar coverage area are excluded. The outliers on the maps (e.g., R1 and BR-SRG) are attributed to local radar calibration issues, which are currently being improved.

Text S5. Spatial performance of blended product on daily gauges

Figure S11 provides an additional independent evaluation of the spatial performance of the blended product (BR-SRG), using correlation and RMSE metrics calculated from daily gauge observations across Australia. This analysis complements the hourly gauge evaluation and offers an independent perspective on the performance of the product at the daily scale. The spatial distribution of correlation (Fig. S11a) shows widespread high values, particularly in eastern and southeastern Australia, where gauge density is higher. The longitudinal variation of correlation (Fig. S11b) reflects generally consistent performance, with peaks in areas with better observational coverage. The latitudinal variation of correlation (Fig. S11c) remains stable across most regions, with a slight increase from low to high latitudes. The spatial distribution of RMSE (Fig. S11d) indicates better performance in southeastern regions and other areas with dense gauge networks. The longitudinal variation of RMSE (Fig. S11e) follows a similar trend. The latitudinal variation of RMSE (Fig. S11f) mirrors the correlation trend, with smaller errors in temperate regions and larger errors in tropical regions. Overall, Fig. S11 demonstrates that BR-SRG performs well at the daily scale, especially in regions with dense observational coverage, further supporting the reliability of the blended product. These patterns are largely shaped by precipitation magnitude, as larger precipitation amounts in lower latitudes introduce to greater variability and error, resulting in lower correlations and larger RMSE values.

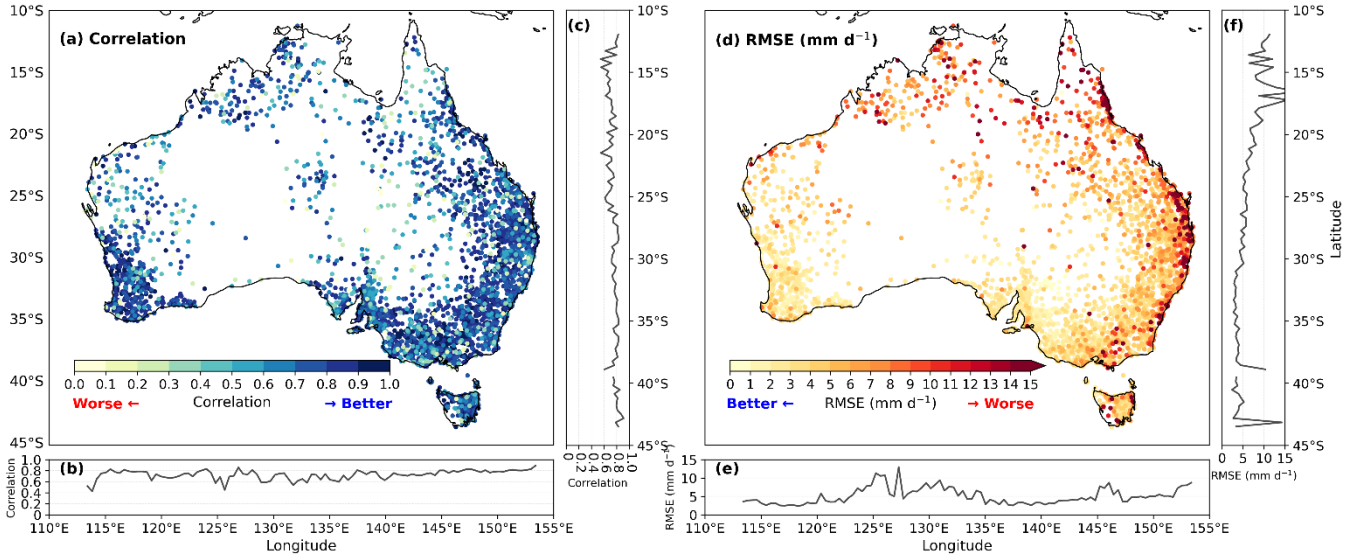

Figure S11. Spatial performance of the blended product (BR-SRG) on daily gauges across Australia. (a) Spatial distribution of correlation between BR-SRG and daily gauge observations, with higher values (blue) indicating better performance. (d) Spatial distribution of RMSE (mm/day), with lower values (red) indicating better performance. (b, e) Longitudinal variations of (b) correlation and (e) RMSE, averaged across gauges. (c, f) Latitudinal variations of (c) correlation and (f) RMSE, averaged across gauges.

Text S6. Temporal performance of blended product across months

Figure S12 evaluates the monthly performance of the blended product (BR-SRG) to provide a clear sense of its reliability throughout the year. This analysis, based on correlation and RMSE metrics using both hourly and daily gauges, assesses whether the blending process maintains consistent quality across different seasons. Correlation results for hourly gauges (Fig. S12a) show stable and strong performance across all months, with only minor variations. Daily gauge results (Fig. S12b) follow a similar pattern, with slightly lower median values but equally consistent trends, confirming the robustness of the blending process under various precipitation regimes. RMSE results for hourly gauges (Fig. S12c) reveal slight increases in error during summer months (e.g., January and February), when convective rainfall is more intense and variable, and lower errors in cooler months (e.g., June to August). Daily gauge RMSE (Fig. S12d) shows a similar seasonal pattern.

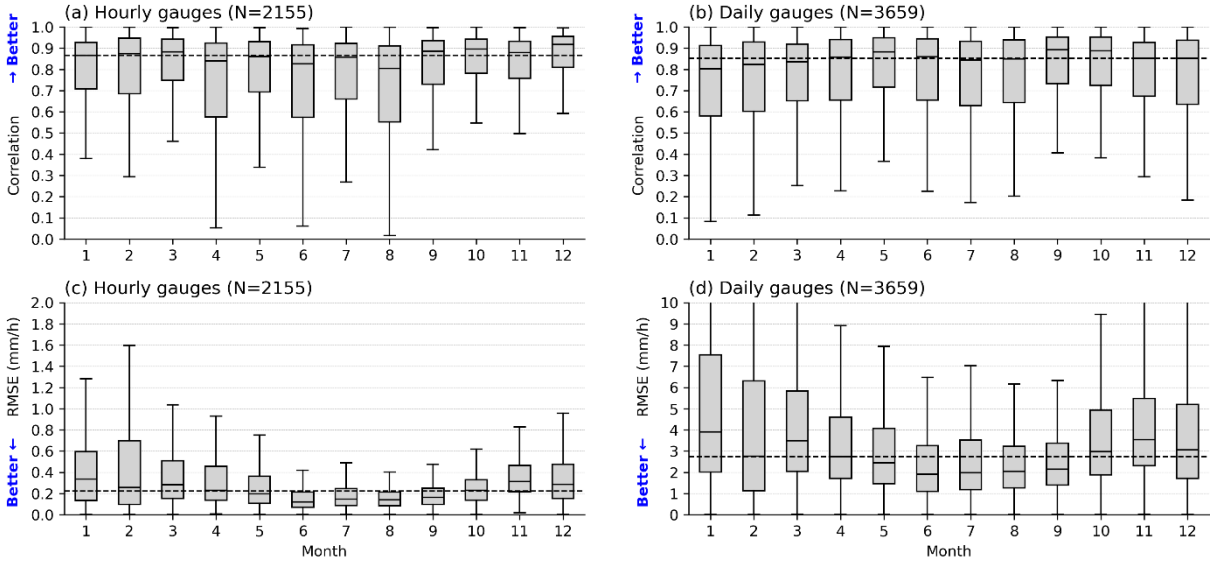

Figure S12. Monthly evaluation of the blended product (BR-SRG) performance using (a, c) hourly gauges and (b, d) daily gauges. The dashed lines in each panel represent the average value across all months, providing a visual reference for overall consistency in performance. Higher correlation and lower RMSE indicate better agreement with gauge observations.

Text S7. Additional independent hourly dataset and evaluation

An additional hourly gauge dataset, comprising approximately 2000 stations, was collated after completion of the main analyses and used for supplementary evaluation. This data is independent of the gauge observations used in the bias correction, blending and cross-validation experiments. The supplementary evaluation provides an external consistency check on the main results and supports the robustness of the overall performance of the blended BRAIN product.

Figure S13 shows the spatial distribution of the additional independent hourly rain gauge stations used for supplementary evaluation. These gauges are fully independent of the hourly gauge network (Fig. 1a) used in the blending process and were not blended or otherwise involved in product generation. Compared with the primary hourly gauge network shown in Fig. 1a, this supplementary network exhibits a different spatial configuration, with variations in station density and regional coverage. While many stations remain located within radar coverage, a subset lies outside radar range, enabling evaluation of blended product performance both with and without direct radar influence. The stations are mainly concentrated along coastal and populated regions, particularly in eastern and south-eastern Australia, with sparse coverage across the interior.

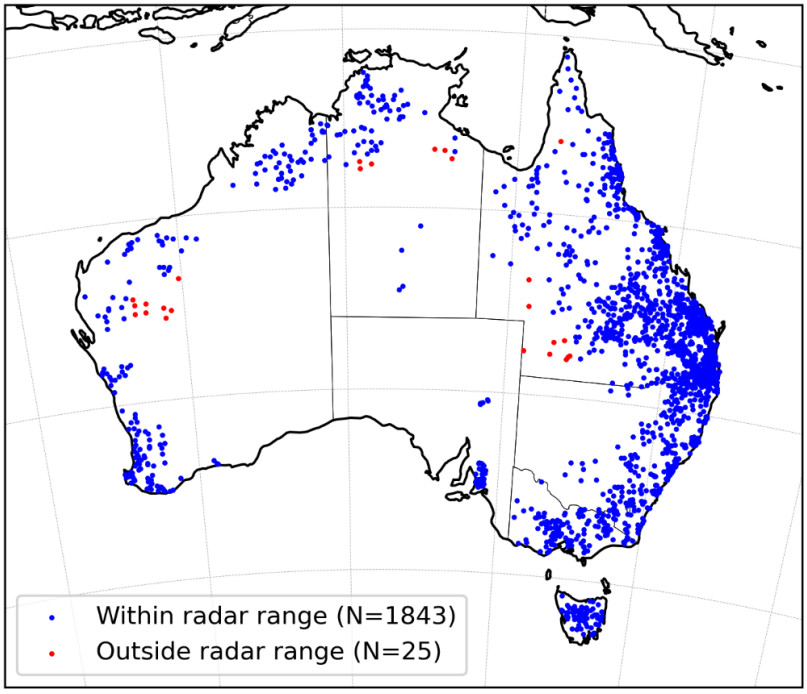

Figure S13. Spatial distribution of additional hourly rain gauge stations used for supplementary evaluation. These gauges are independent of the hourly gauge network used in the blending process and were not blended in the product generation. Stations are classified according to whether they are located within (blue) or outside (red) radar coverage.

Figure S14 presents an independent hourly evaluation using a gauge dataset that was not involved in the blending process, thereby providing an additional and complementary perspective to the main analysis. While the overall performance

hierarchy remains consistent, several notable differences emerge that further strengthen the conclusions. Within radar-covered regions, the multi-source blended analysis (BR-SRG) exhibits lower RMSE than the bias-corrected radar product (R1), indicating that the integration of satellite, radar, and gauge information yields more accurate rainfall magnitude estimates than radar-dominant approaches alone. Outside radar coverage, BR-SRG also achieves lower RMSE than the operational IDW method, demonstrating improved performance even under sparse observational conditions. As in the main analysis, BR-SRG consistently outperforms the satellite-only products (S0 and S1) across all metrics, while its performance relative to BR-SG remains comparable. The use of an additional independent hourly gauge network for evaluation highlights that these results are not specific to a single observational configuration, but instead reflect the general advantage of the multi-source BRAIN framework over individual data sources. Taken together, the complementary evidence from this supplementary evaluation further supports the conclusion that BRAIN provides the most robust and consistently performing precipitation estimates across differing data availability scenarios.

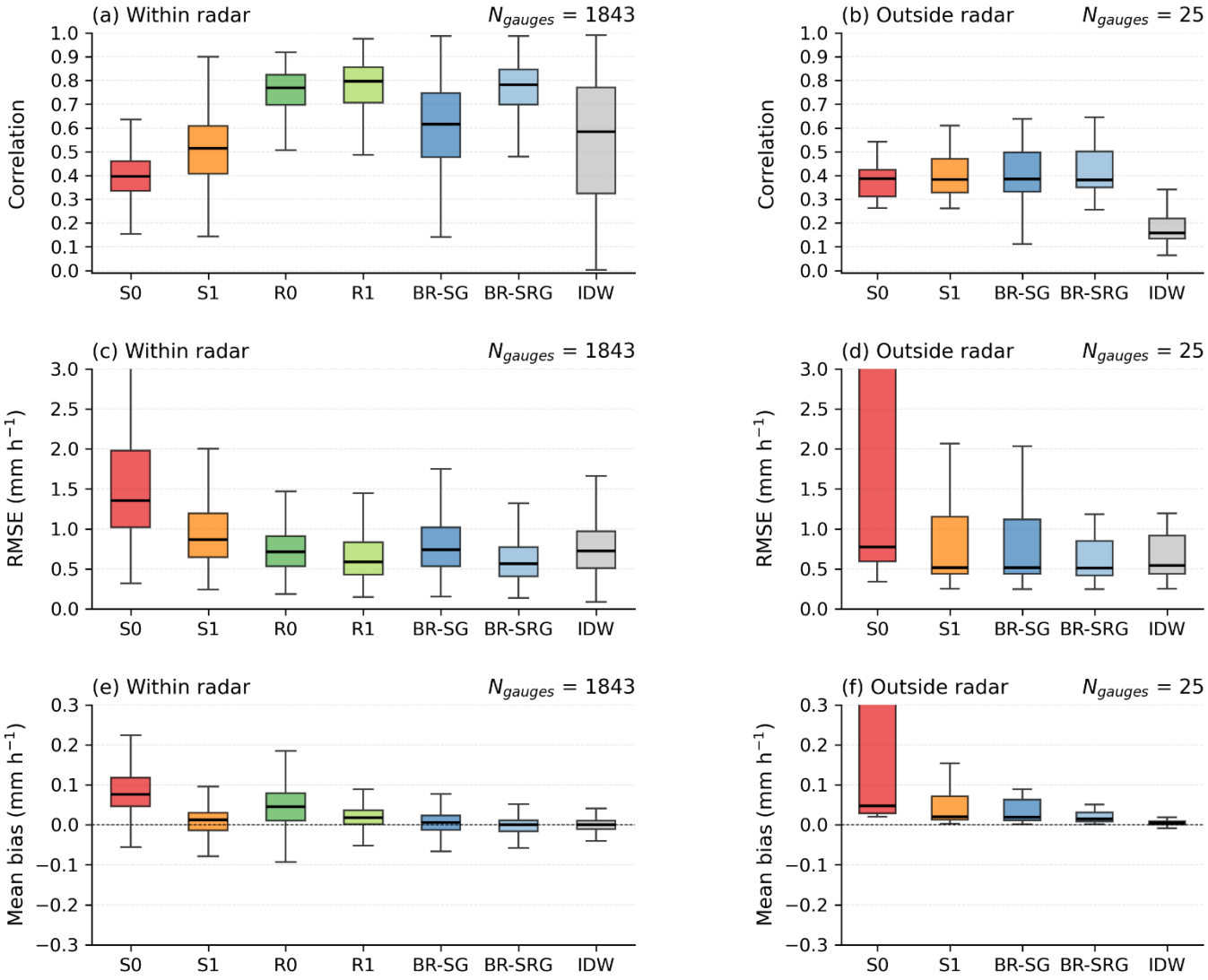

Figure S14. Evaluation of hourly precipitation estimates using the independent hourly rain gauge dataset shown in Fig. S13. Boxplots show the distribution of correlation (a–b), RMSE (c–d), and mean bias (e–f) for gauges located within radar coverage (left column) and outside radar coverage (right column). Higher correlation, lower RMSE, and bias closer to zero indicate better performance. Products include raw satellite (S0), bias-corrected satellite (S1), original radar (R0), bias-corrected radar (R1), two-source blend (BR-SG; BRAIN with satellite and gauge), three-source blend (BR-SRG; BRAIN with satellite radar and gauge), and gauge-based interpolation (IDW).