Supplementary Material for "SUPER v2: A 3-Hourly Global Precipitation Dataset Optimized for Sparse Data Challenges"

S1. Datasets involved in the construction and verification of SUPER v2.

IMERG V07 Final Run: The Integrated Multi-satellitE Retrievals for GPM (IMERG V07 Final Run) is a NASA product delivering high-resolution precipitation estimates (0.1°, 30-min) from 2000 to the present. Operating under the joint NASA-JAXA Global Precipitation Measurement (GPM) framework, the algorithm utilizes the GPM Core Observatory as a calibration anchor to harmonize precipitation retrievals from a diverse international satellite constellation. IMERG Early, Late and Final Run data are made available, with the Final Run incorporating all available data, including adjusted gauge information, to provide the most refined precipitation estimates (Huffman, 2023). The IMERG dataset is available at: https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_07.

CMORPH CDR: The NOAA Climate Data Record (CDR) of CMORPH (CPC Morphing Technique) generates daily, high-resolution precipitation fields (8 km × 8 km) covering the latitude band 60° S–60° N from 1998 to the present. It is generated by integrating passive microwave (PMW) measurements to construct satellite-based global precipitation fields. To mitigate biases, the raw estimates are rigorously corrected against CPC daily gauge analysis over land and the GPCP pentad product over the ocean (Joyce et al., 2004; Xie et al., 2017). The CMORPH dataset is available at: https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/.

PERSIANN CDR: NOAA Climate Data Record of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR) is a daily precipitation dataset developed by NOAA, offering a spatial resolution of 0.25° across 60° S–60° N from 1983 to the present. The algorithm applies artificial neural networks to GridSat-B1 infrared data to estimate rainfall. Crucially, the dataset is adjusted using the Global Precipitation Climatology Project (GPCP) monthly product to maintain consistency at the monthly scale, making it highly suitable for long-term hydro-climatic studies (Ashouri et al., 2015). The PERSIANN dataset is available at: https://www.ncei.noaa.gov/data/precipitation-persiann/access/.

CHIRPS-2.0: Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS-2.0) is a quasi-global (50° S–50° N) high-resolution precipitation dataset spanning from 1981 to the present. It employs a unique fusion approach that integrates satellite-based Cold Cloud Duration imagery with 0.05° climatology (CHPclim) and in-situ station data (Funk et al., 2015). The CHIRPS dataset is available at: https://data.chc.ucsb.edu/products/CHIRPS-2.0/.

CPC: The CPC Unified Gauge-Based Analysis of Global Daily Precipitation dataset, developed by NOAA's Climate Prediction Center, offers daily land-surface precipitation fields at a 0.5° resolution, covering the period from 1979 to the present. CPC utilizes the optimal interpolation (OI) technique, which integrates multiple data sources to produce reliable gridded precipitation fields (Chen et al., 2008; Xie et al., 2007). The CPC dataset is available at: https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html.

ERA5: The global reanalyzed precipitation product from the fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis system (ERA5) provides hourly global precipitation estimates at a spatial resolution of 0.25°. Based on the IFS Cy41r2 system, ERA5 employs an advanced data assimilation framework along with improved model physics and dynamics, leading to significantly enhanced precipitation estimates with higher temporal and spatial resolution (Hersbach et al., 2020). The ERA5 dataset is available at: https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview.

MSWEP V2: Multi-Source Weighted-Ensemble Precipitation (MSWEP V2) is included in this study as a representative gauge-corrected merged product for comparative analysis. It offers global daily precipitation estimates at 0.1° resolution by integrating nine remote sensing and reanalysis datasets. The merging weights are calibrated using gauge observations with intelligent interpolation applied to ungauged regions to ensure global consistency (Beck et al., 2017; Beck et al., 2019). The MSWEP V2 dataset is available at: https://www.gloh2o.org/mswep/.

Gauge data: Note that the majority of RS, reanalysis, and merged precipitation products are heavily calibrated/validated using US and European gauge networks (Kang et al., 2024). Therefore, we focus on gauge data collected from mainland China and Africa, where precipitation observations are less commonly incorporated into existing precipitation products. A total of 5,972 independent gauges are collected from multiple sources to support validation at both daily and sub-daily scales across diverse climatic zones. Among them, 3,298 are daily gauges, including 814 from the National Meteorological Information Center of the China Meteorological Administration (CMA), 235 from the Hydrological Yearbook of China, 168 from the Haihe River Basin (2018–2023), 1,884 from northwestern China (2020–2021), and 197 from Africa (2015–2022) (van de Giesen et al., 2014). Additionally, hourly observations from 2,674 national stations, provided by the Hubei Provincial Meteorological Observatory, are used for validation at a 3-hourly resolution.

S2. Implementation details of QC and CTC-M algorithms

Quadruple collocation (QC)

To simultaneously estimate the error variances σ^2 and the inter-product error covariances required for precipitation merging, we utilizes the Quadruple collocation (QC) technique (Gruber et al., 2016). Following the standard QC framework, the relationship between the precipitation estimate from the i-th product (x_i) and the unknown truth (p) is expressed via a linear additive error model:

$$x_i = \beta_i p + \varepsilon_i \qquad (S.1)$$

Here, β_i represents the additive systematic bias while ε_i denotes the time-varying random error.

The QC analysis involves four precipitation datasets (x_1 through x_4), The method relies on the fundamental assumption that the error terms (ε_i) are statistically independent across different products ($\overline{\varepsilon_1 \varepsilon_2} = 0$ for $i \neq j$) and are uncorrelated with the true precipitation signal ($\mathrm{i.e.}$, $\overline{\varepsilon_\iota p} = 0$). Whiles standard Triple Collocation (TC) assumes zero error cross-correlation between all products, QC accommodates the existence of cross-correlated errors between a specific pair of products, such as x_3 and x_4 (i.e., $\overline{\varepsilon_3 \varepsilon_4} \neq 0$). Under these conditions, the system of equations involving the true precipitation variance and error variances can be solved linearly.

$$\boldsymbol{a}=\begin{bmatrix}\beta_1^2C_{pp}\\ \beta_2^2C_{pp}\\ \beta_3^2C_{pp}\\ \beta_4^2C_{pp}\\ C_{\varepsilon1\varepsilon1}\\ C_{\varepsilon2\varepsilon2}\\ C_{\varepsilon3\varepsilon3}\\ C_{\varepsilon4\varepsilon4}\\ C_{\varepsilon3\varepsilon4}\end{bmatrix}\mathbf{A}=\begin{bmatrix}1&0&0&0&0&1&0&0&0&0\\ 0&1&0&0&0&0&1&0&0&0\\ 0&0&1&0&0&1&0&1&0&0\\ 0&0&0&1&0&1&0&0&1&0\\ 0&0&0&0&1&0&0&0&0&1\\ 1&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0\end{bmatrix}\boldsymbol{b}=\begin{bmatrix}C_{11}\\ C_{22}\\ C_{33}\\ C_{44}\\ C_{34}\\ \dfrac{C_{12}C_{13}}{C_{23}}\\ \dfrac{C_{12}C_{23}}{C_{13}}\\ \dfrac{C_{23}C_{13}}{C_{12}}\\ C_{12}C_{14}/C_{24}\\ C_{12}C_{24}/C_{14}\\ C_{14}C_{24}/C_{12}\\ C_{13}C_{24}/C_{12}\\ \dfrac{C_{14}C_{23}}{C_{12}}\end{bmatrix} \tag{S.2}$$

where C_{xx} denote the covariance matrix of the input datasets. where individual elements such as C_{12} represent the covariance between x_1 and x_2. The coefficient vector $\boldsymbol{a}$ is estimated via a least-squares solution, denoted as $\tilde{\boldsymbol{a}}$:

$$\tilde{\boldsymbol{a}}=(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\boldsymbol{b} \tag{S.3}$$

By designating x_1 as the reference dataset for scaling purposes, we can derive the error variance of the rescaled products directly from the statistics within the $\tilde{\boldsymbol{a}}$ vector:

$$\sigma_{i,\mathrm{qc}}^2=C_{\varepsilon_i\varepsilon_i}\frac{\beta_1^2C_{pp}}{\beta_i^2C_{pp}} \tag{S.4}$$

Similarly, the error covariance between the rescaled x_3 and x_4 is given by:

$$\sigma_{34}^2=C_{\varepsilon_3\varepsilon_4}\frac{\beta_1^2C_{pp}}{\beta_i^2C_{pp}} \tag{S.5}$$

Furthermore, the correlation coefficient between the product and the truth (R_i^{qc}) is calculated as:

$$R_i^{\mathrm{qc}}=\sqrt{\frac{\beta_1^2C_{pp}}{\beta_i^2C_{pp}+C_{\varepsilon_i\varepsilon_i}}} \tag{S.6}$$

Prior to applying QC, a monthly climatological correction (see Sect. 2.1) is applied to all products to mitigate the impact of systematic biases (β_i).

Categorical Triple Collocation-based Merging (CTC-M)

CTC-M provides a superior binary (rain/no-rain) time series by utilizing three mutually independent precipitation products to quantify their relative detection skills and optimally merge their categorical information. This process first employs Categorical Triple Collocation (CTC, McColl et al., 2016) to quantify the relative detection skills of the input products. Subsequently, a probabilistic merging scheme combines the collocated binary series to maximize the likelihood of correct classification at each time step.

In this framework, the binary detection status d_i from product i relates to the ground truth event P through the following model:

$$d_i = P + e_i \tag{S.7}$$

where P represents the unknown true binary state (with +1 indicating rain and -1 indicating no-rain), d_i is the estimated state, and e_i signifies the classification error. The detection capability of d_i is quantified using the balanced accuracy metric, π_i:

$$\pi_i = 0.5(A_i + D_i) \tag{S.8}$$

Here, A_i and D_i correspond to the probability of d_i being correct when P is +1 and -1, respectively (McColl et al., 2016). Assuming that the classification errors are independent among the three products, the inter-product covariance (Q) of of the binary time series is formulated as:

$$Q_{12} = Cov(x_1, x_2) = f(I)(2\pi_1 - 1)(2\pi_2 - 1) \tag{S.9}$$

$$Q_{13} = Cov(x_1, x_3) = f(I)(2\pi_1 - 1)(2\pi_3 - 1) \tag{S.10}$$

$$Q_{23} = Cov(x_2, x_3) = f(I)(2\pi_2 - 1)(2\pi_3 - 1) \tag{S.11}$$

where $f(P)$ relates to the statistical properties of P and is typically unknown. However, by leveraging the inter-product covariances, the relative detection skill (v_i) for each product can be isolated:

$$v_1 = \sqrt{\frac{Q_{12}Q_{13}}{Q_{23}}} \tag{S.12}$$

$$v_2 = \sqrt{\frac{Q_{12}Q_{23}}{Q_{13}}} \tag{S.13}$$

$$v_3 = \sqrt{\frac{Q_{13}Q_{23}}{Q_{12}}} \tag{S.14}$$

where $v_i = \sqrt{f(P)}(2\pi_i - 1)$. Using the v_i values derived above, the merged binary time series (d_m) with optimal detection skill is obtained via the following weighted combination:

$$d_m = sign(w_1\, d_1 + w_2 d_2 + w_3 d_3) \tag{S.15}$$

The weight w_i for each contributing product is calculated as:

$$w_i = \frac{v_i^n}{v_1^n + v_2^n + v_3^n} \tag{S.16}$$

where n serves as a tuning parameter that adjusts the weights according to the relative skill differences among products. Previous analytical derivations (Dong et al., 2020) indicate that setting $n = 1.5$ is a robust choice, generally ensuring that the classification skill of the merged output meets or exceeds that of the best individual parent product (i.e., $\pi_m \geq max\,(\pi_1, \pi_2, \pi_3)$).

References

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., & Prat, O. P. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, *96*(1), 69-83. https://doi.org/https://doi.org/10.1175/BAMS-D-13-00068.1

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, *21*(1), 589-615. https://doi.org/10.5194/hess-21-589-2017

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., & Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473-500. https://doi.org/10.1175/bams-d-17-0138.1

Chen, M. Y., Shi, W., Xie, P. P., Silva, V. B. S., Kousky, V. E., Higgins, R. W., & Janowiak, J. E. (2008). Assessing objective techniques for gauge-based analyses of global daily precipitation. *JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES*, *113*(D4), Article D04110. https://doi.org/10.1029/2007JD009132

Dong, J., Crow, W. T., & Reichle, R. (2020). Improving Rain/No-Rain Detection Skill by Merging Precipitation Estimates from Different Sources. *Journal of Hydrometeorology*, *21*(10), 2419-2429. https://doi.org/https://doi.org/10.1175/JHM-D-20-0097.1

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., & Wagner, W. (2016). Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *Journal of Geophysical Research: Atmospheres*, *121*(3), 1208-1219. https://doi.org/https://doi.org/10.1002/2015JD024027

Huffman, G. J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan. (2023). *GPM IMERG Final Precipitation L3 1 day 0.1 degree V07* Version V07).

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, *5*(3), 487-503. https://doi.org/https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kang, X. Q., Dong, J. Z., Crow, W. T., Wei, L. N., & Zhang, H. W. (2024). The Conditional Bias of Extreme Precipitation in Multi-Source Merged Data Sets. *Geophysical Research Letters*, *51*(22), Article e2024GL111378. https://doi.org/10.1029/2024gl111378

McColl, K. A., Roy, A., Derksen, C., Konings, A. G., Alemohammed, S. H., & Entekhabi, D. (2016). Triple collocation for binary and categorical variables: Application to validating landscape freeze/thaw retrievals. *Remote Sensing of Environment*, *176*, 31-42. https://doi.org/https://doi.org/10.1016/j.rse.2016.01.010

van de Giesen, N., Hut, R., & Selker, J. (2014). The Trans-African Hydro-Meteorological Observatory (TAHMO). *WIREs Water*, *1*(4), 341-348. https://doi.org/https://doi.org/10.1002/wat2.1034

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C. (2007). A Gauge-Based Analysis of Daily Precipitation over East Asia. *Journal of Hydrometeorology*, *8*(3), 607-626. https://doi.org/https://doi.org/10.1175/JHM583.1

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., & Lin, R. (2017). Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. *Journal of Hydrometeorology*, *18*(6), 1617-1641. https://doi.org/https://doi.org/10.1175/JHM-D-16-0168.1