EGU26-9714, updated on 08 Jun 2026
https://doi.org/10.5194/egusphere-egu26-9714
EGU General Assembly 2026

MSWEP V3: Machine Learning-Powered Global Precipitation Estimates at 0.1° Hourly Resolution (1979–Present)

Hylke Beck[1], Xuetong Wang[1], Raied Alharbi[2], Oscar Baez-Villanueva[3], Diego Miralles[3], Jun Ma[1], Shiqin Xu[1], Matthew McCabe[1], Florian Pappenberger[4], Albert van Dijk[5], Tim McVicar[6], Lanka Karthikeyan[7,8], Hayley Fowler[9], Ming Pan[10], and Solomon Gebrechorkos[11,12]

[1]Thuwal, Saudi Arabia (hylke.beck@gmail.com)
[2]Department of Civil Engineering, College of Engineering, King Saud University, Riyadh, Saudi Arabia
[3]Hydro-Climate Extremes Lab (H-CEL), Ghent University, Ghent, Belgium
[4]ECMWF, Reading, United Kingdom
[5]Fenner School of Environment & Society, Australian National University, Canberra, ACT, Australia
[6]CSIRO Environment, Canberra, Australian Capital Territory, Australia
[7]Centre of Studies in Resources Engineering, IIT Bombay, Mumbai, India
[8]Centre for Climate Studies, IIT Bombay,Mumbai, India
[9]School of Engineering, Newcastle University, Newcastle upon Tyne, UK
[10]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, San Diego, California
[11]School of Geography and the Environment, University of Oxford, Oxford, UK
[12]School of Geography and Environmental Science, University of Southampton, Southampton, UK

We introduce Version 3 (V3) of the gridded near real-time Multi-Source Weighted-Ensemble Precipitation (MSWEP) product—the first fully global, machine learning-powered precipitation (*P*) dataset, developed to meet the growing demand for timely and accurate *P* estimates amid escalating climate challenges. MSWEP V3 provides hourly data at 0.1° resolution from 1979 to the present, continuously updated with a latency of approximately two hours. Development follows a two-stage process. First, baseline *P* fields are generated using machine learning model stacks that integrate satellite- and (re)analysis-based *P* and air-temperature products, along with static variables. The models are trained using hourly and daily observations from 15,959 *P* gauges worldwide. Second, these baseline *P* fields are corrected using daily and monthly gauge observations from 57,666 and 86,000 stations globally, using a method that accounts for gauge proximity, reporting times, inter-gauge dependencies, and correlation lengths. To assess MSWEP V3's baseline performance, we evaluated 19 (quasi-) global gridded *P* products—including both uncorrected and gauge-based products—using observations from an independent set of 15,958 gauges excluded from the first training stage. The MSWEP V3 baseline achieved a median daily Kling-Gupta Efficiency (KGE) of 0.69, outperforming all evaluated products. Other uncorrected products achieved median KGE values of 0.61 (ERA5), 0.46 (IMERG-L V7), 0.38 (GSMaP V8), and 0.31 (CHIRP). Notably, the MSWEP V3 baseline also outperformed several gauge-based products, including IMERG-F V7 (0.62), CPC Unified (0.54), and CHIRPS (0.36). Using leave-one-out cross-

validation, the daily gauge correction was found to improve the median daily correlation by 0.09, constrained by the already strong baseline performance. We anticipate that MSWEP V3 will substantially advance data-driven decision-making in hydrology and climate science, by enabling more reliable monitoring, forecasting, and management of water-related risks in a variable and changing climate.