https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Scalable radar-driven approach with compact gradient-boosting models for gap filling in high-resolution precipitation measurements

Peter Lünenschloß[1,2], Antje Claussnitzer[3], Thomas Schartner[3], Mirjam Brunner[1,2], Timo Houben[1,2], David Schäfer[1,2], and Jan Bumberger[1,2,4]

[1]Helmholtz Centre for Environmental Research – UFZ, Research Data Management - RDM, Permoserstraße 15, Leipzig 04318, Germany

[2]Helmholtz Centre for Environmental Research – UFZ, Department Monitoring and Exploration Technologies, Permoserstraße 15, Leipzig 04318, Germany

[3]German Meteorological Service (DWD), Department for Quality Assurance of Measurement and Observation Data, Michendorfer Chaussee 23, 14473 Potsdam, Germany

[4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstraße 4, Leipzig 04103, Germany

Correspondence: Peter Lünenschloß (peter.luenenschloss@ufz.de)

Abstract.

High-frequency precipitation records are essential for hydrological modeling, weather forecasting, and ecosystem research. Unfortunately, they usually exhibit data gaps originating from sensor malfunctions, significantly limiting their usability. We present a framework to reconstruct missing data in precipitation measurements sampled at 10 min frequency using radar-based, gauge independent, precipitation estimates as the only predictor. We fit gradient-boosting models to the statistical relationships between radar-based precipitation fields and collocated rain gauges. The obtained models allow for the filling of data gaps of arbitrary length and additionally provide confidence interval approximations. We evaluate the method using the rain gauge network of the German Weather Service (DWD), which roughly covers the entirety of Germany. The results show robust performance across diverse climatic and topographic conditions at a high level, with the coefficient of determination averaging at around 0.7. The framework is computationally very cheap, relying on a single CPU core only. This makes scaling easy and integration into operational gap filling of extensive sensor networks feasible.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

1 Introduction

High-resolution precipitation data are a cornerstone for accurate weather forecasting, hydrological modelling, and climate monitoring (Kraft et al., 2022; Li et al., 2023). Particularly in densely populated or flood-prone regions, 10-minute rainfall measurements provide critical information for early warning systems, runoff simulations, and the representation of fast hydrological processes (Rafieeinasab et al., 2015; Ming et al., 2020; Piadeh et al., 2022). Their relevance, however, extends well beyond meteorology. Environmental research infrastructures rely on high-frequency rainfall to capture rapid changes in soil moisture, infiltration, and biogeochemical fluxes (Mollenhauer et al., 2018; Schädler et al., 2019; Schnepper et al., 2023; Zacharias et al., 2024; Ohnemus et al., 2025); agricultural applications depend on sub-hourly observations for irrigation scheduling, crop-stress assessment, and yield forecasting (Lagasio et al., 2025); urban hydrology uses them to anticipate sewer surcharge and pluvial flooding (Thorndahl et al., 2017; Piadeh et al., 2022); and water-resource management requires them for reservoir operation, groundwater-recharge estimation, and flash-flood preparedness (Tashie et al., 2016; Yu et al., 2016). The increasing deployment of environmental Digital Twins and AI-driven decision-support systems further amplifies the demand for continuous, gap-free precipitation records to maintain stable system states and ensure reliable model performance (Li et al., 2023; Yang et al., 2024; Hazeleger et al., 2024).

However, such temporal resolution comes at a cost: dense gauge networks are inherently susceptible to interruptions and quality issues. Data streams are frequently affected by missing or erroneous values due to sensor failures, telemetry outages, maintenance cycles, or related quality-control flags. These discontinuities reduce the AI readiness of the data, complicate both operational forecasting and scientific analysis, and limit the applicability of learning algorithms that depend on consistent spatiotemporal input sequences.

In response, many observatories and environmental research infrastructures have implemented automated quality-assurance and control (QA/QC) pipelines that combine rule-based tests, metadata-aware workflows, and scalable software frameworks to ensure traceable and reproducible data streams (Schmidt et al., 2023; Zacharias et al., 2024). Recent developments range from generic systems such as the System of automated Quality Control (SaQC) and digital ecosystems for sensor data management which explicitly exploit possible external reference information for screening both professional and citizen-science gauge networks (Horsburgh et al., 2025; Bumberger et al., 2025). These frameworks not only support continuous validation and curation but increasingly integrate routines for automated gap detection and reconstruction. Within SaQC, for example, rule-based imputation of missing or flagged intervals is directly embedded into the reproducible data-quality workflows, highlighting the growing convergence of QC and time-series gap-filling methodologies. Together with emerging machine-learning approaches – such as graph neural networks for anomaly detection in environmental sensor networks (Lasota et al., 2025) – these systems contribute to a more resilient and interoperable handling of environmental time series.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Despite these advances, high-temporal-resolution precipitation data remain particularly challenging. Even short gaps can compromise downstream applications or data assimilation. Traditional interpolation methods, such as temporal interpolation or inverse distance weighting (IDW), fail to reproduce the localized and intermittent nature of rainfall fields (Vicente-Serrano et al., 2003; Haberlandt, 2007). More advanced statistical and regression-based techniques have been developed to exploit auxiliary information, for example from neighbouring gauges or meteorological covariates such as temperature, humidity, and wind (Portuguez-Maurtua et al., 2022; Tang et al., 2021; Faramarzzadeh et al., 2023). However, their performance tends to deteriorate at high temporal resolution, where spatial correlations weaken and network density becomes critical.

Building on these findings, our earlier work demonstrated that local meteorological predictors – such as temperature, relative humidity, wind direction, and neighbouring-gauge precipitation – provide only limited skill for high-frequency (≤ 30 min) imputation, with performance degrading sharply for convective events and at 10-minute sampling rates (Lünenschloß et al., 2022). Moreover, these approaches depended strongly on dense gauge constellations, limiting their applicability in regions where inter-station distances exceed 15 km, as is typical for parts of the DWD network. These findings directly motivated the present study, which aims to overcome these structural limitations by relying exclusively on radar-derived quantitative precipitation estimates as predictors and thereby avoiding the density constraints inherent to gauge-based auxiliary variables. In the following, we therefore use the German Weather Service (DWD) national gauge network primarily as a large-scale, heterogeneous testbed to evaluate whether such a radar-only strategy can robustly reconstruct 10-minute precipitation across the main hydroclimatic and topographic settings in Germany, from coastal lowlands and river plains to low mountain ranges and densely urbanised areas.

Ensemble and deep-learning frameworks have recently demonstrated strong potential for precipitation estimation and nowcasting (Ravuri et al., 2021; Mital et al., 2020; Qiu et al., 2024), yet most of these models were designed for daily or hourly scales and depend on extensive feature engineering and dense observation networks. To date, no systematically evaluated and scalable framework exists for imputing 10-minute precipitation data and reconstructing missing intervals of arbitrary length in sparse gauge networks. This methodological gap motivates the present study, which introduces a radar-driven machine-learning approach designed to efficiently reconstruct high-frequency precipitation records from radar-based quantitative precipitation estimates (QPE).

Among existing radar products, most national or continental QPE frameworks – such as the U.S. Multi-Radar Multi-Sensor (MRMS), MeteoSwiss' CombiPrecip, and the UK high-resolution gauge–radar–satellite product (UKGrsHP), the Dutch operational gauge-adjusted radar product, and the pan-European EURADCLIM dataset – apply gauge-based bias correction or merging to improve operational accuracy (Erdin et al., 2012; Zhang et al., 2016; Yu et al., 2020; Overeem et al., 2023, 2025). Numerous such systems implement systematic bias adjustments based on national gauge networks, which complicates their use as independent predictors for gauge-imputation due to potential calibration leakage. Similarly, the DWD's operational RADOLAN product performs hourly gauge adjustment of radar composites using approximately 1300 surface gauges from

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

the national network (Winterrath et al., 2012). To avoid calibration leakage when targeting that same gauge network, we use the intermediate, ungauged QPECommon radar product instead (Trömel et al., 2024). Using RADOLAN as a predictor for imputing the very gauges that enter its adjustment would be methodologically inconsistent, because the radar fields already incorporate information from these target stations via the gauge-correction step. This allows a strict separation between predictors and targets and enables a rigorous assessment of the incremental value contained in radar-only features. Since gauge-adjusted products are structurally similar to QPECommon – with calibration primarily increasing correlation with ground truth – we expect our results to be transferable (as a baseline) to additional gauge networks when using RADOLAN or comparable composites, provided the respective gauges are not part of the calibration. From this perspective, the DWD gauge network serves as a comprehensive, well-documented testbed, while the methodological focus lies on a leakage-free coupling between gauge networks and ungauged or non-overlapping radar QPE products, which can in principle be replicated for other national or research networks.

For model design, we employ a compact gradient boosting approach (XGBoost) that is computationally efficient, requires minimal hyperparameter tuning, and has demonstrated strong generalisation capabilities across diverse climatic and geographic contexts. Our approach offers several advantages that distinguish it from previous studies. First, it is applicable across a wide range of climatic and geographic settings, which we empirically demonstrate by applying the framework to the complete DWD national network as a single, heterogeneous test case. Second, it operates at a high temporal resolution of 10 minutes while still enabling reconstruction of missing intervals from single-step gaps to arbitrarily long outages. Third, because it is independent of station density, performance persists even in sparse clusters of a network where inter-station distances exceed 15 km, as is typical for the DWD network. Neighbourhood-based methods rapidly lose reliability under such conditions. Furthermore, instead of training a single global model, we employ lightweight, station-specific learners, which simplifies operational scaling across large national networks while maintaining site-level adaptability. Unlike earlier studies that incorporated temperature, humidity, or wind as auxiliary features, our model relies exclusively on precipitation-related predictors, reflecting our own findings that non-precipitation variables contribute little predictive power at 10-minute resolution.

A further strength lies in the temporal architecture of the model: by restricting predictors to a short temporal horizon around a target gap (+/- 30 minutes), the framework avoids reliance on fragile long-term dependencies and remains suitable for real-time applications. Crucially, the use of the radar-only QPECommon product prevents circular validation and enables a clean assessment of radar-based predictive skill. Finally, the model demonstrates operational readiness: we show reliable performance over extended gaps, making the method suitable for integration into automated quality-control systems and real-time forecasting pipelines.

In summary, our framework introduces three key aspects:

1. it relies exclusively on radar-based predictors from the DWD QPECommon product, avoiding gauge-based inputs or bias corrections;

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

2. it enables the imputation of high-frequency (10-minute) precipitation data while supporting reconstruction of intervals of arbitrary length using a compact gradient-boosting design; and

3. it employs station-specific learners, allowing the method to scale to large national networks while preserving local adaptability.

Together, these features provide a radar-only yet gauge-validated approach that bridges the gap between radar-based precipitation estimation and the operational correction of observational datasets and enhance the AI readiness of the resulting high-frequency precipitation records. This methodological development is relevant for data-quality control and real-time monitoring applications as well as AI-based forecasting and decision-support systems, all of which rely on complete and consistent high-frequency precipitation records. In this study, we therefore interpret the DWD network as an exemplar of a large environmental observation system and use it to demonstrate that a compact, radar-driven imputation framework can operate reliably across the full diversity of settings encountered in such networks, providing a blueprint for extensions to other research infrastructures and operational gauge systems.

2 Data

2.1 Rain gauge network

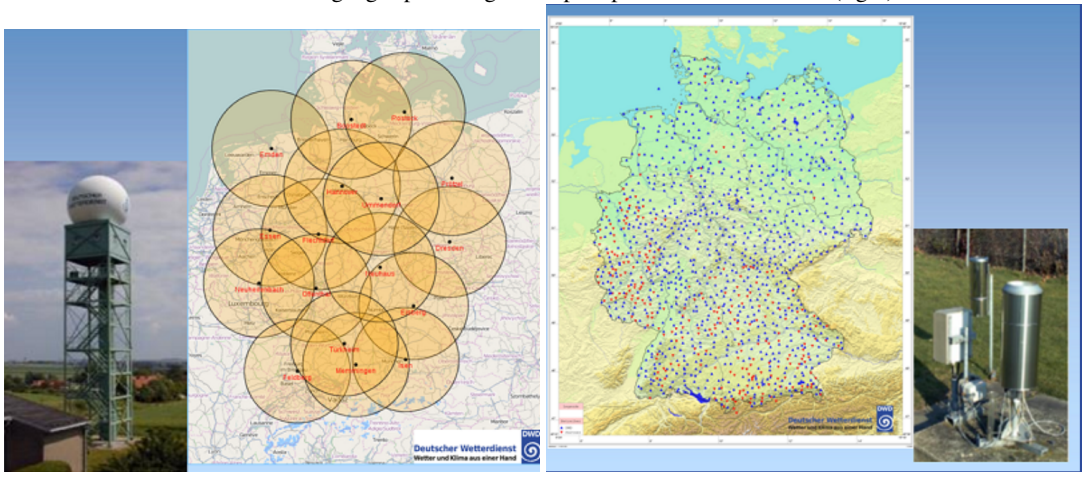
The German Weather Service (DWD) and its partners operate an extensive network of more than 1700 measurement stations distributed across Germany. Figure 1 provides an overview of the DWD observation infrastructure used in this study. Approximately 1100 of these provide, in addition to other meteorological parameters, high-resolution precipitation measurements expressed as water column height accumulated per minute. Precipitation data from this sub-network, which provides near-complete spatial coverage of Germany, were selected as the target variable for this study (see Fig. 1).

The instruments installed at these stations are calibrated at the DWD's national calibration laboratories, which operate in accordance with DIN EN ISO/IEC 17025:2018 and ensure traceability to national standards of the Physikalisch-Technische Bundesanstalt (PTB) (Deutscher Wetterdienst (DWD), 2025). All data undergo a basic near-real-time quality-control process before publication. To align with the available QPECommon radar composites, the analysis period was restricted to 2021–2024. Depending on the station, longer historical records are available, but the chosen period provides consistent coverage across almost all stations. Only a few gauges had to be excluded due to data gaps.

The spatial distribution of the network exhibits strong regional variability, with higher station densities in urbanised and western parts of Germany and sparser coverage in mountainous or rural regions. Distances between neighbouring gauges range from about 10 km to 20 km, averaging roughly 15 km.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Figure 1. Overview of the DWD observation infrastructure used in this study: locations and coverage of the 17 dual-polarisation weather radars (left), and distribution of the ∼1100 rain gauges providing 1 min precipitation measurements (right). Source: DWD



2.2 Quantitative precipitation estimates (QPE)

As the sole predictor variable for the machine-learning model, we used the DWD's intermediate Quantitative Precipitation Estimate (QPECommon). QPECommon represents an areal estimate of precipitation intensity with a spatial resolution of 1 km × 1 km. It is derived from the DWD precipitation (PCP) scan produced by the DWD's network of 17 dual-polarisation weather radars (Deutscher Wetterdienst (DWD), 2015). The PCP scan is a plan-position indicator (PPI) with azimuth-dependent elevation angles chosen such that the radar beam just clears the local horizon in each direction. Rainfall intensities are computed using a multi-stage reflectivity–rain-rate (Z–R) relationship of the form $Z = aR^b$, where Z is radar reflectivity, R is surface rain rate, and a and b are empirical coefficients that vary with hydrometeor class.

For the derivation of QPECommon, the DWD's Hydrometeor Classification Algorithm (Hymec) is applied to determine the spatial extent and vertical position of the atmospheric melting layer (Steinert et al., 2021). The Z–R relationship for returns originating from this layer is then adjusted to mitigate the typical overestimation of reflectivity-to-rain-rate conversion. Subsequently, rain rates from all radar sites are combined into a unified composite using a polar-stereographic projection on the WGS84 ellipsoid, with curated weighting of overlapping radar beams (Mott and Schultze, 2024).

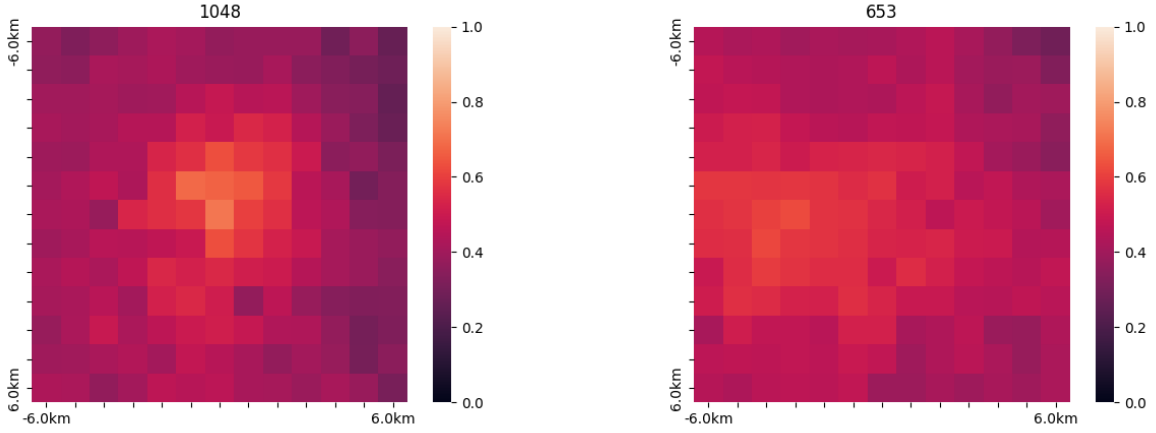
Due to projection-related distortions and the fully radar-based nature of the dataset, the QPECommon composite is not always perfectly spatially aligned with the corresponding rain-gauge positions. Figure 2 illustrates examples of well- and poorly aligned QPE coverage around individual stations, showing correlations between the radar-based estimates and the precipitation measured on the ground at the centre of each map (Trömel et al., 2024). An objective of this study is to demonstrate that statistical relationships learned solely from the training data can already produce accurate mappings between QPECommon

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

and gauge observations, enabling reliable predictions for unseen data.

Importantly, the generation of precipitation intensities – as opposed to quantities – does not incorporate rain-gauge data for bias correction or calibration. This ensures complete independence from the ground observations that serve as regression targets in our study.

Figure 2. Examples of radar–gauge correlation maps illustrating alignment quality between the DWD QPECommon radar composite and gauge observations. The left figure shows a station with high spatial correlation (ID 1048), whereas the right figure represents a case with weaker alignment caused by local displacement or beam geometry effects (ID 653).



3 Related Work

This section situates our contribution in three method families: (i) conventional spatial interpolation for precipitation gap-filling, (ii) regression-based and statistical imputation, and (iii) machine-learning and modern data-driven approaches. Throughout, we emphasize temporal scale, network density, and the role of radar-derived predictors, linking directly to the scope of our study (10-minute resolution; sparse-to-moderate gauge spacing; QPE-driven features).

3.1 Conventional spatial interpolation for precipitation gap-filling

Classical gap-filling relies on spatial interpolation between gauges (Vicente-Serrano et al., 2003; Haberlandt, 2007). Inverse Distance Weighting (IDW) remains a popular deterministic baseline. Several studies have optimized IDW parameters (e.g., search radii of 10–30 km) and reported high agreement between estimated and observed precipitation, though with reduced accuracy during wetter periods compared to drier conditions (Chen and Liu, 2012). Extensions such as Modified-IDW and related weighting refinements have also been evaluated for daily rainfall reconstruction (Teegavarapu and Chandramouli, 2005), underscoring the long-standing role of distance-based schemes in gauge-data imputation. Thiessen polygons and simple gauge

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

averages are easy to implement and frequently serve as baselines, but their performance depends strongly on station density and homogeneity of rainfall fields (Tavares et al., 2025).

Geostatistical methods such as Kriging and variants exploit spatial covariance. Several studies documented systematic distortions for rainfall when using Kriging (e.g., overestimation of rainy days and underestimation of event intensities), motivating the use of Co-Kriging with auxiliary variables (elevation, terrain indices) where appropriate. Comparative evaluations also highlight substantial climate- and network-dependence: for example, Sattari et al. (2017) found strong method sensitivity when filling daily precipitation gaps in arid environments, with no universally best-performing kriging or interpolation variant. Hybrid improvements that embed learning or optimization inside geostatistical schemes have also been explored: e.g., Adhikary et al. (2016) coupled genetic programming with Kriging to enhance interpolation skill. Overall, these methods provide robust spatial baselines but may struggle with highly intermittent, small-scale rainfall at sub-hourly resolution or in complex terrain, which is central to our application.

3.2 Regression-based and statistical imputation

An established class of approaches models cross-station dependence explicitly using regression and statistical imputation. Multiple linear regression (MLR) and related formulations have long been used to reconstruct missing precipitation from neighboring gauges and geographic covariates (Tang et al., 2021; Portuguez-Maurtua et al., 2022; Faramarzzadeh et al., 2023). At monthly scales, Chutsagulprom et al. (2022) used MLR as a strong baseline against Kriging and artificial neural networks (ANNs). Radar-informed regression has also been explored at daily scales for producing gauge-guided gridded fields (e.g. DeGaetano and Wilks, 2009), although these approaches target spatial climatologies rather than sub-hourly gauge time series.

Distribution-aware strategies aim to preserve the empirical properties of rainfall despite gaps. Simolo et al. (2010) proposed a daily imputation scheme that preserves the empirical rainfall probability distribution function (PDF), maintaining rainfall/no-rain frequencies and intensity distributions and mitigating biases that arise with naïve fillers. Such biases can be consequential for climate applications: for example, Zolina et al. (2009) showed that even modest fractions of missing data may distort extreme-precipitation indices in European climate records, underscoring the need for careful treatment of missingness.

Beyond single-imputation, likelihood-based and multiple-imputation frameworks are frequently evaluated. Hırca and Eryılmaz Türkkan (2024) compared mean/median, regression, hot-deck, k-nearest neighbours (k-NN), and expectation–maximization (EM)-based strategies across diverse missingness scenarios and found EM to minimize errors under their assumptions, highlighting its value as a non-ML reference. Region-specific, cluster-aware regressions – e.g., Bagirov et al. (2017) using clustered linear models – can further improve accuracy by tailoring relations to homogeneous subregions. Recent systematic reviews emphasize the breadth of statistical and regression-based gap-filling approaches and their generally strong performance at daily and monthly scales, while noting clear limitations at high temporal resolution where intermittency dominates (e.g. Abdillah et al., 2023). For daily precipitation, Kim and Ryu (2015) further showed that imputation accuracy deteriorates markedly once

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

missingness exceeds approximately 15%, underscoring the sensitivity of classical statistical methods to gap structure and data completeness.

These strands collectively underscore that well-specified statistical baselines can be competitive at daily-to-monthly scales, yet they require careful assumptions and tend to be less effective at capturing the intermittency and sharp gradients of 10-minute rainfall.

3.3 Machine learning and data-driven approaches

Recent work leverages nonlinearity, interactions, and flexible feature sets to address the fine-scale variability of precipitation (Mital et al., 2020; Ravuri et al., 2021; Qiu et al., 2024). At sub-hourly resolution, Chivers et al. (2020) evaluated regression and classification models for 30-minute imputation on a dense UK network and reported $R^2 \approx 0.4$–0.6, with diminishing marginal benefits from distant neighbors (see also Lünenschloß et al., 2022). Pushing temporal resolution further, Vidal-Paz et al. (2023) studied 10-minute imputation using multiple-imputation frameworks across Spain; despite cross-correlation alignment and rich data, performance remained modest ($R^2 < 0.5$), illustrating the intrinsic difficulty at such fine scales.

On daily data, Bellido-Jiménez et al. (2021) compared multilayer perceptrons (MLP), support vector machines (SVM), and random forests (RF) for Andalusia and found MLP particularly strong in coastal, topographically complex regions (nonlinear effects), whereas gains inland were smaller. Conversely, at monthly scales, Chutsagulprom et al. (2022) observed ANNs trailing distance-based and regression baselines, underlining how optimal method choice shifts with temporal aggregation. A recent synthesis by Abdillah et al. (2023) similarly highlights the strong dependence of ANN performance on temporal scale, climate regime, and missingness structure, and concludes that classical machine-learning models often outperform deep architectures for short time-step precipitation imputation.

Ensemble tree learners consistently appear among the top performers for tabular hydro-meteorological data (Toure et al., 2025). Zhang et al. (2019) introduced SMILES, an XGBoost-based imputation framework that surpassed classical multiple-imputation-by-chained-equations (MICE) approaches, including a three-dimensional extension (3D-MICE), on high-dimensional time series, demonstrating that boosted trees can effectively integrate temporal context and cross-series information – a principle we operationalize here for precipitation. In rainfall prediction and related tasks, boosted trees repeatedly outperform conventional ML baselines: for instance, Sharma et al. (2021) report XGBoost achieving the highest accuracy among seven ensemble and non-ensemble classifiers for daily rainfall occurrence prediction, while Setya et al. (2023) find k-nearest neighbours (k-NN) outperforming multiple linear regression in short-term rainfall forecasting, underscoring the competitiveness of nonparametric learners under limited feature engineering. Daily imputation studies such as Teegavarapu et al. (2018) further demonstrate that optimized weighting schemes, k-NN corrections, and ANN-based estimators can outperform classical interpolation, although performance gains diminish at higher temporal resolutions. Studies on related tasks (e.g., daily gap-filling or prediction) also report advantages for gradient boosting and RF under appropriate tuning and QC (e.g., Qiu et al., 2024;

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Portuguez-Maurtua et al., 2022). Simpler nonparametric baselines such as k-NN remain competitive for certain stations or regimes and have been improved via distance reweighting (e.g. Huang et al., 2017); see also broader comparisons that include k-NN, ANN, and hybrids (e.g. Gorshenin et al., 2019; Rodrigues et al., 2023).

Deep learning (DL) has advanced radar-based nowcasting (Wani et al., 2024), but evidence for DL superiority in imputation of gauge time series is mixed and often scale-dependent. A recent review highlights generative adversarial network (GAN)–based imputation as particularly effective for climate time series (Alejo-Sanchez et al., 2025); however, reproducible advantages over compact learners at sub-hourly temporal resolution have not been demonstrated. Overall, comparative studies suggest that compact, well-regularized tree-based models (e.g., XGBoost) provide an excellent balance between accuracy and robustness, particularly under heterogeneous network configurations and limited feature engineering-conditions that characterize our setting.

3.4 Radar-driven imputation and inverse fusion

Radar-gauge fusion has a long history in quantitative precipitation estimation (QPE). Most operational national and continental radar products – such as the U.S. Multi-Radar Multi-Sensor (MRMS), MeteoSwiss' CombiPrecip, the UK high-resolution gauge–radar–satellite product (UKGrsHP), the Dutch operational gauge-adjusted radar product, and the pan-European EURADCLIM dataset – apply systematic gauge-based bias correction or merging to improve spatial accuracy (Erdin et al., 2012; Zhang et al., 2016; Yu et al., 2020; Overeem et al., 2023, 2025). Similarly, the DWD's RADOLAN composite performs hourly gauge adjustment using approximately 1300 surface stations from the national network (Trömel et al., 2024). These and related QPE frameworks have been extensively evaluated and refined, for example for Germany and central Europe (Paulat et al., 2008; Sideris et al., 2014; Rabiei and Haberlandt, 2015), and are summarised in methodological overviews of radar–gauge merging techniques (e.g. McKee and Binns, 2016). They are explicitly designed to produce the best-possible gridded precipitation fields by combining the spatial coverage of radar with the point accuracy of gauges. Modern ML- and DL-based nowcasting approaches extend the same forward-fusion paradigm by refining radar fields with additional learned corrections (e.g. Qiu et al., 2020; Ravuri et al., 2021; Kim and Hong, 2022; Meuer et al., 2025).

In contrast, the inverse direction – using radar-derived QPE as an independent predictor to impute missing values in gauge time series – has received far less attention. Existing fusion approaches almost universally treat gauges as the correction target and radar as the bias-prone predictor, rather than the reverse. To our knowledge, no systematic evaluation exists on radar-only predictors for sub-hourly gauge imputation and existing methodological reviews of radar/gauge integration focus exclusively on forward merging for QPE production (e.g. Rabiei and Haberlandt, 2015; McKee and Binns, 2016).

Several practical factors explain this gap. First, non-adjusted radar composites are often considered insufficiently aligned with point-scale precipitation to serve as reliable predictors, given displacement errors, beam-blockage effects, and bright-band contamination (e.g. Sideris et al., 2014; Rabiei and Haberlandt, 2015). Second, once radar and gauge data are fused in

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

a forward manner (as in RADOLAN, CombiPrecip, MRMS, or EURADCLIM), the resulting gauge-adjusted QPE cannot be reused as an independent predictor without introducing calibration leakage – an issue that directly motivated our decision to use the ungauged DWD QPECommon product as a leakage-free radar predictor. The same design principle can, in principle, be transferred to other observational networks by pairing them with ungauged or non-overlapping radar composites, or with gauge-adjusted products such as RADOLAN or MRMS, provided that their calibration excludes the target gauges and thus preserves predictor–target independence.

Despite the lack of prior work on this inverse fusion direction, radar-driven imputation is highly relevant for applications requiring complete and temporally coherent gauge records, including operational quality control, real-time monitoring, and the maintenance of environmental Digital Twins. In such settings, gap-free gauge series are a key prerequisite for AI-based forecasting, data assimilation, and decision-support workflows. Assessing the predictive value contained in ungauged radar composites therefore fills a methodological gap between traditional radar–gauge merging and the reconstruction of high-frequency observational datasets with a level of completeness and consistency that enhances their AI readiness.

4 Methods

4.1 Model architecture

We deployed the Extreme Gradient Boosting (XGBoost; Chen and Guestrin (2016)) framework to obtain base learners. We fitted an ensemble of XGBoost regression models to the precipitation (RR) record collected at each rain gauge separately.

These ensembles were generated by stacking XGBoost quantile regressors, each targeting one of the 0.1–0.9 quantiles (in 0.1 increments) of the gauge's precipitation record. Quantile predictions itself were added to the data used for training the final XGBoost regressor. The precipitation predictions of this top-level regressor were then accompanied by uncertainty assessments as provided by the stacked quantile regressors. This procedure is detailed in Sec. 4.3, and the per-timestep evaluation/inference protocol is laid out in Sec. 4.4. See Fig. 3 for a schematic representation of the proposed model architecture.

For each station, several of the introduced ensembles were trained, each constituting an independent imputation head. Each imputation head is specialized and trained to reconstruct precipitation for a certain gap size. All possible gap sizes, up to the model's temporal horizon w, are accounted for. To impute gaps exceeding this horizon, an arbitrary head is trained. This head does not rely on any precipitation records preceding or succeeding the prediction target and ensures the model's applicability in any possible outage constellation.

During inference, the head that maximizes utilization of the available precipitation information is selected automatically. This modular construct ensures optimal results across different gap sizes and is computationally feasible due to the relatively minor amount of resources that the training of any XGBoost model instance in the proposed way, does require.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

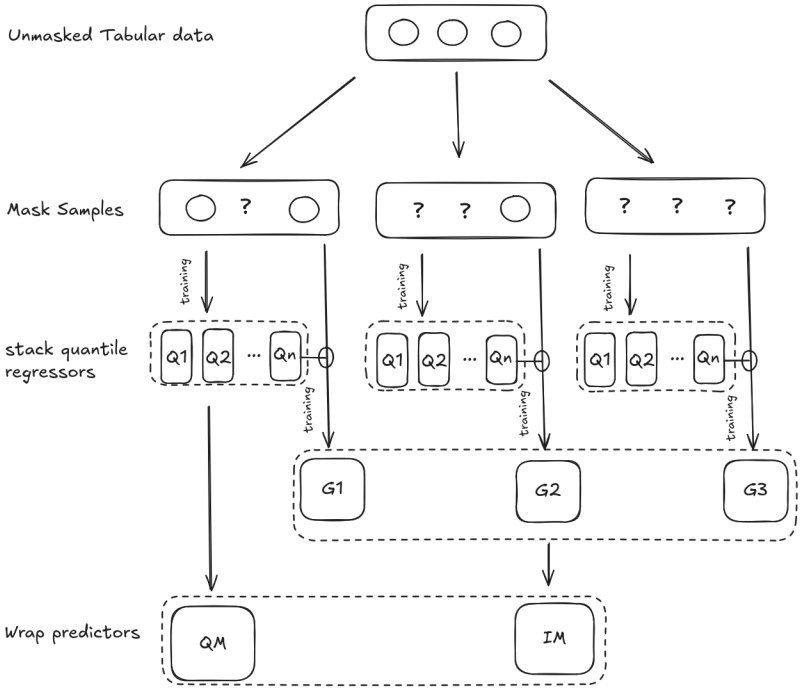


Figure 3. Architecture of the station-specific model ensemble used for precipitation gap filling. The *tabular data* input refers to the processed and enriched feature input (Sec. 4.2). The masking stage refers to the targeted gauges precipitation record. $Q_1 \dots Q_n$ denote the quantile regressors; QM denotes their stacked predictions; IM denotes the final second-stage imputation regressor. The three heads correspond to model heads fitted to different gap shapes and sizes.

4.2 Feature engineering and preprocessing

Preprocessing was applied individually to the QPE data obtained from each station's wider vicinity. The processing procedure aims to enrich the narrower selection of QPE pixels on which training is actually carried out with aggregates that provide a wider spatio-temporal and statistical context.

To identify the QPE pixel that best represents a gauge's position in the overall QPE grid, the temporal cross-correlation between rainfall records collected at the gauge during the training period (2022–2023) and the radar-based estimates collected at all QPE coordinates within the wider range was calculated. The coordinates yielding the highest correlation were selected as the station's reference coordinates and were subsequently used to determine the distance of any other radar pixels to the modeled rain gauge. When processing the test dataset, the reference coordinates calculated in this way were reused to avoid train–test leakage and ensure that evaluation metrics genuinely reflect prediction performance on unseen data.

Subsequently, all QPE pixels within a smaller radius ($r = 3$ km) around the reference were selected as predictor features. By obtaining lagged versions of these selected features, a structured tabular representation of short-term temporal correlations

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

in the radar information is generated.

To engineer variables that reflect wider-range spatial correlations, we derived additional geometric aggregates from QPE pixels available within a wider perimeter of R around the reference pixel. These included the mean, standard deviation, and center of mass along the two main coordinate axes and within the four coordinate quadrants. These aggregates were also lagged both forward and backward by multiples of the window length (l) to generate representations of their short-term variance.

To capture long-term correlations, cyclic seasonality was encoded using sine–cosine representations, adding hour-of-day and day-of-week features. Iterative seasonality, such as month of the year, was encoded using binary features and added as well.

Finally, for training the precipitation regressor, predictions from the quantile regressor models were also added, generating a feature set that combines raw radar signal inputs with geometric summaries, temporal structure representations, and distributional estimates. To further reduce redundancy, inter-feature correlation, and computational load, this feature set is compressed using Partial Least Squares (PLS) decomposition during cross-validation.

The complete feature-engineering pipeline is described in Table 1.

Table 1. Summary of preprocessing operations applied at successive stages of the data processing workflow: (1) complete dataset preparation, (2) processing of train–test splits, and (3) transformations applied to cross-validation folds before model fitting.

1. Processing of wider range QPE data
Reference pixel determination Selection of nearby *QPE* pixels Concatenation with geometric aggregates in wider range
2. Processing of nearby selection
Feature lagging Concatenation with seasonality encoding
3. Processing of cross–validation folds
Concatenation with quantile prediction Dimensionality reduction (PLS)

4.3 Training setup

Model training was conducted using a ten-fold cross-validation (CV) scheme. To preserve temporal structure and avoid data leakage, folds were generated sequentially—without shuffling—so that each fold represented a continuous time segment of the training period. This design ensured that temporal autocorrelation within rainfall sequences was handled realistically and that the model's ability to generalize to unseen future periods was properly captured by the cross-validation.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

XGBoost training was configured using default hyperparameter settings, which are known to provide strong out-of-the-box performance for tabular hydrometeorological data when combined with XGBoost's native early-stopping mechanism. This mechanism halts the training process when the mean validation score across the validation folds no longer improves. The number of non-improving boosting iterations tolerated before early stopping was set to 20 for all stations.

The QPE neighborhood radii were also fixed globally at $r = 3$ km for the close range and $R = 6$ km for the wider geometric aggregates, providing a consistent spatial horizon across all sites. The size of the model's temporal horizon w was fixed globally at 70 minutes, determining the number of forward and backward lags added to the input feature set as $l = 2$.

Training and validation of a site's model were executed on a single CPU core with 30 GB of assigned RAM. End-to-end processing for one station, including feature extraction, model fitting, and evaluation, required roughly 15 minutes of processing time. This modest runtime and independence from GPU availability demonstrate the scalability of the framework: the full network of approximately 1100 gauges can be retrained automatically in shared compute environments, enabling routine updates when new gauge data or revised radar composites become available.

4.4 Testing and Evaluation

For independent testing, all observations obtained in 2024 were excluded from the training procedure and strictly reserved for final model evaluation.

Test scores were obtained by simulating gaps in the test data using a sliding-window masking procedure. For each timestamp, a synthetic gap of width g was created by masking all observations in a window of the same size centered on the target timestamp. The model was then tasked with reconstructing the rainfall intensity at the central timestamp using only the remaining unmasked predictors.

Gap sizes were increased systematically to evaluate reconstruction performance under increasing information loss. The increments consisted of 10-minute, 30-minute, and 50-minute gaps. Finally, the model was applied to gaps of arbitrary length (cf. Fig. 3), simulating reconstruction conditions of long-term, multi-hour outages.

Predictive accuracy metrics were obtained by comparing the values at the center of each reconstructed window with the precipitation actually observed at the same time.

Regression metrics included the coefficient of determination (R^2), the Pearson correlation coefficient, and the root-mean-square error (RMSE). To assess event-detection performance, the F_1 score, F_2 score, and the Heidke Skill Score (HSS) were calculated. In order to do so, boolean event representations for the rain/no-rain classes were derived directly from the regression

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

outputs by using the threshold that maximizes the F-score on the training data as a cutoff. This avoided the need to train a separate classification model.

5 Results and Discussion

5.1 Overall model performance and validation

Table 2 summarises the overall regression and classification metrics obtained across the used 1100 stations of the DWD network. For the 10 min gap length, the model achieved mean correlation values of $r = 0.81$ and $R^2 = 0.66$, indicating strong predictive skill given the high sampling frequency and the spatial intermittency of precipitation at this scale. As the gap length increased to 30, 50 and arbitrary durations, a gradual but moderate decline in performance was observed, with R^2 values remaining above 0.55 even for the most challenging cases. The relatively small variation in RMSE across gap lengths reflects the strong influence of the many low-rain ('zero') periods in the evaluation set, which dominate the error distribution and thereby reduce sensitivity to gap size. Event-detection metrics also remained robust, with F_1 and F_2 scores exceeding 0.8 for gaps up to 50 min. HSS ranged between 0.6 and 0.8, reflecting consistent discrimination between rain and no-rain periods across the network.

Table 2. Regression and classification metrics obtained from aggregation over the network portion owned by the german weather service itself ($\sim$ 1100 stations). Including Partner stations deteriorates the scores by about 10 percent.

w	10 min	30 min	50 min	arbitrary
Pearson	0.81±0.05	0.78±0.06	0.76±0.06	0.75±0.07
R^2	0.66±0.1	0.6±0.1	0.57±0.12	0.55±0.12
$RMSE$	0.07±0.02	0.08±0.02	0.07±0.02	0.08±0.02
F_1	0.88±0.01	0.84±0.02	0.8±0.03	0.75±0.05
F_2	0.89±0.02	0.85±0.02	0.82±0.04	0.75±0.06
HSS	0.81±0.02	0.76±0.04	0.7±0.06	0.61±0.09

Exemplary imputation results for different gap widths are shown in Fig. 4. Even for longer gaps (50 min and arbitrary-length cases), the reconstructed center values capture the observed rainfall dynamics well, illustrating that the quantile-stacked boosting approach maintains accuracy across a wide range of missing-data durations. Taken together, these results demonstrate that the proposed framework provides state-of-the-art imputation performance at 10-minute resolution while remaining computationally compact and robust across heterogeneous network conditions.

5.2 Spatial dependencies and radar-proximity effects

Performance variability across stations shows distinct but interpretable spatial patterns (Fig. 6). While the majority of stations achieve $R^2 > 0.6$, a subset exhibits reduced accuracy, occasionally dropping below 0.5. This reduction correlates with the distance between the gauge and its nearest radar site (Fig. 5), although radar proximity alone does not fully explain the

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

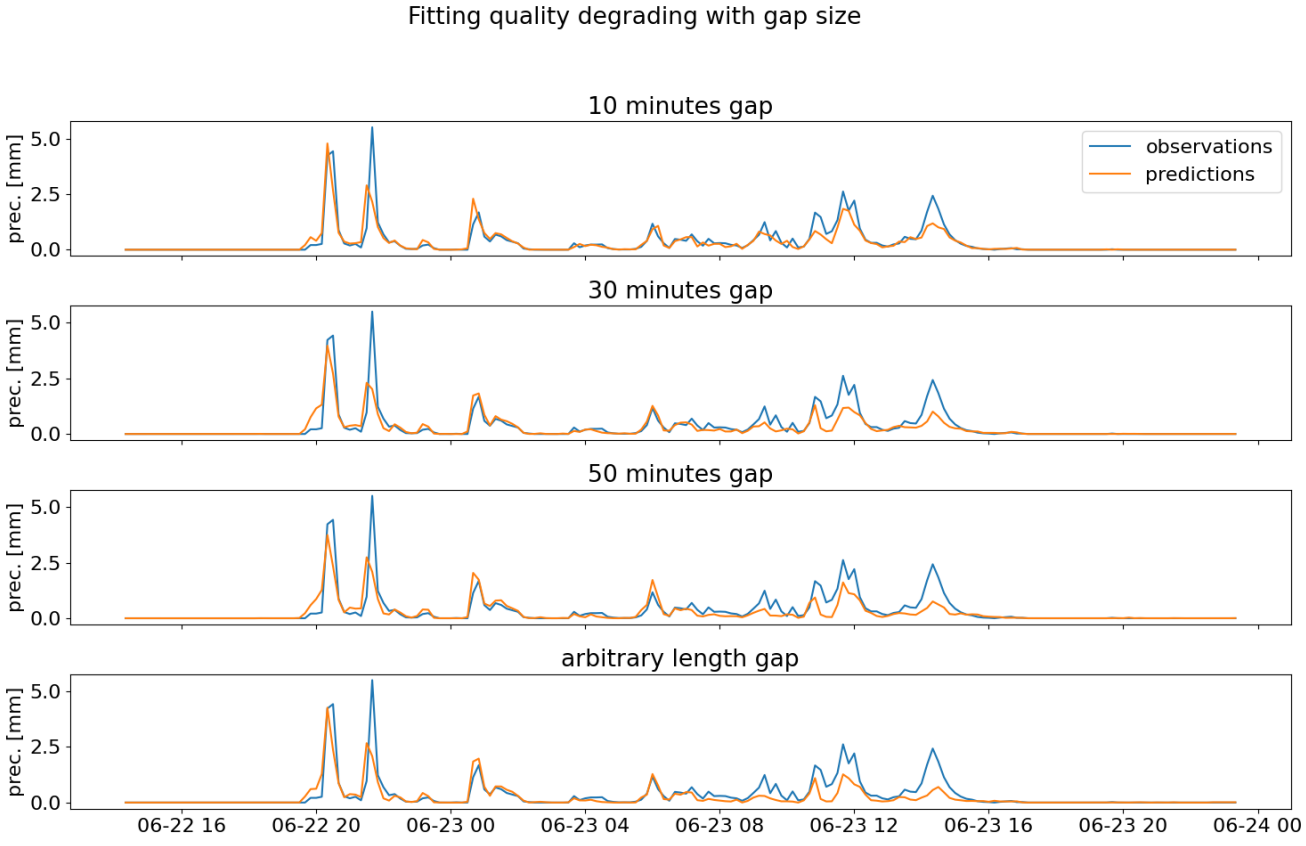


Figure 4. From top to bottom: Exemplary predictions for the center value in synthetically generated gaps of width $w = 10$ min, 30 min, and 50 min, and arbitrary length gaps. Blue shows the observed 10-minute gauge time series (ground truth), while red represents the timeseries obtained by lining up the model predictions for the *central timestamp* in any of the sliding, masking windows. The measurements were obtained from station *3987* (Potsdam, Germany).

spread. Stations located near radar coverage centres generally exhibit more stable performance, suggesting that the spatial representativeness of QPECommon decreases gradually with range and increasingly complex beam geometry.

These findings indicate that the accuracy of radar-derived predictors – particularly from QPECommon – decays smoothly with radar distance, likely due to beam elevation, attenuation, shielding by local topography, and small spatial displacements. These findings indicate that the accuracy of radar-derived predictors – particularly from QPECommon – decays smoothly with radar distance, likely due to beam elevation, attenuation, shielding by local topography, and small spatial displacements. An additional contributor is the range- and azimuth-dependent sampling height of the radar beam relative to station elevation: as beam height increases with distance (and local horizon/terrain effects), the radar increasingly samples precipitation aloft, which can differ from surface rainfall, especially in complex orographic settings. Nevertheless, even stations located in peripheral or complex-beam regions retain acceptable performance ($R^2 \approx 0.5$), underscoring the robustness of the station-specific compact-learning architecture.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

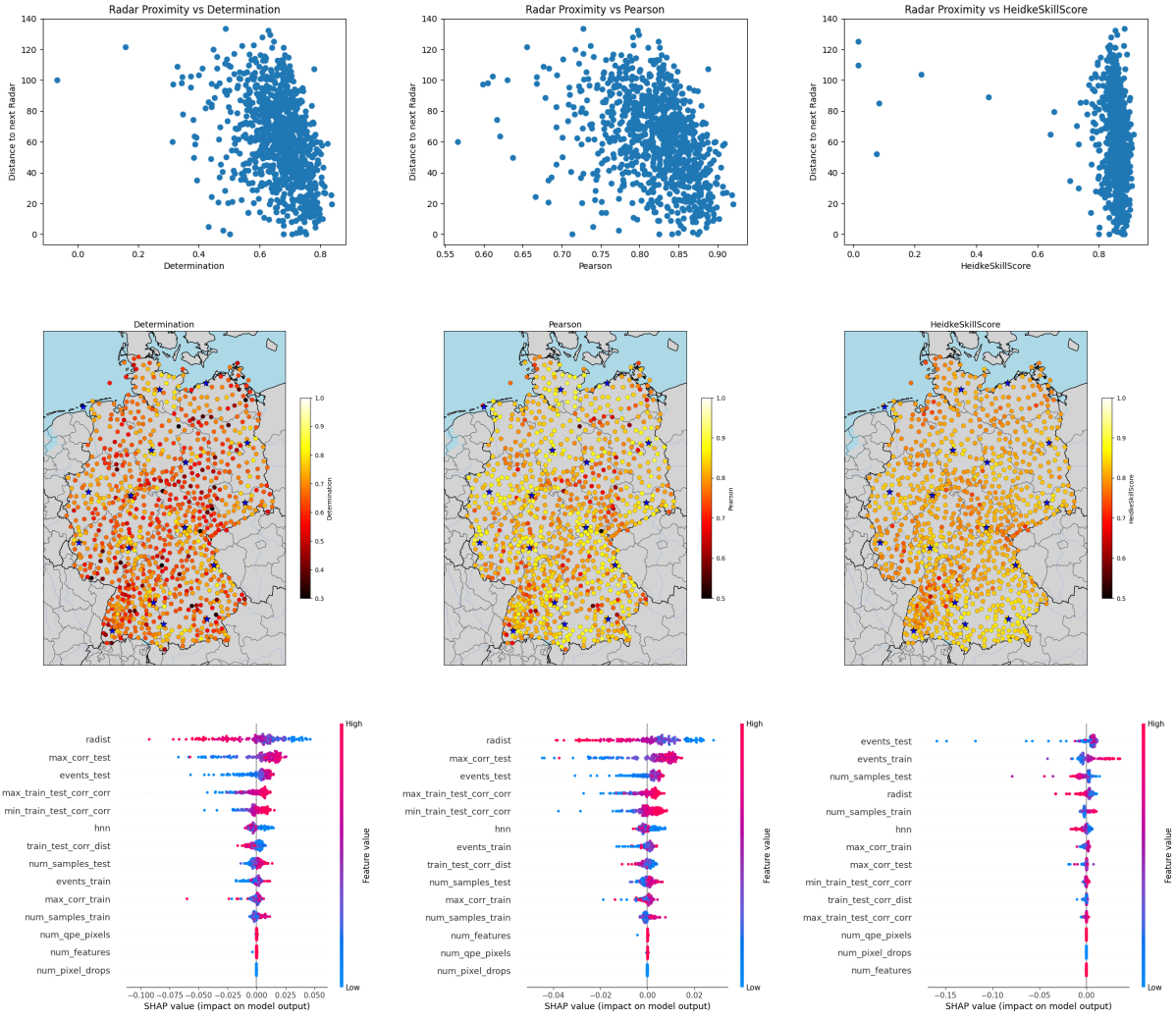
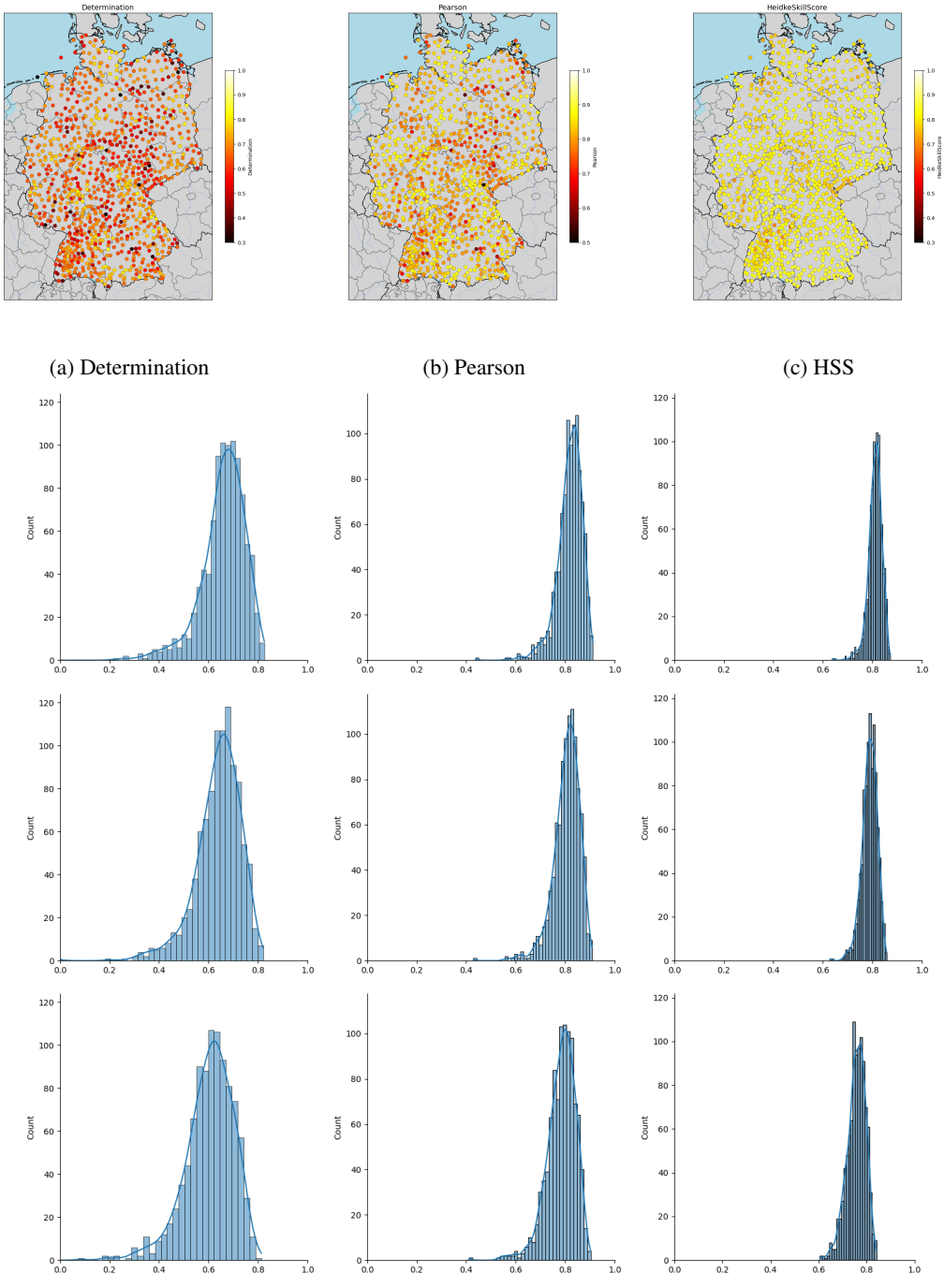


Figure 5. Relationship between radar proximity and model performance across the DWD gauge network.

5.3 Uncertainty and prediction intervals

Because the quantile-stacking architecture naturally provides distributional estimates, prediction intervals can be derived directly from the ensemble output. Figure 7 illustrates 95% prediction intervals for example gaps of 10, 30, and 50 min, as well as for arbitrary-width gaps. The intervals reliably encompass the observations for low to moderate rainfall intensities and remain well-calibrated across the different gap lengths. For high-intensity, short-duration events, the intervals occasionally underestimate the upper tail, indicating that extreme-rainfall behaviour is more challenging to capture with the current quantile configuration. This suggests that the quantile regressors could be further optimised – e.g. through iterative stacking, asymmetric weighting of upper quantiles, or dedicated heavy-rain calibration – to enhance interval width and reliability during extremes.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026



(a) Determination (b) Pearson (c) HSS

Figure 6. Evaluation of regression R (a) and r (b) and classification metrics HSS (c) across the DWD network. Columns refer to increasing gap lengths.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

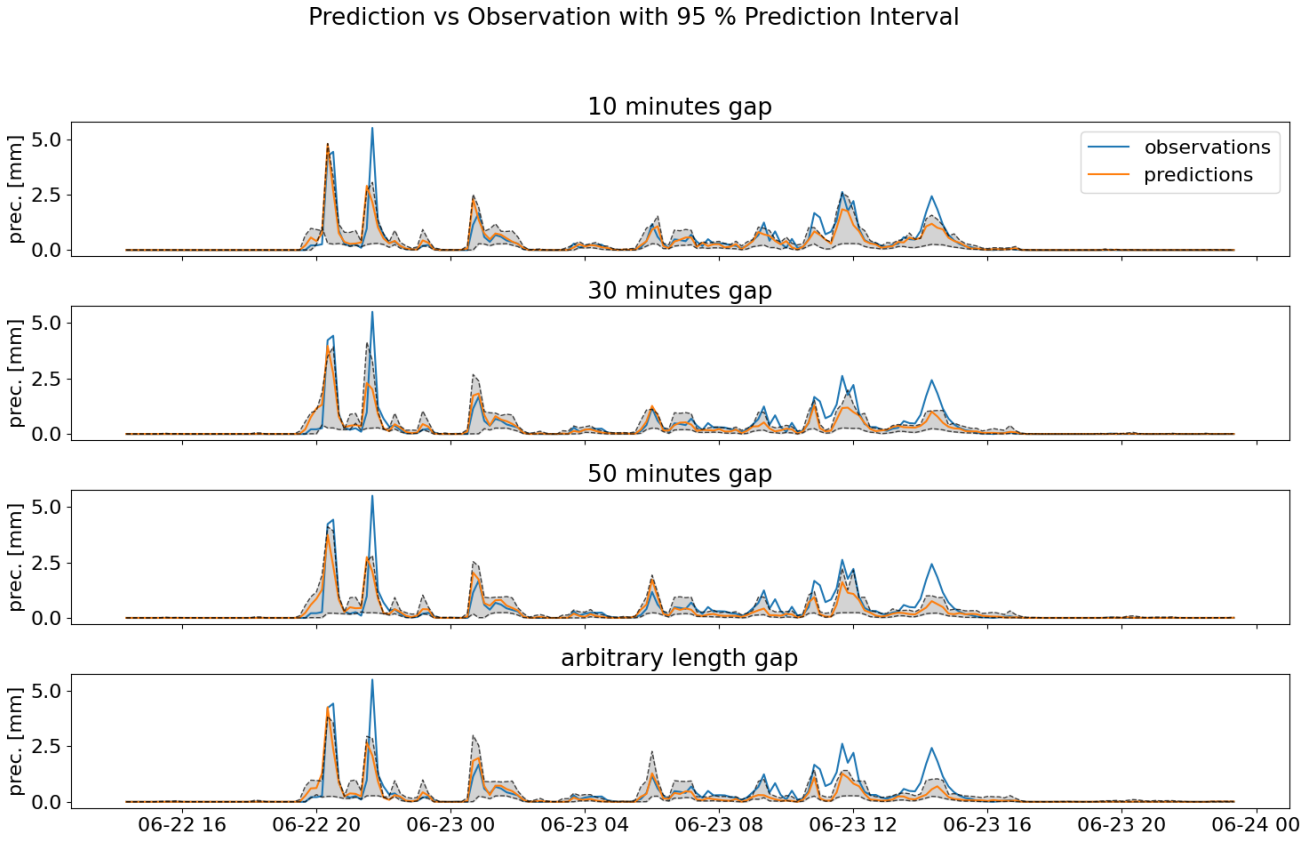


Figure 7. Predictions and estimation of prediction uncertainty. From top to bottom: Exemplary predictions for the center value in synthetically generated gaps of width $w = 10$ min, 30 min, and 50 min, and arbitrary length gaps. Blue shows the observed 10-minute gauge time series (ground truth), while red represents the timeseries obtained by lining up the model predictions for the *central timestamp* in any of the sliding, masking windows. Additionally, a 90-percent confidence interval approximation is displayed. The measurements were obtained from station *3987* (Potsdam, Germany).

5.4 Limitations and improvement potential

Despite robust overall performance, several limitations remain. Capturing highly convective or erratic rainfall events poses challenges, as these are inherently less predictable and often underrepresented in short training periods. Increasing the temporal depth of the training data – e.g. by incorporating additional multi-year archives – would likely improve the model's exposure to rare but influential event types and strengthen its representation of seasonal and convective variability. Future work should also investigate whether the QPECommon radar product itself adequately resolves convective dynamics, or whether complementary radar-based predictors (such as vertical reflectivity gradients, texture measures, or convective indices) could enhance the representation of rapidly evolving storm cells.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Architectural improvements, such as iterative quantile stacking or the inclusion of specialised learners tuned for high-intensity rainfall, could also be explored. Beyond boosted-tree models, convolutional or graph-based neural networks may eventually offer advantages for explicitly structured spatial rainfall fields; however, the simplicity, reproducibility, and low computational cost of the current design remain strong arguments for its operational applicability.

5.5 Applicability and transferability

A key strength of the proposed framework lies in its inherent transferability. Because each station is modelled independently using compact, radar-driven learners, the approach does not rely on any DWD-specific metadata, auxiliary variables, or network geometry. The DWD gauge network therefore acts primarily as a large-scale testbed rather than a structural prerequisite of the method. What matters is the availability of a radar-based QPE product that is not calibrated with the target gauges, ensuring leakage-free predictors during training. For any additional gauge network or observational setup, new station-specific models can be trained using exactly the same workflow.

This property makes the framework directly applicable to any radar-covered region and to a broad spectrum of observational systems, including research-grade precipitation sensors, ecohydrological field stations, agricultural monitoring networks, citizen-science rain gauges, and infrastructure-mounted sensors (e.g. disdrometers, snow pillows, throughfall collectors). Whenever such systems are located within the footprint of a radar composite – such as RADOLAN, MRMS, Combi-Precip, UKGrsHP, or EURADCLIM – station-specific models can be trained in the same way as demonstrated here, provided that the target gauges are not used for gauge adjustment in the composite or that an ungauged intermediate product is available. Note that "ungauged" here refers to the radar input (i.e., not calibrated with the target gauges), not to a site without any ground sensor; in its current station-specific form, the approach still requires a local gauge record (overlap period) for training. The same logic applies to non-DWD gauges in Germany, as long as they are not part of the RADOLAN calibration set. Potential application domains include:

- **Quality control and network maintenance:** Radar-guided imputation provides independent, physics-based references for plausibility checks, fault detection, and automated correction of sensor outages in operational networks. Within modular QA/QC frameworks such as the System for automated Quality Control (SaQC) precipitation measurements are treated as incoming sensor data streams, and station-specific machine-learning models can be executed as additional SaQC processing steps. In this setting, the radar-driven imputed precipitation time series would serve as an external, independent reference signal that stabilises downstream QC workflows and supports reproducible decision rules across large observatory networks (Schmidt et al., 2023; Horsburgh et al., 2025; Bumberger et al., 2025). Radar-based reference information also complements network-centric approaches such as RainGaugeQC and graph neural network classifiers for anomaly detection in commercial microwave-link and soil-moisture networks, thereby improving the automated screening of both professional and opportunistic rain-gauge observations (Ośródka et al., 2025; Lasota et al., 2025).

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

– **Real-time and early-warning systems:** The compact station-specific architecture enables integration into operational chains to provide near-real-time gap filling and stabilised rainfall inputs for environmental Digital Twins and forecasting pipelines. In hydrological and urban-drainage early-warning applications, real-time flood-forecasting systems and digital-twin river basins critically depend on low-latency, high-resolution precipitation forcing to drive coupled hydro – hydraulic models and data-assimilation workflows (Ming et al., 2020; Piadeh et al., 2022; Yang et al., 2024). By turning radar-based QPE streams into locally bias-corrected, station-resolved rainfall series, the framework can supply a continuous hydrometeorological forcing layer for hybrid modelling approaches and Earth-system digital twins that combine big-data processing, machine learning and process-based models (Kraft et al., 2022; Li et al., 2023; Hazeleger et al., 2024).

– **Climate data homogenisation and reanalysis:** Long-term gap filling contributes to more consistent precipitation archives that support trend detection, extreme-rainfall statistics, and regional climate assessments. Recent advances in homogenised rainfall datasets highlight how incomplete and discontinuous observational records can bias climatic indicators and alter the detection of long-term variability (Dumitrescu et al., 2025). At the same time, emerging global synthesis efforts demonstrate the importance of consistent multidecadal hydro-climatic information for constraining large-domain models and advancing reanalysis frameworks (McMillan et al., 2025). Continuous, leakage-free radar-guided imputation can therefore enhance climate-scale datasets in regions where historical gauge archives are sparse or intermittently observed.

– **Ecosystem and ecohydrological research:** Many ecological observatories and experimental infrastructures rely on complete rainfall records to interpret changes in soil moisture, transpiration, nutrient fluxes, or carbon cycling. Radar-driven imputation enables consistent rainfall forcing across heterogeneous sensor networks, lysimeter arrays, and long-term observatory platforms such as TERENO, GCEF, and eLTER, thereby supporting integrated assessments of ecosystem responses to climate and land-use change (Mollenhauer et al., 2018; Schädler et al., 2019; Schnepper et al., 2023; Zacharias et al., 2024; Li et al., 2024; Ohnemus et al., 2025; Huang et al., 2025).

– **Hydrological modelling and flood forecasting, and water-resources management:** Gap-free high-frequency rainfall inputs improve runoff simulations, flash-flood prediction, and data assimilation in hydrological models, especially in sparsely instrumented catchments, and complement existing radar–rain-gauge frameworks for real-time flood forecasting and urban flash-flood prediction (Rafieeinasab et al., 2015; Gabriele et al., 2017; Ming et al., 2020; Piadeh et al., 2022). In addition, they provide a more consistent basis for eco-hydrological flow management, reservoir operation, and event-scale groundwater-recharge estimation in water-scarce regions (Tashie et al., 2016; Yu et al., 2016).

– **Agricultural meteorology and irrigation management:** Gap-free precipitation series support event-based runoff simulations and early-warning systems for fluvial and flash floods, improve short-term reservoir inflow forecasts and operational storage decisions, and provide a more consistent basis for groundwater-recharge estimation and drought-risk assessments in water-scarce regions (Dari et al., 2025; Lagasio et al., 2025).

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

– **Urban hydrology and drainage control:** High-temporal-resolution reconstruction helps detect localised convective events and stabilise rainfall forcing for sewer-system models, thereby supporting combined-sewer-overflow control, pluvial-flood early warning, and stormwater digital twins for real-time drainage management (Thorndahl et al., 2017; Piadeh et al., 2022; Kim et al., 2025; Ge and Qin, 2025; Hlal et al., 2025).

Together, these examples demonstrate that the method is not confined to the DWD network but constitutes a general, radar-based imputation strategy suitable for any precipitation monitoring network exposed to a consistent QPE product. The modular, station-specific architecture is precisely what enables this portability: new learners can be trained for additional stations and networks without altering the overall framework, as long as a suitable, leakage-free radar composite is available for the respective domain.

6 Conclusion

This study presented a scalable, radar-driven framework for imputing missing high-frequency precipitation data using compact gradient-boosting models. By relying exclusively on radar-derived quantitative precipitation estimates (QPE) as predictors, the approach eliminates dependence on neighbouring gauges or additional meteorological variables. This radar-only setup prevents circular validation and isolates the incremental value of QPE for imputation. Using the DWD's ungauged QPECommon product ensured a strictly leakage-free separation between predictor and target domains, enabling statistically independent training and validation.

Across the 1100 DWD stations, the framework achieved mean correlation coefficients of $r \approx 0.8$ and R^2 values between 0.55 and 0.7, with stable performance for gap lengths of up to one hour. These results demonstrate that sub-hourly precipitation series can be reliably reconstructed from radar-based areal estimates using lightweight, station-specific learners. Performance declines with increasing gap length were moderate, and the models retained predictive skill even in regions of weaker radar coverage, highlighting both spatial robustness and computational efficiency. A weak but systematic dependence on radar range was observed, yet predictive skill remained acceptable across the network.

The architecture's quantile-stacking design further allowed the derivation of probabilistic prediction intervals, providing a first step towards uncertainty-aware gap filling. While the current configuration slightly underestimates extreme rainfall events, iterative stacking or specialised learners for high-intensity regimes may improve calibration in future work. Extending the training base across multiple years is expected to strengthen the representation of seasonal and convective variability.

A key operational advantage lies in the modular and resource-efficient design: each station is represented by an independent, easily retrainable learner that can be updated without reprocessing the entire network. This makes the framework directly applicable beyond the DWD network, including research observatories, agricultural and ecohydrological monitoring sites, citizen-science networks, and any other sensor systems located within the footprint of a suitable radar product. In such appli-

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

cations, new station-specific models can be trained on the respective radar-gauge records using the same workflow, and the methodology is compatible with a wide range of radar composites, including gauge-adjusted products, provided that predictor and target data are kept statistically independent to avoid calibration leakage.

In summary, the proposed method bridges radar-based quantitative precipitation estimation and gauge-based data completion by providing an efficient, reproducible, and radar-only imputation framework. It enables consistent, high-frequency precipitation records essential for hydrological modelling, climate analyses, and operational monitoring. Future developments will focus on enlarging the temporal domain, refining uncertainty estimates, and assessing cross-product transferability across diverse radar datasets and climatic regimes.

Code and data availability. Data(Lünenschloß et al., 2025a) und codebasis(Lünenschloß et al., 2025b) implementing model construction and training is published and available through Zenodo.

Author contributions. Peter Lünenschloß, Antje Claussnitzer, Thomas Schartner, Mirjam Brunner, Timo Houben, David Schäfer, and Jan Bumberger conceived the study. Peter Lünenschloß, Mirjam Brunner, Timo Houben, David Schäfer, and Jan Bumberger designed the methodology. Antje Claussnitzer and Thomas Schartner collected and curated the data. Peter Lünenschloß and Mirjam Brunner developed and implemented the computational approach. Peter Lünenschloß and Jan Bumberger led the writing of the manuscript. All authors contributed critically to the drafts and approved the final version for publication.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by a contract research project issued by the German Weather Service (DWD). In the final stage of the review process of this work, the authors used OpenAIs ChatGPT in order to improve the legibility of individual phrases. After using this tool, the authors reviewed and edited the content as required and take full responsibility for the content of the publication.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

References

Abdillah, W., Fauziati, S., and Pratama, A. R.: Utilization of Machine Learning Approaches for Rainfall Data Imputation: A Systematic Literature Review, in: 2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 313–318, https://doi.org/10.1109/IC3INA60834.2023.10285764, 2023.

Adhikary, S. K., Muttil, N., and Yilmaz, A. G.: Genetic Programming-Based Ordinary Kriging for Spatial Interpolation of Rainfall, Journal of Hydrologic Engineering, 21, 04015 062, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001300, 2016.

Alejo-Sanchez, L. E., Márquez-Grajales, A., Salas-Martínez, F., Franco-Arcega, A., López-Morales, V., Acevedo-Sandoval, O. A., González-Ramírez, C. A., and Villegas-Vega, R.: Missing data imputation of climate time series: A review, MethodsX, 15, 103 455, https://doi.org/10.1016/j.mex.2025.103455, 2025.

Bagirov, A. M., Mahmood, A., and Barton, A.: Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach, Atmospheric Research, 188, 20–29, https://doi.org/10.1016/j.atmosres.2017.01.003, 2017.

Bellido-Jiménez, J. A., Gualda, J. E., and García-Marín, A. P.: Assessing Machine Learning Models for Gap Filling Daily Rainfall Series in a Semiarid Region of Spain, Atmosphere, 12, https://doi.org/10.3390/atmos12091158, 2021.

Bumberger, J., Abbrent, M., Brinckmann, N., Hemmen, J., Kunkel, R., Lorenz, C., Lünenschloss, P., Palm, B., Schnicke, T., Schulz, C., van der Schaaf, H., and Schäfer, D.: Digital ecosystem for FAIR time series data management in environmental system science, SoftwareX, 29, 102 038, https://doi.org/10.1016/j.softx.2025.102038, 2025.

Chen, F.-W. and Liu, C.-W.: Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan, Paddy and Water Environment, 10, 209–222, https://doi.org/10.1007/s10333-012-0319-1, 2012.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, https://doi.org/10.1145/2939672.2939785, 2016.

Chivers, B. D., Wallbank, J., Cole, S. J., Sebek, O., Stanley, S., Fry, M., and Leontidis, G.: Imputation of missing sub-hourly precipitation data in a large sensor network: A machine learning approach, Journal of Hydrology, 588, https://doi.org/10.1016/j.jhydrol.2020.125126, 2020.

Chutsagulprom, N., Chaisee, K., Wongsaijai, B., Oonariya, C., and et al.: Spatial interpolation methods for estimating monthly rainfall distribution in Thailand, Theoretical and Applied Climatology, 148, 317–328, https://doi.org/10.1007/s00704-022-03927-7, 2022.

Dari, J., Lo Presti, S., and Brocca, L.: Irrigation monitoring from satellite at hyper-high resolution: Paving the way for remote-sensing-based agricultural water management support services, Agricultural Water Management, 317, 109 627, https://doi.org/10.1016/j.agwat.2025.109627, 2025.

DeGaetano, A. T. and Wilks, D. S.: Radar-guided interpolation of climatological precipitation data, International Journal of Climatology, 29, 185–196, https://doi.org/10.1002/joc.1714, 2009.

Deutscher Wetterdienst (DWD): Radarniederschlag. Prinzip der Niederschlagsbestimmung mit Radar inklusive Umrechnung der Radarreflektivitäten in Momentanwerte des Niederschlages, Tech. rep., https://www.dwd.de/DE/leistungen/radarniederschlag/rn_info/download_niederschlagsbestimmung.pdf, [in German], 2015.

Deutscher Wetterdienst (DWD): Leistungsangebot der Kalibrierlaboratorien des Deutschen Wetterdienstes (DG1540 DW14), Tech. rep., https://www.dwd.de/DE/derdwd/messnetz/kalibrierlaboratorien/kalibrierlaboratorien_dwd.pdf, [in German], 2025.

Dumitrescu, A., Micu, D., Guijarro, J., Manea, A., and Cheval, S.: Long-term homogenized air temperature and precipitation datasets in Romania, 1901–2023, Scientific Data, 12, 1116, https://doi.org/10.1038/s41597-025-05371-4, 2025.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Erdin, R., Frei, C., and Künsch, H. R.: Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges, Journal of Hydrometeorology, 13, 1332 – 1346, https://doi.org/10.1175/JHM-D-11-096.1, 2012.

Faramarzzadeh, M., Ehsani, M. R., Akbari, M., et al.: Application of Machine Learning and Remote Sensing for Gap-filling Daily Precipitation Data of a Sparsely Gauged Basin in East Africa, Environmental Processes, 10, 8, https://doi.org/10.1007/s40710-023-00625-y, 2023.

Gabriele, S., Chiaravalloti, F., and Procopio, A.: Radar–rain-gauge rainfall estimation for hydrological applications in small catchments, Advances in Geosciences, 44, 61–66, https://doi.org/10.5194/adgeo-44-61-2017, 2017.

Ge, C. and Qin, S.: Urban flooding digital twin system framework, Systems Science & Control Engineering, 13, 2460432, https://doi.org/10.1080/21642583.2025.2460432, 2025.

Gorshenin, A., Lebedeva, M., Lukina, S., and Yakovleva, A.: Application of Machine Learning Algorithms to Handle Missing Values in Precipitation Data, in: Distributed Computer and Communication Networks - DCCN 2019, Lecture Notes in Computer Science, Springer, https://doi.org/10.1007/978-3-030-36614-8_43, 2019.

Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, Journal of Hydrology, 332, 144–157, https://doi.org/10.1016/j.jhydrol.2006.06.028, 2007.

Hazeleger, W., Aerts, J. P. M., Bauer, P., Bierkens, M. F. P., Camps-Valls, G., Dekker, M. M., Doblas-Reyes, F. J., Eyring, V., Finkenauer, C., Grundner, A., Hachinger, S., Hall, D. M., Hartmann, T., Iglesias-Suarez, F., Janssens, M., Jones, E. R., Kölling, T., Lees, M., Lhermitte, S., van Nieuwpoort, R. V., Pahker, A.-K., Pellicer-Valero, O. J., Pijpers, F. P., Siibak, A., Spitzer, J., Stevens, B., Vasconcelos, V. V., and Vossepoel, F. C.: Digital twins of the Earth with and for humans, Communications Earth Environment, 5, 463, https://doi.org/10.1038/s43247-024-01626-x, 2024.

Hlal, M., Baraka Munyaka, J.-C., Chenal, J., Azmi, R., Diop, E. B., Bounabi, M., Ebnou Abdem, S. A., Almouctar, M. A. S., and Adraoui, M.: Digital Twin Technology for Urban Flood Risk Management: A Systematic Review of Remote Sensing Applications and Early Warning Systems, Remote Sensing, 17, https://doi.org/10.3390/rs17173104, 2025.

Horsburgh, J. S., Lippold, K., Slaugh, D. L., and Ramirez, M.: HydroServer: A software stack supporting collection, communication, storage, management, and sharing of data from in situ environmental sensors, Environmental Modelling & Software, 193, 106637, https://doi.org/https://doi.org/10.1016/j.envsoft.2025.106637, 2025.

Huang, J., Sehgal, V., Alvarez, L. V., Brocca, L., Cai, S., Cheng, R., Cheng, X., Du, J., El Masri, B., Endsley, K. A., Fang, Y., Hu, J., Jampani, M., Kibria, M. G., Koren, G., Li, L., Liu, L., Mao, J., Moreno, H. A., Rigden, A., Shi, M., Shi, X., Wang, Y., Zhang, X., and Fisher, J. B.: Remotely Sensed High-Resolution Soil Moisture and Evapotranspiration: Bridging the Gap Between Science and Society, Water Resources Research, 61, e2024WR037929, https://doi.org/https://doi.org/10.1029/2024WR037929, 2025.

Huang, M., Lin, R., Huang, S., and Xing, T.: A novel approach for precipitation forecast via improved K-nearest neighbor algorithm, Advanced Engineering Informatics, 33, 89–95, https://doi.org/10.1016/j.aei.2017.05.003, 2017.

Hırca, T. and Eryılmaz Türkkan, G.: Assessment of Different Methods for Estimation of Missing Rainfall Data, Water Resources Management, 38, 5945–5972, https://doi.org/10.1007/s11269-024-03936-3, 2024.

Kim, J. and Ryu, J. H.: Quantifying a Threshold of Missing Values for Gap Filling Processes in Daily Precipitation Series, Water Resources Management, 29, 4173–4184, https://doi.org/10.1007/s11269-015-1052-5, 2015.

Kim, Y. and Hong, S.: Very Short-Term Rainfall Prediction Using Ground Radar Observations and Conditional Generative Adversarial Networks, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–8, https://doi.org/10.1109/TGRS.2021.3108812, 2022.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Kim, Y., Oh, J., and Bartos, M.: Stormwater digital twin with online quality control detects urban flood hazards under uncertainty, Sustainable Cities and Society, 118, 105 982, https://doi.org/10.1016/j.scs.2024.105982, 2025.

Kraft, B., Jung, M., Körner, M., Koirala, S., and Reichstein, M.: Towards hybrid modeling of the global hydrological cycle, Hydrology and Earth System Sciences, 26, 1579–1614, https://doi.org/10.5194/hess-26-1579-2022, 2022.

Lagasio, M., Barindelli, S., Chitu, Z., Contreras, S., Fernández-Rodríguez, A., de Klerk, M., Fumagalli, A., Gatti, A., Hammerschmidt, L., Haskovic, D., Milelli, M., Oberto, E., Ontel, I., Orensanz, J., Ramelli, F., Uboldi, F., Validi, A., and Realini, E.: Integrating Advanced Sensor Technologies for Enhanced Agricultural Weather Forecasts and Irrigation Advisories: The MAGDA Project Approach, Remote Sensing, 17, https://doi.org/10.3390/rs17111855, 2025.

Lasota, E., Houben, T., Polz, J., Schmidt, L., Glawion, L., Schäfer, D., Bumberger, J., and Chwala, C.: Interpretable Quality Control of Sparsely Distributed Environmental Sensor Networks Using Graph Neural Networks, Artificial Intelligence for the Earth Systems, 4, e240 032, https://doi.org/10.1175/AIES-D-24-0032.1, 2025.

Li, X., Feng, M., Ran, Y., Su, Y., Liu, F., Huang, C., Shen, H., Xiao, Q., Su, J., Yuan, S., and Guo, H.: Big Data in Earth system science and progress towards a digital twin, Nature Reviews Earth & Environment, 4, 319–332, https://doi.org/10.1038/s43017-023-00409-w, 2023.

Li, X., Xia, K., Wu, T., Wang, S., Tang, H., Xiao, C., Tang, H., Xu, N., and Jia, D.: Increased precipitation has not enhanced the carbon sequestration of afforestation in Northwest China, Communications Earth & Environment, 5, 619, https://doi.org/10.1038/s43247-024-01733-9, 2024.

Lünenschloß, P., Claussnitzer, A., Schartner, T., Brunner, M., Houben, T., Schäfer, D., and Bumberger, J.: Scalable radar-driven approach with compact gradient-boosting models for gap filling in high-resolution precipitation measurements, https://doi.org/10.5281/zenodo.17937464, [Data set], 2025a.

Lünenschloß, P., Claussnitzer, A., Schartner, T., Brunner, M., Houben, T., Schäfer, D., and Bumberger, J.: Scalable radar-driven approach with compact gradient-boosting models for gap filling in high-resolution precipitation measurements, https://doi.org/10.5281/zenodo.17940311, [Software], 2025b.

Lünenschloß, P., Schäfer, D., Gransee, F., Claußnitzer, A., Schartner, T., and Bumberger, J.: ML Driven Imputation of Precipitation Data Collected at High Sampling Rates, in: EMS Annual Meeting 2022, https://doi.org/10.5194/ems2022-467, 2022.

McKee, J. L. and Binns, A. D.: A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41, 186–203, https://doi.org/10.1080/07011784.2015.1064786, 2016.

McMillan, H., Araki, R., Bolotin, L., Kim, D.-H., Coxon, G., Clark, M., and Seibert, J.: Global patterns in observed hydrologic processes, Nature Water, 3, 497–506, https://doi.org/10.1038/s44221-025-00407-w, 2025.

Meuer, J., Bouwer, L. M., Kaspar, F., Lehmann, R., Karl, W., Ludwig, T., and Kadow, C.: Infilling of missing rainfall radar data with a memory-assisted deep learning approach, Hydrology and Earth System Sciences, 29, 3687–3701, https://doi.org/10.5194/hess-29-3687-2025, 2025.

Ming, X., Liang, Q., Xia, X., Li, D., and Fowler, H. J.: Real-Time Flood Forecasting Based on a High-Performance 2-D Hydrodynamic Model and Numerical Weather Predictions, Water Resources Research, 56, e2019WR025 583, https://doi.org/10.1029/2019WR025583, 2020.

Mital, U., Dipankar, D., Brown, J. B., Faybishenko, B., Painter, S. L., and Steefel, C. I.: Sequential imputation of missing spatio-temporal precipitation data using random forests, Frontiers in Water, 2, https://doi.org/10.3389/frwa.2020.00020, 2020.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Mollenhauer, H., Kasner, M., Haase, P., Peterseil, J., Wohner, C., Frenzel, M., Mirtl, M., Schima, R., Bumberger, J., and Zacharias, S.: Long-term environmental monitoring infrastructures in Europe: observations, measurements, scales, and socio-ecological representativeness, Science of The Total Environment, 624, 968–978, https://doi.org/10.1016/j.scitotenv.2017.12.095, 2018.

Mott, M. and Schultze, M.: Kompositerstellung und Verlagerung, promet - Meteorologische Fortbildung, 107, https://doi.org/10.5676/DWD_pub/promet_107_03, [in German], 2024.

Ohnemus, T., Dirnböck, T., Bäck, J., Gaube, V., Kühn, I., Mirtl, M., Mollenhauer, H., Vereecken, H., and Zacharias, S.: Fitness for future: eLTER RI's representation of climate and land use change, Ecological Indicators, 171, 113 159, https://doi.org/10.1016/j.ecolind.2025.113159, 2025.

Ośródka, K., Szturc, J., Jurczyk, A., and Kurcz, A.: Adaptation of RainGaugeQC algorithms for quality control of rain gauge data from professional and non-professional measurement networks, Atmospheric Measurement Techniques, 18, 3229–3245, https://doi.org/10.5194/amt-18-3229-2025, 2025.

Overeem, A., van den Besselaar, E., van der Schrier, G., Meirink, J. F., van der Plas, E., and Leijnse, H.: EURADCLIM: the European climatological high-resolution gauge-adjusted radar precipitation dataset, Earth System Science Data, 15, 1441–1464, https://doi.org/10.5194/essd-15-1441-2023, 2023.

Overeem, A., Leijnse, H., Veldhuizen, M., and Anker, B.: The Dutch real-time gauge-adjusted radar precipitation product, Earth System Science Data, 17, 4715–4736, https://doi.org/10.5194/essd-17-4715-2025, 2025.

Paulat, M., Frei, C., Hagen, M., and Wernli, H.: A gridded dataset of hourly precipitation in Germany: Its construction, climatology and application, Meteorologische Zeitschrift, 17, 719–732, https://doi.org/10.1127/0941-2948/2008/0332, 2008.

Piadeh, F., Behzadian, K., and Alani, A. M.: A critical review of real-time modelling of flood forecasting in urban drainage systems, Journal of Hydrology, 607, 127 476, https://doi.org/10.1016/j.jhydrol.2022.127476, 2022.

Portuguez-Maurtua, M., Arumi, J. L., Lagos, O., Stehr, A., and Montalvo Arquiñigo, N.: Filling Gaps in Daily Precipitation Series Using Regression and Machine Learning in Inter-Andean Watersheds, Water, 14, https://doi.org/10.3390/w14111799, 2022.

Qiu, H., Chen, H., Xu, B., Liu, G., Huang, S., Nie, H., and Xie, H.: Multiple Types of Missing Precipitation Data Filling Based on Ensemble Artificial Intelligence Models, Water, 16, 3192, https://doi.org/10.3390/w16223192, 2024.

Qiu, Q., Liu, J., Tian, J., Jiao, Y., Li, C., Wang, W., and Yu, F.: Evaluation of the Radar QPE and Rain Gauge Data Merging Methods in Northern China, Remote Sensing, 12, https://doi.org/10.3390/rs12030363, 2020.

Rabiei, E. and Haberlandt, U.: Applying bias correction for merging rain gauge and radar data, Journal of Hydrology, 522, 544–557, https://doi.org/10.1016/j.jhydrol.2015.01.020, 2015.

Rafieeinasab, A., Norouzi, A., Kim, S., Habibi, H., Nazari, B., Seo, D.-J., Lee, H., Cosgrove, B., and Cui, Z.: Toward high-resolution flash flood prediction in large urban areas – Analysis of sensitivity to spatiotemporal resolution of rainfall input and hydrologic modeling, Journal of Hydrology, 531, 370–388, https://doi.org/10.1016/j.jhydrol.2015.08.045, 2015.

Ravuri, S., Lenc, K., Willson, M., et al.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672–677, https://doi.org/10.1038/s41586-021-03854-z, 2021.

Rodrigues, D. T., Gonçalves, W. A., Silva, C. M. S. E., Spyrides, M. H. C., and Lúcio, P. S.: Imputation of precipitation data in northeast Brazil, Anais da Academia Brasileira de Ciencias, 95, e20210 737, https://doi.org/10.1590/0001-3765202320210737, 2023.

Sattari, M.-T., Rezazadeh-Joudi, A., and Kusiak, A.: Assessment of different methods for estimation of missing data in precipitation studies, Hydrology Research, 48, 1032–1044, https://doi.org/10.2166/nh.2016.364, 2017.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Schmidt, L., Schäfer, D., Geller, J., Lünenschloss, P., Palm, B., Rinke, K., Rebmann, C., Rode, M., and Bumberger, J.: System for automated Quality Control (SaQC) to enable traceable and reproducible data streams in environmental science, Environmental Modelling Software, 169, 105 809, https://doi.org/10.1016/j.envsoft.2023.105809, 2023.

Schnepper, T., Groh, J., Gerke, H. H., Reichert, B., and Pütz, T.: Evaluation of precipitation measurement methods using data from a precision lysimeter network, Hydrology and Earth System Sciences, 27, 3265–3292, https://doi.org/10.5194/hess-27-3265-2023, 2023.

Schädler, M., Buscot, F., Klotz, S., Reitz, T., Durka, W., Bumberger, J., Merbach, I., Michalski, S. G., Kirsch, K., Remmler, P., Schulz, E., and Auge, H.: Investigating the consequences of climate change under different land-use regimes: a novel experimental infrastructure, Ecosphere, 10, e02 635, https://doi.org/10.1002/ecs2.2635, 2019.

Setya, B., Nurhidayatullah, R. A., Hewen, M. B., and Kusrini, K.: Comparative Analysis Of Rainfall Value Prediction In Semarang Using Linear And K-Nearest Neighbor Algorithms, in: 2023 5th International Conference on Cybernetics and Intelligent System (ICORIS), pp. 1–5, https://doi.org/10.1109/ICORIS60118.2023.10352274, 2023.

Sharma, A., Khanna, A., Bhargava, M., and Pendse, R.: Rainfall Prediction: Analysis of Machine Learning Algorithms and Ensemble Techniques, in: 2021 7th International Conference on Signal Processing and Communication (ICSC), pp. 234–240, https://doi.org/10.1109/ICSC53193.2021.9673275, 2021.

Sideris, I. V., Gabella, M., Erdin, R., and Germann, U.: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland, Quarterly Journal of the Royal Meteorological Society, 140, 1097–1111, https://doi.org/10.1002/qj.2188, 2014.

Simolo, C., Brunetti, M., Maugeri, M., and Nanni, T.: Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach, International Journal of Climatology, 30, 1564–1576, https://doi.org/10.1002/joc.1992, 2010.

Steinert, J., Tracksdorf, P., and Heizenreder, D.: Hymec: Surface Precipitation Type Estimation at the German Weather Service, Weather and Forecasting, 36, 1611 – 1627, https://doi.org/10.1175/WAF-D-20-0232.1, 2021.

Tang, G., Clark, M. P., and Papalexiou, S. M.: The Use of Serially Complete Station Data to Improve the Temporal Continuity of Gridded Precipitation and Temperature Estimates, Journal of Hydrometeorology, 22, 1553 – 1568, https://doi.org/10.1175/JHM-D-20-0313.1, 2021.

Tashie, A. M., Mirus, B. B., and Pavelsky, T. M.: Identifying long-term empirical relationships between storm characteristics and episodic groundwater recharge, Water Resources Research, 52, 21–35, https://doi.org/10.1002/2015WR017876, 2016.

Tavares, C. d. M. G., Armond, N. B., Ferreira, C. d. C. M., and Guerra, A. J. T.: Filling gaps of daily precipitation data at rain gauges in a tropical mountainous region: a case study of the municipality of petrópolis, rj, Brazil, Theoretical and Applied Climatology, 156, 178, https://doi.org/10.1007/s00704-025-05397-z, 2025.

Teegavarapu, R. S. and Chandramouli, V.: Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records, Journal of Hydrology, 312, 191–206, https://doi.org/10.1016/j.jhydrol.2005.02.015, 2005.

Teegavarapu, R. S. V., Aly, A., Pathak, C. S., Ahlquist, J., Fuelberg, H., and Hood, J.: Infilling missing precipitation records using variants of spatial interpolation and data-driven methods: use of optimal weighting parameters and nearest neighbour-based corrections, International Journal of Climatology, 38, 776–793, https://doi.org/10.1002/joc.5209, 2018.

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R., and Molnar, P.: Weather radar rainfall data in urban hydrology, Hydrology and Earth System Sciences, 21, 1359–1380, https://doi.org/10.5194/hess-21-1359-2017, 2017.

https://doi.org/10.5194/egusphere-2025-6349
Preprint. Discussion started: 26 January 2026

Toure, M., Klutse, N. A. B., Sarr, M. A., Bhuiyan, M. A. E., Kenne, A. D., Thiaw, W. M., Badiane, D., Gaye, A. T., Ndiaye, O., and Mbow, C.: Machine learning approaches for imputing missing meteorological data in Senegal, Applied Computing and Geosciences, 27, 100 281, https://doi.org/10.1016/j.acags.2025.100281, 2025.

Trömel, S., Chen, J., Chwala, C., Gottschalk, M., Weigl, E., and Winterrath, T.: Quantitative Niederschlagsschätzung (QPE), promet - Meteorologische Fortbildung, 107, https://doi.org/10.5676/DWD_pub/promet_107_05, [in German], 2024.

Vicente-Serrano, S. M., Saz-Sánchez, M. A., and Cuadrat, J. M.: Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature, Climate Research, 24, 161–180, https://doi.org/10.3354/cr024161, 2003.

Vidal-Paz, J., Rodríguez-Gómez, B. A., and Orosa, J. A.: A Comparison of Different Methods for Rainfall Imputation: A Galician Case Study, Applied Sciences, 13, 12 260, https://doi.org/10.3390/app132212260, 2023.

Wani, O. A., Mahdi, S. S., Yeasin, M., Kumar, S. S., Gagnon, A. S., Danish, F., Al-Ansari, N., El-Hendawy, S., and Mattar, M. A.: Predicting rainfall using machine learning, deep learning, and time series models across an altitudinal gradient in the North-Western Himalayas., Scientific Reports, 14, 27 876, https://doi.org/10.1038/s41598-024-77687-x, 2024.

Winterrath, T., Rosenow, W., and Weigl, E.: On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management, in: Weather Radar and Hydrology: Proceedings of a Symposium held in Exeter, UK, April 2011, edited by Moore, R. J., Cole, S. J., and Illingworth, A. J., no. 351 in IAHS Publication, pp. 323–329, IAHS Press, 2012.

Yang, Y., Xie, C., Fan, Z., Xu, Z., Melville, B. W., Liu, G., and Hong, L.: Digital twinning of river basins towards full-scale, sustainable and equitable water management and disaster mitigation, npj Natural Hazards, 1, 43, https://doi.org/10.1038/s44304-024-00047-2, 2024.

Yu, C., Yin, X., Yang, Z., Cai, Y., and Sun, T.: A shorter time step for eco-friendly reservoir operation does not always produce better water availability and ecosystem benefits, Journal of Hydrology, 540, 900–913, https://doi.org/10.1016/j.jhydrol.2016.07.008, 2016.

Yu, J., Li, X. F., Lewis, E., et al.: UKGrsHP: a UK high-resolution gauge–radar–satellite merged hourly precipitation analysis dataset, Climate Dynamics, 54, 2919–2940, https://doi.org/10.1007/s00382-020-05144-2, 2020.

Zacharias, S., Loescher, H. W., Bogena, H., Kiese, R., Schrön, M., Attinger, S., Blume, T., Borchardt, D., Borg, E., Bumberger, J., Chwala, C., Dietrich, P., Fersch, B., Frenzel, M., Gaillardet, J., Groh, J., Hajnsek, I., Itzerott, S., Kunkel, R., Kunstmann, H., Kunz, M., Liebner, S., Mirtl, M., Montzka, C., Musolff, A., Pütz, T., Rebmann, C., Rinke, K., Rode, M., Sachs, T., Samaniego, L., Schmid, H. P., Vogel, H.-J., Weber, U., Wollschläger, U., and Vereecken, H.: Fifteen Years of Integrated Terrestrial Environmental Observatories (TERENO) in Germany: Functions, Services, and Lessons Learned, Earth's Future, 12, e2024EF004 510, https://doi.org/10.1029/2024EF004510, 2024.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., and Kitzmiller, D.: Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities, Bulletin of the American Meteorological Society, 97, 621 – 638, https://doi.org/10.1175/BAMS-D-14-00174.1, 2016.

Zhang, X., Yan, C., Gao, C., Malin, B., and Chen, Y.: XGBoost Imputation for Time Series Data, in: 2019 IEEE International Conference on Healthcare Informatics (ICHI), pp. 1–3, https://doi.org/10.1109/ICHI.2019.8904666, 2019.

Zolina, O., Simmer, C., Belyaev, K., Kapala, A., and Gulev, S.: Improving Estimates of Heavy and Extreme Precipitation Using Daily Records from European Rain Gauges, Journal of Hydrometeorology, 10, 701 – 716, https://doi.org/10.1175/2008JHM1055.1, 2009.