Near-Real-Time Error Correction of Satellite Precipitation Products Using a U-Net-Based Deep Learning Approach

Hua Wang , Zhenyu Yu * , Hanqing Chen *

Posted Date: 27 February 2026

doi: 10.20944/preprints202602.1763.v1

Keywords: precipitation; IMERG; U-Net; deep learning methods; error

 doi:10.20944/preprints202602.1763.v1

Article

Near-Real-Time Error Correction of Satellite Precipitation Products Using a U-Net-Based Deep Learning Approach

Hua Wang [1], Zhenyu Yu [2,*] and Hanqing Chen [3,*]

[1] School of Business Administration, Zhejiang University of Finance and Economics Dongfang College, Jiaxing 314000, China

[2] Faculty of Computer Science and Information Technology, University Malaya, Kuala Lumpur 50603, Malaysia

[3] College of Ocean Engineering and Energy, Guangdong Ocean University, Zhanjiang 524005, China

* Correspondence: yuzhenyuyxl@foxmail.com (Z.Y.); hanqingchen1007@163.com (H.C.)

Abstract

This study investigates the application of a U-Net deep learning model for real-time error correction of satellite precipitation products, using IMERG-Late data over mainland China during 2015–2017. The performance of U-Net is systematically compared with five widely used machine learning approaches, including Decision Tree Regressor, Gradient Boosting Decision Tree, Multilayer Perceptron, Random Forest, and Extreme Gradient Boosting. The results indicate that U-Net consistently improves overall precipitation accuracy and is particularly effective in reducing false precipitation events, outperforming the comparative methods across most evaluation metrics. However, a relatively lower probability of detection (POD) is observed for U-Net, which is primarily associated with over-correction effects that suppress weak or marginal precipitation signals. Despite this limitation, the overall performance demonstrates that the U-Net-based approach provides a robust and reliable solution for real-time error correction of satellite precipitation products, with clear potential for hydrological and meteorological applications.

Keywords: precipitation; IMERG; U-Net; deep learning methods; error correction

1. Introduction

Satellite precipitation products provide near-global, high-resolution estimates of precipitation in near real time and play a critical role in hydrological applications such as flood forecasting, landslide early warning, and typhoon monitoring [1–4]. Despite these advantages, extensive validation studies have shown that the accuracy of satellite-based precipitation estimates remains limited, particularly under complex terrain and heterogeneous climatic conditions. Large uncertainties have been reported in mountainous regions [5–7], in the detection of snowfall and light precipitation events [8–10], and in the balance between false alarms and missed precipitation across many regions worldwide [11]. These deficiencies constrain the reliability of satellite precipitation products in regional hydrological applications and highlight the need for effective error correction strategies.

In practice, the accuracy of satellite precipitation products is often improved through integration with ground-based observations. However, such approaches are not always feasible in real-time applications due to delays in data availability and the uneven spatial distribution of rain gauges. When near-real-time satellite precipitation products exhibit substantial uncertainties, correction methods that do not rely on concurrent ground observations become particularly important. Existing real-time correction approaches can be broadly categorized into two groups: methods that

 doi:10.20944/preprints202602.1763.v1

incorporate auxiliary variables synchronized with satellite observations (e.g., soil moisture and temperature), and methods that establish statistical relationships between historical satellite estimates and gauge-based precipitation to adjust future real-time products.

For mature satellite precipitation products, commonly used correction techniques include statistical regression approaches [12], error–factor function modeling [13], and methods that exploit satellite-derived soil moisture information for real-time adjustment [14–17]. In recent years, machine learning and deep learning methods have been increasingly applied to precipitation error correction. However, many existing approaches remain limited in their ability to represent the spatial structure of precipitation fields, particularly the distinction between true precipitation events and false detections [18].

From an error composition perspective, satellite precipitation uncertainties arise from a combination of hit errors, missed detections, and false alarms [19]. Previous studies have demonstrated that, over large portions of the globe, false precipitation detections constitute a dominant source of error [13,20]. Nevertheless, most correction methods primarily focus on improving the estimation of correctly detected precipitation events, while their effectiveness in identifying and suppressing false precipitation remains limited. For example, Tian et al. [21] emphasized the correction of systematic biases in precipitation amounts for hit events but reported limited improvement in false precipitation detection.

Motivated by these limitations, this study focuses on the real-time correction of both false precipitation and hit-related errors in satellite precipitation products. By building upon existing correction frameworks, we investigate the potential of a U-Net-based deep learning approach to enhance the reliability of near-real-time satellite precipitation estimates at the regional scale.

2. Materials

2.1. Study Area and Datasets

This study focuses on mainland China, located in the southeastern part of the Eurasian continent (73°41′-135°02′E, 18°10′-53°33′N). The region exhibits pronounced climatic and topographic heterogeneity under the influence of the East Asian monsoon system. Terrain elevation generally decreases from west to east, with the Qinghai–Tibet Plateau representing the most prominent topographic feature. Mountainous areas, plateaus, and hills account for approximately 67% of the total land area, while plains and basins constitute about 33%.

Climatic conditions vary substantially across the region. Eastern and southeastern China are characterized by humid subtropical and temperate monsoon climates, influenced by the western Pacific Ocean. Northwestern China is dominated by arid and semi-arid climates associated with mountainous and basin terrains, whereas southwestern China is largely controlled by the cold and high-altitude climate of the Tibetan Plateau. The coexistence of diverse climatic regimes and complex terrain makes mainland China a representative and challenging region for evaluating satellite precipitation products and their error correction performance (Figure 1).

 doi:10.20944/preprints202602.1763.v1

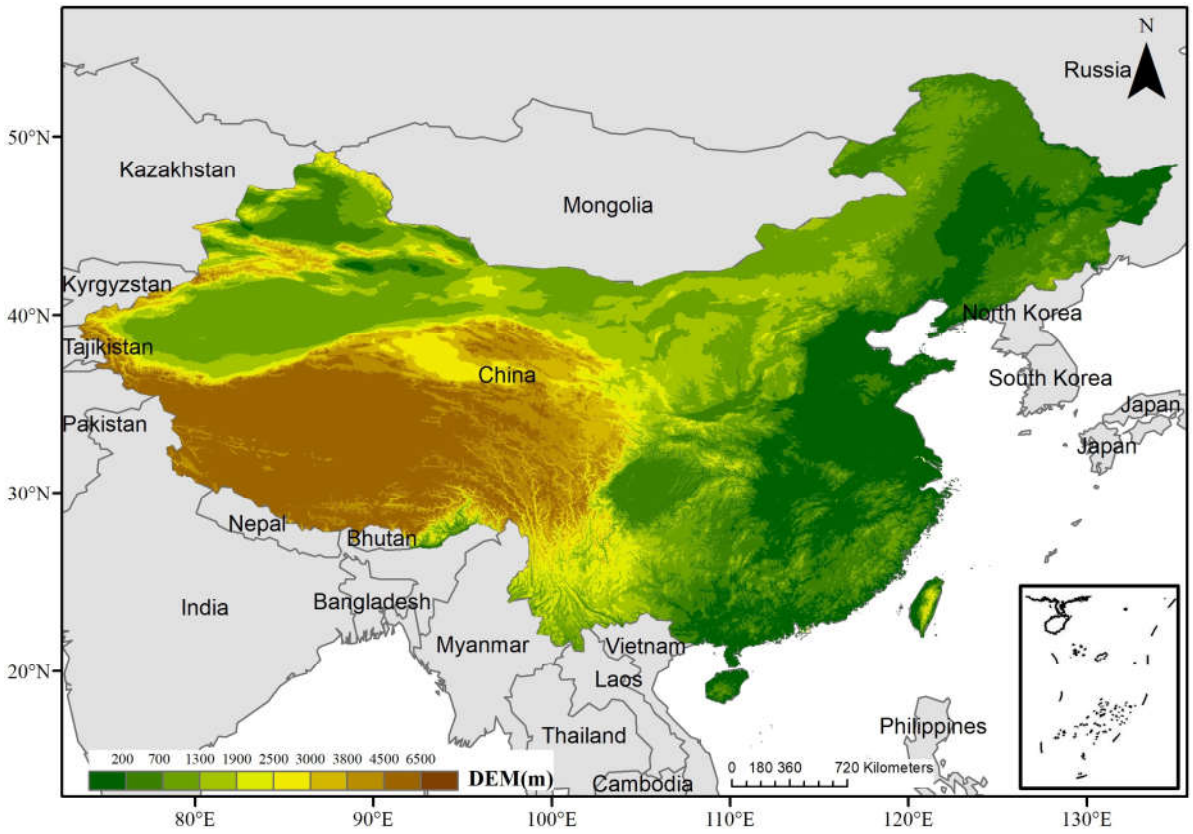

Figure 1. Study area and topography of mainland China.

2.2. Data Source

Two precipitation datasets were employed in this study, including a near-real-time satellite precipitation product and a gauge-based reference dataset. Both datasets were harmonized to the same spatial and temporal resolution to ensure consistency in model training and evaluation. The study period covers 2015-2017, with a spatial resolution of 0.1° and a temporal resolution of 1 hour. To facilitate clarity and reproducibility, the main characteristics of the datasets used in this study are summarized in Table 1.

2.2.1. IMERG-Late Satellite Precipitation

The satellite precipitation data used in this study were obtained from the Integrated Multi-satellitE Retrievals for GPM (IMERG) Late Run product, developed under the Global Precipitation Measurement (GPM) mission. IMERG provides three product versions—Early, Late, and Final—designed to balance latency and accuracy requirements. Among them, IMERG-Late applies forward and backward morphing techniques based on cloud motion vectors and incorporates monthly climatological adjustments, but does not assimilate concurrent ground-based gauge observations. This design makes IMERG-Late particularly suitable for near-real-time hydrological applications, while also leading to systematic and random errors that motivate further correction.

2.2.2. CMPA Gauge-Based Reference Precipitation

The reference precipitation data were obtained from the China Multi-source Precipitation Analysis (CMPA) dataset. CMPA integrates hourly precipitation observations from approximately 30,000 automatic weather stations across mainland China with satellite-derived precipitation estimates from CMORPH. The dataset is generated using a combination of probability density matching and optimal interpolation methods, resulting in high spatial continuity and accuracy. Previous validation studies have shown that CMPA provides reliable precipitation estimates over

 doi:10.20944/preprints202602.1763.v1

China, with overall errors generally within 10%, and within 20% in sparsely gauged western regions. Due to its high data quality and extensive validation, CMPA was adopted as the reference precipitation dataset in this study.

Table 1. Summary of precipitation datasets used in this study.

Dataset	Data Type	Temporal Resolution	Spatial Resolution	Spatial Coverage	Primary Purpose
IMERG-Late [22]	Satellite-based	1 hour	0.1° × 0.1°	Global (90°S-90°N)	Near-real-time precipitation estimation
CMPA [23,24]	Gauge–satellite merged	1 hour	0.1° × 0.1°	Mainland China	Reference precipitation dataset

3. Methods

3.1. Problem Setup

This study addresses near-real-time error correction of gridded satellite precipitation by learning a mapping from the IMERG-Late product to a gauge-based reference field over mainland China. Let $P_s(\mathbf{x},t)$ denote the satellite-derived precipitation at grid cell $\mathbf{x}$ and time t, and $P_g(\mathbf{x},t)$ denote the corresponding reference precipitation from CMPA. The objective is to estimate a corrected precipitation field $\hat{P}(X, t)$ such that $\hat{P}(X, t)$ approximates $P_g(X,t)$ while retaining realistic spatial structures in precipitation patterns. Paired samples were constructed by collocating IMERG-Late and CMPA at the same spatial grid (0.1°) and temporal resolution (hourly). The study period spans 2015-2017. For model development, a subset of paired data from 2015 was used for training and validation, and the trained models were then applied for correction and evaluation over the target period.

3.2. Error Correction Models

3.2.1. U-Net-Based Spatial Error Correction Model

Following the problem formulation in Section 3.1, the objective of the error correction model is to estimate a corrected precipitation field $\hat{P}(X, t)$ that approximates the gauge-based reference precipitation $P_g(X,t)$ from the satellite-derived input $P_s(X,t)$, while preserving realistic spatial structures in precipitation patterns. To explicitly account for the spatial organization of satellite precipitation errors, a U-Net-based deep learning model is adopted as the primary correction framework.

Unlike grid-wise regression approaches that treat each grid cell independently, the U-Net model performs field-to-field regression at each time step. Specifically, for a given time t, the model takes the entire gridded satellite precipitation field as input and outputs a corrected precipitation field with the same spatial resolution and extent. This mapping can be expressed as

$$\hat{P}(\cdot, t) = f_\theta(P_s(\cdot,t))$$

where $f_\theta(\cdot)$ denotes the nonlinear mapping learned by the U-Net model with parameters θ. The corrected precipitation value at a specific grid cell $\mathbf{x}$ is then obtained by

$$\hat{P}(X, t) = [f_\theta(P_s(\cdot,t))](X)$$

The U-Net architecture follows a symmetric encoder–decoder structure with multi-scale skip connections, as illustrated in Figure 2. In the encoding path, the input precipitation field is progressively transformed through convolutional layers and downsampling operations, enabling the extraction of hierarchical spatial features. As spatial resolution decreases, the number of feature channels increases, allowing the model to capture precipitation characteristics ranging from local

 doi:10.20944/preprints202602.1763.v1

intensity variations to regional-scale organization. In the decoding path, these features are progressively upsampled back to the original spatial resolution to reconstruct the corrected precipitation field. Skip connections concatenate feature maps from corresponding encoder levels to the decoder, ensuring that fine-scale spatial information—such as precipitation boundaries and localized gradients—is preserved during reconstruction.

Model training aims to minimize the discrepancy between the corrected precipitation field and the reference precipitation field over the training samples. Let Ω denote the set of grid cells in the spatial domain and $\mathcal{T}_{tr}$ denote the set of training time indices. The training objective is defined as a mean squared error over all grid cells and training times:

$$\min_{\theta} \mathcal{L}_{U-Net}(\theta) = \frac{1}{|\mathcal{T}_{tr}||\Omega|} \sum_{t \in \mathcal{T}_{tr}} \sum_{X \in \Omega} \left(\left[f_{\phi}(P_s(\cdot, t)) \right](X) - P_g(X, t) \right)^2$$

By learning this spatially explicit mapping, the U-Net model corrects not only precipitation magnitude at individual grid cells but also spatially coherent error structures. This property is particularly relevant for near-real-time satellite precipitation products, where false precipitation and other biases often appear as spatially clustered patterns rather than isolated pointwise errors. At the same time, the reliance on spatial consistency implies an inherent trade-off, whereby weak or spatially fragmented precipitation signals may be partially suppressed during correction. This characteristic provides methodological context for the performance patterns observed in subsequent evaluations.

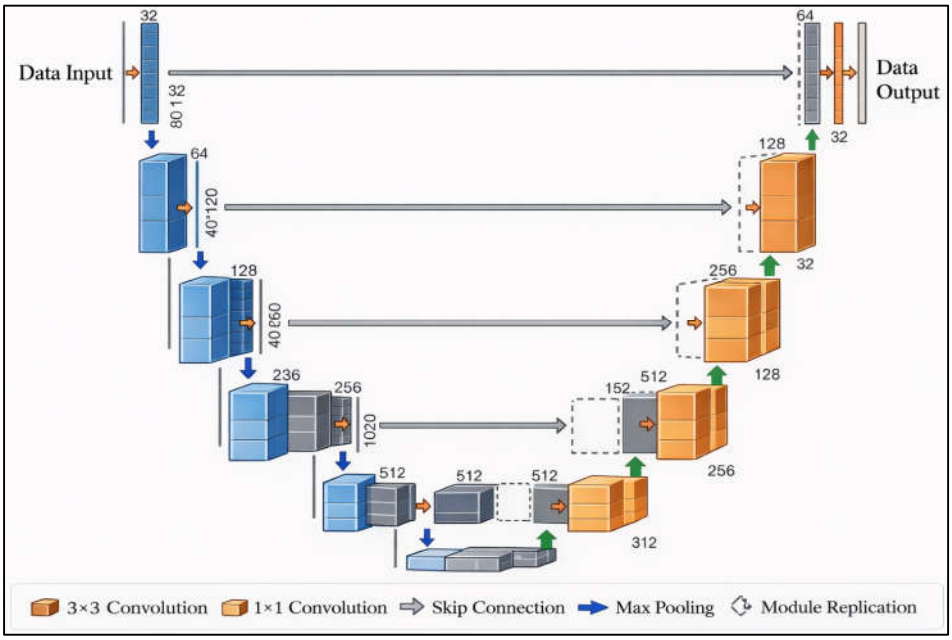

Figure 2. Architecture of the U-Net model used for precipitation error correction.

3.2.2. Grid-Wise Machine Learning Benchmark Models

To benchmark the U-Net correction, we implement several widely used machine learning models as comparative approaches. In contrast to the field-to-field formulation above, these baselines perform grid-wise regression, treating each grid cell X and time t as an independent sample. Each model learns a scalar mapping $g_{\phi}(\cdot)$ parameterized by ϕ such that

$$\hat{P}(X,\ t) = g_{\phi}(P_s(X, t))$$

i.e., the corrected precipitation at a given location depends only on the satellite estimate at that location, without explicit conditioning on surrounding grid cells.

 doi:10.20944/preprints202602.1763.v1

Consistent with the supervised objective in Section 3.1, the baseline models are trained by minimizing a pointwise mean squared error over all training samples:

$$\min_{\phi} \mathcal{L}_{ML}(\phi) = \frac{1}{|\mathcal{T}_{tr}||\Omega|} \sum_{t \in \mathcal{T}_{tr}} \sum_{X \in \Omega} (g_{\phi}(P_s(X,t)) - P_g(X,t))^2$$

In this study, the benchmark set includes representative tree-based and boosting-based regressors as well as a feedforward neural network regressor, namely Decision Tree Regressor (DTR), Random Forest (RF), Gradient Boosting Decision Tree (GBDT), Multilayer Perceptron (MLP), and Extreme Gradient Boosting (XGBoost). These models were selected as representative nonlinear regression approaches that have been widely used in precipitation estimation and bias correction studies.

These grid-wise models are generally effective at reducing systematic bias and random errors in precipitation magnitude, especially when the dominant discrepancies are locally predictable. However, because spatial dependencies among neighboring grid cells are not explicitly modeled, grid-wise correction may have limited ability to enforce spatial coherence or suppress spatially clustered false precipitation patterns.

3.2.3. Model Comparison Strategy

All models are trained to approximate the same target mapping defined in Section 3.1, using identical satellite–reference paired samples and the same training/validation split. For consistency, the U-Net model is trained on precipitation fields $P_s(\cdot,t)$ with field-level supervision $P_g(\cdot,t)$, while the machine learning baselines are trained on pointwise pairs $(P_s(X,t), P_g(X,t))$. After training, each method produces corrected precipitation estimates $\hat{P}(X,\ t)$ over the evaluation period, and all corrected outputs are assessed using the same statistical and event-detection metrics described in Section 3.3.

This unified formulation ensures that performance differences can be attributed to modeling strategy—field-to-field spatial correction versus grid-wise pointwise correction—rather than differences in data pairing, target definition, or evaluation protocol.

3.3. Model Training, Inputs and Hyperparameter Selection

To ensure a consistent and fair comparison among different correction models, all experiments were conducted using the same satellite–reference precipitation pairs and an identical training and validation strategy. A subset of paired IMERG-Late and CMPA precipitation data from 2015 was used for model development, while the trained models were subsequently applied to the full evaluation period.

Specifically, 10% of the paired CMPA precipitation data in 2015 was randomly sampled and collocated with IMERG-Late at the same spatial grid and time stamps to construct supervised learning samples. These samples were further divided into 80% for training and 20% for validation. This sampling strategy was adopted to reflect realistic data availability in near-real-time applications while maintaining sufficient data for model training and validation.

For the U-Net model, the input consists of gridded IMERG-Late precipitation fields at a spatial resolution of 0.1°, and the output is a corrected precipitation field with the same spatial dimensions. Model training aims to minimize the discrepancy between the corrected output and the reference precipitation field, following the field-wise objective defined in Section 3.2.1. Model selection was based on validation performance to avoid overfitting.

For the grid-wise machine learning benchmark models, the input is the satellite precipitation value at each grid cell and time step, and the output is the corresponding reference precipitation value from CMPA. Hyperparameters for all baseline models were tuned using the validation set to achieve a balance between model capacity and generalization. For tree-based models, tuning focused on parameters controlling model complexity, such as tree depth, minimum samples per split or leaf, and the number of estimators. Boosting-based models additionally considered learning rate and

 doi:10.20944/preprints202602.1763.v1

subsampling-related parameters. For the multilayer perceptron model, network capacity, activation functions, and regularization settings were adjusted. For XGBoost, key parameters related to boosting iterations, tree depth, learning rate, and regularization were tuned following standard practice.

For each model, the final configuration was selected based on validation performance, primarily evaluated using RMSE and Bias. The selected models were then used to generate corrected precipitation fields for subsequent evaluation.

3.4. Evaluation Metrics

The performance of the corrected satellite precipitation products was evaluated using gauge-based precipitation observations as reference values. Both continuous statistical metrics and categorical event-based metrics were employed to comprehensively assess the accuracy of precipitation estimation and the ability to detect precipitation events.

3.4.1. Continuous Statistical Metrics

To evaluate the quantitative accuracy of precipitation estimates, three widely used statistical metrics were adopted: the correlation coefficient (CC), relative bias (Bias), and root mean square error (RMSE). These metrics describe the consistency, systematic deviation, and overall magnitude of errors between satellite-derived precipitation and reference observations.

The correlation coefficient is defined as

$$CC = \frac{\sum_{i=1}^{N}(P_i - \bar{P})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^{N}(P_i - \bar{P})^2}\sqrt{\sum_{i=1}^{N}(G_i - \bar{G})^2}}$$

where N denotes the total number of samples, P_i represents the evaluated precipitation (satellite or corrected product), G_i represents the reference precipitation from gauges, and $\bar{P}$ and $\bar{G}$ are their respective mean values.

The relative bias is used to quantify systematic overestimation or underestimation and is expressed as

$$Bias = \frac{\sum_{i=1}^{N}(P_i - G_i)}{\sum_{i=1}^{N} G_i}$$

where positive values indicate overestimation and negative values indicate underestimation.

The root mean square error is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - G_i)^2}$$

which reflects the overall magnitude of estimation errors.

3.4.2. Categorical Event-Based Metrics

In addition to continuous metrics, categorical statistics were employed to evaluate the capability of precipitation products to detect precipitation events. Based on a predefined precipitation threshold, precipitation estimates were classified into hit, missed, and false events, and event-detection metrics were calculated accordingly.

The probability of detection (POD) and the false alarm ratio (FAR) were used to quantify precipitation detection performance. These metrics are defined as

$$POD = \frac{H}{H + M}$$

 doi:10.20944/preprints202602.1763.v1

$$FAR = \frac{F}{H + M}$$

where H denotes the number of hit events, M denotes the number of missed events, and F denotes the number of false alarm events. A hit event occurs when both the evaluated precipitation and the reference precipitation exceed the threshold. A missed event occurs when the reference precipitation exceeds the threshold but the evaluated precipitation does not. A false alarm event occurs when the evaluated precipitation exceeds the threshold while the reference precipitation does not.

Following common practice in satellite precipitation evaluation, a daily precipitation threshold of 0.2 mm day^{-1} was adopted to define precipitation occurrence. Hourly precipitation estimates were aggregated to daily totals prior to the calculation of categorical metrics. POD ranges from 0 to 1, with higher values indicating better detection capability, while FAR ranges from 0 to 1, with lower values indicating fewer false precipitation detections.

3. Results and Discussion

3.1. Overall Error Assessment

As shown in Figure 3, box plots of multiple evaluation metrics illustrate the overall error correction performance of U-Net and the benchmark machine learning models. The results indicate that different correction methods exhibit distinct behaviors across accuracy- and event-based metrics, reflecting inherent trade-offs in near-real-time precipitation error correction.

In terms of quantitative precipitation estimation, U-Net exhibits higher correlation coefficients and lower RMSE values than most grid-wise machine learning methods, indicating improved agreement with gauge-based reference precipitation. In contrast, Decision Tree Regressor (DTR) shows relatively lower CC and higher RMSE, suggesting limited capability in correcting precipitation magnitude. Overall, U-Net demonstrates competitive performance in representing the variability of total precipitation after correction. In addition to CC and RMSE, relative Bias further reflects the systematic deviation of corrected precipitation estimates. All correction methods show negative Bias values, indicating a general tendency to underestimate precipitation, which suggests that the overestimation present in the original satellite product has been reduced to varying degrees. Among the evaluated methods, U-Net and Random Forest (RF) exhibit Bias distributions closer to zero, implying relatively better control of systematic deviation.

Event-based metrics further highlight differences in precipitation detection characteristics among the models. The Multilayer Perceptron (MLP) achieves relatively higher median values of the probability of detection (POD), indicating stronger capability in identifying precipitation occurrences. In contrast, U-Net shows lower POD values, reflecting a more conservative detection behavior. However, this reduction in POD is accompanied by a pronounced decrease in the false alarm ratio (FAR). Compared with the benchmark models, U-Net yields substantially lower FAR values, indicating improved suppression of false precipitation signals.

Taken together, these results reveal a clear trade-off between precipitation detection and false alarm suppression. While U-Net effectively reduces false precipitation events, it may also remove some weak or spatially fragmented precipitation signals, leading to lower detection rates. Nevertheless, considering its performance across all evaluated metrics, U-Net provides a balanced correction behavior that is particularly advantageous for applications sensitive to false precipitation, such as near-real-time hydrological monitoring and early warning systems.

 doi:10.20944/preprints202602.1763.v1

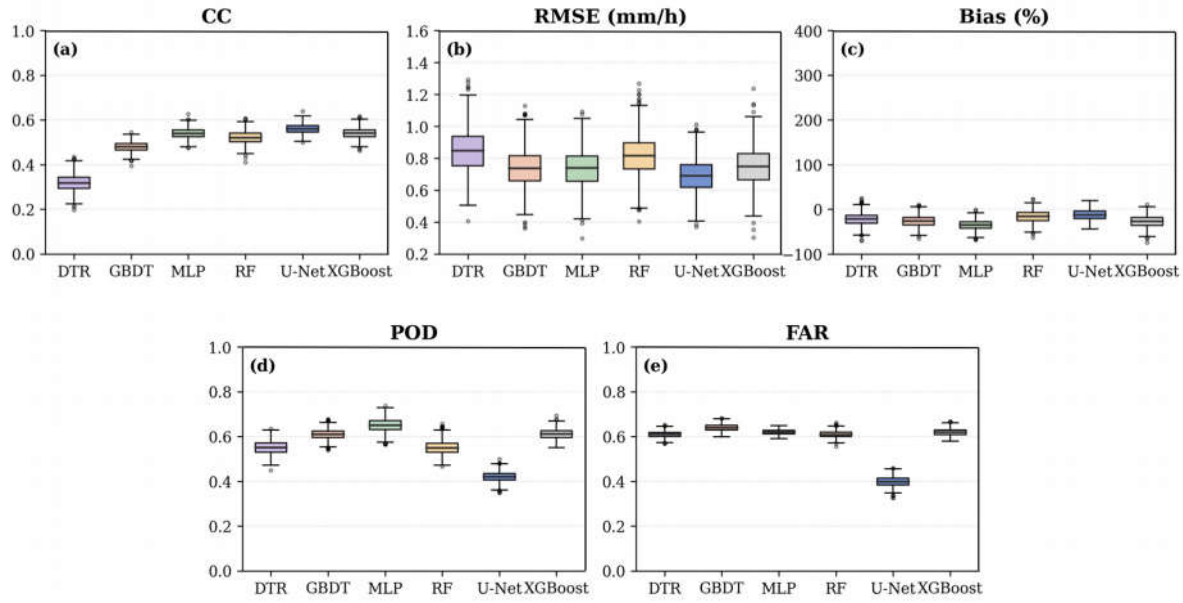

Figure 3. Boxplots of five evaluation metrics (CC, RMSE, Bias, POD, and FAR) for U-Net and five benchmark machine-learning models (DTR, GBDT, MLP, RF, and XGBoost) over Mainland China (2015–2017).

3.2. Spatial Performance Comparative Analysis

The spatial performance of different correction methods was further examined to assess their ability to represent regional precipitation characteristics across mainland China. Spatial distributions of evaluation metrics provide complementary insights beyond overall statistics by revealing regional heterogeneity and spatially organized correction behavior.

As shown in Figures 4 and 5, the spatial patterns of the correlation coefficient (CC) and RMSE exhibit clear differences among the evaluated methods. Decision Tree Regressor (DTR) and Gradient Boosting Decision Tree (GBDT) generally show lower CC values and higher RMSE over large areas, indicating weaker spatial consistency with gauge-based observations. Random Forest (RF) also exhibits relatively elevated RMSE in many regions. In contrast, U-Net and Multilayer Perceptron (MLP) display higher CC values and comparatively lower RMSE across broad regions, suggesting improved spatial agreement and reduced magnitude errors. This improvement is particularly evident in regions with frequent precipitation, where spatial continuity plays an important role in correction performance.

 doi:10.20944/preprints202602.1763.v1

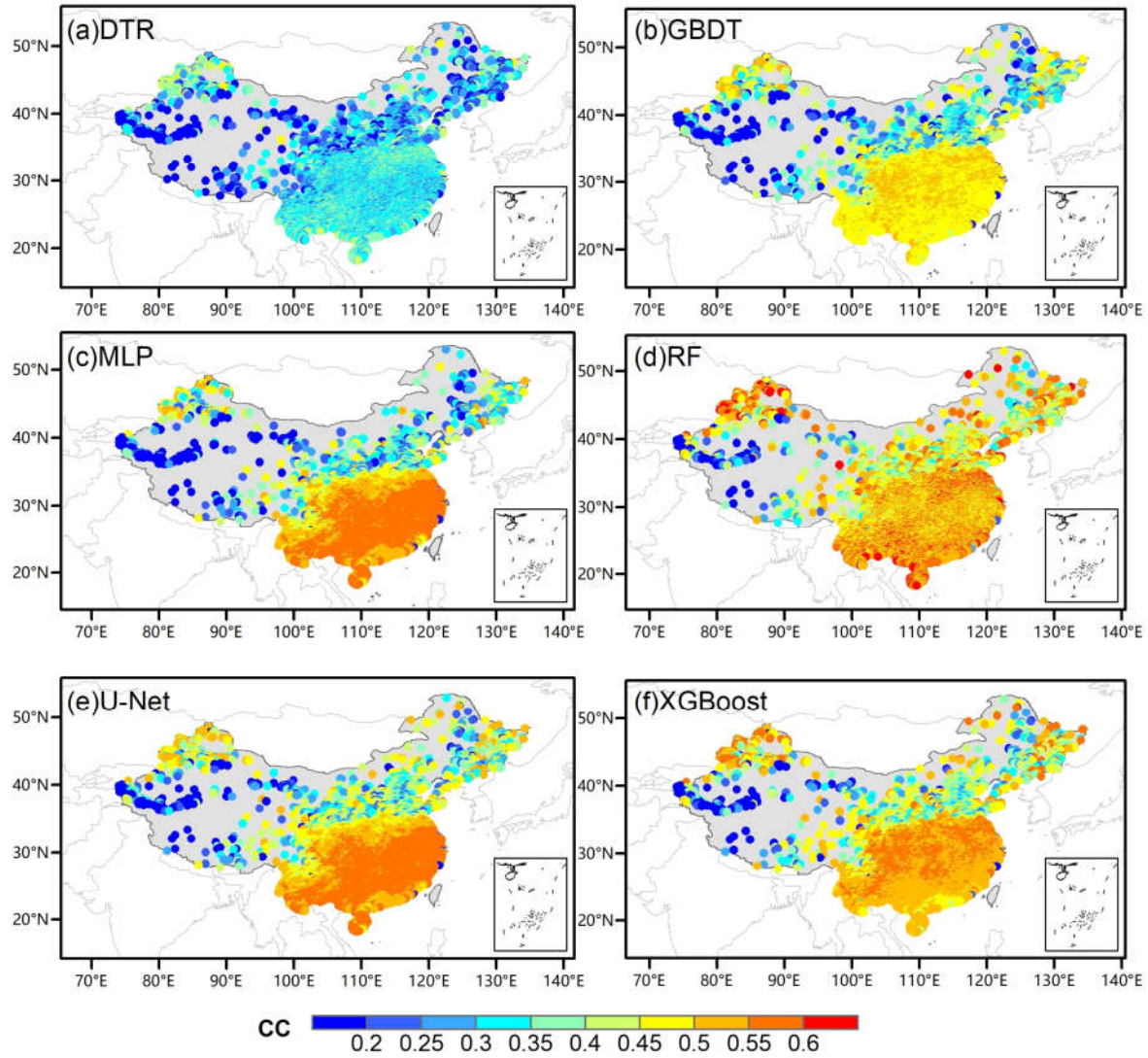

Figure 4. Spatial distribution of the correlation coefficient (CC) across the study area for six machine learning correction methods.

 doi:10.20944/preprints202602.1763.v1

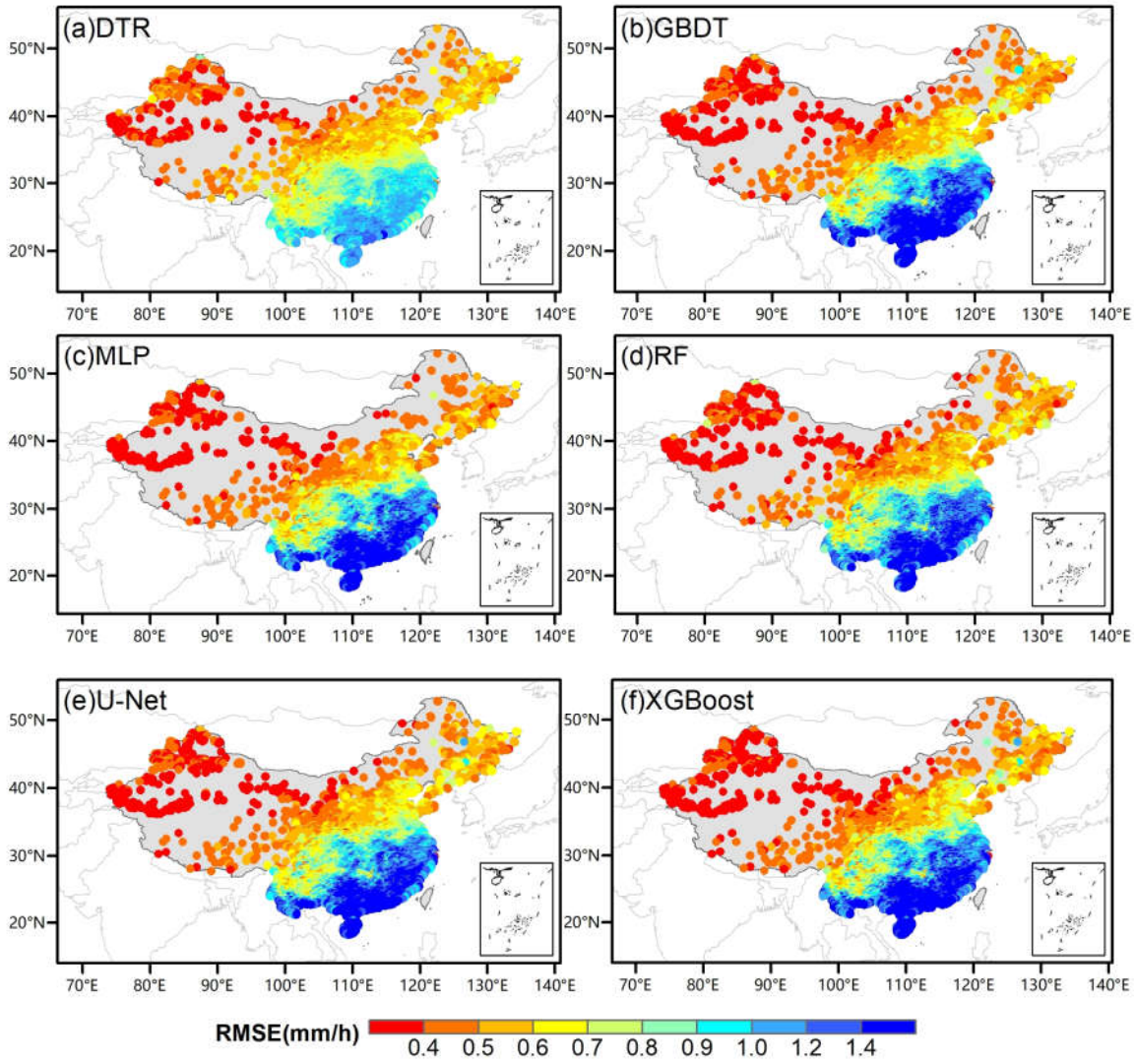

Figure 5. Spatial distribution of the root mean square error (RMSE) across the study area for six machine learning correction methods.

The spatial distribution of relative Bias is illustrated in Figure 6. All correction methods exhibit predominantly negative Bias values across the study area, indicating an overall tendency to underestimate precipitation. However, the magnitude and spatial patterns of Bias vary substantially among methods. MLP shows relatively strong underestimation in many regions, while DTR and RF exhibit similar Bias patterns, and GBDT and XGBoost display comparable spatial characteristics. In comparison, U-Net generally produces Bias values closer to zero over most regions, indicating more effective control of systematic deviation in a spatially coherent manner.

 doi:10.20944/preprints202602.1763.v1

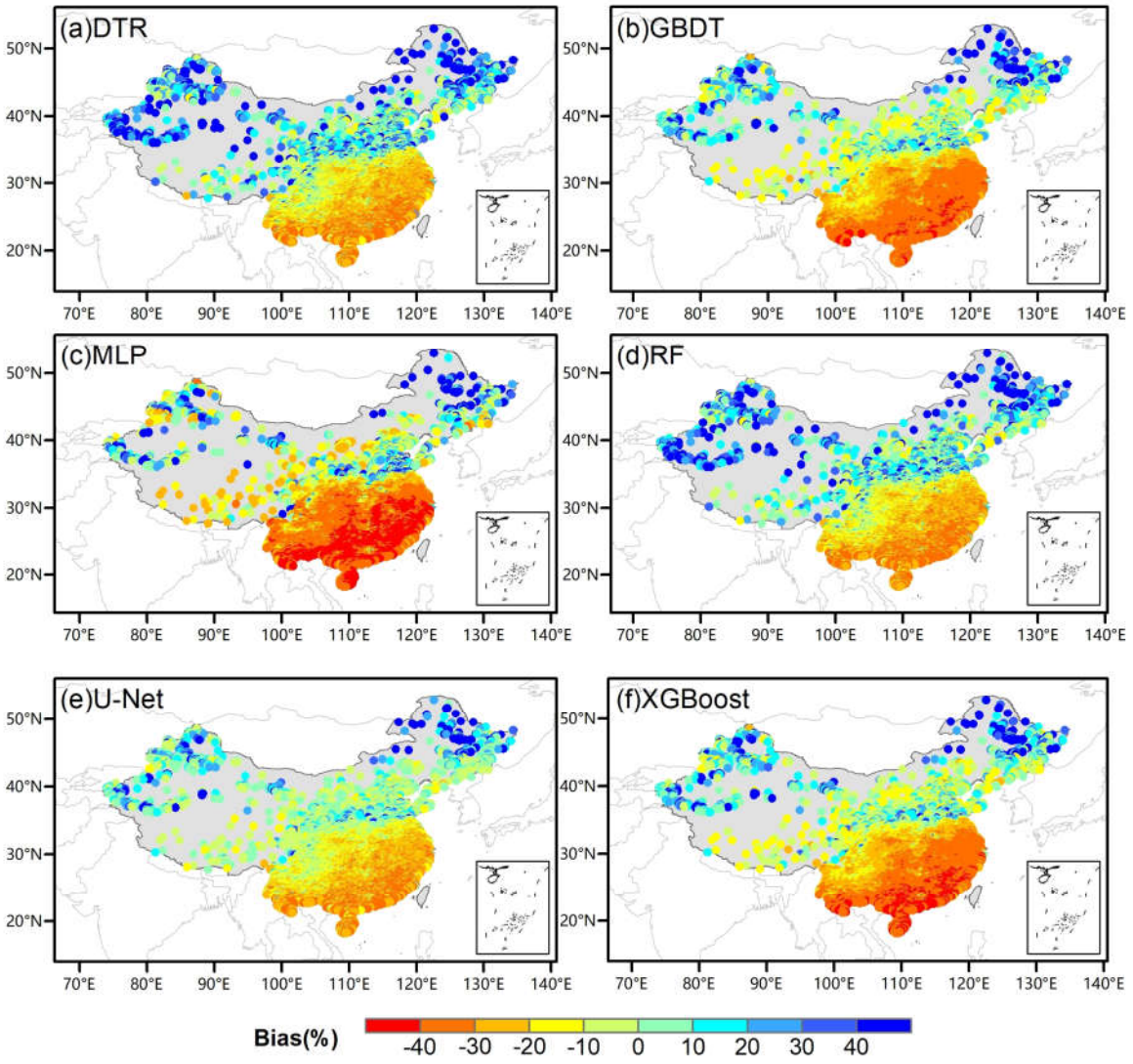

Figure 6. Spatial distribution of relative bias (Bias) across the study area for six machine learning correction methods.

Spatial patterns of precipitation detection performance are shown in Figures 7 and 8 for POD and FAR, respectively. Grid-wise machine learning models such as GBDT, MLP, and XGBoost tend to maintain relatively higher POD values across many regions, indicating stronger precipitation detection capability. DTR and RF exhibit moderate POD performance. In contrast, U-Net shows lower POD values over large portions of the study area, reflecting a more conservative detection behavior. However, this reduction in POD is accompanied by a clear decrease in FAR. As shown in Figure 8, U-Net consistently yields lower FAR values than the grid-wise correction methods, indicating effective suppression of spatially clustered false precipitation events.

 doi:10.20944/preprints202602.1763.v1

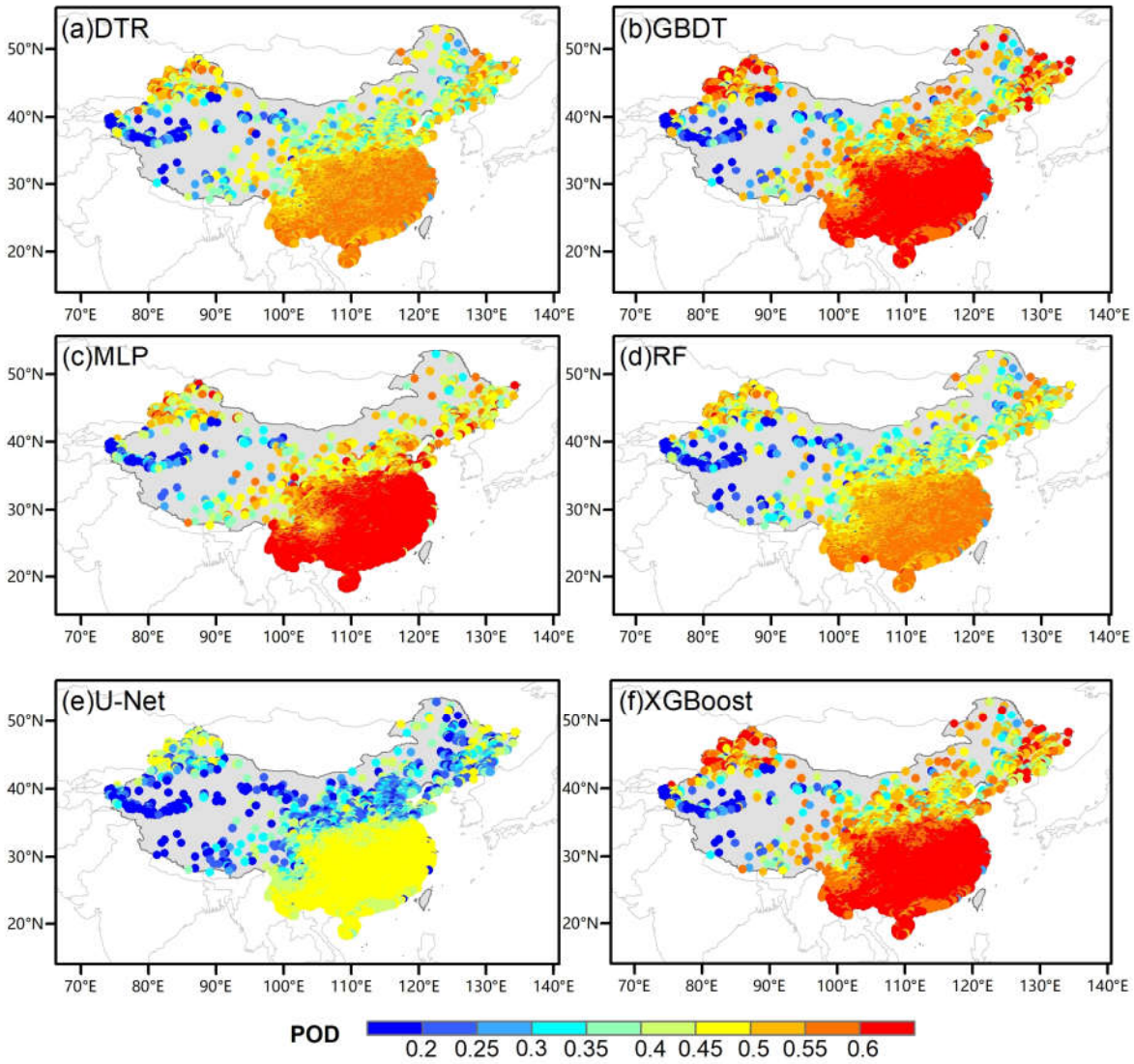

Figure 7. Spatial distribution of the probability of detection (POD) across the study area for six machine learning correction methods.

 doi:10.20944/preprints202602.1763.v1

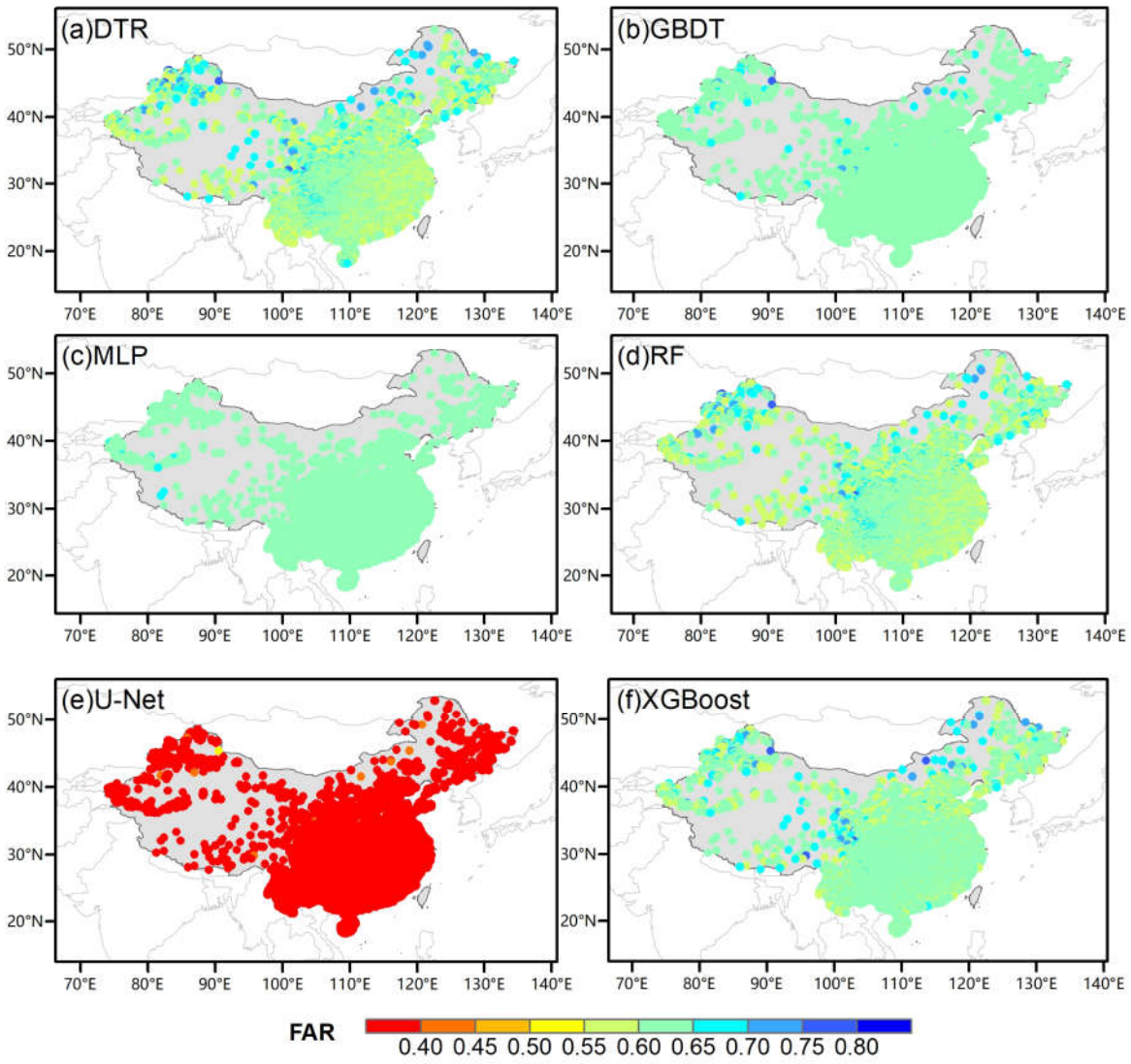

Figure 8. Spatial distribution of the false alarm ratio (FAR) across the study area for six machine learning correction methods.

Overall, the spatial analysis highlights a consistent trade-off between precipitation detection and false alarm suppression across different correction strategies. Grid-wise machine learning models tend to preserve higher detection rates but exhibit limited ability to suppress false precipitation in a spatially consistent manner. In contrast, the spatially explicit U-Net model demonstrates improved control of false precipitation through spatially coherent correction, albeit at the cost of reduced detection of weak or fragmented precipitation signals. These results suggest that incorporating spatial context is particularly beneficial for near-real-time precipitation applications where false precipitation poses a critical concern.

4. Conclusion

Despite the continuous improvement of satellite precipitation observations and correction techniques under the Global Precipitation Measurement (GPM) framework, near-real-time precipitation products still exhibit considerable errors and uncertainties, particularly related to false precipitation. To address this issue, this study proposed a U-Net-based deep learning approach for real-time error correction of satellite precipitation products and systematically evaluated its performance against several widely used machine learning methods over mainland China. The results indicate that the proposed method achieves improved overall accuracy and spatial

 doi:10.20944/preprints202602.1763.v1

consistency, with clear advantages in reducing false precipitation events, as reflected by lower FAR, reduced Bias, and improved CC and RMSE. At the same time, the method exhibits a more conservative precipitation detection behavior, leading to relatively lower POD values, which suggests a trade-off between false precipitation suppression and precipitation detection. This behavior may be associated with the spatially coherent correction strategy of the U-Net model and highlights the need for further optimization in balancing detection sensitivity and false alarm control. Notably, satisfactory correction performance was achieved using only a limited fraction of the available training data, indicating that the proposed method has relatively low dependence on training sample size and good applicability for near-real-time scenarios. Overall, this study demonstrates that incorporating spatial context through deep learning provides a promising direction for improving the accuracy and reliability of real-time satellite precipitation products.

Author Contributions: Hua Wang: Data Curation, Validation, Writing - Original Draft, Writing - Review & Editing. Zhenyu Yu: Conceptualization, Methodology, Writing - Review & Editing. Hanqing Chen: Resources, Writing - Review & Editing.

Funding: This research was funded by Zhejiang Provincial Soft Science Research Program Project, grant number 2025C25051.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. JIANG L, BAUER-GOTTWEIN P. How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China [J]. Journal of Hydrology, 2019, 572: 486-500.
2. TANG G, ZENG Z, MA M, et al. Can Near-Real-Time Satellite Precipitation Products Capture Rainstorms and Guide Flood Warning for the 2016 Summer in South China? [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14: 1208-1212.
3. HE K, CHEN X, ZHAO D, et al. Precipitation-induced landslide risk escalation in China's urbanization with high-resolution soil moisture and multi-source precipitation product [J]. Journal of Hydrology, 2024, 638: 131536.
4. LI W, KANG Y, LI L, et al. Comprehensive assessment of five near-real-time satellite precipitation products in the Lower Yangtze River Basin and the Lixiahe region, China: Dual perspectives from time series and extreme events [J]. Atmospheric Research, 2024, 308: 107520.
5. LU X, TANG G, WANG X, et al. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China [J]. Journal of Hydrology, 2019, 575: 1239-1252.
6. JIA J, HE Y, ZHANG B, et al. Evaluation of hourly summer precipitation products over the Tibetan Plateau: A comparative analysis of IMERG, CMORPH, and TPHiPr [J]. Atmospheric Research, 2025, 316: 107955.
7. ROJAS Y, MINDER J R, CAMPBELL L S, et al. Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central Chile [J]. Atmospheric Research, 2021, 253: 105454.
8. TANG G, CLARK M P, PAPALEXIOU S M, et al. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets [J]. Remote Sensing of Environment, 2020, 240: 111697.
9. BECK H E, VAN DIJK A I J M, LEVIZZANI V, et al. MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data [J]. Hydrology and Earth System Sciences, 2017, 21(1): 589-615.
10. PANEGROSSI G, RYSMAN J-F, CASELLA D, et al. CloudSat-Based Assessment of GPM Microwave Imager Snowfall Observation Capabilities [J]. Remote Sensing, 2017, 9(12): 1263.
11. CHEN H, YONG B, WANG L, et al. Global component analysis of errors in five satellite-only global precipitation estimates [M]. 2020.

 doi:10.20944/preprints202602.1763.v1

12. TIAN Y, PETERS-LIDARD C D, EYLANDER J B. Real-Time Bias Reduction for Satellite-Based Precipitation Estimates [J]. Journal of Hydrometeorology, 2010, 11(6): 1275-1285.
13. CHEN H, YONG B, GOURLEY J J, et al. A Novel Real-Time Error Adjustment Method With Considering Four Factors for Correcting Hourly Multi-satellite Precipitation Estimates [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 4105211.
14. ZHANG Z, WANG D, WANG G, et al. Use of SMAP Soil Moisture and Fitting Methods in Improving GPM Estimation in Near Real Time [J]. Remote Sensing, 2019, 11: 368.
15. CROW W T, HUFFMAN G J, BINDLISH R, et al. Improving Satellite-Based Rainfall Accumulation Estimates Using Spaceborne Surface Soil Moisture Retrievals [J]. Journal of Hydrometeorology, 2009, 10(1): 199-212.
16. ZHAN W, PAN M, WANDERS N, et al. Correction of real-time satellite precipitation with satellite soil moisture observations [J]. Hydrol Earth Syst Sci, 2015, 19(10): 4275-4291.
17. ROMáN-CASCóN C, PELLARIN T, GIBON F, et al. Correcting satellite-based precipitation products through SMOS soil moisture data assimilation in two land-surface models of different complexity: API and SURFEX [J]. Remote Sensing of Environment, 2017, 200: 295-310.
18. Lyu Yi, Yong Bin, Shen Zhehui, et al. A rapid correction method for near real-time FY-4A retrieved precipitation based on ensemble learning. Journal of Remote Sensing, 2024, 28: 677–688.
19. TIAN Y, PETERS-LIDARD C D, EYLANDER J B, et al. Component analysis of errors in satellite-based precipitation estimates [J]. Journal of Geophysical Research: Atmospheres, 2009, 114(D24).
20. CHEN H, YONG B, KIRSTETTER P-E, et al. Global component analysis of errors in three satellite-only global precipitation estimates [J]. Hydrology and Earth System Sciences, 2021, 25(6): 3087-3104.
21. TIAN Y, PETERS-LIDARD C D. A global map of uncertainties in satellite-based precipitation measurements [J]. Geophysical Research Letters, 2010, 37(24).
22. KIDD C, HUFFMAN G. Global precipitation measurement [J]. Meteorological Applications, 2011, 18(3): 334-353.
23. SHEN Y, ZHAO P, PAN Y, et al. A high spatiotemporal gauge-satellite merged precipitation analysis over China [J]. Journal of Geophysical Research: Atmospheres, 2014, 119(6): 3063-3075.
24. Yu Jingjing, Shen Yan, Pan Yang, et al. Comparative evaluation of daily merged precipitation datasets over China and international precipitation products. Acta Meteorologica Sinica, 2015, 73(2): 394-410.
25. RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation [J]. 2015, 9351: 234-241.
26. BREIMAN L. Random Forests [J]. Machine Learning, 2001, 45(1): 5-32.
27. CHEN T, GUESTRIN C. XGBoost [M]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 785-794.
28. CAIRES S, STERL A. Validation of ocean wind and wave data using triple collocation [J]. Journal of Geophysical Research: Oceans, 2003, 108(C3): 3098.
29. YONG B, REN L, HONG Y, et al. First evaluation of the climatological calibration algorithm in the real-time TMPA precipitation estimates over two basins at high and low latitudes [J]. Water Resources Research, 2013, 49(5): 2461-2472.