remote sensing

MDPI

Article

Artificial Intelligence-Based Precipitation Estimation Method Using Fengyun-4B Satellite Data

Nianqing Liu [1,2,3], Jianying Jiang [1,2,3,*], Dongyan Mao [1,2,3], Meng Fang [1,2,3], Yun Li [1,2,3], Bowei Han [1,2,3] and Suling Ren [1,2,3]

[1] National Satellite Meteorological Center (National Centre for Space Weather), Beijing 100081, China; liunq@cma.gov.cn (N.L.); maody@cma.gov.cn (D.M.); fangmeng@cma.gov.cn (M.F.); liyun@cma.gov.cn (Y.L.); hanbw@cma.gov.cn (B.H.); rensl@cma.gov.cn (S.R.)
[2] Innovation Center for FengYun Meteorological Satellite (FYSIC), Beijing 100081, China
[3] Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites/Key Laboratory of Space Weather, CMA, Beijing 100081, China
* Correspondence: jiangjy@cma.gov.cn

Abstract: This paper proposes a novel precipitation estimation method based on FY-4B meteorological satellite data (FY-4B_AI). This method facilitates the spatiotemporal matching of 125 features derived from the multi-temporal and multi-channel data of the FY-4B satellite with precipitation data at stations. Subsequently, a precipitation model was constructed using the light gradient boosting machine (LGBM) algorithm. A comparative analysis of FY-4B_AI and GPM/IMERG-L products for over 450 million station cases throughout 2023 revealed the following: (1) The results demonstrate that FY-4B_AI is more accurate than GPM/IMERG-L. Six of the eight evaluation indices exhibit superior performance for FY-4B_AI in comparison to GPM/IMERG-L. These indices include the mean absolute error (MAE), root mean square error (RMSE), relative error (RE), correlation coefficient (CC), probability of detection (POD), and critical success index (CSI). As for the MAE, the results are 1.67 (FY-4B_AI) and 1.92 (GPM/IMERG-L), respectively. The RMSEs are 3.68 and 4.07, respectively. The REs are 17.72% and 26.28%, respectively. The CCs are 0.44 and 0.36, respectively. The PODs are 61.84% and 47.31%, respectively. The CSIs are 0.30 and 0.27, respectively. However, with regard to the mean errors (MEs) and false alarm rates (FARs), FY-4B_AI (−0.88 and 62.85%, respectively) displays a slight degree of inferiority in comparison to GPM/IMERG-L (−0.80 and 62.21%, respectively). (2) An evaluation of two strong weather events to represent the spatial distribution of precipitation in different climatic zones revealed that both FY-4B_AI and GPM/IMERG-L are equally capable of accurately representing these phenomena, irrespective of whether the region in question is humid, as is the case in the southeast, or dry, as is the case in the northwest.

Keywords: FY-4B; artificial intelligence; multi-temporal; precipitation estimation; evaluation

check for updates

Citation: Liu, N.; Jiang, J.; Mao, D.; Fang, M.; Li, Y.; Han, B.; Ren, S. Artificial Intelligence-Based Precipitation Estimation Method Using Fengyun-4B Satellite Data. *Remote Sens.* **2024**, *16*, 4076. https://doi.org/10.3390/rs16214076

Academic Editor: Ismail Gultepe

Received: 15 August 2024
Revised: 16 October 2024
Accepted: 29 October 2024
Published: 31 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Precipitation plays a pivotal role in the global water cycle, exerting a significant influence on the global water and energy balance [1,2]. The accurate estimation of precipitation is of great importance for the formulation of effective environmental policies and the implementation of appropriate disaster mitigation strategies. The prevailing approach to precipitation observation is based on direct measurements from ground-based rain gauges. While ground observations offer high-precision precipitation data, their spatial distribution is constrained by topography, which presents a challenge in accurately reflecting the spatiotemporal distribution patterns of precipitation in data-scarce regions (mountainous areas, oceans, etc.) [3,4]. The majority of traditional precipitation estimation methods rely on ground observations and radar data. However, the deployment of weather radars is constrained in complex terrain areas due to a number of factors, including terrain blockage,

uncertainty in the relationship between rainfall radar reflectivity and rainfall intensity, and the limitations of the telemetry range. Consequently, precipitation observation using this method is not sufficiently precise [5]. In contrast, satellite observation offers a number of advantages over ground-based observation, including all-weather, continuous observation and high spatial and temporal resolution. This makes it an effective means of monitoring precipitation in a variety of environments [6].

Over the past few decades, scientists have been engaged in research aimed at developing methods for utilising satellite data for precipitation estimation. In the initial stages of research, the utilisation of visible and infrared window channels was the primary focus of the investigation. However, recent studies have increasingly focused on the application of infrared split-window channels and other multi-channel infrared data. Infrared split-window channels have been extensively employed for precipitation estimation, largely due to their capacity to furnish data regarding cloud top temperature and atmospheric vertical structure. This information is of great consequence for the identification of precipitation areas and the estimation of precipitation intensity. For example, lower cloud top temperatures are frequently associated with more intense precipitation events. In addition to infrared split-window channels, other infrared channels, such as infrared water vapour channels, have also been employed for precipitation estimation. These channels can provide information about the water vapour content in the atmosphere, which is of great consequence for the comprehension of precipitation system dynamics.

In comparison to infrared and visible channels, microwave channels demonstrate superior penetration capabilities, thereby facilitating the acquisition of structural information within precipitation cloud systems [7]. Grody was the first to develop a methodology for estimating land surface precipitation based on microwaves [8]. The Tropical Rainfall Measuring Mission (TRMM) satellite, which is a joint mission by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), is equipped with the world's inaugural spaceborne precipitation radar (PR), which is capable of providing a three-dimensional representation of heavy rainfall [9]. Moreover, the TRMM satellite was equipped with microwave instruments that were capable of providing high-quality precipitation estimates based on microwave technology. On the basis of scattering theory, utilising the dynamic clustering method and in accordance with the brightness temperature of nine microwave channels of the TRMM microwave imager (TRMM/TMI) and their combinations, Li et al. [10] established a rainfall intensity retrieval algorithm with scattering index and polarisation-corrected temperature (PCT) as parameters. Based on data from the microwave radiation imager (MWRI) of the TRMM/TMI and FY-3B, Li et al. [11] generated a cloud-radiation dataset that is capable of describing MWRI observation features and retrieving total rainfall intensity. The Global Precipitation Measurement (GPM) mission, which succeeded TRMM, is a new-generation global satellite precipitation product that provides robust data support for research on global climate change, drought monitoring, and flood monitoring, among other fields [5,12].

Recent studies have demonstrated that the combination of satellite data from multiple frequency bands can enhance the accuracy of precipitation estimation. This integration encompasses not only the maximisation of information extraction from individual frequency bands but also the synergistic effects between different frequency bands, which are known as 'multi-band synergy' (MBS). This approach is more effective at capturing the intricacies of precipitation-related phenomena, thereby enhancing the precision and dependability of precipitation estimates.

Despite notable advancements, precipitation estimation based on satellite data still presents challenges. These include the identification of precipitation regions with greater accuracy, the improvement of the spatial and temporal resolution of precipitation estimation, and the reduction of cloud opacity. These are all areas that require further investigation.

In recent years, there has been a notable increase in the application of machine learning (ML) techniques in meteorology [13–15], particularly in precipitation estimation using satellite remote sensing data. The deployment of machine learning methodologies, encompassing deep learning and attention mechanisms, has illustrated the potential to augment the precision and expediency with which intricate weather systems and precipitation patterns can be addressed. In particular, convolutional neural networks (CNN) and recurrent neural networks (RNN) have demonstrated significant advancement in the domain of precipitation forecasting through the application of deep learning techniques [16,17]. To illustrate, the U-Net architecture has been employed for near-real-time precipitation estimation, thereby demonstrating its considerable capabilities in processing image data, which is of particular importance for satellite remote sensing data. Attention mechanisms [18] represent a class of methods that permit models to focus on the most pertinent information in input data. They have been demonstrated to have significant potential in improving precipitation estimation accuracy. The application of attention mechanisms to satellite remote sensing data allows researchers to more effectively identify and predict precipitation events, which in turn leads to an enhanced level of precipitation estimation precision.

The advent of meteorological satellites, spaceborne sensors, and artificial intelligence has resulted in considerable advances in precipitation estimation based on satellite data. This is primarily accomplished through precipitation inversion based on channel radiation characteristics and satellite precipitation inversion based on artificial intelligence machine learning [19,20]. In a study conducted by Liu et al. [21], a precipitation estimation product that integrated data from the FY-4A meteorological satellite, and artificial intelligence was used to analyse a shear line event over the Tibetan Plateau. A comparison of single-temporal and multi-temporal geostationary satellite precipitation estimation products revealed that the latter exhibited superior precipitation inversion efficiency. The algorithms integrate the latest ground precipitation data iteratively without reliance on simultaneous ground-based observations, thereby enhancing the effectiveness of the inversion product. This is of particular significance in the context of issuing catastrophic weather warnings.

This paper represents an updated and further refinement of the reference [21], which is cited for the purposes of providing context and establishing the current state of knowledge in this field. Firstly, the FY-4B is used in place of the FY-4A. The FY-4B meteorological satellite was successfully launched on 3 June 2021. In comparison to the FY-4A meteorological satellite, the FY-4B features an additional channel for the detection of low-level water vapour in the 7.24–7.6 μm range, which is of particular interest in meteorological studies. The rapid fluctuations in low-level water vapour are a key indicator of the formation and development of local sudden convective systems. Secondly, in view of the greater practicality of multi-temporal satellite precipitation estimation products in real-time applications (see [21]), this paper exclusively utilises the multi-temporal FY-4B satellite data precipitation estimation algorithm (FY-4B_AI), excluding single-frame satellite precipitation estimation products. The research scope has been expanded from the Tibetan Plateau [21] to the entire mainland of China, with the objective of evaluating the accuracy of the products in question. This is to establish a suitable disaster weather FY-4B satellite precipitation inversion model for mainland China.

2. Verification Area and Data

2.1. Verification Area

The validation area is defined as the entire mainland of China, extending from 70°E to 135°E longitude and 15°N to 55°N latitude, as illustrated in Figure 1.

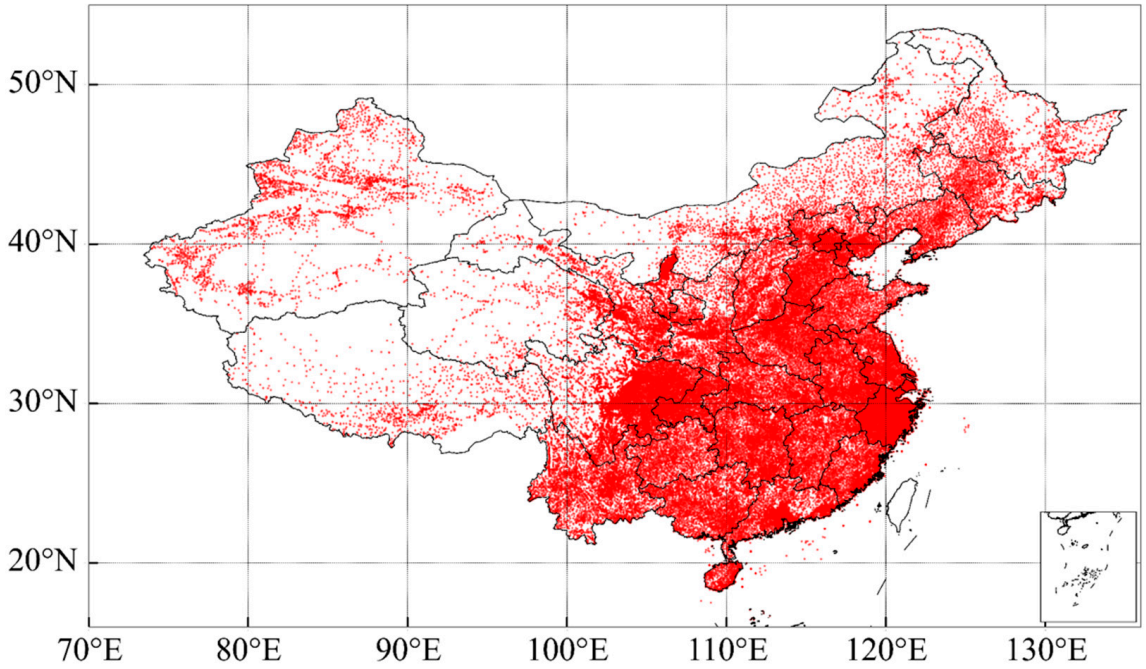


Figure 1. Distribution of ground-based precipitation observation stations in mainland China (red dots) [21].

2.2. FY-4B/AGRI Data

The Fengyun-4B (FY-4B) satellite, which was launched on 3 June 2021, represents the inaugural operational satellite in China's nascent Fengyun-4 series of geostationary meteorological satellites [22]. The satellite is equipped with four instruments, three of which are meteorological in nature. These are the advanced geostationary radiation imager (AGRI), the geostationary high-speed imager (GHI), and the geostationary interferometric infrared sounder (GIIRS). The FY-4B/AGRI has 15 channels (Table 1), with a resolution of 0.5–1 km for the visible and near-infrared channels, 2 km for the shortwave infrared channels, 2–4 km for the mid-wave infrared channels, and 4 km for the far-infrared channels. This study utilises the radiance brightness temperature values and derived variables from channels 07 to 15 (a total of nine channels) in the mid-wave infrared to far-infrared bands of FY-4B/AGRI as input. The original horizontal resolution of channel 7 is 2 km. The input data are reprocessed with an upscaled resolution of 4 km, ensuring consistent spatial resolution across all channels.

Table 1. FY-4B/AGRI performance parameters.

Band	Central Wavelength (μm)	Bandwidth (μm)	Spatial Resolution (km)	Main Purpose
1	0.47	0.45~0.49	1	Aerosol
2	0.65	0.55~0.75	0.5	Fog, cloud
3	0.825	0.75~0.90	1	Vegetation
4	1.379	1.371~1.386	2	Cirrus
5	1.61	1.58~1.64	2	Cloud, snow
6	2.25	2.10~2.35	2	Cirrus, aerosol
7	3.75	3.50~4.0 (high)	2	Fire

Table 1. *Cont.*

Band	Central Wavelength (μm)	Bandwidth (μm)	Spatial Resolution (km)	Main Purpose
8	3.75	3.50~4.0 (low)	4	Land surface
9	6.25	5.80~6.70	4	Upper-level water vapour
10	6.95	6.75~7.15	4	Mid-level water vapour
11	7.42	7.24~7.60	4	Low-level water vapour
12	8.55	8.3~8.8	4	Cloud
13	10.80	10.30~11.30	4	Surface temperature
14	12.00	11.50~12.50	4	Surface temperature
15	13.3	13.00~13.60	4	Clouds and water vapour

This study employs data from the Northern Hemisphere region, with a 15 min observation frequency. The data were observed at the start of every 00, 15, 30, and 45 min of each hour, in accordance with the specified observation schedule.

2.3. Ground-Based Rain Gauge Precipitation Data in Mainland China

The ground-based rain gauge precipitation data employed in this study are hourly observations from China's ground meteorological stations. These comprise both national-level ground meteorological stations (exceeding 10,000) and regional meteorological automatic stations (exceeding 50,000), amounting to approximately 70,000 stations in total. The distribution of rain gauges (illustrated by red dots in the figure) is shown in Figure 1. As can be observed from the figure, ground rain gauges are densely distributed in the low-altitude areas of eastern and central China, while they are relatively sparse in the high-altitude areas of western China and northern regions like Inner Mongolia.

The ground precipitation observation data employed in this study comprise the following variables: observation time, accumulated rainfall over the preceding hour (in millimetres), the longitude and latitude of the station, and data quality control code. The lowest recorded rainfall value was 0.1 mm/h. The data are updated in near real-time. The data underwent quality control procedures, and only those with a quality control code of 0, indicating no data anomalies, were selected for analysis. Over the course of the research period, which commenced at 00:00 on 1 January 2023 and concluded at 23:00 on 31 December 2023, a total of 70,000 data points were collected per hour, day, and year. Following the removal of data from satellite orbit control and stations that failed quality control, 453 million station times of data were employed for testing purposes.

2.4. GPM/IMERG-L Dataset

The integrated multi-satellite retrievals for GPM (IMERG) is the unified U.S. algorithm that provides the multi-satellite precipitation product for the U.S. GPM team [23]. The late precipitation estimates provided by IMERG are derived from the various precipitation-relevant satellite passive microwave (PMW) sensors that comprise the GPM constellation. These estimates are computed using the 2017 version of the Goddard profiling algorithm (GPROF2017) and then gridded, intercalibrated with the GPM combined Ku radar-radiometer algorithm (CORRA), and merged into half-hourly $0.1^\circ \times 0.1^\circ$ fields. The 'late' multi-satellite product is generated approximately 14 h after observation time, employing both forward and backward morphing techniques. The hourly precipitation data

utilised in the present study is derived from the period between 1 January and 31 December 2023. The data are presented in units of millimetres per hour.

3. Research Methods

3.1. Spatiotemporal Matching Method for the Evaluation of the Precision of Satellite-Derived Precipitation Estimates

Prior research [21,24–26] has demonstrated that the PERSIANN-CCS data exhibit suboptimal performance in the accuracy assessment. Accordingly, this study is limited to a comparison of the AI-based FY-4B meteorological satellite precipitation estimation (FY-4B_AI) with that of GPM/IMERG-L. The spatial resolution of the FY-4B_AI products is 0.04° longitude by 0.04° latitude, with a temporal resolution of 0.25 h, and the unit of measurement is mm/h. The spatial resolution of GPM/IMERG-L is 0.1° longitude by 0.1° latitude, with a temporal resolution of 0.5 h, and the unit is mm/h. The precipitation estimation product accuracy verification employs ground-based rain gauge observations as the truth value.

The data are subjected to a process of spatiotemporal matching whereby the data are aligned with respect to both space and time. The method of temporal matching is as follows: The FY-4B_AI data from 00:00 to 01:00 are matched with the ground rain gauge data from 00:00 to 01:00. This process is repeated for the GPM/IMERG-L data. The data from 00:00 to 00:30 and 00:30 to 01:00 are averaged and subsequently matched with the ground rain gauge data from 00:00 to 01:00. The spatial matching method entails the selection of the nearest grid point of the satellite precipitation product based on the location of the ground rain gauge station.

3.2. Satellite Precipitation Estimation Product Evaluation Methods

This study employs eight evaluation indices, namely, mean error (ME), mean absolute error (MAE), root mean square error (RMSE), relative error (RE), correlation coefficient (CC), probability of detection (POD), false alarm rate (FAR), and critical success index (CSI).

See Equations (1)–(5) for the calculation of the mean error (ME), mean absolute error (MAE), root mean square error (RMSE), relative error (RE), and correlation coefficient (CC), respectively, as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i) \tag{1}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Y_i - X_i| \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2} \tag{3}$$

$$\mathrm{RE} = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\frac{Y_i - X_i}{X_i}\right| \times 100\right) \tag{4}$$

$$\mathrm{CC} = \sum_{i=1}^{n}\left[\left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right)\right] / \sqrt{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \tag{5}$$

where Y is the satellite precipitation retrieval, X is the precipitation from the ground observation stations, n is the matching sample number, $\overline{Y}$ is the average of satellite precipitation retrievals of n samples, and $\overline{X}$ is the average of the precipitation of n samples from ground observation stations.

Formulas for the calculation of the probability of detection (POD), false alarm rate (FAR), and critical success index (CSI) are shown in Equations (6)–(8), respectively. The POD reflects the satellite's ability to identify precipitation, the FAR represents the ratio of the satellite's false alarm of precipitation, and the CSI indicates the proportion of the satellite's successful precipitation retrievals.

$$\mathrm{POD} = N_a/(N_a + N_c) \tag{6}$$

$$FAR = N_b/(N_a + N_b) \quad (7)$$

$$CSI = N_a/(N_a + N_b + N_c) \quad (8)$$

where N_a represents the number of stations where both the satellite and the ground rainfall station have observed precipitation (hit), N_b is the number of stations where the satellite shows there is precipitation but the ground rainfall station shows no precipitation (false alarm), and N_c is the number of stations where the satellite shows no precipitation but the ground rainfall station has observed precipitation (failure to report).

The performance of the two satellite precipitation products, FY-4B_AI and GPM/IMERG-L, in precipitation events will be discussed in detail in the following sections.

3.3. Intelligent Precipitation Estimation Algorithm Based on Multi-Temporal Satellite Data

The precipitation data recorded at a meteorological ground station represent the accumulated precipitation over the past hour. In circumstances in which meteorological conditions are undergoing rapid change, the precipitation rate within a single hour can vary considerably. Accordingly, this study employs multi-temporal satellite observation data in order to derive more accurate precipitation estimation outcomes. The time-frequency of FY-4B/AGRI in the Northern Hemisphere observation is 15 min. This study selects data at 15 min intervals within one hour as input, derives multiple features based on channel information, and trains a model using the light gradient boosting machine (LGBM) algorithm to achieve 1 h satellite precipitation inversion. The LGBM is a decision tree-based ensemble algorithm that is known for its fast training speed, low memory consumption, high accuracy, and support for distributed processing. It is commonly used to address various classification and regression problems.

3.3.1. Feature Construction for FY-4B_AI Satellite Precipitation Estimation

The FY-4B_AI method provides an illustrative example of the 01:00 satellite precipitation estimation algorithm. The model receives input from five time points: 00:00, 00:15, 00:30, 00:45, and 01:00. Each time point comprises nine channels of radiance brightness temperature data. The features derived from these data include the time rate of change in radiation values and the radiation differences between different channels at the same time.

The constructed features are illustrated in Figure 2. These comprise the maximum and minimum brightness temperatures of the nine channels at the five time points, resulting in a total of 18 features; the maximum and minimum brightness temperature differences of the nine channels at 15 min intervals, resulting in a total of 18 features; and the maximum and minimum brightness temperature differences of the nine channels at 30 min intervals. In total, 18 features are generated. A further 18 features are derived from the maximum and minimum brightness temperature differences of the nine channels at 45 min intervals. Furthermore, nine additional features are derived from the differences between the nine channels at 60 min intervals. Finally, the maximum and minimum differences between the various channels at the same time point are calculated. Given that channels 7 and 8, which are sensitive to the high- and low-temperature ends of the mid-infrared spectrum, respectively, contain solar radiation information during daylight hours, the differential features between channels at a given time point are calculated exclusively for channels 7 and 8. Consequently, channels 9 to 15 are excluded from the subtraction process, resulting in a total of 44 features. Consequently, 18, 18, 18, 18, 9, and 44 features are derived, resulting in a total of 125 features. These features capture a substantial amount of satellite data within a single hour, thereby establishing the foundation for training a precipitation estimation model with high precision.

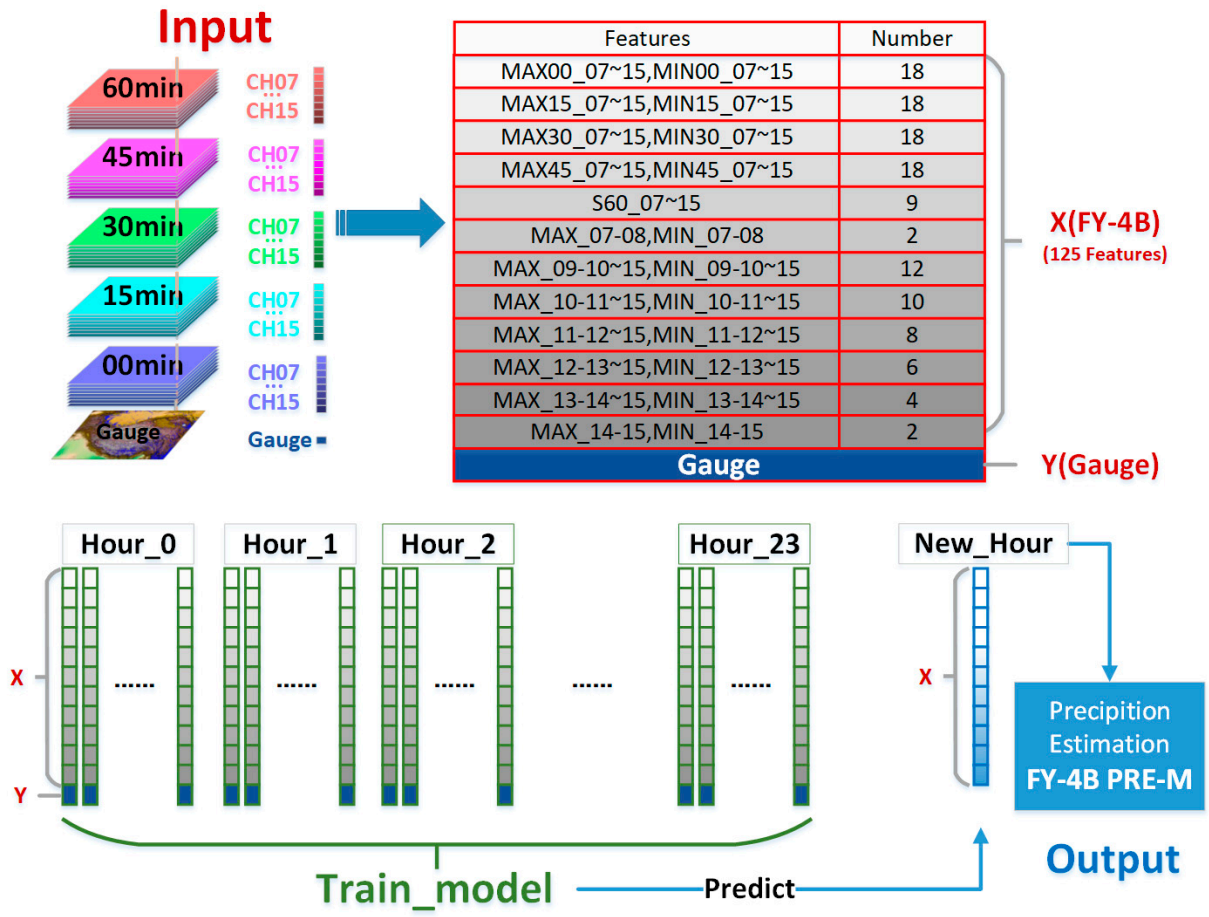


Figure 2. Data processing flow and algorithm model for multi-temporal FY-4B meteorological satellite precipitation estimation product based on artificial intelligence.

3.3.2. Precipitation Estimation Algorithm Model

The FY-4B_AI utilises a light gradient boosting machine (LGBM) algorithm. The construction process is illustrated with reference to the 01:00 FY-4B_AI precipitation estimation algorithm model. Based on the satellite data at 00:00, 00:15, 00:30, 00:45, and 01:00, the 125 features (X) mentioned above are calculated. The 1 h rainfall data from the ground rain gauge at 01:00 (Y, gauge) are selected as the truth value, and a training data point is established through the spatial matching method between satellite data and ground rain gauge precipitation data. In the event that there are N ground rain gauge observation stations situated within the designated study area, N training data points are established. In each instance of precipitation estimation derived from satellite data, the model is trained exclusively using the satellite data and precipitation observation data available within the preceding 24 h. To illustrate, in order to estimate precipitation at 01:00 today, satellite data from yesterday's 01:00 to today's 00:00 and precipitation observation data are employed as the input for training the model. It is not necessary to input the current one-hour ground precipitation observation data into the model. Immediate precipitation estimation can be performed following the acquisition of the current satellite data, thereby enhancing the timeliness of FY-4B_AI precipitation estimation. This is of particular importance for the monitoring and warning of rapidly developing and changing strong convective weather. Furthermore, even in the event of a small amount of missing rain gauge precipitation data or satellite data in the previous 24 h, the model can still be trained. Therefore, the model exhibits excellent compatibility. The model establishment and data processing flow are illustrated in Figure 2.

In the algorithmic model, each input feature is subjected to an analysis and evaluation in order to ascertain its relative importance. A higher score is indicative of greater importance. The analysis of the 125 features on an hourly basis from 1 January to 31 December 2023 (the ranking details have been omitted) reveals that 13 of the top 20 most important features are brightness temperature differences between different channels at the same time point, 6 are single-channel brightness temperature values, and 1 is the maximum value of a single channel across multiple time points. The results demonstrate that the most significant information for satellite precipitation inversion is derived from the differences between various channels at the same time point, with crucial insights present in the mid-wave infrared to far-infrared bands. The features ranked 21–40 are primarily composed of brightness temperature differences between the same channel at different times. Overall, the importance of brightness temperature differences at 15 min, 30 min, 45 min, and 60 min decreases in a sequential manner. This information is crucial for improving the quality of satellite precipitation estimation products.

4. Accuracy Verification of Precipitation Estimation Products from Meteorological Satellites Based on Artificial Intelligence

The methodology employed in this study for precipitation estimation inversion utilising multi-temporal satellite data based on artificial intelligence requires only historical 1 h ground rain gauge observations as input data for training the model. In other words, precipitation estimation at 01:00 is not based on the rain gauge data collected at that specific time point. This indicates that satellite precipitation estimation does not incorporate rain gauge observation data. In order to evaluate the accuracy of the FY-4B_AI satellite precipitation estimation product, the 1 h precipitation data from ground rain gauge stations within the study area is used as the truth value for the test. This study analyses and compares the FY-4B_AI and GPM/IMERG-L satellite-derived data, conducting statistical analysis by time (month) and region (along the 400 mm rainfall line), respectively.

4.1. Overall Accuracy Assessment

From 1 January to 31 December 2023, two types of satellite precipitation estimation (FY-4B_AI and GPM/IMERG-L) and ground rain gauge precipitation were conducted for the purpose of evaluating the accuracy of the former two through the time-space matching method. The findings of the precipitation accuracy assessment are presented in Table 2.

Table 2. Overall evaluation indices of satellite-derived precipitation products by FY-4B_AI (15 min/0.04°) and GPM/IMERG-L (30 min/0.1°) in China in 2023.

Index	FY-4B_AI	GPM/IMERG-L
ME	−0.88	−0.80
MAE	1.67	1.92
RMSE	3.68	4.07
RE	17.72%	26.28%
CC	0.44	0.36
POD	61.84%	47.31%
FAR	62.85%	62.21%
CSI	0.30	0.27

As illustrated in Table 2, the results demonstrate that six of the eight evaluation indices exhibit superior performance for FY-4B_AI in comparison to GPM/IMERG-L. These indices include the mean absolute error (MAE), root mean square error (RMSE), relative error (RE), correlation coefficient (CC), probability of detection (POD), and critical success index (CSI). In particular, the relative error (RE) of FY-4B_AI is 17.72%, which is significantly lower than that of GPM/IMERG-L (26.28%).

4.2. Monthly Variation Characteristics of Satellite-Derived Precipitation Accuracy

The monthly distribution maps of the average precipitation inversion evaluation for FY-4B_AI and GPM/IMERG-L from January to December 2023 (Figure 3) demonstrate that a number of indicators display distinctive annual variation characteristics. The monthly average indicators for the two types of satellite-derived precipitation evaluation demonstrate that FY-4B_AI is superior to GPM/IMERG-L. In particular, FY-4B_AI demonstrates superior performance in terms of MAE, RMSE, CC, and POD when compared to GPM/IMERG-L. The MAE and RMSE are observed to be larger during the summer half-year (April to September) and smaller during the winter half-year (October to March). The MAE and RMSE of FY-4B_AI are superior to those of GPM/IMERG-L, reaching a minimum in December 2023 and a maximum in July 2023, indicating a potential relationship between these variables and seasonal variations in precipitation. The CC and POD also demonstrate that FY-4B_AI outperforms GPM/IMERG-L, although the correlation with the season is not evident.

The ME of GPM/IMERG-L is observed to be the smallest in March 2023 and the largest in August 2023. The ME of FY-4B_AI is superior to that of GPM/IMERG-L in January, February, October, and December 2023. As stated by Zhao et al. [27], GPM/IMERG-L lacks the requisite sensitivity to accurately detect precipitation events during the winter season. It may therefore be assumed that the ME of FY-4B_AI is superior to that of GPM/IMERG-L in the context of winter precipitation. However, this is based on a single year of data (2023), and more years of data are required to verify this hypothesis. The distribution of the RE is similar to that of the ME. With regard to the CSI, GPM/IMERG-L outperforms FY-4B_AI only in July and August 2023, while FY-4B_AI outperforms GPM/IMERG-L in all other months. The FAR is approximately equivalent for FY-4B_AI and GPM/IMERG-L, which is consistent with the results shown in Table 2.

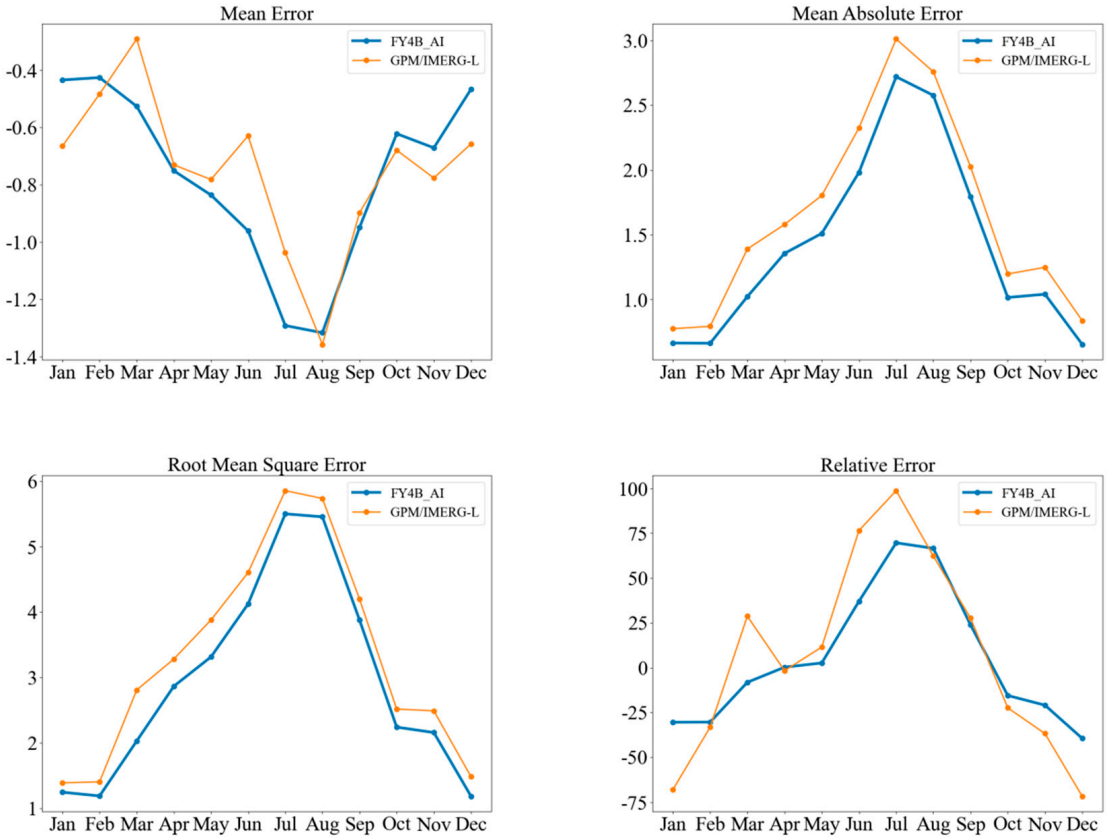


Figure 3. *Cont.*



Figure 3. Monthly distribution of satellite−derived precipitation accuracy evaluation indices from January to December 2023.

4.3. Spatial Variation Characteristics of Satellite-Derived Precipitation Accuracy (Northwest Dry Region/Southeast Humid Region)

As illustrated in Figure 1, ground rain gauge stations are densely distributed in the low-altitude regions of eastern and central China, whereas they are relatively sparse in the high-altitude areas of western China and northern regions such as Inner Mongolia. The advantages of satellite observation of precipitation include continuous monitoring, suitability for all weather conditions, and high spatial and temporal resolution. Consequently, satellite precipitation products are well-suited to high-altitude areas, such as western China and northern Inner Mongolia. However, it is important to assess the accuracy of these products in different regions of China, including the northwest dry region and the southeast humid region.

The modern 400 mm isohyet from the Greater Khingan Mountains to the southwest passes through Zhangjiakou in Hebei Province, Lanzhou, in Gansu Province, Lhasa, in Tibet, and reaches the eastern Himalayas. It can be statistically identified as a significant boundary separating monsoon and non-monsoon regions, as well as the arid and semi-arid regions in China. Additionally, it demarcates the transition between farming and nomadic herding practices. Furthermore, the correlation between climate and population density is evident along this line, with the semi-arid region situated on one side and the humid region situated on the other (as indicated by the red line in Figure 4).

Following an analysis of the overall and monthly accuracy characteristics of the two precipitation products, FY-4B_AI and GPM/IMERG-L, in 2023, this subsection employs the 400 mm average rainfall line in 2023 as a boundary (Figure 4) to differentiate between the northwest dry region and the southeast humid region. This investigation examines the precipitation distribution characteristics on either side of the aforementioned boundary line and compares the accuracy of two distinct precipitation products within these two regions. The results demonstrate that FY-4B_AI outperforms GPM/IMERG-L in seven

out of eight evaluation indices in both the northwest dry region and the southeast humid region (Tables 3 and 4). Notably, the RE of FY-4B_AI is considerably lower than that of GPM/IMERG-L.

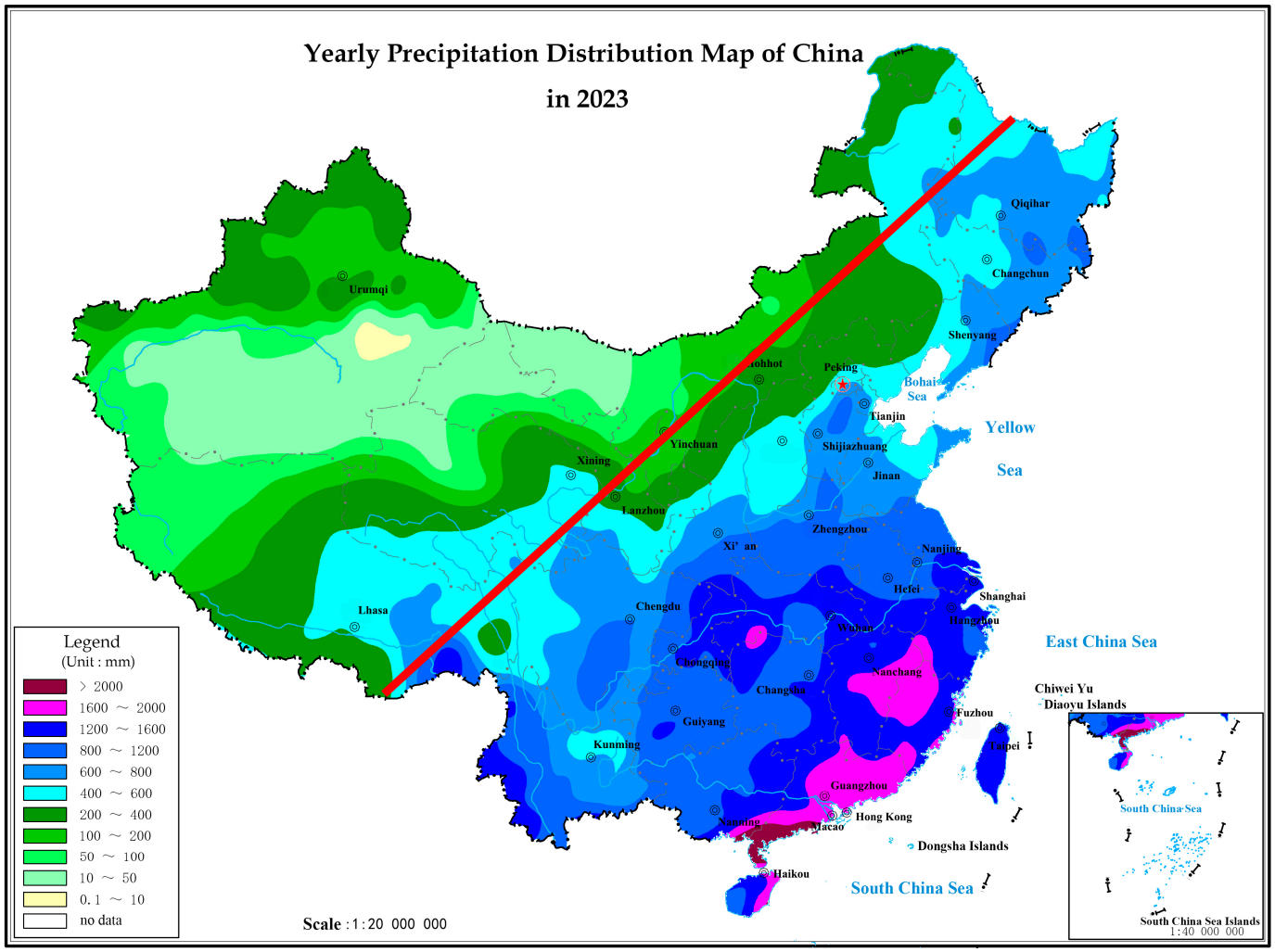


Figure 4. Yearly precipitation distribution map of China in 2023 (unit: mm) (cited from Figure 1.15 of the 'China Climate Bulletin (2023)').

Table 3. Overall evaluation indices of satellite-derived precipitation products by FY-4B_AI (15 min/0.04°) and GPM/IMERG-L (30 min/0.1°) in China's northwest dry region in 2023.

Index	FY-4B_AI	GPM/IMERG-L
ME	−0.73	−0.79
MAE	1.24	1.37
RMSE	2.64	2.83
RE	−3.78%	−20.73%
CC	0.35	0.32
POD	59.05%	38.25%
FAR	72.89%	65.45%
CSI	0.23	0.22

Table 4. Overall evaluation indices of satellite-derived precipitation products by FY-4B_AI (15 min/0.04°) and GPM/IMERG-L (30 min/0.1°) in China's southeast humid region in 2023.

Index	FY-4B_AI	GPM/IMERG-L
ME	−0.91	−0.80
MAE	1.76	2.03
RMSE	3.85	4.27
RE	21.95%	35.53%
CC	0.45	0.36
POD	62.39%	49.09%
FAR	60.10%	61.66%
CSI	0.32	0.27

5. Application Evaluation of Strong Weather Events

The preceding analysis demonstrates that the FY-4B_AI satellite precipitation inversion is, in general, more accurate than the GPM/IMERG-L product in six out of the eight evaluation indices. To assess the applicability of these two satellite precipitation inversion products in typical strong weather events over mainland China, this study evaluates their performance in two strong weather events that occurred in the southeast humid region in June 2023 and the northwest dry region in July 2023.

5.1. Application Evaluation of the Strong Weather Event in Guizhou Province on 18 June 2023 (Southeast Humid Region)

From 00:00 to 00:00 (UTC, as defined below) on 18–19 June 2023, Guizhou Province was subjected to a substantial precipitation event. It is of note that Puding Station in Guizhou Province recorded a 24 h rainfall of 194.4 mm, representing a heavy rainstorm and ranking first in the 24 h rainfall list. This is the second strongest rainfall in June locally. The heavy rainfall at Puding Station was primarily concentrated between 13:00 and 14:00 on 18 June, with a brief period of intense rainfall reaching 105.1 mm/h. This section employs the 1 h rainfall data from Puding Station between 13:00 and 14:00 on 18 June as a case study to assess the efficacy of the proposed methodology.

Figure 5 illustrates the overlay of infrared cloud images and rainfall at five consecutive time points (13:00, 13:15, 13:30, 13:45, and 14:00) for FY-4B_AI, ground station truth values (13:00 and 14:00), and GPM/IMERG-L (13:00, 13:30, and 14:00). The results indicate the following: (1) In comparison to the ground station truth values, both FY-4B_AI and GPM/IMERG-L demonstrate an accurate reproduction of the spatial distribution characteristics of precipitation. However, due to the spatial resolution of FY-4B_AI being 0.04° and that of GPM IMERG-L being 0.1°, FY-4B_AI provides a more detailed distribution of precipitation. (2) Additionally, the temporal resolution of FY-4B_AI is 15 min, whereas that of GPM IMERG-L is 30 min. Consequently, during the one-hour period from 13:00 to 14:00, FY-4B_AI (15 min/time) updates five times, while GPM/IMERG-L (30 min/time) updates three times. In contrast, the station updates only twice (60 min/time). This clearly demonstrates the high temporal resolution characteristics of FY-4B_AI. (3) Both FY-4B_AI and GPM/IMERG-L are capable of capturing the heavy precipitation zone located in western Guizhou province (Puding station); however, the centre of the precipitation zone in GPM/IMERG-L is shifted. A similar conclusion can be drawn with regard to the distribution of heavy precipitation zones in eastern Hunan Province.

As mentioned before, we also calculated the eight indices during this heavy rainfall event (Table 5). The results showed similar conclusions as in Table 2. Six of the eight evaluation indices exhibit superior performance for FY-4B_AI in comparison to GPM/IMERG-L. These indices include MAE, RMSE, RE, CC, POD, and CSI. In particular, the RE of FY-4B_AI is 142.013%, which is significantly lower than that of GPM/IMERG-L (255.706%) and much greater than that in Table 2.

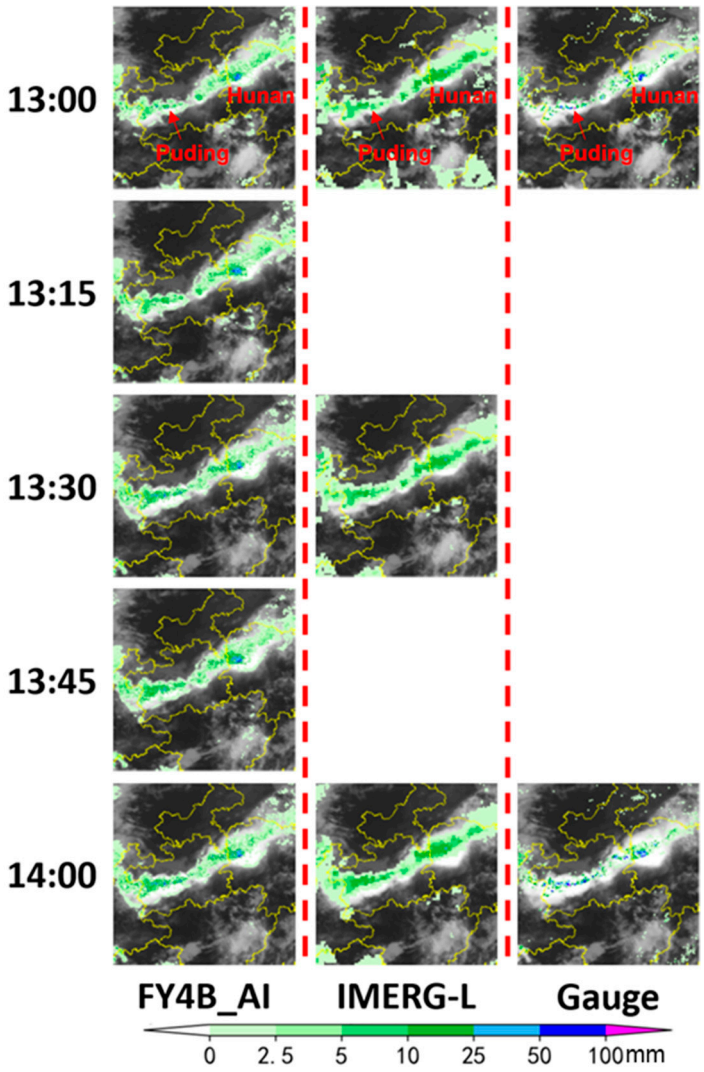


Figure 5. Overlay of FY-4B_AI, GPM/IMERG-L, ground station truth values, and infrared cloud images from 13:00 to 14:00 (UTC) on 18 June 2023.

Table 5. Overall evaluation indices of satellite-derived precipitation products by FY-4B_AI (15 min/0.04°) and GPM/IMERG-L (30 min/0.1°) in South China from 13:00 to 14:00 (UTC) on 18 June 2023.

Index	FY-4B_AI	GPM/IMERG-L
ME	−0.843	−0.288
MAE	1.951	2.591
RMSE	4.988	6.029
RE	142.013%	255.706%
CC	0.438	0.307
POD	62.129%	56.709%
FAR	60.962%	60.726%
CSI	0.315	0.302

5.2. Application Evaluation of the Strong Weather Event in Inner Mongolia on 20 July 2023 (Northwest Dry Region)

From the late afternoon of 20 July 2023 onwards, a number of robust convective cloud clusters emerged in the central and southeastern regions of Inner Mongolia. As they proceeded in an eastward direction, these clusters merged and intensified, resulting

in the occurrence of strong convective weather phenomena, including brief but heavy precipitation and winds reaching speeds of 8 to 10 on the Beaufort scale.

Figure 6 presents an overlay of infrared cloud images and rainfall at five consecutive hours (09:00, 10:00, 11:00, 12:00, and 13:00) on 20 July 2023, for FY-4B_AI, ground station truth values, and GPM/IMERG-L. The results indicate that in the sparsely populated eastern and central Inner Mongolia, FY-4B_AI once again demonstrates its high spatial resolution characteristics. Both FY-4B_AI and GPM/IMERG-L accurately reproduce the spatial distribution characteristics of precipitation. However, the GPM/IMERG-L product displays a tendency to overrepresent the range and magnitude of precipitation compared to the ground truth, with some deviation in capturing the centre of precipitation. In contrast, the FY-4B_AI precipitation estimations are closer to the actual situation.

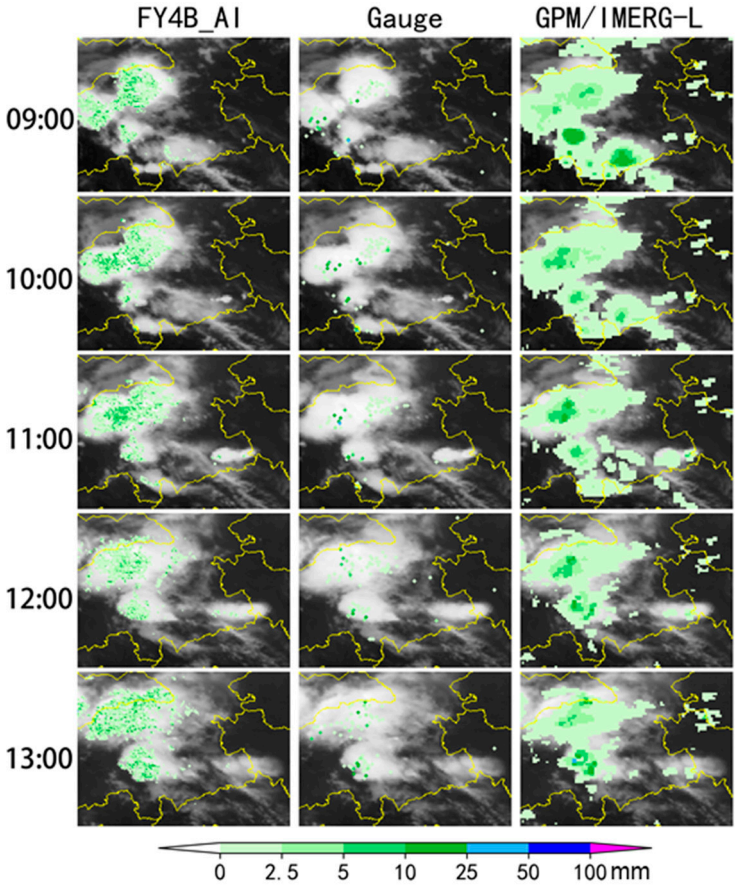


Figure 6. Overlay of FY-4B_AI, ground station truth values, and GPM/IMERG-L with infrared cloud images from 09:00 to 13:00 on 20 July 2023.

Also, we calculated the eight indices during this strong convective weather event (Table 6). FY-4B_AI and GPM/IMERG-L performed equally. For each, four of the eight evaluation indices are superior to those of the other, which is vastly different from the results that appeared in Tables 2 and 5. This may be related to the weak rain and strong wind in this weather event.

Although the evaluation indices of the two cases are not the same, the results of the evaluation of the two strong weather events indicate that both FY-4B_AI and GPM/IMERG-L are capable of accurately reproducing the spatial distribution characteristics of precipitation, irrespective of whether the region in question is the southeast humid region or the northwest dry region. However, due to its higher spatiotemporal resolution, FY-4B_AI not only provides more detailed distribution characteristics of precipitation but also updates more frequently, thereby making it a valuable indicator of strong precipitation.

Table 6. Overall evaluation indices of satellite-derived precipitation products by FY-4B_AI (15 min/0.04°) and GPM/IMERG-L (30 min/0.1°) in North China from 9:00 to 13:00 (UTC) on 20 July 2023.

Index	FY-4B_AI	GPM/IMERG-L
ME	−0.483	−0.536
MAE	0.599	0.598
RMSE	2.556	2.570
RE	−39.461%	−62.465%
CC	0.012	0.006
POD	20.000%	21.111%
FAR	99.687%	99.339%
CSI	0.003	0.006

6. Discussions

This study examined a multi-temporal FY-4B meteorological satellite precipitation estimation method based on artificial intelligence (FY-4B_AI) and trained a precipitation inversion model for mainland China. This satellite precipitation inversion model is founded upon findings pertaining to the characteristics of precipitation inversion in satellite channels. By analysing the channel characteristics of the multi-channel radiometer and its temporal changes regarding potential contributions to precipitation inversion, this study initially constructed multiple derived features based on satellite channel brightness temperature. Subsequently, a regional precipitation estimation model was constructed using the LGBM algorithm, thereby enhancing the efficiency of precipitation inversion. This algorithm continuously integrates the latest available ground precipitation data to update the model iteratively, without relying on simultaneous ground-based precipitation observations. This significantly enhances the effectiveness of the inversion product, which is of crucial importance for disaster weather warnings.

The data study used in this study are based on the year 2023; thus, it is essential to consider the limitations of the data and to conduct further verification using a larger number of individual cases. Also, different rainfall intensities, such as light, moderate, heavy, and torrential rain, are not classified.

7. Conclusions

The present study investigated a multi-temporal FY-4B meteorological satellite precipitation estimation method based on artificial intelligence and derived a total of 125 features from five time points. The methodology employed for the derivation of features from satellite channel brightness temperature data, along with the selection process for training model data, was elucidated in detail. The method was subsequently applied to estimate precipitation from the FY-4B meteorological satellite for the period between 1 January and 31 December 2023, with the resulting precipitation products then evaluated in terms of their accuracy. This study employed ground-based 1 h precipitation observation data within the study area as the verification truth, comparing and analysing two types of satellite-derived precipitation products, FY-4B_AI and GPM/IMERG-L. Furthermore, this study conducted application assessments of two strong convecive weather events. The main conclusions are as follows:

(1) The comprehensive accuracy assessment demonstrated that six of the eight evaluation indices exhibited superior performance in FY-4B_AI compared to GPM/IMERG-L, namely, MAE, RMSE, RE, CC, POD, and CSI. Of particular note is the significantly lower RE of FY-4B_AI in comparison to GPM/IMERG-L.

(2) The monthly distribution of precipitation accuracy revealed that a number of indicators displayed distinctive annual variation characteristics. The FY-4B_AI model displays superior performance in MAE, RMSE, CC, and POD. It can be observed that MAE and RMSE are larger during the summer half-year and smaller during the winter half-year. This suggests a correlation between these features and seasonal variations

in precipitation. Furthermore, the CC and POD demonstrate that the FY-4B_AI model outperforms GPM/IMERG-L, although the correlation with the season is not evident. The ME and RE of the FY-4B_AI are superior to those of the GPM/IMERG-L in the context of winter precipitation.

(3) The spatial distribution of precipitation accuracy demonstrates that FY-4B_AI outperforms GPM/IMERG-L in seven out of eight evaluation indices across both the northwest dry region and the southeast humid region. Of particular note is the significantly lower RE exhibited by FY-4B_AI in comparison to GPM/IMERG-L.

Above all, FY-4B_AI exhibited a superior capacity to delineate the distribution of precipitation with greater precision. Additionally, given its more frequent updates, FY-4B_AI demonstrated a noteworthy degree of indicative significance in accurately representing strong precipitation events.

Author Contributions: Conceptualization, J.J., N.L. and S.R.; methodology, N.L. and J.J.; software, N.L.; validation, N.L., J.J. and Y.L.; formal analysis, N.L. and J.J.; investigation, D.M. and M.F.; resources, N.L. and J.J.; data curation, N.L. and Y.L.; writing—original draft preparation, N.L. and B.H.; writing—review and editing, J.J.; visualization, N.L.; supervision, D.M., M.F. and S.R.; project administration, D.M. and M.F.; funding acquisition, D.M. and M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Policy Research on Meteorological Soft Science Project of China Meteorological Administration in 2024 'Study on Strategies for Enhancing the Supporting Capacity of Ocean Meteorological Services Based on Fengyun Satellites' (2024ZDIANXM08).

Data Availability Statement: The FY-4B data used in this study can be downloaded from https://satellite.nsmc.org.cn (accessed on 1 June 2022).

Acknowledgments: We would like to express our gratitude to Dapeng Huang of the National Climate Center for his assistance. Also, this manuscript has benefited greatly from the constructive comments and helpful suggestions of four anonymous reviewers.

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

References

1. Xu, R.; Tian, F.Q.; Yang, L.; Hu, H.; Lu, H.; Hou, A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over the southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.* **2017**, *122*, 910–924. [CrossRef]
2. Zambrano-Bigiarini, M.; Nauditt, A.; Birkel, C.; Verbist, K.; Ribbe, L. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1295–1320. [CrossRef]
3. Baroentti, A.; Acquaotta, F.; Fratianni, S. Rainfall variability from a dense rain gauge network in north-west Italy. *Clim. Res.* **2018**, *75*, 201–213. [CrossRef]
4. Tang, G.Q.; Behrangi, A.; Long, D.; Li, C.; Hong, Y. Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *J. Hydrol.* **2018**, *559*, 294–306. [CrossRef]
5. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]
6. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Woleff, D.B. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scale. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]
7. Zhang, M.; Lu, N.; Gu, S.; Zhang, W. Temperature-sounding Microwave Channels for FY-3(02). *J. Appl. Meteorol. Sci.* **2012**, *23*, 223–230. (In Chinese)
8. Grody, N.C. Classification of snow cover and precipitation using the Special Sensor Microwave Imager. *J. Geophys. Res. Atmos.* **1991**, *96*, 7423–7435. [CrossRef]
9. Kummerow, C.; Barnes, W.; Kozu, T.; Shiue, J.; Simpson, J. The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. Ocean. Technol.* **1998**, *15*, 809–817. [CrossRef]
10. Li, W.; Chen, Y.; Zhu, Y.; Zhao, B. Retrieval of rain over land by using TRMM/TMI measurements. *Acta Meteorol. Sin.* **2001**, *59*, 591–601. (In Chinese)
11. Li, X.; Yang, H.; You, R.; Zhao, F.; Qiao, Y. Remote sensing typhoon Songda's rainfall structure based on Microwave Radiation Imager of FY-3B satellite. *Chin. J. Geophys.* **2012**, *55*, 2844–2853. (In Chinese)
12. Li, P.; Xu, Z.; Ye, C.; Ren, M.; Chen, H.; Wang, J.; Song, S. Assessment on IMERG V06 Precipitation Products Using Rain Gauge Data in Jinan City, Shandong Province, China. *Remote Sens.* **2021**, *13*, 1241. [CrossRef]

13. Xie, S.; Sun, X.; Zhang, S.; Xiong, Z.; Wei, X.; Cui, C. Precipitation forecast correction in South China based on SVD and machine learning. *J. Appl. Meteor. Sci.* **2022**, *33*, 293–304. (In Chinese) [CrossRef]
14. Zhou, K.; Zheng, Y.; Han, L.; Dong, W. Advances in Application of Machine Learning to Severe Convective Weather Monitoring and Forecasting. *Meteorol. Mon.* **2021**, *47*, 274–289. (In Chinese)
15. Li, D.; Lin, W.; Liu, Q.; Feng, H.; Hu, S.; Wang, Z. Application of machine learning to statistical evaluation of artificial rainfall enhancement. *J. Appl. Meteor. Sci.* **2024**, *35*, 118–128. (In Chinese)
16. Zhang, Y.; Zhang, Y.; Zhai, D.; Liu, B.; Zhou, G. Thoughts on application and improvement of deep learning in severe precipitation nowcasting technology. *Torrential Rain Disasters* **2022**, *41*, 506–514. (In Chinese)
17. Chen, Y.; Cao, Y.; Sun, J.; Fu, J.; Dong, Q.; Yu, C.; Liu, C.; Tang, J.; Guo, Y. Progress of Fine Gridded Quantitative Precipitation Forecast Technology of National Meteorological Centre. *Meteorol. Mon.* **2021**, *47*, 655–670. (In Chinese)
18. Fang, W.; Shen, L.; Zou, L.; Pang, L. Extrapolation method of precipitation nowcasting radar echo based on GCA-ConvLSTM prediction network. *Torrential Rain Disasters* **2023**, *42*, 427–436. (In Chinese)
19. Moraux, A.; Dewitte, S.; Cornelis, B.; Munteanu, A. Deep learning for precipitation estimation from satellite and rain gauge measurements. *Remote Sens.* **2019**, *11*, 2463. [CrossRef]
20. Kidd, C.; Matsui, T.; Ringerud, S. Precipitation retrievals from passive microwave cross-track sensors: The precipitation retrieval and profiling scheme. *Remote Sens.* **2021**, *13*, 947. [CrossRef]
21. Liu, N.; Ren, S.; Jiang, J.; Qin, D.; Han, B. AI-based estimation of precipitation in the Tibetan Plateau using multi-temporal FY-4A satellite data. *Int. J. Remote Sens.* **2023**, *44*, 6523–6547. [CrossRef]
22. Xian, D. Fengyun-4B. *Appl. Satell.* **2021**, *7*, 4. (In Chinese) [CrossRef]
23. Huffman, G.J.; Stocker, E.F.; Bolvin, D.T.; Nelkin, E.J.; Tan, J. *GPM IMERG Late Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06*; Goddard Earth Sciences Data and Information Services Center (GES DISC): Greenbelt, MD, USA, 2019. [CrossRef]
24. Wang, J.; Wang, H.; Hong, Y. Comparison of satellite-estimated and model-forecasted rainfall data during a deadly debris-flow event in Zhouqu, Northwest China. *Atmos. Ocean. Sci. Lett.* **2016**, *9*, 139–145. [CrossRef]
25. He, S.; Wang, J.; Wang, H. Hindcast study of "6.18" Mentougou debris-flow event based on satellite rainfall and WRF forecasted rainfall. *Chin. J. Atmos. Sci.* **2018**, *42*, 590–606. (In Chinese)
26. Zhang, L.; Kang, Y.; Yue, Q.; Tang, J.; Xu, J.; Wang, J.; Hao, Z. Analysis of the applicability of various satellite-based precipitation in the source region of Yellow River. *Yellow River* **2021**, *43*, 29–33. (In Chinese)
27. Zhao, B.; Hudak, D.; Rodriguez, P.; Mekis, E.; Brunet, D.; Eckert, E.; Melo, S. Assessment of IMERGv06 Satellite Precipitation Products in the Canadian Great Lakes Region. *J. Hydrometeorol.* **2023**, *24*, 1017–1037. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.