geomatics

MDPI

Article

A Hybrid ConvNeXt, Dual-Transformer, and U-Net Architecture for Direct Satellite-Based Precipitation Estimation from MSG/SEVIRI

Mounir Sehad [1,*], Yacine Mohia [1], Mourad Lazri [1], Fethi Ouallouche [1], Karim Labadi [2], Rafik Absi [2] and Samir Hamaci [2]

[1] LAMPA Laboratory— Laboratoire d'Analyse et de Modelisation des Phenomènes Aleatoires, Department of Electronics, Faculty of Electrical Engineering and Computer Science, Mouloud MAMMERI University of Tizi Ouzou, Tizi Ouzou 15000, Algeria; yacine.mohia@ummto.dz (Y.M.)
[2] LR2E Laboratory Quartz, ECAM-EPMI, Ecole Supérieure d'Ingénieurs, 95092 Cergy Pontoise, France
* Correspondence: mounir.sehad@ummto.dz

Abstract

Accurate quantification of precipitation is critical for anticipating extreme cli mate events such as droughts and flash floods, yet ground rain gauges and weather radars in Northern Algeria suffer from limited spatial coverage and high maintenance costs. Satellite remote sensing via Meteosat Second Generation (MSG/SEVIRI) multispectral imagery offers optimal spatio-temporal coverage, but converting indirect cloud-top properties into direct rainfall rates remains challenging. To overcome the limitations of standard machine learning and conventional convolutional neural networks in modeling long-range atmospheric dependencies, this study proposes a novel end-to-end hybrid deep learning architecture, combining ConvNeXt, Dual-Transformer, and U-Net models, for direct pixel-wise precipitation regression without intermediate classification. The model leverages a ConvNeXt backbone with large-kernel depthwise convolutions for local spatial feature extraction, a Dual-Transformer module utilizing parallel spatial and spectral attention to capture non-local cloud structures, and a U-Net decoder with multi-scale skip connections for high-resolution rainfall reconstruction. Trained and validated against reference target fields constructed from weather radar and rain gauge fusion covering 2006–2012, the proposed ConvNeXt + Dual-Transformer + U-Net framework demonstrates superior performance over traditional machine learning baselines (RF, MLP, SVM) and standard CNN architectures (U-Net, ConvNeXt + U-Net). The model effectively reduces heteroscedastic error dispersion, achieving Pearson correlation coefficients (R) of 0.78 for instantaneous estimates, 0.84 for 3 h accumulations, 0.89 for daily totals, and 0.96 at the weekly scale, alongside lower root mean square errors (RMSE = 12.4 mm/day) and a minimal bias (0.1 mm/week). These results confirm that integrating global attention mechanisms with convolutional backbones enhances the modeling of complex convective rainfall, providing a robust operational tool for continuous satellite-based precipitation retrieval over Northern Algeria.

Keywords: precipitation estimation; deep learning; ConvNeXt; Dual-Transformer; U-Net; MSG-SEVIRI; satellite meteorology

Check for updates

Academic Editor: Alessandro Vitale

Received: 13 August 2026
Revised: 28 September 2026
Accepted: 1 October 2026
Published: 4 October 2026

Copyright: © 2026 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.

1. Introduction

In recent years, the planet has been facing increasing climate instability, leading to extreme weather events manifested by severe droughts in some regions while others experi-

ence devastating floods. A key element in understanding and anticipating these situations lies in the accurate quantification of precipitation. However, measuring precipitation remains a complex task due to its high spatial and temporal variability [1,2]. Northern Algeria is a topographically complex and climatically vulnerable region, characterized by a semi-arid to Mediterranean climate with high spatio-temporal rainfall variability. The area is increasingly exposed to extreme hydrological events, ranging from severe droughts that threaten water security to localized, intense convective storms that cause devastating flash floods across major urban centers and agricultural basins. Quantifying precipitation accurately in this region is particularly challenging due to the complex terrain formed by the Tellian and Saharan Atlas mountain ranges, which strongly modulates cloud dynamics and rainfall distribution. Ground-based observation networks (rain gauges) in Northern Algeria remain sparse and unevenly distributed, particularly over high-altitude and complex topographies. Furthermore, weather radar coverage is spatially constrained and subject to beam blockage in mountainous zones. In this context, high-resolution satellite remote sensing using Meteosat Second Generation (MSG/SEVIRI) data provides an indispensable alternative for continuous, regional-scale precipitation monitoring over Northern Algeria.

Traditionally, precipitation is measured using ground-based rain gauges. Although these measurements are considered highly reliable at a point scale, they lack spatial representativeness, leaving entire regions—often difficult to access—without data [3]. While weather radars provide near real-time estimates of precipitation intensity, their coverage is limited by geographical constraints and high maintenance costs [3,4].

In this context, satellite remote sensing offers an essential complementary solution for regional-scale precipitation estimation. Meteorological satellites, particularly the Meteosat Second Generation (MSG) series equipped with the SEVIRI radiometer, provide optimal temporal and spatial coverage over North Africa [5]. The visible and infrared spectral channels of SEVIRI enable the extraction of cloud microphysical and optical properties, which are crucial for modeling precipitation processes [6]. However, these observations provide indirect signals, such as cloud-top brightness temperature and reflectance, requiring mathematical models to relate these parameters to rainfall rates [7,8].

Early methods based on thermal thresholding [1,7] have evolved toward advanced supervised machine learning techniques. Various approaches have demonstrated their effectiveness, including artificial neural networks [9], support vector machines (SVM) [10–13], random forests [14,15], and K-nearest neighbors (KNN) algorithms [16]. These models have been combined into hybrid structures or multi-classifier systems to improve estimation accuracy [17,18]. Nevertheless, their implementation heavily relies on the availability and quality of labeled datasets. Insufficient or poorly labeled training data may lead to overfitting or introduce biases that limit the generalization capability of the model [19,20].

To overcome these limitations, research has moved toward methods capable of leveraging large volumes of data. While semi-supervised approaches, such as S3VM optimized with metaheuristic algorithms like the Firefly Algorithm, have enabled the integration of unlabeled data to stabilize results [21–23], a new technological breakthrough has emerged with the advent of deep learning. Recent advances show that model performance no longer depends solely on data quantity, but also on the ability of the architecture to extract hierarchical and global contextual features [24].

Despite the progress of standard convolutional neural networks (CNNs), transitioning from discrete classification to continuous rainfall regression remains a major challenge. Conventional approaches often struggle to capture long-range dependencies within cloud structures. To address this issue, hybrid architectures combining modern convolutional structures and attention mechanisms have recently emerged [25,26]. In particular, the integration of the ConvNeXt backbone, which simulates a large spatial receptive field [27,28],

together with Transformers for spatio-spectral contextual modeling, provides a more suitable framework for pixel-wise regression.

Recent advances in deep learning have seen ConvNeXt, Vision Transformers, and U-Net architectures applied independently across various satellite remote sensing and precipitation retrieval tasks. Standalone ConvNeXt models leverage large depthwise convolutions to effectively capture local cloud textures and microphysical signatures; however, their purely localized receptive fields limit their ability to model long-range spatial interactions across distant atmospheric systems. Conversely, standalone Vision Transformers excel at capturing global spatio-spectral relationships, but they incur high computational costs and often struggle with fine pixel-level localization required for dense continuous rainfall regression. Meanwhile, traditional U-Net backbones remain popular for multi-scale spatial reconstruction, yet standard convolutional layers struggle to resolve the highly non-linear dynamics of localized convective storms. Integrating ConvNeXt for local feature extraction, Dual-Transformer modules for global contextual modeling, and U-Net for multi-scale reconstruction into a unified hybrid framework directly bridges these gaps, providing a complementary mechanism capable of simultaneously handling local spatial precision and broad regional dependencies.

The objective of this study is to propose a novel hybrid architecture, namely ConvNeXt + Dual-Transformer + U-Net, for direct precipitation estimation over Northern Algeria. Unlike approaches based on prior classification or conventional semi-supervised models, our contribution lies in the design of an end-to-end regression model. This model employs a ConvNeXt encoder for local feature extraction, enhanced by a Dual-Transformer module to capture global dependencies, all integrated within a U-Net structure for accurate multi-scale reconstruction. Validation is conducted through a rigorous comparison with radar and rain gauge measurements, demonstrating the model's ability to generalize effectively across different temporal scales (from hourly to weekly).

The manuscript is organized into five sections: Introduction, Method, Results, Discussion, and Conclusion. Section 2 (Method) describes the study area, the MSG-SEVIRI, rain gauge, and radar data; the preprocessing steps; and the detailed architecture of the developed model, including the mathematical foundations of ConvNeXt, the Dual-Transformer, and the U-Net architecture. Section 3 (Results) presents the validation results and their statistical analysis. Section 4 (Discussion) provides a discussion of the results. Finally, Section 5 (Conclusion) concludes the paper and outlines perspectives for future work.

2. Methods

2.1. Data and Preprocessing

2.1.1. Study Area

The study focuses on a portion of the territory located in North–East Algeria (see Figure 1). This region enjoys a Mediterranean climate characterized by the joint presence of both convective and stratiform precipitation mechanisms. The climate is of a warm temperate to subtropical type, with dry, scorching summers and mild, rainy winters. Summer precipitation is very low, often negligible or even non-existent. The northward migration of the subtropical anticyclone during this season generates convective-type rainy episodes. Conversely, winter constitutes the wettest period of the year, with the bulk of precipitation being provided by well-organized stratiform systems.

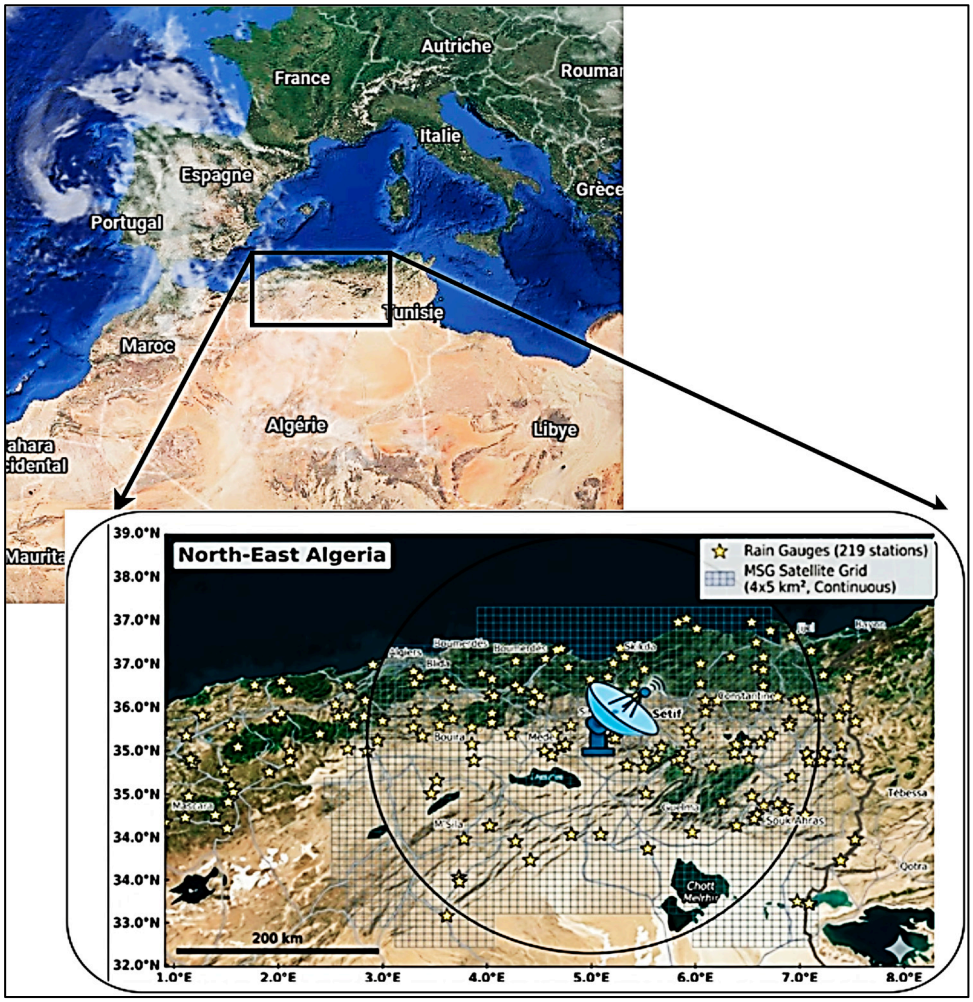


Figure 1. Geographic location of the study area in North–East Algeria, showing the distribution of the 219 rain gauges and the spatial coverage of the Sétif meteorological radar.

2.1.2. Data Used

This study uses datasets combining information from the MSG satellite with data obtained from weather radar and rain gauges. MSG data are exploited for precipitation estimation over the study region using the proposed methodology.

The fusion of radar and rain gauge data provides a major operational advantage, as it combines the complementary strengths of these two sources while overcoming their respective limitations. Rain gauges provide direct and highly accurate point measurements of precipitation at the ground level. However, their spatial distribution is often sparse, which limits their ability to accurately capture the spatial variability of precipitation, particularly during intense and localized convective events. In contrast, weather radar offers wide spatial coverage and high spatio-temporal resolution, enabling the monitoring of the structure, movement, and evolution of precipitating systems. Nevertheless, radar estimates are indirect measurements that may suffer from systematic errors (biases) due to several factors, such as signal attenuation, ground clutter, and uncertainties in the relationship between reflectivity and rainfall rate (Z–R relationship).

By merging these two types of data, a reliable dataset is obtained for both training and validation of the developed model, benefiting from:

(1) The quantitative accuracy of rain gauges, which serve as “ground truth” to correct systematic biases in radar estimates;

(2) The detailed spatial structure provided by radar, enabling a more realistic mapping of precipitation between ground-based measurement stations.

Radar and rain gauge data collected from January 2006 to November 2011 were used for training (80%), while additional datasets at different temporal scales (instantaneous, 3-hourly, 24-hourly, and 7-day accumulations) collected from December 2011 to December 2012 were used for model validation (20%). The choice of the 2006–2012 period was deliberate. During this period, meteorological radar data from the Sétif region were reliable because the radar was well calibrated. From 2013 onward, however, the only radar that remained operational in the study region provided unreliable data. For this reason, the post-2012 period was not used in our experiments. This selection ensures that the precipitation estimates are based on high-quality radar observations, which is essential for the validity and robustness of our results.

2.1.3. MSG Satellite Data

The data provided by the MSG satellite implicitly describe the optical and microphysical properties of cloud systems. These features are exploited in this study for precipitation characterization. The main instrument onboard the MSG satellites is the SEVIRI imaging radiometer, which acquires measurements across 12 spectral bands and scans the Earth every 15 min. Depending on the channel used, the spatial resolution at the ground ranges between 1 and 3 km.

The spectral set includes two visible bands (VIS 0.6 and VIS 0.8), one near-infrared band (NIR 1.6), and eight channels covering the thermal infrared and water vapor domains (IR 3.9, WV 6.2, WV 7.3, IR 8.7, IR 9.7, IR 10.8, IR 12.0, and IR 13.4). A broadband High Resolution Visible (HRV) channel is also available (Table 1).

Table 1. Input parameters of the model developed for precipitation estimation.

Channel	Role in Precipitation Estimation
VIS 0.6 µm	Used during daytime only, the reflectance measured in the visible channel at 0.6 µm (VIS0.6) makes it possible to estimate cloud optical thickness, a key indicator of precipitation potential (optically thick clouds being associated with a higher probability of precipitation).
NIR 1.6 µm	The near-infrared channel at 1.6 µm (NIR1.6) is used to discriminate cloud phase: clouds composed of liquid water exhibit high reflectance, whereas ice clouds appear significantly darker in this spectral band.
IR 3.9 µm	Day and night, excellent for detecting low clouds and fog. At night, it provides implicit information on droplet size at the cloud top.
IR 10.8 µm (CTT), IR 8.7 µm, IR 12.0 µm	These are the key channels. IR 10.8 µm is the reference; its brightness temperature is directly related to cloud altitude (colder = higher = more intense rainfall). IR 12.0 µm and IR 8.7 µm are used to discriminate cloud phase (ice/water) and composition.
WV 6.2 µm, WV 7.3 µm	Probe humidity in the upper (6.2 µm) and (7.3 µm) troposphere, key indicators of atmospheric instability. Temporal differences between these channels are related to rainfall occurrence.
ΔT(8.7–10.8) (CP) (IR8.7−IR10.8)	Excellent indicator of cloud phase. This difference is significantly higher for ice clouds than for liquid water clouds, helping to identify potentially precipitating systems.
ΔT(10.8–12.0) (CP) (IR10.8−IR12.0)	Used in combination to confirm cloud phase and is sensitive to water vapor in the lower layers.
ΔT(3.9–10.8) (CWP) (IR3.9−IR10.8)	Highly sensitive to droplet size at the cloud top. The larger the difference, the larger the droplets, indicating precipitating clouds.
ΔT(3.9–7.3) (CWP) (IR3.9−WV7.3)	Combines information on droplet size with mid-tropospheric humidity (7.3 µm channel) to refine detection.
ΔT(6.2–10.8) (CTH) (WV6.2−IR10.8)	Often positive for convective clouds; it indicates very cold cloud tops in a humid upper troposphere, revealing atmospheric instability favorable to heavy rainfall.

The pixels are encoded on 10 bits, corresponding to digital counts ranging from 0 to 1023 for all channels. These raw digital values can be converted into radiances and, for infrared and water vapor channels, into brightness temperatures. For visible channels, the conversion results in reflectance values [29].

2.1.4. Radar Measurements

In this study, radar observations are considered as a reliable reference for both the training and validation phases of the developed models. These data are acquired from the Sétif weather radar (SR), with a spatial resolution of 1 km × 1 km and a temporal resolution of 15 min.

The SR instrument operates in the C-band frequency range, centered at 5.6 GHz. It is located near the city of Sétif, at 36°01′ N, 5°25′ E, and at an altitude of 1700 m above sea level. Its main technical specifications include a wavelength of 5.5 cm, a maximum transmitted power of 250 kW, a pulse repetition frequency of 250 Hz, and a pulse width of 4 µs (the study area, located in northeastern Algeria, is shown in Figure 1).

The radar provides images in PPI (Plan Position Indicator) format, consisting of a 512 × 512 pixel matrix. Due to 4-bit encoding, each pixel is assigned one of 16 possible reflectivity levels, expressed in dBZ, according to the range (4, 12, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70) (an example of a radar image is given in Figure 2).

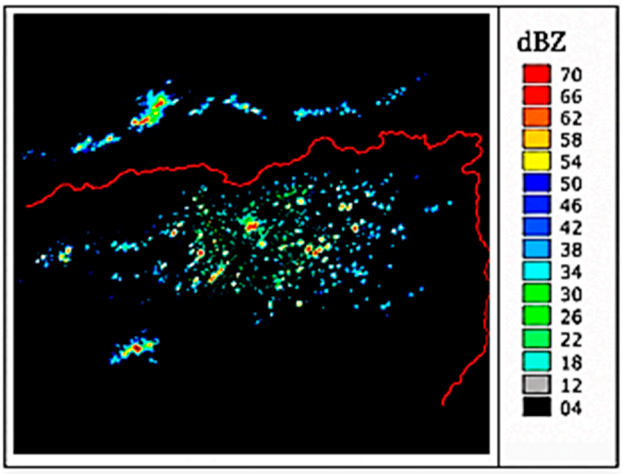


Figure 2. Image taken by the Sétif radar on 5 January 2009 at 3:33 a.m. shows the 16 reflectivity levels (dBZ).

The reflectivity factor Z represents the raw physical variable measured by the radar. Using a locally calibrated Z–R relationship [5], this parameter can be converted into rainfall rate R, expressed in millimeters per hour (mm/h), as given by the following relationship:

(a) For convective precipitation with $Z \geq 42$ Dbz

$$Z = 248.28 \times R^{1.58} \quad (1)$$

(b) For stratiform precipitation with $Z \leq 38$ dBZ

$$Z = 412.35 \times R^{1.49} \quad (2)$$

2.1.5. Rain Gauge Data

A network of approximately 219 rain gauge stations was used across the entire study area. These stations provide direct measurements of precipitation at a daily time scale. From these daily observations, it is possible to compute accumulated precipitation over a given period. Rain gauge data are used in this study for fusion with radar observations.

2.1.6. Fusion of Radar and Rain Gauge Data Using ARBF

Although radar data provide continuous spatial coverage, they are affected by biases and random errors, particularly during extreme events [30]. To overcome these limitations, radar and rain gauge data are fused.

Among the fusion methods developed in the literature (e.g., kriging, RBF, quantile mapping) [31,32], comparative studies have shown that geostatistical approaches such as kriging require more complex calibration compared to statistical methods [33].

The Radial Basis Function (RBF) method has recently demonstrated superior performance in terms of accuracy (lower RMSE and better spatial correlation) while requiring minimal calibration effort [34,35].

The radar data, denoted as $P_{\text{radar}}(x)$, where "x" represents the precipitation intensity at each pixel of the radar grid, whereas rain gauges provide point-based measurements at specific locations $P_{\text{gauge}}(x_i)$.

The residual (or error) at a rain gauge location x_i is defined as:

$$r(x_i) = P_{\text{gauge}}(x_i) - P_{\text{radar}}(x_i) \tag{3}$$

2.1.7. Adaptive Gaussian Radial Basis Function

The performance of the standard RBF method depends heavily on the scale parameter "σ" of the Gaussian function, which controls the size of the influence zone for each rain gauge. A single "σ" for the entire study area may not be optimal, especially if the rain gauge density is variable (e.g., denser in the plains, sparser in the mountains).

To address this limitation, we employ the Adaptive Radial Basis Function (ARBF) method. The core principle consists of varying the scale parameter "σ" spatially, adapting it to the local density of the measurement points:

$$\phi(x, x_j) = exp\left(-\frac{\| x - x_j \|^2}{\sigma(x)^2}\right) \tag{4}$$

with

$$\sigma(x) = \alpha \cdot d_k(x) \tag{5}$$

where

α is a global scaling factor (to be optimized) and

$d_k(x)$ is the average distance from station x to its k nearest neighbors.

For each interpolation point "x", a value of "σ" is defined according to the density of rain gauges in its surrounding neighborhood. In areas where stations are dense (and thus residuals are well constrained), a small value of "σ" is used. This allows the model to capture fine and highly localized variations in radar error, which are typical of convective storm events. Conversely, in regions where the network is sparse (i.e., where rain gauges are widely separated), a larger value of "σ" is applied. This enables the influence of each station to extend over a broader spatial area, thereby avoiding the introduction of artifacts or unrealistic discontinuities in the interpolated precipitation field.

Interpolation ARBF. The estimated residual at any point x is given by the following ARBF interpolation formulation, written using Equation (6)

$$\hat{r}(x) = \sum_{j=1}^{n} w_j \, \phi(x, x_j) \tag{6}$$

The weights w_j are obtained by imposing that the interpolation is exact at the measurement points:

$$\hat{r}(x_i) = r(x_i), \forall i = 1, \ldots, n \tag{7}$$

This leads to the linear system:

$$\Phi \mathrm{w} = \mathrm{r} \tag{8}$$

with:
Φ: Matrix $n \times n$ such that $\Phi_{ij} = \phi(|x_i - x_j|)$,
$w = [w_1, \ldots, w_n]^T$,
$r = [r(x_1), \ldots, r(x_n)]^T$.

Corrected precipitation map. Once the residuals have been interpolated over the entire radar grid, the fused precipitation map (ground truth) is obtained by:

$$P_{\text{fusion}}(x) = P_{\text{radar}}(x) + \hat{r}(x) \tag{9}$$

This map serves as the target (ground truth) for the training and validation of the model developed for precipitation estimation.

Parameter optimization by LOOCV cross-validation. The two key parameters (k, α) are determined by leave-one-out cross-validation [36,37]. For each tested pair (k, α) each rain gauge is successively removed, the residual at its location is interpolated using the remaining $n - 1$ stations, and the squared error is computed. The selected pair minimizes the average MSE:

$$\mathrm{MSE}(k, \alpha) = \frac{1}{n} \sum_{i=1}^{n} \left(r(x_i) - \hat{r}_{-i}(x_i)\right)^2 \tag{10}$$

Implementation. The ARBF is implemented according to the following steps:

(1) Explore a grid of values for k (for example $k = 3, 4, \ldots, 10$);
(2) Explore a grid for α (for example $\alpha = 0.5,\ 1.0,\ 1.5,\ 2.0,\ 2.5$);
(3) For each (k, α), compute the MSE using LOOCV;
(4) Retain the pair yielding the lowest MSE.

ARBF Fusion algorithm [38]. The complete algorithm can be summarized as follows:

(a) Residual computation $r(x_i)$ for each station i;
(b) Optimization: determination of the parameters (k, α) of the ARBF using LOOCV;
(c) ARBF interpolation: construction of the matrix "Φ" with the adaptive radial function, and solution of the system $\Phi \mathrm{w} = \mathrm{r}$ to obtain the weights w_j;
(d) Spatial correction: for each pixel "x" of the radar grid, computation of the estimated residual $\hat{r}(x)$ according to the ARBF equation;
(e) Fusion: $P_{\text{fusion}}(x) = P_{\text{radar}}(x) + \hat{r}(x)$.

The complete algorithm can be summarized as follows: We applied the ARBF algorithm in the study area using 219 rain gauges. The optimal values of $(k = 4,\ \alpha = 2)$ were obtained for: a minimum error of $MSE = 1.25\ \mathrm{mm}^2/\mathrm{h}^2$.

2.1.8. Matching Between Satellite and Radar Measurements

To resolve the spatial scale disparity between high-resolution weather radar data (1 km × 1 km) and coarser SEVIRI MSG imagery (4 km × 5 km over Northern Algeria), a two-step spatial preprocessing strategy was implemented:

(i) Spatial Resampling: High-resolution radar observations were reprojected onto the SEVIRI satellite grid using spatial area-weighted averaging, converting 1 km × 1 km radar pixels to match the satellite resolution, yielding a uniform grid of 128 × 102 pixels covering the study region; and (ii) Parallax Bias Mitigation: Due to the geostationary viewing angle, cloud tops appear spatially shifted relative to ground precipitation. To correct this geometric overlay error, an adaptive spatial smoothing filter was applied based

on cloud-top altitude: a 5 × 5 pixel neighborhood window was used for high-altitude convective clouds (where parallax displacement is severe), while a 3 × 3 pixel window was applied to low-altitude clouds.

2.1.9. Model Input Parameters

Different spectral channels and derived combinations from the SEVIRI radiometer onboard the MSG satellite have been widely used in the literature and have demonstrated their effectiveness in precipitation classification and estimation [14,39,40]. Among the 12 spectral bands (Table 1), we selected an optimal subset of 8 channels (Infrared: IR 8.7 µm, IR 10.8 µm, IR 12.0 µm, IR 3.9 µm; Water vapor: WV 6.2 µm and 7.3 µm; Visible: VIS 0.6 µm and NIR 1.6 µm), and 5 brightness temperature differences to characterize the radiative signatures associated with precipitation, defined as follows:

(1) **Cloud Top Temperature (CTT)** is directly assimilated to the brightness temperature measured in the thermal infrared channel centered at 10,8 µm (denoted $T_{10.8}$) [39,40].
(2) **Cloud Top Height (CTH)** is a reliable indicator for discriminating convective systems, which are characterized by a significantly higher top altitude than stratiform formations. This height is estimated from brightness temperature differences $\Delta T(WV_{6.2}–IR_{10.8})$ and $\Delta T(WV_{7.3}–IR_{12.0})$ [41,42].
(3) **Cloud Phase (CP)** represents a critical parameter in the microphysical processes of precipitation formation, particularly with regard to the presence of ice particles in the upper parts of clouds. It is derived from brightness temperature differences $\Delta T(IR_{8.7}–IR_{10.8})$ and $\Delta T(IR_{10.8}–IR_{12.0})$ [41,43–45].
(4) **Cloud Water Path (CWP)** provides information on potential precipitation rates. During daytime, it depends on optical thickness, derived from reflectance in the visible channel at 0.6 µm, and on the effective particle radius, estimated from the near-infrared channel at 1.6 µm. At night, its estimation relies on brightness temperature differences ΔT $(IR_{3.9}–IR_{10.8})$ and ΔT $(IR_{3.9}–WV_{7.3})$ [14,40,41].

The 13 spectral channels used as input parameters for the initial layer (Stem) of the ConvNeXt architecture are summarized in Table 1.

Meteosat Second Generation (MSG) SEVIRI imagery was selected over alternative satellite databases (such as LEO-based GPM/TRMM or reanalysis products) due to three key advantages: (i) Geostationary Footprint: Positioned at 0° longitude, MSG provides direct viewing geometry over Northern Algeria, minimizing geometric distortions over its complex coastal and mountainous topography; (ii) Operational Continuity: Unlike polar-orbiting satellites that exhibit temporal coverage gaps between overpasses, MSG provides continuous 24/7 observations, crucial for tracking fast-developing convective storms; and (iii) Sensor Consistency: Utilizing a single, stable geostationary platform eliminates cross-sensor calibration biases and interpolation noise common in multi-platform merged precipitation products.

2.1.10. Normalization of MSG Data

The brightness temperatures associated with the different spectral channels of the SEVIRI instrument, initially expressed in Kelvin (infrared and water vapor), were standardized using Z-score transformation, in accordance with Equation (1). For each channel c, the statistical parameters μ_c (mean) and σ_c (standard deviation) were estimated using only the training dataset. The resulting transformation was then applied uniformly to the training and validation datasets.

The combined channels, corresponding to brightness temperature differences, were processed using a similar procedure, with their own statistical parameters computed independently. This normalization, which is weakly sensitive to outliers, helps stabilize the con-

vergence of the ConvNeXt architecture while preserving the relative distributional structure of the spectral variables. The normalization of reflectance from the visible ($VIS_{0.6}$ µm) and near-infrared ($NIR_{1.6}$ µm) channels is performed using the same equation (Equation (11)), replacing brightness temperatures (T) with reflectance (R).

$$T'_{b,c} = (T_{b,c} - \mu_c)/\sigma_c \quad (11)$$

2.2. Proposed Architecture

In this study, we propose a hybrid deep learning architecture for pixel-scale precipitation estimation from MSG satellite images. The model is based on an end-to-end structure combining three main components:

1. A ConvNeXt backbone for local spatial feature extraction;
2. A Dual-Transformer module for global context modeling;
3. A U-Net architecture for reconstructing precipitation fields at the pixel scale.

Unlike conventional approaches, the proposed method enables direct estimation of precipitation intensity without an intermediate classification step (convective/stratiform), thereby reducing error propagation and improving the spatial consistency of the results. For a better description, the input data correspond to a multispectral satellite image denoted as:

$$X \in \mathbb{R}^{H \times W \times C} \quad (12)$$

where H, W and C represent respectively the height, the width, and the number of spectral channels. The model learns a regression function: $f_0 : X \rightarrow Y$, where $Y \in \mathbb{R}^{H \times W}$ corresponds to the estimated precipitation field.

The block diagram of our model is defined in Figure 3.

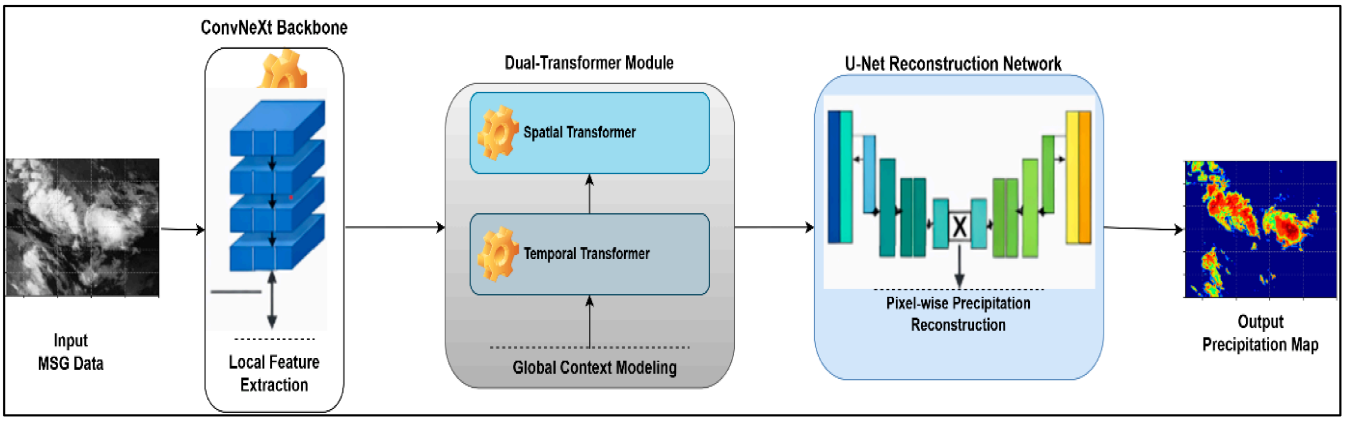


Figure 3. A Hybrid ConvNeXt–Dual-Transformer U-Net Framework for Pixel-Wise Precipitation Estimation from MSG Satellite Imagery.

The details of the proposed hybrid model ConvNeXt + Dual-Transformer + U-Net are described as follows.

2.2.1. ConvNeXt for Feature Extraction

The ConvNeXt backbone processes the input image $X \in \mathbb{R}^{HxWxC}$ through a series of modified residual blocks. The main innovation of ConvNeXt lies in the use of depthwise convolutions and an "inverted bottleneck" design.

Depthwise Convolution. ConvNeXt uses a 7 × 7 per-channel convolution to capture a larger receptive field, mimicking the spatial attention mechanism of Transformers F_l. At layer l, the operation is given by:

$$F_l' = DWConv_{7x7}(F_l) \tag{13}$$

where each channel is filtered independently, reducing the number of parameters while increasing spatial efficiency.

A standard 3 × 3 convolution does not sufficiently capture the spatial structure of the information. The 7 × 7 kernel used in ConvNeXt captures the "shape" of the cloud, which is crucial for regression tasks. The large 7 × 7 receptive field is particularly well suited to MSG data, as it allows the model to integrate the spatial organization of cloud tops, a key indicator of convective intensity.

ConvNeXt block (Inverted Bottleneck). After spatial convolution, the block applies an expansion of the channel dimension, followed by a reduction. The forward pass through a ConvNeXt block can be modeled by the following sequence of operations:

1. Normalization and Expansion

   This step can be described by the following expression

$$F_{ext} = Linear_{C \to 4C}(LayerNorm(F_l')) \tag{14}$$

2. Non-linear activation (GELU)

   The non-linear activation is represented by the following Equation (15)

$$F_{act} = \sigma_{GELU}(F_{ext}) \tag{15}$$

3. Compression and projection

   The compression and projection step is explicitly defined by the following relation 16.

$$F_{proj} = Linear_{4C \to C}(F_{act}) \tag{16}$$

4. Residual connection and block output

To facilitate gradient flow and enable the learning of identity functions, the final output of the block F_{l+1} is defined as:

$$F_{l+1} = F_l + \gamma . F_{proj} \tag{17}$$

where γ is a Layer Scale parameter (a learnable scalar) that stabilizes training by dynamically adjusting the magnitude of residual contributions. The use of LayerNorm and GELU allows the model to better converge toward accurate regression solutions. These blocks are organized into four hierarchical stages, separated by stride-2 convolution downsampling layers, enabling the model to capture features ranging from local cloud microphysics to synoptic-scale cloud system structures. The detailed ConvNeXt block is illustrated in Figure 4.

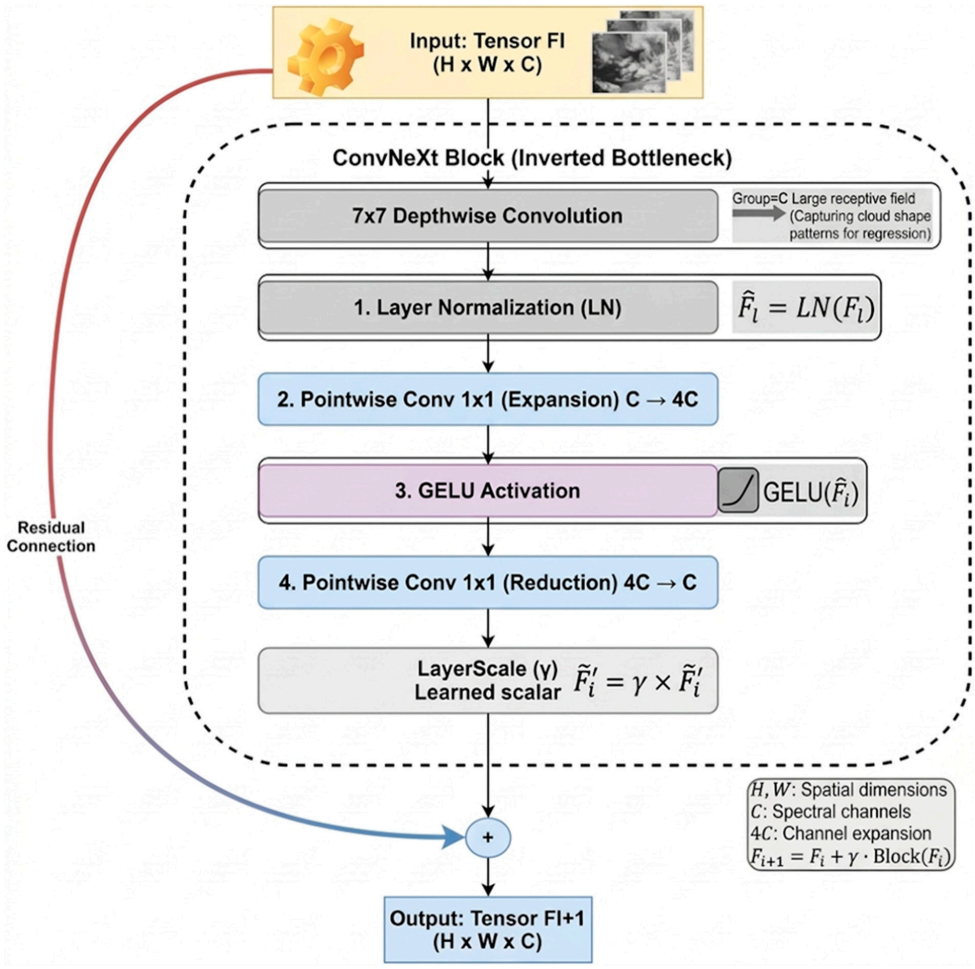


Figure 4. ConvNeXt Residual Block: Spatial Feature Extraction via Large-Kernel Depthwise Convolution.

2.2.2. Dual-Transformer for Contextual Modeling

The Dual-Transformer module processes the feature maps produced by the ConvNeXt backbone in order to capture non-local dependencies. It consists of two attention branches operating either in parallel or sequentially: Spatial Attention (SA) and Spectral/Channel Attention (CA).

Self-Attention mechanism. For an input $F \in R^{N \times D}$ (where $N = H \times W$ is the number of pixels and D is the feature dimension), the module computes three matrices: Queries (Q), Keys (K), and Values (V), obtained through linear projections $W_{_q}$, W_k, W_v:

$$Q = FW_q,\ K = FW_k\ ,\ V = FW_v \tag{18}$$

The attention is then computed using the scaled dot-product:

$$Attention(Q,\ K,\ V) = \ \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{19}$$

where $\sqrt{d_k}$ is a scaling factor used to stabilize gradients.

Duality of the module. The term "Dual" refers to the application of this mechanism along two axes:

(a) Spatial Attention (Spatial-wise) enables the correlation of each pixel with all other pixels in the image, regardless of distance. This is crucial for identifying whether an isolated convective cell is part of an organized system (e.g., a squall line). The complexity of pixel-to-pixel relationships is given by the equation $\mathbb{R}^{(HW)\ \times (HW)}$.

(b) The Channel Attention (Spectral-wise) models dependencies between different physical parameters (ΔT, R, water vapor). The model learns how combinations of these variables and their interactions relate to precipitation intensity. The complexity is described along the depth of the channel dimension $\mathbb{R}^{D\times D}$.

Integration into the Transformer Block. Each Dual-Transformer block follows a residual learning structure with layer normalization (LN) and a feed-forward network (MLP):

(a) Attention step: Z′ = F + Dual-Attention (LN(F));
(b) MLP step: Z = Z′ + MLP(LN(Z′)).

The MLP consists of two linear layers with a GELU activation function, enabling the transformation of attention-based representations into potential precipitation intensities.

Spectral Attention mechanism (Efficient Attention). Unlike spatial attention, which computes pixel-to-pixel relationships, spectral attention focuses on interactions between the *D* feature channels (physical signatures). To optimize computation, we use a linearized version of attention.

Linear transformation. Instead of computing a spatial affinity matrix of size (HW × HW), we first compute a global context matrix by multiplying Keys (K) and Values (V). The spectral context matrix M_{spec} is defined as:

$$M_{spec} = softmax_{row}\,(K)^T\,V \tag{20}$$

where $M_{spec} \in \mathbb{R}^{D\times D}$

This matrix represents a global knowledge base describing how the different spectral channels interact with each other across the entire image.

Attention application. The output of the spectral attention for the Queries (Q) is then obtained by:

$$\text{Spectral} - \text{Attention}(Q,\ K,\ V)\ = softmax_{col}(Q)\ M_{spec} \tag{21}$$

A schematic illustrating the Dual-Transformer is shown in Figure 5.

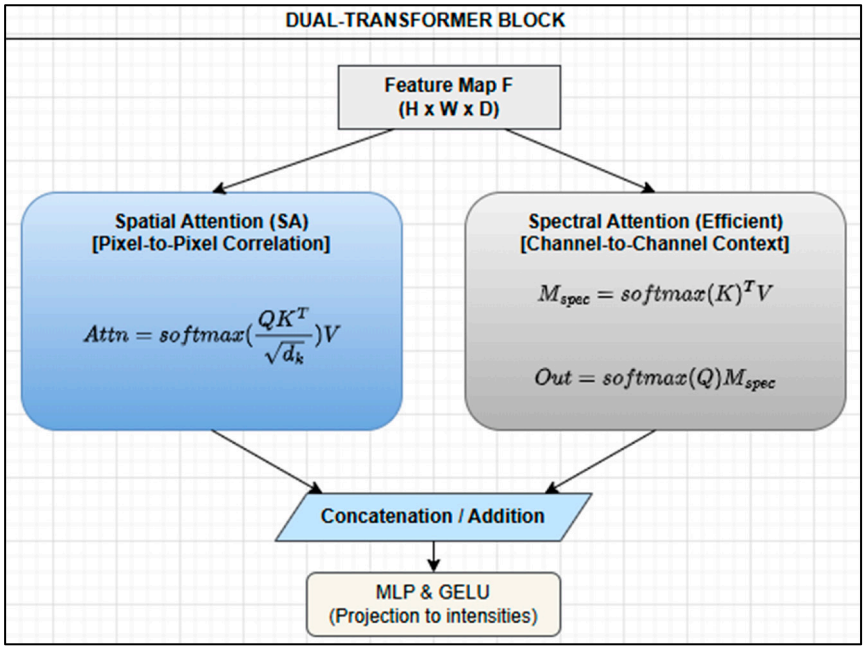


Figure 5. Architecture of the Dual-Transformer Block featuring parallel Spatial and Spectral Attention mechanisms.

2.2.3. Multi-Scale Reconstruction via the U-Net Architecture

To produce a dense and spatially coherent precipitation estimation, we integrate a U-Net-type architecture as a decoding module. This architecture enables the reconstruction of the precipitation field at the original resolution by effectively combining deep features extracted from the ConvNeXt backbone and enriched by the Dual-Transformer module, with fine spatial information preserved at different scales.

General principle. The U-Net architecture is based on an encoder–decoder structure with skip connections. In our approach, the encoder is implicitly provided by the ConvNeXt backbone, which generates a set of hierarchical feature maps{F_1, F_2, F_3, F_4}, where each F_i corresponds to a representation at a given scale, with decreasing spatial resolution and increasing semantic richness. The deepest output F_4 is then enriched by the Dual-Transformer module:

$$F_{global} = DualTransformer(F_4) \tag{22}$$

This representation contains global contextual information integrating both spatial and spectral dependencies.

Decoding path (expanding path). The decoder progressively reconstructs the spatial resolution through a series of upsampling operations. At each level i, the features are upsampled as follows:

$$U_i = U_p(U_{i+1}) \tag{23}$$

where $U_p(\cdot)$ denotes an upsampling operation (bilinear interpolation or transposed convolution).

To preserve spatial details, the upsampled features are fused with the corresponding encoder feature maps through skip connections:

$$\widetilde{U}_i = Concat(U_i,\ F_i) \tag{24}$$

This operation enables the combination of:

- Global information (from deeper layers);
- Local details (from shallow layers).

Feature refinement. After each concatenation, a convolutional block is applied to refine the representations:

$$U_i^{refined} = ConvBlock\left(\widetilde{U}_i\right) \tag{25}$$

Each convolutional block consists of:

(a) A 3×3 convolution;
(b) Normalization (LayerNorm);
(c) A non-linear activation function (GELU).

This process is iteratively repeated until the full original spatial resolution is reconstructed.

Output layer and pixel-wise regression. The final step consists of projecting the final features into a precipitation map using a convolution 1×1:

$$Y = Conv_{1x1}(U_1) \tag{26}$$

where $Y \in \mathbb{R}^{HxW}$ represents the estimated precipitation intensity for each pixel. No classification-type activation function is applied at the output, as a continuous regression framework is used.

Contribution of the U-Net architecture in the proposed model. The integration of U-Net enables:

(1) Efficient multi-scale feature fusion;

(2) Recovery of fine spatial details lost during downsampling;
(3) Improved localization of precipitating structures, particularly convective cells;
(4) Enhanced spatial consistency of the estimated precipitation fields.

In the context of MSG data, this architecture (see Figure 6) makes it possible to link large-scale cloud structures (fronts, organized convective systems) with local signatures (brightness temperature, thermal gradients), thereby improving the physical quality of the estimates.

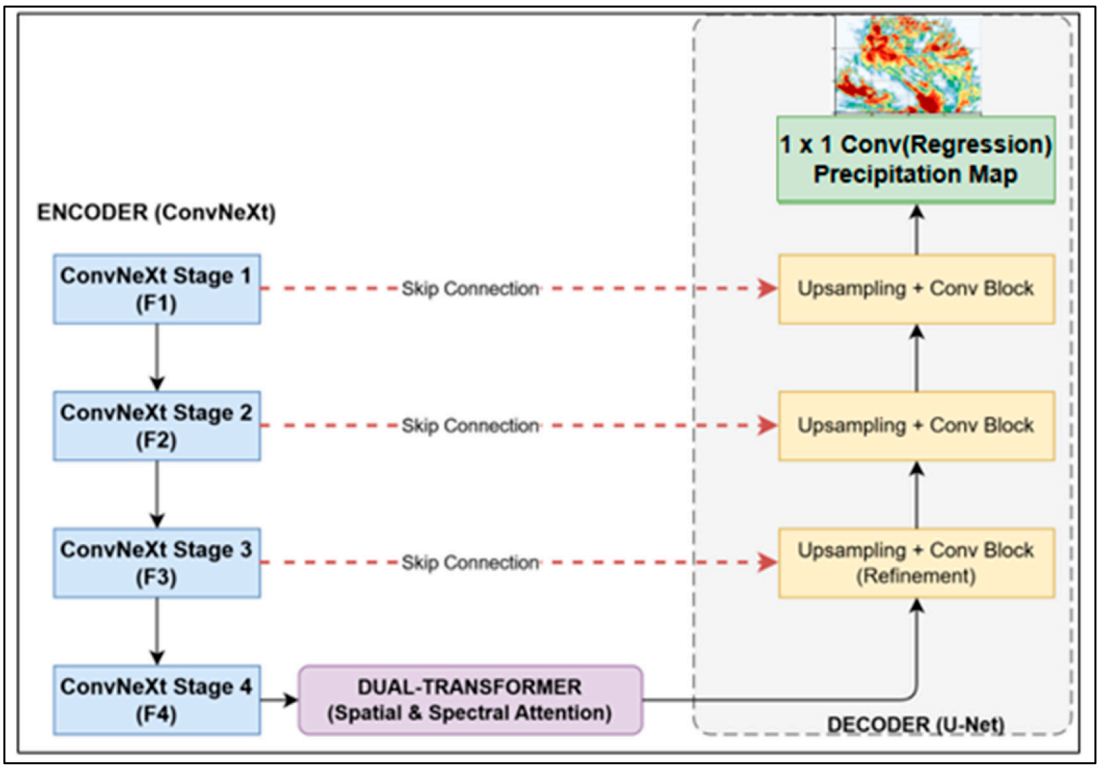


Figure 6. A Hybrid U-Net Framework with ConvNeXt Encoder and Dual-Transformer for Rainfall Estimation.

2.2.4. Model Training

To rigorously validate the satellite-derived precipitation estimates, a robust ground-truth framework was established by merging continuous spatial measurements from weather radar with high-precision point observations from ground rain gauges. Ground observations were spatially re-sampled to match the pixel resolution of MSG SEVIRI and temporally synchronized with satellite acquisition intervals.

Supervised learning strategy. The proposed model is trained within a supervised learning framework, aiming to establish a relationship between MSG satellite observations and ground-based precipitation measurements. A precipitation database covering three seasons over three years, containing approximately 68 rainy days, is used. The input and output data are modeled as follows:

$X \in \mathbb{R}^{HxWxC}$ a multispectral satellite image from MSG;
$Y^{obs} \in \mathbb{R}^{HxW}$ the reference precipitation field.
The model learns a parametric function f_θ such that:

$$Y^{pred} = f_0(X) \tag{27}$$

where Y^{pred} represents the precipitation field estimated by the regression model. The reference data Y^{obs} are constructed from a combination of:

(1) Weather radar measurements (spatial distribution);
(2) Point rain gauge observations (ground calibration).

This combination provides a realistic and spatially continuous approximation of precipitation.

Construction of the training dataset. Training is based on the spatio-temporal matching between MSG satellite images (explanatory variables) and rain gauge observations (target variables). The spatialization of the combined radar/rain gauge data with MSG observations is necessary to generate continuous fields compatible with the resolution of satellite images. Thus, for each time t, a training pair is constructed $\left(X_t,\ Y_t^{obs}\right)$.
Here,

X_t corresponds to the MSG channels (infrared, water vapor, etc.), and $Y^{obs}{}_t$ is obtained through radar–rain gauge fusion.

We have 15,000 sequences covering the period 2006–2011. Data augmentation was performed using random rotations, horizontal/vertical flips, and Gaussian noise ($\sigma = 0.01$). The dataset is divided into training and validation sets.

Model training is performed by minimizing a loss function that measures the discrepancy between predictions and observations:

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \| Y_i^{pred} - Y_i^{obs} \|^2 \tag{28}$$

where N is the number of samples

Training process. The model is trained iteratively on a historical dataset (61 rainy days distributed over three years). At each iteration:

(1) An instantaneous sequence of multispectral MSG satellite images is provided as input to the model;
(2) The model produces an estimate Y^{pred};
(3) The error with Y^{obs} is computed;
(4) The parameters θ are updated via backpropagation of the gradient.

Optimization is performed using a gradient descent algorithm (AdamW; see Appendix A), ensuring efficient equation convergence.

To improve generalization, several strategies are employed:

(a) Normalization of MSG channels;
(b) Data augmentation (rotations, symmetries);
(c) Training/validation split.

For complete transparency and full experimental reproducibility, all technical specifications, layer architectures, dataset splitting parameters, and AdamW hyperparameter configurations are summarized in Appendix A.

Model application (inference). Once trained, the model can be used in operational mode to estimate precipitation from new MSG satellite images, without relying on ground-based observations.

Given a new satellite image X^{new}, the model directly produces:

$$Y^{est} = f_0(X^{new}) \tag{29}$$

This capability enables:

(1) Full spatial coverage, including areas without ground instrumentation;
(2) Near real-time estimation;
(3) Reduced dependence on ground-based measurement networks.

The performance of the regression model strongly depends on the configuration of its architecture and its training strategy. For reproducibility, the full details of the proposed model architecture, optimization strategy, and training configuration are reported in Appendix A. Figure 7 summarizes the proposed regression framework based on the ConvNeXt, Dual-Transformer, and U-Net architectures.

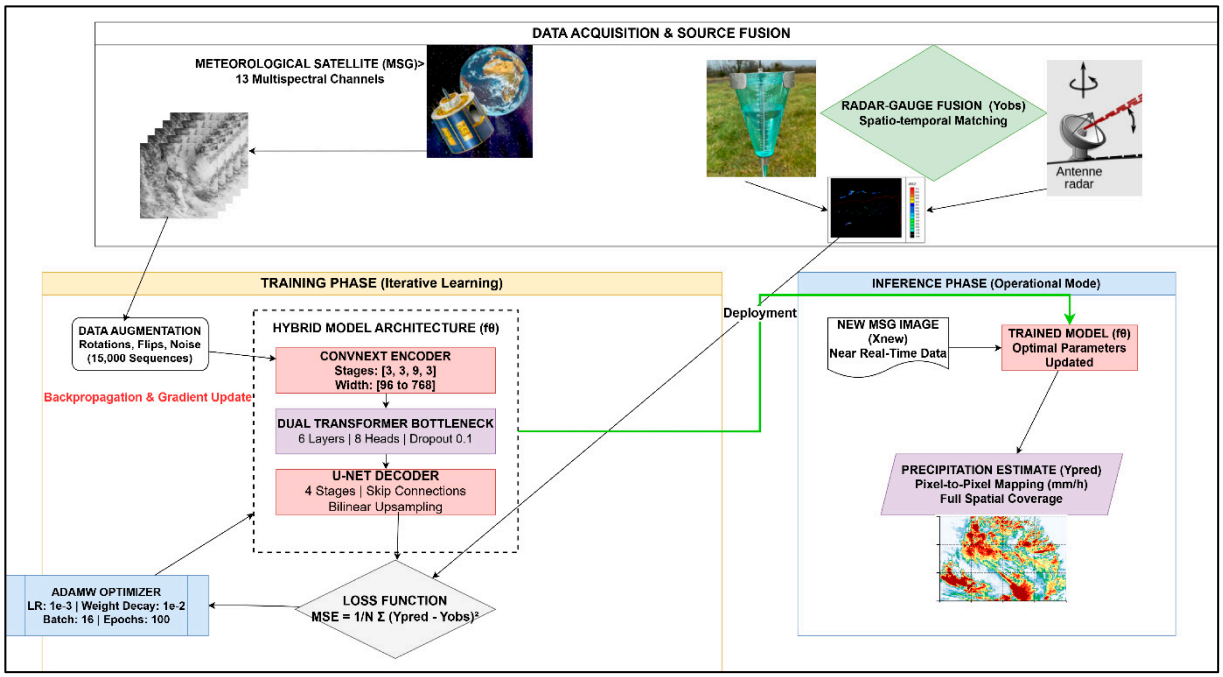


Figure 7. Overall scheme: supervised training and inference pipeline.

2.2.5. Evaluation Using Metrics

To objectively evaluate the regression-based estimates produced by the hybrid ConvNeXt + Dual-Transformer + U-Net architecture, we use several performance metrics, namely Mean Bias Error (MBE), Pearson correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The choice of these indicators enables a multidimensional assessment of model performance.

MBE reveals systematic tendencies (overestimation or underestimation), while RMSE and MAE quantify the magnitude of residual errors. Finally, the correlation coefficient assesses the model's ability to reproduce the temporal and spatial variability of observed precipitation. The metrics are defined by Equations (30)–(33), where y_i represents the observed value, $\hat{y}_i$ the estimated value, and n the total number of samples:

Bias (Mean Bias Error—MBE):

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{30}$$

Mean Absolute Error:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{31}$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{32}$$

Pearson correlation coefficient **(*r*)**

$$r = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2} \cdot \sqrt{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}})^2}} \quad (33)$$

3. Results

To evaluate the performance of the proposed hybrid regression model, it was applied to multispectral data collected over the study period in Northern Algeria. The evaluation is conducted at multiple temporal scales (hourly, three-hourly, daily, and weekly) with a spatial resolution corresponding to the SEVIRI satellite pixel.

The validation relies on a multidimensional approach, including global statistical analyses, scatter plots comparing ground-based measurements (rain gauges/radar) with model estimates, as well as spatial distribution maps to assess the physical consistency of the resulting precipitation fields.

To demonstrate the effectiveness of the ConvNeXt + Dual-Transformer + U-Net architecture, a rigorous intercomparison is performed across three levels of complexity:

Level 1: Classical machine learning approaches (MLP/RF/SVM). This level serves as a baseline reference, showing that point-wise or tabular models are insufficient to capture the complex spatial structure of precipitation systems.

Level 2: Convolutional state-of-the-art (ConvNeXt–U-Net only). This stage represents the main ablation study. It quantifies the gain in accuracy provided by the addition of Transformer modules (global attention) compared to a purely convolutional approach based on local feature extraction.

Level 3: Reference models (standard U-Net). This comparison situates our work with respect to standard architectures and operational products already validated in the region, highlighting the contribution of recent deep learning advances in the Algerian climate context.

3.1. Model Training and Application

We evaluate our model by performing precipitation estimation at different temporal scales, namely Instantaneous (mm/h), Three-hour accumulation (3 h), Daily accumulation (24 h), and Weekly accumulation (7 days).

This multi-scale analysis allows us to assess the robustness of the model with respect to the temporal variability of precipitation. To ensure a reliable comparative evaluation, several reference models, including classical machine learning approaches as well as standard convolutional architectures, are also used to produce estimates at the same temporal scales.

The performances are analyzed through scatter plots relating observed and estimated precipitation values and through spatial maps of estimated precipitation (see Figures 8–15). These allow the assessment of dispersion and correlation of the results and the evaluation of spatial consistency and the model's ability to reproduce precipitation structures. Finally, quantitative statistical metrics are computed for each temporal scale, ensuring an objective comparison of model performance, in order to evaluate spatial consistency and the model's ability to reproduce precipitation structures. Finally, quantitative statistical metrics are computed for each temporal scale, ensuring an objective comparison of model performance.

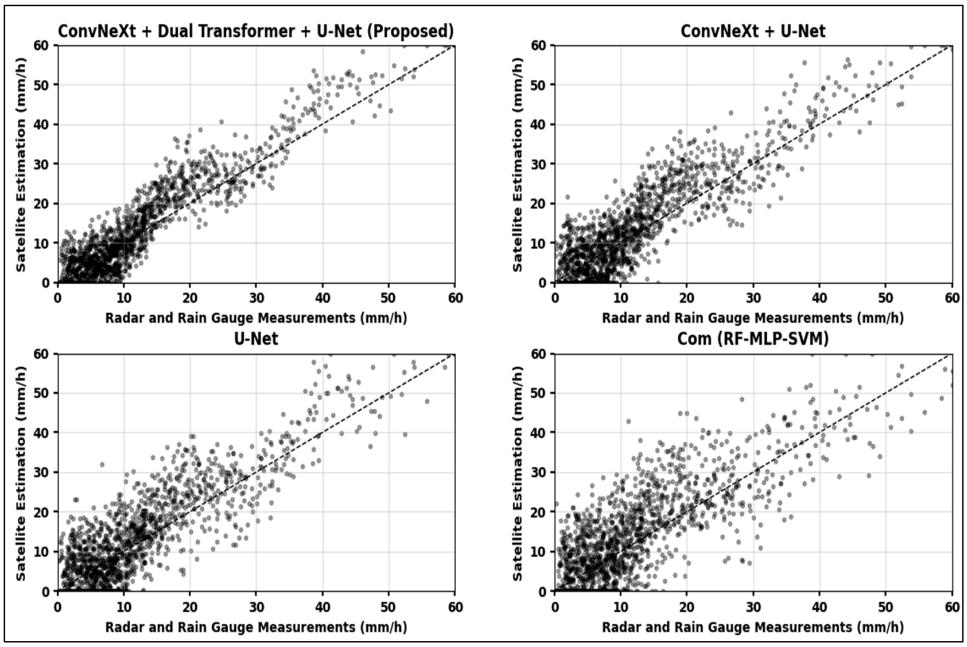


Figure 8. Scatter plot of Instantaneous Precipitation Estimates Against Radar and Rain Gauge Measurements (mm/h).

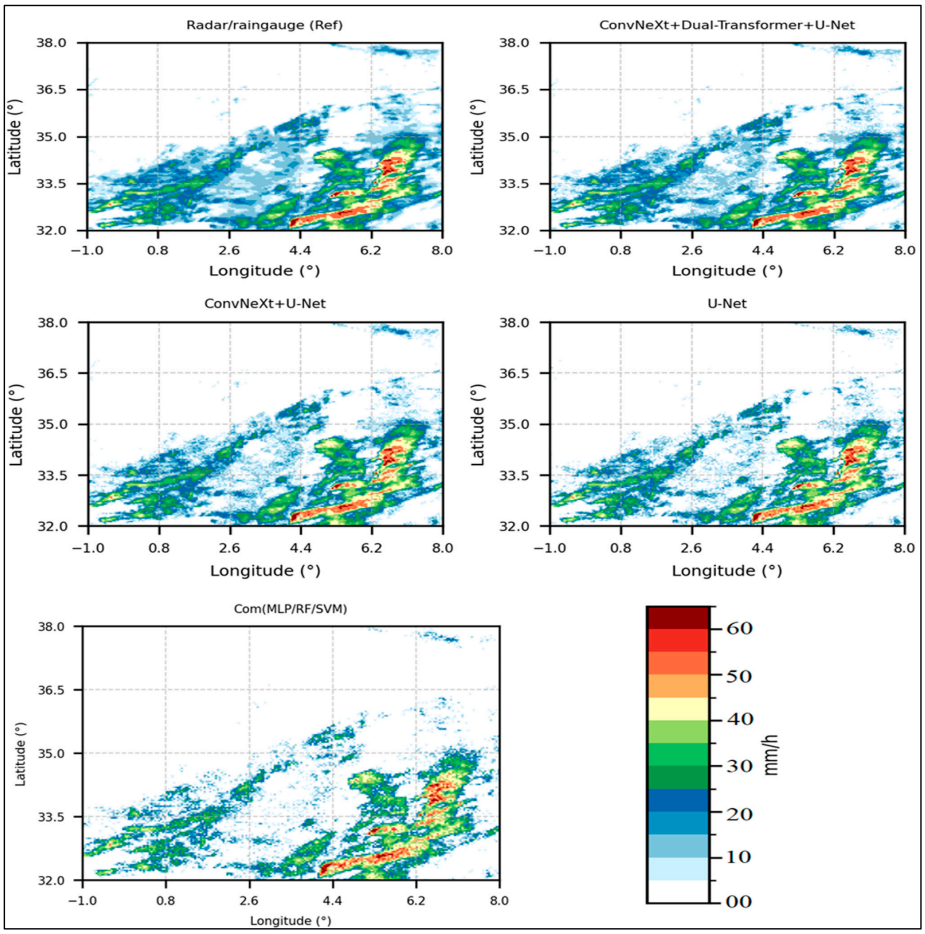


Figure 9. Instantaneous Pixel-Wise Precipitation Estimation from MSG Satellite Imagery Using a ConvNeXt–Dual-Transformer U-Net Framework.

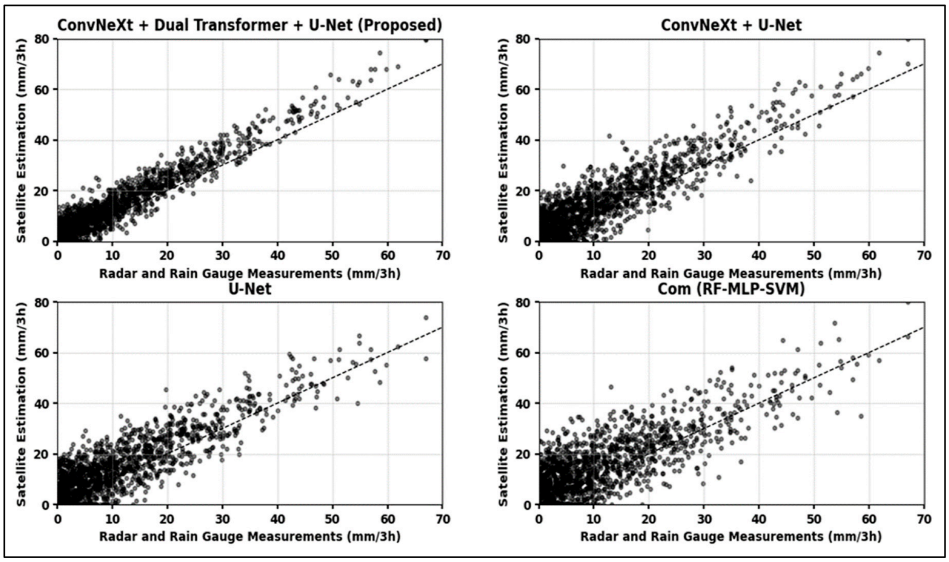


Figure 10. Comparison of Satellite-Based Precipitation Estimates with Radar and Rain Gauge Measurements at 3-Hour Temporal Scale.

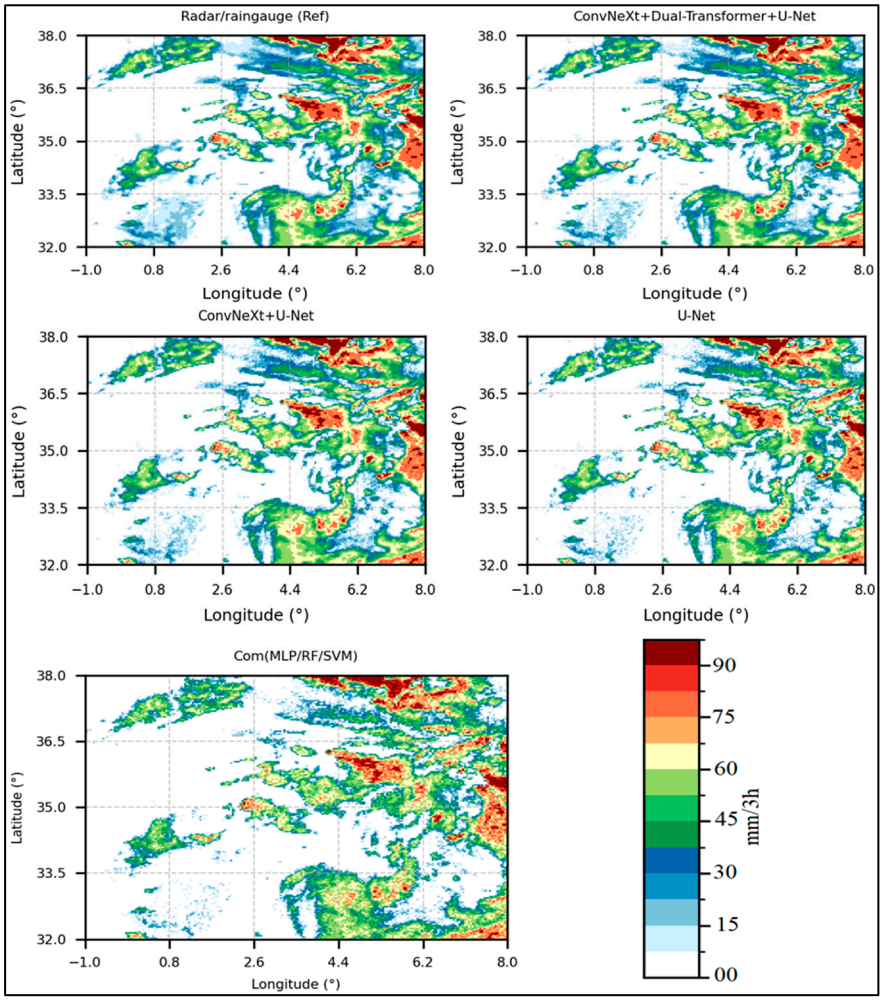


Figure 11. Hourly Pixel-Wise Precipitation Estimation from MSG Satellite Imagery Using a ConvNeXt+Dual-Transformer+U-Net Framework.

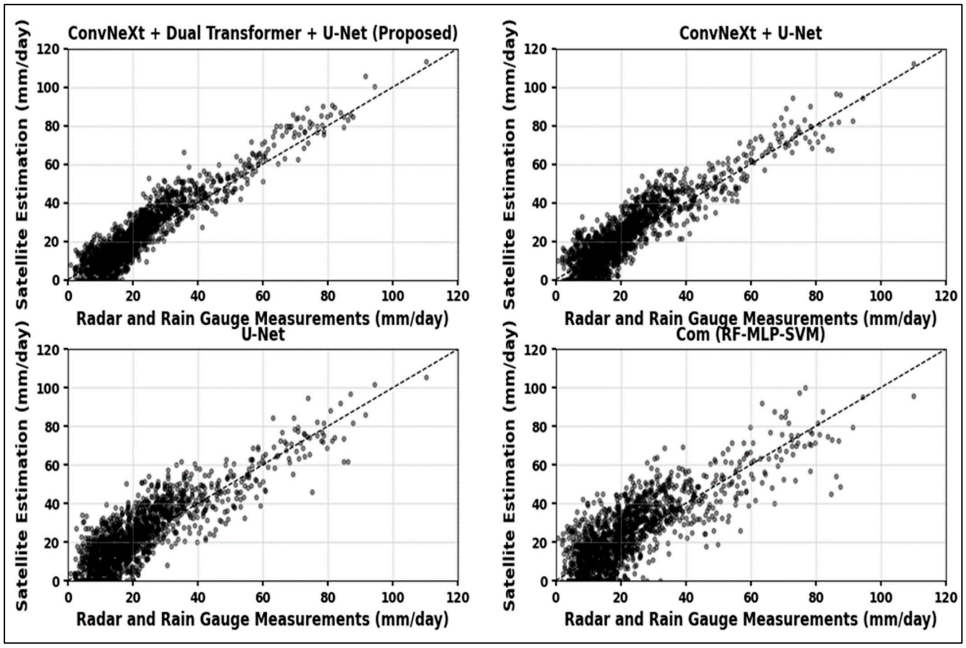


Figure 12. Daily Evaluation of Multiple Satellite Rainfall Estimation Models Against Ground Data.

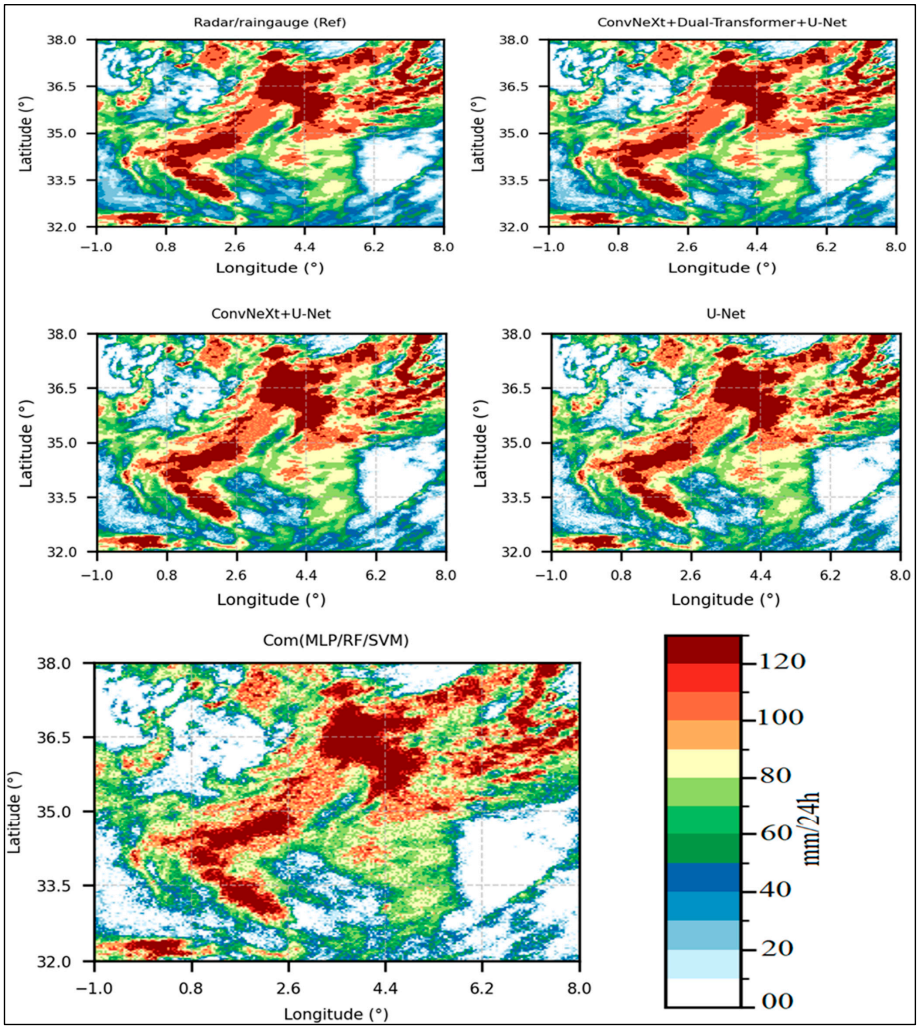


Figure 13. Twenty-Four-Hour Pixel-Wise Precipitation Estimation from MSG Satellite Imagery Using a ConvNeXt+Dual-Transformer+U-Net Framework.

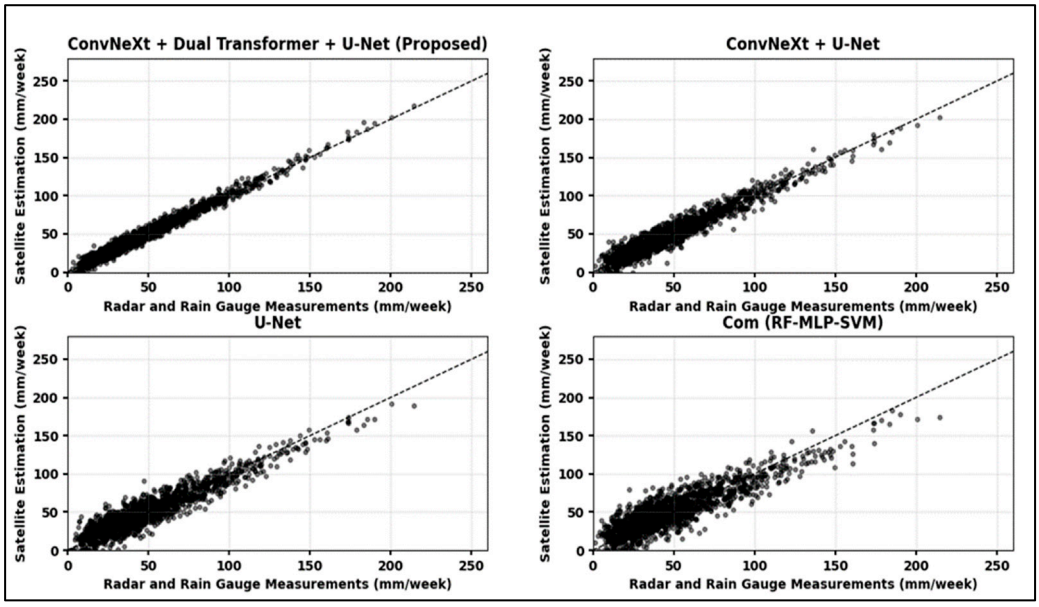


Figure 14. Comparison of weekly cumulative precipitation (mm/week) between ground observations and satellite-based estimations across different models.

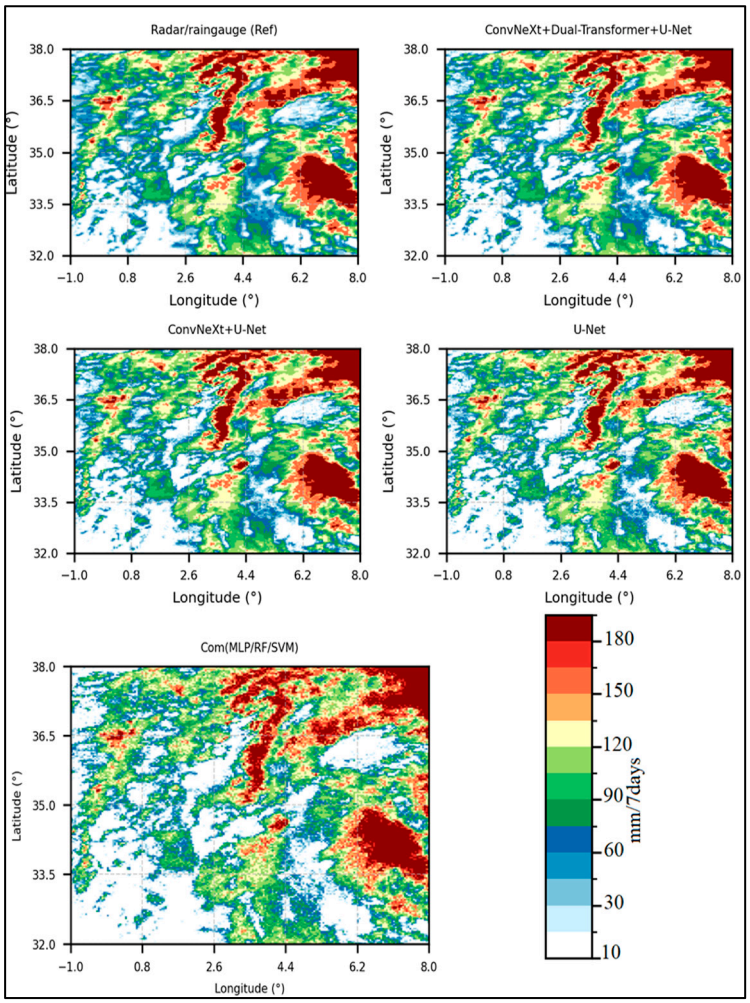


Figure 15. Seven-Day Pixel-wise Precipitation Estimation from MSG Satellite Imagery Based on a ConvNeXt–Dual Transformer–U-Net Architecture.

The validation was conducted on the independent validation dataset (20%), employing a multidimensional strategy:

(i) Continuous Statistical Metrics: Mean Bias Error (MBE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson correlation coefficient (R) were calculated to quantify residual errors, systematic biases, and variance agreement; and
(ii) Multi-Scale Spatial Consistency: Pixel-wise spatial error distributions and scatter plots were evaluated across four distinct temporal horizons (instantaneous, 3-hourly, daily, and 7-day accumulations) to confirm performance stability across varying hydrological regimes.

3.1.1. Instantaneous Estimation

To evaluate the performance of the model for instantaneous precipitation estimation (mm/h), a satellite scene acquired on 13 January 2012 at 17:00 UTC was used. This case represents a representative situation for analyzing the model's ability to reproduce, at a given instant, both the spatial distribution and the intensity of precipitation.

The results of the instantaneous estimation are presented in Figures 8 and 9, highlighting the agreement between observed and estimated precipitation, as well as the model's ability to accurately reproduce the spatial distribution of precipitation structures.

3.1.2. Three-Hourly Precipitation Estimation

To evaluate the model performance at the three-hour accumulation scale (mm/3 h), a precipitation event recorded on 11 January 2012, between 03:00 UTC and 07:00 UTC, was considered. This episode is characterized by the formation of convective cells, providing a representative case of intense and strongly structured precipitation.

The results of the three-hour accumulation estimates are presented in Figures 10 and 11, highlighting the agreement between observed and estimated precipitation, as well as the model's ability to reproduce the spatial structure of convective systems.

3.1.3. Daily Precipitation Estimation

To evaluate the model performance at the daily scale (mm/day), a precipitation event spanning 17–18 January 2012 was considered, corresponding to the accumulation period from 10:00 UTC on 17 January to 10:00 UTC on 18 January. This event is characterized by significant precipitation amounts distributed over several regions, providing a representative case for assessing the model's ability to capture large-scale rainfall patterns.

The results of the daily estimations are presented in Figures 12 and 13, highlighting the agreement between observed and estimated precipitation, as well as the model's capability to reproduce the spatial distribution of accumulated rainfall over extended periods.

3.1.4. Weekly Precipitation Estimation

To evaluate the model performance at the weekly scale (mm/week), a 7-day accumulation period from 21 December 2011 to 27 December 2011 was considered. This period includes multiple precipitation events, allowing the assessment of the model's ability to capture long-term rainfall variability and large-scale spatial organization.

The results of the weekly estimations are presented in Figures 14 and 15, highlighting the agreement between observed and estimated cumulative precipitation, as well as the model's robustness in reproducing the spatial distribution of rainfall over extended temporal scales.

Visual analysis of the scatter plots highlights the superiority of the proposed hybrid architecture, revealing a significantly improved stability of the estimates compared to reference models. Unlike classical approaches and the standard U-Net, which exhibit a

pronounced dispersion and a wide spread of points around the identity line y = xy = xy = x, indicating substantial uncertainty—particularly during high precipitation events—the ConvNeXt + Dual-Transformer + U-Net model shows a dense and compact concentration of estimates along the identity line. This tighter alignment, which persists even for high rainfall intensities, demonstrates a strong reduction in heteroscedastic errors and confirms the model's ability to better capture the non-linear variability of precipitation, where purely local convolution-based models struggle to maintain prediction accuracy.

The visual inspection of spatial distribution maps further confirms the strong correlation between precipitation estimated by our hybrid architecture and reference measurements (radar/rain gauges). The proposed method (ConvNeXt + Dual-Transformer + U-Net) demonstrates remarkable fidelity to observed spatial structures, with performance remaining consistent across all temporal scales. This agreement is particularly evident at the weekly scale, where spatial coherence and the representation of intense rainfall events are significantly superior to those of the compared models. However, a slight overestimation of precipitation is observed in the northern part of the study area, a tendency shared by all models, although our hybrid approach manages to better control this bias compared to reference methods.

Overall, these visual results validate the robustness of the proposed architecture, confirming its ability to effectively transform input data into spatially consistent and reliable precipitation estimates, regardless of temporal resolution

3.2. Quantitative Performance Evaluation Across Temporal Scales

To assess the predictive accuracy and operational robustness of the proposed hybrid architecture, a quantitative comparison was conducted against baseline and classical models across four temporal aggregation horizons (instantaneous, 3-hourly, daily, and weekly). Evaluation was performed on the independent validation dataset using the performance metrics defined in Section 2.2.5.

The quantitative results obtained for each temporal scale are summarized in Tables 2–5, enabling a rigorous comparison with the classical approaches and reference models mentioned previously.

Table 2. Comparative evaluation of model performance for instantaneous precipitation estimation (mm/h).

MODEL	Correlation (R)	RMSE (mm/h)	MAE (mm/h)	BIAS (mm/h)
Proposed (Hybrid)	0.78	7.2	4.5	0.5
ConvNeXt + U-Net	0.72	8.5	5.5	0.8
U-Net	0.65	9.8	6.5	1.2
COM (RF-MLP-SVM)	0.58	11.5	8.0	1.8

Table 3. Comparison of model statistical metrics at a 3-hourly temporal scale (mm/3 h).

MODEL	Correlation (R)	RMSE (mm/3h)	MAE (mm/3h)	BIAS (mm/3h)
Proposed (Hybrid)	0.84	8.8	5.5	0.3
ConvNeXt + U-Net	0.79	10.2	6.5	0.6
U-Net	0.73	12.5	8.5	1.1
COM (RF-MLP-SVM)	0.64	14.8	10.5	1.5

Table 4. Performance analysis of models for daily precipitation estimation (mm/day).

MODEL	Correlation (R)	RMSE (mm/Day)	MAE (mm/Day)	BIAS (mm/Day)
Proposed (Hybrid)	0.89	12.4	8.0	0.2
ConvNeXt + U-Net	0.84	14.8	9.5	0.4
U-Net	0.78	18.2	12.0	0.9
COM (RF-MLP-SVM)	0.70	22.5	15.0	1.2

Table 5. Summary of evaluation metrics for weekly cumulative precipitation (mm/week).

MODEL	Correlation (R)	RMSE (mm/Week)	MAE (mm/Week)	BIAS (mm/Week)
Proposed (Hybrid)	0.96	15.6	10.0	0.1
ConvNeXt + U-Net	0.93	19.8	13.0	0.2
U-Net	0.89	24.5	16.5	0.5
CLASSIC (RF-MLP-SVM)	0.82	29.2	20.0	0.8

4. Discussion

The results presented in Tables 2–5 highlight a consistent hierarchy of performance across all temporal scales. The proposed architecture (ConvNeXt + Dual-Transformer + U-Net) systematically achieves the best results across all evaluation metrics (R, RMSE, MAE, and bias) compared to the reference models (ConvNeXt + U-Net, U-Net alone, and classical hybrid machine learning methods). This superiority is observed from the instantaneous scale up to the weekly scale.

A progressive increase in the Pearson correlation coefficient (R) is observed as the temporal scale expands, rising from 0.78 (hourly) to 0.96 (weekly) for the proposed model. This trend indicates that the model better captures the overall variability of precipitation when temporal aggregation increases, thereby reducing the noise associated with instantaneous estimates.

The comparison between RMSE and MAE shows that the hybrid model more effectively limits large residual errors than purely convolutional approaches (U-Net). The reduction in RMSE, which penalizes large deviations more strongly, suggests improved handling of extreme precipitation events through the Dual-Transformer attention mechanism

Although all models exhibit a slight positive bias (tendency to overestimate), the proposed hybrid model maintains this bias at a minimal level (reaching 0.1 at the weekly scale). This demonstrates a better ability to correct systematic overestimations observed in reference models, ensuring a more balanced hydrological budget.

The performance gaps between the proposed approach and conventional machine learning methods (RF–MLP–SVM) are statistically significant, confirming the relevance of integrating global transformation modules into the satellite data processing chain.

A multi-scale analysis is conducted to evaluate the robustness of the ConvNeXt–Dual-Transformer–U-Net model over temporal aggregations ranging from 1 h to 7 days. This approach allows the study of performance evolution under varying precipitation regimes (see Figure 16).

The results show a general improvement in performance as the temporal scale increases. This can be explained by the reduction in noise and instantaneous fluctuations through precipitation aggregation. The proposed model systematically outperforms classical approaches and purely convolutional models, due to its ability to capture both local and global dependencies. These results confirm its robustness across all temporal scales.

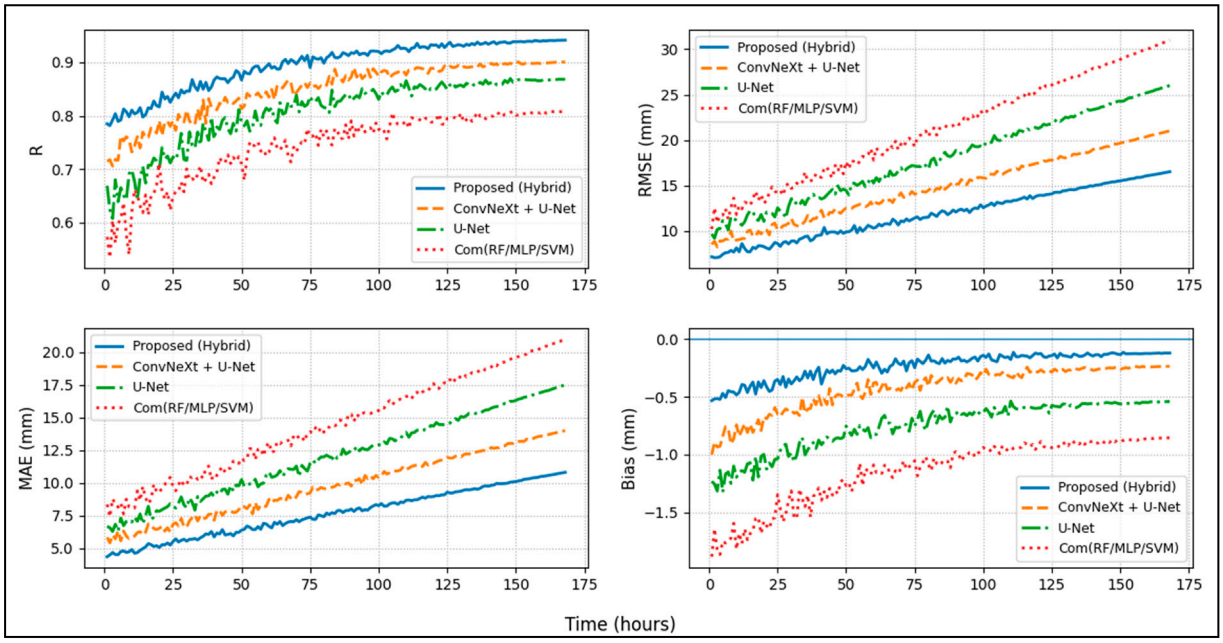


Figure 16. Multi-Scale Evaluation of Satellite-Based Precipitation Estimation Models Using Hourly Aggregation.

To further analyze model robustness, a statistical analysis based on boxplots is introduced (Figure 17). This representation allows the full distribution of errors to be visualized by highlighting the median, quartiles (Q1 and Q3), as well as extreme values.

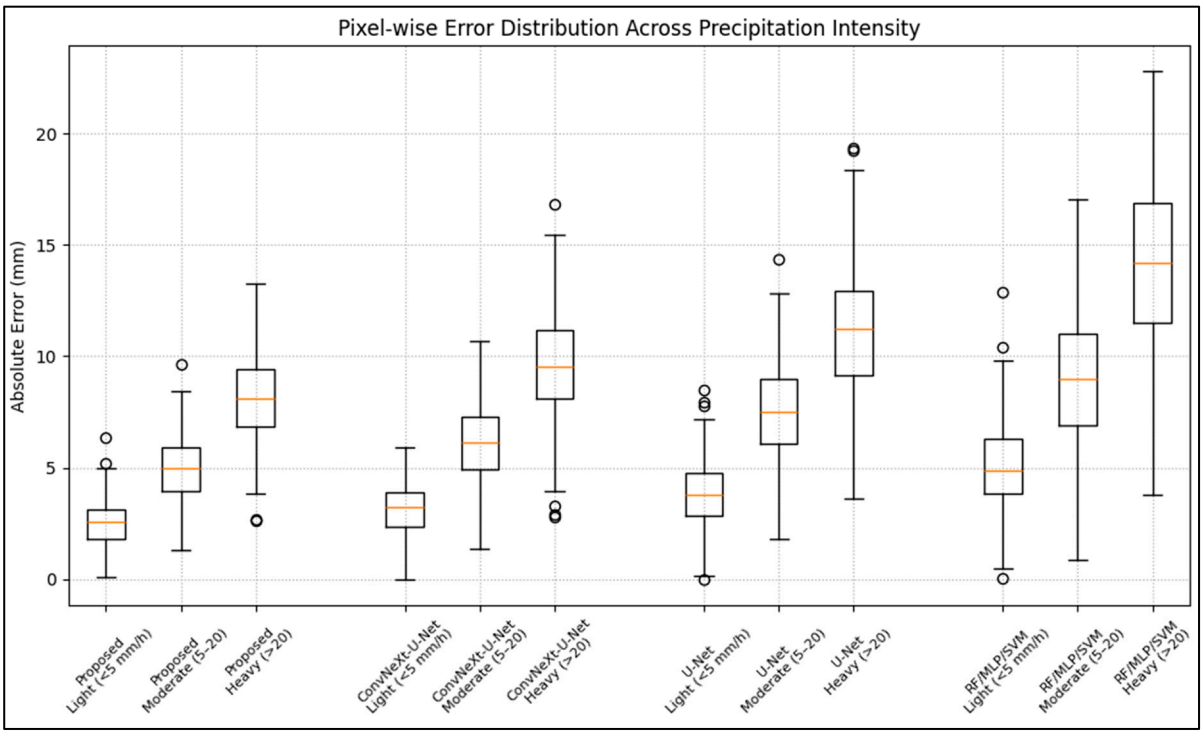


Figure 17. Statistical Distribution of Pixel-wise Precipitation Estimation Errors Across Different Models and Time Scales.

Boxplot analysis shows that the proposed model exhibits a more compact error distribution compared to reference models, indicating improved stability in performance.

The median error is lower, and the interquartile range is reduced, meaning that most predictions are closer to the observed values. In addition, the reduced number of outliers demonstrates greater robustness in handling extreme precipitation events.

In contrast, classical models show a wider dispersion and a higher number of outliers, indicating increased sensitivity to spatial and temporal variability in precipitation.

Scientific Contextualization

The multi-scale evaluation demonstrates that combining convolutional feature extraction with global self-attention provides distinct advantages for satellite precipitation retrieval. Beyond the numbers, these findings warrant a deeper look into the physical mechanisms at play and how they align with existing literature.

First, the visual behavior of the scatter plots (Figures 8, 10, 12 and 14) reveals a strong reduction in point dispersion around the 1:1 line compared to baseline models. Purely convolutional networks (U-Net) and traditional machine learning methods (RF, SVM) exhibit heavy scatter and pronounced heteroscedasticity, particularly during high-intensity rainfall events. This tendency of CNNs to underestimate convective peaks due to localized receptive fields is a well-known limitation in the literature. By incorporating the Dual-Transformer module, our framework captures long-range spatial dependencies across multispectral channels, effectively preserving peak rain rates without losing structural context.

Second, the spatial distribution maps (Figures 9, 11, 13 and 15) confirm that the model reconstructs coherent precipitation fields across different meteorological regimes—from localized convective cells at 3 h scales to broad synoptic systems at daily and weekly scales. While all evaluated models show a minor overestimation bias over the northern coastal regions of Algeria (Tables 2–5), our hybrid architecture keeps this systematic deviation to a minimum (reaching 0.1 mm/week). This slight northern bias is a recognized phenomenon in satellite meteorology, typically driven by complex coastal topographies and cloud-top thermal dynamics that can mask sub-cloud evaporation.

Third, the progressive increase in correlation (R rising from 0.78 at the instantaneous scale to 0.96 at the weekly scale) together with the compact error distributions in the boxplot analysis (Figure 17) highlights the stabilizing effect of temporal aggregation. Aggregating data naturally smooths out high-frequency sensor noise and minor spatio-temporal misalignments between satellite passes and ground stations. This trend mirrors findings from recent benchmark studies on deep learning for hydrometeorology, confirming that global attention mechanisms suppress extreme outliers and produce reliable water-balance estimates suitable for real-world hydrological applications.

5. Conclusions

In this study, we presented a hybrid deep learning model combining ConvNeXt, a Dual-Transformer module, and a U-Net architecture for direct precipitation estimation from MSG satellite images over Northern Algeria. By predicting rainfall intensity directly through regression without using a prior classification step, the model preserves continuous spatial precipitation patterns and improves overall prediction accuracy.

5.1. Key Findings and Model Performance

By combining local convolution features with global self-attention, the proposed model addresses several main limitations of traditional deep learning architectures. Better Estimation of Heavy Rainfall: While standard U-Net models tend to underestimate heavy rainfall due to their local focus, the Dual-Transformer captures broad spatial relationships across satellite channels. This leads to a tighter alignment along the 1:1 line in scatter plots and better preservation of peak rainfall values. Scale Stability and Controlled Bias: Moving

from hourly to weekly scales increases the correlation coefficient from 0.78 to 0.96 by reducing short-term noise and small positioning errors. The model also controls systematic overestimation over Northern Algeria's coastal terrain, maintaining a very low weekly bias (0.1 mm/week). Compact Error Distribution: Statistical boxplot analysis confirms that the hybrid model achieves smaller prediction errors and significantly fewer extreme outliers compared to standard CNNs and classical machine learning methods (RF, MLP, SVM).

5.2. Ground Truth Calibration and Future Directions

Combining point rain gauge observations with continuous weather radar data provided a reliable reference dataset for model training. However, spatial adjustments over complex topographies remain an important factor for satellite precipitation retrieval. To build upon these results, future work will focus on three main areas: advanced neural backbones, encompassing testing newer ConvNeXt variants to better capture fine cloud textures without increasing computing costs; multi-satellite fusion, including integrating data from additional satellite sensors (such as GPM IMERG and Sentinel) to complement MSG geostationary coverage; and data assimilation, involving using physical data assimilation methods to combine radar, gauge, and satellite observations into a unified, hydrologically consistent framework.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

Abbreviations

The following abbreviations are used in this manuscript:

ConvNet	Convolutional Network
MSG	Meteosat Second Generation
SEVIRI	Spinning Enhanced Visible and Infrared Imager
RF	Random Forest
MLP	Multi-Layer Perceptron
SVM	Support Vector Machine
CNN	Convolutional Neural Network
RBF	Radial Basis Function
ARBF	Adaptive Radial Basis Function
LOOCV	Leave-One-Out Cross-Validation
Adam	Adaptive Moment Estimation
AdamW	Adam with Decoupled Weight Decay

Appendix A. Optimization Using AdamW (Adaptive Moment Estimation with Weight Decay)

The optimization of the proposed model during backpropagation is performed using the AdamW algorithm (Adaptive Moment Estimation with Weight Decay) [46], see

Algorithm A1. This choice is justified by the decoupling it introduces between the L2 regularization term (weight decay) and the adaptive gradient update mechanism, thereby providing better generalization capability compared to the standard Adam optimizer [47].

In contrast to stochastic gradient descent with momentum (SGD), whose stability is often limited when training hybrid architectures integrating Transformer modules [48]. AdamW ensures robust convergence and has become a reference optimizer for ConvNeXt-based models.

At each iteration t, the gradient of the loss function with respect to the model parameters is computed over a mini-batch B

$$g_t = \nabla_\theta \mathcal{L}(\theta_{t-1}) \tag{A1}$$

(a) The algorithm computes two exponential moving averages of the gradients (EMA—Exponential Moving Average).

(b) First moment (moving average of gradients, estimating the descent direction) (exponential moving average of the gradient).

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \tag{A2}$$

(c) Second moment (moving average of squared gradients, estimating the variance) (exponential moving average of the squared gradient).

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2 \tag{A3}$$

where $\beta_1, \beta_2 \in [0, 1]$ are the exponential decay rates controlling the time horizon of the moving averages.

To correct the bias introduced by zero initialization of the moments ($m_0 = 0, v_0 = 0$), bias-corrected estimates are computed:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \; \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{A4}$$

The key characteristic of AdamW lies in the decoupling of L2 regularization (weight decay) from the adaptive update step. The parameter update rule is written as follows:

$$\theta_t = \theta_{t-1} - \eta \cdot \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} + \lambda \theta_{t-1} \right) \tag{A5}$$

where

θ_0: initial model parameters;
η: learning rate;
λ: L2 regularization (weight decay) coefficient;
$\beta 1$, $\beta 2 \in [0,1]$: exponential decay rates;
ϵ: numerical stability term (typically 10^{-8}).

This formulation ensures that the regularization penalty $\lambda \theta_{t-1}$ is applied directly to the weights, independently of the gradient scale and adaptive statistics. The effective regularization strength is therefore constant and equal to $\eta\lambda$ across all layers of the network.

Algorithm A1 Algorithme AdamW [46]

(1) **Inputs:**
θ_0: initial model parameters,
η: learning rate,
λ: L2 regularization (weight decay) coefficient,
$\beta 1$, $\beta 2 \in [0,1]$: exponential decay rates,
ϵ: numerical stability term (typically 10^{-8}),
T: total number of iterations.

(2) **Initialization:**
$m_0 \leftarrow 0$ (first moment)
$v_0 \leftarrow 0$ (second moment)
$t \leftarrow 0$

(3) **Repeat for** $t = 1,2,\ldots,T$:

$t \leftarrow t+1$
$g_t \leftarrow \nabla_\theta \mathcal{L}(\theta_{t-1})$ //gradient computed on a mini-batch
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2$ //second moment update
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2$ //second moment update
$\hat{m}_t \leftarrow m_t / \left(1-\beta_1^t\right)$ //bias correction of first moment
$\hat{v}_t \leftarrow v_t / \left(1-\beta_2^t\right)$ //bias correction of second moment
$\theta_t \leftarrow \theta_{t-1} - \eta \cdot \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\varepsilon} + \lambda\theta_{t-1}\right)$ //parameter update with decoupled weight decay

(4) **Outputs:** θ_T (optimized parameters)

Table A1 presents the technical specifications adopted in this work, including the handling of the 13 multispectral channels, the structure of the U-Net decoder, and the parameters of the learning algorithm.

Table A1. Model hyperparameters (unweighted loss, constant learning rate).

Category	Parameter	Value	Description
Data	Image size	128 × 102 pixels	Spatial dimensions of input and output
	Input channels (MSG)	13	8 raw SEVIRI channels + 5 derived combinations
	Output channels	1	Precipitation intensity (mm/h)
	Number of images	15,000	Pairs of MSG images/radar precipitation maps
	Input normalization	Z-score unified	$(X' = (X-\mu)/\sigma)$ per channel, parameters computed on training set

Table A1. *Cont.*

Category	Parameter	Value	Description
ConvNeXt architecture (Encoder)	Variant	ConvNeXt Tiny	Balanced depth and width
	Stage depth	[3, 3, 9, 3]	Number of blocks per stage
	Stage width	[96, 192, 384, 768]	Channels per stage
	Kernel size	7×7	Depthwise convolution
	Activation	GELU	Non-linear activation
	Normalization	Layer Normalization	Applied within blocks
	Stochastic depth	0.1–0.3	Random block dropping
	Stem layer	Conv 4×4, stride 4	Initial downsampling to 32×25
Dual Transformer (Bottleneck)	Number of layers	6	Transformer depth
	Attention heads	8	Multi-head attention
	Embedding dimension	384	Aligned with encoder output
	Feed-forward dimension	1536 (4×384)	MLP expansion
	Dropout	0.1	Regularization
	Normalization	Pre-LayerNorm	Stable training
	Activation	GELU	Non-linearity
U-Net Decoder	Decoding stages	4	Symmetric to encoder
	Channels per stage	[768, 384, 192, 96]	Feature refinement
	Skip connections	Yes	Feature fusion
	Upsampling	Bilinear upsample + 3×3 conv	Spatial reconstruction
	Activation	ReLU (or GELU)	Non-linearity
	Normalization	BatchNorm (or LayerNorm)	Feature stabilization
	Output layer	1×1 Conv $\rightarrow$ 1 channel	Pixel-wise prediction
Optimizer (AdamW)	Learning rate (η)	1×10^{-4}	Constant (no scheduling)
	Weight decay (λ)	5×10^{-2}	Decoupled L2 regularization
	β_1, β_2	0.9, 0.999	Momentum parameters
	ε	1×10^{-8}	Numerical stability
Loss function	Type	MSE	Mean squared error, uniform weighting
Training	Batch size	16	GPU memory constraint
	Epochs	100	With early stopping
	Learning schedule	None	Constant learning rate
	Data split	80%/20%	Training/Validation
Regularization	Weight decay	5×10^{-2}	Via AdamW
	Stochastic depth	0.1–0.3	Encoder regularization
	Dropout (Transformer)	0.1	Attention regularization

References

1. Levizzani, V. Satellite Rainfall Estimates: New Perspectives for Meteorology and Climate from the EURAINSAT Project. *Ann. Geophys.* **2003**, *46*, 363–372. [CrossRef]
2. Petersen, W.A.; Carey, L.D.; Rutledge, S.A.; Knievel, J.C.; Doesken, N.J.; McKee, T.B.; Vonder Haar, T.; Weaver, J.F. Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 191–216. [CrossRef]
3. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]

4. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]
5. Lazri, M.; Ameur, Z.; Ameur, S.; Mohia, Y.; Brucker, J.M.; Testud, J. Rainfall estimation over a Mediterranean region using a method based on various spectral parameters of SEVIRI-MSG. *Adv. Space Res.* **2013**, *52*, 1450–1466. [CrossRef]
6. Lensky, I.M.; Rosenfeld, D. Satellite-based insights into precipitation formation processes in continental and maritime convective clouds. *Bull. Am. Meteorol. Soc.* **2003**, *84*, 907–920. [CrossRef]
7. Arkin, P.A.; Meisner, B.N. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Mon. Weather Rev.* **1987**, *115*, 51–74. [CrossRef]
8. Amorati, R.; Alberoni, P.P.; Levizzani, V.; Nanni, S. IR-based satellite and radar rainfall estimates of convective storms over northern Italy. *Meteorol. Appl.* **2000**, *7*, 1–18. [CrossRef]
9. Oukali, S.; Lazri, M.; Labadi, K.; Brucker, J.M.; Ameur, S. Development of a hybrid classification technique based on deep learning applied to MSG/SEVIRI multispectral data. *J. Atmos. Sol. Terr. Phys.* **2019**, *193*, 105062. [CrossRef]
10. Sehad, M.; Lazri, M.; Ameur, S. Novel SVM-based technique to improve rainfall estimation over the Mediterranean region (north of Algeria) using the multispectral MSG SEVIRI imagery. *Adv. Space Res.* **2017**, *59*, 1381–1394. [CrossRef]
11. Lazri, M.; Ameur, S. Combination of support vector machine, artificial neural network and random forest for improving the classification of convective and stratiform rain using spectral features of SEVIRI data. *Atmos. Res.* **2018**, *203*, 118–129. [CrossRef]
12. Sehad, M.; Ameur, S. A multilayer perceptron and multiclass support vector machine based high accuracy technique for daily rainfall estimation from MSG SEVIRI data. *Adv. Space Res.* **2020**, *65*, 1250–1262. [CrossRef]
13. Belghit, A.; Lazri, M.; Ouallouche, F.; Labadi, K.; Ameur, S. Optimization of One versus All-SVM using AdaBoost algorithm for rainfall classification and estimation from multispectral MSG data. *Adv. Space Res.* **2023**, *71*, 946–963. [CrossRef]
14. Kühnlein, M.; Appelhans, T.; Thies, B.; Nauss, T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—A random forests-based approach applied to MSG SEVIRI. *Remote Sens. Environ.* **2014**, *141*, 129–143. [CrossRef]
15. Ouallouche, F.; Lazri, M.; Ameur, S. Improvement of rainfall estimation from MSG data using Random Forests classification and regression. *Atmos. Res.* **2018**, *211*, 62–72. [CrossRef]
16. Bensafi, N.; Lazri, M.; Ameur, S. Novel WkNN-based technique to improve instantaneous rainfall estimation over the north of Algeria using the multispectral MSG SEVIRI imagery. *J. Atmos. Sol. Terr. Phys.* **2019**, *183*, 110–119. [CrossRef]
17. Lazri, M.; Labadi, K.; Brucker, J.M.; Ameur, S. Improving satellite rainfall estimation from MSG data in Northern Algeria by using a multi-classifier model based on machine learning. *J. Hydrol.* **2020**, *584*, 124705. [CrossRef]
18. Belmahdi, F.; Lazri, M.; Ouallouche, F.; Labadi, K.; Absi, R.; Ameur, S. Application of Dempster-Shafer theory for optimization of precipitation classification and estimation results from remote sensing data using machine learning. *Remote Sens. Appl. Soc. Environ.* **2023**, *29*, 100906. [CrossRef]
19. Jackson, Q.; Landgrebe, D.A. An adaptive classifier design for high-dimensional data analysis with a limited training data set. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2664–2679. [CrossRef]
20. Boukendour, O.; Lazri, M.; Absi, R.; Ouallouche, F.; Labadi, K.; Attaf, Y.; Belghit, A.; Ameur, S. A Semi-Supervised SVM-Firefly Hybrid for Rainfall Estimation from MSG Data. *Atmosphere* **2026**, *17*, 133. [CrossRef]
21. Su, H.; Cai, Y.; Du, Q. Firefly-Algorithm-Inspired Framework with Band Selection and Extreme Learning Machine for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *10*, 309–320. [CrossRef]
22. Ding, S.; Zhu, Z.; Zhang, X. An overview on semi-supervised support vector machine. *Neural Comput. Appl.* **2017**, *28*, 969–978. [CrossRef]
23. Van Engelen, J.E.; Hoos, H.H. A survey on semi supervised learning. *Mach. Learn.* **2020**, *109*, 373–440. [CrossRef]
24. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021*; IEEE: New York, NY, USA, 2021; pp. 10012–10022. [CrossRef]
25. Gao, Y.; Guan, J.; Zhang, F.; Wang, X.; Long, Z. Attention-Unet-Based Near-Real-Time Precipitation Estimation from Fengyun-4A Satellite Imageries. *Remote Sens.* **2022**, *14*, 2925. [CrossRef]
26. Yang, J.; Wan, H.; Shang, Z. Enhanced hybrid CNN and transformer network for remote sensing image change detection. *Sci. Rep.* **2025**, *15*, 10161. [CrossRef] [PubMed]
27. Yang, X.-S.; He, X. Firefly Algorithm: Recent Advances and Applications. *Int. J. Swarm Intell.* **2013**, *1*, 36–50. [CrossRef]
28. Yang, X.-S. Firefly Algorithms for Multimodal Optimization. In *Stochastic Algorithms: Foundations and Applications, Proceedings of the 5th International Symposium on Stochastic Algorithms, Foundations and Applications (SAGA 2009), Sapporo, Japan, 23−25 October 2009*; Watanabe, O., Zeugmann, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5792, pp. 169–178. [CrossRef]
29. Murphy, D. EUMETSAT Geostationary Meteorological Satellite Programs. In *Handbook of Satellite Applications*; Springer: New York, NY, USA, 2013; pp. 991–1019. ISBN 978-1-4419-7670-3. [CrossRef]
30. Goudenhoofdt, E.; Delobbe, L. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 195–203. [CrossRef]

31. McKee, J.L.; Binns, A.D. A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology. *Can. Water Resour. J.* **2016**, *41*, 186–203. [CrossRef]
32. Ochoa Rodriguez, S.; Wang, L.P.; Willems, P.; Onof, C. A review of radar rain gauge data merging methods and their potential for urban hydrological applications. *Water Resour. Res.* **2019**, *55*, 6356–6391. [CrossRef]
33. Li, W.; Jiang, H.; Li, D.; Bedient, P.B.; Fang, Z.N. Evaluation of radar precipitation products and assessment of the gauge radar merging methods in Southeast Texas for extreme precipitation events. *Remote Sens.* **2023**, *15*, 2033. [CrossRef]
34. Le, X.-H.; Koyama, N.; Kikuchi, K.; Yamanouchi, Y.; Fukaya, A.; Yamada, T. Evaluating geostatistical and statistical merging methods for radar–gauge rainfall integration: A multi-method comparative study. *Remote Sens.* **2025**, *17*, 2622. [CrossRef]
35. Ryu, S.; Song, J.J.; Lee, G. Radar–rain gauge merging for high-spatiotemporal-resolution rainfall estimation using radial basis function interpolation. *Remote Sens.* **2025**, *17*, 530. [CrossRef]
36. Cavoretto, R.; Haider, A.; Lancellotti, S.; Mezzanotte, D.; Noorizadegan, A. Adaptive residual subsampling algorithms for kernel interpolation based on cross validation techniques. *Constr. Math. Anal.* **2024**, *7*, 76–92. [CrossRef]
37. Qiu, Q.; Liu, J.; Tian, J.; Jiao, Y.; Li, C.; Wang, W.; Yu, F. Evaluation of the radar QPE and rain gauge data merging methods in Northern China. *Remote Sens.* **2020**, *12*, 363. [CrossRef]
38. Gao, K.; Mei, G.; Cuomo, S.; Piccialli, F.; Xu, N. ARBF: Adaptive radial basis function interpolation algorithm for irregularly scattered point sets. *Soft Comput.* **2020**, *24*, 17693–17704. [CrossRef]
39. Feidas, H.; Giannakos, A. Identifying precipitating clouds in Greece using multispectral infrared Meteosat Second Generation satellite data. *Theor. Appl. Climatol.* **2011**, *104*, 25–42. [CrossRef]
40. Lazri, M.; Ameur, S.; Mohia, Y. Instantaneous rainfall estimation using neural network from multispectral observations of SEVIRI radiometer and its application in estimation of daily and monthly rainfall. *Adv. Space Res.* **2014**, *53*, 138–155. [CrossRef]
41. Thies, B.; Nauss, T.; Bendix, J. Precipitation process and rainfall intensity differentiation using Meteosat Second Generation Spinning Enhanced Visible and Infrared Imager data. *J. Geophys. Res.* **2008**, *113*, D23206. [CrossRef]
42. Schmetz, J.; Tjemkes, S.A.; Gube, M.; van de Berg, L. Monitoring deep convection and convective overshooting with METEOSAT. *Adv. Space Res.* **1997**, *19*, 433–441. [CrossRef]
43. Ackerman, S.A.; Strabala, K.I.; Menzel, W.P.; Frey, R.A.; Moeller, C.C.; Gumley, L.E. Discriminating clear sky from clouds with MODIS. *J. Geophys. Res.* **1998**, *103*, 32141–32157. [CrossRef]
44. Strabala, K.I.; Ackerman, S.A.; Menzel, W.P. Cloud properties inferred from 8–12 μm data. *J. Appl. Meteorol. Climatol.* **1994**, *33*, 212–229. [CrossRef]
45. Ouallouche, F.; Lazri, M.; Labadi, K.; Alouache, D.; Mohia, Y.; Sehad, M.; Ameur, S. Hybrid RF–ConvLSTM Approach for Rainfall Estimation from MSG Data over Northern Algeria. *Atmosphere* **2026**, *17*, 296. [CrossRef]
46. Visweshkishan, V.; Santhosh Kumar, N.; Sasikala, M. Sentiment Analysis using Optimal Algorithm (AdamW). *Int. J. Res. Appl. Sci. Eng. Technol.* **2025**, *13*, 1367–1372. [CrossRef]
47. Shakya, S. Smart Fault Diagnostics using Convolutional Neural Network and Adam Stochastic Optimization. *J. Soft Comput. Paradig.* **2021**, *3*, 38–46. [CrossRef]
48. Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 19–24 June 2022*; IEEE: New York, NY, USA, 2022; pp. 11976–11986. [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.