Fırat Üniversitesi Deneysel ve Hesaplamalı Mühendislik Dergisi

FUJECE
Fırat University Journal of Experimental and Computational Engineering

Derin Öğrenme, Makine Öğrenmesi ve İstatistiksel Modellerin Yağış ve Sıcaklık Tahmin Performanslarının Karşılaştırılması: Hakkâri İli Örneği

Ertuğrul GÜL[1]*

[1]Çevre Sağlığı Programı, Tıbbi Hizmetler ve Teknikler Bölümü, Sağlık Hizmetleri MYO, Hakkari Üniversitesi, Hakkari, Türkiye.
[1]ertugrulgul@hakkari.edu.tr

Geliş Tarihi: 24.03.2026
Kabul Tarihi: 16.05.2026
Düzeltme Tarihi: 25.04.2026
doi: https://doi.org/10.62520/fujece.1915506
Araştırma Makalesi

Alıntı: E Gül, "Derin öğrenme, makine öğrenmesi ve istatistiksel modellerin yağış ve sıcaklık tahmin performanslarının karşılaştırılması: Hakkâri ili örneği", Fırat Üni. Deny. Müh. Derg., vol. 5, no 2, pp. 496-518, Haziran 2026.

Öz

Bu çalışma, Türkiye'nin güneydoğusundaki Hakkâri bölgesinde aylık ortalama sıcaklık ve toplam yağış tahmini için derin öğrenme (LSTM), makine öğrenmesi (Random Forest) ve istatistiksel (Prophet) modellerin karşılaştırmalı performansını değerlendirmektedir. 1940–2025 dönemine ait ERA5 reanaliz verileri kullanılmış; model doğruluğu hata kareler ortalamasının karekökü (RMSE), ortalama mutlak hata (MAE), belirtme katsayısı (R^2), Nash–Sutcliffe verimliliği (NSE) ve Kling–Gupta verimliliği (KGE) olmak üzere beş metrikle karşılaştırılmıştır. 85 yıllık tarihsel iklim kaydının zamansal analizi, tüm mevsimlerde istatistiksel olarak anlamlı ısınma eğilimleri (0,221°C/10 yıl, $p < 0{,}001$) ortaya koymaktadır; ilkbahar ısınması 1,578°C/10 yıl ($p < 0{,}001$) ile en belirgin mevsimsel eğilimi oluşturmaktadır. Toplam yağıştaki uzun dönemli değişim ihmal edilebilir düzeydedir (+0,60 mm/10 yıl). Kar örtüsü ve evapotranspirasyon doğrudan modellenmediğinden, kar-yağmur faz geçişi ve kuraklık riskine ilişkin yorumlar ampirik sonuç değil, fiziksel olarak olası hidrolojik çıkarımlar olarak sunulmaktadır. Sızıntı önleyici değerlendirme çerçevesinde sıcaklık tahmininde Prophet en yüksek doğruluğu sağlamıştır (RMSE = 1,51°C, MAE = 1,15°C, R^2 = 0,976, NSE = 0,976, KGE = 0,952). Random Forest ikinci sıradadır (RMSE = 1,96°C, MAE = 1,59°C, R^2 = 0,960, NSE = 0,960, KGE = 0,839); LSTM daha düşük performans göstermiştir (RMSE = 2,52°C, MAE = 2,08°C, R^2 = 0,933, NSE = 0,933, KGE = 0,791). Yağış tahmini, sıcaklık tahminine göre belirgin biçimde daha zorlu kalmıştır. Prophet yağışta en iyi performansı sergilemiştir (RMSE = 50,15 mm, MAE = 33,88 mm, R^2 = 0,601, NSE = 0,601, KGE = 0,681). Random Forest (RMSE = 52,26 mm, MAE = 35,56 mm, R^2 = 0,567, NSE = 0,567, KGE = 0,672) ve LSTM (RMSE = 53,80 mm, MAE = 36,80 mm, R^2 = 0,542, NSE = 0,542, KGE = 0,680) benzer ve sınırlı tahmin becerisi göstermiştir. Bu bulgular, sıcaklık gibi güçlü mevsimsellik gösteren değişkenlerin farklı modelleme yaklaşımlarıyla tahmin edilebilirliğini koruduğunu; buna karşılık aylık yağışın, aynı ay bilgisi dışlandığında ve gecikmeli öngörücüler kullanıldığında, sızıntı önleyici koşullar altında yalnızca orta düzeyde tahmin edilebilirlik sergilediğini göstermektedir. Gelecekteki çalışmalar, ek mekânsal öngörücüler, karla ilişkili değişkenler ve hidrometeorolojik göstergelerin dâhil edilmesinden ve ayrıca hibrit derin öğrenme ile topluluk öğrenmesi mimarilerinin test edilmesinden yarar sağlayabilir.

Anahtar kelimeler: İklim modelleme, ERA5 yeniden analiz, Zaman serisi analizi, Makine öğrenimi

*Yazışılan Yazar

İntihal Kontrol: Evet – Turnitin
Şikayet: fujece@firat.edu.tr

Firat Univ Jour. of Exp. and Comp. Eng., 5(2), 496-518, 2026

Firat University Journal of Experimental and Computational Engineering

Comparison of Deep Learning, Machine Learning, and Statistical Models for Precipitation and Temperature Forecasting Performance: A Case Study of Hakkari Province

Ertuğrul GÜL[1]*

[1]Environmental Health Program, Department of Medical Services and Techniques, Vocational School of Health Services, Hakkari University, Hakkari, Türkiye.
[1]ertugrulgul@hakkari.edu.tr

Received: 24.03.2026
Accepted: 16.05.2026
Revision: 25.04.2026
doi: https://doi.org/10.62520/fujece.1915506
Research Article

Citation: E Gül, "Comparison of deep learning, machine learning, and statistical models for precipitation and temperature forecasting performance: a case study of Hakkari province", Firat Univ. Jour.of Exper. and Comp. Eng., vol. 5, no 2, pp. 496-518, June 2026.

Abstract
This study compares three forecasting approaches for monthly mean temperature and total precipitation over Hakkari, southeastern Turkey: LSTM (deep learning), Random Forest (machine learning), and Prophet (statistical). ERA5 reanalysis data spanning 1940–2025 served as input, and model performance was evaluated using five metrics (RMSE, MAE, R^2, NSE, and KGE). Analysis of the 85-year ERA5 record reveals statistically significant warming across all seasons (0.221°C/10 years, $p < 0.001$), with spring warming being the strongest (1.578°C/10 years, $p < 0.001$). Long-term precipitation change remains negligible (+0.60 mm/10 years). Because snow cover and evapotranspiration were not directly modeled, interpretations related to snow-rain phase shifts and drought risk are treated as physically plausible implications rather than direct empirical conclusions. Under the leakage-safe evaluation, Prophet produced the best temperature performance (RMSE = 1.51°C, MAE = 1.15°C, R2 = 0.976, NSE = 0.976, KGE = 0.952), followed by Random Forest (RMSE = 1.96°C, MAE = 1.59°C, R2 = 0.960, NSE = 0.960, KGE = 0.839) and LSTM (RMSE = 2.52°C, MAE = 2.08°C, R^2 = 0.933, NSE = 0.933, KGE = 0.791). Precipitation forecasting was consistently more demanding than temperature forecasting. Prophet achieved the best precipitation performance (RMSE = 50.15 mm, MAE = 33.88 mm, R^2 = 0.601, NSE = 0.601, KGE = 0.681), while Random Forest (RMSE = 52.26 mm, MAE = 35.56 mm, R^2 = 0.567, NSE = 0.567, KGE = 0.672) and LSTM (RMSE = 53.80 mm, MAE = 36.80 mm, R^2 = 0.542, NSE = 0.542, KGE = 0.680) showed similar limited skill. These findings suggest that strongly seasonal variables such as temperature remain tractable across modeling frameworks, whereas monthly precipitation exhibits only moderate predictability under leakage-safe conditions when same-month information is excluded and lagged predictors are used. Future studies may benefit from incorporating additional spatial predictors, snow-related variables, and hydrometeorological indicators, as well as testing hybrid deep learning and ensemble learning architectures.

Keywords: Climate forecasting, ERA5 reanalysis, Time-series analysis, Machine learning

*Corresponding author

Plagiarism Checks: Yes – Turnitin
Complaints: fujece@firat.edu.tr

1. Introduction

Climate change has become one of the defining challenges of the contemporary era, with far-reaching consequences for environmental, economic, and social systems across the globe [1–3]. The World Economic Forum has identified extreme weather events as the second most pressing global risk for 2025, with their frequency and intensity projected to escalate over the coming decade. The most recent IPCC report warns that the +1.5°C warming threshold may be crossed sooner than previously projected, marking a critical turning point in global climate stability. In recent years, numerous historical records for air temperature, sea surface temperature, and ice extent have been broken in succession; July 22, 2024, was documented as the hottest day ever recorded in human history, and 2024 became the first calendar year in which the global mean temperature exceeded 1.5°C signaling a measurable new phase in anthropogenic climate change [4]. Against this backdrop, the accurate forecasting of meteorological variables carries considerable practical weight, particularly for water resource management, agricultural planning, and disaster risk reduction [5, 6]. At regional scales, reliable modeling of temperature and precipitation variability is central to sustaining water resources, maintaining agricultural productivity, and preserving ecosystem balance. More broadly, the capacity to model and project climate components with precision has become indispensable for both regional and global environmental governance and policy formulation [7, 8]. Although traditional statistical approaches have long underpinned this field, artificial intelligence-based deep learning and machine learning methods have seen rapid and widespread adoption in recent years, owing to their ability to handle complex, high-dimensional data and deliver improved predictive performance[1–3,8–13].

Traditional statistical methods such as ARIMA, SARIMA, and Prophet have long been employed for capturing temporal dependencies in climate data, yielding robust results particularly for stationary time series [3,14–16]. Their ability to represent nonlinear climate processes, however, remains inherently limited [17–19]. Machine learning models, notably Random Forest (RF) and XGBoost, have shown considerable promise in learning complex relationships within large datasets [20–22]. Deep learning approaches, particularly Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, have likewise demonstrated notable success in capturing long-term temporal dependencies in climate time series [23–25]. Comparative studies conducted across diverse geographic regions and climate regimes have generally found that deep learning models attain lower error metrics (RMSE, MAE) than their statistical and machine learning counterparts [21, 26, 27], though their computational demands remain substantially higher [28, 29]. Despite these global advancements, existing literature has remained limited regarding comparative modeling studies in Turkey’s high mountainous regions, particularly in the Southeastern and Eastern Anatolia areas. In regions such as Hakkari, characterized by complex topography and sparse meteorological observation networks, the comparative performance of different model types (statistical, machine learning, and deep learning) has not been sufficiently examined. In areas with complex terrain and high climatic variability, uncertainties persist about which model type can provide more reliable predictions. Furthermore, the performance differences of these models for various meteorological variables (precipitation and temperature) and their interactions with regional climatic characteristics have not been adequately explored. The accuracy of ERA5 data over complex terrain and the effectiveness of integrating these datasets with different modeling approaches require further investigation, particularly in high-altitude regions characterized by complex topography and the presence of glaciers, such as Hakkari.

This study aims to evaluate the comparative performance of deep learning (LSTM), machine learning (Random Forest), and statistical (Prophet) models in forecasting precipitation and temperature over the Hakkari region. Drawing on ERA5 reanalysis datasets, the predictive accuracy of these frameworks is systematically assessed through multiple performance metrics, including RMSE, MAE, R^2, NSE, and KGE. A central objective is to identify the most suitable modeling approach for climate prediction in complex mountainous terrain while clarifying the relative strengths and limitations of each model type. The results are intended to offer a replicable methodological framework for climate forecasting in data-scarce mountain environments, with practical implications for regional water resource management, agricultural planning, and the formulation of climate adaptation strategies. Beyond these applied dimensions, the study seeks to strengthen the scientific basis for developing early warning systems and decision-support tools in climatically sensitive regions such as Hakkari.

2. Materials and Methods

2.1. Study area

Hakkari, located in the southeastern region of Turkey, is a mountainous province situated directly adjacent to the borders of Iran and Iraq. The primary mountain ranges, collectively known as the Hakkari Mountains, extend in an east-west direction, shaping the area's topography. Approximately 88% of the province's terrain is mountainous and rugged. The provincial capital is positioned at an elevation of about 1720 meters above sea level and lies at approximately 37.58°N latitude and 43.73°E longitude (Figure 1). Hakkari exhibits characteristics of a continental climate, with cold and harsh winters and hot, dry summers. The region's high altitude and rugged terrain present challenges but also make it a significant area for climate modeling studies. Additionally, the province is notable for its natural resources, lakes, and rich wildlife. These geographic and climatic features collectively establish Hakkari as an important site for climate modeling research in high mountainous areas.

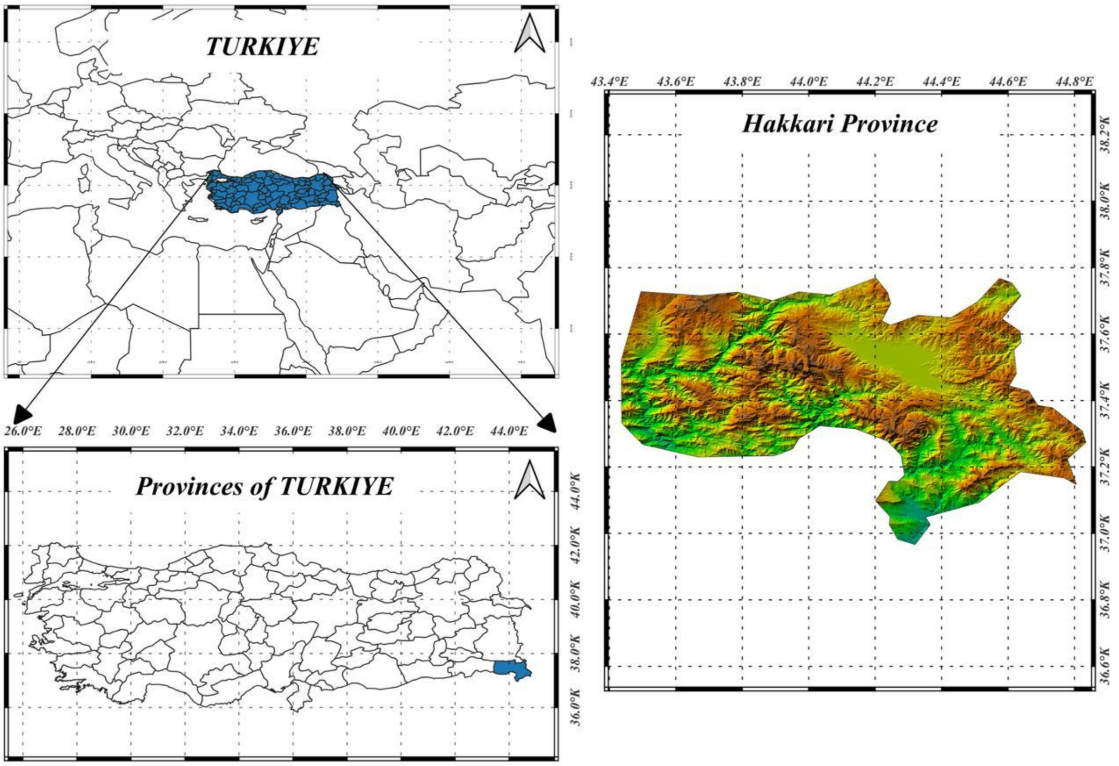

Figure 1. Geographical location and topography of Hakkari Province

2.2. Data collection and processing

In machine learning models, data preprocessing is a critical step that transforms raw meteorological data into a clean, consistent, and model-ready format, significantly enhancing model performance [30]. This study used ERA5 monthly reanalysis data for the period from January 1940 to May 2025 [31]. The total number of observations was 1025 monthly records, calculated as 85 complete years from 1940 to 2024 (85 x 12 = 1020 months) plus the first five months of 2025. The raw dataset included 33 meteorological variables, including wind components, temperature, pressure, humidity, radiation, and precipitation. The analysis was based on the ERA5 grid point closest to the center of Hakkari Province, located at approximately 37.5°N and 44.0°E. ERA5 has a regular latitude-longitude grid with a horizontal spatial resolution of 0.25° x 0.25°. No station-based meteorological observations were used in this study. Therefore, the dataset should be interpreted as a single ERA5 grid-point time series representative of Hakkari Province center. No in-situ meteorological station data were used; all analyses are based exclusively on ERA5 reanalysis output for this grid point. Since the analysis was based on a single ERA5 grid point rather than multiple spatial grid cells, spatial interpolation, or area-averaged fields, the study should therefore be understood as a univariate time-series forecasting analysis for a single representative grid point, not as a spatiotemporal or multi-site modeling framework. Extending the analysis to a spatial grid of ERA5 points or incorporating multi-site station data would be a natural direction for future work. Prior to modeling, meteorological variables were converted into physically meaningful units: temperature values in Kelvin were converted to Celsius, precipitation from meters to millimeters, and surface pressure from Pascals to hectopascals. Dew point temperature was converted to Celsius, followed by computation of saturation and actual vapor pressures. Relative humidity was subsequently calculated as the ratio of actual to saturation vapor pressures, bounded between 0 and 100 percent, and appended as a new dataset variable. Additional feature engineering included dew-point depression and cyclic month encodings using sin_month and cos_month.

The target variables were monthly mean 2-m air temperature (t2m_C) and monthly total precipitation (tp_mm). For the LSTM model, a 24-month look-back window was used, and each target month was predicted from the preceding 24 monthly records. Min-Max scalers were fitted only on the training sequences and then applied to the test sequences. For Random Forest and Prophet, dynamic meteorological predictors were converted into one-month lagged variables so that the prediction for month t used information available up to month t-1. Random Forest was implemented as a multi-output regression model, while Prophet was fitted separately for t2m_C and tp_mm with lagged external regressors. Correlation analysis was used as an exploratory screening step to examine redundancy among meteorological predictors: variables exhibiting Pearson $|r| > 0.90$ with another predictor were flagged as potentially redundant, and the less physically interpretable of the pair was excluded from the final feature set. This procedure was conducted solely on training-set data to avoid information leakage. For Random Forest and Prophet, all dynamic meteorological predictors were shifted by one month (lag-1) before being matched to the target month, ensuring that no same-month information was available to these models at inference time. For LSTM, the 24-month sequence window itself enforces temporal precedence, so no separate lag transformation was applied to the sequence inputs. The chronological order of the data was preserved throughout the workflow, and no random shuffling was applied during train-test splitting.

2.3. Models

2.3.1. Long-short term memory (LSTM)

The LSTM is a particular kind of RNN introduced by Hochreiter and Schmidhuber [32]. The standard RNN model cannot learn long-term dependence. Accordingly, LSTM was created to solve the problem of long-term dependence and to deal with the vanishing gradient issue. The LSTM architecture contains memory cells and gates that regulate data flow within the network and enable the retention of information over extended periods [33]. An LSTM network fundamentally comprises memory units referred to as cells. Every cell possesses two states: the cell state and the hidden state. The cells in the LSTM network facilitate critical decision-making by retaining or disregarding information regarding significant elements. These elements are referred to as gates, categorized into forget gates, input gates, and output gates [34]. The LSTM model functions in three phases, as seen in Figure 2.

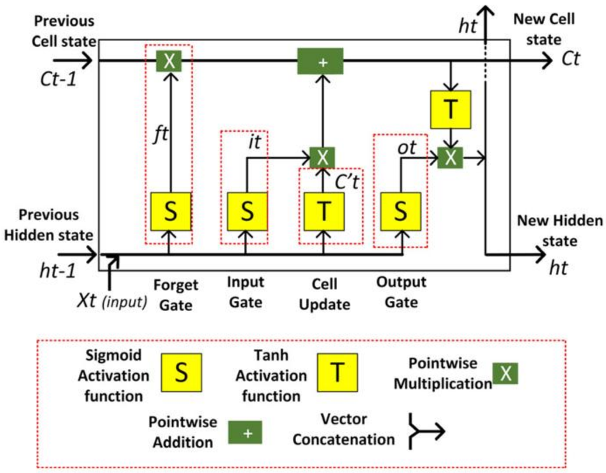

Figure 2. The architecture of LSTM network [33–36]

In the initial phase, the network uses the forget gate to determine which information should be disregarded or retained for the cell state. The calculation starts by considering the input at the current time step (x_t) and the previous value of the hidden state ($h_{(t-1)}$) using the sigmoid function (S) [35]. The calculation formula for the forgetfulness threshold is given in equation 1.

$$F_t = S\,(wf\,.\;[h_{(t\text{-}1)},\,x_t] + b_f) \tag{1}$$

During the second phase, the network's computation progresses by transforming the previous cell state ($C_{(t-1)}$) into a new cell state (C_t). This procedure determines the new information to be incorporated into long-term memory (cell state). To derive the new cell state value, the computation must consider the reference value from the forgetting gate, the input gate, and the cell update gate value. The equations for this phase are presented below.

$$i_t = \sigma\,(W_i\,.\;[h_{t\text{-}1}, x_t] + b_i) \tag{2}$$
$$C'_t = tanh(W_C\,.[h_{t-1}, x_t] + b_C\,) \tag{3}$$

In the third step, the previous cell state, C_{t-1}, is updated. The previous cell state is first multiplied by the forget gate, f_t, to remove unnecessary information. Subsequently, the product of the input gate, i_t, and the candidate cell state, C'_t, is added to obtain the updated cell state. This update is calculated as follows:

$$C_t = (f_t * C_{t\text{-}1}) + (i_t * C'_t) \tag{4}$$

In the final stage, the model output is determined. A sigmoid layer is first applied to identify which components of the cell state should be output. The cell state is then passed through the tanh function and multiplied by the output of the sigmoid gate. The computational formulas are given as follows:

$$o_t = \sigma\,(W_0\,.\;[h_{t-1},\,x_t] + b_0) \tag{5}$$
$$h_t = o_t * \tanh(C_t) \tag{6}$$

2.3.2. Random forest (rf)

Random Forest (RF), proposed by Breiman [37], is a non-parametric supervised learning method used for classification and regression tasks. The algorithm constructs an ensemble of multiple decision trees, each trained on a bootstrap sample randomly drawn—with replacement—from the original dataset, using different subsets of features for each tree. This randomness increases model diversity and reduces correlation among the trees. The final prediction is obtained by majority voting in classification or by averaging in regression. Validation is performed using out-of-bag (OOB) samples, which are the data points not included in the training subsets, eliminating the need for a separate validation dataset. The randomization in both data and feature selection enhances the model's generalization capability while substantially reducing the risk of overfitting. Parallel construction of decision trees improves computational efficiency, whereas ensemble aggregation enhances predictive accuracy. Owing to its ability to model nonlinear relationships and rank variable importance, Random Forest is particularly suitable for complex and high-dimensional environmental and climatic datasets [38–42]. The algorithm flow is shown in Figure 3.

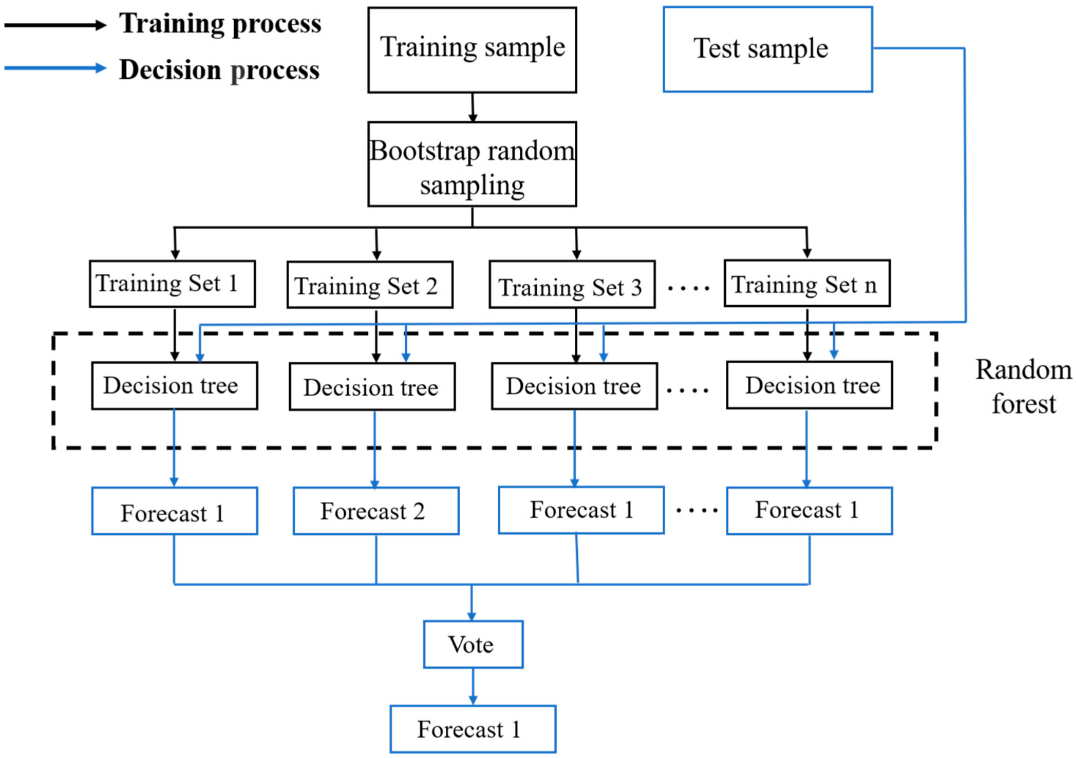

Figure 3. Flow chart of the random forest algorithm

2.3.3. Prophet

Facebook-Prophet is an open-source time-series forecasting model created by Facebook, intended to accommodate diverse time-series patterns, including seasonality and trend fluctuations [43–45]. The main advantages of Prophet include its ability to automatically handle missing values and outliers, its simple parameter tuning, high computational efficiency, and full automation. Therefore, it is particularly effective for environmental and climatic time series characterized by irregular or discontinuous observations [40,46,47]. The Prophet model operates on the notion of decomposition and subsequent recombination. It disaggregates the time series into four components: trend elements, seasonal factors, holiday influences, and stochastic noise. The forecast is subsequently integrated using either an additive or multiplicative model. The additive model is articulated as;

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \tag{7}$$

Here, y(t) represents the observed value of the time series at time t, while g(t) denotes the trend component that captures long-term non-periodic variations and can be modeled using either a piecewise linear or logistic growth function. The term s(t) represents the seasonal component, which accounts for periodic fluctuations such as daily, weekly, or annual cycles and is typically modeled using Fourier series. The component h(t) represents the effects of holidays or events that occur on irregular schedules and may cause sudden changes in the time series. Finally, ϵ_t denotes the error term, which captures random noise or idiosyncratic variations not explained by the model and is assumed to follow a normal distribution.

2.4. Model implementation and validation setup

To improve methodological transparency and reproducibility, the implementation details and hyperparameter settings used in the LSTM, Random Forest, and Prophet models are explicitly reported in Table 1. The same chronological training and testing partitions were used for all models to ensure a consistent comparison. The LSTM configuration includes the look-back window, network architecture, optimizer, batch size, epoch number, and early stopping criterion; the Random Forest configuration includes the number of trees, tree depth, and leaf-size settings; and the Prophet configuration includes the seasonality mode, yearly seasonality, and changepoint settings.

Table 1. Main hyperparameters used in the forecasting models

Model	Main implementation settings
LSTM	24-month look-back window; two LSTM layers (100 units → Dropout 0.2 → 50 units → Dropout 0.2); Dense layers (25 units, ReLU; output); batch size = 32; maximum epochs = 100; Adam optimizer; learning rate = 0.001; early stopping patience = 10 (monitor: val_loss).
Random Forest	n_estimators = 300; max_depth = 15; min_samples_leaf = 3; random_state = 42; multi-output regression used for the two target variables.
Prophet	Default Prophet settings; separate models for temperature and precipitation; lagged external regressors added using add_regressor.

2.5. Model performance evaluation

Model performance was evaluated using RMSE, MAE, R^2, NSE, and KGE. In addition, residual diagnostics, scatter plots, monthly RMSE patterns, Taylor diagrams, and Friedman tests with post-hoc Wilcoxon comparisons were used to support model interpretation. The equations for the performance metrics are as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y_i})^2} \tag{8}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \overline{y_i}| \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \overline{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{10}$$

$$NSE = 1 - \frac{MSE}{\sigma^2} \tag{11}$$

$$KGE = 1 - \sqrt{(r-1)^2(\alpha-1)^2(\beta-1)^2} \tag{12}$$

where n is the number of data point computed, y_i is the observed value, $\overline{y_i}$ is the predicted value, $\overline{y}$ is the mean observed value. RMSE is employed to demonstrate the disparity between predicted and actual values. Reduced values of MSE and RMSE signify superior predictive performance by machine learning models. The R^2 coefficient indicates the extent of correlation between expected and actual values. An R^2 score approaching 1 signifies that the model is an excellent match. NSE is directly associated with MSE in Equation (11) divided by σ^2, the variance of the true value. KGE comprises three separate components that indicate correlation, bias, and a measure of relative variability between simulated and actual values. r represents the linear correlation between observations and simulations, α quantifies the flow variability error, and β signifies bias [27,40,41].

3. Results and Discussion

3.1. Temporal analysis of historical climate data (1940–2025)

Historical climate data analysis for Hakkari reveals significant temporal patterns in key meteorological variables over the 85-year study period (Table 2), indicating a statistically significant warming trend across the region's climate system accompanied by consistent signals of progressive drying in all seasons. Monthly climatological records reveal the pronounced seasonal contrasts characteristic of Hakkari's mountainous Eastern Anatolia region. Winter conditions (monthly means ranging from −7.22°C to −5.23°C) exhibit substantial thermal variability (standard deviation ±2.03–2.13°C) and account for approximately 42.0% (428.33 mm) of the annual precipitation. Spring constitutes a transitional phase marked by rapid aridification from March (173.86 mm) to May (82.96 mm), coinciding with a pronounced 10.34°C increase in temperature (from −1.13°C in March to +9.21°C in May). The summer season (June–August) represents an extremely dry period with only 27.13 mm of total precipitation, while mean temperatures stabilize around ~18.89°C with low variability; monthly rainfall reaches just 5.21 mm in July and 3.95 mm in August. Autumn reflects a hydrological and thermal transition, characterized by the onset of moisture recovery in October (58.72 mm) and intensified precipitation in November (102.97 mm) accompanying rapid cooling (0.89°C). The annual precipitation sum is 1020.35 mm, of which 38.7% occurs in spring and 16.6% in autumn.

Table 2. Descriptive statistics of monthly rainfall and temperature extremes

Month	mean		std		min		max	
	(°C)	(mm)	(°C)	(mm)	(°C)	(mm)	(°C)	(mm)
Jan (Winter)	-7.22	146.99	2.03	74.51	-13.69	11.10	-1.75	346.01
Feb (Winter)	-5.23	148.66	2.06	64.99	-10.29	35.34	-1.23	334.06
Mar (Spring)	-1.13	173.86	2.11	62.18	-5.96	45.26	3.85	351.60
Apr (Spring)	3.42	138.37	1.62	55.23	0.29	38.82	8.16	284.24
May (Spring)	9.21	82.96	1.71	42.19	4.98	7.51	14.38	200.04
Jun (Summer)	15.93	17.97	1.38	12.99	13.15	0.5	19.22	59.06
Jul (Summer)	20.52	5.21	1	6.87	18.39	0	22.68	38.07
Aug (Summer)	20.21	3.95	1.16	5.32	17.86	0	22.99	32.61
Sep (Autumn)	15.7	8.01	1.14	7.80	13.46	0	18.58	40.31
Oct (Autumn)	8.68	58.72	1.41	40.37	5.11	4.47	11.86	175.26
Nov (Autumn)	0.89	102.97	2.15	62.09	-4.26	0	5.72	294.30
Dec (Winter)	-5.45	132.68	2.13	71.60	-10.06	11.86	-1.3	368.37

Seasonal trend analysis was performed using the non-parametric Mann-Kendall test for significance assessment and Sen's slope estimator for the magnitude of change, both applied to monthly and seasonally aggregated time series. As shown in Table 3, seasonal decomposition of mean temperatures demonstrates statistically significant upward trends across all seasons. Warming rates were observed in spring (MAM: 1.578°C/10 years, $p < 0.001$), summer (JJA: 0.759°C/10 years, $p < 0.001$), autumn (SON: 0.503°C/10 years, $p = 0.002$), and winter (DJF: 0.687°C/10 years, $p = 0.001$). These results indicate year-round warming, with the strongest increase occurring in spring, where the warming rate is approximately 2.1 to 3.1 times higher than in the other seasons. This enhanced spring warming coincides with simultaneous increases in both precipitation (+10.791 mm/10 years, $p = 0.100$) and temperature (+1.578°C/10 years, $p < 0.001$) during the wettest period of the year (March mean ~174 mm), suggesting a possible shift from snow-dominated to rain-dominated conditions. Such pronounced seasonal asymmetry may contribute to earlier snowmelt dynamics in mountainous regions, as earlier and more rapid melt of the snowpack shifts the traditional spring–summer runoff peak forward, effectively moving snow-derived water limitation into an earlier part of the hydrological year. Additionally, the systematic winter drying trend (−6.501 mm/10 years) further supports the concern that the region's snowpack accumulation potential is increasingly at risk.

Table 3. Seasonal trend analysis results

	Temperatures		**Precipitation**	
	Trend (°C/10 years)	P-value	Trend (mm/10 years)	P-value
Winter	0.687	0.001	-6.501	0.362
Spring	1.578	0.000	10.791	0.100
Summer	0.759	0.000	-0.358	0.750
Autumn	0.503	0.002	-4.590	0.438
	0.221	**0.000**	**0.595**	**0.551**

The 85-year (1940-2025) ERA5 grid-point record representative of Hakkari Province center exhibits statistically significant warming accompanied by seasonally variable precipitation behavior across the regional climate system. The temperature anomaly time series (relative to the 1981-2010 baseline) reveals a marked regime transition beginning in the late 1980s: while anomalies fluctuate around zero between 1940 and 1980, the post-1990 period is characterized by a persistent shift into the positive anomaly domain. Following 2010, temperature anomalies frequently exceed +2°C and approach +2.5°C toward 2025. The precipitation series shows only a negligible long-term change (+0.60 mm per 10 years). Therefore, the results support a robust warming signal, whereas statements about reduced snow persistence, evapotranspiration-driven drying, or hydrological drought risk should be understood as plausible implications that require direct confirmation with snow, runoff, evapotranspiration, or drought-index data (Figure 4).

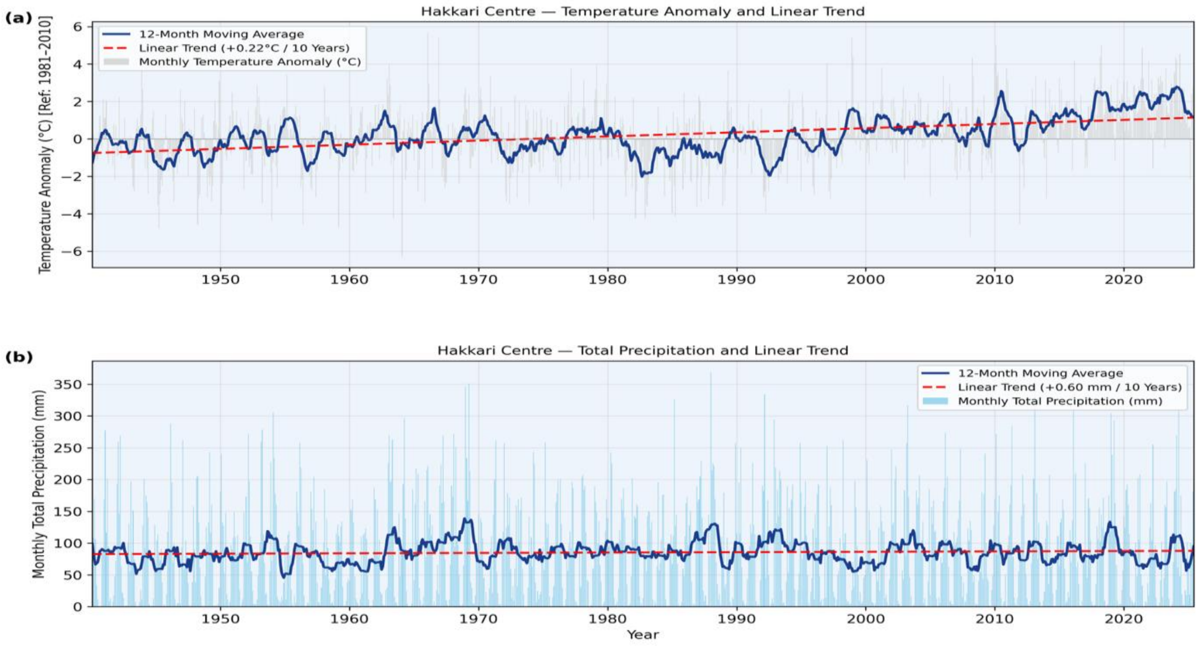

Figure 4. Temperature anomaly and total precipitation in Hakkari

3.2. Performance analysis of forecasting models

The comparative performance analysis of three models developed to predict monthly mean temperature (t2m_C) and total precipitation (tp_mm) for Hakkari is presented within a leakage-safe feature processing framework. Correlation analysis was used as an exploratory screening step to examine redundancy among meteorological predictors and guide the final feature set rather than as a standalone inferential procedure (Figure 5). The final modeling workflow used lagged dynamic predictors for Random Forest and Prophet, while LSTM used a 24-month look-back sequence.

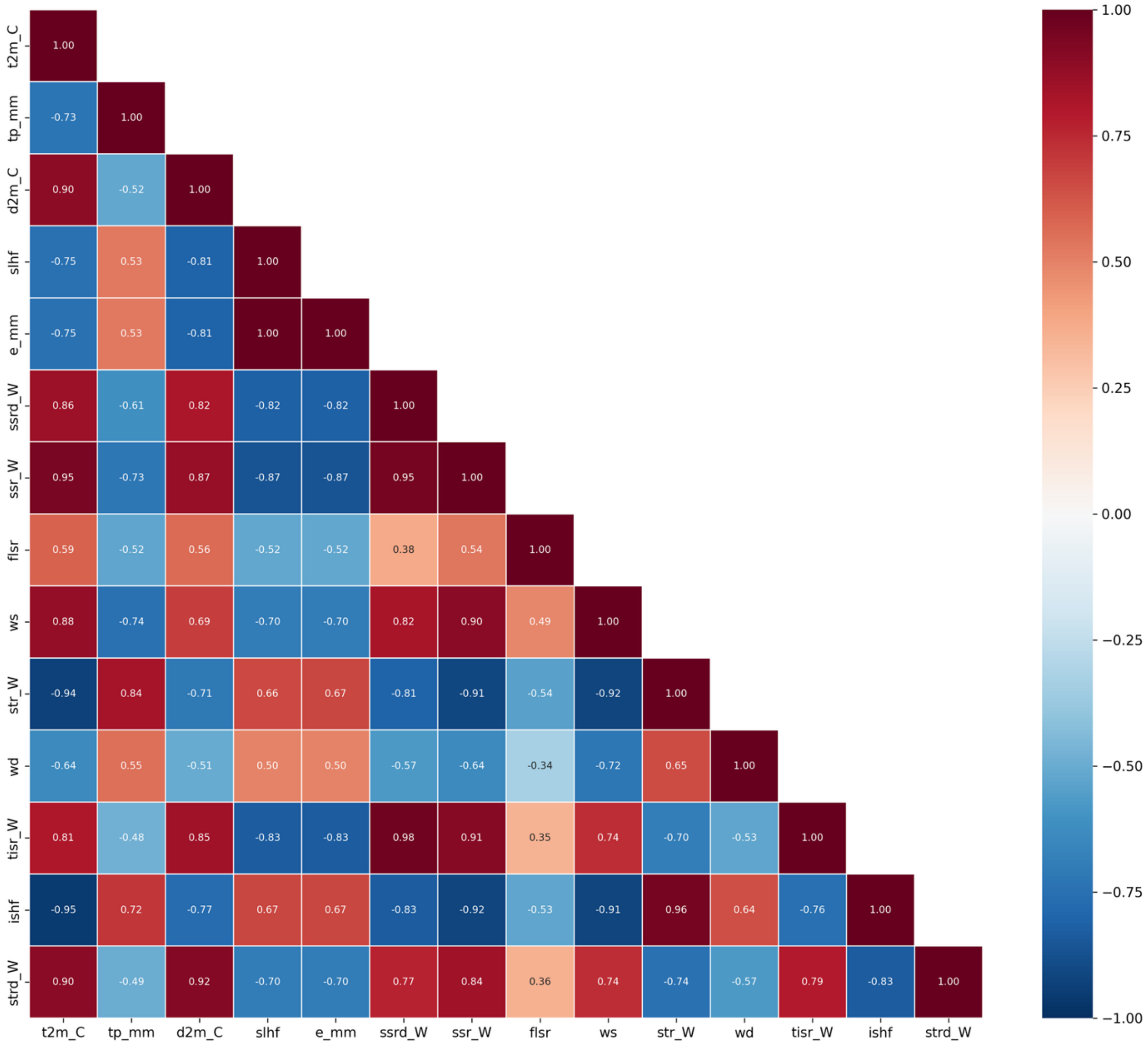

Figure 5. Correlation matrix of final selected features

The model-ready dataset included selected meteorological variables and derived predictors such as dew_point_depression, sin_month, and cos_month. For Random Forest and Prophet, dynamic meteorological variables were shifted by one month to avoid using information from the target month. For LSTM, the sequence input consisted of the preceding 24 monthly records. This design allowed the models to use temporally available information while reducing the risk of information leakage.

3.2.1. LSTM

Figure 6 illustrates the residual distribution for temperature (a) and precipitation (b), revealing substantial differences in the statistical behavior of the two variables. For temperature, the histogram displays a unimodal and nearly symmetric structure, and the QQ-plot aligns closely with the theoretical normal distribution line. This indicates that the residuals are well-behaved, largely normally distributed, and free from systematic bias, reflecting a stable and reliable predictive performance. In contrast, the precipitation residuals show a highly skewed and heavy-tailed distribution, with a dense clustering of small errors near zero and pronounced deviations at the positive tail. The corresponding QQ-plot exhibits substantial departures from normality, particularly in extreme quantiles, confirming the model's difficulty in representing high-intensity precipitation events. These results collectively highlight that while the LSTM model captures the general structure of temperature dynamics [23,48,49], precipitation forecasts are hindered by the inherent variability and extremes present in rainfall data [50,51].

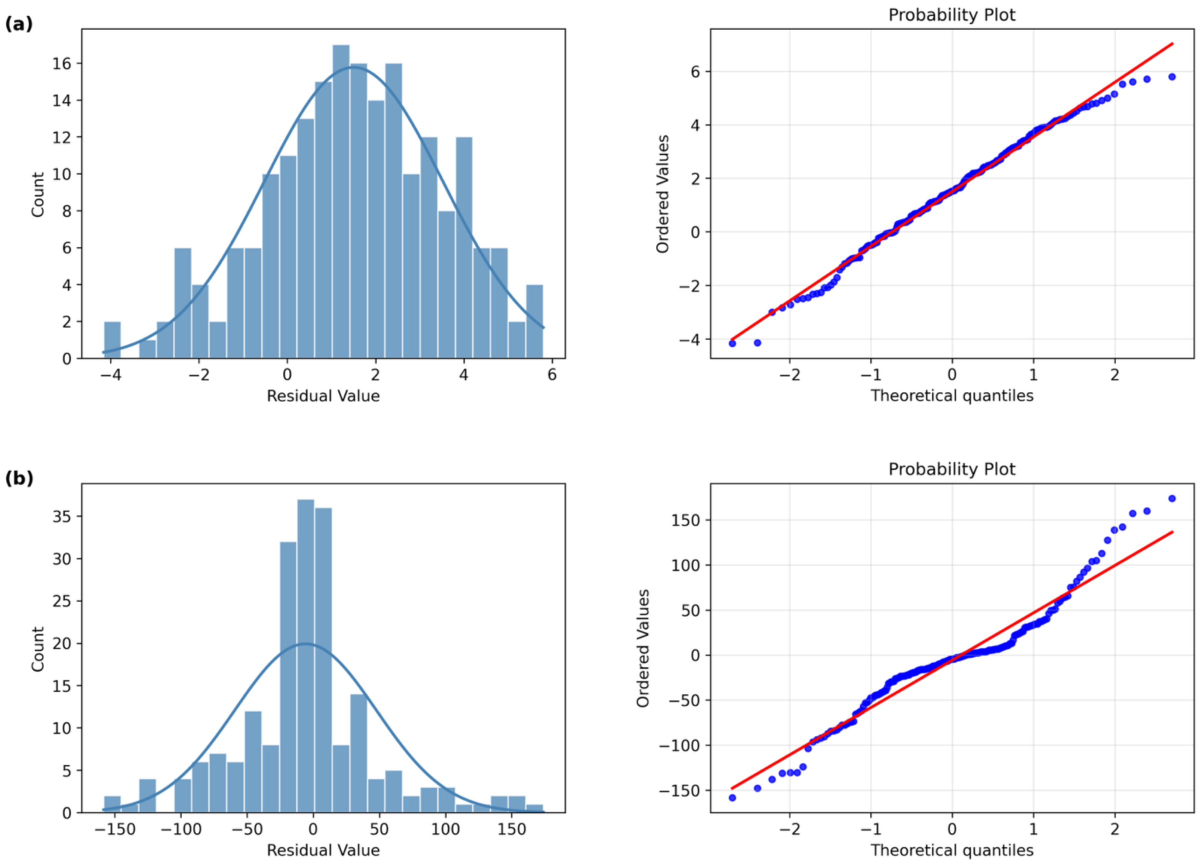

Figure 6. Residual distribution and normality assessment of LSTM predictions: (a) temperature (t2m) and (b) precipitation (tp)

Figure 7 provides a direct comparison between actual and predicted values, further emphasizing the divergent performance of the model for temperature (a) and precipitation (b). The temperature scatter plot exhibits a strong linear alignment along the 1:1 reference line, demonstrating high predictive accuracy and minimal dispersion around the ideal trajectory. This pattern confirms that the model successfully learns the temporal structure of temperature and generalizes well across the test samples. Conversely, the precipitation scatter plot shows a much wider spread with a clear tendency toward systematic underestimation of high precipitation values. While predictions are reasonably accurate at low and moderate rainfall amounts, the model fails to track the magnitude of heavy precipitation events. The substantial deviation from the reference line reflects the challenges posed by the non-linear and extreme-event-dominated nature of precipitation, which the LSTM architecture struggles to fully represent [52, 53].

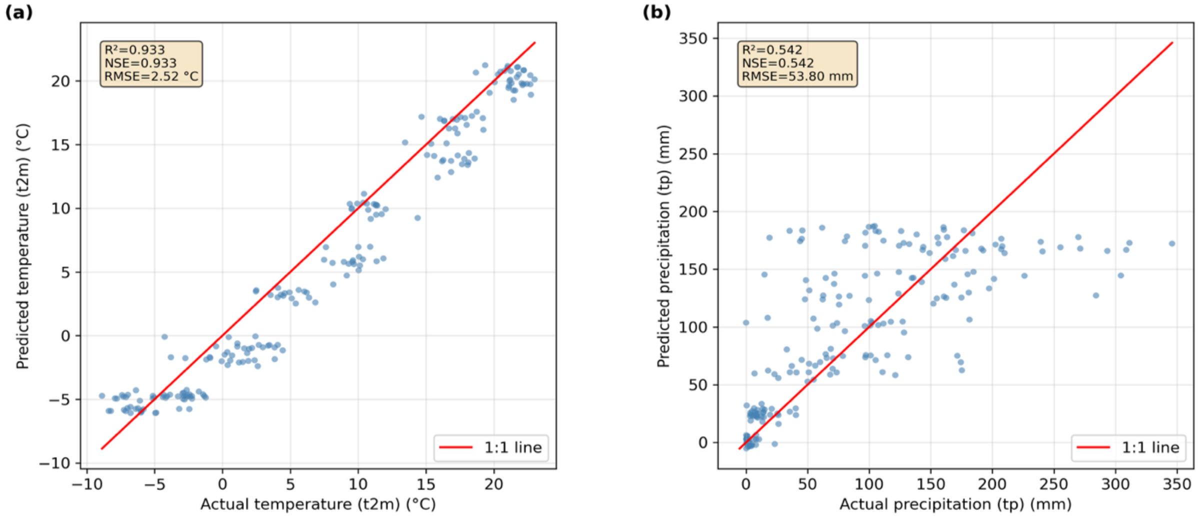

Figure 7. Scatter plots comparing actual and predicted values obtained from the LSTM model: (a) temperature (t2m) and (b) precipitation (tp)

Figure 8 demonstrates the seasonal evolution of prediction errors and reinforces the insights obtained from the previous analyses. For temperature, monthly RMSE values remain consistently low throughout the year, with noticeable increases only in March, October, and November—months characterized by transitional meteorological conditions and rapid temperature fluctuations. The minimal error observed during the summer months indicates that the model performs particularly well under stable atmospheric regimes. In contrast, precipitation RMSE exhibits a pronounced seasonal pattern, with the highest errors recorded in January, March, and December. These months coincide with periods of intense winter precipitation and frequent snow–rain transitions in Hakkari, which introduce abrupt hydrometeorological variability that the model is unable to capture adequately. During the dry summer season, RMSE values decrease sharply, reflecting the reduced complexity of the precipitation signal. Overall, Figure 8 confirms that temperature predictions are robust across seasons, whereas precipitation forecasts deteriorate significantly during high-variability winter periods.

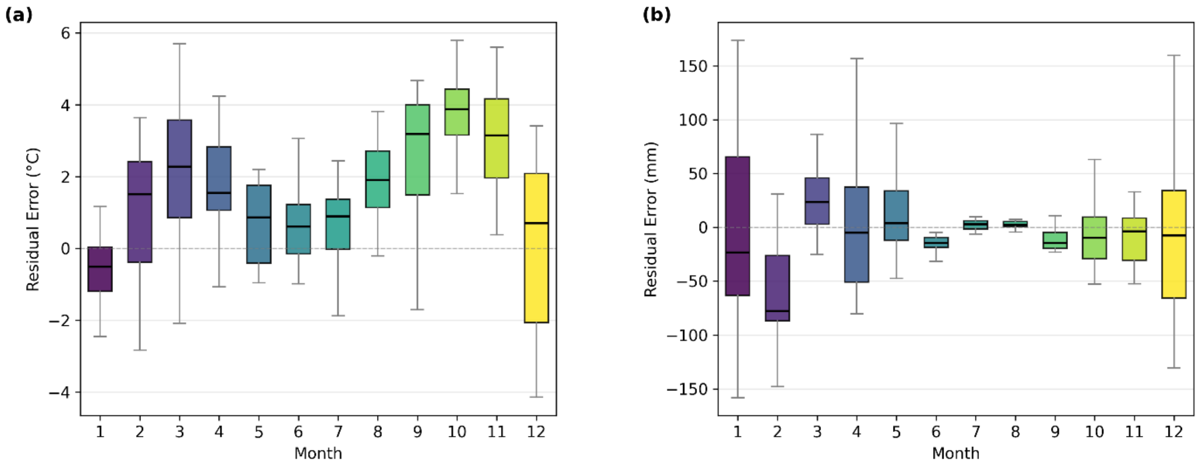

Figure 8. Monthly RMSE values for the LSTM model: (a) temperature (t2m) and (b) precipitation (tp)

3.2.2. Random forest

The residual distribution presented in Figure 9 reveals that the Random Forest model exhibits markedly different behavior for temperature and precipitation, closely paralleling the patterns observed in the LSTM model. For temperature, the histogram shows a nearly symmetric and unimodal distribution, and the QQ-plot aligns well with the theoretical normal line, indicating that residuals are statistically well-behaved and free from systematic bias. This mirrors the LSTM results, where temperature errors were also normally distributed with minimal tail deviations. In contrast, precipitation residuals display a heavily right-skewed, heavy-tailed structure dominated by extreme values. The QQ-plot demonstrates substantial divergence in the upper quantiles, confirming that, like LSTM, the Random Forest model struggles to represent high-intensity precipitation events. These findings collectively show that both models produce stable and reliable temperature predictions yet face inherent limitations when modeling the extreme-event–dominated nature of precipitation.

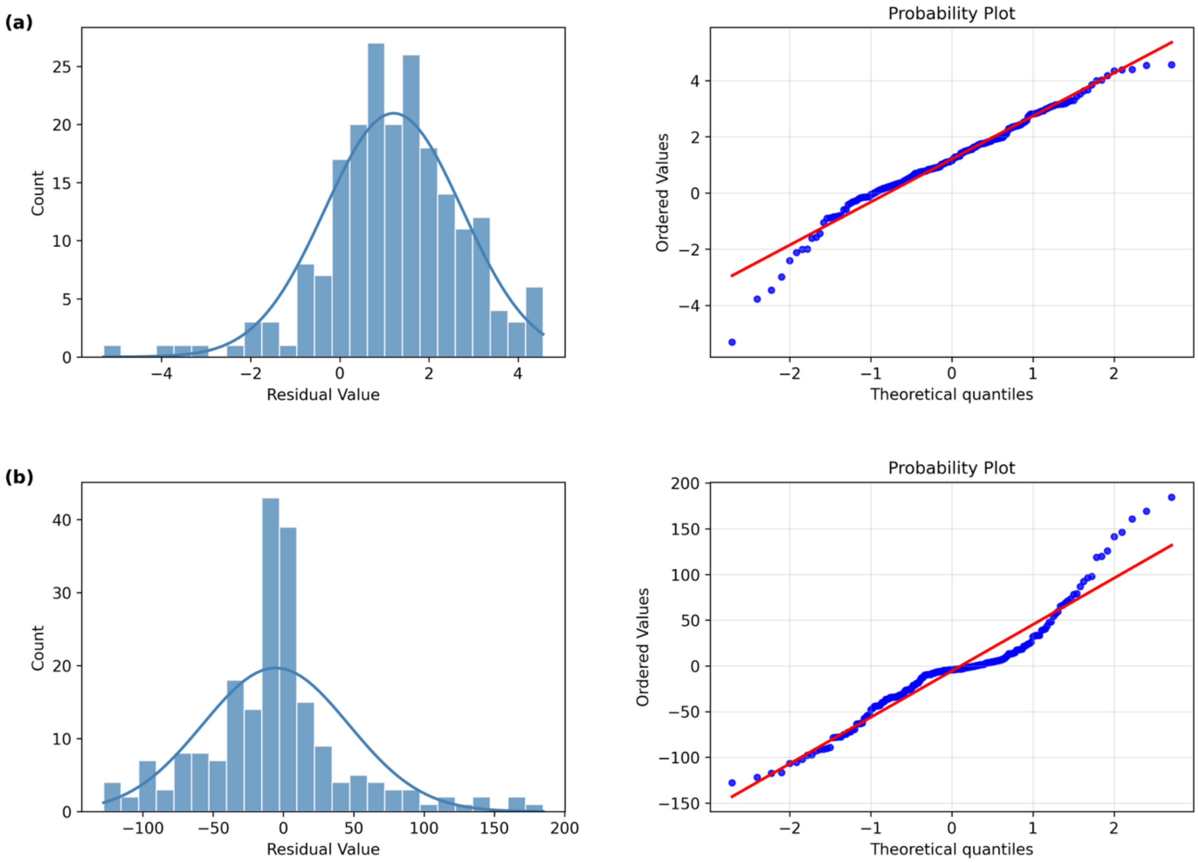

Figure 9. Residual distribution and normality assessment of Random Forest predictions: (a) temperature (t2m) and (b) precipitation (tp)

For temperature, Random Forest predictions show a generally coherent relationship with observed values, although the leakage-safe evaluation indicates slightly weaker performance than Prophet but better than LSTM (R^2 = 0.9602). For precipitation, Random Forest performance is moderate (R^2 = 0.5665) and slightly better than LSTM (R^2 = 0.5421). The precipitation scatter plot therefore indicates that both models struggle to represent high-magnitude monthly precipitation events when same-month meteorological information is excluded (Figure 10).

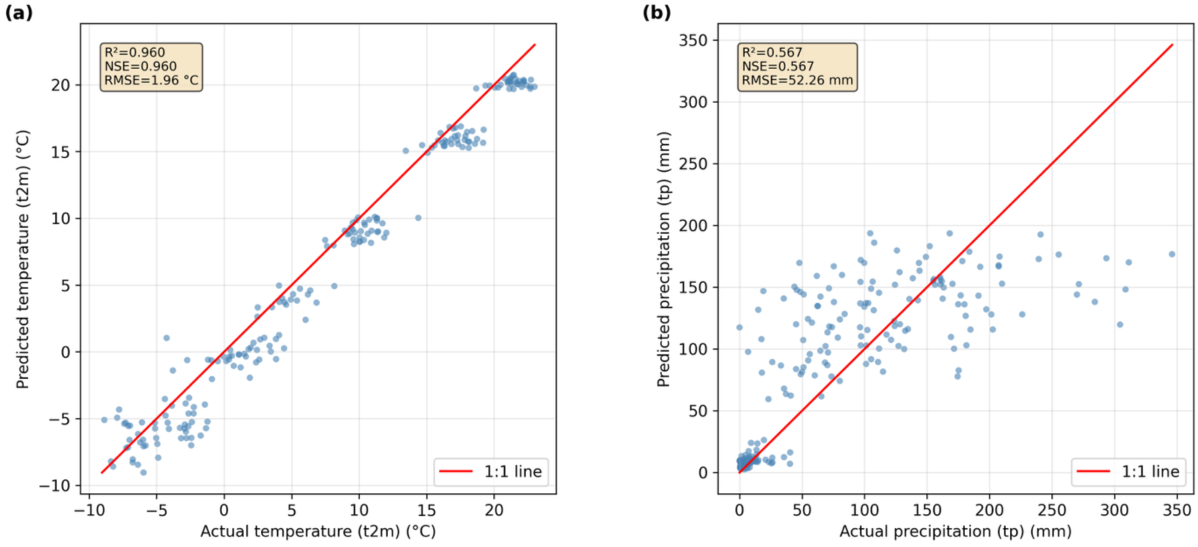

Figure 10. Scatter plots comparing actual and predicted values obtained from the Random Forest model: (a) temperature (t2m) and (b) precipitation (tp)

Figure 11 shows the seasonal characteristics of Random Forest prediction errors under the leakage-safe setup. Temperature RMSE remains relatively low across the year, whereas precipitation RMSE is consistently higher and more seasonally variable. The error structure indicates that Random Forest provides only moderate monthly precipitation skill when lagged predictors are used, with larger errors during wetter months and smaller errors during the dry summer period.

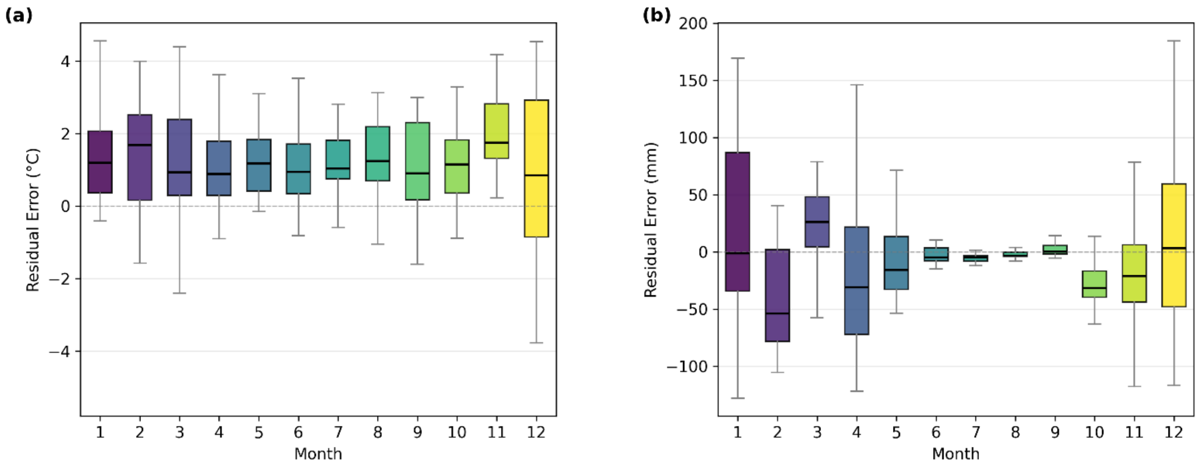

Figure 11. Monthly RMSE values for the Random Forest model: (a) temperature (t2m) and (b) precipitation (tp)

3.2.3. Prophet

The residual distributions further highlight the differences in Prophet model performance (Figure 12). For temperature, Prophet residuals remain relatively concentrated around zero, indicating that the model captures the dominant seasonal structure of monthly temperature. In contrast, precipitation residuals are more dispersed and heavy-tailed, with larger deviations during high-precipitation months. These results indicate that Prophet, despite achieving the best precipitation performance among the tested models, still has limited skill in representing abrupt and high-magnitude monthly precipitation events.

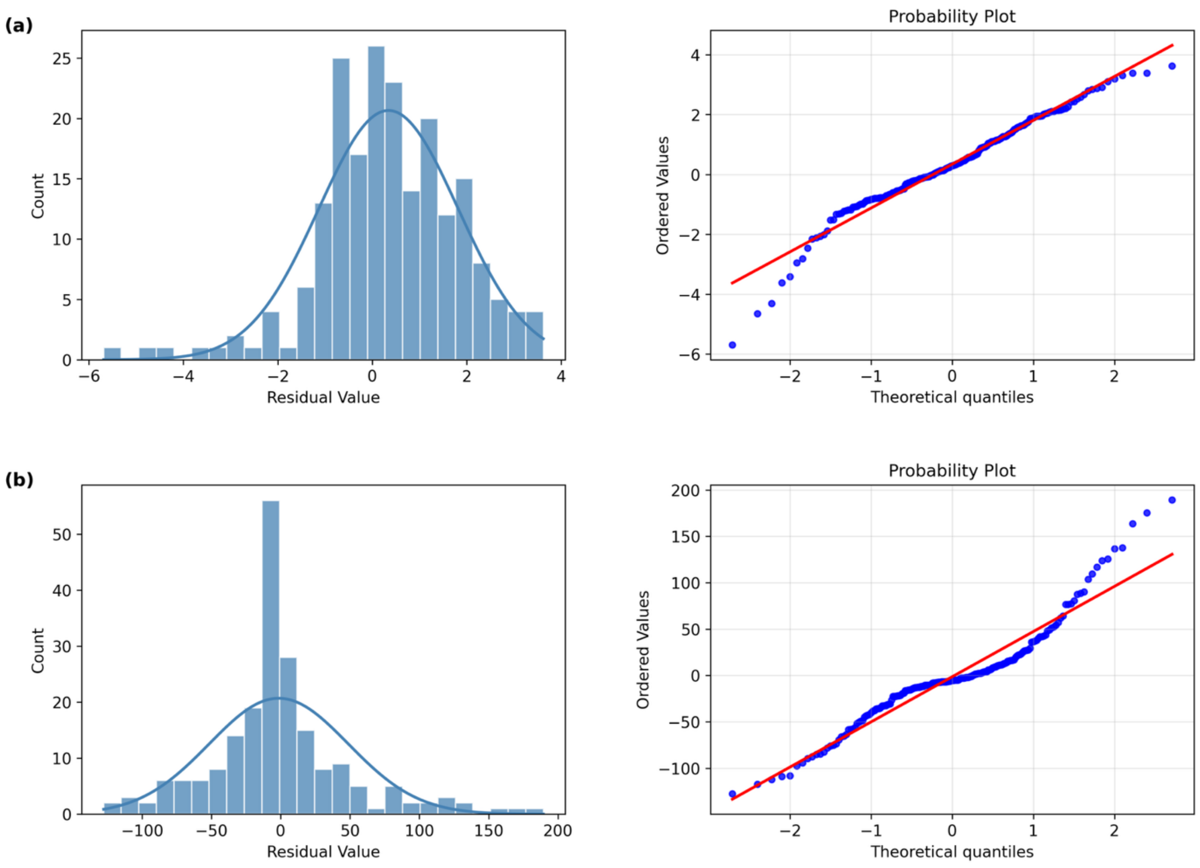

Figure 12. Residual distribution and normality assessment of Prophet predictions: (a) temperature (t2m) and (b) precipitation (tp)

Prophet demonstrates the strongest temperature forecasting performance among the three models, with predicted monthly temperatures clustering closely around the 1:1 line (Figure 13(a)). In contrast, precipitation forecasts show a wider scatter (Figure 13(b)), especially at higher precipitation values. This pattern confirms that monthly precipitation remains substantially more difficult to predict than temperature, even when Prophet is supplied with lagged external regressors.

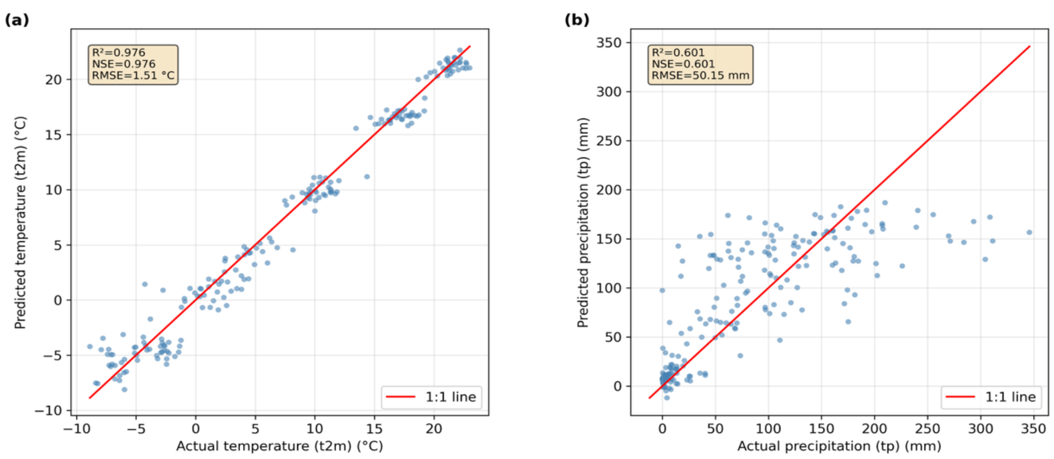

Figure 13. Scatter plots comparing actual and predicted values obtained from the Prophet model: (a) temperature (t2m) and (b) precipitation (tp)

Prophet's forecast accuracy varies seasonally, especially for precipitation (Figure 14). Temperature errors remain comparatively low across the year, whereas precipitation errors increase during wetter months and decline during the dry summer period. This seasonal pattern suggests that Prophet captures broad monthly seasonality but remains limited in predicting the magnitude of high-precipitation events under leakage-safe evaluation.

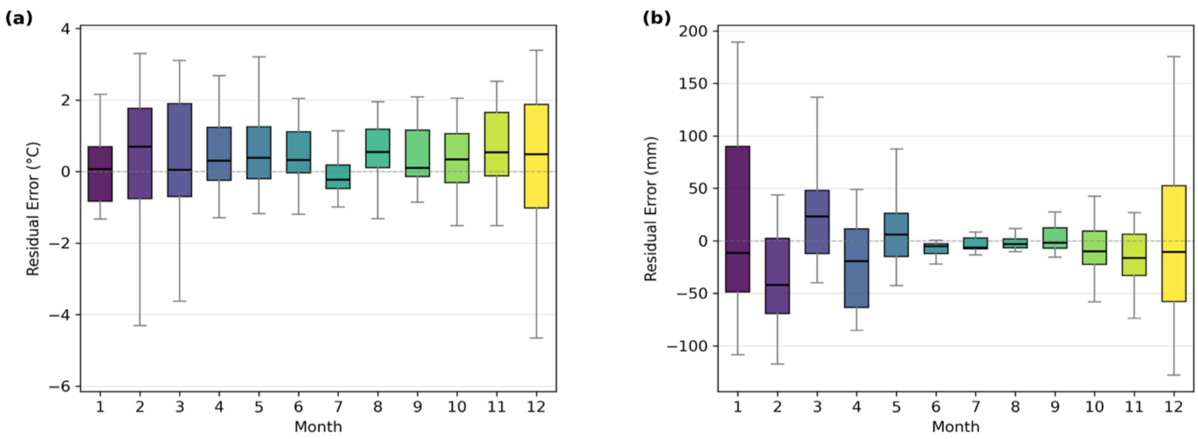

Figure 14. Monthly RMSE values for the Prophet model: (a) temperature (t2m) and (b) precipitation (tp)

3.2.4. Comparative performance analysis of forecasting models

The comparative evaluation of the LSTM, Random Forest, and Prophet models demonstrates distinct performance characteristics shaped by both the statistical structure of the predictors and the intrinsic modeling strategies of each method. Across all three models, temperature (t2m) emerges as a substantially more predictable variable, while precipitation (tp) exhibits pronounced non-linearity, intermittency, and extreme-event dominance, resulting in lower model fidelity and higher uncertainty (Table 4).

Table 4. Comparative performance metrics of LSTM, Random Forest, and Prophet models for monthly mean temperature (t2m) and total precipitation (tp) forecasting

Model	RMSE		MAE		R^2		NSE		KGE	
	t	TP	t	TP	t	TP	t	TP	T	TP
LSTM	2.52	53.80	2.08	36.80	0.93	0.54	0.93	0.54	0.79	0.68
Random Forest	1.96	52.26	1.59	35.56	0.96	0.57	0.96	0.57	0.84	0.67
Prophet	1.51	50.15	1.15	33.88	0.98	0.60	0.98	0.60	0.95	0.68

*t: temperature, TP: precipitation

For temperature forecasting, the three models showed differing levels of predictive accuracy under the leakage-safe evaluation framework. Prophet produced the best overall temperature forecasts, with the lowest error values (RMSE = 1.5136°C; MAE = 1.1485°C) and the highest goodness-of-fit statistics (R^2 = 0.9761; NSE = 0.9761; KGE = 0.9517). LSTM showed moderate performance (RMSE = 2.5227°C; MAE = 2.0845°C; R^2 = 0.9334; NSE = 0.9334; KGE = 0.7911), suggesting that while the 24-month look-back structure partially captured the temporal persistence and seasonal behavior of temperature, its generalization was weaker than the other approaches under the leakage-safe evaluation. Random Forest yielded a better performance (RMSE = 1.9559°C; MAE = 1.5855°C; R^2 = 0.9602; NSE = 0.9602; KGE = 0.8388). Overall, these results confirm that monthly temperature is a relatively predictable variable in the Hakkari ERA5 grid-point series, even when leakage-safe preprocessing and lagged predictors are used.

In contrast, precipitation forecasting remained substantially more difficult. All models produced only moderate predictive skill for monthly total precipitation, reflecting the intermittent and event-driven nature of rainfall in the mountainous Hakkari region. Prophet achieved the best precipitation performance among the three models (RMSE = 50.1468 mm; MAE = 33.8796 mm; R^2 = 0.6009; NSE = 0.6009; KGE = 0.6814), but its accuracy remained much lower than for temperature. Random Forest and LSTM showed very similar precipitation performance, with Random Forest producing RMSE = 52.2578 mm and R^2 = 0.5665, while LSTM produced RMSE = 53.8049 mm and R^2 = 0.5421. These results indicate that predictive skill remains limited for precipitation when same-month information is excluded and only lagged predictors are used.

Seasonal RMSE patterns further support this contrast between temperature and precipitation predictability. Temperature errors remained relatively low across the year for all models, with moderate increases during transitional months when atmospheric conditions change more rapidly. Precipitation errors, however, showed stronger seasonal clustering, with larger errors during wet-season and winter months and smaller errors during the dry summer period. This pattern suggests that monthly precipitation forecasting in Hakkari is strongly constrained by seasonal intermittency and high-magnitude precipitation events. Because snow cover, precipitation phase, and snow-water equivalent were not directly modeled, interpretations related to snowfall-rainfall transitions should be treated as physically plausible explanations rather than direct empirical conclusions.

To assess the statistical significance of performance differences, a Friedman test was applied to the absolute residuals of all three models. For temperature forecasting, the Friedman test yielded a statistically significant result ($X^2 = 66.03$, df = 2, $p < 0.001$), and post-hoc pairwise Wilcoxon tests with Bonferroni correction ($\alpha = 0.0167$) confirmed that all model pairs differed significantly (LSTM vs. Random Forest: $p < 0.001$; LSTM vs. Prophet: $p < 0.001$; Random Forest vs. Prophet: $p < 0.001$). For precipitation forecasting, the Friedman test showed no statistically significant differences among the three models ($X^2 = 0.67$, df = 2, $p = 0.717$), indicating that their precipitation skill was equivalent. Taylor diagram analysis (Figure 15) further illustrates the relative performance of each model: for temperature, Prophet exhibited the highest correlation with observations and the closest standard deviation to the reference, followed by Random Forest and then LSTM; for precipitation, all three models clustered at comparable distances from the reference point, consistent with the non-significant Friedman result. These pairwise comparisons are illustrated in Figure 16.

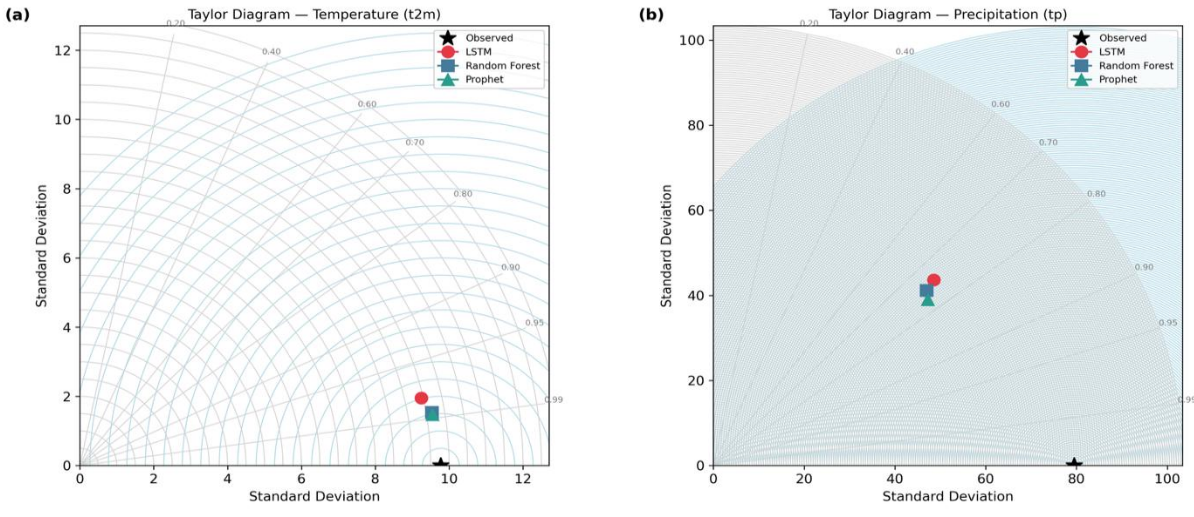

Figure 15. Taylor diagrams comparing LSTM, Random Forest, and Prophet against ERA5 observations for (a) temperature (t2m) and (b) precipitation (tp)

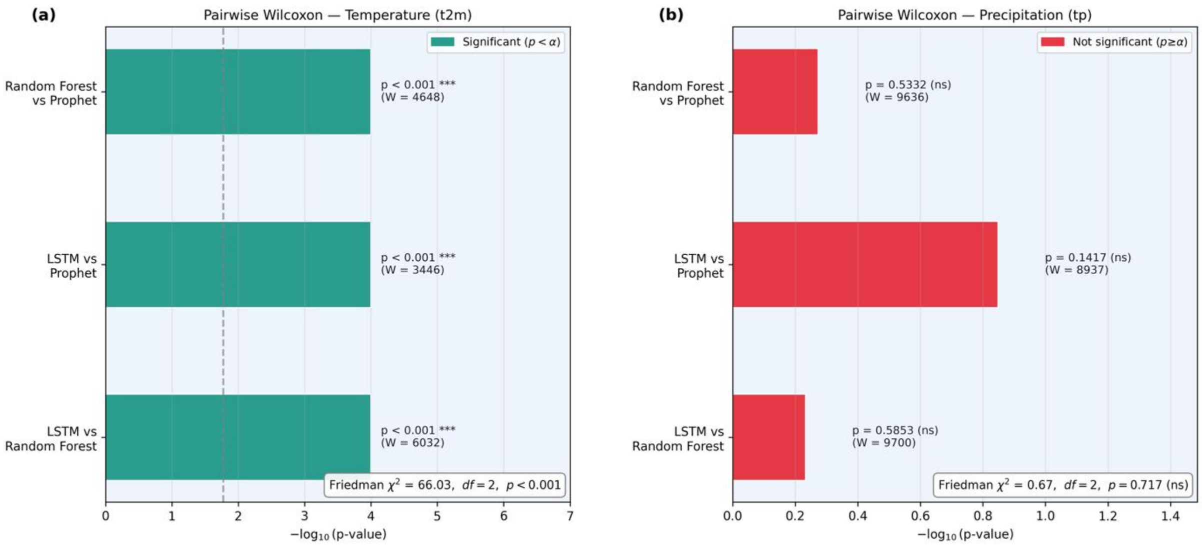

Figure 16. Post-hoc pairwise Wilcoxon signed-rank tests of model residuals with Bonferroni correction (α = 0.0167) for (a) temperature and (b) precipitation

In conclusion, the leakage-safe results of this study show that monthly precipitation forecasting in high-elevation mountainous regions such as Hakkari remains substantially more challenging than temperature forecasting. This difficulty is likely related to the intermittent, seasonal, and high-variability nature of precipitation in complex mountainous terrain. Several limitations of the current study should be acknowledged. The validation framework employed a single chronological 80/20 train-test split, which provides an initial assessment of out-of-sample performance but may not fully capture temporal variability or the long-term generalizability of the fitted models. More robust strategies such as walk-forward validation or rolling-origin evaluation are recommended in future work to provide stronger evidence of model stability across different time periods. Furthermore, the analysis is based on a single ERA5 grid point, which may not fully represent the spatial heterogeneity of Hakkari's complex mountainous terrain; incorporating multi-point or spatially distributed input data could improve representativeness. Future studies may benefit from testing hybrid deep learning architectures, such as CNN–LSTM [54–57], ConvLSTM [58,59], QLSTM [50], and CNN–XGBoost [60], which have been reported in the literature to capture both short-term fluctuations and longer temporal dependencies. In addition, ensemble gradient boosting algorithms such as XGBoost [61–64], LightGBM [65–68], and CatBoost [69–72] should be further investigated as complementary modeling frameworks for representing non-linear precipitation processes under challenging orographic conditions.

4. Acknowledgements

This study was supported by the Scientific Research Projects Coordination Office of Hakkari University, Türkiye (Project No: FM19BAP).

5. Ethics Committee Approval and Conflict of Interest

“There is no conflict of interest with any person/institution in the prepared article”
“The authors declare no conflict of interest with any individual or institution regarding the prepared manuscript.”

6. Ethical Statement Regarding the Use of Artificial Intelligence

“During the preparation of this manuscript, the artificial intelligence tool "Gemini" developed by "Google" was used to a limited extent solely for linguistic refinement. The scientific content, analysis, and results are entirely the responsibility of the author(s).”

7. References

[1] J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses, "Air Temperature Forecasting Using Machine Learning Techniques: A Review," *Energies*, vol. 13, p. 4215, 2020.
[2] F. Shah and A. Sharifi, "Climate Models for Predicting Precipitation and Temperature Trends in Cities: A Systematic Review," *Sustain. Cities Soc.*, vol. 120, p. 106171, 2025.
[3] V. S. Yavuz, "Forecasting Monthly Rainfall and Temperature Patterns in Van Province, Türkiye, Using ARIMA and SARIMA Models: A Long-Term Climate Analysis," *J. Water Clim. Change*, vol. 16, pp. 800–818, 2025.
[4] L. Adil, D. Eckstein, V. Künzel, and L. Schäfer, *Climate Risk Index: Who Suffers Most from Extreme Weather Events?* Bonn, Germany: Germanwatch, 2025, p. 72.
[5] International Water Management Institute (IWMI), *IWMI Annual Report 2023*. Colombo, Sri Lanka: IWMI, 2024.
[6] M. Manucharyan, "Climate Change Impacts on Sustainable Agriculture: Evidence from Armenia," *Unconv. Resour.*, vol. 6, p. 100159, 2025.
[7] OECD, *Climate Change, Water and Agriculture: Towards Resilient Systems*, OECD Studies on Water. Paris, France: OECD Publishing, 2014.
[8] M. M. Raghuwanshi, Y. Katre, A. Sahu, D. Sharma, A. Udapure, and C. Lonarkar, "Weather Prediction with Machine Learning," *Int. J. Innov. Sci. Res. Technol.*, pp. 3271–3275, 2024.
[9] P. Fan, J. Yang, Z. Zhang, N. Zang, Y. Li, and G. Feng, "Summer Precipitation Prediction in Eastern China Based on Machine Learning," *Clim. Dyn.*, vol. 60, pp. 2645–2663, 2023.
[10] S. Gui, Z. Yang, H. Wu, R. Yang, Z. Dong, Y. Yang, and Y. Lian, "Improving the Hindcast of Late Spring-Early Summer Heatwaves in Southeast Asia with Machine Learning Models," *Clim. Dyn.*, vol. 63, p. 302, 2025.
[11] D. Kumar, S. Shekhar, and T. Tewary, "Editorial: AI and Data Analytics for Climate Data Management," *Front. Environ. Sci.*, vol. 13, 2025.
[12] V. Ramachandra, *Artificial Intelligence in Climate Science: A State-of-the-Art Review (2020–2025)*, 2025.
[13] R. Sha and T. Guha, "Climate Time Series Prediction with Deep Learning and LSTM," in *Proc. 4th Int. Conf. Smart Electron. Commun. (ICOSEC)*, Sep. 2023, pp. 1631–1637.
[14] K. E. Fahim, L. C. De Silva, and H. Yassin, "Time Series Forecasting of Active Power Using ARIMA, SARIMA and Hybrid Models," in *Proc. Smart Sustain. Develop. Conf. (SSD 2025)*, S. Singh, B. Dowdeswell, J. Sun, I. Ardekani, F. Mehdipour, and W. Hewage, Eds. Dordrecht, Netherlands: Atlantis Press Int. BV, 2025, vol. 2, pp. 203–219.
[15] S. Kumari and P. Muthulakshmi, "SARIMA Model: An Efficient Machine Learning Technique for Weather Forecasting," *Procedia Comput. Sci.*, vol. 235, pp. 656–670, 2024.
[16] Y. Sabzevari and S. Eslamian, "Climate Warming Prediction Using Time Series Analysis: Case Study, Western Iran," *AUT J. Civ. Eng.*, 2024.
[17] T. Balasubramaniam, W. A. Mohotti, K. Sabir, and R. Nayak, "Feature Engineering on Climate Data with Machine Learning to Understand Time-Lagging Effects in Pasture Yield Prediction," *Ecol. Inform.*, vol. 86, p. 103011, 2025.
[18] Y.-J. Kim, S.-J. Lee, H.-S. Jin, I.-A. Suh, and S.-Y. Song, "Comparison of Linear and Nonlinear Statistical Models for Analyzing Determinants of Residential Energy Consumption," *Energy Build.*, vol. 223, p. 110226, 2020.
[19] R. Yang, J. Hu, Z. Li, J. Mu, T. Yu, J. Xia, X. Li, A. Dasgupta, and H. Xiong, "Interpretable Machine Learning for Weather and Climate Prediction: A Review," *Atmos. Environ.*, vol. 338, p. 120797, 2024.
[20] X. Cui, Y. Xu, L. Sun, and T. Yao, "Machine Learning Models for Predicting Rural Residential Carbon Emissions and Optimising Spatial Forms," *Sci. Rep.*, vol. 15, p. 33111, 2025.
[21] T. Hovorushchenko, V. Alekseiko, and V. Levashenko, "Machine Learning Methods' Comparison for Land Surface Temperatures Forecasting Due to Climate Classification," in *Proc. 1st Int. Workshop Adv. Appl. Inf. Technol. (AdvAIT 2024)*, CEUR Workshop Proc., Zilina, Slovakia, 2024, vol. 3899, pp. 66–72.
[22] S. Saiful and N. B. Wibisono, "Crop Yield Prediction Using Random Forest Algorithm and XGBoost Machine Learning Model," *Int. J. Res. Innov. Soc. Sci.*, vol. IX, pp. 1983–1994, 2025.

[23] Q. Guo, Z. He, and Z. Wang, "Monthly Climate Prediction Using Deep Convolutional Neural Network and Long Short-Term Memory," *Sci. Rep.*, vol. 14, p. 17748, 2024.
[24] Q. Liu, M. Yang, K. Mohammadi, D. Song, J. Bi, and G. Wang, "Machine Learning Crop Yield Models Based on Meteorological Features and Comparison with a Process-Based Model," *Artif. Intell. Earth Syst.*, vol. 1, 2022.
[25] M. Waqas and U. W. Humphries, "A Critical Review of RNN and LSTM Variants in Hydrological Time Series Predictions," *MethodsX*, vol. 13, p. 102946, 2024.
[26] S. N. H. Bukhari and K. A. Ogudo, "Forecasting Temperature and Rainfall Using Deep Learning for the Challenging Climates of Northern India," *PeerJ Comput. Sci.*, vol. 11, p. e3012, 2025.
[27] F. Z. Che Rose, N. A. K. Rosili, and M. F. Marsani, "Comparison of Machine Learning Model Performance for Predicting the Climate Variables in Johor Bahru, Malaysia," *Sci. Rep.*, vol. 15, p. 23465, 2025.
[28] C. O. De Burgh-Day and T. Leeuwenburg, "Machine Learning for Numerical Weather and Climate Modelling: A Review," *Geosci. Model Dev.*, vol. 16, pp. 6433–6477, 2023.
[29] R. Ranjan and A. Shrivastava, "Evaluating Deep Learning Models for Predicting Climate Change Trends: A Review," *Int. J. Innov. Res. Technol.*, vol. 11, pp. 2326–2329, 2025.
[30] S. Khatri, "Advanced Weather Forecasting with Machine Learning: Leveraging Meteorological Data for Improved Predictions," *Commun. Appl. Nonlinear Anal.*, vol. 32, pp. 1–16, 2025.
[31] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, *et al*., "The ERA5 Global Reanalysis," *Q. J. R. Meteorol. Soc.*, vol. 146, pp. 1999–2049, 2020.
[32] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, pp. 1735–1780, 1997.
[33] F. Farhangi, A. Sadeghi-Niaraki, J. Safari Bazargani, S. V. Razavi-Termeh, D. Hussain, and S.-M. Choi, "Time-Series Hourly Sea Surface Temperature Prediction Using Deep Neural Network Models," *J. Mar. Sci. Eng.*, vol. 11, p. 1136, 2023.
[34] X. Zhang, Q. Zhang, G. Zhang, Z. Nie, Z. Gui, and H. Que, "A Novel Hybrid Data-Driven Model for Daily Land Surface Temperature Forecasting Using Long Short-Term Memory Neural Network Based on Ensemble Empirical Mode Decomposition," *Int. J. Environ. Res. Public Health*, vol. 15, p. 1032, 2018.
[35] F. Aksan, Y. Li, V. Suresh, and P. Janik, "CNN-LSTM vs. LSTM-CNN to Predict Power Flow Direction: A Case Study of the High-Voltage Subnet of Northeast Germany," *Sensors*, vol. 23, p. 901, 2023.
[36] F. Shafiq, A. Zafar, M. U. Ghani Khan, S. Iqbal, A. S. Albesher, and M. N. Asghar, "Extreme Heat Prediction through Deep Learning and Explainable AI," *PLoS One*, vol. 20, p. e0316367, 2025.
[37] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
[38] G. Ali and W. Robert, "Machine Learning for Temperature Analysis in Ouagadougou: A Random Forest Perspective," *EDRAAK*, vol. 2024, pp. 101–105, 2024.
[39] Y. He, J. Li, F. Bi, Y. Mao, B. Li, C. Chen, and Y. Chu, "Prediction of Near-Surface Air Temperature at Glacier Region Using ERA5 Data Based on Random Forest Regression Method," in *Proc. 3rd Int. Conf. Geol., Mapping Remote Sens. (ICGMRS)*, Zhoushan, China, Apr. 2022, pp. 179–182.
[40] Q. Ji, X. Jia, L. Jiang, M. Xie, Z. Meng, Y. Wang, and X. Lin, "Contribution of Atmospheric Factors in Predicting Sea Surface Temperature in the East China Sea Using the Random Forest and SA-ConvLSTM Model," *Atmosphere*, vol. 15, p. 670, 2024.
[41] T. Magesh, C. Sai Pratap, A. Vishal, T. Naveen Kumar, M. Karthikeyan, and S. Iniyan, "Enhanced Accuracy of Weather Forecasting Using Hybrid LSTM and Random Forest Models," in *Proc. Int. Conf. Power, Energy, Control Transm. Syst. (ICPECTS)*, Oct. 2024, pp. 1–5.
[42] W. Zhang, Z. Li, Y. Tian, and L. Liu, "Research on Temperature Prediction Based on RF-LSTM Modeling," in *Proc. 6th Int. Acad. Exchange Conf. Sci. Technol. Innov. (IAECST)*, Dec. 2024, pp. 873–876.
[43] M. G. S. Kenyi and K. Yamamoto, "A Hybrid SARIMA-Prophet Model for Predicting Historical Streamflow Time-Series of the Sobat River in South Sudan," *Discov. Appl. Sci.*, vol. 6, p. 457, 2024.

[44] M. J. A. Shohan, M. O. Faruque, and S. Y. Foo, “Forecasting of Electric Load Using a Hybrid LSTM-Neural Prophet Model,” *Energies*, vol. 15, p. 2158, 2022.
[45] J. Tang, L. Xu, C. Fan, C. Xu, and Y. Ning, “A Short-Term Prediction of Ionospheric TEC Using the RF-Prophet Model for GNSS Stations in China,” *Adv. Space Res.*, vol. 76, pp. 3654–3669, 2025.
[46] C. B. Aditya Satrio, W. Darmawan, B. U. Nadia, and N. Hanafiah, “Time Series Analysis and Forecasting of Coronavirus Disease in Indonesia Using ARIMA Model and PROPHET,” *Procedia Comput. Sci.*, vol. 179, pp. 524–532, 2021.
[47] J. Tang, Y. Li, D. Yang, and M. Ding, “An Approach for Predicting Global Ionospheric TEC Using Machine Learning,” *Remote Sens.*, vol. 14, p. 1585, 2022.
[48] G. Codeluppi, A. Cilfone, L. Davoli, and G. Ferrari, “AI at the Edge: A Smart Gateway for Greenhouse Air Temperature Forecasting,” in *Proc. IEEE Int. Workshop Metrol. Agric. Forestry (MetroAgriFor)*, Nov. 2020, pp. 348–353.
[49] Y.-S. Lin, S.-L. Fang, L. Kang, C.-C. Chen, M.-H. Yao, and B.-J. Kuo, “Combining Recurrent Neural Network and Sigmoid Growth Models for Short-Term Temperature Forecasting and Tomato Growth Prediction in a Plastic Greenhouse,” *Horticulturae*, vol. 10, p. 230, 2024.
[50] Y. Dong and H. Ding, “Bayesian Optimization of Hybrid Quantum LSTM in a Mixed Model for Precipitation Forecasting,” *Mach. Learn. Sci. Technol.*, vol. 6, p. 015065, 2025.
[51] Y. Li and S. Dev, “Short-Term Precipitation Forecasting Based on BU-LSTM Model in Dublin City, Ireland,” in *Proc. 17th Int. Congr. Image Signal Process., BioMed. Eng. Informatics (CISP-BMEI)*, Oct. 2024, pp. 1–6.
[52] A. H. Saputra, I. M. A. Satya, F. P. Sari, and A. Mulya, “Spatial Rain Probabilistic Prediction Performance Using Cost-Sensitive Learning Algorithm,” *E3S Web Conf.*, vol. 464, p. 19001, 2023.
[53] Y. Wang and L. Zhang, “Using Long Short-Term Memory (LSTM) to Merge Precipitation Data over Mountainous Area in Sierra Nevada,” 2024.
[54] C. J. Gamboa-Villafruela, J. C. Fernández-Alvarez, M. Márquez-Mijares, A. Pérez-Alarcón, and A. J. Batista-Leyva, “Convolutional LSTM Architecture for Precipitation Nowcasting Using Satellite Data,” in *Proc. 4th Int. Electron. Conf. Atmos. Sci.*, Jun. 2021, p. 33.
[55] S. Salma, “Hybrid CNN-LSTM Model: Rainfall Analysis and Prediction for Karnataka Region,” 2022.
[56] Y. Wang, P. Jia, Z. Shu, K. Liu, and A. R. M. Shariff, “Multidimensional Precipitation Index Prediction Based on CNN-LSTM Hybrid Framework,” 2025.
[57] M. Fan, O. Imran, A. Singh, and S. A. Ajila, “Using CNN-LSTM Model for Weather Forecasting,” in *Proc. IEEE Int. Conf. Big Data*, Dec. 2022, pp. 4120–4125.
[58] W. Liu, Y. Wang, D. Zhong, S. Xie, and J. Xu, “ConvLSTM Network-Based Rainfall Nowcasting Method with Combined Reflectance and Radar-Retrieved Wind Field as Inputs,” *Atmosphere*, vol. 13, p. 411, 2022.
[59] Y. Ding, B. Wang, Q. Yang, S. Ren, F. Yang, and L. Sun, “High-Resolution Climate Prediction in Mountainous Terrain Using a ConvLSTM-XGBoost Hybrid Model with Dynamic Bayesian Weighting,” *Sci. Rep.*, vol. 15, p. 36923, 2025.
[60] M. S. Islam, M. Shafiuzzaman, G. Mahmud, N. Nowshin, P. Reza, J. Hasan, M. F. Ahamed, M. Nahiduzzaman, M. A. Ayari, and A. Khandakar, “Explainable Deep Learning for Rainfall Prediction: A CNN-XGBoost Hybrid Approach in the Northern Region of Bangladesh,” *Neural Comput. Appl.*, vol. 37, pp. 28125–28160, 2025.
[61] V. Khettar, N. Kitratporn, S. Lerk-u-suke, J. Chaiwongsai, P. Jeefoo, and C. Sukawattanavijit, “Indices for Extreme Rainfall Risk Mapping in Thailand Using XGBoost,” 2025.
[62] S. A. Putri, G. Darmawan, R. Arisanti, and C. C. Clorinda, “Extreme Gradient Boosting Method Forecasting Rainfall in Lembang District, West Java Province,” *Indo-MathEdu Intellect. J.*, vol. 4, pp. 1959–1972, 2023.
[63] T. Tang, D. Jiao, T. Chen, and G. Gui, “Medium- and Long-Term Precipitation Forecasting Method Based on Data Augmentation and Machine Learning Algorithms,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 1000–1011, 2022.
[64] X. Lin, J. Fan, Z. J. Hou, and J. Wang, “Machine Learning of Key Variables Impacting Extreme Precipitation in Various Regions of the Contiguous United States,” *J. Adv. Model. Earth Syst.*, vol. 15, p. e2022MS003334, 2023.

[65] D. S. Chadha, R. Chauhan, and N. Gupta, "Flood Prediction and Rainfall Analysis Using Light Gradient Boosted Machine," *NeuroQuantology*, vol. 20, 2022.

[66] L. Pang, Q. Fan, H. Zhang, B. Gong, H. Ye, J. He, J. Wu, H. Li, Q. Zhang, and H. Zhang, "Research on Rainfall and Temperature Forecasting and Correction Technology in Mountainous Microterrain," in *Proc. 13th Annu. Int. Conf. Mater. Sci. Environ. Eng. (MSEE 2025)*, SPIE, Sep. 2025, vol. 13811, pp. 304–312.

[67] T. Tang, Y. Wu, Y. Li, L. Xu, X. Shi, H. Zhao, and G. Gui, "Advanced Machine Learning Ensembles for Improved Precipitation Forecasting: The Modified Stacking Ensemble Strategy in China," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 18, pp. 4242–4254, 2025.

[68] H. Zhuang, F. Lehner, and A. T. DeGaetano, "Improved Diagnosis of Precipitation Type with LightGBM Machine Learning," *J. Appl. Meteorol. Climatol.*, vol. 63, pp. 437–453, 2024.

[69] T. Babu, R. R. Nair, and P. K, "Optimized Rainfall Prediction Model Using Optune CatBoost: An Innovative Approach," in *Proc. 3rd Odisha Int. Conf. Electr. Power Eng., Commun. Comput. Technol. (ODICON)*, Nov. 2024, pp. 1–6.

[70] R. Szczepanek, "Daily Streamflow Forecasting in Mountainous Catchment Using XGBoost, LightGBM and CatBoost," *Hydrology*, vol. 9, p. 226, 2022.

[71] V. Prashanthi, S. Kanakala, D. Borgaonkar, and D. S. Babu, "Rainfall Prediction Using CatBoost Machine Learning Algorithm," in *Proc. Int. Conf. Netw., Multimedia Inf. Technol. (NMITCON)*, Sep. 2023, pp. 1–5.

[72] D. Niu, L. Diao, Z. Zang, H. Che, T. Zhang, and X. Chen, "A Machine-Learning Approach Combining Wavelet Packet Denoising with CatBoost for Weather Forecasting," *Atmosphere*, vol. 12, p. 1618, 2021.