Clean Water and Sanitation
https://www.sciltp.com/journals/cws

*Article*

# Forecasting Hydroclimatic Variables in the Kokcha River Basin Afghanistan: A Time Series Modeling Approach

Abbas Gulistani * and Mehmet Kazim Yetik

Department of Environmental Engineering, Karabuk University, 78050 Karabuk, Turkey
* Correspondence: ab.gulistani22@gmail.com

**How To Cite:** Gulistani, A.; Yetik, M.K. Forecasting Hydroclimatic Variables in the Kokcha River Basin Afghanistan: A Time Series Modeling Approach. *Clean Water and Sanitation* **2026**, *1*(1), 1.

Received: 16 December 2025
Revised: 19 April 2026
Accepted: 28 April 2026
Published: 12 May 2026

**Abstract:** In this study, a statistical modeling approach was developed to monitor climate change over time in the Kokcha River Basin, located in northeastern Afghanistan. Monthly meteorological data from 1979 to 2023 were analyzed using ARIMA and Seasonal ARIMA (SARIMA) models to identify trends and make future projections. The initial phase involved stationarity testing of temperature and precipitation time series, confirming the presence of significant trends. Regression analyses were conducted for 1980–2009 and 2012–2023. Results indicate a noticeable recent increase in average temperature, surpassing 6 °C, and heightened variability in precipitation patterns. The correlation between temperature and precipitation has become more moderate over time. Specifically, the Pearson correlation coefficient declined from −0.5913 (1980–2009), indicating a strong negative relationship, to −0.1593 (2012–2023), which is statistically insignificant. Furthermore, the slope of the regression model decreased from −44.66 to −3.04, suggesting a reduced sensitivity of precipitation to temperature fluctuations, likely due to increasingly complex atmospheric processes. Forecasts generated by SARIMA models for precipitation demonstrated high accuracy, supported by low Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. Projections extending to 2030 suggest a continued rise in temperatures and increased unpredictability in precipitation events. These findings highlight the growing climate risk in the Kokcha River Basin and underscore the urgent need for proactive water resource management. Statistical forecasting methods employed in data-scarce and climate-sensitive regions, such as Afghanistan, can inform policy development and adaptation strategies in similarly vulnerable contexts.

**Keywords:** ARIMA & SARIMA; time series analysis; climate change; Kokcha River Basin; water resources

## 1. Introduction

Studies addressing the impacts of climate change in Afghanistan remain limited. However, available research indicates that parts of the country have experienced significant warming, while precipitation trends remain uncertain [1,2].

Aich et al. conducted one of the first comprehensive analyses of Afghanistan's climate, reporting an average temperature increase of 1.8 °C between 1951 and 2010. This warming has impacted key sectors, including water resources, agriculture, and energy, leading to droughts, floods, and land degradation. Given that over 80% of the population depends directly on natural resources particularly river basins these changes have serious socio-economic implications [3].

Climate change has also altered hydrological processes in major basins. Studies report glacier retreat and permafrost loss in the Hindu Kush and surrounding regions, affecting river flows in the Amu Darya system, including the Kokcha Basin [4,5].

Environmental degradation, driven by prolonged conflict and recurring drought, has further reduced resilience and water availability [6,7]. Findings were reported by Shams and Alawi, confirming a consistent warming pattern but uncertain rainfall changes in northeastern Afghanistan [8,9]. Basin-scale analyses consistently show rising temperatures but weak or insignificant precipitation trends. For example, Akhundzadah reported warming of +1.46 °C across Afghanistan and +1.20 °C in the Amu Darya Basin (1980–2022) .with no significant precipitation trend [10].

The Kokcha River Basin is highly vulnerable to hydroclimatic extremes, including seasonal floods, prolonged droughts, and accelerated glacier melting. Snowmelt-driven hydrology makes the basin particularly sensitive to rising temperatures, which alter runoff timing and reduce natural water storage. Recent flood events highlight these vulnerabilities. In May 2024, severe flooding in northeastern Afghanistan affected more than 60,000 people across Badakhshan, Baghlan, and Takhar provinces. Thousands of homes were damaged, infrastructure was disrupted, and significant casualties were reported MoRR and OCHA, 2024. increasing flood frequency, while declining snowpack and glacier retreat threaten long-term water availability. Despite these risks, basin-specific hydroclimatic assessments remain limited, especially those integrating both statistical modeling and recent observational data [11,12].

These climatic changes have important implications for water resources. Afghanistan relies heavily on snowpack and glaciers above 2000 m, which act as natural storage systems and regulate seasonal water availability [7]. Rising temperatures are accelerating glacier melt, reducing long-term water storage, while increasing variability in river discharge [4]. This leads to increased risks of water scarcity, flash floods, and soil erosion, particularly in downstream communities.

These events underscore the urgent need for improved water resource management and climate adaptation strategies. Despite these challenges, hydrometeorological data in Afghanistan remain limited due to long-term political instability, resulting in significant data gaps [9].

The Kokcha River Basin, a major tributary of the Amu Darya, is strategically important for Afghanistan's water management. However, no comprehensive study has yet examined the combined effects of hydroclimatic variables in this basin using all available observational data. This study addresses this gap by analyzing long-term changes in temperature, precipitation, and river discharge from 1970 to 2023 from the Near Keshem station, despite gaps in 2010–2011. To achieve this, the study applies time series analysis and forecasting models, including ARIMA and SARIMA, to evaluate trends and predict future climate patterns. By integrating these approaches, the research aims to improve understanding of climate impacts on water resources, agriculture, and livelihoods, and to support integrated water resources management (IWRM) and climate adaptation strategies in northeastern Afghanistan.

This study aims to analyze long-term hydroclimatic variability and forecast future trends in the Kokcha River Basin using time series modeling. The key research questions are:

(1) How have temperature and precipitation patterns changed over time.
(2) Has the relationship between these variables shifted in recent decades.
(3) Can ARIMA/SARIMA models reliably forecast future hydroclimatic conditions.

The novelty of this study lies in combining long-term observational data with statistical forecasting in a data-scarce basin, while also examining temporal shifts in climate relationships, which are rarely explored in Afghanistan.

## 2. Methodology Approach

### *2.1. Study Area and Gauging Stations*

Afghanistan is a landlocked country situated in South and Central Asia [13]. Its neighbours include China in the far northeast, Turkmenistan, Uzbekistan, and Tajikistan in the north, Iran in the west, and Pakistan in the east and south. It is the 41st largest country in the world, with a total area of 652,000 km$^2$ and a population of about 31 million [10,11]. The Himalayas, Pamirs, and Hindukush, which rise to elevations exceeding 7500 m, are among the world's tallest mountain ranges in Afghanistan's untamed landscape. Approximately 75% of Afghanistan's land is covered by mountains, running in a northeast-southwest direction [12,13].

The Kokcha River Basin is being studied in northeastern Afghanistan between latitudes 35°26′35″ N and 37°28′14″ E and longitudes 68°32′18″ E to 71°38′12″ E, covering an area of approximately 20,600 km$^2$ (as seen in Figure 1). Most of the water in the wide Kokcha River comes from snowmelt in the spring and summer [3,9].

In the lower reaches of the river, the freshwater required by farmers during the spring and summer seasons primarily originates from melting snow in the surrounding mountains. Therefore, monitoring the regional water balance is essential, and this process should begin with an assessment of snow cover statistics within the basin [3,9]. Mountain glaciers in the region contribute not only to the local terrain but also feed the rivers, thereby playing a crucial role in supplying water to the local population.

Afghanistan is characterized by substantial climatic variability over relatively small geographic areas. The Kokcha Basin, where this study is conducted, is situated within the Hindu Kush Mountains. Between 1951 and 2010, the national average temperature in Afghanistan increased by approximately 1.8 °C, with the highest temperature rise observed in the eastern regions (2.4 °C), and the lowest in the Hindu Kush region (0.6 °C) [12]. Within the Panj-Amu Darya Basin, the Kokcha River Basin represents a critical hydrological sub-basin [14]. This basin must supply water for forestry, industry, energy production, agricultural irrigation, domestic use, rural communities, and towns [15]. The Kokcha River Basin has a complex geomorphology shaped by tectonic activity, erosion, and climate. Elevation ranges from about 818 m to over 5647 m, creating steep terrain in the central and eastern areas, with high erosion and runoff risk, while gentler western and northern areas are more suitable for agriculture. The basin shows a dendritic drainage pattern, indicating geological uniformity. Curvature and contour analyses highlight zones of erosion, deposition, and a highly dissected landscape, as shown in Figure 2. Overall, steep relief and active processes strongly influence hydrology and land use.

Temperature and precipitation data collected from 1979 to 2009 and from 2012 to 2023 have been examined using observational records from measurement stations within the Kokcha River Basin. Due to political instability in the country, there is a data gap for the period between 2009 and 2012. Despite certain exceptions, recent data for the period 2012–2023 are available. This study utilizes data from stations that have been monitoring the regional climate for at least 39 years (statistical Data are presented in Table 1). Meteorological and hydrological information is provided by the Afghanistan Meteorological Department and the Ministry of Energy and Water (MEW).

The division into two periods (1980–2009 and 2012–2023) is primarily due to data unavailability during 2010–2011. In addition, this separation allows investigation of potential nonstationarity and recent climate shifts. To support this, trend consistency and structural differences between the two periods were evaluated.

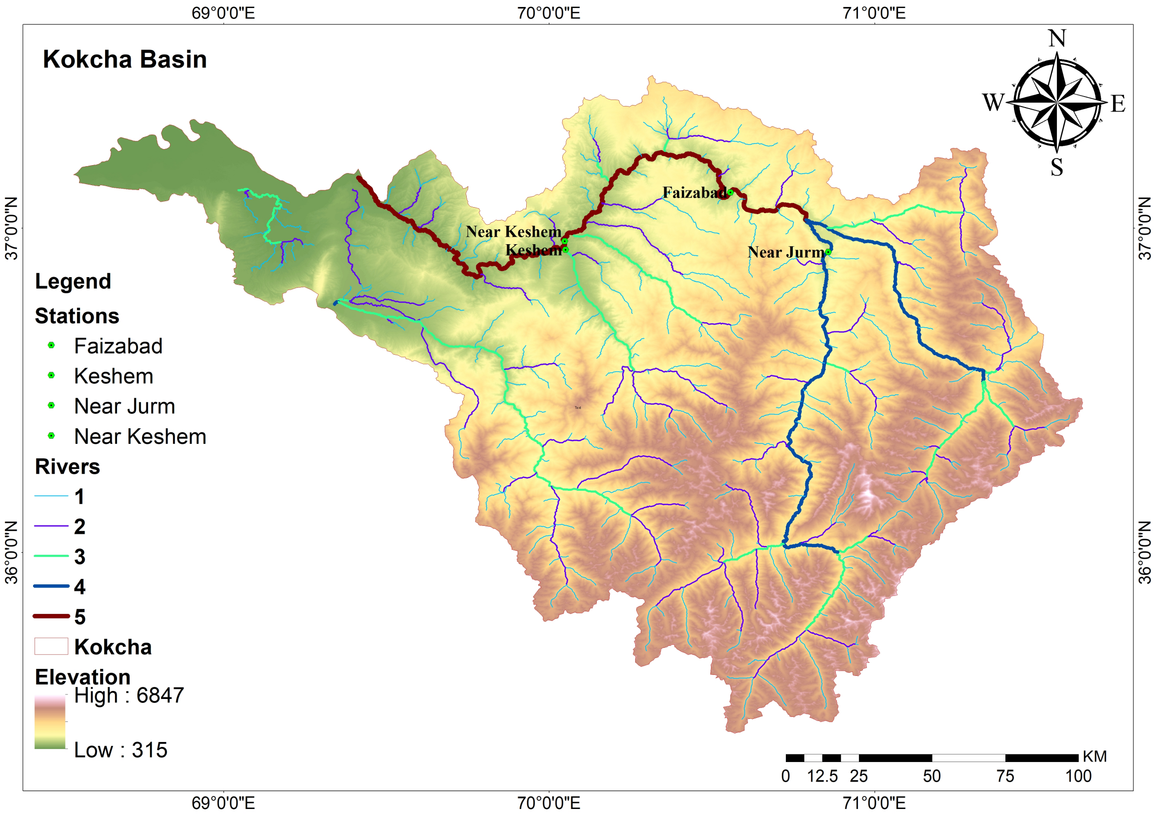



**Figure 1.** Study gauging station is on the Kokcha River Basin.

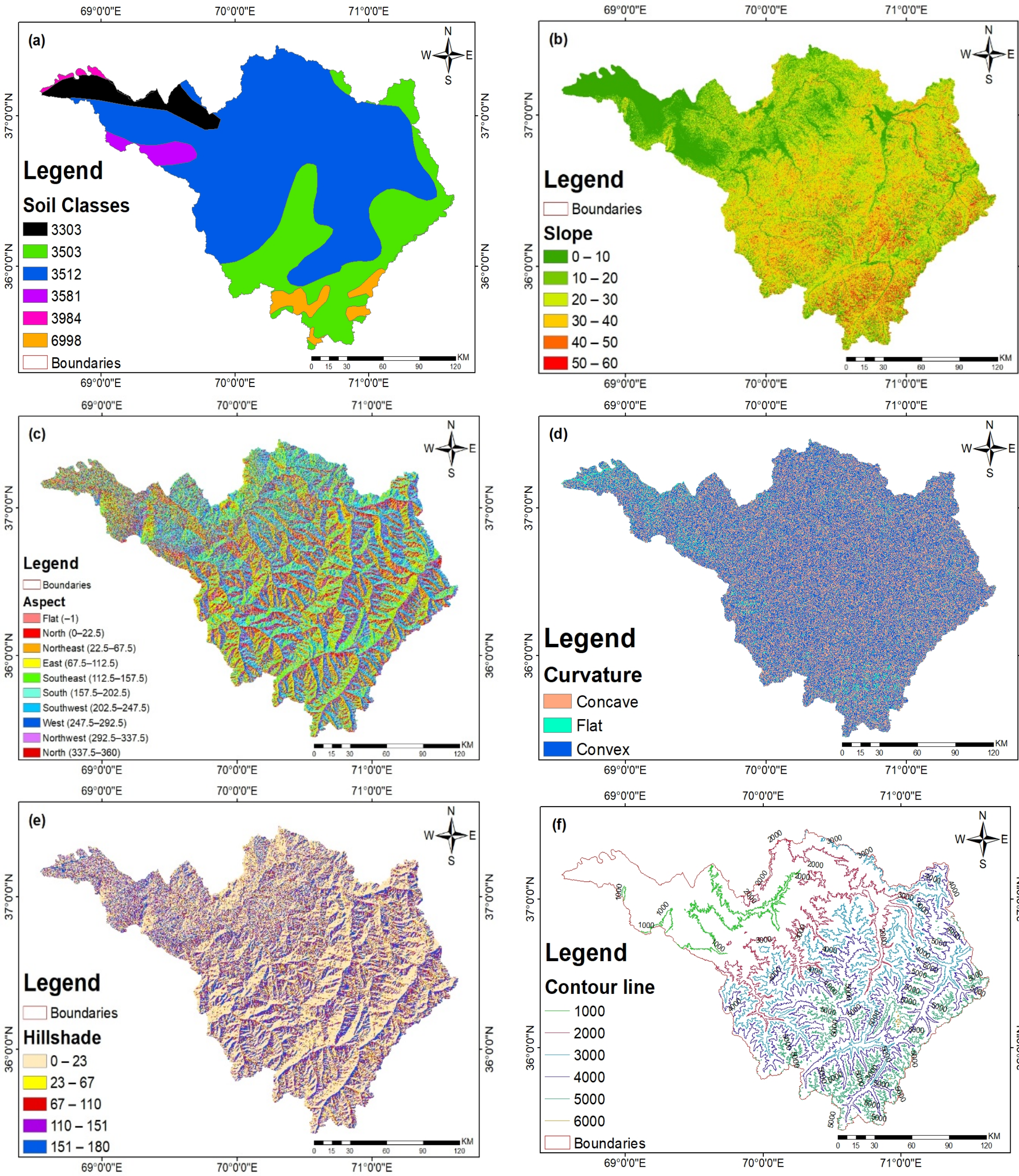



**Figure 2.** Geomorphological characteristics maps of the Kokcha River Basin: (**a**) Soil Class; (**b**) Slope; (**c**) Aspect; (**d**) Curvature; (**e**) Hill Shade; (**f**) Contour Line.

**Table 1.** Statistics of Rainfall and Temperature in the Kokcha River Basin years(1979–2023).

| Parameters | Observation | Unit | Min | Max | Mean | Standard Deviation |
|---|---|---|---|---|---|---|
| Rainfall | 43 | mm | 147.050 | 436.380 | 289.654 | 68.473 |
| Temperature | 42 | °C | 7.902 | 19.100 | 12.128 | 3.629 |

This reduction emphasizes the effects of climate change on cryospheric resources, which have direct implications for river discharge, downstream water availability, and long-term water resource sustainability, and the comparative changes in land cover categories within the Kokcha River Basin from 2017 to 2024. Figure 3 shows a significant drop in snow and ice cover, as well as a noticeable increase in rangeland area. Urban areas and croplands both show slight increases, whereas water, trees, vegetation, and barren ground show only tiny shifts.

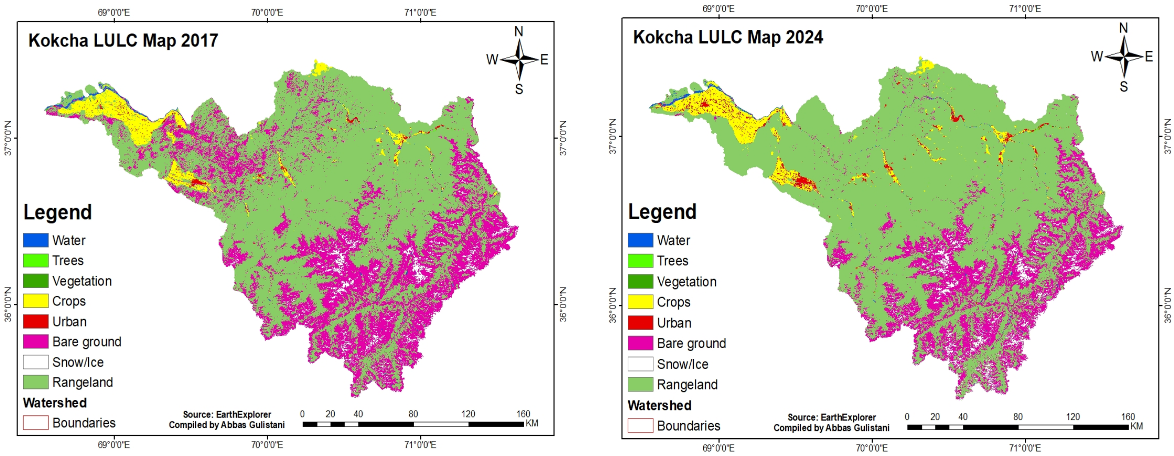



**Figure 3.** Land cover classes of Kokcha River Basin.

### *2.2. Stationarity Test*

Time series analysis may reveal the trends in time, seasonal variation and repetitions in climate data. The attention to ARIMA and SARIMA models in this inquiry was pegged back due to their reported capacity to handle trend lines and seasonality in time series records. SARIMA works best with datasets that have significant seasonal components, which makes it especially applicable for studying monthly climate data with recurrent annual trends, while ARIMA works best with non-seasonal data [16,17].

For ARIMA modeling to be effective, the time series data must be stationary. If the series is not stationary, differencing is required prior to ARIMA modeling to transform it into a stationary series [14]. The stationarity of a time series can be examined using the Augmented Dickey-Fuller (ADF) test. When the *p*-value obtained from the ADF test is less than 5% ($p < 0.05$), it indicates that the differenced data are stationary.

The ADF test is based on the following regression equation:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_p \Delta y_{t-p} + \varepsilon_t$$

In this equation:

- $\Delta y_t$ denotes the difference between consecutive observations in the series,
- $t$ represents the time trend (optional),
- $\gamma$ is the coefficient of interest (if $\gamma = 0$ the series contains a unit root and is non-stationary),
- $p$ indicates the number of lagged difference terms included in the test

Based on this framework, a *p*-value less than 0.05 following differencing confirms the stationarity of the series and its suitability for ARIMA modeling [15,18].

### *2.3. ARIMA Model*

The process for applying any of the four stochastic models (AR, MA, ARMA, or ARIMA) to time series data is substantially the same. This study adopts the stochastic modeling approach specifically through the ARIMA model, as it incorporates the fundamental components of the other models (AR, MA, and ARMA). Furthermore, the ARIMA(p, d, q) model can be easily reduced to AR, MA, or ARMA models by appropriately adjusting its parameters.

The stochastic modeling of time series typically involves four essential steps [18,19].

(1) Model identification or selection
(2) Estimating model parameters
(3) Model evaluation and diagnostic check
(4) Forecasting

Before applying stochastic modeling, it is necessary to ensure that the time series is stationary, detrended, and adjusted for seasonality.

### *2.4. SARIMA Model*

ARIMA (Autoregressive Integrated Moving Average) is a commonly used model for univariate time forecasting. The seasonal patterns can be identified in the data so that such data is quite often represented with the help of the Seasonal ARIMA (SARIMA) model that captures both trend and seasonal representation. In the event of developing a SARIMA model, the seasonality of the series will actually be considered. Every SARIMA (p, d, q) (P, D, Q) model has a set of parameters containing non-seasonal and seasonal aspects [19,20].

$$\varphi P(B_s)\varphi p(1-B)(1-B_s)\Delta yt = \theta(B)\theta(B_s)\varepsilon_t$$

Where;

$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$ (The order $p$ of AR term)

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ (The order $q$ of the MA term)

$\varphi(B^s) = 1 - \varphi_1 B^s - \varphi_2 B^{2s} - \cdots - \varphi_P B^{Ps}$ (The order $P$ of the seasonal AR term)

$\theta(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_Q B^{QS}$ (Seasonal MA term)

where $yt$ the original time series at time $t$, $\varepsilon_t$ white noise (error term at time $t$), $s$ length of the seasonal period (for monthly data with annual seasonality) [20–23].

ACF and PACF: The data must be analysed to determine the appropriate values for the SARIMA parameters. The proper value of $q$ in the Moving Average feature can be found using the Autocorrelation function. The Partial Autocorrelation function can extract the suitable $p$ value from the Autoregressive component [5,24].

### *2.5. Regression Model Development and Performance Assessment*

Forecasting stage involves inputting the computed parameters in order to generate future values of time series along with confidence interval in predictions once the selected ARIMA model has passed the analysis stage. These predictions usually extend more than the original data. In case the model was used on a converted or differenced series, the forecasts need to be re-integrated in order to give the forecasts in the original scale of the data. Residual(s) The opposite of differencing, integration ensures that the forecasts are in accord with the input series of time. Reintegration can also be depicted by the letter (I) in the acronym used to describe ARIMA (Autoregressive Integrated Moving Average) [24,25]. SARIMA model is the extension of the ARIMA model constructed on the time series data having seasonality. It is composed of two terms as the non-seasonal parameters of ARIMA (p, d, q) and seasonal parameters (P, D, Q, s) [23,26,27].

### *2.6. Software and Tools*

The data analysis and modeling process was conducted using the Python programming language, with a particular emphasis on the statsmodels library and ARIMA and SARIMA models. These advanced time series analysis tools were employed to understand the structure of the data, perform stationarity tests, and develop appropriate forecasting models. In addition, Excel and MATLAB were utilized for statistical analyses and for plotting regression lines, which allowed for a more detailed and comparative examination of climate trends. The integration of these software platforms enhanced model accuracy and reinforced the consistency of the results obtained through various analytical approaches. Consequently, the use of multiple tools made it easier to analyze the dynamics of climate change and to generate reliable forecasts, thereby ensuring that the study remains both scientifically sound and transparent.

## 3. Results and Discussion

The results of time series and statistical analyses utilized to anticipate hydroclimatic variables and assess climate changes in the Kokcha River Basin are shown in this section. The findings, which show changes in the relationship between temperature and precipitation across two different time periods (1980–2009 and 2012–2023), include regression and correlation studies. In order to provide insight into the accuracy and dependability of climate forecasts in the face of data shortages and growing variability, the performance of the ARIMA and SARIMA models is also evaluated. These findings are explored in the context of climate adaptation and water resource management in a vulnerable and data-constrained region. While SARIMA models provide reliable short-term forecasts, they do not explicitly incorporate physical climate drivers such as snowmelt, evapotranspiration, or atmospheric circulation. Future research should integrate downscaled climate model outputs (e.g., GCM/RCM) with bias correction techniques to improve long-term projections and better capture hydroclimatic processes.

### *3.1. Statistical Analysis of Precipitation Based on Temperature*

Linear regression analysis was performed to examine the impact of temperature on annual precipitation patterns for two distinct time periods, 1980–2009 and 2012–2023 Table 2. A straightforward linear model was employed to evaluate the model's statistical performance on both historical and more recent datasets, aiming to determine whether air temperature (T) could accurately predict annual precipitation (P). Fit a linear regression model:

$$P = aT + bP = aT + b$$

- Predict annual precipitation values based on the observed temperature data.
- Calculate the coefficient of determination ($R^2$) and root mean square error (RMSE) to assess model accuracy.
- Plot observed vs. predicted precipitation as a time series for visual inspection.

**Table 2.** Linear regression performance of precipitation (1980–2009 and 2012–2023).

| Period | $R^2$ | RMSE (mm) | Regression Equation |
|---|---|---|---|
| 1980–2009 | 0.349 | ~59.8 | $P = -44.66 \times T + 726.92$ |
| 2012–2023 | 0.025 | ~67.3 | $P = -3.04 \times T + 326.89$ |

#### 3.1.1. Descriptive Analysis of Climate Variables

Data on yearly and monthly average temperatures and precipitation were examined for the Kokcha River Basin during 1980–2009 and 2012–2023, but no data were available for 2010 and 2011.

- The annual temperatures in 1980–2009 were between 7.9 °C and 11.2 °C, and precipitation ranged between 147 mm and 436 mm.
- Between 2012 and 2023, the temperature went up considerably, from 15.2 °C to 19.1 °C, which shows a clear rise in recent years. The quantity of precipitation during this period changed, from 193 mm to 432 mm.

#### 3.1.2. Predicting Annual Temperature Based on Precipitation

To further explore the climate dynamics within the Kokcha River Basin, a reverse regression analysis was performed to predict mean annual air temperature (T) based on annual precipitation (P). The analysis was divided into two temporal phases(1980–2009) and (2012–2023) to assess whether precipitation can serve as a reliable predictor of temperature trends across different climate periods (Table 3). It means temperature is now the dependent variable, and precipitation is the predictor.

- Linear regression model of the form: $T = a \cdot P + bT = aP + b$
- Calculation of the coefficient of determination ($R^2$) to evaluate model performance.
- Computation of the Root Mean Square Error (RMSE) to assess prediction accuracy.
- Visualization of observed vs. predicted temperature using time series plots.

**Table 3.** Linear regression performance of temperature (1980–2009 and 2012–2023).

| Period | $R^2$ | RMSE (°C) | Regression Equation |
|---|---|---|---|
| 1980–2009 | 0.208 | ~0.73 | $T = -0.0066 \times P + 12.053$ |
| 2012–2023 | 0.010 | ~1.02 | $T = -0.0017 \times P + 17.164$ |

### *3.2. Correlation and Regression Analysis*

Both correlation and regression analyses were conducted to examine the relationship between the variables (Figure 4). Pearson's correlation coefficient helped determine the type and intensity of the linear relationship, and simple linear regression was used to estimate the impact of one variable on another. The findings, as shown in Table 4, indicate the degree to which changes in the independent variable can explain changes in the dependent variable. Correlation coefficients within the range $-0.3 \leq R \leq 0.3$ are considered indicative of weak or no linear association. However, this interpretation applies only to linear relationships, and variables may still exhibit non-linear dependence not captured by the Pearson correlation coefficient [5,23].

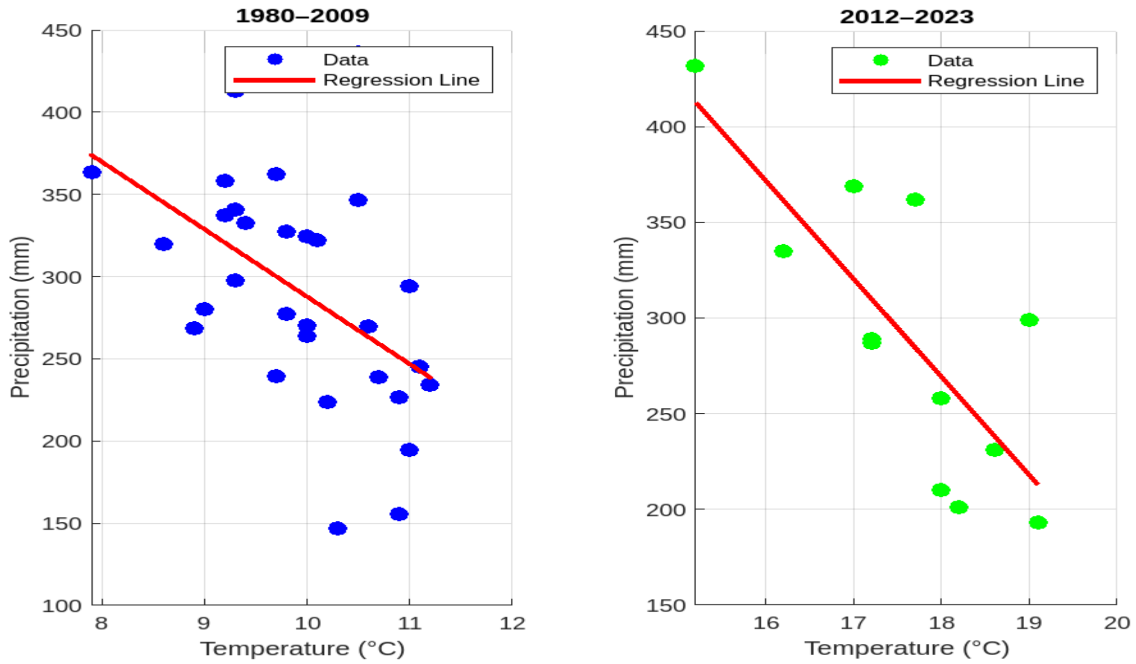



**Figure 4.** Regression analysis period 1980–2009 and 2012–2023.

**Table 4.** Relationship between correlation and regression analysis.

| Period | 1980–2009 | 2012–2023 |
|---|---|---|
| Correlation Coefficient (R) | −0.5913 | −0.1593 |
| Regression Equation | P = −44.6579 × T +726.9152 | P = −3.0427 × T + 326.8915 |
| *p*-value | 0.0007 | 0.6220 |
| Correlation | Stronger Negative | Much Weaker |

3.2.1. Predicting Precipitation from Temperature

Keeping the temperature values, the anticipated precipitation was calculated by applying the regression-based linear models. In the early period, the model performed admirably (Figure 5).

- With the earlier model, about 35% of the changes in precipitation could be explained.
- Between 2012 and 2023, small amounts of solar radiation reached the Earth which may be because precipitation was becoming more erratic.

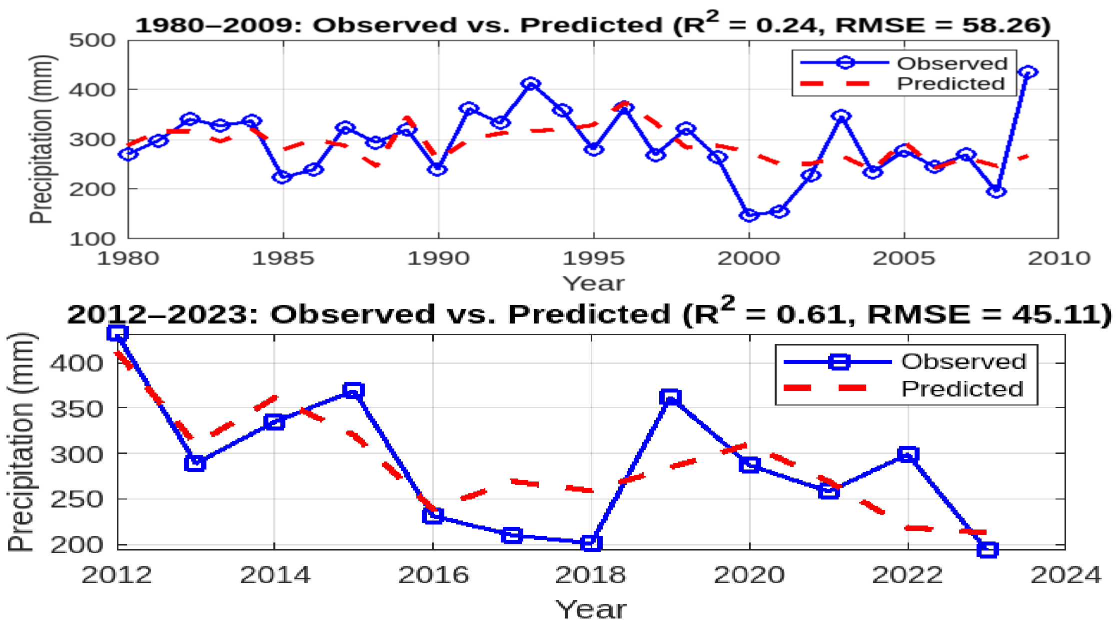



**Figure 5.** Observed and predicted Rainfall based on Temperature between 1980–2009 and 2012–2023.

3.2.2. Predicting Temperature from Precipitation

The temperature was calculated from the precipitation numbers using a reverse regression. (Figure 6) The regression model is indicated:

- During 1980–2009, 21% of the ups and downs in temperature were caused by changes in precipitation which indicates a gentle linear pattern.
- The 2012–2023 period showed poor predictive power, much like the earlier result.

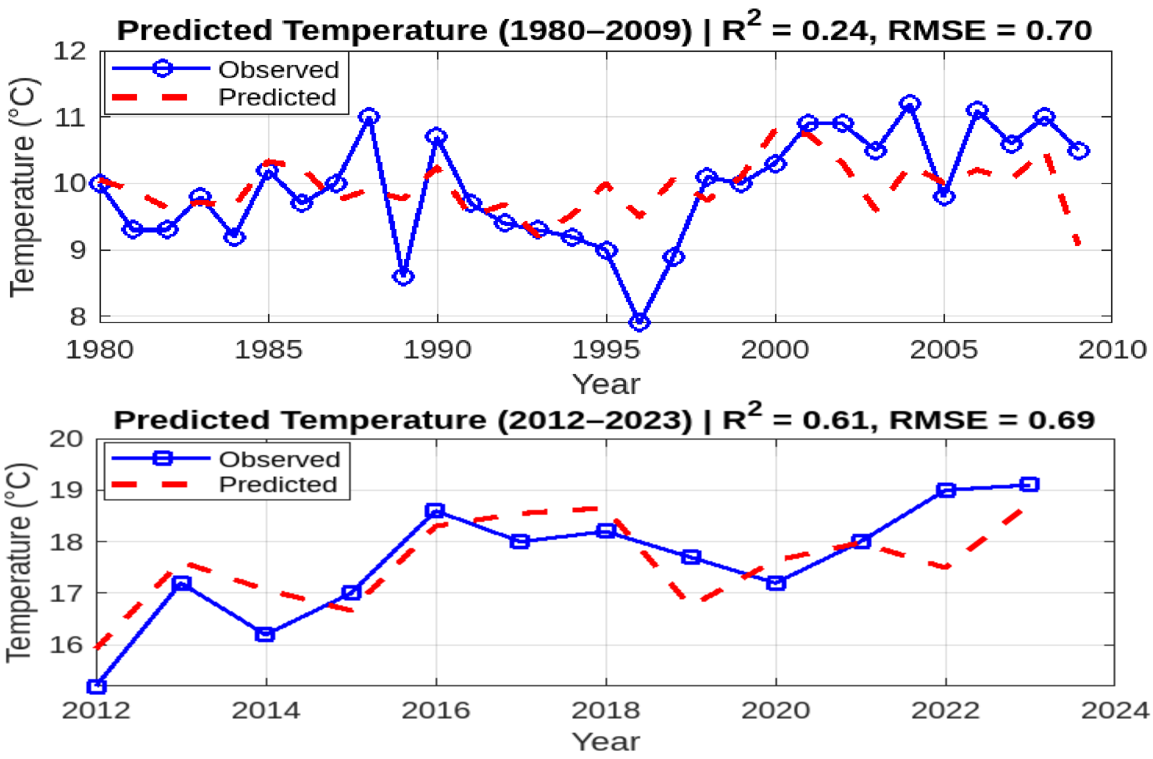



**Figure 6.** Observed and predicted Temperature based on Rainfall between 1980–2009 and 2012–2023.

### *3.3. Visual Comparison*

Time series plots of observed vs. predicted values illustrate:

(1) Observed and predicted precipitation or temperature were in good agreement during the years 1980–2009.
(2) The years 2012–2023 have seen big differences between the results, pointing to lower reliability of linear models with current climate change.

### *3.4. Interpretation and Implications*

Results show there is a change in the climate in the Kokcha River Basin. After 2012, the recorded temperatures went up significantly, at least 6 °C above the baseline period. There is evidence that precipitation happens less steadily these days, which may have to do with changes in air patterns, the melting of Arctic glaciers, or people changing land use and the cover of plants. Recent years have shown that models should consider more factors, for example, humidity, pressure, evapotranspiration, and land surface changes. It is confirmed by the analysis that the historic link between temperature and precipitation is not as strong as it used to be these days. The weakening relationship between temperature and precipitation may be linked to changing atmospheric circulation patterns, reduced snowpack contribution, and increased evapotranspiration under warming conditions. These processes reduce the direct coupling between temperature and precipitation, leading to greater climate variability in the basin.

It suggests the climate in the area is becoming more variable, which makes basic regression models less appropriate for forecasting the climate (Figure 7). The use of machine learning and similar techniques, as well as modeling of time-series data and climates, might give better information about future water conditions.

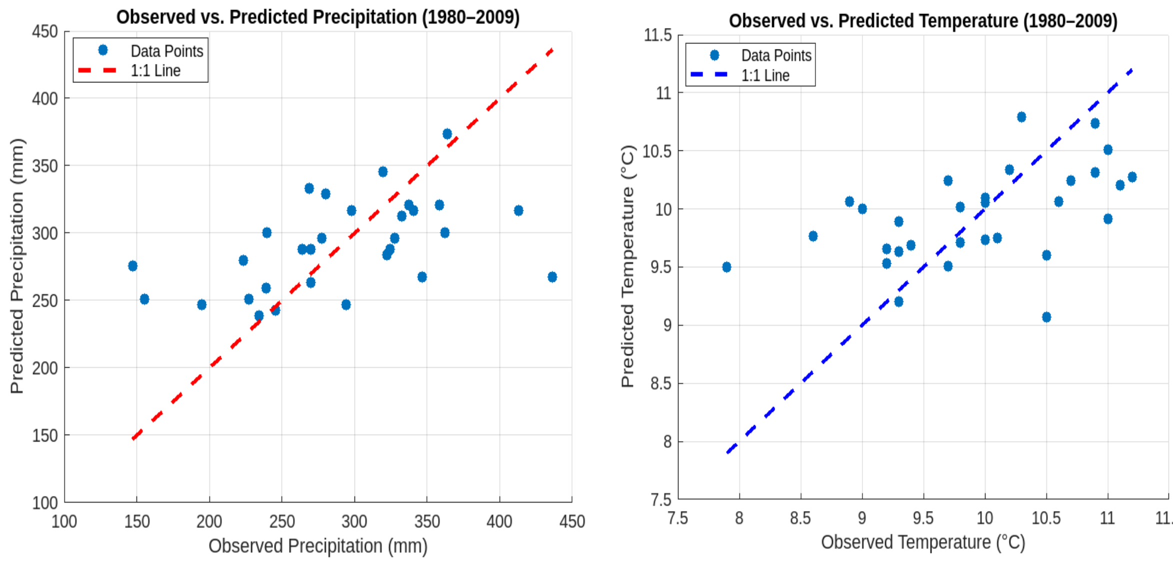



**Figure 7.** Observed and predicted relationship between 1980–2009 and 2012–2023.

### *3.5. ARIMA and SARIMA Model Forecasting*

The SARIMA (2,0,1)(0,1,1)12 model was successfully fitted to the seasonally differenced time series. The model presented a high degree of prediction accuracy, as evidenced by the RMSE of 1.403 and the MAPE of 5.738% on the original scale. Model selection criteria like AIC (95.636) and BIC (102.449) were used to assess the model's adequacy. Although the MAPE on the differenced series was high (378.767%), this is mainly due to small values in the differenced data and does not accurately reflect the overall quality of the model.

The data was fitted to another seasonal ARIMA model of order (0,1,1)(0,1,0)12 to account for both seasonal and non-seasonal variance. On the original scale, the goodness-of-fit statistics reveal a mean absolute percentage error (MAPE) of roughly 20.8%, displaying a reasonable level of forecast accuracy. The residuals' variability was indicated by the root mean square error (RMSE), which was 104.32. For model comparison, the information criteria values (AIC = 283.43, BIC = 285.70) were utilized. Based on residual analysis, the model's assumptions were judged sufficient.

### *3.6. Model Identification*

The original data were first subjected to an augmented Dickey-Fuller (ADF) test to establish stationarity using the unit root and stationary tests; if the results revealed a non-stationary series, stationarity would be established by a differentiated transformation. to ascertain the initial values for the seasonal parameters P, D, and Q, as well as the autoregressive order p, the order of differencing d, and the moving average order q. The Inverse Autocorrelation Function (IACF), Partial Autocorrelation Function (PACF), and Autocorrelation Function (ACF) are the three key diagnostic tools utilized in time series analysis. These functions are presented for precipitation in Figure 8 and for temperature in the context of SARIMA model evaluation in Figure 9.

The ACF in a time series with lag q checks how much the observed values depend linearly on each other. The PACF helps figure out how many auto-regressive terms (p) are required. When there are obvious differences in the data, the IACF is like an ACF from a non-stationary process. The Likelihood and Akaike Information Criterion (AIC) helped identify the top proposal models, and the model with the lowest AIC was chosen as the best. Mathematically, the Akaike Information Criterion (AIC) is written as:

$AIC = -2\log L + 2m$ (number of model parameters), where m equals the sum of p, q, P, and Q, which is the number of terms, and L represents the likelihood function used in ARIMA. It is a formula based on models, and its value keeps dropping as the sum of squared residuals goes down. To understand SBC, we must know its mathematical formulation, which is $SBC\ -2\log L + m\ln(n)$. The letter n refers to the number of observations.

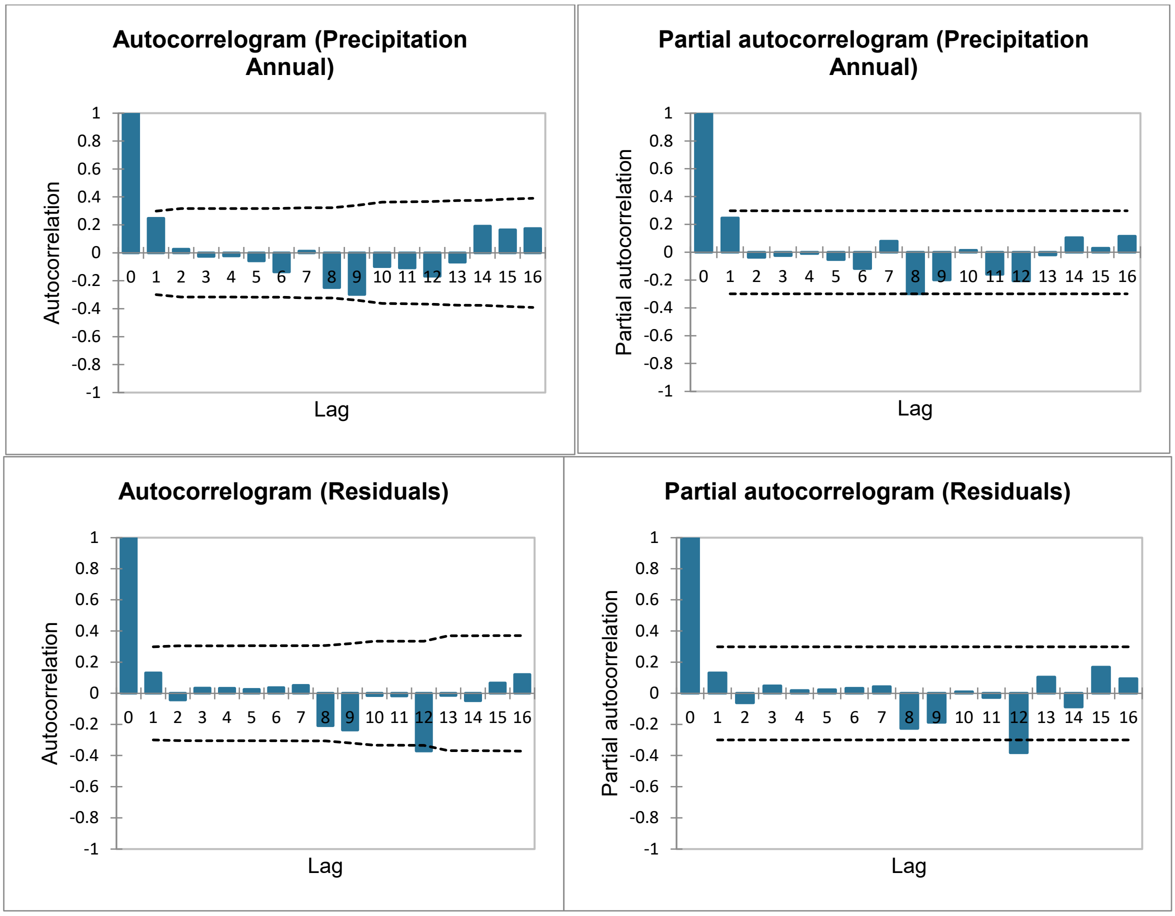



**Figure 8.** Autocorrelogram and Partial Autocorrelogram of SARIMA models for Precipitation.

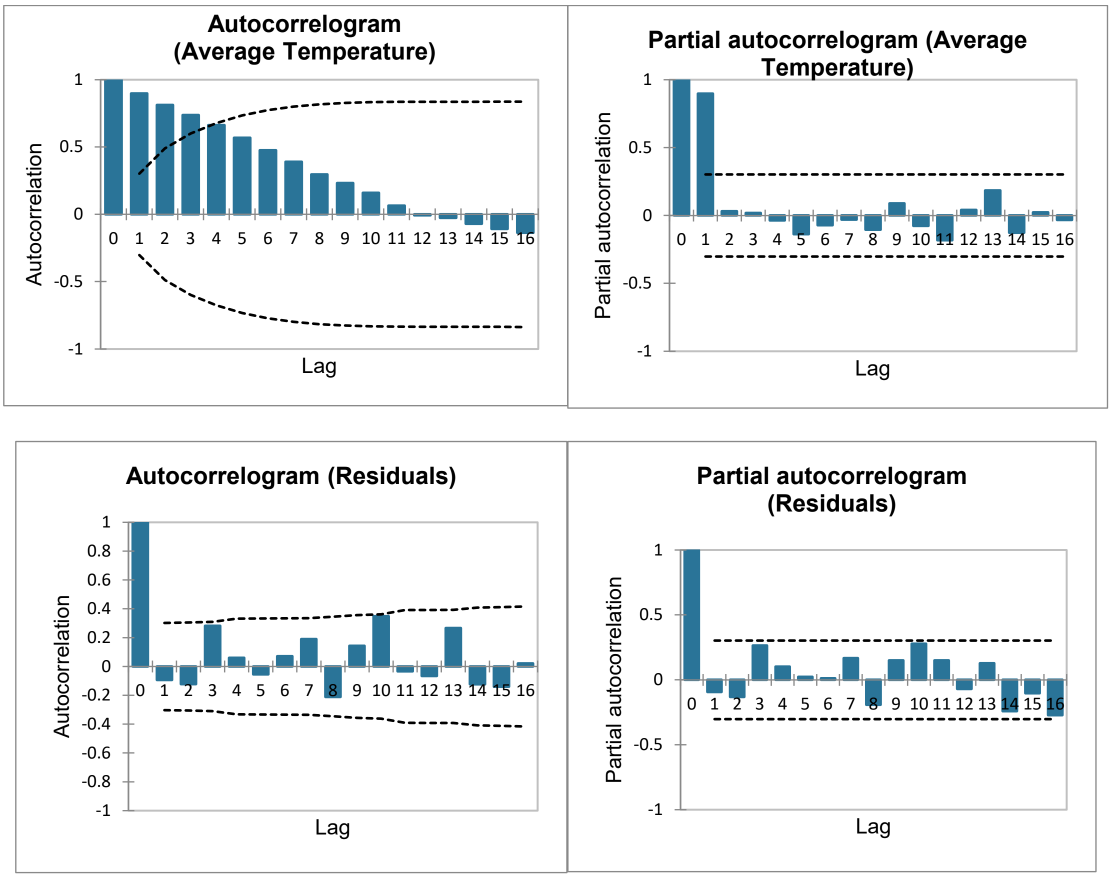



**Figure 9.** Autocorrelogram and Partial Autocorrelogram of SARIMA models for Temperature.

*3.7. Stationarity Analysis of Rainfall and Temperature Series*

The initial rainfall and temperature data had regularities and significant autocorrelation, showing the data was non-stationary. Both datasets were processed using first-order differencing to try to make the trend data stationary.

As illustrated in Figure 10, second-order differencing enhanced the stability of the rainfall series, while first-order differencing significantly reduced the trend. Furthermore, the application of seasonal differencing smoothed the data and improved stationarity by filtering out annually recurring trends. Similarly, for the temperature series, first-order differencing proved beneficial, as indicated by lower autocorrelation values at key lags observed after differencing. To enable the application of the ARIMA model, a stationary series was obtained through a combination of first order and annual seasonal differencing, as shown in Figure 11.

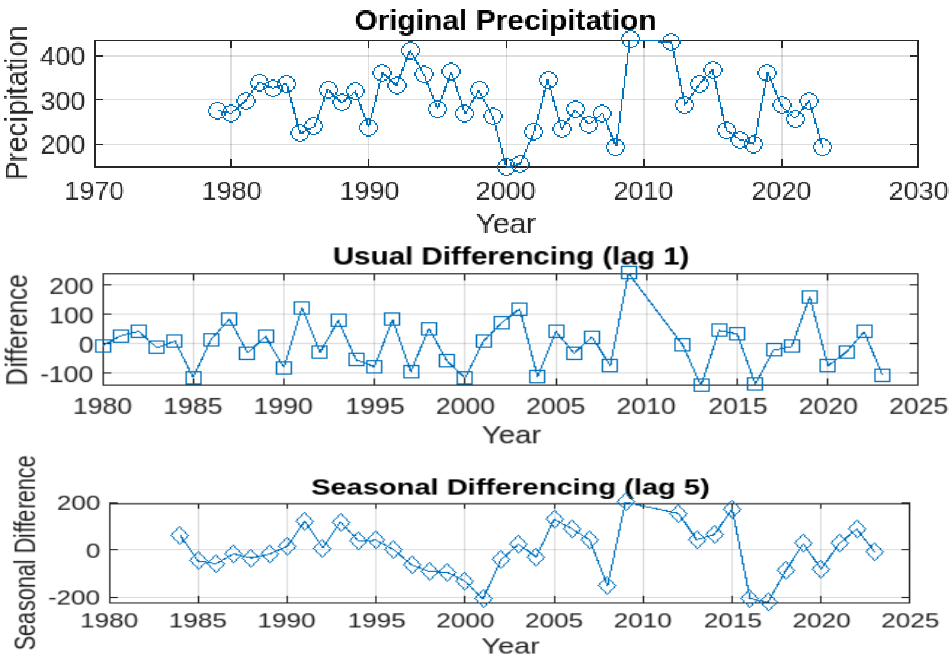



**Figure 10.** Rainfall series.

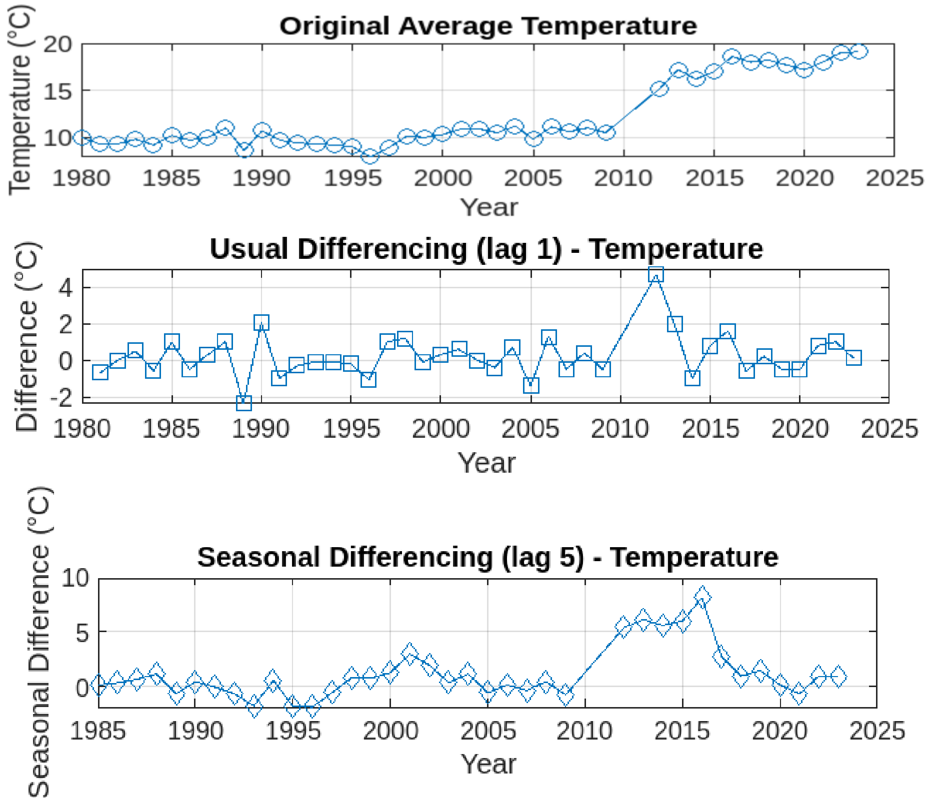



**Figure 11.** Temperature series.

*3.8. Diagnostic Model for Precipitation and Temperature Forecasts*

The diagnostic results for rainfall are shown in Figure 12, and temperature results are in Figure 13 as correlograms, histogram density, Q–Q plots, and normalized residuals. According to the Q–Q plots, the residuals follow a normal distribution, showing that the models capture the right pattern in the data. Because the residuals had little autocorrelation, it was confirmed that the model was capable of predicting temperature and rainfall.

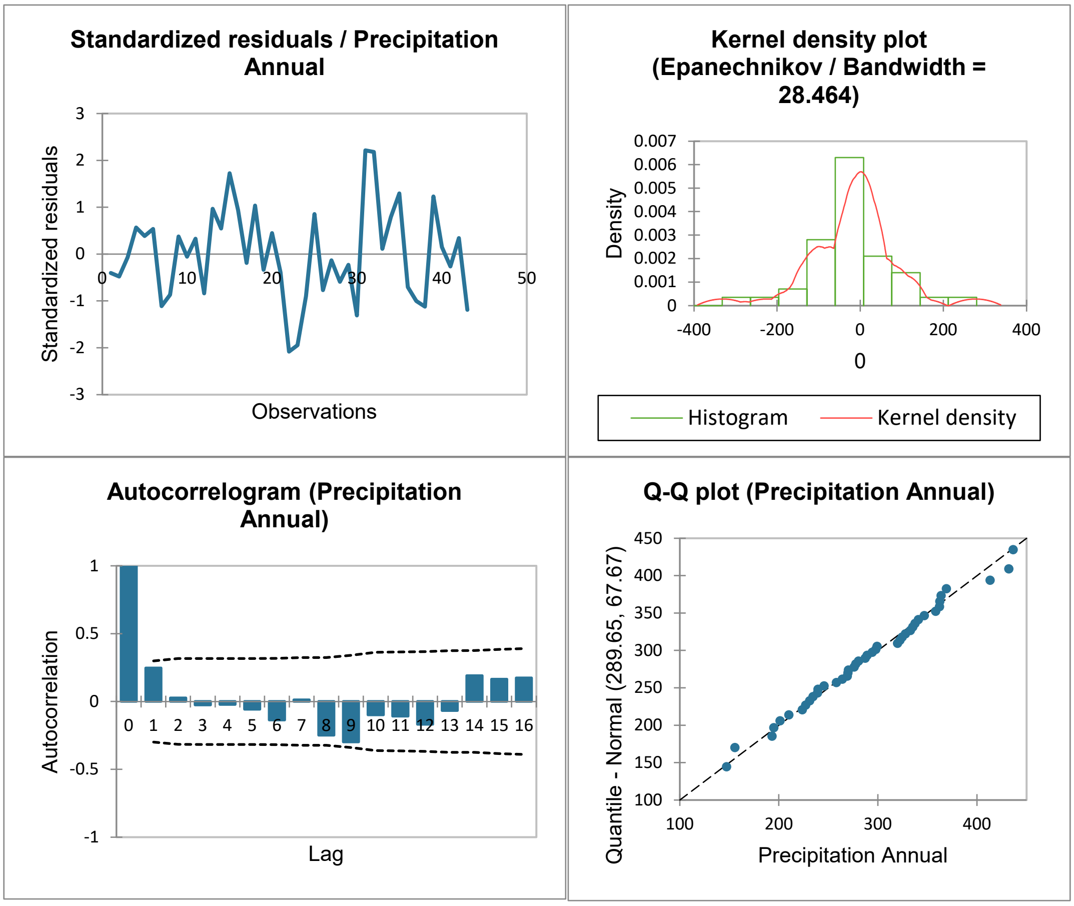
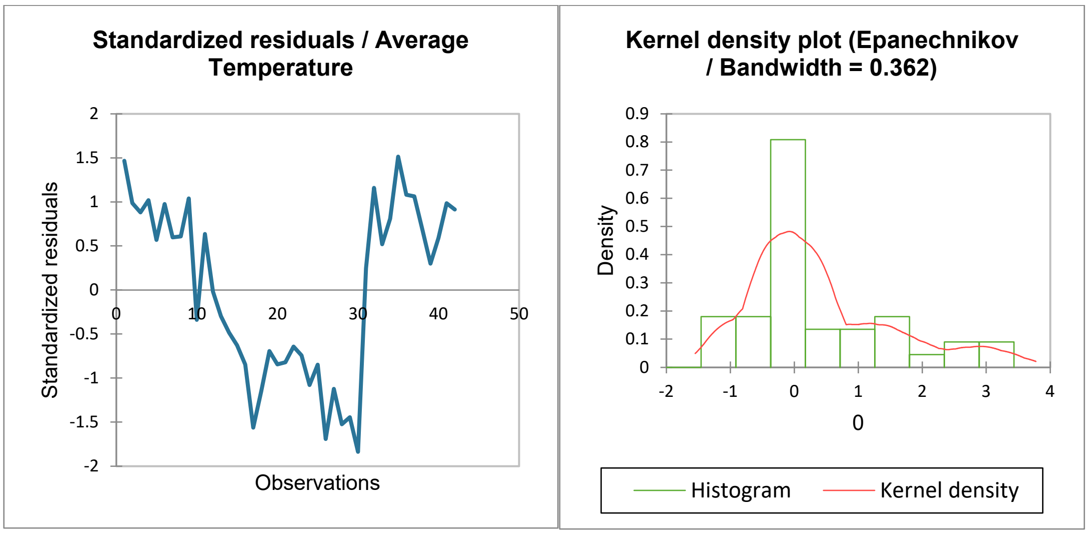



**Figure 12.** Rainfall diagnostics model.



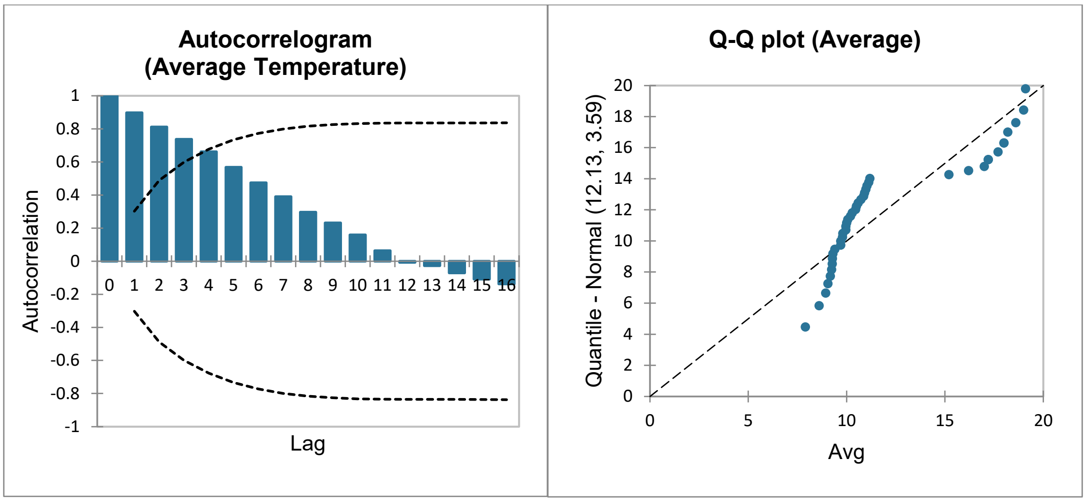



**Figure 13.** Temperature diagnostics model.

## 4. Forecasting and Projections

The annual precipitation trends and forecasts produced by a SARIMAX(0,1,1)(0,1,0){12} model are shown in Figure 14. The model was fitted to historical precipitation data (1979–2023), and the forecasts include the years 2024–2030. The prediction interval (95% confidence bounds) shows that future predictions will be more unclear, demonstrating how well the model captures both trend and seasonality components. The model's predictive ability appears to be sufficient for short- to medium-term forecasting, as seen by the pretty good alignment between the validation data (gray squares) and the predicted values (red dots).

It is clear that the uncertainty and variation in long-term precipitation forecasts are growing because the ranges have become wider. Figure 15 shows predictions for average, minimum and maximum temperatures made by SARIMAX models up to 2030. Forecasts made by each model cover a 95% confidence range and also mimic developments seen in the past. Projections show that both the average and the extreme temperatures are on a general rising trend. With increasing uncertainty, the intervals becoming wider, forecaster results usually comply with what is predicted.

The SARIMAX model is used to long-term forecast average temperature and monthly precipitation as seen in Figure 16. The upper panel shows there are clear seasonal trends and moderate changes in monthly precipitation, modeled by SARIMAX(2,0,2)(1,1,1){12}. According to a SARIMAX(2,1,2)(0,1,1){12} model, the figure below reveals that the average temperature is generally going up with some seasonal patterns. Both models show more uncertainty about the long-term forecasts by providing results with 95% confidence ranges up to 2030.

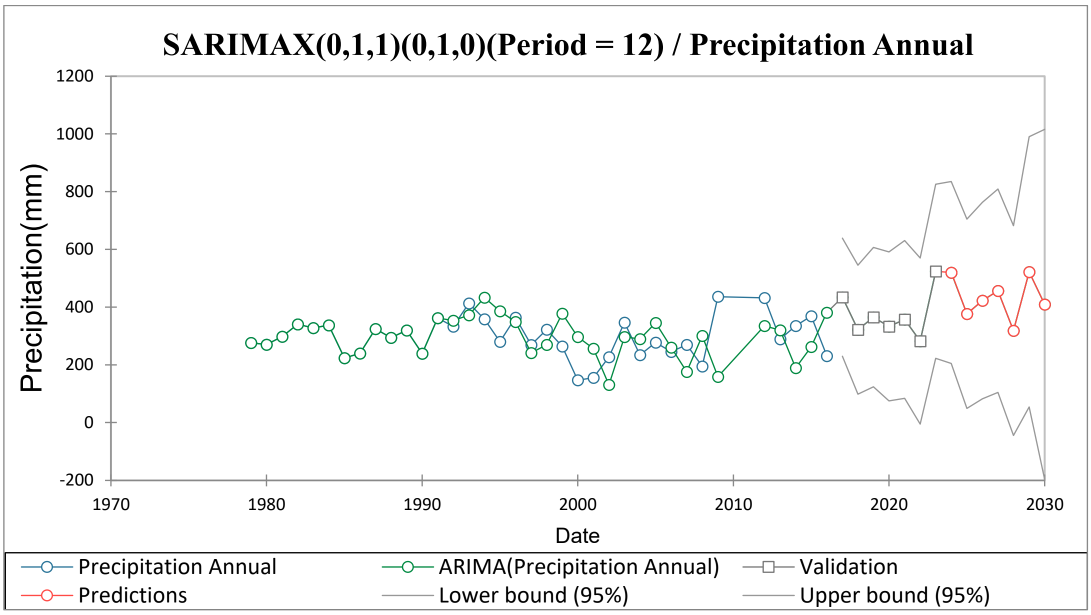



**Figure 14.** Long-term forecast for Annual Rainfall.

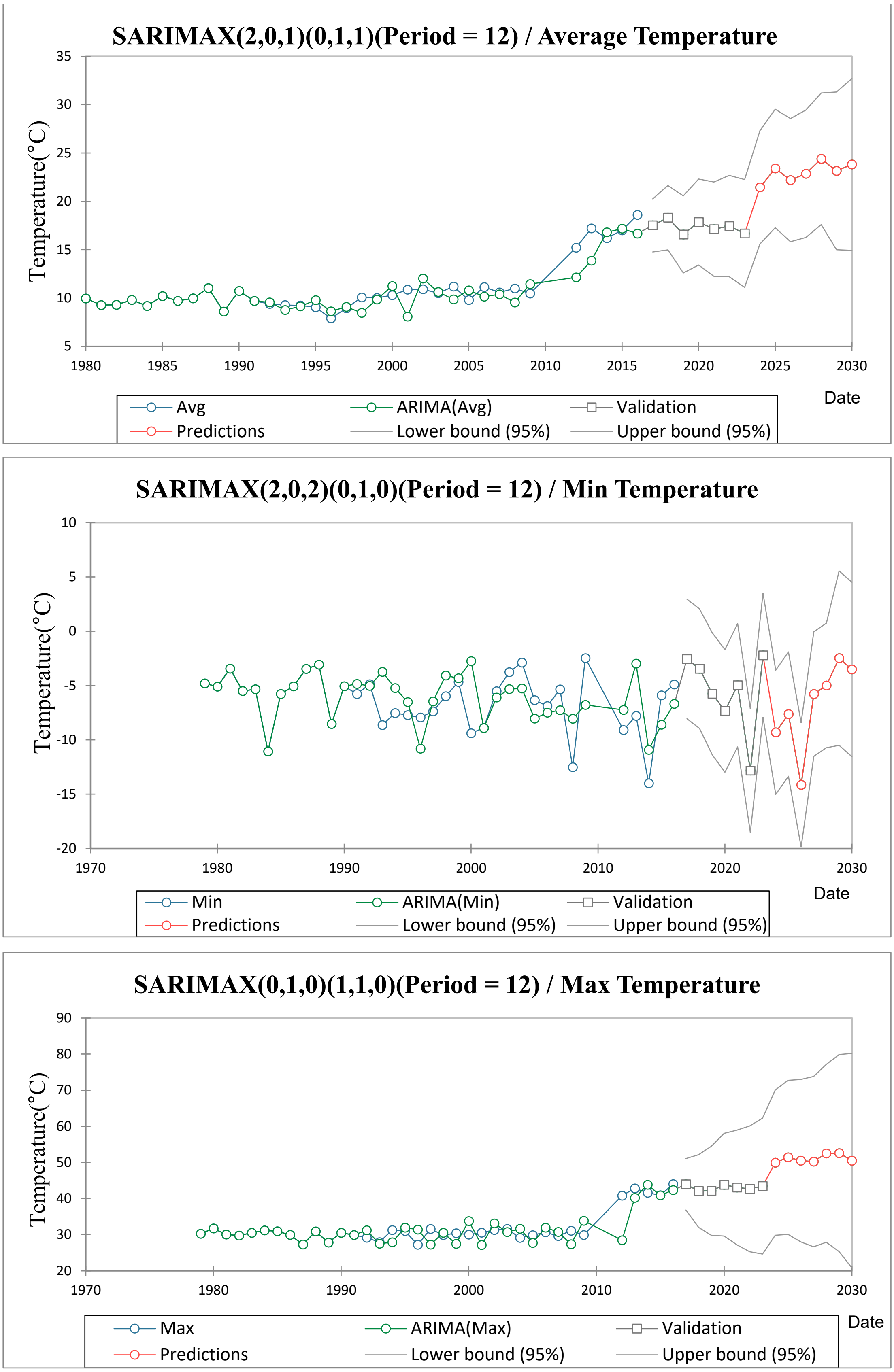
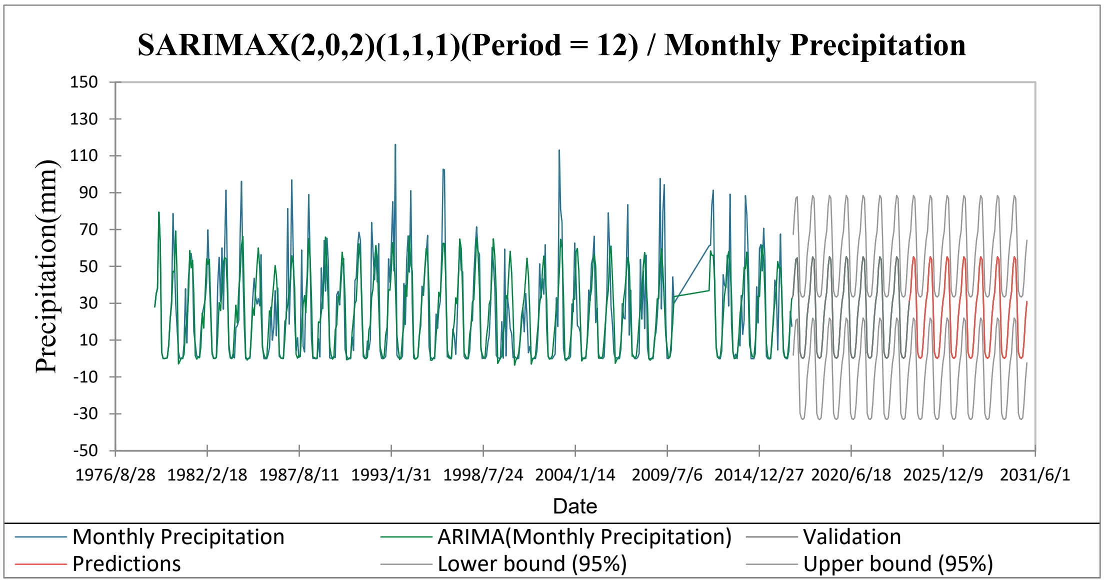



**Figure 15.** Predictions of Average, Min, and Max Temperature using SARIMA models.



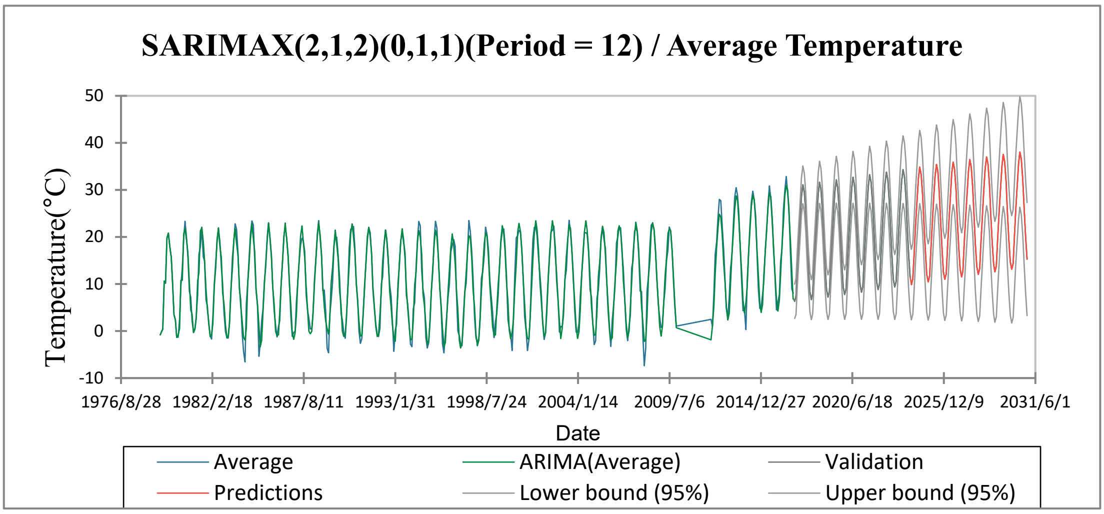




**Figure 16.** Long-term forecast for Annual Monthly Rainfall and Average Temperature.

## 5. Conclusions

This study applied ARIMA and SARIMA models to analyze and forecast hydroclimatic variables in the Kokcha River Basin using long-term observational data (1979–2023). The results reveal a clear warming trend alongside increasing variability in precipitation, indicating a shift toward more unstable hydroclimatic conditions in recent years. The statistical relationship between temperature and precipitation has weakened over time, suggesting growing complexity in climate dynamics.

The SARIMA models demonstrated good performance for short-term forecasting, as reflected by low RMSE and MAPE values, confirming their suitability for data-scarce regions. However, increasing uncertainty in long-term projections highlights the limitations of purely statistical approaches. The observed changes are likely driven by physical processes such as reduced snowpack, accelerated glacier melt in the Hindu Kush region, and increased evapotranspiration, all of which affect runoff generation and seasonal water availability.

These findings have important implications for water resources management in the basin. Rising temperatures and more variable precipitation are expected to alter streamflow patterns, increase the risk of floods and droughts, and reduce long-term water availability. Therefore, integrating statistical models with physically based approaches such as hydrological modeling and climate scenario analysis using downscaled GCM/RCM data is recommended for future research. This study provides valuable insights into hydroclimatic trends and forecasting in the Kokcha River Basin and supports the development of informed, climate-resilient water management strategies in Afghanistan.

## Author Contributions

A.G.: Writing—original draft, resources, formal analysis, methodology, and investigation; M.K.Y.: Writing—review and editing, supervision, investigation, formal analysis, and conceptualization. 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

All relevant data are included in the paper.

## Acknowledgments

We would like to thank Eng. Mohammad Hassan Eskandari (Head of the Hydrometeorological Database and Coordinator of the Panj-Amu River Basin) for providing the hydrometeorological data for the study area.

## Conflicts of Interest

The authors declare no conflict of interest.

## Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT to improve language, grammar, and clarity of the manuscript. After using this tool the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## References

1. Ghulami, M. Assessment of Climate Change Impacts on Water Resources and Agriculture in Data-Scarce Kabul Basin, Afghanistan. Available online: https://theses.hal.science/tel-01737052v1 (accessed on 9 May 2026).
2. Aich, V.; Akhundzadah, N.A.; Knuerr, A.; et al. Climate Change in Afghanistan Deduced from Reanalysis and Coordinated Regional Climate Downscaling Experiment (CORDEX)-South Asia Simulations. *Climate* **2017**, *5*, 38. https://doi.org/10.3390/cli5020038.
3. Akhundzadah, N.A.; Soltani, S.; Aich, V. Impacts of Climate Change on the Water Resources of the Kunduz River Basin, Afghanistan. *Climate* **2020**, *8*, 102. https://doi.org/10.3390/cli8100102.
4. Wester, P.; Mishra, A.; Mukherji, A.; et al. *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People*; Springer: Berlin/Heidelberg, Germany, 2019. https://doi.org/10.1007/978-3-319-92288-1.
5. Machiwal, D.; Jha, M.K. *Hydrologic Time Series Analysis: Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 1–303.
6. UNEP. *Environment and Security in the Amu Darya Basin*; UNEP: Nairobi, Kenya, 2011. Available online: https://wedocs.unep.org/handle/20.500.11822/7517 (accessed on 9 May 2026).
7. Moorthy, R.; Bibi, S. Water Security and Cross-Border Water Management in the Kabul River Basin. *Sustainability* **2023**, *15*, 792. https://doi.org/10.3390/su15010792.
8. Nori, S.M. Challenges of Transboundary Water Governance in Afghanistan. *Cent. Asian J. Water Res.* **2020**, *6*, 18–38. https://doi.org/10.29258/CAJWR/2020-R1.V6-1/18-38.ENG.
9. Alawi, S.A.; Özkul, S. Impact Assessment of Climate Change on Water Resources of the Kokcha Watershed: A Sub-Basin of the Amu Darya River Basin in Afghanistan. *J. Water Clim. Change* **2023**, *14*, 1656–1671. https://doi.org/10.2166/wcc.2023.504.

10. Akhundzadah, N.A. Analyzing Temperature, Precipitation, and River Discharge Trends in Afghanistan's Main River Basins Using Innovative Trend Analysis, Mann-Kendall, and Sen's Slope Methods. *Climate* **2024**, *12*, 196. https://doi.org/10.3390/cli12120196.
11. Yang, X.; Guo, L. A Study of the Extension of the China-Pakistan Economic Corridor to Afghanistan in the Context of the Taliban's Return to Power: A Response from Sichuan. *World J. Soc. Sci. Res.* **2022**, *9*, 43. https://doi.org/10.22158/wjssr.v9n2p43.
12. Sarkar, P.; Kumar, P.; Vishwakarma, D.K.; et al. Watershed Prioritization Using Morphometric Analysis by MCDM Approaches. *Ecol. Inform.* **2022**, *70*, 101763. https://doi.org/10.1016/j.ecoinf.2022.101763.
13. Shroder, J.F.; Ahmadzai, S.J. *Transboundary Water Resources in Afghanistan: Climate Change and Land-Use Implications*; Elsevier: Amsterdam, The Netherlands, 2016. Available online: https://books.google.com/books/about/Transboundary_Water_Resources_in_Afghani.html?hl=tr&id=kdyoBAAAQBAJ (accessed on 9 May 2026).
14. Afrifa-Yamoah, E.; Saeed, B.I.I.; Karim, A. Sarima Modelling and Forecasting of Monthly Rainfall in the Brong Ahafo Region of Ghana. *World Environ.* **2016**, *6*, 1–9. https://doi.org/10.5923/j.env.20160601.01.
15. Zhang, L.; Xin, F. Prediction Model of River Water Quality Time Series Based on ARIMA Model. *Commun. Comput. Inf. Sci.* **2019**, *980*, 127–133. https://doi.org/10.1007/978-981-13-7025-0_13.
16. Ray, S.; Das, S.S.; Mishra, P.; et al. Time Series SARIMA Modelling and Forecasting of Monthly Rainfall and Temperature in the South Asian Countries. *Earth Syst. Environ.* **2021**, *5*, 531–546. https://doi.org/10.1007/s41748-021-00205-w.
17. Abebe, T.H. Time Series Analysis of Monthly Average Temperature and Rainfall Using Seasonal ARIMA Model (in Case of Ambo Area, Ethiopia). *Int. J. Theor. Appl. Math.* **2020**, *6*, 76–87. https://doi.org/10.11648/j.ijtam.20200605.13.
18. Katimon, A.; Shahid, S.; Mohsenipour, M. Modeling Water Quality and Hydrological Variables Using ARIMA: A Case Study of Johor River, Malaysia. *Sustainable Water Resour. Manag.* **2018**, *4*, 991–998. https://doi.org/10.1007/s40899-017-0202-8.
19. Wilson, G.T. Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. *J. Time Ser. Anal.* **2016**, *37*, 709–711. https://doi.org/10.1111/jtsa.12194.
20. Dimri, T.; Ahmad, S.; Sharif, M. Time Series Analysis of Climate Variables Using Seasonal ARIMA Approach. *J. Earth Syst. Sci.* **2020**, *129*, 149. https://doi.org/10.1007/s12040-020-01408-x.
21. Abbas, S.A.; Xuan, Y.; Bailey, R.T. Assessing Climate Change Impact on Water Resources in Water Demand Scenarios Using SWAT-MODFLOW-WEAP. *Hydrology* **2022**, *9*, 164. https://doi.org/10.3390/hydrology9100164.
22. Rojas, I.; Pomares, H.; Herrera, L.J.; et al. Precipitation Time Series Analysis and Forecasting for Italian Regions. *Eng. Proc.* **2023**, *39*, 23. https://doi.org/10.3390/engproc2023039023.
23. Yetik, M.K. A Statistical Approach to Estimating Water Quality Parameter: A Case Study on Turkey's Rivers. *Global NEST J.* **2025**, *27*, gnest_06086. https://doi.org/10.30955/gnj.06086.
24. Achite, M.; Bazrafshan, O.; Azhdari, Z.; et al. Forecasting of SPI and SRI Using Multiplicative ARIMA under Climate Variability in a Mediterranean Region: Wadi Ouahrane Basin, Algeria. *Climate* **2022**, *10*, 36. https://doi.org/10.3390/cli10030036.
25. Gulistani, A.; Yetik, M.K. To Examine the Effects of Flow Analysis and Climate Change on Integrated Water Resources of the Kokcha River Basin in Afghanistan Using GIS Software. *Nat. Hazards* **2026**, *122*, 50. https://doi.org/10.1007/s11069-025-07795-y.
26. Tofallis, C. A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation. *J. Oper. Res. Soc.* **2015**, *66*, 1352–1362. https://doi.org/10.1057/jors.2014.103.
27. Ghimire, B.N.; Chandra Kaphle Devendra Adhikari Deependra Parajuli, G. Application of ARIMA Model for River Discharges Analysis. *J. Nepal Phys. Soc.* **2017**, *4*, 27–32. https://doi.org/10.3126/jnphyssoc.v4i1.17333.