CHAPTER 6
FORECASTING VARIABLES USING ARIMA: COMPARATIVE ANALYSIS OF INTER-REGIONAL CHANGES IN GUJARAT

6.1 Introduction

Forecasting variables, particularly in climate studies, hold significant importance as they allow for predicting future climatic conditions (Huangpeng et al., 2021), aiding in effective resource management and planning. With the advancement of computational methods, many models have emerged, offering a range of options for forecasting tasks. Various models have been developed to enhance the accuracy and reliability of these predictions, each with its unique strengths and limitations. Among these models, the Autoregressive Integrated Moving Average (ARIMA) is one of the most powerful and widely accepted tools (Chandran et al., 2023). Developed by George E. P. Box and Gwilym M. Jenkins (Box & Jenkins, 1976), the ARIMA model has a rich history in time series analysis (Khashei et al., 2012). It has gained global acceptance due to its robust mathematical foundation and proven effectiveness in a wide range of forecasting applications (Rizeei et al., 2018). The ARIMA model has defined by the three parameters: p, d, and q. These parameters represent the order of the autoregressive component, the number of differencing steps required to make the time series stationary, and the order of the moving average component, respectively (Sirisha et al., 2022). The ability to handle non-stationary data by integrating differencing techniques makes it particularly valuable for forecasting climatic variables, which often exhibit trends and seasonal patterns (Dimri et al., 2020). Central to the Box-Jenkins approach is its systematic methodology, which involves, the identification, estimation, and diagnostic checking of observed time series data (Unnikrishnan & Jothiprakash, 2020). This process effectively incorporates forecasting errors (noise) into the model, allowing for continuous adjustments and refinements.

ARIMA models address linear correlation among sequential observations, while SARIMA models expand this capability to effectively capture seasonal variations and non-stationarity within and across seasons in time series data (Wilson, 2016). The ability of (S)ARIMA models to integrate with diagnostic tools like the Partial Autocorrelation Function (PACF), Autocorrelation Function (ACF), and Augmented Dickey-Fuller (ADF) test is one of its main advantages (Dimri et al., 2020). The ADF test is crucial for testing the stationarity of time series data. Stationarity is a fundamental assumption in time series analysis, and the ADF test helps determine whether differencing is required to achieve it. The ACF and PACF are equally

essential in the Box-Jenkins methodology for identifying the appropriate model parameters (Adineh et al., 2021). The ACF measures the correlation between the time series and its lagged values, helping to determine the number of lag terms to include in the model (Schaffer et al., 2021). Meanwhile, the PACF provides information on the correlation of the series with its lagged values, controlling for the values of the intermediate lags (Mestre et al., 2021). These helps identify the specific lags that significantly affect the time series, which is essential for accurately specifying the ARIMA model. Using these tools, the model reflects the underlying patterns and structures within the data and further enhances the model's predictive power and accuracy, making it an indispensable tool in climate forecasting (Mestre et al., 2021). The resulting forecasts allow researchers and policymakers to anticipate and plan for future climatic conditions more effectively. Additionally, the integration of an exogenous predictor transforms the SARIMA model into SARIMAX (Alharbi & Csala, 2022; Sirisha et al., 2022), allowing for the inclusion of external factors influencing the time series beyond its inherent components. This enhancement broadens SARIMAX's utility in diverse forecasting applications. To effectively incorporate an exogenous variable into the forecast, it is necessary to have access to the variable's values for the forecast period as well (Sirisha et al., 2022).

6.2 Methodology

The methodology begins with thorough data preprocessing in Python, utilizing pandas to handle rainfall and temperature datasets. Ensuring accurate date/time formatting and numeric consistency is a priority, followed by data resampling and normalization where necessary. The stationarity analysis uses the ADF test, with differencing applied to eliminate trends. ACF and PACF plots assist in selecting parameters for ARIMA (p, d, q) models, guiding the determination of AR and MA orders. Based on the analysis of the plots and the minimum Akaike Information Criterion (AIC) values, appropriate ARIMA models have been chosen using suitable parameter combinations. For datasets exhibiting seasonality, SARIMA models are applied to improve predictive performance. After preprocessing, the dataset is divided into training and testing sets to validate the model. The selected model is trained on the training set and evaluated on the testing set using metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to assess forecast accuracy. The trained model is then used to project future rainfall and temperature trends. These projections also serve as inputs to calculate Potential Evapotranspiration (PET) using the Hargreaves equation. Further identification of water surplus or deficit can be achieved by subtracting the forecasted PET from the rainfall across regions. Figure 6.1 outlines the adopted methodology for the analysis.

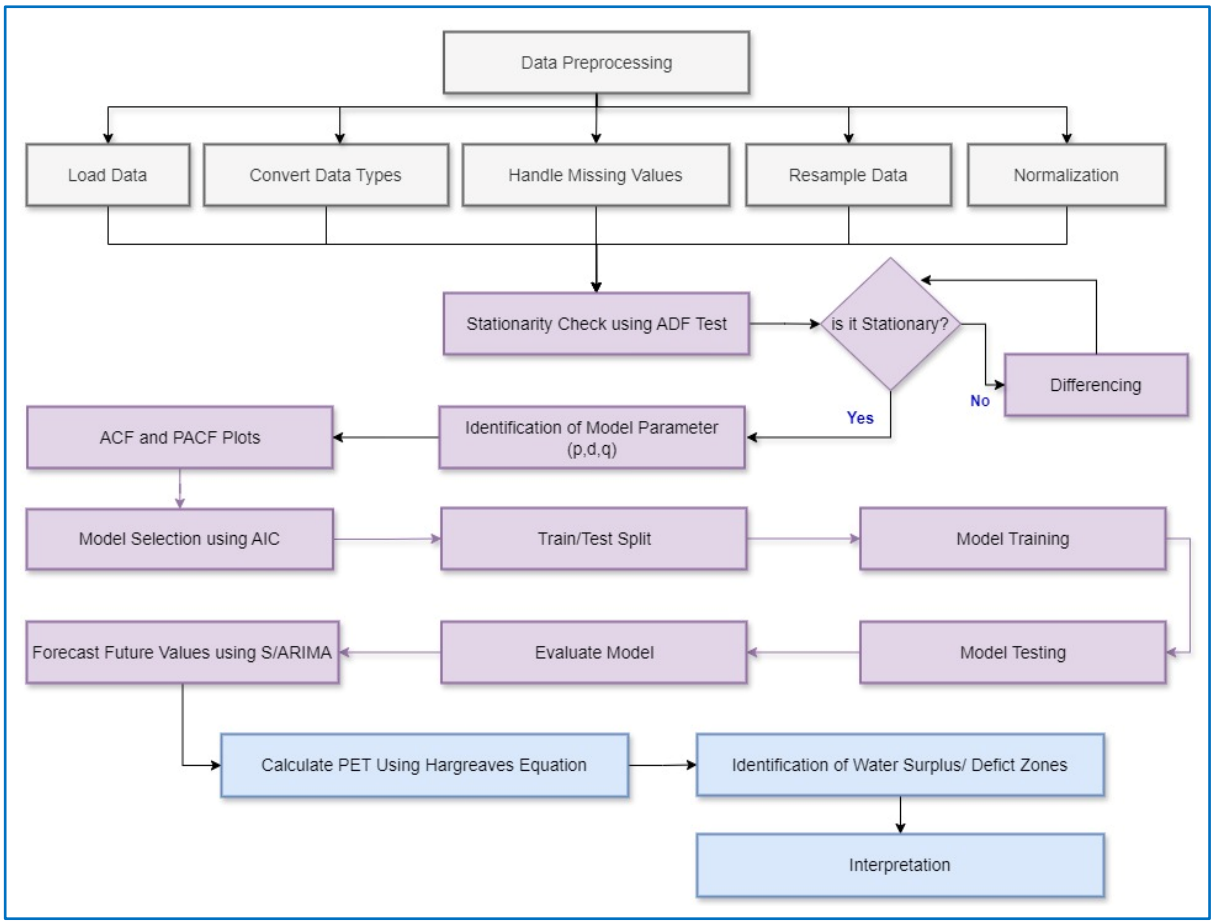


Figure 6.1 Flowchart of the Methodology Adopted.

6.2.1 Augmented Dickey-Fuller Test

The ADF test is employed to assess whether a time series is stationary, meaning its statistical properties remain constant over time. Stationarity is essential for many time series modeling techniques, including ARIMA (Al Sayah et al., 2021). The ADF (Dickey & Fuller, 1979) test is an extension of the Dickey-Fuller test, adding lagged differences of the series as predictors to account for higher-order correlation. The primary advantage of the ADF test lies in its ability to handle high variability in data, making it robust in real-world applications where data often exhibit complex patterns and noise (Gianfreda et al., 2023). By including lagged differences, the ADF test can better address autocorrelation issues, providing a more reliable assessment of stationarity. Performing the ADF test involves estimating the following regression model:

$$\Delta_{yt} = \alpha + \beta t + \gamma_{yt-1} + \sum_{i=1} \delta_i \Delta_{yt-i} + \epsilon_t \quad (1)$$

Where Δ_{yt} is the first difference of the series y_t ($y_t - y_{t-1}$); α is a constant; βt is a trend term; γ_{yt-1} is the lagged level of the series; $\delta_i \Delta_{yt-i}$ is the lagged differences; and ϵ_t is the error term.

The null hypothesis of the ADF test is that the series has a unit root, i.e., non-stationarity (More et al., 2015). If the test statistic is less than the critical value, the null hypothesis can be rejected, indicating that the series is stationary. In the context of ARIMA modeling, achieving stationarity is crucial, as the ARIMA framework assumes that the time series is stationary (Schaffer et al., 2021). If the ADF test indicates non-stationarity, differencing is applied to the series. Differencing involves subtracting the previous observation from the current observation:

$$y_t' = y_t - y_{t-1} \quad (2)$$

If the initial differencing does not achieve stationarity, additional differencing may be applied to ensure the time series meets the stationarity requirement. The order of differencing (d) is a key parameter, representing the number of times differencing is performed to achieve stationarity. The differenced series is subsequently utilized to fit the ARIMA model, with the ACF and PACF plots guiding the selection of the AR and MA components, respectively (Schaffer et al., 2021).

6.2.2 Autoregressive Integrated Moving Average Model

The ARIMA model is a robust tool for forecasting time series data (Chandran et al., 2023), making it highly suitable for predicting rainfall and temperature patterns. One of the main advantages of using ARIMA in such forecasting is its ability to handle various types of data patterns, including trends and seasonality (Adineh et al., 2021), by incorporating differencing and moving average components. This flexibility allows ARIMA to model complex climate data accurately, making it an ideal choice for meteorological and environmental forecasting (Al Balasmeh et al., 2019). The general form of an ARIMA model can be represented by the equation (Nguyen & Pham, 2024):

$$yt = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (3)$$

Where y_t is the value at time t, c is a constant, ϕ represents the autoregressive coefficients, θ represents the moving average coefficients, and ϵ_t is the error term (Essefiani et al., 2024).

When seasonality is present in the data, the Seasonal ARIMA (SARIMA) model is used, which extends ARIMA by incorporating seasonal parameters: P, D, Q, and s (Adams et al., 2020). The general form of a SARIMA model can be represented as (Ma et al., 2021):

$$SARIMA(p, d, q)(P, D, Q)[s] \quad (4)$$

Where P, D, and Q represent the seasonal AR, differencing, and MA orders, respectively, and s represents the length of the seasonal cycle.

A crucial step in ARIMA modeling is determining the parameters p, d, and q through ACF and PACF analysis. The ACF plot identifies the moving average order (q), while the PACF plot determines the autoregressive order (p). Differencing (d) is applied to achieve stationarity, verified by the ADF test. The ACF measures the correlation between observations of a time series that are separated by k time units (Mestre et al., 2021). The ACF at lag k, denoted as ρ(k), is given by:

$$\rho(k) = \frac{\sum_{t=1}^{N-K} (X_t - \ddot{X})(X_{t+k} - \ddot{X})}{\sum_{t=1}^{N} (X_t - \ddot{X})^2} \quad (5)$$

where X_t is the value of the time series at time t, $\ddot{X}$ is the mean of the time series, and N is the total number of observations. While The PACF measures the correlation between observations of a time series separated by k time units, after removing the effects of shorter lags (Mestre et al., 2021). The PACF at lag k, denoted as ϕ(k), is computed by regressing X_t on X_{t-1}, $X_{t-2..,}X_{t-k.}$

The selection of the best-fit (S)ARIMA model is guided by the Akaike Information Criterion (AIC) (Adams et al., 2020). The model with the lowest AIC value is preferred, as it strikes a balance between model complexity and accuracy, thereby minimizing the risk of overfitting. The equation for AIC is:

$$AIC = 2k - 2\ln(L) \quad (6)$$

Where K is the number of estimated parameters in the model; and L is the maximum value of the likelihood function for the model.

The selected model undergoes training on the designated training set, and its performance is rigorously evaluated on the testing set using the metrics of MAE and RMSE. These metrics are employed to analyse the predictive accuracy of the model (Z. Zhao et al., 2023). MAE measures the absolute differences between SARIMA model forecasts and the actual values in the test set (Kabbilawsh et al., 2022). It is calculated as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \quad (7)$$

where y_i represents the actual value, $\hat{y}_i$ represents the predicted value, and n is the number of data points (Kabbilawsh et al., 2022). RMSE, on the other hand, quantifies the differences between the values predicted by a SARIMA model and the actual values in the test set (Z. Zhao et al., 2023). The formula for RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (8)$$

where yi represents the projected values, $\hat{y}_i$ corresponds to the observed or test value and n denotes the total number of data points. Both RMSE and MAE are scale-dependent metrics used in this study to assess the accuracy of the SARIMA models (Kabbilawsh et al., 2022).

6.2.3 Potential Evapotranspiration (PET)

PET quantifies the atmosphere's capacity to remove water through evaporation and transpiration (Zomer et al., 2022), reflecting the atmospheric demand for moisture. Additionally, it refers to the maximum ET achieved in a particular region under the assumption that water supply is not restricted (Y. Li et al., 2022). Numerous methods for calculating PET exist, each with its own set of advantages and limitations. Among these methods, the Hargreaves equation (Hargreaves & Samani, 1985) is particularly advantageous due to its simplicity and less data requirements. Unlike more complex methods that require extensive meteorological data, the Hargreaves equation primarily relies on temperature data, making it accessible and practical for regions with limited data availability (Y. Li et al., 2022). The Hargreaves Equation is expressed as:

$$PET = 0.0023 \times (T_{avg} + 17.8) \times (T_{max} - T_{min}) 0.5 \times R_a \quad (9)$$

Where T_{avg} is the average monthly temperature (°C), T_{max} is the maximum monthly temperature (°C), T_{min} is the minimum monthly temperature (°C), and R_a is the extraterrestrial radiation.

To identify regions of water surplus and deficit, the water balance is calculated by subtracting PET from rainfall (X. Li & Huang, 2021). A positive water balance indicates a surplus, suggesting that rainfall exceeds atmospheric demand. Conversely, a negative water balance indicates a deficit, potentially leading to drought conditions.

6.3 Results and Discussion

The preceding chapter systematically addressed data pre-processing and trend analysis, setting the stage for this chapter, which focuses on assessing stationarity and selecting the best-fit model for accurate forecasting. In this chapter, a thorough examination of stationarity is conducted, ensuring that the time series data meet the necessary conditions for reliable modeling. The Augmented Dickey-Fuller (ADF) test was utilised to assess stationarity. In this analysis, the critical values were -3.43, -2.86, and -2.56 at the 1%, 5%, and 10% significance levels, respectively. Table 6.1 shows that for the time series of Rainfall, T_{max}, and T_{min} across regions, the ADF test statistics consistently fell below the critical thresholds, confirming stationarity at the 1% significance level. This result highlights the robustness and stability of the dataset. Additionally, stationarity observed post-differencing further validates the data.

Table 6.1 ADF Test Statistics: Before and After Differencing.

		Before Differencing		After Differencing	
Region	Variable	Test Statistic	p-value	Test Statistic	p-value
Central Gujarat	Rainfall	-6.64	0.00	-16.08	0.00
	T_{max}	-7.09	0.00	-8.51	0.00
	T_{min}	-3.93	0.00	-7.88	0.00
Kutch	Rainfall	-6.01	0.00	-20.35	0.00
	T_{max}	-5.68	0.00	-7.88	0.00
	T_{min}	-4.36	0.00	-8.04	0.00
North Gujarat	Rainfall	-6.93	0.00	-21.90	0.00
	T_{max}	-6.59	0.00	-8.06	0.00
	T_{min}	-4.98	0.00	-7.91	0.00
Saurashtra	Rainfall	-5.19	0.00	-22.68	0.00
	T_{max}	-6.49	0.00	-8.30	0.00
	T_{min}	-3.52	0.00	-7.85	0.00
South Gujarat	Rainfall	-5.74	0.00	-13.10	0.00
	T_{max}	-7.60	0.00	-8.66	0.00
	T_{min}	-4.51	0.00	-7.39	0.00

The results indicate that the time series are well-suited for further analysis and reliable forecasting, as the risk of non-stationarity affecting model accuracy is low. However, since the ADF test might not reveal all underlying patterns, using the AIC is essential to determine whether differencing is required (Cavaliere et al., 2015). This ensures that any remaining non-stationarity is addressed, enhancing the accuracy and reliability of the forecasting model (L. Zhao et al., 2022). Following the assessment of stationarity with the ADF test, the analysis continued with examining ACF and PACF plots to investigate autocorrelations and identify potential ARIMA model components. Figures 6.2 to 6.4 show the ACF and PACF plots for Rainfall, Tmax, and Tmin, showing both before and after differencing analyses across the distinct physiographic regions of Gujarat. These plots reveal changes in autocorrelation patterns and highlight the role of differencing in reducing temporal dependencies, which aids in identifying suitable ARIMA model components. These plots are essential for achieving stationarity and selecting ARIMA model orders: the ACF plot indicates the q parameter by showing significant autocorrelations at specific lags, while the PACF plot identifies the p parameter by revealing significant lags after accounting for intervening effects . Figure 6.2 shows the ACF and PACF plots of the annual rainfall time series across the physiographic regions of Gujarat, comparing the patterns before and after first-order differencing. This analysis is crucial for selecting an appropriate modeling approach and ensuring the accuracy of the forecasts.

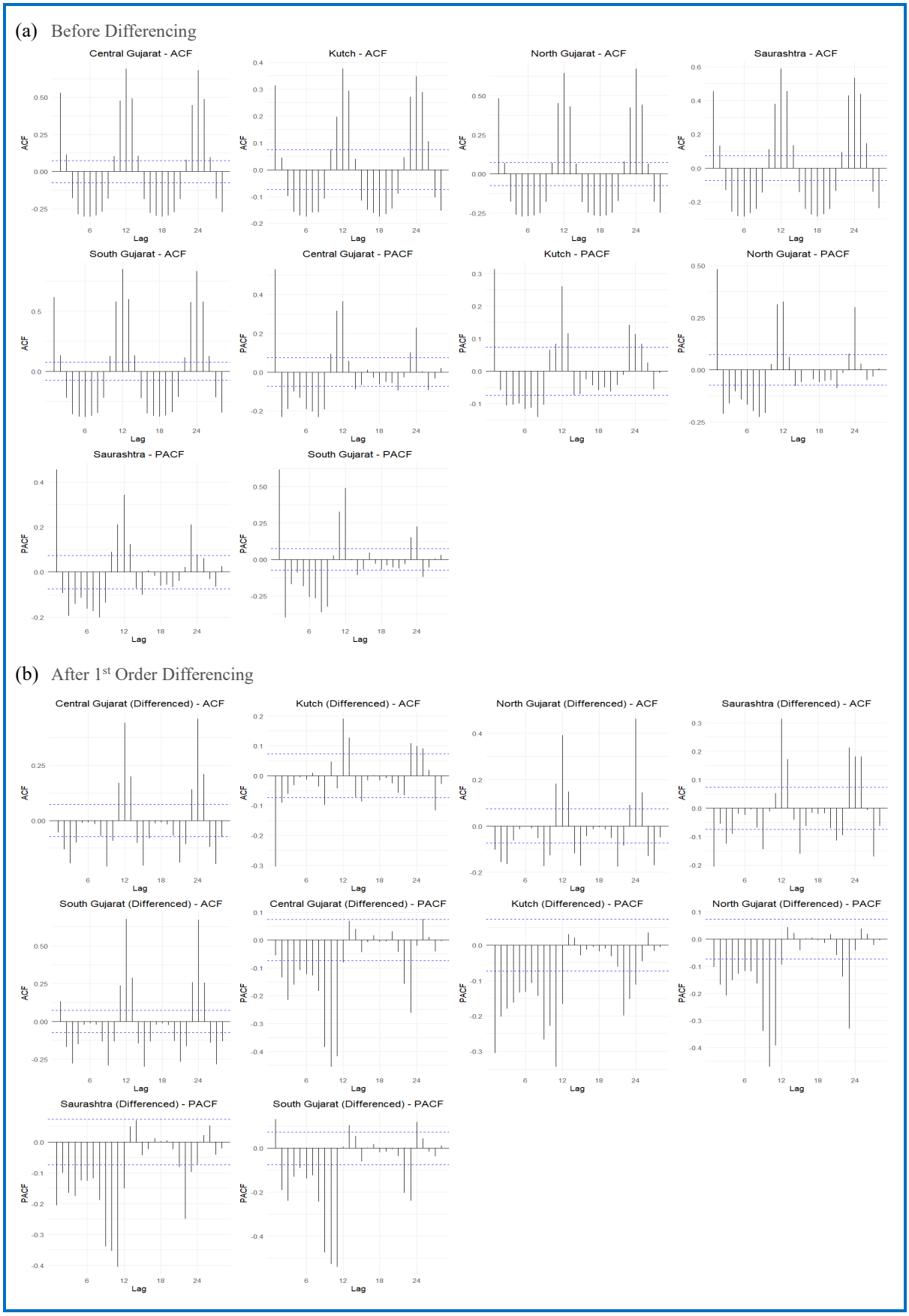


Figure 6.2 ACF and PACF Plots of Annual Rainfall Across Physiographic Regions of Gujarat: Comparative Analysis (a) Before and (b) After First-Order Differencing.

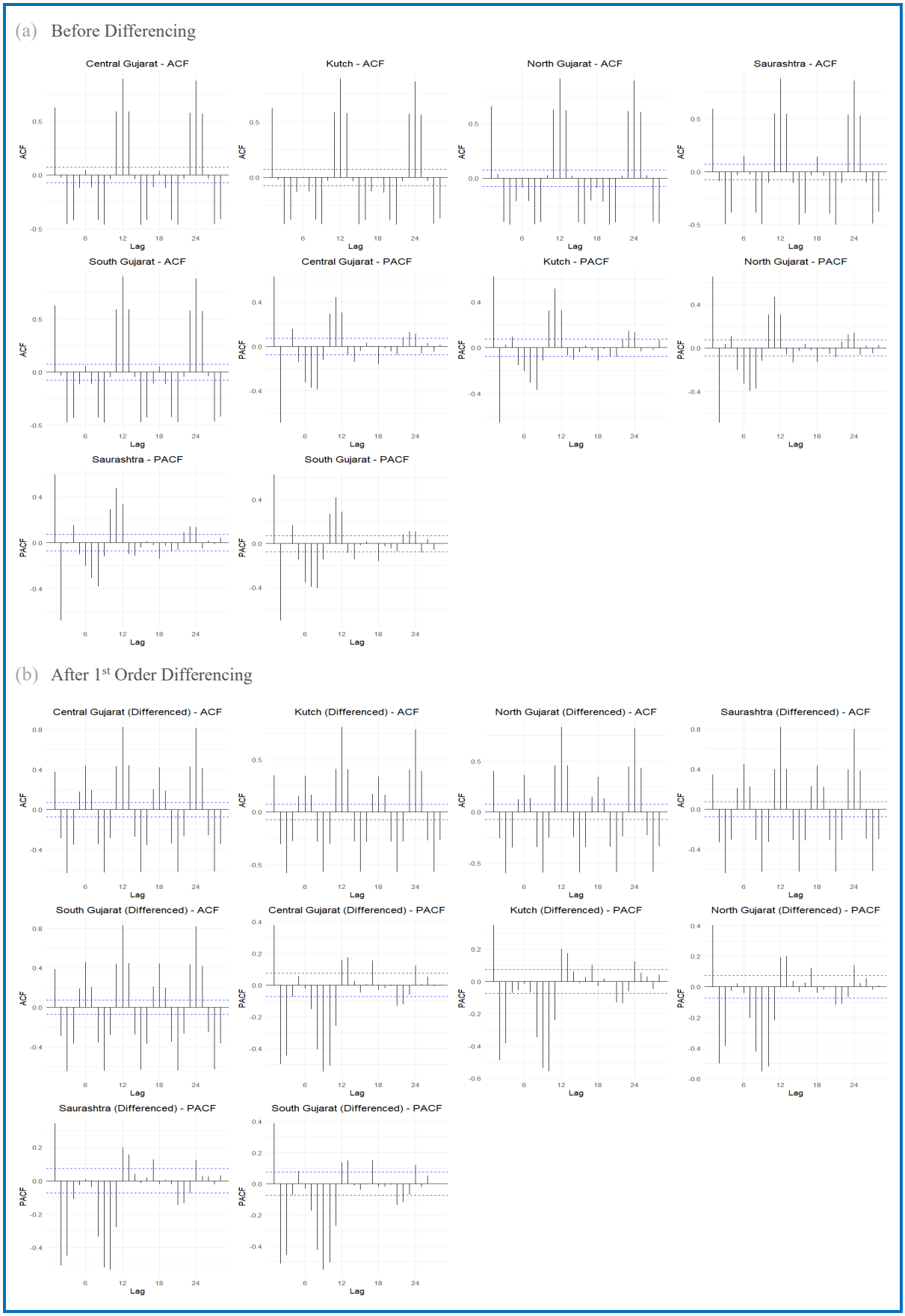


Figure 6.3 ACF and PACF Plots of Annual T_{max} Across Physiographic Regions of Gujarat: Comparative Analysis (a) Before and (b) After First-Order Differencing.

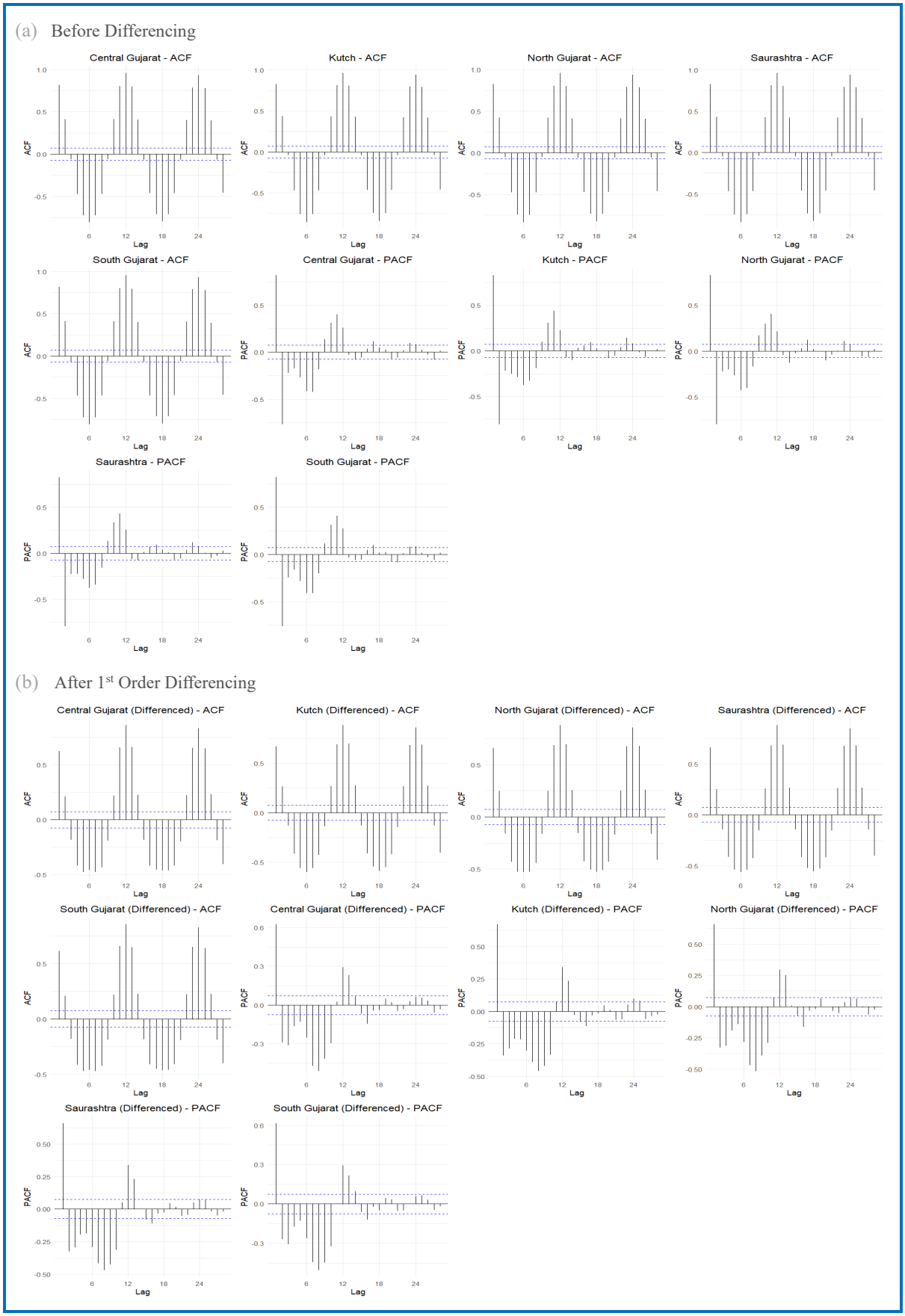


Figure 6.4 ACF and PACF Plots of Annual T_{min} Across Physiographic Regions of Gujarat: Comparative Analysis (a) Before and (b) After First-Order Differencing.

Figures 6.2 to 6.4 show that the ACF and PACF plots of the annual rainfall and temperature (T_{max} and T_{min}) series reveal seasonality across variables for the distinct physiographic regions of Gujarat, highlighting the need for employing a SARIMA model for analysis. The ACF plots show recurring spikes at regular seasonal lags, indicating strong autocorrelation at these intervals. Similarly, the PACF plots show significant correlations at corresponding lags, confirming the presence of seasonal patterns in the data. First-order differencing was applied to achieve stationarity in the seasonal component across regions, except for Kutch's rainfall data, which exhibits stationarity without requiring differencing. Additionally, no further differencing was needed, as the data achieved stationarity through the initial adjustments.

The significant spikes observed in the ACF and PACF plots across the physiographic regions provide crucial insights for selecting the parameters of the SARIMA (p, d, q)(P, D, Q)[s] model. The repeated spikes in the ACF indicate the order of the MA component (q), reflecting the extent to which past forecast errors influence current observations and suggesting the inclusion of lagged values to capture autocorrelation. Similarly, the significant correlations observed in the PACF indicate the order of the AR component (p), highlighting the relationship between current observations and their past values (Adams et al., 2020). These insights show that the SARIMA model integrates historical errors and past observations to enhance forecasting accuracy. While the ACF and PACF plots aid in identifying model parameters, the optimal model is eventually selected based on the minimum AIC value (L. Zhao et al., 2022) and log-likelihood measures. This approach ensures that the chosen model captures the underlying patterns and balances the complexity and predictive accuracy (Chandran et al., 2023).

6.3.1 Best-Fit Model Selection and Forecast Trends

Following the assessment of stationarity and identification of model parameters, the best-fit model was selected based on the AIC, which balances goodness of fit with model complexity, incorporating log-likelihood measures (Kabbilawsh et al., 2022). A lower AIC indicates a more efficient model that captures the underlying patterns without overfitting, while a higher log likelihood suggests a better fit to the observed data (Ziegel et al., 1995). Balancing these metrics ensures the model is accurate and not overly complex, improving its ability to generalize to new data (Merabet et al., 2021). This approach helps identify models that accurately reflect the data while minimizing the risk of overfitting, ensuring reliable forecasts. Table 6.1 shows the best-fit SARIMA models selected for each variable across the regions based on the AIC. This selection ensures an optimal balance between model complexity and explanatory power, with model performance validated through MAE and RMSE metrics. These metrics quantify the

deviations between predicted and observed values, with lower values indicating high accuracy (Z. Zhao et al., 2023). Additionally, these metrics affirm that the selected models show a robust fit and reliable forecasting performance, rendering them well-suited for future predictions.

Table 6.2 Best-Fit SARIMA Models for Annual Rainfall and Temperature (T_{min} and T_{max}) with Performance Metrics Across Distinct Physiographic Regions of Gujarat.

Region	Variable	Best-Fit SARIMA Model	AIC	Log-Likelihood	MAE	RMSE
Central Gujarat	Rainfall	SARIMA(1, 0, 0)(1, 1, 0)[12]	8146.8	-4070.4	33.01	68.61
	T_{max}	SARIMA(2, 0, 0)(2, 1, 0)[12]	2177.3	-1083.6	0.89	1.12
	T_{min}	SARIMA(1, 0, 3)(1, 1, 0)[12]	2001.8	-993.9	0.68	1.00
Kutch	Rainfall	SARIMA(1, 0, 2)(2, 0, 0)[12]	8085.1	-4035.5	53.89	92.92
	T_{max}	SARIMA(2, 0, 1)(1, 1, 0)[12]	2038.7	-1013.3	0.90	1.15
	T_{min}	SARIMA(1, 0, 2)(2, 1, 0)[12]	1783.2	-884.6	0.58	0.80
North Gujarat	Rainfall	SARIMA(1, 0, 0)(1, 1, 0)[12]	8062.3	-4027.1	39.66	76.77
	T_{max}	SARIMA(2, 0, 0)(2, 1, 0)[12]	2227.3	-1108.6	0.90	1.15
	T_{min}	SARIMA(1, 0, 3)(2, 1, 0)[12]	1966.8	-975.4	0.59	0.83
Saurashtra	Rainfall	SARIMA(0, 0, 1)(2, 1, 1)[12]	7838.2	-3914.1	41.86	85.94
	T_{max}	SARIMA(2, 0, 0)(2, 1, 0)[12]	2227.3	-1108.6	0.76	0.96
	T_{min}	SARIMA(1, 0, 2)(2, 1, 0)[12]	1760.4	-873.2	0.56	0.78
South Gujarat	Rainfall	SARIMA(1, 0, 0)(2, 1, 0)[12]	8068.5	-4029.2	38.61	86.04
	T_{max}	SARIMA(2, 0, 0)(2, 1, 0)[12]	2227.3	-1108.6	0.87	1.09
	T_{min}	SARIMA(2, 0, 2)(1, 1, 0)[12]	1996.2	-992.1	0.71	1.10

Table 6.1 provides a comprehensive summary of the selected SARIMA models at the annual scale, highlighting their effectiveness in capturing the seasonal dynamics of the data. Based on AIC and log-likelihood measures, SARIMA(2, 0, 0)(2, 1, 0)[12] is identified as the most appropriate model for T_{max}, while Kutch requires an additional MA term in the non-seasonal component. For T_{min}, the optimal models across regions effectively capture seasonal dynamics, with identical models chosen (SARIMA(1, 0, 2)(2, 1, 0)[12]) for Kutch and Saurashtra, (SARIMA(1, 0, 3)(1, 1, 0)[12]) for Central Gujarat and North Gujarat, with the latter requiring an additional seasonal AR term to capture distinctive seasonal patterns. South Gujarat incorporates both AR and MA terms to improve adaptability to seasonal trends. For rainfall, the models chosen for Central Gujarat and North Gujarat are identical (SARIMA(1, 0, 0)(1, 1, 0)[12]), employing a simple structure that effectively captures seasonal patterns. Kutch enhances its model with an additional MA term in the seasonal component. Saurashtra highlights the importance of the MA term in effectively addressing seasonal dynamics. In contrast, South Gujarat emphasizes the role of seasonal terms in capturing rainfall trends.

Following model selection, validation using MAE and RMSE metrics confirms the accuracy of the models. The T_{min}, T_{max}, and rainfall models show strong fit across regions, with consistently low error values, strengthening the reliability of the selected SARIMA models in capturing seasonal dynamics and providing accurate forecasts. Meanwhile, Table 6.2 shows the coefficients and standard errors for the selected SARIMA model parameters across Gujarat's physiographic regions. This table shows insights into the dynamics of each variable, showing how AR and MA terms contribute to the models. The coefficients indicate the strength and direction of relationships, while the standard errors assess the reliability of these estimates, providing insight into the model's effectiveness in capturing each region's seasonal dynamics.

Table 6.3 Coefficients and Standard Errors for Selected SARIMA Model Parameters.

Region	Variable	AR1	AR2	MA1	MA2	MA3	SAR1	SAR2	SMA1
Central Gujarat	Rainfall	0.157	-	-	-	-	-0.497	-	-
		0.037		-	-	-	0.033	-	-
	T_{max}	0.313	0.072	-	-	-	-0.682	-0.344	-
		0.038	0.038	-	-	-	0.036	0.036	-
	T_{min}	0.772	-	-0.544	-0.164	0.081	-0.450	-	-
		0.110	-	0.114	0.049	0.038	0.034	-	-
Kutch	Rainfall	0.909	-	-0.774	-0.180	-	0.283	0.207	-
		0.045	-	0.060	0.040	-	0.038	0.039	-
	T_{max}	0.999	-0.190	-0.675	-	-	-0.480	-	-
		0.235	0.096	0.229	-	-	0.034	-	-
	T_{min}	0.848	-	-0.562	-0.169	-	-0.621	-0.326	-
		0.079	-	0.089	0.049	-	0.037	0.036	-
North Gujarat	Rainfall	0.122	-	-	-	-	-0.547	-	-
		0.038	-	-	-	-	0.032	-	-
	T_{max}	0.321	0.045	-	-	-	-0.680	-0.331	-
		0.038	0.038	-	-	-	0.037	0.036	-
	T_{min}	0.841	-	-0.588	-0.119	0.067	-0.631	-0.338	-
		0.067	-	0.076	0.048	0.041	0.037	0.036	-
Saurashtra	Rainfall	-	-	0.077	-	-	-0.058	-0.178	-0.854
		-	-	0.036	-	-	0.065	0.059	0.066
	T_{max}	0.321	0.045	-	-	-	-0.680	-0.331	-
		0.038	0.038	-	-	-	0.037	0.036	-
	T_{min}	0.850	-	-0.574	-0.158	-	-0.598	-0.362	-
		0.075	-	0.085	0.048	-	0.036	0.036	-
South Gujarat	Rainfall	0.151	-	-	-	-	-0.675	-0.338	-
		0.038	-	-	-	-	0.037	0.036	-
	T_{max}	0.321	0.045	-	-	-	-0.680	-0.331	-
		0.038	0.038	-	-	-	0.037	0.036	-
	T_{min}	0.490	0.298	-0.268	-0.384	-	-0.438	-	-
		0.200	0.150	0.193	0.121	-	0.034	-	-

**SAR Seasonal Autoregressive*
**SMA Seasonal Moving Average*

Coefficient Standard Error

Figure 6.5 Mean Monthly Rainfall Forecasts (2021-2050) and Deviations from Observed Averages Across Gujarat's Physiographic Regions: Projected Variability in Rainfall Patterns.

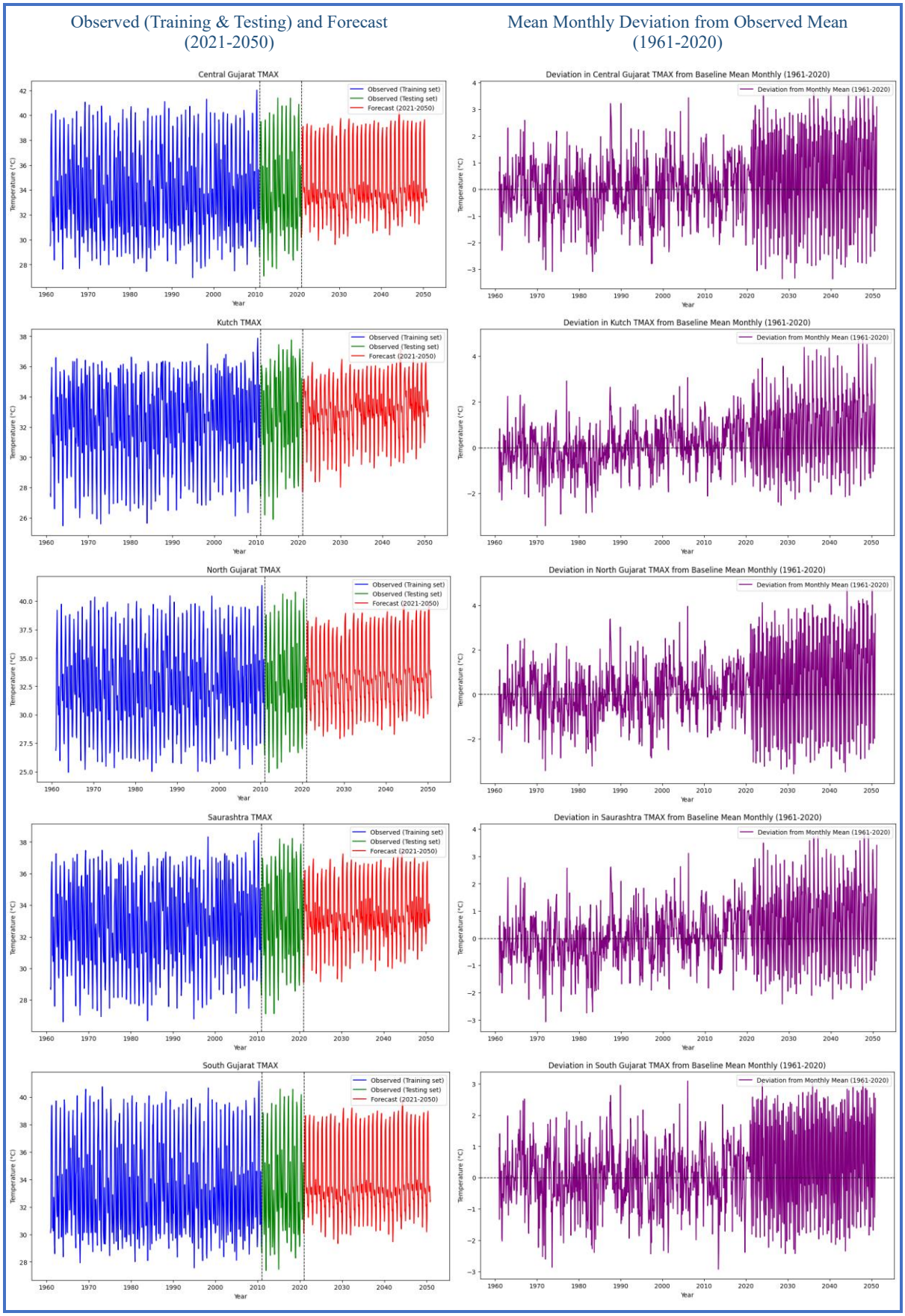


Figure 6.6 Mean Monthly T_{max} Forecasts (2021-2050) and Deviations from Observed Averages Across Gujarat's Physiographic Regions: Projected Variability in T_{max} Patterns.

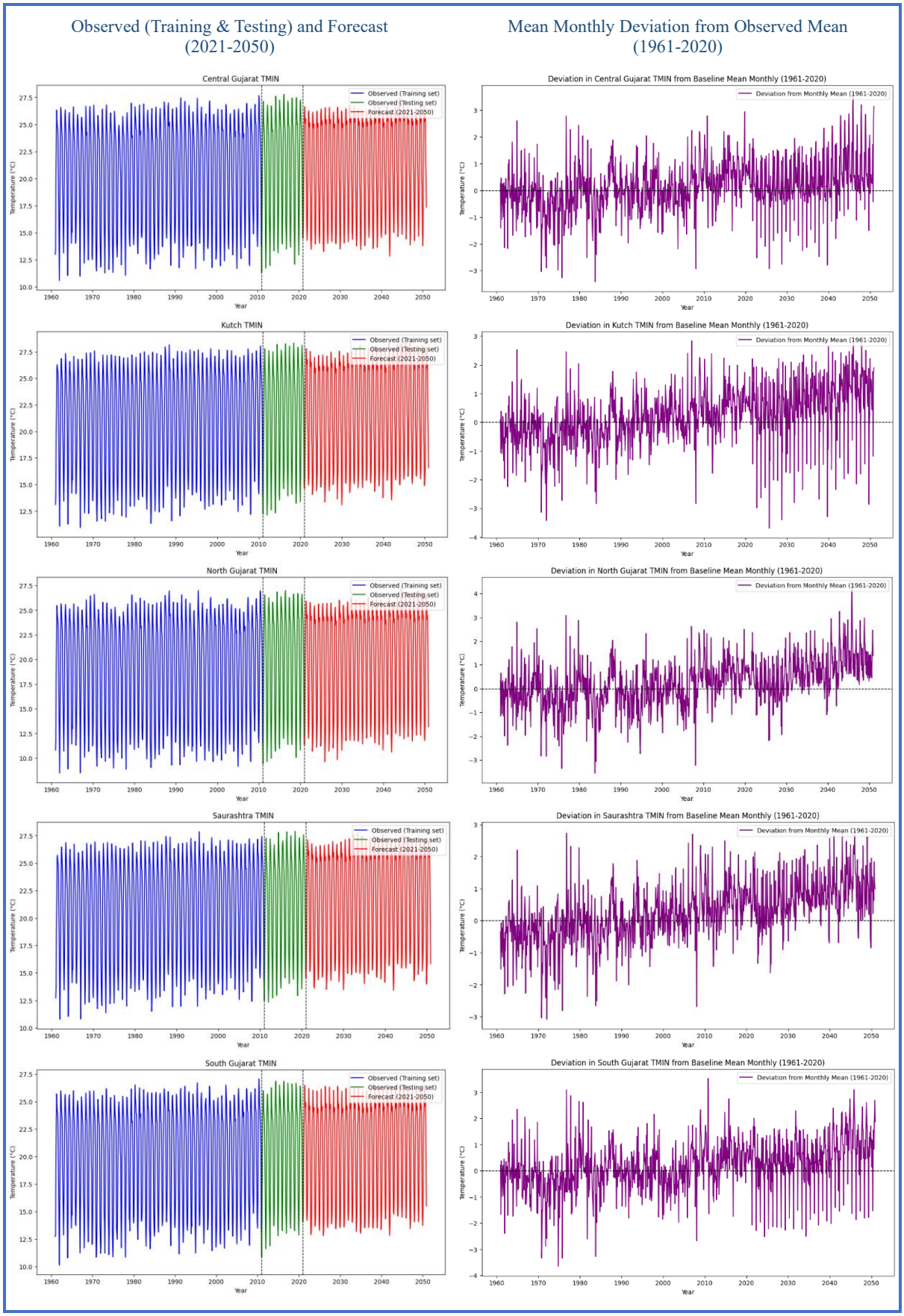


Figure 6.7 Mean Monthly T_{min} Forecasts (2021-2050) and Deviations from Observed Averages Across Gujarat's Physiographic Regions: Projected Variability in T_{min} Patterns.

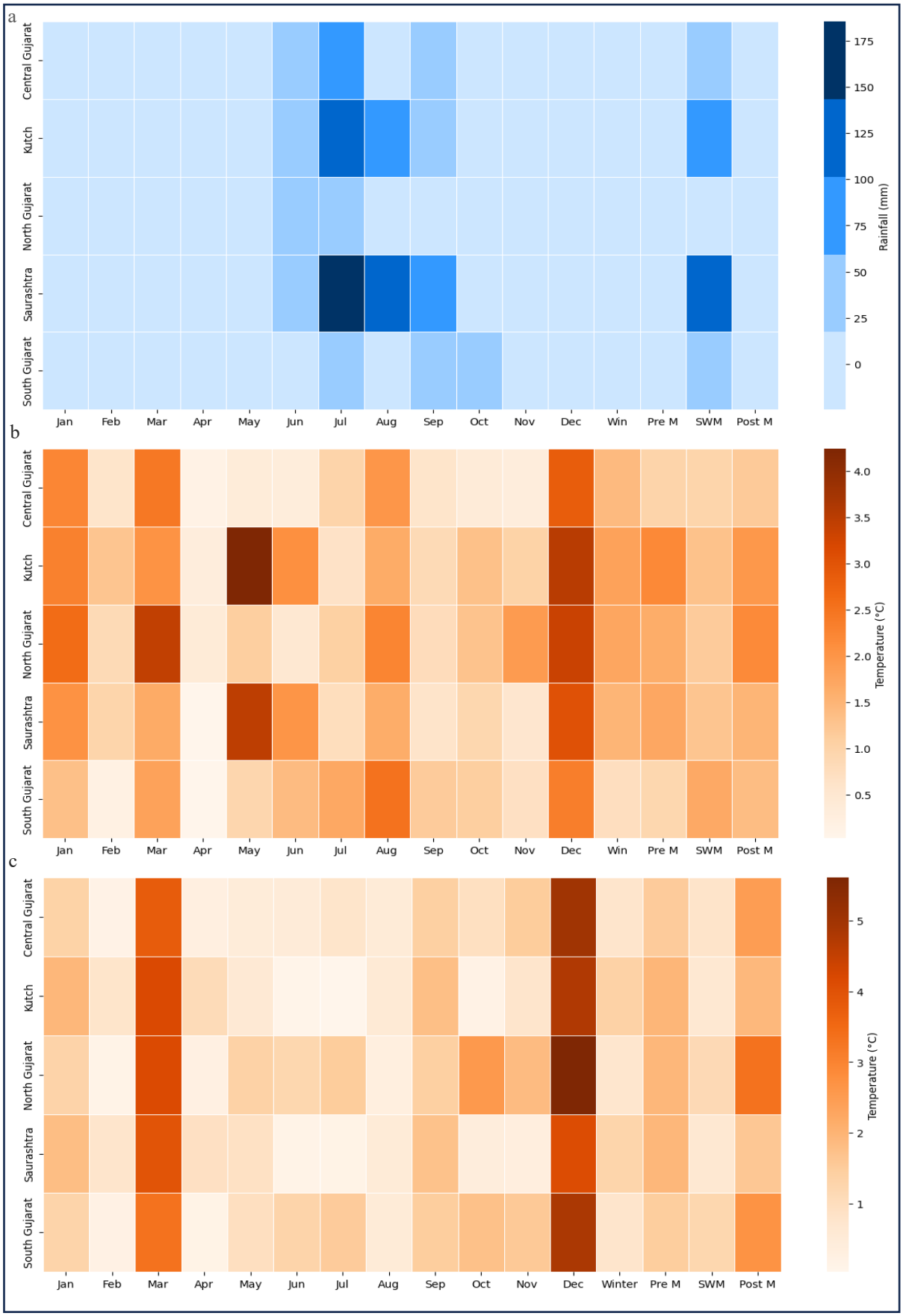


Figure 6.8 Comparative Analysis of Mean Monthly Differences in Forecasted and Observed (a) Rainfall, (b) T_{max}, and (c) T_{min} Across Physiographic Regions of Gujarat.

Figures 6.5 to 6.7 show mean monthly forecasts and deviations from 60-year observed averages for rainfall, T_{max}, and T_{min} across Gujarat's physiographic regions. Figure 6.8 compares mean monthly differences between observed and forecasted variables. The projected increase in SWM rainfall (Figure 6.5), which accounts for over 96% of the region's annual rainfall, reveals an increase of 69.3% for Saurashtra, followed by Kutch at 49.1% from the 60-year observed mean. In contrast, Central Gujarat, North Gujarat, and South Gujarat are projected to experience a moderate increase of 9.6%, 7.2%, and 4.7%, respectively. Although winter and pre-monsoon seasons account for less than 2% of the total observed annual rainfall, they exhibit high variability across distinct regions. Kutch is projected to experience a substantial 85.6% increase in winter rainfall, followed by Saurashtra, and North Gujarat with increases of 42.7%, and 32.4%, respectively. Conversely, South Gujarat is expected to see a decline in winter rainfall of 14.5%. In the pre-monsoon season, most regions are expected to see a decrease in rainfall, with Kutch and Saurashtra as exceptions, projected an increase of 347.1% and 58.2%, respectively. Meanwhile, projections for post-monsoon rainfall, which contributes 3-4% to the annual total, indicate that South Gujarat will witness an 87.4% increase in post-monsoon rainfall, followed by Kutch at 39% and Saurashtra at 29.8%. In contrast, Central Gujarat and North Gujarat are expected to experience declines of 35.9% and 15.3%, respectively. While Figure 8 shows the most significant mean monthly differences between observed and forecast rainfall in July, particularly in Saurashtra. Table 6.3 shows the statistics of changes in mean observed and forecast rainfall across regions, highlighting absolute and percentage variations.

Table 6.4 Comparative Analysis of Observed and Forecasted Mean Rainfall Across Regions.

Region	Season	Observed Mean (1961-2020)	Forecast Mean (2021-2050)	Changes (mm)	Changes (%)
Central Gujarat	Winter	1.14	1.14	0.00	0.00
	Pre-Monsoon	7.41	6.53	-0.88	-11.88
	SWM	818.44	897.03	78.59	9.60
	Post-Monsoon	26.28	16.83	-9.45	-35.96
	Annual	853.28	921.50	68.22	8.00
Kutch	Winter	1.81	3.36	1.55	85.64
	Pre-Monsoon	5.71	25.53	19.82	347.11
	SWM	409.12	610.18	201.06	49.14
	Post-Monsoon	20.04	27.87	7.83	39.07
	Annual	436.67	666.94	230.27	52.73
North Gujarat	Winter	2.50	3.31	0.81	32.40
	Pre-Monsoon	7.91	3.39	-4.52	-57.14
	SWM	696.70	747.10	50.40	7.23
	Post-Monsoon	22.10	18.70	-3.40	-15.38
	Annual	729.21	772.58	43.37	5.95

Saurashtra	Winter	0.96	1.37	0.41	42.71
	Pre-Monsoon	5.20	8.23	3.03	58.27
	SWM	640.57	1084.86	444.29	69.36
	Post-Monsoon	28.88	37.49	8.61	29.81
	Annual	675.62	1131.90	456.28	67.54
South Gujarat	Winter	0.96	0.82	-0.14	-14.58
	Pre-Monsoon	7.53	8.06	0.53	7.04
	SWM	1361.55	1426.26	64.71	4.75
	Post-Monsoon	40.73	76.36	35.63	87.48
	Annual	1410.76	1511.52	100.76	7.14

Rainfall projections for Gujarat indicate varying increases across regions, with Central Gujarat expecting an 8% increase from the observed annual mean. Saurashtra is projected to experience a substantial increase of 67.5%, followed by Kutch with a 52.7% increase, stressing the need for strategic planning to accommodate these changes. Meanwhile, South Gujarat and North Gujarat are expected to have a moderate increase of 7.1% and 5.9%, respectively, indicating a positive shift compared to the observed annual mean. This upward shift in rainfall patterns carries significant implications for water management and agriculture, necessitating adaptive strategies to ensure sustainable water resource allocation.

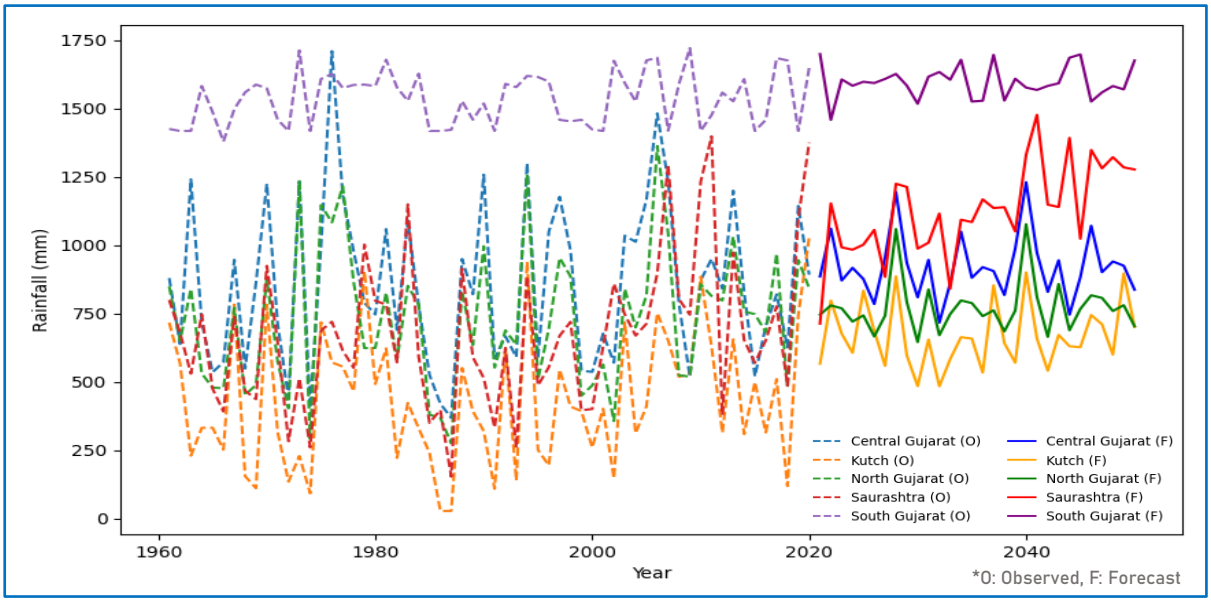


Figure 6.9 Average Annual Rainfall for Observed and Forecasted Periods Across Regions.

Figure 6.9 shows the average annual rainfall variability across different regions, comparing observed data with forecasted projections. South Gujarat is projected to consistently maintain the highest rainfall levels throughout the next three decades. In contrast, Kutch, historically a region with low rainfall, is projected to experience increasing rainfall. Overall, the forecast suggests a general trend of increasing rainfall across regions, with substantial increases projected in Saurashtra and Kutch, exhibiting potential shifts in regional rainfall patterns.

Additionally, Figures 6.6 and 6.7 show projections for mean monthly T_{max} and T_{min} across regions, indicating an increase compared to the observed mean. The most substantial increase in T_{max} is expected during the winter season, with Kutch projected to experience an increase of 1.8°C compared to the observed mean. North Gujarat, Saurashtra, and Central Gujarat are also projected to experience an increase of 1.7°C, 1.5°C, and 1.4°C, respectively, during winter. This further suggests an increase in the frequency of warm days during winter months. During the pre-monsoon, North Gujarat and Kutch are expected to see increases of 1.0°C and 0.7°C, respectively. In the post-monsoon, Kutch and Saurashtra are projected to experience an increase of 0.7°C and 0.5°C, respectively. Although the projected changes in T_{max} during the pre-monsoon and post-monsoon periods are less pronounced than those in winter, they still raise concerns across the regions due to extended periods of higher temperatures beyond the usual patterns. Meanwhile, the projections indicate a maximum increase in T_{max} of 0.9°C in South Gujarat during the SWM season, followed by increases of 0.3°C in Kutch and Saurashtra.

For T_{min}, the most substantial changes compared to the observed mean are expected during the post-monsoon season, with a projected increase of 3.2°C in North Gujarat, followed by 2.6°C in South Gujarat, 2.4°C in Central Gujarat, and 1.8°C in Kutch. During the pre-monsoon and winter seasons, Kutch and Saurashtra are expected to experience increases of 1.9°C and 1.3°C, respectively. Additionally, South Gujarat is projected to experience the maximum increase of 1.2°C from the observed mean during the SWM season. However, while South Gujarat shows relative stability in projected T_{max}, it is expected to experience immense increases in T_{min} across seasons. Meanwhile, Figure 6.8b, c shows the maximum mean monthly differences between observed and forecast T_{max} and T_{min} occur in December, followed by March (for T_{min}) and May (for T_{max}). Table 6.4 shows the observed and forecasted changes in T_{max} and T_{min} across the physiographic regions of Gujarat.

Table 6.5 Comparative Analysis of Observed and Forecasted Mean T_{max} and T_{min} Across Regions.

Region	Season	Observed Mean (1961-2020)		Forecast Mean (2021-2050)		Changes (°C)		Changes (%)	
		T_{max}	T_{min}	T_{max}	T_{min}	T_{max}	T_{min}	T_{max}	T_{min}
Central Gujarat	Winter	30.13	13.79	31.60	14.58	1.47	0.79	4.88	5.76
	Pre-Monsoon	38.32	22.98	38.92	24.51	0.60	1.53	1.57	6.67
	SWM	33.35	25.01	33.61	25.82	0.26	0.81	0.78	3.26
	Post-Monsoon	32.86	17.95	33.14	20.37	0.28	2.42	0.85	13.46
	Annual	33.95	20.89	34.23	22.26	0.28	1.37	0.82	6.55
Kutch	Winter	28.30	14.08	30.12	15.47	1.82	1.39	6.43	9.85
	Pre-Monsoon	34.78	22.74	35.52	24.70	0.74	1.96	2.13	8.62

	SWM	32.73	25.94	33.2	26.56	0.47	0.62	1.44	2.39
	Post-Monsoon	31.92	18.68	32.65	20.52	0.73	1.84	2.29	9.85
	Annual	32.32	21.37	32.95	22.74	0.63	1.37	1.95	6.39
North Gujarat	Winter	27.87	11.63	29.65	12.39	1.78	0.76	6.39	6.56
	Pre-Monsoon	37.07	21.77	38.11	23.70	1.04	1.93	2.81	8.88
	SWM	33.04	24.20	33.43	25.32	0.39	1.12	1.18	4.62
	Post-Monsoon	31.35	15.88	31.59	19.14	0.24	3.26	0.77	20.50
	Annual	32.78	19.44	33.29	21.21	0.51	1.77	1.56	9.13
Saurashtra	Winter	29.31	14.06	30.86	15.39	1.55	1.33	5.29	9.46
	Pre-Monsoon	35.75	22.50	36.25	24.43	0.5	1.93	1.40	8.57
	SWM	32.67	25.43	33.06	26.05	0.39	0.62	1.19	2.45
	Post-Monsoon	32.49	18.63	33.02	20.18	0.53	1.55	1.63	8.33
	Annual	32.85	21.12	33.34	22.40	0.49	1.28	1.49	6.07
South Gujarat	Winter	30.45	13.52	31.00	14.31	0.55	0.79	1.81	5.85
	Pre-Monsoon	38.10	22.54	38.34	24.02	0.24	1.48	0.63	6.56
	SWM	32.44	24.26	33.34	25.47	0.90	1.21	2.77	5.00
	Post-Monsoon	32.46	17.20	32.58	19.85	0.12	2.65	0.37	15.39
	Annual	33.54	20.30	33.62	21.84	0.08	1.54	0.24	7.60

Additionally, projections indicate an increase in average annual temperatures across regions by the mid-21st century, with Kutch and Saurashtra expected to experience the most substantial increases. Figure 6.10 shows mean annual temperature projections over the next 30 years, revealing that Kutch is projected to exceed the 1°C threshold above the 1961-2020 observed mean in 11 instances, including 8 occurrences after 2040. This alarming trend poses serious risks for the region. Saurashtra and North Gujarat are also projected to surpass this threshold in 3 instances each, while Central Gujarat is expected to exceed it once. In contrast, South Gujarat shows no projections for increases exceeding the 1°C threshold above the observed mean, indicating relative stability in that region.

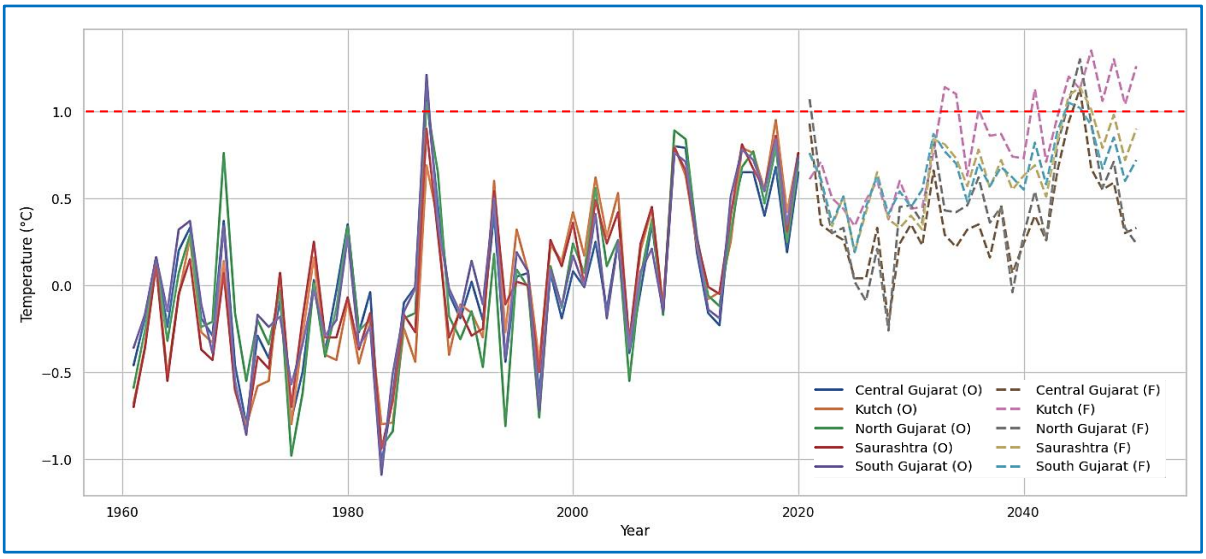


Figure 6.10 Average Annual Temperature Anomalies Compared to the 1961-2020 Observed Mean.

6.3.2 Assessment of Hydrological Excess and Deficit

Projections indicate increases in T_{max}, T_{min}, and rainfall across the physiographic regions of Gujarat; however, these changes do not ensure that a region will experience a water surplus or deficit. The relationship between temperature, rainfall, and water availability is complex, and an increase in rainfall may not necessarily lead to a surplus if PET also increases concurrently. Figure 6.11a shows an annual PET cycle, steadily increasing from January and reaching its apex during the pre-monsoon months, particularly in May, followed by a sharp decline with the onset of the SWM. December records the lowest PET across regions, primarily due to relatively low temperatures and decreased evaporation rates during winter. This pattern reflects the climatic rhythm of regions, with increased PET demands leading up to the onset of the SWM, followed by reduced levels during the SWM months. Additionally, Figure 6.11a shows a comparative analysis of the mean monthly observed and forecasted PET cycle, highlighting variations across regions. The projected changes reveal distinct PET patterns in each region, emphasizing the disparities in how each physiographic region responds to seasonal shifts.

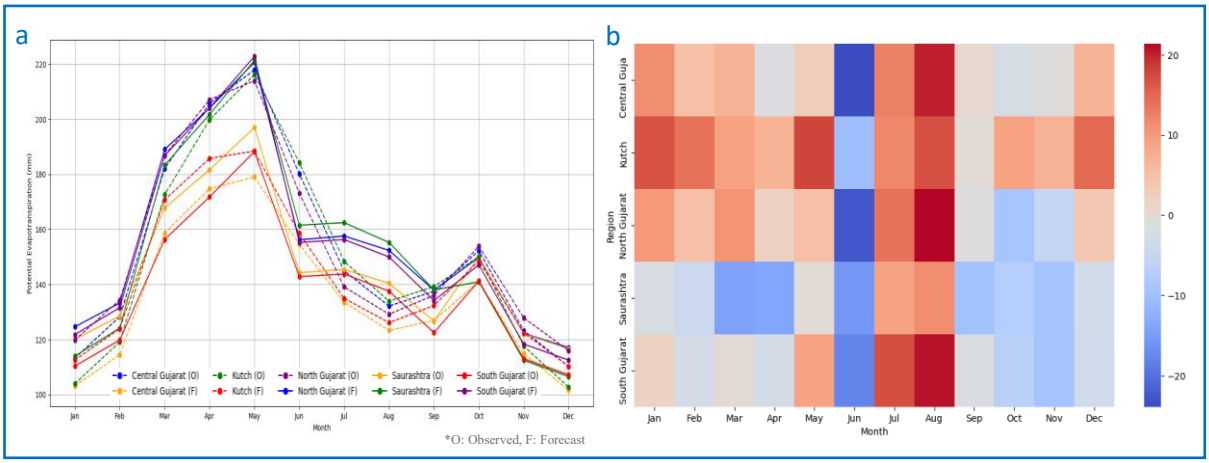
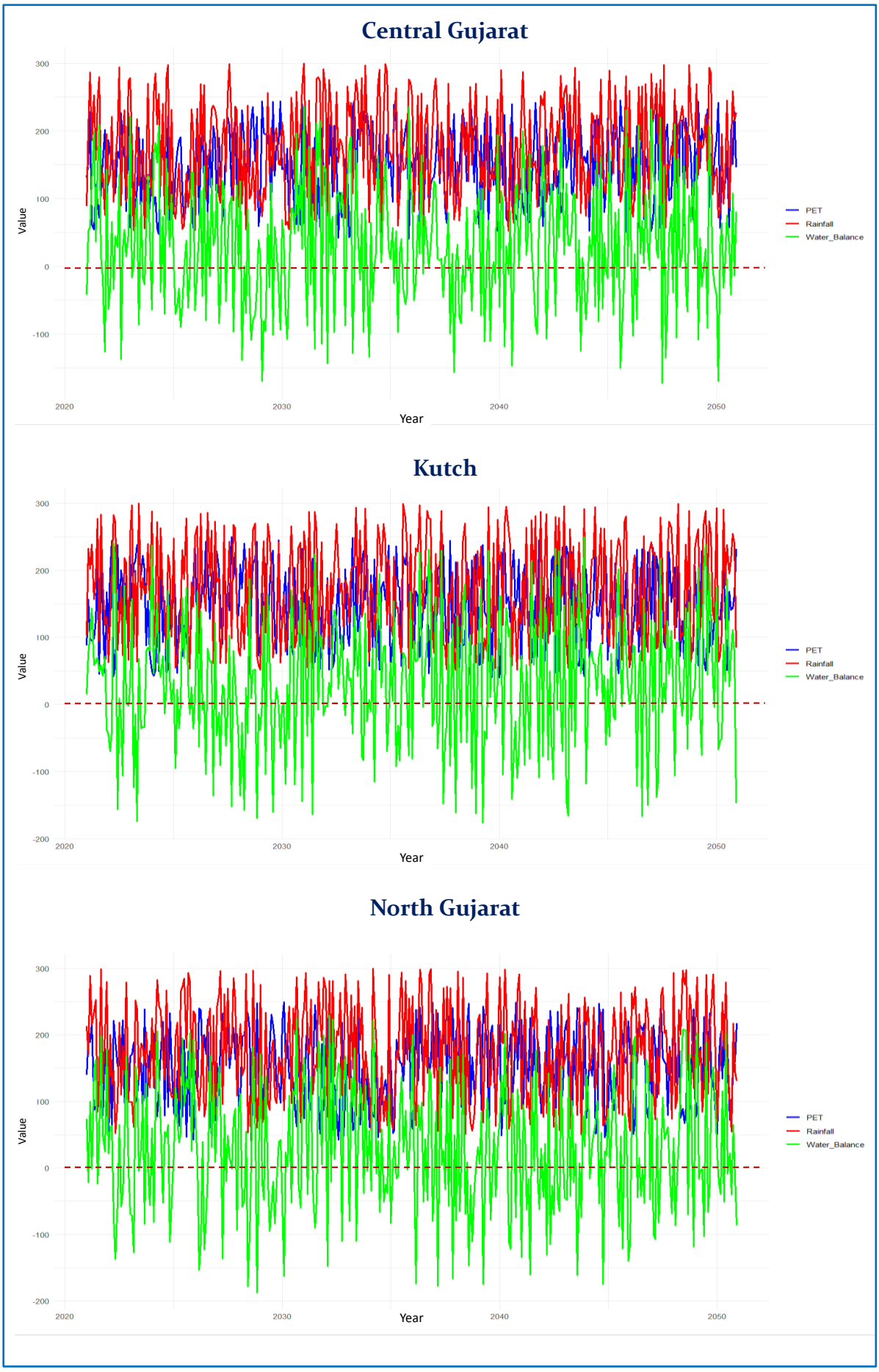


Figure 6.11 a) Annual Cycle of Observed and Forecasted PET, b) Mean Monthly Differences Between Forecasted and Observed PET Across Physiographic Regions of Gujarat.

However, Figure 6.11b shows the mean monthly differences between observed and forecasted PET across regions, indicating that the most substantial negative differences were projected in June across regions, particularly pronounced in Central Gujarat and North Gujarat, followed by Saurashtra and South Gujarat during the post-monsoon season. In contrast, the maximum positive changes were observed in August, followed by July, highlighting a period of increased PET across regions. Additionally, Figure 6.12 exhibits the projected water balance for 2021 to 2050, showing the difference between forecasted rainfall and PET, while incorporating forecasted T_{min} and T_{max}. The assessment highlights crucial shifts in water availability, which could substantially impact water resource management and agricultural practices.


Central Gujarat
300
200
100
0
-100
Value
2020
2030
2040
2050
Year
PET
Rainfall
Water_Balance
Kutch
300
200
100
0
-100
-200
Value
2020
2030
2040
2050
Year
PET
Rainfall
Water_Balance
North Gujarat
300
200
100
0
-100
-200
Value
2020
2030
2040
2050
Year
PET
Rainfall
Water_Balance

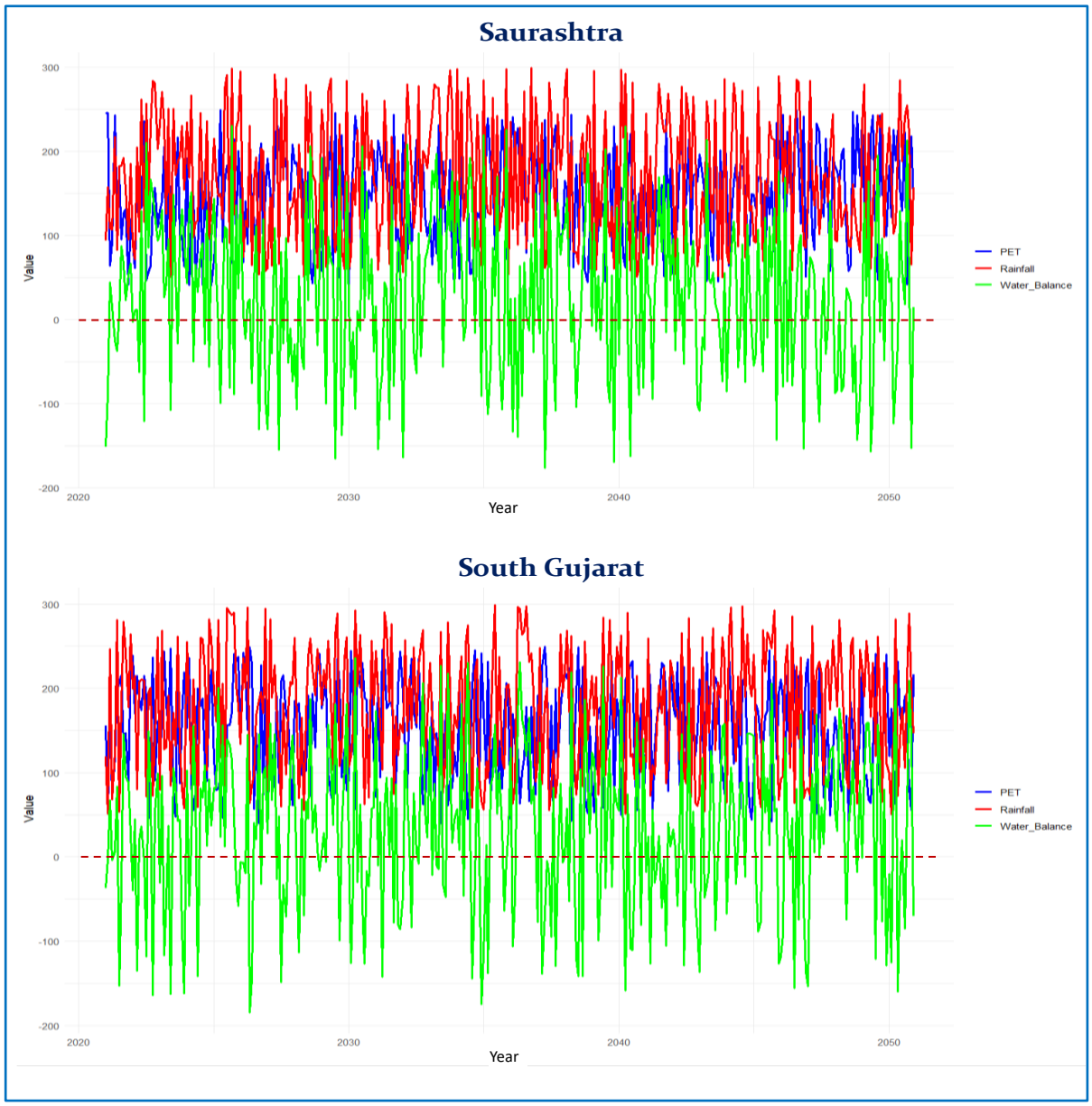


Figure 6.12 Projected Water Balance (2021-2050) Across Physiographic Regions of Gujarat.: Assessment of Potential Surpluses and Deficits.

Figure 6.12 shows the seasonal shifts in water balance forecasted for the distinct physiographic regions of Gujarat from 2021 to 2050, highlighting both water surplus and deficit periods. The analysis reveals that while the regions experience a surplus during the SWM, especially in July and August, significant regional disparities exist. Kutch and North Gujarat experience limited surpluses, mainly confined to July and August, and endure prolonged deficits, particularly beyond the SWM months. Central Gujarat shows a restricted surplus, with positive water balance occurring only in July, August, and September. In contrast, Saurashtra and South Gujarat benefit from a more consistent surplus throughout the SWM, enhancing their resilience during dry periods. These patterns emphasize the need for region-specific water management

strategies. Additionally, Figure 6.13 exhibits the mean monthly water balance forecast across distinct physiographic regions from 2021 to 2050, showing the variations in water surplus and deficit throughout the year and crucial insights into seasonal trends.

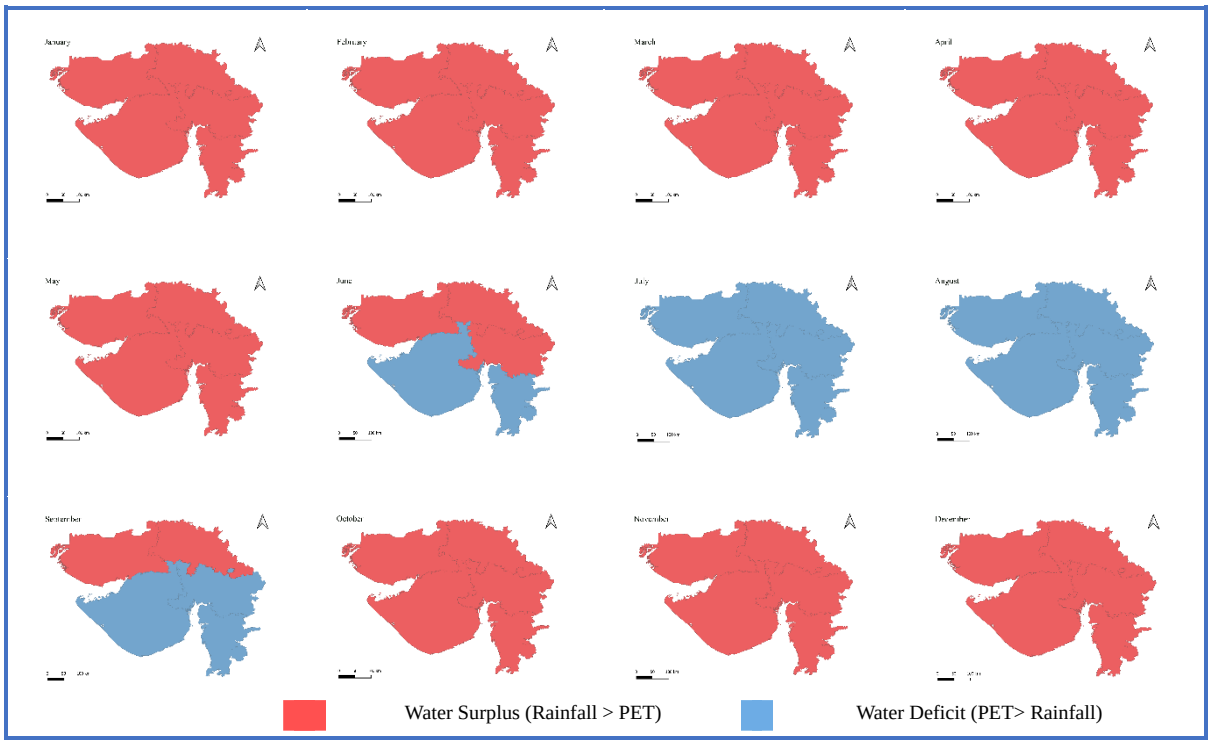


Figure 6.13 Mean Monthly Water Balance Forecast (2021-2050) for Gujarat's Physiographic Regions: Analysis of Water Surpluses and Deficits.

Figure 6.13 shows that, apart from the SWM season (JJAS), there are no projected water surpluses across regions. Projections show that during the SWM, South Gujarat and Saurashtra will consistently experience water surpluses due to abundant rainfall. In contrast, Kutch and North Gujarat are expected to have surpluses only in July and August. However, the variation is not invariant; it changes over the forecasted period, exhibiting shifts in water balance patterns as climatic factors, such as rainfall patterns and temperature, evolve. This dynamic variability highlights the necessity for adaptive water management strategies that can accommodate the projected shifts in water availability, ensuring resilience against potential shortages in drier months and supporting sustainable agricultural practices across the regions. Figure 6.14 shows the decadal changes in projected water balance across distinct regions, emphasizing the mean monthly variations in water surpluses and deficits. The findings reveal marked changes, especially in Saurashtra and Kutch, where water balance patterns fluctuate substantially, potentially challenging future water availability. In contrast, South Gujarat is expected to experience relatively low changes, showing a more stable water balance in the coming decades.

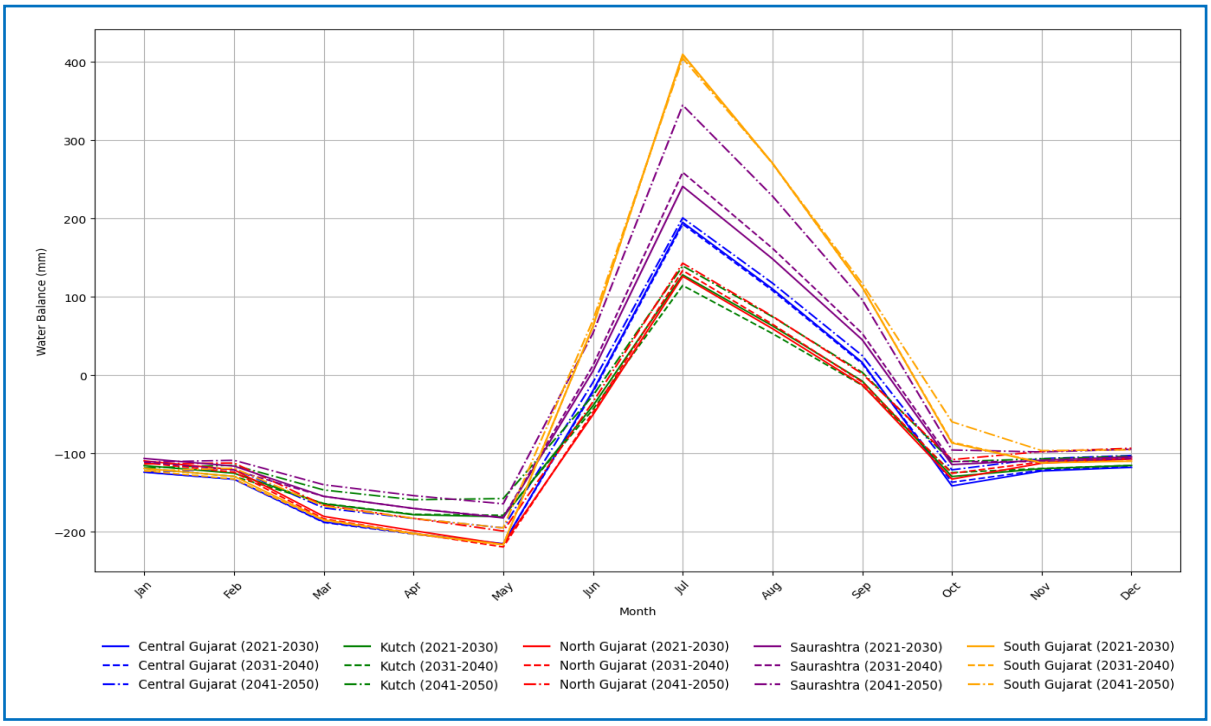


Figure 6.14 Decadal Changes in Projected Water Balance Across Physiographic Regions of Gujarat: Mean Monthly Analysis of Surpluses and Deficits

The findings further highlight the need for adaptive water management strategies that account for regional differences, ensuring that regions experiencing higher variability, such as Saurashtra and Kutch, are better equipped to manage the impacts of these shifts while maintaining water security across the state. However, the decadal assessment shows a consistent surplus in July, followed by August across regions. The disparity mainly exists for September, where Kutch and North Gujarat are projected to transition from a negative water balance during the first two decades of the forecast to a positive water balance in the latter decade. This shift indicates a potential improvement in water availability for these regions, emphasizing the importance of monitoring and adjusting water management practices to respond to evolving climatic conditions and ensure sustainable resource allocation.

6.4 Conclusion

The study employs the SARIMA model, optimized based on AIC and log-likelihood measures, to effectively forecast rainfall, T_{max}, and T_{min} across the distinct physiographic regions of Gujarat. Utilizing 60 years of observed data, the dataset was divided into training sets of 50 years and testing sets of 10 years, facilitating a thorough evaluation of the model's predictive accuracy. The robustness of the model was assessed through performance metrics such as MAE and RMSE, which quantify the model's forecasting errors. The findings indicate a substantial increase in SWM rainfall across regions, with Saurashtra and Kutch experiencing increases of

69.3% and 49.1%, respectively, from the observed mean. Although the winter and pre-monsoon seasons contribute negligible, they exhibit high variability. Kutch, Saurashtra, and North Gujarat are projected to experience increases of 85.6%, 42.7%, and 32.4%, respectively, during winter, while South Gujarat is projected to see a decline of 14.5%. During the pre-monsoon season, most regions are expected to experience a decline from the observed mean, except for Kutch and Saurashtra, with projected increases of 347.1% and 58.2%, respectively. Meanwhile, projections for post-monsoon rainfall indicate that South Gujarat will experience an 87.4% increase, followed by Kutch at 39% and Saurashtra at 29.8%. In contrast, Central Gujarat and North Gujarat are expected to experience declines of 35.9% and 15.3%, respectively. Despite seasonal variations, annual rainfall projections indicate overall increases across regions. Saurashtra and Kutch are expected to see substantial rises of 67.5% and 52.7%, respectively, while Central Gujarat, North Gujarat, and South Gujarat are projected to experience more moderate increases of 8%, 7.1%, and 5.9%, respectively, compared to their observed annual means. Meanwhile, the most significant mean monthly differences between observed and forecast rainfall in July occur in Saurashtra, followed by Kutch. This shift in rainfall patterns carries important implications for water management and agriculture, necessitating adaptive strategies to ensure sustainable water resource allocation.

Additionally, the temperature projections indicate an increase in T_{min} and T_{max} across regions, with the most substantial increases in T_{max} expected during the winter season in Kutch at 1.82°C, followed by North Gujarat at 1.7°C, and Saurashtra at 1.5°C, from the observed mean, suggesting frequent warm winter days. Although projected T_{max} increases during the pre-monsoon and post-monsoon periods are less pronounced, they signal extended periods of elevated heat and extremes. Meanwhile, South Gujarat is projected to experience the maximum T_{max} increase during the SWM season. For T_{min}, the primary concern lies in the projected increase during the post-monsoon period, particularly in North Gujarat, where T_{min} is expected to increase by 3.2°C, followed by South Gujarat at 2.6°C and Central Gujarat at 2.4°C from the observed mean. Kutch and Saurashtra are projected to experience increases of 1.3°C and 1.9°C during the winter and pre-monsoon seasons, respectively. South Gujarat is expected to experience an increase in T_{min} of 1.2°C during the SWM season compared to the observed mean. Meanwhile the results indicate maximum mean monthly differences between observed and forecast T_{max} and T_{min} occur in December, followed by March for T_{min} and May for T_{max}. Additionally, projections indicate an increase in average annual temperatures across regions from observed at least 0.5-0.8°C by the mid-21st century, with Kutch and Saurashtra are

expected to exceed the 1°C threshold for mean annual temperatures frequently by mid-century, highlighting serious climate risks for these regions. The combined impact of increasing temperatures and altered rainfall patterns highlights the need for urgent, region-specific strategies to mitigate these effects. The alarming trend of increasing temperatures demands immediate and robust action to address the growing climate risks facing the region.

Further, to identify the water balance and potential water deficits or surpluses, forecasted rainfall was subtracted from PET, incorporating projections for T_{min} and T_{max}. The findings indicate that while Gujarat generally experiences a water surplus during the SWM months, significant regional disparities persist. Moreover, the projections reveal evolving water balance patterns across regions. While July and August consistently show a surplus, September shows a disparity, particularly in Kutch and North Gujarat, which are projected to shift from a negative water balance in the early decades to a positive one in the latter, highlighting the urgent need for region-specific water management strategies to address emerging challenges effectively.