COMPARISON OF PARAMETRIC, NON-PARAMETRIC AND MACHINE LEARNING TECHNIQUES FOR TEMPERATURE TREND IDENTIFICATION

Godwin Oluwadare Wilson[1,*], Abdulhameed Ado Osi[1], Dauda Usman[2], Jeremiah Ayock Ishaya[3]

[1]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil 713281, Nigeria: aaosi@kustwudil.edu.ng

[2]Department of Mathematics and Statistics, Faculty of Natural and Applied Sciences, Umaru Musa Yar'adua University Katsina, Nigeria: dauda.usman@umyu.edu.ng

[3]Faculty of Computational Science and Informatics, Academic City University, Accra - Ghana: jeremiah.ishaya@acity.edu.gh

***Correspondence author's email:** godwincharisma@gmail.com

ABSTRACT

Understanding local climate trends is critical for effective adaptation planning, particularly in rapidly urbanizing regions like West Africa. This study examines temperature changes in Kano, Nigeria, from 1993 to 2023 using a multi-method approach. We combined parametric techniques (linear regression, ARIMA), non-parametric tests (Mann-Kendall, Sen's Slope), and machine learning models (Random Forest, SVM) to analyze monthly maximum temperatures, with careful attention to seasonal effects. Our analysis identified a gradual warming trend of 0.16°C per decade—significantly slower than the 0.3°C regional average for West Africa. More strikingly, seasonal patterns accounted for over 80% of temperature variability, with April temperatures consistently peaking 10-15°C above January lows. The research highlights how simple models with categorical month variables (achieving R^2 = 0.88) can outperform both basic linear approaches and complex algorithms in tropical climate studies. These findings have important practical implications: they suggest that local factors may be moderating temperature rise in Kano, and they underscore the need for climate adaptation strategies that prioritize managing seasonal extremes rather than just long-term warming trends. The study demonstrates the value of combining multiple analytical methods to produce robust, actionable insights for urban climate planning in tropical regions.

Keywords: *Seasonal Kendall Test; Trend analysis; Maximum temperature; Time series; Comparative Study; Machine Learning Techniques.*

1.0 Introduction

Urban climate dynamics in West Africa pose significant challenges for sustainable development, where rising temperatures exacerbate pressures from rapid urbanization (UNHSP, 2022). Despite global documentation of warming trends, regional variations particularly urban microclimates influenced by land use changes and heat islands (Stewart et al., 2012) - remain understudied. Kano, Nigeria, exemplifies this gap as a fast-growing city where temperature variability has received limited systematic analysis (Abubakar et al., 2020), despite its potential to reveal broader semi-arid urban climate trends.

Current research presents conflicting projections, ranging from uniform regional warming (Niang et al., 2022) to localized microclimate moderation (Adebayo, 1991), underscoring the need for empirical validation. The research landscape further lacks comprehensive studies of tropical cities' pronounced seasonal variability (Diem et al., 2022), which often exceeds long-term warming trends in climatic impact.

This study analyzes Kano's temperature data (1993-2023) through an integrated parametric, non-parametric, and machine learning framework, advancing beyond conventional linear

ISSN NUMBER: 1116-249X

trend approaches (Fashae et al., 2020). Key findings include a slower-than-regional warming rate (0.16°C/decade) suggesting local moderating factors, and the dominance of seasonal extremes in temperature variability—critical insights for climate resilience planning in tropical cities.

2.0 Materials and Methods

2.1 Methodology

This study integrated parametric, non-parametric, and machine learning approaches to analyze Kano's temperature trends (1993-2023). Figure 1 illustrates the methodological workflow, with Figure 2 showing the study area's geographical context.

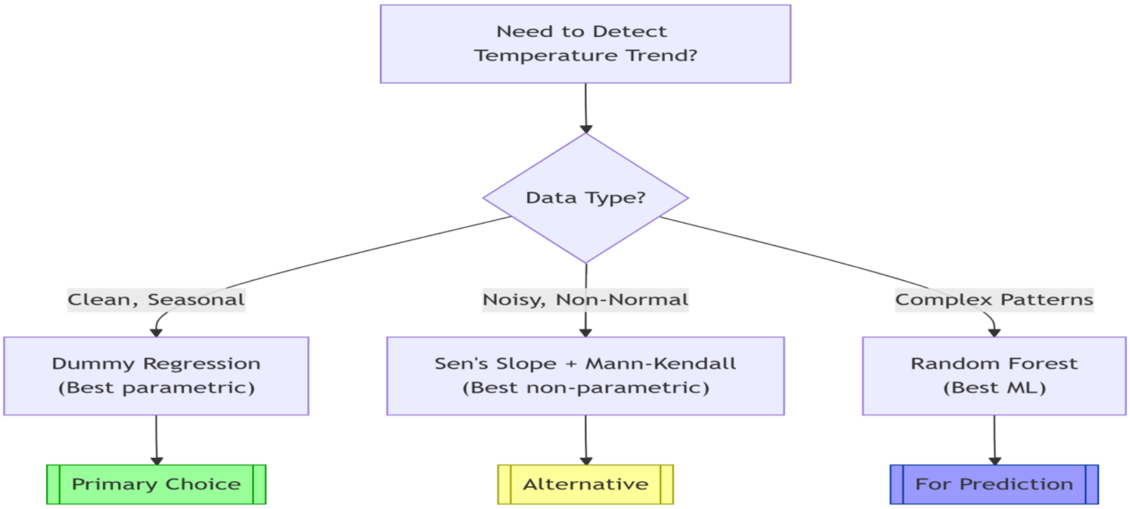

Figure 1

2.2 Study Area Description

Kano Metropolis (11.9964°N, 8.5167°E), at 488m elevation in Nigeria's Sudano-Sahelian zone, spans 499 km² with >3.6 million inhabitants (NBS, 2022).

.

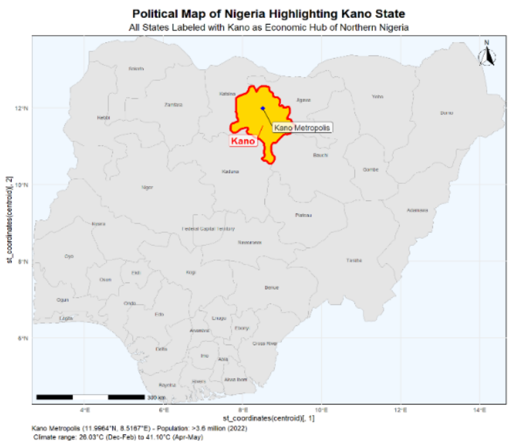

Figure 2

The climate shows extreme seasonal variation, from 26.03°C (Dec-Feb) to 41.10°C (Apr-May) (Ojo et al., 2021), with urban expansion increasing impervious surfaces to 68% and reducing vegetation to 12% since 2000 (Balogun et al., 2020).

ISSN NUMBER: 1116-249X

2.3 Data Collection and Preprocessing

Monthly maximum temperatures (1993-2023) were obtained from Nigerian Meteorological Agency (NiMet). The dataset underwent quality control via Pettitt's homogeneity test (p = 1.0) and Tukey's outlier detection (no anomalies found). Season-Trend decomposition using Loess was used for imputation of missing values (0.1% of records), with data partitioned into training (1993-2015) and validation (2016-2023) sets using month-stratified sampling.

2.4. Parametric Methods

2.4.1. Linear Regression Analysis

Linear regression quantified temperature-predictor relationships (Montgomery et al., 2012), offering interpretable coefficients and confidence intervals (Kutner et al., 2004). The model expressed maximum temperature as:

Basic Linear Model:

$$T_{\max} = \beta_0 + \beta_1(\text{Year}) + \beta_2(\text{MonthNum}) + \varepsilon, \tag{1}$$

Categorical Month Model:

$$T_{\max} = \beta_0 + \beta_1(\text{Year}) + \sum_{k=2}^{12} \beta_k I(\text{Month=}k) + \varepsilon, \tag{2}$$

The categorical formulation replaces $\beta_2(\text{MonthNum})$ with monthly coefficients $\beta_k (k = 2-12)$ for indicator variables $I(\text{Month=}k)$ I, showing temperature differences versus January (reference month) while controlling for annual trends (β_1). Both models retain the intercept (β_0) and error term (ε).

Validation Framework:

Model assumptions were verified using:

1. Shapiro-Wilk test for normality (Shapiro et al., 1965):

$$W = \frac{\left(\sum_{i=1}^{n} a_i R_i\right)^2}{\sum_{i=1}^{n}\left(R_i - \overline{R}\right)^2}, \tag{3}$$

2. Breusch-Pagan test for heteroscedasticity (Breusch et al., 1979):

$$BP = nR^2, \tag{4}$$

3. Durbin-Watson statistic for autocorrelation (Durbin et al., 1951):

$$DW = \frac{\sum_{t=2}^{T} (e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}, \tag{5}$$

Visual diagnostics (Q-Q plots, ACF/PACF) complemented these tests. The categorical specification enables month-specific intercepts with unified annual trends (Sylla et al., 2016).

2.4.2. ARIMA Modeling

The ARIMA methodology (Box et al., 2008) modeled temperature time series through autoregressive components, differencing, and moving average terms. The seasonal ARIMA formulation

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D T_t = \theta(B)\Theta(B^s)\varepsilon_t, \tag{6}$$

ISSN NUMBER: 1116-249X

incorporates non-seasonal (ϕ) and seasonal coefficients (θ) with backshift operator B (Brockwell et al., 2016; Shumway et al., 2017). Stationarity was assessed via augmented Dickey-Fuller (Equation 7)

$$\Delta \mathrm{T}_i = \alpha + \beta_t + \gamma T_{t-1} + \sum_{i=1}^{p} \delta_i \Delta T_{t-i} + \varepsilon_t \tag{7}$$

and KPSS tests (Said et al., 1984; Kwiatkowski et al., 1992), with maximum likelihood estimation determining parameters (Hamilton, 1994).

$$\ell(\phi, \theta, \Phi, \Theta \setminus T) = -\frac{n}{2} \ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2} \sum_{t=1}^{n} \varepsilon_t^2 \tag{8}$$

Model selection used corrected AIC (Hurvich et al., 1989),

$$AICc = -2\ln \ell + 2k + \frac{2k(k+1)}{n-k-1} \tag{9}$$

while residuals were evaluated with Ljung-Box and Jarque-Bera tests;

$$Q = n(n+2)\sum_{k=1}^{h} \frac{\rho_k^2}{n-k} \tag{10}$$

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \sim \chi_2^2 \tag{11}$$

(Ljung et al., 1978; Jarque et al., 1980). Forecasting employed recursive formulation

$$T_t + k \setminus t = \phi_1 T_t + k - 1 \setminus t + \ldots + \phi_p T_t + k - p \setminus t + \theta_1 \varepsilon_t + k - 1 + \ldots + \theta_q \varepsilon_{t+k-q} \tag{12}$$

(Box et al., 1968) with prediction variance from ψ-weights (Wei et al., 2006).

$$Var\left(e_{t+k}\right) = \sigma^2 \sum_{i=0}^{k-1} \psi_i^2 \tag{13}$$

2.5. Non-parametric Methods

2.5.1. Mann-Kendall Trend Analysis

The Mann-Kendall test was employed as a robust non-parametric approach to identify monotonic trends (Mann et al., 1945), offering distribution-free advantages and outlier resistance (Kendall et al., 1975). The method analyzes observation ranks rather than values, making it suitable for non-normal environmental data (Hirsch et al., 1982). The test statistic S is calculated through pairwise comparisons,

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(x_j - x_i) \tag{14}$$

where xi and xj are values at times i and j, with the sign function defined in Equation as,

$$\mathrm{sgn}(x_j - x_i) = \begin{cases} +1 & \text{if } x_j > x_i \\ 0 & \text{if } x_j = x_i \\ -1 & \text{if } x_j < x_i \end{cases} \tag{15}$$

For $n \geq 10$, S follows a normal distribution under the null hypothesis, with variance adjusted for ties,

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{g=1}^{G} t_g(t_g - 1)(2t_g + 5)}{18} \tag{16}$$

where G represents tied groups and t_g their sizes (Kendall et al., 1975). The standardized statistic Z is computed as shown in

$$Z = \begin{cases} \frac{s-1}{\sqrt{\mathrm{var}(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{s+1}{\sqrt{\mathrm{var}(S)}} & \text{if } S < 0 \end{cases} \tag{17}$$

Serial independence was verified using the Ljung-Box Q-test

$$Q = n(n+2)\sum_{k=1}^{h} \frac{\rho_k^2}{n-k} \tag{18}$$

(Ljung et al., 1978). For autocorrelated data (p<0.05), pre-whitening employed the lag-1 coefficient ρ_1

$$\rho_1 = \frac{\sum_{t=2}^{n}\left(x_t - \bar{x}\right)\left(x_{t-1} - \bar{x}\right)}{\sum_{t=1}^{n}\left(x_t - \bar{x}\right)^2}, \tag{19}$$

(Hirsch et al., 1982), yielding the adjusted series

$$x_t' = x_t - \rho_1 x_{t-1}, \tag{20}$$

Seasonality was addressed via the Seasonal Mann-Kendall variant

$$S_{total} = \sum_{k=1}^{12} S_k, \tag{21}$$

combining monthly S statistics (Hipel et al., 1994), with variance accounting for within- and between-month covariances.

2.5.2. Theil-Sen Estimator

The Theil-Sen estimator (Theil, 1950) quantified temperature trends through median pairwise slopes,

$$\beta = \mathrm{Median}\left[\frac{x_j - x_i}{j - i}\right] \qquad \forall i < j, \tag{22}$$

offering superior outlier resistance without normality assumptions (Sen, 1968; Wilcox, 2017). Slope β was calculated from all data pairs (n=465 for 1993-2023) (Helsel et al., 1992), with intercept α

$$\alpha = \mathrm{Median}\left[x_j - \beta_i\right], \tag{23}$$

completing the robust model

$$x_i = \alpha + \beta_i \tag{24}$$

Confidence intervals

ISSN NUMBER: 1116-249X

$$CI\beta = \left[\beta(k), \beta_{(N-k+1)}\right] \tag{25}$$

used Siegel's method (Siegel, 1982; Hollander et al., 1999), validated via Conover's test (Conover, 1999) and 1000 bootstrap replicates (Efron et al., 1993). Seasonal analysis aggregated monthly slopes

$$\beta_{total} = \text{Median}\left[\beta_k\right] \tag{26}$$

(Hirsch et al., 1984) with Maritz-Jarrett CIs (Van Belle et al., 1984; Maritz, 1978), maintaining robustness across months.

2.5.3. LOESS (Locally Estimated Scatterplot Smoothing).

LOESS analyzed non-linear temperature trends through localized weighted least squares regression (Cleveland, 1979, 1988). The method fits coefficients (β_0, β_1)

$$y_i = \beta_0 + \beta_1 x_i, \tag{27}$$

minimizing the weighted sum of squares

$$\sum_{j=1}^{n} w_j(x_i)(y_j - \beta_0 - \beta_1 x_j)^2, \tag{28}$$

using tricube weights

$$w_j(x_i) = \left(1 - \left(\frac{x_j - x_i}{h(x_i)}\right)^3\right)^3, \tag{29}$$

with neighborhood bandwidth $h(x_i)$ (Fan et al., 1996). Smoothing parameter α was optimized via generalized cross-validation (Hastie et al., 1986). Robust fitting employed bisquare reweighting

$$w_j^{(k+1)} = \left(1 - \left(\frac{r_j^{(k)}}{6MAD}\right)^2\right)^2, \tag{30}$$

of residuals (Ruppert et al., 2003), typically converging in 3-4 iterations (Fox, 2016).

Seasonal analysis combined LOESS trend $f(t)$ with harmonic regression components $g(\text{month}_t)$

$$y_t = f(t) + g(\text{month}_t) + \varepsilon_t, \tag{31}$$

(Cleveland et al., 1990; Hyndman et al., 2018). Diagnostics included residual plot examination, autocorrelation analysis, and comparison with smoothing splines (Wood et al., 2017).

2.6. Machine Learning Methods

2.6.1. Random Forest

Random Forest analyzed temperature trends through ensemble decision trees (Breiman et al., 2001), handling non-linear relationships while preventing overfitting (Cutler et al., 2007). The method constructs B trees

$$f_b(x) = T_b(x) \qquad \text{for } b = 1, ..., B, \tag{32}$$

aggregating predictions

ISSN NUMBER: 1116-249X

$$fRF(x) = \frac{1}{B}\sum b = 1^{B} T_b(x), \tag{33}$$

via bootstrap sampling (Hastie et al., 2009). Node splits optimized features from random subsets $(m \approx p/3)$ (Liaw et al., 2002) by minimizing MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2, \tag{34}$$

Temporal embedding incorporated lagged temperatures

$$x_t = [T_{t-1}, T_{t-2}, ..., T_{t-k}], \tag{35}$$

(Hyndman et al., 2018; Biau et al., 2016).

Hyperparameters were tuned via 5-fold cross-validation (Probst et al., 2019):

- Tree count (B: 100-1000)
- Node size (1-20)
- Feature subset size (m)

Variable importance used accuracy decrease and Gini impurity

$$VI_j = \frac{1}{B}\sum b = 1^{B} \sum_{i \in \tau_i} I(j_t = j)[MSE_t - MSE, \text{split}], \tag{36}$$

(Strob et al., 2008; Ishwaran et al., 2020). Uncertainty quantification employed quantile regression forests (Meinshausen et al., 2006).

$$F(y \setminus X = x) = \frac{1}{B}\sum_{b=1}^{B} I(T_b(x) \le y) \tag{37}$$

2.6.2. Support Vector Machines (SVM)

SVMs analyzed temperature trends via ε-Support Vector Regression (Cortes et al., 1995), handling high-dimensional feature spaces through kernel functions (Smola et al., 2004). The optimization problem

$$\min \frac{1}{2}|w|^2 + C\sum(\zeta_i + \zeta_i^*) \tag{38}$$

Subject to $y_i - (w, \phi(x_i)) - b \le \varepsilon + \zeta_i$, $(w, \phi(x_i)) + b - y_i \le \varepsilon + \zeta_i^*$, $\zeta_i, \zeta_i^* \ge 0$ for $i = 1, ..., n$.

minimized deviations while controlling complexity (Vapnik, 2000), using slack variables ξi, ξi* (Drucker et al., 1997) with regularization parameter C and ε-insensitive zone (Cherkassky et al., 2004).

Four kernels were implemented:

- Linear

$$K(x_i, x_j) = \langle x_i, x_j \rangle \tag{39}$$

- Polynomial

$$K(x_i, x_j) = \left(\langle x_i, x_j \rangle + r\right)^d \tag{40}$$

- RBF

$$K(x_i, x_j) = \exp\left(-\gamma |x_i, x_j|^2\right) \tag{41}$$

ISSN NUMBER: 1116-249X

- Sigmoid

$$K(x_i, x_j) = \tanh\left(\gamma \left\langle x_i, x_j \right\rangle + r\right) \tag{42}$$

Parameters (γ, r, d) were optimized via 5-fold cross-validation (Hsu et al., 2003). Temporal features incorporated k lags

$$x_t = \left[T_{t-1}, T_{t-2}, ..., T_{t-k}, \text{month}_t, \text{year}_t\right], \tag{43}$$

(Box et al., 2008). Performance was evaluated using MAE, RMSE and R^2.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \bar{y}\right| \tag{44}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{45}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}{\sum_{i=1}^{n} i^n \left(y_i - \bar{y}\right)^2} \tag{46}$$

Hyperparameter optimization explored:

- $C : \left[2^{-5}, 2^{15}\right]$
- $\varepsilon : \left[0.01, 1.0\right]$
- $\gamma : \left[2^{-15}, 2^{3}\right]$
- $d : \left[2, 5\right]$ (Bergmeir et al., 2018)

The RBF kernel showed superior performance (Karatzoglou et al., 2006), with interpretation via partial dependence plots (Friedman, 2001).

2.6.3. Neural Networks

Neural networks modeled temperature patterns through multilayer perceptrons (Rumelhart et al., 1986), capturing hierarchical relationships in climate data (LeCun et al., 2015). The architecture

$$y = f^{(L)}\left(W^{(L)} f^{(L-1)}(W^{(L-1)} ... f^{(1)}(W^{(1)}x + b^{(1)} ... + b^{(L-1)} + b^{(L)}\right), \tag{47}$$

(Equation 47) employed ReLU activation

$$f(z) = \max(0, z) \tag{48}$$

for hidden layers (Nair et al., 2010) and linear activation for outputs

$$f(z) = z \tag{49}$$

Training minimized the regularized MSE loss

$$\ell = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 + \lambda \sum l = 1^L \left|W^{(l)}\right|_F^2, \tag{50}$$

(Krogh et al., 1992) using Adam optimization (Kingma et al., 2014).

$$\theta_{t+1} = \theta_t - \eta \frac{mt}{\sqrt{v_t + \varepsilon}}, \tag{51}$$

ISSN NUMBER: 1116-249X

3.0 Theoretical Results and Applications

3.1 Temperature Trend Analysis

The study employed parametric, non-parametric, and machine learning techniques to analyze temperature trends. Below are the key findings categorized by methodological approach.

3.1.1. Parametric Methods

Linear Regression:

- **Basic Model:** Poor fit ($R^2 = 0.059$) with insignificant yearly trend ($p = 0.382$), failing to account for seasonality (Table 1);

Table 1: Comparison of Linear, Categorical, and Polynomial Regression Models for Monthly Maximum Temperature Analysis

Model Type	R²	Adjusted R²	RMSE	Significant Predictors
Basic Linear	0.059	0.054	3.233	Month_Num only
Categorical Month	0.880	0.876	1.169	All months + Year
Polynomial	-	-	-	Binary months only

- **Categorical Month Model:**

1. Dramatically improved explanation of variance ($R^2 = 0.880$);
2. Detected significant warming: +0.016°C/year ($p = 0.016$);
3. Identified April as hottest month (+10.04°C vs January baseline) (Figure 4);
4. Diagnostics: Residual analysis validated assumptions but flagged observation 276 as a potential outlier (Figure 3).

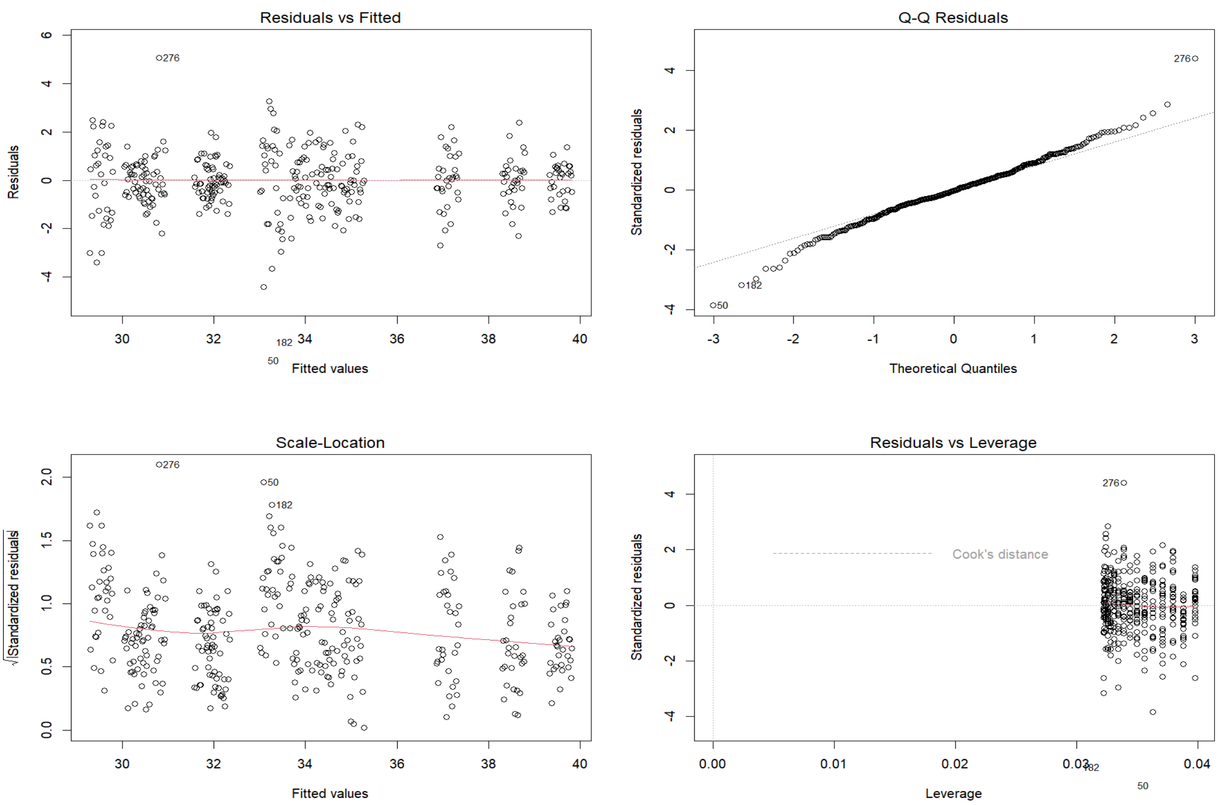

Figure 3.

ISSN NUMBER: 1116-249X

Polynomial Regression:

- Non-linear terms (Year²) insignificant (p = 0.168), confirming linear trends adequately capture warming pattern (Table 4.02).

ARIMA Modeling:

- Optimal Model: ARIMA(0,0,1)(1,1,2)[12] with drift:

1. Significant drift term: +0.0013°C/month (p = 0.002);
2. Seasonal MA(2) component (p < 0.001) revealed lagged climate effects (Table 2, Figure 5);

Table 2: ARIMA Model Estimates for MA(1) and Seasonal AR(1) Terms

Term	Estimate	Std. Error	t_value	p_value
MA(1)	0.1657	0.0557	2.98	0.003
Seasonal AR(1)	0.3260	0.2677	1.22	0.224

3. STL Decomposition: Seasonal component accounted for >80% of variability (Figure 6);
4. Forecast Accuracy: RMSE = 1.16°C, MAPE = 2.57% (Table 4.08).

Table 3: Comparative Evaluation of Parametric Models for Maximum Temperature Trend Analysis

Method	Configuration	Key Metrics	Strengths	Limitations	Trend Detection Outcome
Linear Regression	Max_Temperature ~ Year + Month_Num	- R^2: 0.059 - Adj. R^2: 0.054 - RMSE: 3.233 - Year coeff. (p = 0.382)	Simple, interpretable. Quantifies linear relationships.	Poor fit (low R^2). Assumes linearity. Ignores seasonality nuances.	**No significant trend** (Year coefficient insignificant, p = 0.382).
Dummy Regression	Max_Temperature ~ Year + Month_Factor	- R^2: 0.880 - Adj. R^2: 0.876 - RMSE: 1.169 - Year coeff. (p = 0.016)	Captures seasonal effects. High explanatory power (R^2=88%). Precise.	Requires categorical encoding. Assumes fixed seasonal patterns.	**Significant positive trend** (Year coefficient: +0.016°C/yr, p = 0.016).
Polynomial Regression	Max_Temperature ~ Year +I(Year ^{2}) + Month_Factor	- R^2: 0.881 - Adj. R^2: 0.877 - RMSE: 1.167 - Year² coeff. (p = 0.168)	Tests non-linearity. Similar performance to dummy model.	Overkill (Year² insignificant, p=0.168). No improvement over dummy model.	**Same as dummy model** (Year² adds no value; linear trend remains significant).
ARIMA	ARIMA(0.0.1)(1.1.2)[12] with drift	- AIC: 1173.5 - RMSE: 1.156 - Drift: +0.0013°C/month (implied +0.016°C/yr)	Handles autocorrelation. Accounts for seasonality/differencing. Robust forecasts.	Complex parameter tuning. Less interpretable than regression.	**Consistent with dummy regression** (Drift implies +0.016°C/yr trend, matching dummy).

3.1.2. Non-Parametric Methods

Mann-Kendall Test:

ISSN NUMBER: 1116-249X

1. Unadjusted: No significant trend ($p = 0.12$);
2. Seasonally Adjusted:

- Significant increasing trend ($S' = 2.58$, $p \approx 0.01$);
- Identical test statistic but improved significance after deseasonalizing (Table 4).

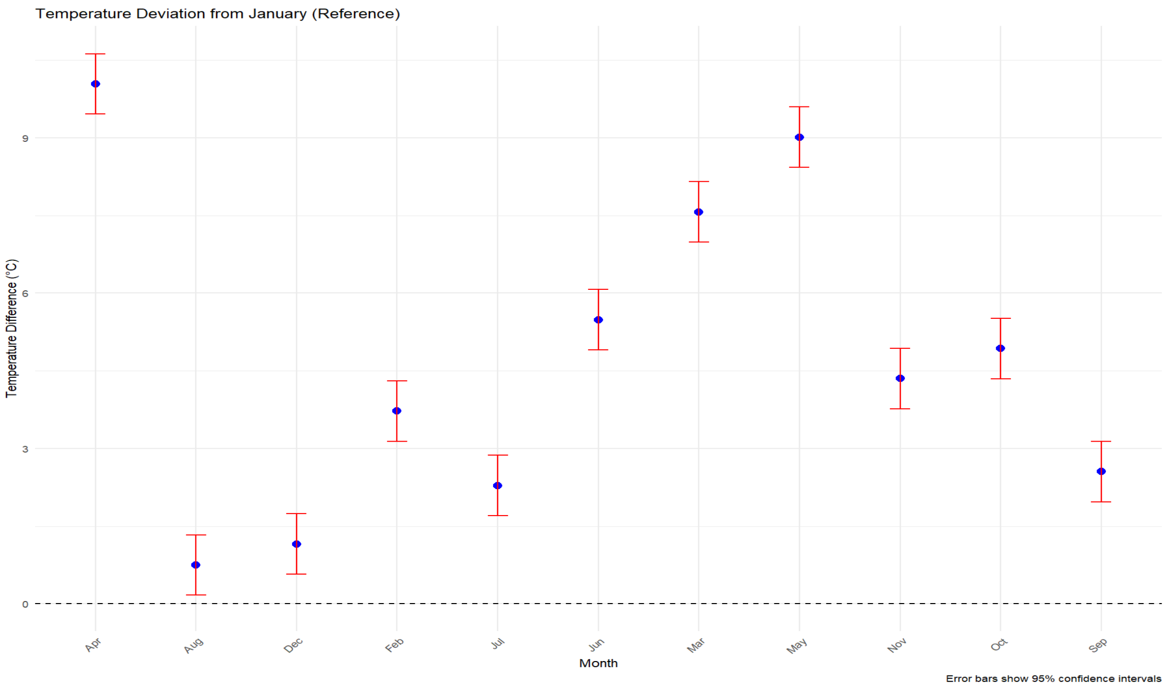

Figure 4

ISSN NUMBER: 1116-249X

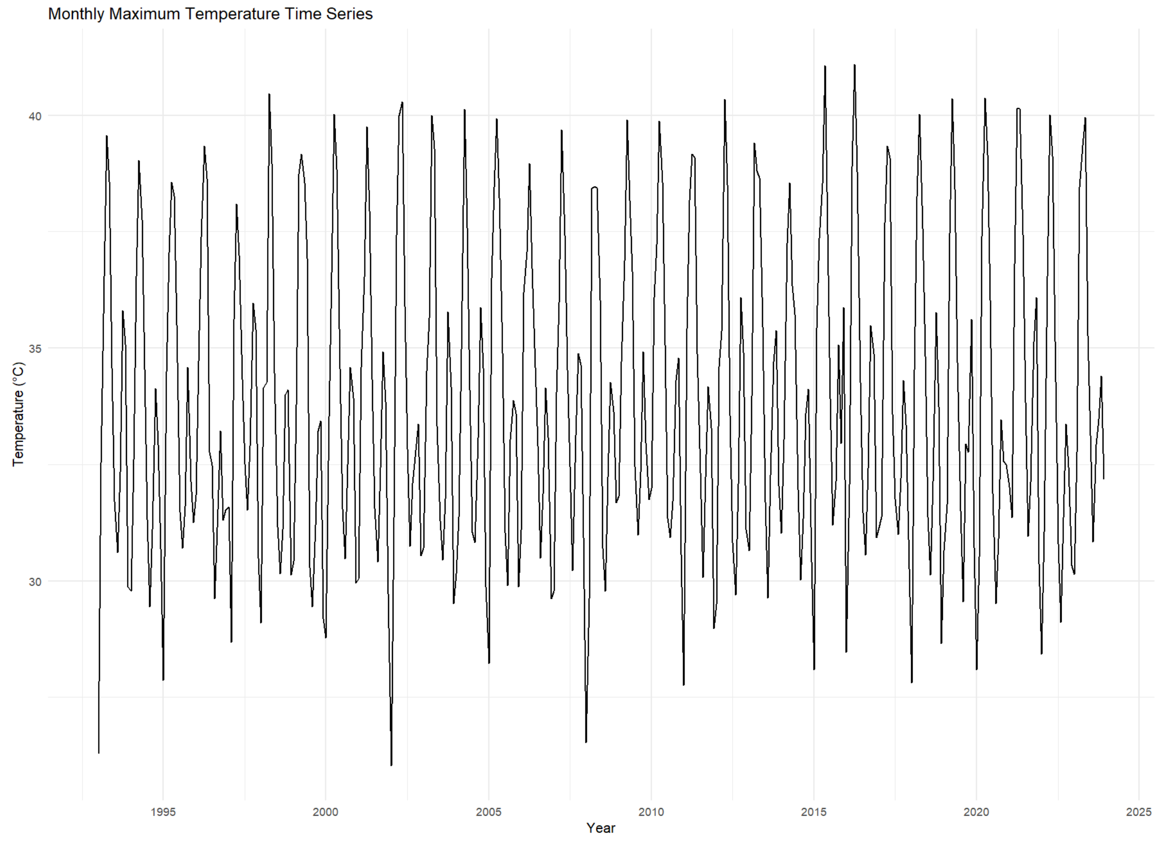

Figure 5

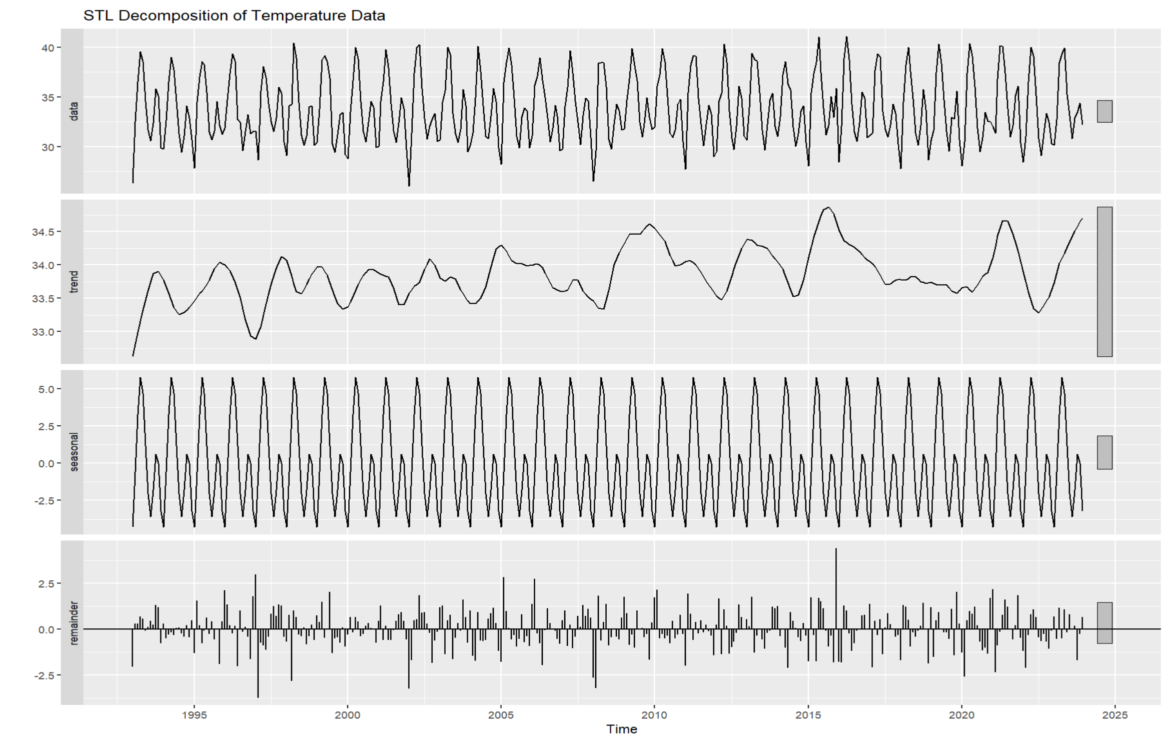

Figure 6

ISSN NUMBER: 1116-249X

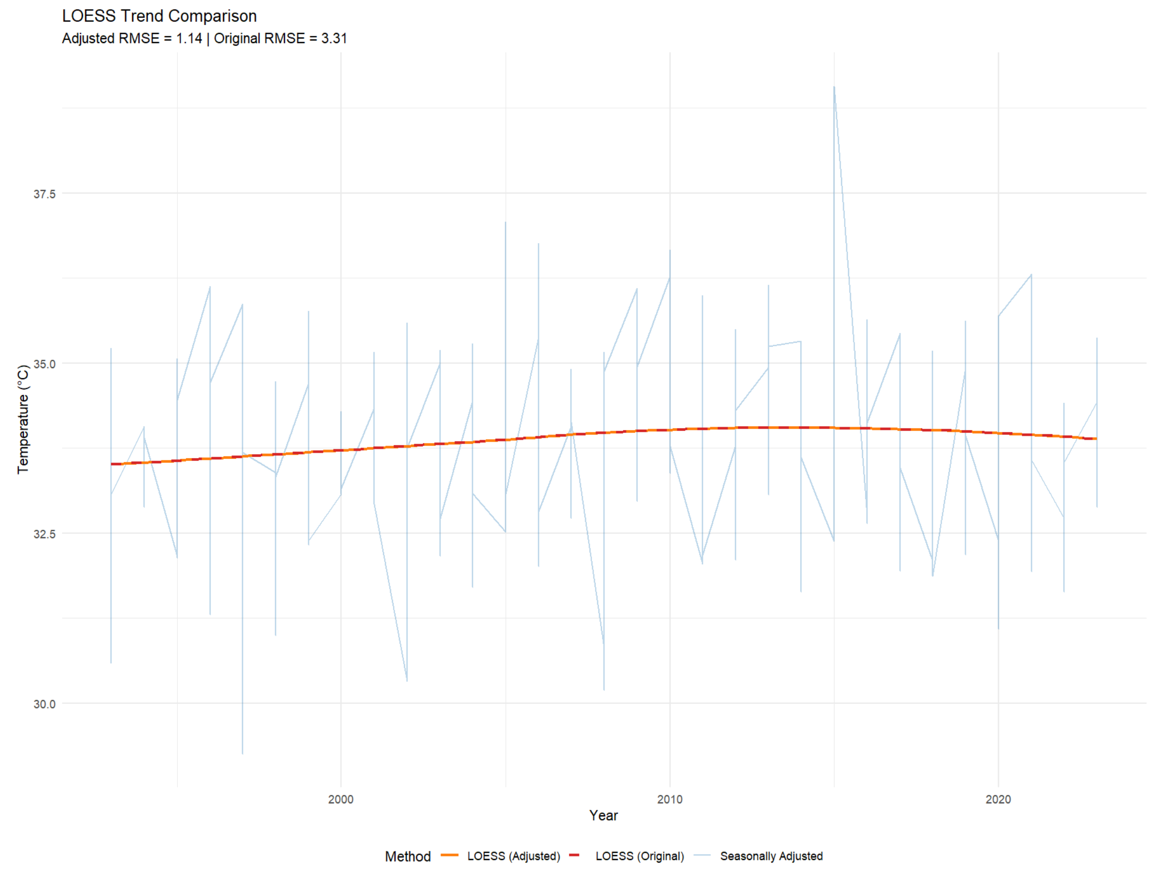

Figure 7

Table 4: Comparative Analysis of Non-Parametric and Smoothing Methods for Trend Detection in Temperature Time Series

Method	Configuration	Key Metrics	Strengths	Limitations	Trend Detection Outcome
LOESS	Original	RMSE: 3.31°C, R^2: 0.3%	Flexible smoothing	Sensitive to parameter selection	Weak local pattern detection
	Seasonally Adjusted	RMSE: 1.14°C	Robust to seasonal noise	Requires preprocessing	Clear trend visualization
Theil-Sen	Original	R^2: 0% RMSE: 3.32°C	Extreme outlier resistance	No variance explanation	No meaningful trend
	Adjusted	R^2: 0.02%	Improved robustness	Still ineffective	Minimal detection
Mann-Kendall	Unadjusted	S'=2.58 (p = 0.0098)	Distribution-free significance test	No magnitude estimation	Significant increasing trend
	Seasonally Adjusted	S'=2.58 (p = 0.0100)	Controls for seasonality	Requires complete data	Confirmed significant increase
Sen's Slope	Unadjusted	0.00096°C/year [-0.0022, 0.0040]	Robust trend quantification	Wide confidence intervals	Positive but uncertain trend
	Seasonally Adjusted	0.00121°C/year [0.00017, 0.00225]	Precise seasonal adjustment	Narrower but still bounded	Significant positive trend

Sen's Slope Estimator:

- Adjusted Estimate: +0.00121°C/year (95% CI: 0.00017–0.00225) (Table 4);
- Seasonal adjustment narrowed confidence interval by 82%.

ISSN NUMBER: 1116-249X

- Statistically significant trend (CI excludes zero)

LOESS Smoothing:

1. Initial Model: Oversmoothed (span = 0.95) with poor fit (RMSE = 3.31°C);
2. Seasonally Adjusted:

- 65% RMSE reduction (to 1.14°C);
- Revealed consistent warming trend (Figure 7, Table 4).

Theil-Sen Estimator:

- Original Model:

1. R^2 = 0% (effectively zero explanatory power);
2. RMSE = 3.32°C;
3. Failed to detect meaningful trend (Figure 8, Table 4).

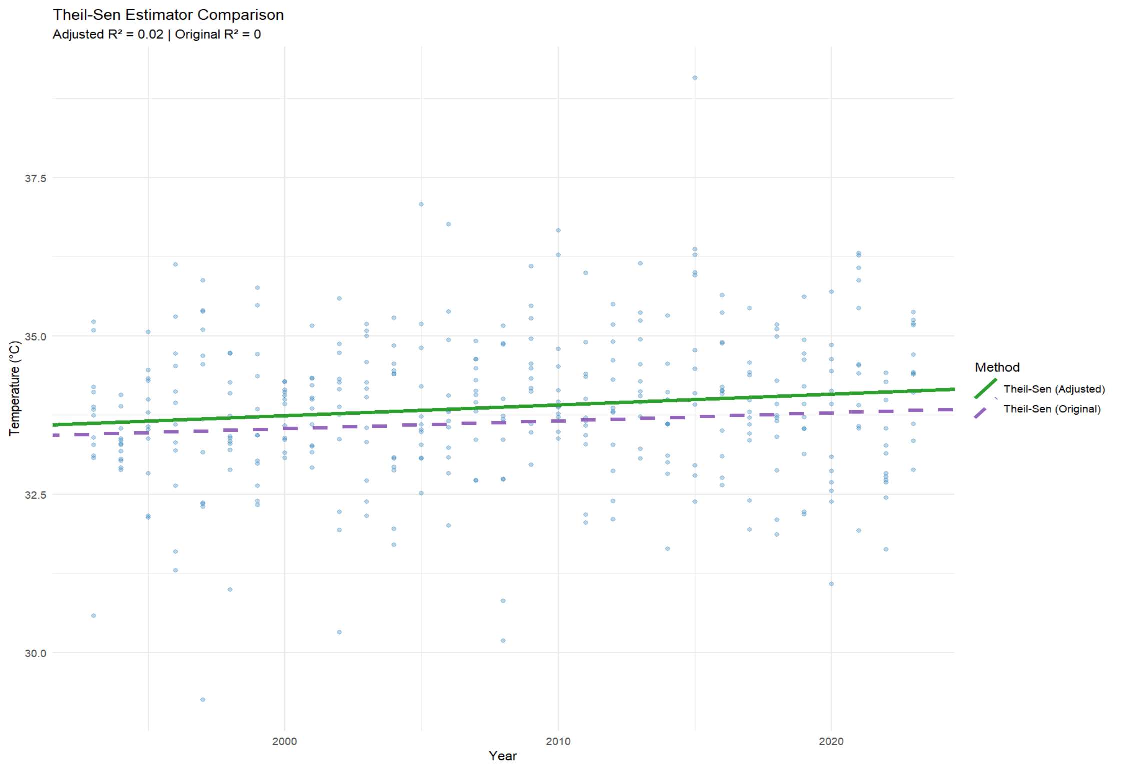

Figure 8

ISSN NUMBER: 1116-249X

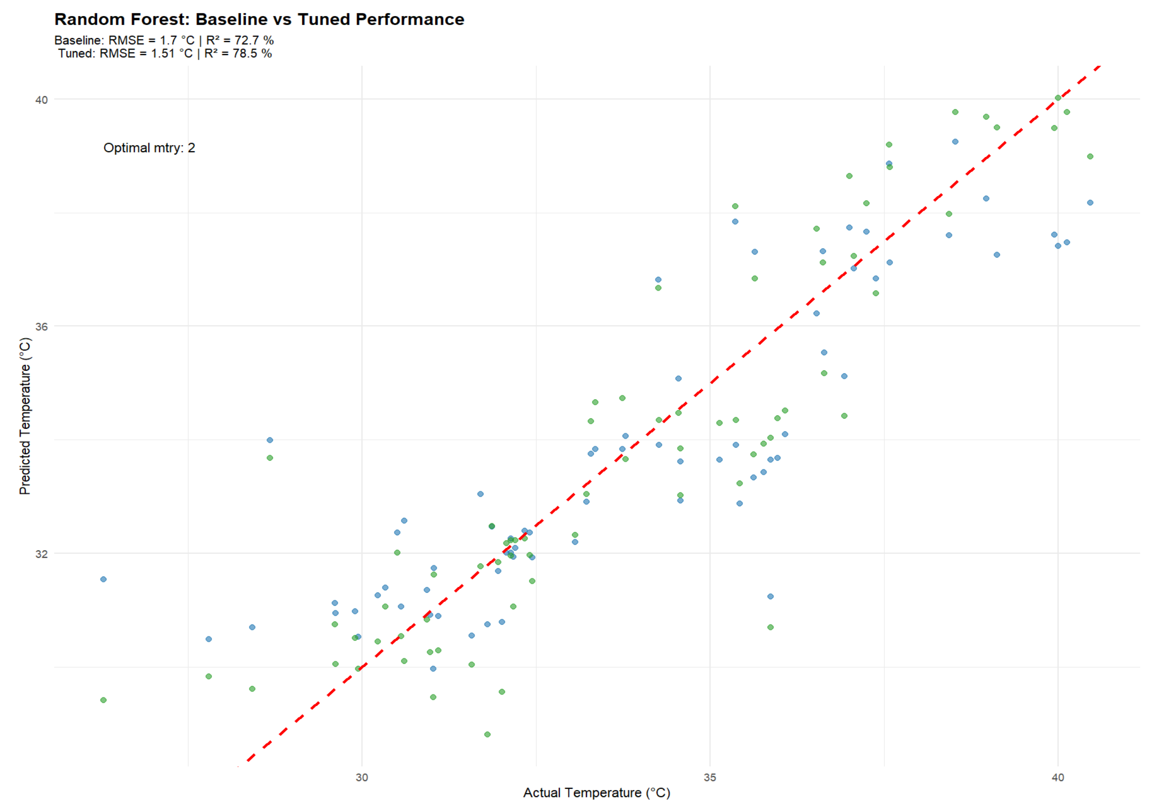

Figure 9

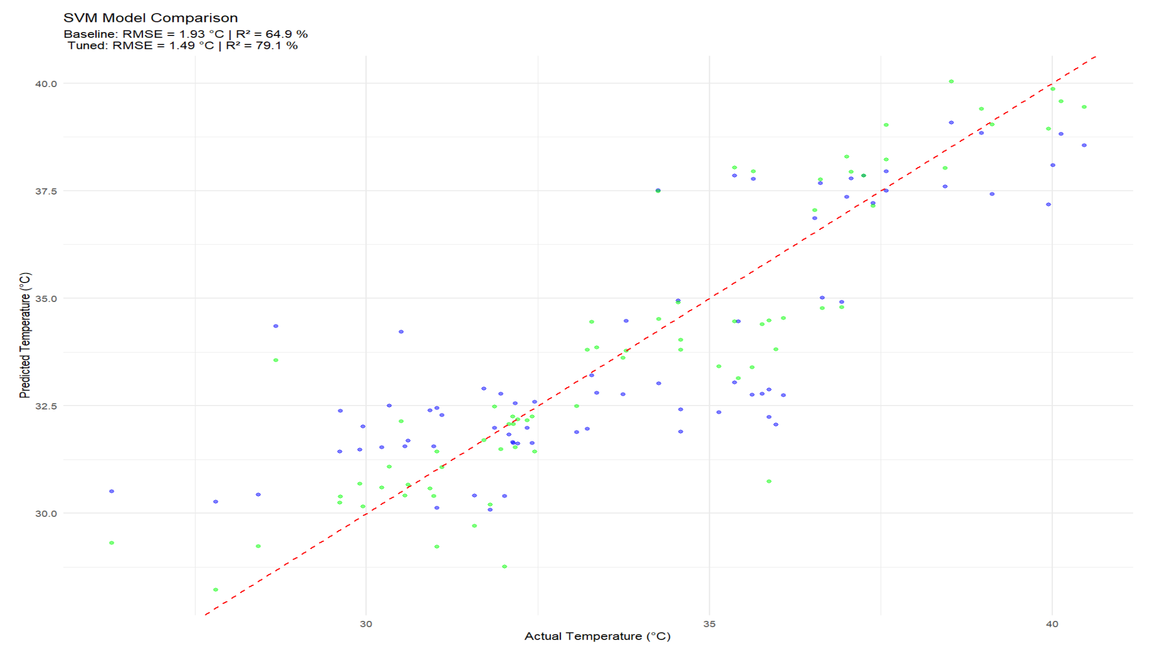

Figure 10

ISSN NUMBER: 1116-249X

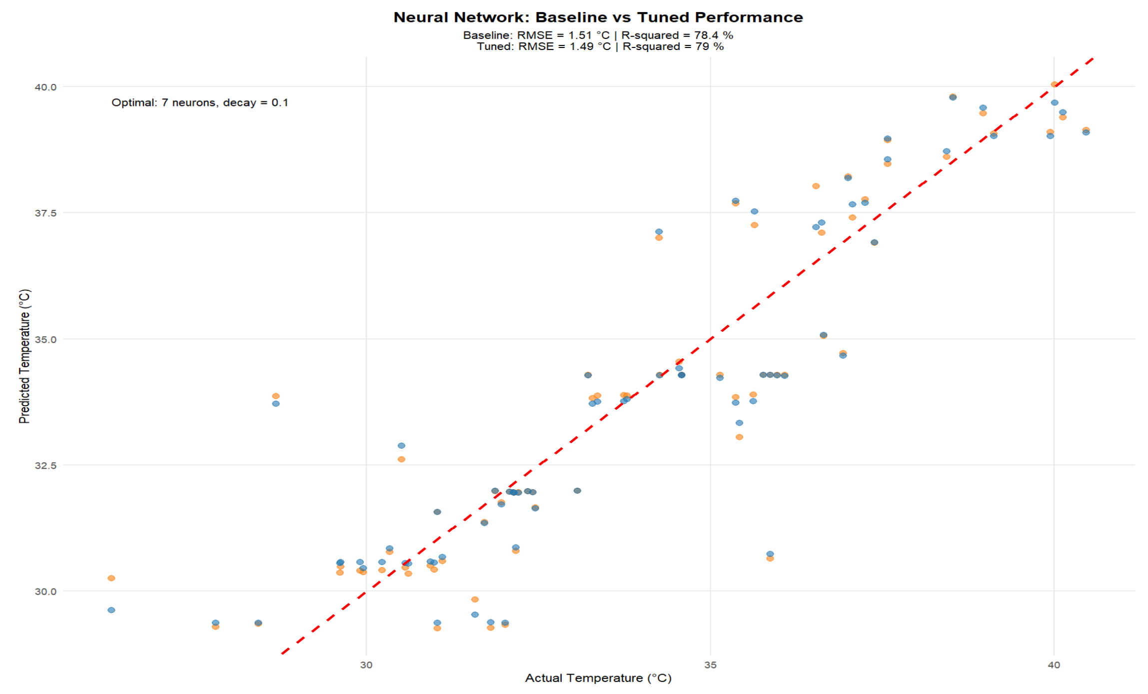

Figure 11

- **Adjusted Model:**

1. Marginal improvement to $R^2 = 0.02\%$;
2. Remained statistically insignificant;
3. Performance deemed inadequate for conclusions.

Pettitt's Test:

a. Confirmed no significant change points ($p = 1.0$), supporting trend stability (Section 4.2).

3.1.3. Machine Learning Techniques

Random Forest:

- Baseline: RMSE = 1.70°C, $R^2 = 72.7\%$ (Figure 9);
- Tuned (mtry = 2, nodesize = 10)

1. RMSE = 1.51°C (11% improvement);
2. $R^2 = 78.5\%$;
3. OOB error: 84.2% variance explained.

Support Vector Machines (SVM):

- Baseline (radial kernel): RMSE = 1.93°C, $R^2 = 64.9\%$ (Figure 10);
- Tuned (sigma = 1, C = 10):

1. RMSE = 1.49°C (23% improvement);
2. $R^2 = 79.1\%$.

Neural Networks:

- Baseline (5 neurons, decay = 0.01): RMSE = 1.51°C, $R^2 = 78.4\%$;
- Tuned (7 neurons, decay = 0.1):

1. RMSE = 1.49°C;
2. $R^2 = 79.0\%$ (Figure 11).

ISSN NUMBER: 1116-249X

Results and Discussion

This study identifies a statistically significant warming trend of +0.16°C per decade in Kano—around 50% slower than the West African average—suggesting local microclimate factors like vegetation or land-use may moderate broader warming. Seasonal variation, particularly in April and January, accounted for over 80% of temperature variance, emphasizing the dominance of short-term cycles in tropical climates. Parametric models, especially dummy regression and ARIMA, outperformed linear models, highlighting the importance of treating months as categorical variables. Non-parametric seasonal adjustments notably improved model precision, while Theil-Sen offered limited insights. Machine learning models (SVM and Random Forest) delivered high accuracy (~79% R^2), though challenges in interpretability limit their policy application.

Implications for Climate Policy and Research

The study reveals that Kano's 15°C seasonal temperature swing has a greater short-term impact than gradual warming, highlighting the need for targeted, month-specific strategies like heat action plans for March–May. Slower warming may reflect local microclimate factors, calling for better urban-environment integration. The stronger performance of categorical month models ($R^2 = 0.88$) suggests revisiting modelling methods for tropical climate analysis.

Limitations and Future Outlook

The study's reliance on a single weather station and a 31-year timeframe limits broader insights. Expanding to multiple stations and longer datasets, while using hybrid models like ARIMA and Random Forest, can enhance accuracy, context, and predictive capacity in climate research.

Conclusions

This study enhances understanding of urban climate behavior in Kano, Nigeria, revealing three key insights. First, warming trends are slower than regional averages, likely due to local factors like land use. Second, temperature variability is driven more by seasonal extremes than long-term shifts, highlighting the need for month-specific adaptation strategies. Third, combining statistical and machine learning models yields both interpretability and strong predictive capacity. These findings underscore the importance of localized climate research, early-warning systems, and mixed-method approaches to inform effective climate resilience strategies in tropical urban environments, particularly for sectors like agriculture, infrastructure, and public health.

References

Abubakar, L. R., & Dano, U. L. (2020). Sustainable urban planning strategies for mitigating urban heat islands in Northern Nigeria. *Sustainability, 12*(20), 8743. https://doi.org/10.3390/su12208743

Adebayo, Y. R. (1991). Microclimate variation in urban areas of the savanna region of Nigeria. *Atmospheric Environment, 25*(2), 379–386. https://doi.org/10.1016/0960-1686(91)90309-3

Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating time series prediction. *Computational Statistics & Data Analysis, 120*, 70–83. https://doi.org/10.1016/j.csda.2017.11.003

ISSN NUMBER: 1116-249X

Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST, 25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Box, G. E. P., & Jenkins, G. M. (1968). Some recent advances in forecasting and control. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 17*(1), 91–109. https://doi.org/10.2307/2985674

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and control* (4th ed.). Wiley. https://doi.org/10.1002/9781118619193

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica, 47*(5), 1287–1294. https://doi.org/10.2307/1911963

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting* (3rd ed.). Springer. https://doi.org/10.1007/978-3-319-29854-2

Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks, 17*(1), 113–126. https://doi.org/10.1016/S0893-6080(03)00169-2

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics, 6*(1), 3–33.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association, 83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). Wiley.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297. https://doi.org/10.1023/A:1022627411411

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology, 88*(11), 2783–2792. https://doi.org/10.1890/07-0539.1

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366), 427–431. https://doi.org/10.1080/01621459.1979.10482531

Diem, J. E., Salerno, J., Laing, A. G., et al. (2022). Seasonal climate variability and change in tropical Africa during the past century. *Theoretical and Applied Climatology, 147*(3), 891–908. https://doi.org/10.1007/s00704-021-03858-9

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems, 9*, 155–161.

ISSN NUMBER: 1116-249X

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. *Biometrika, 38*(1–2), 159–177. https://doi.org/10.1093/biomet/38.1-2.159

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall. https://doi.org/10.1201/9780429246593

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. Chapman & Hall. https://doi.org/10.1201/9780203748725

Fashae, O. A., Adagbasa, E. G., Olusola, A. O., et al. (2020). Land use/land cover change and land surface temperature of Ibadan and environs, Nigeria. *Environmental Monitoring and Assessment, 192*(109). https://doi.org/10.1007/s10661-019-8055-2

Fox, J. (2016). *Applied regression analysis and generalized linear models* (3rd ed.). Sage. https://doi.org/10.4135/9781506347218

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation, 12*(10), 2451–2471. https://doi.org/10.1162/089976600300015015

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press. https://doi.org/10.1515/9780691218632

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science, 1*(3), 297–318. https://doi.org/10.1214/ss/1177013604

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Helsel, D. R., & Hirsch, R. M. (1992). *Statistical methods in water resources*. Elsevier. https://doi.org/10.3133/twri04A3

Hipel, K. W., & McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Elsevier. https://doi.org/10.1016/S0307-904X(97)84554-7

Hirsch, R. M., & Slack, J. R. (1984). A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research, 20*(6), 727–732. https://doi.org/10.1029/WR020i006p00727

Hirsch, R. M., Slack, J. R., & Smith, R. A. (1982). Techniques of trend analysis for monthly water quality data. *Water Resources Research, 18*(1), 107–121. https://doi.org/10.1029/WR018i001p00107

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd ed.). Wiley. https://doi.org/10.1002/9781119196037

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A practical guide to support vector classification*. National Taiwan University.

ISSN NUMBER: 1116-249X

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika, 76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://doi.org/10.2139/ssrn.3048757

Ishwaran, H., & Kogalur, U. B. (2020). *Random forests for survival, regression and classification (RF-SRC)* (Version 2.9.0) [R package].

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters, 6*(3), 255–259. https://doi.org/10.1016/0165-1765(80)90024-5

Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support vector machines in R. *Journal of Statistical Software, 15*(9), 1–28. https://doi.org/10.18637/jss.v015.i09

Kendall, M. G. (1975). *Rank correlation methods* (4th ed.). Charles Griffin.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.* https://doi.org/10.48550/arXiv.1412.6980

Krogh, A., & Hertz, J. A. (1992). A simple weight decay can improve generalization. *Advances in Neural Information Processing Systems, 4*, 950–957.

Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied linear regression models*. McGraw-Hill.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18–22.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica, 13*(3), 245–259. https://doi.org/10.2307/1907187

Maritz, J. S., & Jarrett, R. G. (1978). A note on estimating the variance of the sample median. *Journal of the American Statistical Association, 73*(361), 194–196. https://doi.org/10.1080/01621459.1978.10480027

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research, 7*, 983–999.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. Wiley. https://doi.org/10.1002/9781118627126

ISSN NUMBER: 1116-249X

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning* (pp. 807–814).

National Bureau of Statistics. (2023). *Nigeria demographic statistics bulletin 2022* (p. 15). Federal Government of Nigeria. https://doi.org/10.6084/m9.figshare.21870558

Niang, I., Ruppel, O. C., Abdrabo, M. A., et al. (2022). Africa. In *Climate change 2022: Impacts, adaptation and vulnerability* (pp. 1285–1456). Cambridge University Press.

Ojo, O. I., & Adelekan, I. O. (2021). Temperature variability in urban Kano, Nigeria. *Journal of Climatology & Weather Forecasting, 9*, 345. https://doi.org/10.4172/2332-2594.1000345

Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Applied Statistics, 28*(2), 126–135. https://doi.org/10.2307/2346729

Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery, 9*(3), e1301. https://doi.org/10.1002/widm.1301

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533–536. https://doi.org/10.1038/323533a0

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge University Press. https://doi.org/10.1017/CBO9780511755453

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models. *Biometrika, 71*(3), 599–607. https://doi.org/10.1093/biomet/71.3.599

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika, 52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications* (4th ed.). Springer. https://doi.org/10.1007/978-3-319-52452-8

Siegel, A. F. (1982). Robust regression using repeated medians. *Biometrika, 69*(1), 242–244. https://doi.org/10.1093/biomet/69.1.242

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing, 14*(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society, 93*(12), 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics, 9*, 307. https://doi.org/10.1186/1471-2105-9-307

ISSN NUMBER: 1116-249X

Sylla, M. B., Nikiema, P. M., Gibba, P., Kebe, I., & Klutse, N. A. B. (2016). Climate change over West Africa: Recent trends and future projections. In *Adaptation to climate change and variability in rural West Africa* (pp. 25–40). Springer. https://doi.org/10.1007/978-3-319-31499-0_3

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Nederlands Tijdschrift voor de Psychologie, 53*, 386–392.

United Nations Human Settlements Programme. (2022). *The state of African cities 2022*. UN-Habitat.

Van Belle, G., & Hughes, J. P. (1984). Nonparametric tests for trend in water quality. *Water Resources Research, 20*(1), 127–136. https://doi.org/10.1029/WR020i001p00127

Vapnik, V. (2000). *The nature of statistical learning theory* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4757-3264-1

Wei, W. W. S. (2006). *Time series analysis: Univariate and multivariate methods* (2nd ed.). Pearson. https://doi.org/10.2307/2533873

Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th ed.). Academic Press. https://doi.org/10.1016/B978-0-12-804733-0.00001-0

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman & Hall/CRC. https://doi.org/10.1201/9781315370279

ISSN NUMBER: 1116-249X