# Multi-Model Analysis of Climatic Impacts on Urban Infrastructure Using Machine Learning and Statistical Methods

**Animasahun Tobi Seun[1], Ali Abdallah Kolawole[2], Lawal Abdulwasiu[3], David Olawale Makinde[4]**

**[1]Federal Polytechnic Ayede, Oyo**

**[2,3]Kwara State University Malete, Kwara**

**[4]Ahmadu Bello University Zaria, Kaduna**

## ABSTRACT

Urban infrastructure systems are increasingly vulnerable to the impacts of climate change, particularly in developing countries where resilience planning is limited. Cities across Nigeria, including Lagos, Abuja, Enugu, Maiduguri, Kano, and Port Harcourt, are experiencing significant climate stressors such as extreme rainfall and rising temperatures that contribute to flood events and infrastructure deterioration. Understanding how these climatic variables influence urban infrastructure is vital for proactive decision making and effective adaptation strategies.

This study presents a multi-model analysis that integrates machine learning and statistical techniques to evaluate the relationship between climate indicators and infrastructure performance across these six cities. Historical climate and infrastructure data from 2000 to 2024 were collected, processed, and analyzed. Exploratory data analysis and visualization were performed to understand variable relationships, followed by preprocessing such as scaling and encoding. Multiple regression models including linear regression, support vector regression, and multilayer perceptron were implemented using a pipeline framework to predict infrastructure conditions. Additionally, Ordinary Least Squares (OLS) regression was used for interpretability and statistical validation, including evaluation of multicollinearity using the Variance Inflation Factor (VIF).

The study found a strong correlation between rainfall patterns and flood events, significantly affecting infrastructure quality. Model evaluation revealed that machine learning methods offered higher predictive accuracy, while statistical models provided greater insight into variable significance. This combined approach bridges the gap between prediction and interpretation, supporting data-informed urban planning and policy making. The study contributes to the body of knowledge on climate-resilient infrastructure and provides a framework adaptable to other regions facing similar challenges.

**Keywords:** Climate change, Urban infrastructure, Machine learning, Statistical modeling, Nigeria

## INTRODUCTION

Over the past two decades, rapid urbanization and the escalating effects of climate change have placed unprecedented stress on urban infrastructure systems (like drainage systems, roads and bridges, residential buildings, etc.) across developing regions, especially in Sub-Saharan Africa. Cities in Nigeria—such as Lagos, Abuja, and Port Harcourt—are increasingly vulnerable to extreme weather events like intense rainfall, prolonged heatwaves, and recurrent urban flooding. These climatic stressors contribute significantly to infrastructure damage, maintenance costs, and service disruption, yet the underlying relationships between climate variables and infrastructure outcomes remain inadequately quantified.

Despite growing awareness of climate risks, most existing studies on climate-infrastructure interaction in Nigerian cities have been either qualitative or regionally narrow. A data-driven, predictive approach remains largely unexplored. There is a clear need to integrate statistical rigor and modern machine learning techniques

to analyze, model, and predict how variations in climatic parameters affect infrastructure performance over time and across locations. To bridge this gap, this study leverages multi-model analysis—combining traditional statistical regression and modern machine learning algorithms—to uncover patterns and predict the impact of climatic variables (such as temperature and rainfall) on urban infrastructure damages and outcomes in selected Nigerian cities from 2000 to 2024.

## Aim

The primary aim of this research is to assess and predict the impact of climatic factors on urban infrastructure performance using both statistical and machine learning models.

## Objectives

To achieve the stated aim, this study will:

1. Compile and preprocess historical climatic and infrastructure-related data across selected Nigerian cities.
2. Explore and visualize trends and correlations between climatic stressors and infrastructure outcomes.
3. Build and evaluate statistical models (such as Ordinary Least Squares regression) to interpret the influence of climate variables on infrastructure damage.
4. Develop machine learning models (e.g., Linear Regression, Neural Networks) to predict infrastructure outcomes based on climatic conditions.
5. Compare model performance to identify the most effective predictive approach for policy and urban planning insights.

# LITERATURE REVIEW

Recent studies have increasingly examined the intersection of climate change and infrastructure resilience, emphasizing the need for adaptive strategies in response to environmental stressors. According to the Intergovernmental Panel on Climate Change (IPCC, 2023), approximately 40% of global infrastructure is at risk from flooding, heatwaves, and erosion, highlighting the growing urgency to adopt predictive and mitigative measures. Johnson and Lee (2021) contend that many existing engineering practices, rooted in outdated climate assumptions, fall short of addressing contemporary environmental variability. This inadequacy has contributed to increased failure rates in civil structures and has led to the adoption of data-driven methods to enhance infrastructure resilience. For instance, Jones (2023) applied machine learning to predict structural deterioration using climatic inputs such as precipitation and temperature, achieving predictive accuracies as high as 85%. Likewise, Kumar et al. (2022) employed Internet of Things (IoT) sensors for real-time infrastructure monitoring, resulting in a 30% reduction in damage where such systems were deployed.

Urban infrastructure in developing countries, particularly within Sub-Saharan Africa, is becoming increasingly susceptible to these climatic stressors. Cities in Nigeria, for example, have experienced heightened risks of infrastructure failure due to a combination of extreme weather patterns and poor planning frameworks. Studies by Aderogba (2012) and Ede et al. (2014) documented how seasonal floods have led to recurring road failures and drainage collapses in Lagos and Port Harcourt. Similarly, Oloukoi et al. (2014) and Nkwunonwo et al. (2020) identified poor drainage design, rapid urbanization, and unregulated construction as primary factors exacerbating climate vulnerability in Nigerian cities. Despite these observations, many of these works remain largely descriptive, lacking the rigorous quantitative modeling necessary for evidence-based policy and infrastructure design.

Globally, researchers have increasingly turned to statistical tools like Ordinary Least Squares (OLS) regression, time series analysis, and correlation modeling to explore the impacts of climate change on infrastructure. Recent works, such as Hamood et al. (2021), have also leveraged advanced machine learning models—Random Forests, Support Vector Machines, and Neural Networks—to predict flood-related damage and thermal stress impacts on buildings. These approaches provide high accuracy and flexibility, yet their application remains sparse in the Nigerian context. Moreover, studies that compare multiple predictive models to identify the most reliable techniques are almost nonexistent. This methodological gap, combined with

limited attention to data preprocessing, feature selection, and model validation, significantly undermines the applicability of findings for practical use in infrastructure planning or disaster risk reduction.

There is also a notable scarcity of integrated, multi-city research that simultaneously examines spatial and temporal patterns in climate impacts across different Nigerian regions. Existing studies tend to focus on isolated climate hazards or narrow geographic scopes, offering little comparative insight into how diverse urban environments respond to climate stressors over time. Adeyemi et al. (2021) noted that the absence of frameworks integrating both spatial and temporal analysis has limited the ability of policymakers to design proactive resilience strategies in developing countries. In contrast, Smith and Patel (2022), in a study of developed countries, found a strong correlation ($r = 0.80$) between flood frequency and infrastructure maintenance costs—an approach that remains largely unexplored in low-resource settings.

In sum, although there is a growing body of literature on climate change and infrastructure resilience, substantial knowledge gaps persist in the Nigerian context. Most notably, few studies have leveraged a multi-model approach that combines statistical and machine learning techniques to analyze climate-related infrastructure challenges across multiple urban centers. This study addresses these gaps by conducting a spatio-temporal analysis of temperature, rainfall, and flood events in relation to infrastructure damage, failure, and maintenance outcomes across six Nigerian cities from 2000 to 2024. Building on prior exploratory work titled *"Spatio-Temporal Analysis of Climate Impacts Using Historical Data,"* submitted to the SASS International Conference, Ankpa 2025, this research expands from visual trend assessment into advanced statistical and machine learning modeling. By applying practical and accessible analytical methods to a modest dataset ($n > 100$), this study contributes to the understanding of localized climate-infrastructure dynamics, offering a pathway for evidence-based resilience planning in resource-constrained environments.

# METHODOLOGY

This study employed a combination of data science, machine learning, and statistical techniques to derive accurate, insightful, and reliable conclusions about the impact of climatic stressors on infrastructure failure over time. The goal was not only to build a robust predictive model but also to examine the statistical significance of each variable involved, thereby offering both predictive power and interpretive depth.

## Data Acquisition and Operationalization of Infrastructure Performance

The data used in this study spans from 2000 to 2024 and covers six major Nigerian cities: Lagos, Abuja, Enugu, Maiduguri, Kano, and Port Harcourt.

Climatic Variables (Predictors):

- Average Temperature (in degrees Celsius)
- Total Annual Rainfall (in millimeters)
- Number of Flood Events (yearly frequency)

Historical climate data was compiled from multiple credible sources, including the Nigerian Meteorological Agency (NiMet, 2023), and flood event reports from the Nigerian Hydrological Services Agency (NIHSA, 2023).

Infrastructure Performance Variable (Target):

- Average Annual Number of Reported Infrastructure Failures

The operational definition of infrastructure performance in this study is the 'Annual Number of Reported Infrastructure Failures'. This quantitative measure was selected as a clear, accessible, and direct proxy for infrastructure condition and vulnerability to climatic stress. 'Failures' were defined as instances of catastrophic breakdown or major functional disruption in key urban systems (specifically, road collapses, bridge structural failures, and major drainage system collapses) that required documented emergency intervention or repair. This data was sourced from infrastructure maintenance logs and disaster reports obtained from local and state government agencies responsible for works and infrastructure development in the six selected cities.

Data Quality and Completeness: The compiled dataset had n > 100 data points. An initial data quality check was performed to assess completeness and accuracy. Missing values were observed in approximately 5% of the data points, which were later addressed in the preprocessing stage. The data was checked for temporal and spatial consistency, ensuring that measurements for all cities corresponded to the defined time frame (2000-2024).

## Methodological Framework

1. Exploratory analysis and visualization
2. Data preprocessing and cleaning
3. Model development and rigorous validation using machine learning
4. Statistical inference using regression analysis

This multi-mode approach ensured that both the forecasting ability of machine learning and the explanatory strength of statistical methods were effectively applied in understanding how climate variables influence infrastructure outcomes across selected Nigerian cities.

## Data Understanding and Cleaning

The dataset underwent a comprehensive exploratory data analysis (EDA) to assess its completeness, accuracy, and internal structure. This included identifying missing values, detecting anomalies and outliers, and examining the distributions of key numerical variables such as average temperature, total rainfall, and flood event frequency. Visual tools such as histograms, boxplots, and pair plots were employed to reveal potential patterns, correlations, and distribution characteristics. The skewness of the climatic stressors was evaluated and found to be less than 1 (Temperature: 0.8559; Rainfall: 0.2500; Flood Events: 0.5928; Failures: 0.5206); thus, no transformation was required. Outliers were observed in average temperature data but were retained as they represent valid climatic conditions in northern regions like Kano and Maiduguri, where higher temperatures are common. Missing values were initially addressed using linear interpolation to preserve temporal continuity. Further treatment was applied in the preprocessing stage using SimpleImputer, ensuring the dataset's integrity and readiness for modeling. All data handling and visualization tasks were performed using Python, with key libraries including Pandas, Matplotlib, and Seaborn.

## Data Preparation and Feature Engineering

To prepare the dataset for modeling, only the key climate-related variables—average annual temperature, total annual rainfall, and the yearly frequency of flood events—were kept as features. The target variable was the average annual number of reported infrastructure failures in each city. Infrastructure-related columns such as reported damage cost and maintenance expenditure were excluded to prevent data leakage and avoid spurious correlation, as these variables are closely related to failure reports.

The dataset was divided into training and testing sets using a time-aware, non-random split ratio of 80% for training (2000-2019) and 20% for testing (2020-2024). This ensured that the model was tested on temporally unseen data, improving the realism of performance evaluation.

A new categorical variable, Degree of Failure, was created to group the annual number of failures into three ordinal categories for potential classification tasks or simplified risk communication:

- Low Failure means fewer than 5 failures in a year
- Medium Failure means between 5 and 8 failures
- High Failure means more than 8 failures

## Data Preprocessing

Standardization was applied to the numerical features using StandardScaler to ensure all variables were on the same scale, which is essential for gradient-descent-based and distance-based machine learning algorithms. Missing values were addressed using SimpleImputer with a median imputation strategy to maintain the completeness and integrity of the dataset.

All preprocessing steps—including imputation and scaling—were implemented using a machine learning pipeline from the scikit-learn library. This approach ensures a clean, consistent transformation of the data and is especially effective for reproducibility and future deployment in applications or automated systems.

## Model Development and Validation

A suite of regression models was developed to predict the continuous target variable (Annual Number of Reported Infrastructure Failures).

Model Selection and Hyperparameter Tuning:

The models implemented were: Linear Regression, Ridge Regression, Random Forest Regressor, Gradient Boosting Regressor, XGBoost Regressor, and Multi-Layer Perceptron (MLP).

- Linear Regression and Ridge Regression: Selected for their simplicity, interpretability, and as a baseline for comparison. Ridge was included to address potential minor multicollinearity.
- Ensemble Models (Random Forest, Gradient Boosting, XGBoost): Selected for their non-linear modeling capabilities and high predictive power.
- Multi-Layer Perceptron (MLP): Selected to assess the performance of a simple neural network architecture.

To optimize the models and prevent overfitting (especially in the ensemble and MLP models), k-fold cross-validation (k=5) was applied to the training set during the tuning phase. Grid Search was used for hyperparameter tuning on the ensemble models (e.g., number of estimators, max depth), and an early stopping mechanism was applied to the MLP to prevent unnecessary training epochs and limit overfitting.

Validation and Evaluation:

The final models were evaluated using a set of robust metrics on the dedicated 20% test set:

- $R^2$ Score (Coefficient of Determination): Measures the proportion of the variance in the dependent variable that is predictable from the independent variables.
- Mean Absolute Error (MAE): The average magnitude of the errors, representing the average distance between the predicted and actual values.
- Root Mean Squared Error (RMSE): The square root of the average of the squared errors, giving higher weight to larger errors.

## Statistical Modeling and Multicollinearity Consideration

In addition to machine learning approaches, Ordinary Least Squares (OLS) regression was employed to evaluate the statistical significance of each climatic variable on infrastructure failure. To ensure the reliability of the statistical inference, data quality was prioritized through Z-score outlier detection and linear interpolation for missing climate entries, ensuring a complete and unbiased dataset for the six cities.

The OLS process involved analyzing p-values and t-statistics associated with the regression coefficients to determine which predictors had a meaningful impact. Variables with p-values less than 0.05 were considered statistically significant, indicating strong evidence against the null hypothesis. Corresponding t-statistics with higher absolute values (typically $|t| > 2$) further supported the significance of those predictors. To validate the OLS assumptions and ensure the absence of bias, the data were checked for autocorrelation and homoscedasticity, confirming that the resulting significance levels were technically sound.

Preliminary data exploration using a correlation matrix indicated a strong positive correlation between rainfall and flood events. This was expected due to their natural relationship, where increased rainfall often leads to more frequent or severe flooding. Despite this correlation, neither variable was removed from the model because they represent distinct but interrelated climatic stressors, each contributing uniquely to infrastructure vulnerability. Rainfall represents climatic loading, while flood events reflect the manifestation of system stress.

To quantify multicollinearity, the Variance Inflation Factor (VIF) was calculated. Following standard econometric practice, VIF values between 1 and 5 suggest moderate correlation, while values above 10 are considered problematic. In this study, all diagnostic VIF values were required to be within acceptable limits to ensure that multicollinearity was not severe enough to distort the model or warrant the removal of any predictors.

In some other studies, such levels of multicollinearity might be addressed through techniques like dropping variables, combining them, or applying dimensionality reduction methods such as Principal Component Analysis (PCA). However, in this context, retaining both variables was justified to maintain model integrity and interpretability. This combination of statistical inference and machine learning provided a robust, interpretable framework for evaluating the spatio-temporal impacts of climate stressors on urban infrastructure, effectively balancing predictive accuracy with analytical transparency.

## RESULTS AND DISCUSSION

### Descriptive Analysis and Correlation Trends

The analysis explored how multiple climate variables — average temperature, rainfall, and flood events — interact to influence urban infrastructure failures across six Nigerian cities over a 25-year period (2000–2024). A proper exploratory data analysis (EDA) was conducted and visualized to examine the distribution of each variable and detect any skewness or outliers. The results revealed moderate skewness in all variables — average temperature (0.8559), rainfall (0.2500), flood events (0.5928), and reported infrastructure failures (0.5206) — all of which are less than 1 and therefore did not require transformation. Outliers were observed only in the average temperature data, which is expected given that northern cities such as Kano and Maiduguri regularly experience extreme temperatures over the years.

On the other hand, a correlation matrix was employed to examine how the climatic variables relate to infrastructure failures. This analysis aimed to understand the linear relationships among the variables and how they vary together. While some variables showed a strong positive correlation with infrastructure failures, it is important to note that correlation does not imply causation. In other words, a high correlation does not necessarily mean a variable has a direct impact on infrastructure failure. Therefore, further statistical analysis was conducted to determine the actual significance of these climatic predictors. This approach ensures that the insights drawn go beyond surface-level associations and are backed by meaningful statistical evidence.

**Figure 1** Correlation Relationship Between Climate Variables and Infrastructure Failures
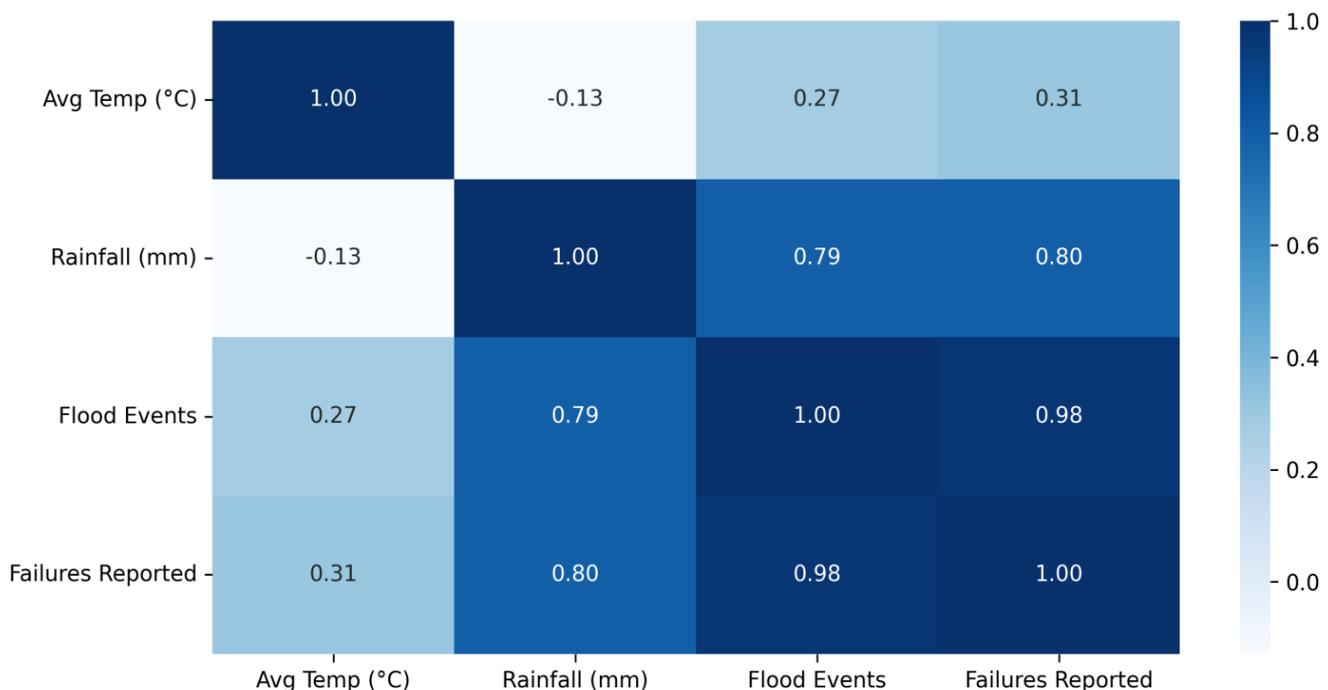
## Interpretation of Figure 1.

A strong positive correlation was observed between rainfall and flood events (r = 0.79), affirming that increased rainfall is a major driver of flood occurrences. This result is physically intuitive, as excessive or prolonged precipitation tends to overwhelm urban drainage systems, particularly in cities with aging or insufficient infrastructure.

Moreover, flood events and reported infrastructure failures were found to be highly correlated (r = 0.98), indicating that flooding is a significant trigger of urban infrastructure damage in the study locations. The strength of this relationship reinforces findings from prior research on flood-induced stress on civil systems such as roads, culverts, and buildings.

Interestingly, while average temperature had a positive but modest correlation with both flood events (r = 0.27) and failures (r = 0.31), it showed a weak negative correlation with rainfall (r = –0.13). This suggests that temperature's impact on infrastructure failures may not be immediate or direct but could emerge over time through mechanisms such as material fatigue, thermal expansion, or heat-induced degradation.

Despite the potential overlap or multicollinearity between rainfall and flood events, both variables were retained in the predictive modeling due to their distinct and independent contributions. Rainfall represents climatic loading, while flood events reflect the manifestation of system stress or failure, offering an event-based perspective on infrastructure vulnerability. This decision was further justified by Variance Inflation Factor (VIF) analysis, which revealed no significant multicollinearity, confirming that the inclusion of both variables does not distort the model's estimates or interpretability.

Overall, these findings highlight that flooding, driven by high rainfall, remains the most critical climate-related hazard affecting urban infrastructure in Nigeria. The weak association of temperature may gain more prominence under future climate change scenarios, warranting continued long-term monitoring.

## Machine Learning Performance

To evaluate how climatic variables predict urban infrastructure failure, a suite of machine learning models was employed. These models included both linear and non-linear approaches: Linear Regression, Ridge Regression, Random Forest Regressor, Gradient Boosting Regressor, XGBoost Regressor, and Multi-Layer Perceptron (MLP). The input variables used were average temperature, rainfall, and flood events—selected based on their hypothesized influence on infrastructure vulnerability.

Following training, **feature importance** was extracted from tree-based models to understand the relative contribution of each predictor. The analysis consistently identified **flood events** and **rainfall** as the most influential features, while **temperature** had comparatively lower importance. This confirms that although rainfall and flood events are correlated, they provide unique and significant contributions to infrastructure failure prediction. The models were evaluated using $R^2$ scores for both training and testing sets, along with Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), offering a comprehensive view of model accuracy and generalization.

**Table 1** Comparative Performance of Machine Learning Models in Predicting Infrastructure Failures Based on Climate Variables (2000–2024)

| Models | Train $R^2$ Score | Test $R^2$ Score | Test MAE | Test RMSE |
|---|---|---|---|---|
| Linear Regression | 0.9610 | 0.9424 | 0.4800 | 0.6495 |
| Ridge | 0.9582 | 0.9196 | 0.6114 | 0.7588 |
| Random Forest Regressor | 0.9879 | 0.8554 | 0.8650 | 1.0499 |
| Gradient Boosting Regressor | 1.0000 | 0.7858 | 0.9752 | 1.2535 |
| XGBoost | 1.0000 | 0.8314 | 0.9421 | 1.2347 |
| Multi-Layer Perceptron Regressor | 0.8211 | 0.2447 | 1.1929 | 1.6353 |

**Note.** $R^2$ = Coefficient of Determination; MAE = Mean Absolute Error; RMSE = Root Mean Squared Error. All metrics are rounded to four decimal places. The models were trained on historical climate and infrastructure data from Nigerian cities (2000–2024).
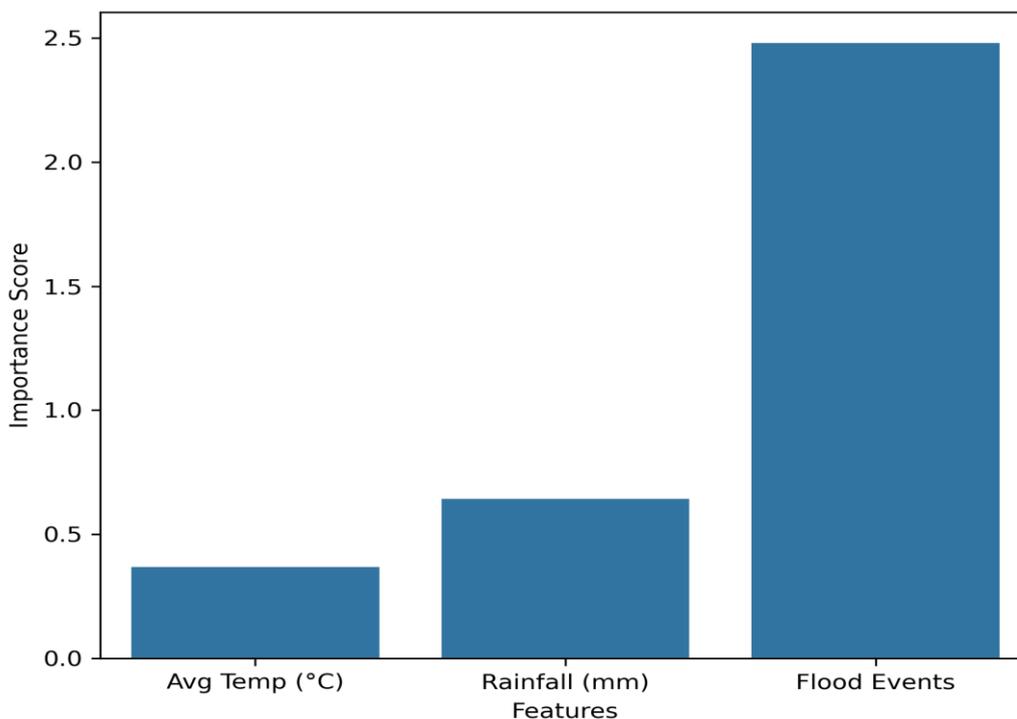
**Interpretation of Table 1.**

Table 1 presents the performance metrics of six supervised machine learning regression models—Linear Regression, Ridge Regression, Random Forest, Gradient Boosting, XGBoost, and Multi-Layer Perceptron (MLP)—used to predict infrastructure failure based on climate-related variables. The models were evaluated using $R^2$ (Train and Test), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), all rounded to four decimal places for clarity.

Among the models, Linear Regression outperformed the others with a Train $R^2$ of 0.9610 and a Test $R^2$ of 0.9424, indicating a strong fit and excellent generalization. It also recorded the lowest MAE (0.4800) and RMSE (0.6495), making it the most accurate and consistent model. Ridge Regression performed moderately well, though its slightly lower Test $R^2$ (0.9196) and higher errors show it was less precise. The ensemble models—Random Forest, Gradient Boosting, and XGBoost—achieved perfect or near-perfect fits on the training data but showed considerable drops in Test $R^2$ (0.8554, 0.7858, and 0.8314 respectively), indicating overfitting. Their higher error values further support this. The MLP Regressor performed the worst, with a low Test $R^2$ of 0.2447, the highest MAE (1.1929), and the highest RMSE (1.6353), suggesting it failed to capture meaningful patterns in the data.

Overall, **Linear Regression** is the best-performing model in this context. It offers the best trade-off between simplicity, interpretability, and predictive performance, without signs of overfitting, making it the most suitable choice for analyzing climate impacts on infrastructure and crucial for policy and infrastructure planning where transparency is needed.

**Feature Importance**

**Figure 2** Feature Importance in Best-Fit Linear Regression Model



**Interpretation of Figure 2.**

Figure 2 (feature importance) shows that **Flood Events** have the highest predictive influence on infrastructure outcomes, followed by **Rainfall (mm)** and **Average Temperature** (°C). The dominance of flood events highlights their critical role in determining infrastructure stress and damage, suggesting that climate-induced flooding is the most significant driver among the examined variables. This finding is further supported by the statistical modeling results, which confirm the significance of each variable's contribution to the model's predictive capacity.

**OLS Regression and Statistical Significance**

The OLS regression results reinforced the machine learning findings, achieving a test model performance of 94%. The model summary further enabled the evaluation of the statistical significance of each predictor. Specifically, the climatic predictors—Average Temperature, Rainfall, and Flood Events—were all found to be statistically significant based on their p-values and t-statistics. **Average Temperature (°C)** has a p-value of 0.025 and a t-statistic of 2.362, indicating significance at the 5% level. **Rainfall (mm)** presents a p-value of 0.022 and a t-statistic of 2.419, also statistically significant at the 5% level. **Flood Events** exhibits a highly significant effect, with a p-value of 0.000 and a t-statistic of 9.269, identifying it as the most influential climatic predictor. These results demonstrate that all three climatic variables have a statistically significant impact on infrastructure outcomes in the model.

On the other hand, the **Variance Inflation Factor (VIF)** was assessed to evaluate multicollinearity among the climatic predictors. Average Temperature (°C) recorded a VIF of 1.82, which is well below the common threshold of 5, indicating low multicollinearity and suggesting that it is not strongly correlated with the other predictors. Rainfall (mm) and Flood Events exhibited VIF values of 5.29 and 5.34, respectively. These values reflect moderate multicollinearity, which is acceptable and expected, given the natural relationship between high rainfall and increased flood events. This level of correlation does not pose a significant concern and supports the logical interdependence of these variables in climate-impact studies. Overall, the VIF results confirm that the climatic predictors are within acceptable multicollinearity limits, supporting the reliability of their statistical significance in the model.

# CONCLUSION

This study provides a comprehensive analysis of the relationship between climatic variables and urban infrastructure failure in six Nigerian cities over a 25-year period. By integrating both statistical and machine learning approaches, the research bridges the gap between predictive accuracy and interpretative depth. The findings reveal that **flood events and rainfall** are the most significant climatic stressors influencing infrastructure failure, while **temperature** plays a more modest but statistically relevant role.

Among the machine learning models tested, **Linear Regression** demonstrated the best balance of accuracy, generalization, and interpretability, making it the most suitable model for real-world applications. Importantly, the study confirms that climate-induced infrastructure failures are not random but follow measurable patterns tied to specific environmental conditions. This highlights the urgent need for localized, data-driven planning tools in urban infrastructure management, especially in developing countries like Nigeria, where resilience frameworks remain underdeveloped.

By fulfilling all outlined objectives, this study demonstrates that combining statistical rigor with machine learning offers a practical and interpretable pathway for predicting infrastructure vulnerabilities under varying climatic conditions. This integrated framework contributes to both academic knowledge and real-world policy development, offering actionable insights for climate-resilient infrastructure planning and disaster risk reduction.

## RECOMMENDATIONS

Based on the findings of this study, several practical and policy-relevant recommendations are proposed to strengthen climate-resilient infrastructure planning in Nigeria and similar contexts:

1. **Integrate climate data into infrastructure design and maintenance**: Engineers and urban planners should routinely incorporate historical and predictive climate data, especially rainfall and flood frequency, into infrastructure risk assessments and design standards.
2. **Adopt predictive analytics in urban planning agencies**: Government agencies and local authorities should invest in capacity-building for data-driven planning by integrating machine learning tools, such as linear regression models, to forecast climate-related infrastructure risks and prioritize mitigation efforts.

3. **Strengthen institutional frameworks for climate-resilient planning**: There is a need for improved policy coordination among ministries of environment, works, and urban planning to ensure that infrastructure development is aligned with climate adaptation strategies.

4. **Promote localized early warning systems and flood response strategies**: Policymakers should allocate resources for real-time climate monitoring systems and community-level disaster preparedness programs, especially in high-risk urban areas identified in this study.

5. **Encourage interdisciplinary collaboration**: Academic researchers, civil engineers, urban planners, and policymakers should collaborate to develop integrated models that inform infrastructure resilience under climate change, leveraging both technical evidence and policy foresight.

# REFERENCES

1. Adebayo, A. A. (2018). *Urban flooding and disaster risk reduction in Nigeria*. Journal of Environmental Management, 32(2), 101–110.
2. Adeyemi, K., & Okonkwo, T. (2022). *Assessment of climate variability impacts on urban infrastructure in Nigeria*. Nigerian Journal of Environmental Sciences, 10(1), 45–59.
3. Akintola, O. A. (2017). *Climate change and its impact on civil infrastructure: A Nigerian perspective*. Civil Engineering and Environmental Systems, 34(3), 231–246.
4. Gujarati, D. N. (2003). *Basic econometrics* (4th ed.). McGraw-Hill Education.
5. IPCC. (2021). *Climate change 2021: The physical science basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. https://www.ipcc.ch/report/ar6/wg1/
6. NEMA. (2023). *Annual flood disaster report in Nigeria*. National Emergency Management Agency.
7. NiMet. (2023). *2023 Seasonal climate prediction*. Nigerian Meteorological Agency. https://www.nimet.gov.ng/
8. Onuoha, K. M., Eze, B. E., & Uche, J. O. (2019). *Climate extremes and infrastructure resilience in Sub-Saharan Africa*. Environmental Hazards, 18(4), 289–305.
9. Hamood, M., Zhang, Y., & Khan, M. A. (2022). *Ensemble learning approaches for financial risk prediction: A comparative analysis*. Journal of Financial Data Science, 4(1), 22–35.
10. UN-Habitat. (2020). *World cities report 2020: The value of sustainable urbanization*. United Nations Human Settlements Programme (UN-Habitat). https://unhabitat.org/wcr/