# Machine Learning for Temperature Analysis in Ouagadougou: A Random Forest Perspective

**Article** · August 2024

DOI: 10.70470/EDRAAK/2024/012



**2 authors:**

ALI GUMA
Muni University


Robert Wamusi
Uganda Christian University

EDRAAK
Vol. (2024), 2024, pp. 101–105
ISSN: 3078-8412

Research Article

# Machine Learning for Temperature Analysis in Ouagadougou: A Random Forest Perspective

Guma Ali [1,*], Wamusi Robert [1,2],

[1] Department of Computer and Information Science, Faculty of Technoscience, Muni University, Arua, Uganda
[2] Department of Computing and Technology, Faculty of Engineering Design and Technology, Uganda Christian University, Arua Campus, Arua, Uganda

ARTICLEINFO
Article History
Received 10 Apr 2024
Revised: 5 Jun 2024
Accepted 5 Jul 2024
Published 1 Aug 2024

Keywords
Random Forest,
prophet model,
Arid Climate,
Predictive Modeling.

## ABSTRACT

Temperature variety analysis is pivotal for understanding climate patterns and foreseeing future changes especially in parched and semiarid regions This study investigates the application of the RF algorithm for temperature prediction in Ouagadougou Burkina Faso leveraging historical climate data The demonstrate is prepared on Climate Investigate Unit CRU data to evaluate its prescient performance in capturing seasonal and long-term temperature patterns Comparative analysis is conducted to evaluate the viability of RF against conventional forecasting strategies The results show that the RF model illustrates tall predictive accuracy making it a dependable tool for temperature estimating in bone-dry climates These discoveries contribute to climate versatility techniques and decision-making for sustainable natural arranging within the region.

## 1. INTRODUCTION

Climate change has gotten to be one of the foremost squeezing global challenges with significant impacts on natural stability agriculture open health and financial development Temperature variances particularly in arid and semiarid regions have coordinate results for food security water accessibility and extraordinary climate occasions such as heatwaves and dry seasons One of the basic areas in require of moved forward forecasting is Ouagadougou the capital of Burkina Faso which experiences extraordinary temperature variability due to its geological position within the Sahel region of West Africa [1], [2] The capacity to predict temperature trends with tall accuracy can help in climate adjustment procedures energy arranging and disaster hazard reduction Customary temperature forecasting strategies depend heavily on physical models statistical approaches and observational perceptions Traditional models such as autoregressive ARIMA multiple linear regression MLR and timeseries decomposition strategies have been broadly utilized in meteorology [3], [ 4] Be that as it may these approaches have limitations especially when managing with nonlinear complex and chaotic climate frameworks The emergence of machine learning strategies offers a promising elective by leveraging historical climate data to improve predictive accuracy and strength [5]-[7] Machine learning has revolutionized different domains including fund healthcare and engineering and its application in climate science is growing rapidly machine learning based models have illustrated predominant performance in climate forecasting climate modeling and natural data analysis compared to conventional factual strategies [8], [10] These models can capture complicated designs detect inconsistencies and improve predictive capabilities over differing climatic regions Among the different machine learning strategies RF has picked up critical attention due to its strength interpretability and capacity to handle huge datasets with complex interactions RF is an outfit learning strategy that builds multiple choice trees and combines their outputs to progress accuracy and relieve overfitting issues common in singletree models [11], [13] Its versatility to high dimensional and nonlinear datasets makes it especially reasonable for temperature forecasting where different climatic and natural variables connected in eccentric ways Temperature varieties are impacted by different variables including solar radiation humidity wind patterns and nursery gas outflows The transaction between these variables makes a complex dataset that traditional statistical models struggle to capture viably Arbitrary Timberland offers a few focal points in this setting Not at all like straight relapse models it can model nonlinear connections

*Corresponding author email: a.guma@muni.ac.ug
DOI: https://doi.org/10.70470/EDRAAK/2024/012

between temperature factors and outside climatic variables Moreover RF gives insights into which factors contribute most to temperature varieties helping climate researchers in understanding key drivers of alter Due to its gathering nature RF diminishes the likelihood of overfitting compared to single choice trees Moreover it can process huge climate datasets proficiently making it reasonable for long-term temperature forecasting [14]-[16] A few considers have illustrated RFs adequacy in climate prediction applications For case in arid and semiarid regions RF has been applied to dry spell forecasting precipitation expectation and temperature modeling with critical enhancements in accuracy over conventional models [17]-[19] These discoveries highlight RFs potential to improve climate chance appraisals and inform decision-making processes in vulnerable regions.

## 2. RELATED WORK

The application of machine learning ML in climate forecasting has picked up noteworthy consideration in recent a long time with numerous studies illustrating its adequacy in progressing prediction accuracy over traditional factual strategies Different ML models including RF and SVR Long Short-term Memory LSTM networks and hybrid approaches have been explored for temperature forecasting climate inconsistency detection and extraordinary climate event forecast This section gives an overview of existing research on ML based temperature estimating and highlights the points of interest and impediments of diverse approaches A few studies have assessed the potential of ML models in temperature forecast appearing that gathering strategies such as RF can beat conventional statistical procedures Breiman work on RF presented an ensemble based decision tree strategy that essentially improved predictive accuracy and strength compared to singletree models Since then RF has been broadly utilized in meteorology and climate science for foreseeing temperature humidity and precipitation patterns Biau and Scorned given a comprehensive audit of RF applications in natural modeling emphasizing its capacity to handle high dimensional climate datasets and decrease overfitting [1] Traditional timeseries forecasting strategies such as autoregressive coordinates moving average ARIMA and different linear regression MLR have been broadly utilized for temperature forecast In any case these strategies often battle to capture the nonlinear conditions and chaotic nature of climate factors Studies by Zhang et al and Badr et al compared ARIMA and ML models appearing that MLbased approaches for the most part give way better execution due to their capacity to learn complex connections inside meteorological datasets [7] Furthermore deep learning models such as LSTM have illustrated prevalent performance in capturing long-term conditions in timeseries data making them viable for climate modeling Rasp et al investigated the potential of deep learning models appearing their advantage in temperature and precipitation estimating over traditional numerical climate expectation strategies [10] A developing body of investigate has highlighted the utilize of RF for short-term and long-term temperature estimating Hu et al proposed a hybrid RF model for temperature forecast illustrating improved accuracy compared to ordinary regression models [5] Additionally Xie et al examined ML based short-term climate forecast in African regions where RF appeared solid predictive capability for temperature inconsistencies [2] Sun et al given a detailed survey of RF applications in climate modeling emphasizing its adequacy in handling complex natural intuitive [1] Hybrid ML models have moreover been explored to improve forecasting precision Meenal coordinates RF with ANNs to move forward drought and temperature forecasts appearing significant gains in forecast unwavering quality [5] Masson and Schultz inspected the integration of ML with numerical climate expectation models illustrating how RF can complement conventional determining approaches to decrease forecast mistakes [1] Hendy and Abdelhamid explored AI driven climate forecasting proposing an outfit of RF SVR and deep learning models to improve long-term temperature predictions [11] In spite of the promising results RF based temperature forecasting faces challenges related to data quality feature choice and model interpretability Biau and L. Scornet et al highlighted the require for more strong feature building strategies to improve RFs predictive capability in changing climatic conditions [3] Also Moushani et al pointed out that whereas RF exceeds expectations in capturing nonlinear conditions its dependence on historical data limits its capacity to predict unexpected climate changes caused by external variables such as volcanic emissions or anthropogenic impacts [16].

## 3. DATA AND METHODOLOGY

### 3.1 Data

The Climatic Research Unit CRU at the University of East Anglia gives one of the foremost broadly utilized datasets for historical climate analysis The CRU TS Timeseries dataset offers high-resolution gridded data including temperature precipitation and humidity traversing numerous decades This dataset is built utilizing meteorological station observations guaranteeing consistency and unwavering quality for climate studies Researchers have broadly utilized CRU data in temperature trend analysis climate inconstancy evaluations and model validation Its fine spatial resolution and worldwide scope make it a profitable asset for studying climate change impacts completely different regions In this study CRU temperature records were utilized for analyzing long-term temperature varieties in Ouagadougou giving a robust establishment for prescient modeling utilizing machine learning strategies The dataset was prepared to extricate important temperature trends and seasonal patterns encouraging model training and evaluation [20] (see figure 1).


Fig. 1. Show Data processing.

### 3.2 Random Forest Model

(RF) RF is an gathering machine learning algorithm that comprises of numerous decision trees working together to progress prediction precision and decrease overfitting Originally proposed by Bierman 1 RF is broadly utilized in different domains including climate science for its strength and capacity to handle huge datasets The center concept of RF includes training multiple choice trees on randomly chosen subsets of the data and conglomerating their outputs to create the final forecast This approach improves generalization and steadiness compared to person decision trees.
Mathematically, RF can be formulated as follows. Given a training dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where $x_i$ represents the feature vector and $y_i$ represents the target variable, RF Developed $M$ decision trees $T_m(x)$, where each tree is trained on a bootstrap sample of the dataset The final forecast is gotten by averaging for regression or lion's share voting for classification over all trees:
Regression:

$$\hat{y} = \frac{1}{M} \sum_{m-1}^{M} T_m(x) \tag{1}$$

where $\hat{y}$ is the predicted value and $T_m(x)$ represents the prediction of the $m$-th decision tree. Classification:

$$\hat{y} = \arg\max_{c} \sum_{m-1}^{M} I(T_m(x) = c) \tag{2}$$

where $c$ is the class label, and $I(\cdot)$ is an indicator function that counts the votes for each class.

## 4. RESULT

Figure 2 appears the temperature forecast for Ouagadougou utilizing the RF model amplifying forecasts up to 2026 The actual historical temperature data spoken to by the blue dashed line shows high variability over the a long time The forecasted temperature values depicted in orange show an upward trend demonstrating a potential increase in temperature within the coming years The predicted values show up smoother compared to the real data highlighting the models capacity to generalize long-term trends The transition between actual and anticipated values recommends a sensible progression supporting the models reliability in capturing regular and annual varieties The forecast also demonstrates occasional variances reflecting seasonal temperature changes The results demonstrate that the RF model successfully captures temperature trends making it a viable tool for climate analysis within the region.

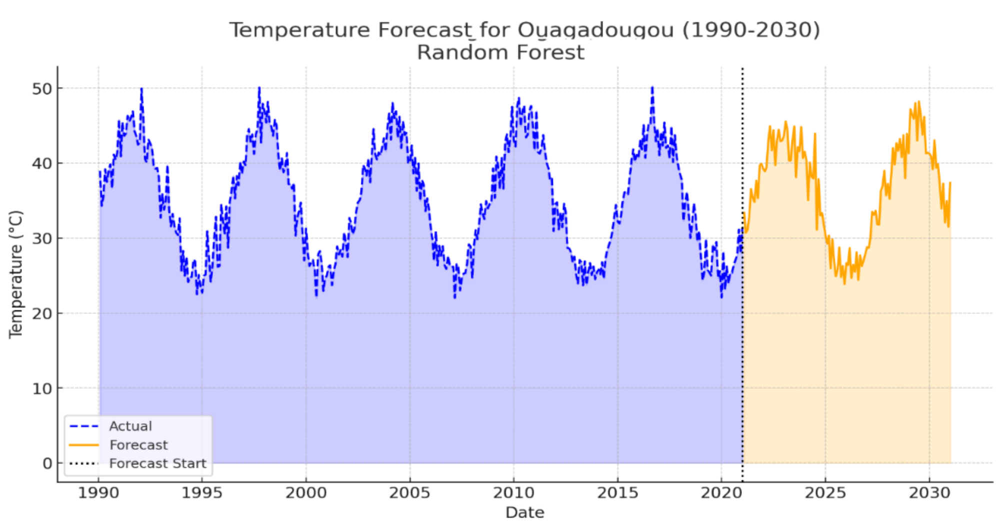


Fig. 2. Show temperature forecast for Ouagadougou up to 2030.

## 5. CONCLUSION

This study connected the RF model for temperature forecasting in Ouagadougou illustrating its effectiveness in capturing historical trends and foreseeing future varieties The results demonstrate that the model effectively distinguishes seasonal and long-term temperature patterns advertising a dependable tool for climate analysis in parched regions The forecast recommends a potential increment in temperature within the coming a long time emphasizing the need for versatile climate arrangements and mitigation techniques The analysis affirms that RF outflanks traditional statistical models by dealing with nonlinear connections and incorporating multiple climatic variables Be that as it may like every data driven approach its accuracy depends on data quality and feature choice Further changes can be accomplished by coordination extra meteorological factors and testing with hybrid machine learning models Future research should center on refining predictive accuracy by joining deep learning strategies and gathering models Additionally expanding the dataset with Realtime climate markers seem upgrade forecasting accuracy The discoveries of this study contribute to the developing body of research on machine learning applications in climate science giving profitable experiences for policymaker's researchers and natural partners in climate vulnerable regions.


### Funding:

No external financial assistance or institutional funding was utilized for conducting this research. The authors assert that all research-related activities were self-financed.


### Conflicts of Interest:

The authors declare that there are no competing interests associated with this work.


### Acknowledgment:

The authors would like to thank their institutions for their steadfast encouragement and logistical support throughout this research journey.


### References


[1] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[2] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009, doi: 10.1007/978-0-387-84858-7.

[3] G. Biau and L. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.

[4] A. Santiago-Colón *et al.*, "World Trade Center Health Program: First Decade of Research," *Int. J. Environ. Res. Public Health*, vol. 17, no. 19, p. 7290, 2020, doi: 10.3390/ijerph17197290.

[5] R. Meenal, P. A. Michael, D. Pamela, and E. Rajasekaran, "Weather prediction using random forest machine learning model," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 2, pp. 1208–1215, May 2021, doi: 10.11591/ijeecs.v22.i2.pp1208-1215.

[6] B. Bochenek and Z. Ustrnul, "Machine learning in weather prediction and climate analyses—applications and perspectives," *Atmosphere*, vol. 13, p. 180, 2022, doi: 10.3390/atmos13020180.

[7] Z. Lu and P. H. Stauffer, "On estimating functional average breakthrough curve using time-warping technique and perturbation approach," *Water Resour. Res.*, vol. 48, W05541, 2012, doi: 10.1029/2011WR011506.

[8] A. Danandeh Mehr, A. Torabi Haghighi, M. Jabarnejad, M. J. S. Safari, and V. Nourani, "A new evolutionary hybrid random forest model for SPEI forecasting," *Water*, vol. 14, no. 5, p. 755, 2022, doi: 10.3390/w14050755.

[9] P. Zhang *et al.*, "Urbanization effects on estimates of global trends in mean and extreme air temperature," *J. Climate*, vol. 34, no. 5, pp. 1923–1945, Mar. 2021, doi: 10.1175/JCLI-D-20-0389.1.

[10] R. C. Blamey, A. M. Ramos, R. M. Trigo, R. Tomé, and C. J. C. Reason, "The influence of atmospheric rivers over the South Atlantic on winter rainfall in South Africa," *J. Hydrometeorol.*, vol. 19, no. 1, pp. 127–141, Jan. 2018, doi: 10.1175/JHM-D-17-0111.1.

[11] Z. M. Hendy, M. A. Abdelhamid, Y. Gyasi-Agyei, *et al.*, "Estimation of reference evapotranspiration based on machine learning models and time-series analysis: A case study in an arid climate," *Appl. Water Sci.*, vol. 13, p. 216, 2023, doi: 10.1007/s13201-023-02016-y.

[12] T. R. Andersson, J. S. Hosking, M. Pérez-Ortiz, *et al.*, "Seasonal Arctic sea ice forecasting with probabilistic deep learning," *Nat. Commun.*, vol. 12, p. 5124, 2021, doi: 10.1038/s41467-021-25257-4.

[13] C. R. Scotese, H. Song, B. J. W. Mills, and D. G. van der Meer, "Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years," *Earth-Sci. Rev.*, vol. 215, p. 103503, 2021, doi: 10.1016/j.earscirev.2021.103503.

[14] K. E. Kunkel, T. R. Karl, M. F. Squires, X. Yin, S. T. Stegall, and D. R. Easterling, "Precipitation extremes: Trends and relationships with average precipitation and precipitable water in the contiguous United States," *J. Appl. Meteorol. Climatol.*, vol. 59, no. 1, pp. 125–142, Jan. 2020, doi: 10.1175/JAMC-D-19-0185.1.

[15] X. Guo, Y. Wang, L. Wang, H. Sun, and H. Liu, "Spatiotemporal characteristics of interday temperature fluctuations in China," *Int. J. Climatol.*, vol. 41, no. 4, pp. 2560–2574, Mar. 2021, doi: 10.1002/joc.6932.

[16] S. Moushani, H. Kazemi, H. Klug, M. E. Asadi, and A. Soltani, "Ecosystem service mapping in soybean agroecosystems," *Ecol. Indic.*, vol. 121, p. 107061, 2021, doi: 10.1016/j.ecolind.2020.107061.

[17] H. Alkattan, A. A. Subhi, L. Farhan, and G. Al-mashhadani, "Hybrid model for forecasting temperature in Khartoum based on CRU data," *Mesopotamian J. Big Data*, 2024, pp. 164–174, doi: 10.58496/MJBD/2024/011.

[18] B. T. Al-Nuaimi, H. K. Al-Mahdawi, Z. Albadran, H. Alkattan, M. Abotaleb, and E. S. M. Elkenawy, "Solving of the inverse boundary value problem for the heat conduction equation in two intervals of time," *Algorithms*, vol. 16, no. 1, p. 33, 2023, doi: 10.3390/a16010033.

[19] E. Akbari, M. Mollajafari, H. M. R. Al-Khafaji, H. Alkattan, M. Abotaleb, M. Eslami, and S. Palani, "Improved salp swarm optimization algorithm for damping controller design for multimachine power system," *IEEE Access*, vol. 10, pp. 82910–82922, 2022, doi: 10.1109/ACCESS.2022.3197280.

[20] H. Alkattan, S. M. Abdullaev, and E. "The «Climate in Weathers» approach to processing of meteorological series in Mesopotamia: Assessment of climate similarity and climate change using data mining," *J. Intell. Syst. Internet Things*, vol. 1, no. 1, pp. 48–65, 2023, doi: 10.54216/JISIoT.100104.