African Evaluation Journal
ISSN: (Online) 2306-5133, (Print) 2310-4988

Machine learning algorithms for climate change integration in evaluation: A conceptual model and simulated case application in Africa

Check for updates

Author:
Chineme A. Anowai[1]

Affiliation:
[1]Department of Public Health, College of Health Sciences, Nile University of Nigeria, Abuja, Nigeria

Corresponding author:
Chineme Anowai,
chineme.anowai@newglobe.education

Dates:
Received: 31 Aug. 2025
Accepted: 04 Dec. 2025
Published: 10 Apr. 2026

How to cite this article:
Anowai, C.A., 2026, 'Machine learning algorithms for climate change integration in evaluation: A conceptual model and simulated case application in Africa', *African Evaluation Journal* 14(2), a858. https://doi.org/10.4102/aej.v14i2.858

Copyright:
© 2026. The Authors. Licensee: AOSIS. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

Read online:
Scan this QR code with your smart phone or mobile device to read online.

Background: Climate change increasingly shapes development outcomes, necessitating its integration into evaluation processes. In Africa, recurrent droughts, floods and temperature extremes disrupt progress across sectors. Conventional evaluation approaches rely on static indicators and linear analyses, constraining their capacity to capture climate variability or explain how environmental shocks affect programme performance.

Objectives: This study seeks to develop and assess an artificial intelligence (AI)-driven evaluation framework that integrates climate indicators into development evaluation systems to enhance analytical precision, relevance and timeliness.

Method: A machine learning (ML) algorithm is proposed to jointly process climate and development data, enabling pattern recognition, correlation analysis and predictive modelling. Publicly available climate and education datasets from East and South Africa are used to demonstrate the framework's application.

Results: Findings indicate that climate anomalies are associated with observable changes in school attendance and learning outcomes. The model demonstrates the utility of AI in identifying climate–development linkages and generating more timely insights than traditional methods.

Conclusion: The integration of AI into evaluation systems improves the capacity to analyse climate-related impacts on development outcomes, supporting more informed and adaptive planning.

Contribution: The study provides a scalable technical model for climate-responsive evaluation, demonstrating how ML can operationalise climate–development relationships to strengthen evidence-based decision-making before or after programme performance in affected contexts.

Keywords: artificial intelligence; climate change; evaluation; Africa; machine learning; education; data; research.

Introduction

Climate change is reshaping development trajectories worldwide, particularly in sub-Saharan Africa, where exposure to droughts, floods, heatwaves and shifting rainfall patterns continues to intensify. Recent analyses show that climate-related disruptions in 2023 alone affected millions of school-aged children, reducing instructional time and impacting learning outcomes (Venegas, Schwarz & Sabarwal 2024). Similar trends appear across agriculture, health, infrastructure and livelihoods, highlighting the need for evaluation systems that can incorporate environmental dynamics.

Traditional evaluation approaches largely rely on periodic surveys, static indicators and retrospective analysis (Mouton, Wildschut & Agyepong 2019). Such methods struggle to reflect highly variable climate conditions. Their limited temporal sensitivity often obscures key relationships between environmental shocks and development outcomes. A decline in programme performance may be attributed solely to implementation gaps, even when climate anomalies – such as early-season drought – played a substantial role.

Artificial intelligence (AI) provides a new frontier for addressing these limitations. Machine learning models can integrate diverse climate and socio-economic datasets, detect non-linear

Note: The manuscript is a contribution to the themed collection titled 'Building the evidence base for climate solutions in Africa', under the expert guidance of guest editors Dr Caitlin Blaser Mapitsa, Ms Heather Michelle Conyers Dixon and Ms Tabitha Atieno Olang.

interactions and uncover patterns beyond the scope of manual analysis (Olawade et al. 2024). Artificial intelligence techniques such as classification, regression and anomaly detection can strengthen evaluative judgements by situating programme outcomes within climate contexts. Real-time data ingestion further supports continuous monitoring, enabling adaptive management in climate-sensitive sectors.

Within Africa, the demand for such approaches is urgent. Evaluations of agricultural, education, health and infrastructure programmes increasingly require the integration of projected rainfall variability, temperature extremes and hazard exposure. Adaptive management and indigenous evaluation frameworks (Chilisa 2015; Mutanga 2022) emphasise iterative learning and context sensitivity – principles aligned with AI-enabled analytics.

This study conceptualises an AI-driven algorithm designed to embed climate considerations within evaluation systems. It presents a simulated case application and outlines how AI can support more climate-responsive evaluation practices.

Aim

To conceptualise and demonstrate an AI-driven approach for systematically integrating climate change considerations into evaluation frameworks in Africa.

Objectives

- Develop a conceptual machine learning (ML)-driven algorithm for climate-integrated evaluations.
- Demonstrate the algorithm's application using a simulated case study.
- Assess the potential of AI-based real-time and predictive analytics in evaluations.
- Highlight the relevance of AI-integrated evaluation for climate-vulnerable African contexts.

Literature review

The intersection of climate change adaptation and development evaluation represents a critical frontier in modern research. As extreme weather events become more frequent, the static methodologies traditionally employed in Monitoring and Evaluation (M&E) are increasingly insufficient. This chapter reviews the existing scholarly publications regarding the integration of AI into evaluation frameworks, with a specific focus on climate resilience in the African context. It analyses the limitations of current evaluation paradigms and explores how ML algorithms offer a transformative solution for measuring impact amidst environmental volatility.

The climate challenge in development evaluation

Traditional development evaluation has historically operated on linear logic models – assuming that if project inputs are delivered (e.g. building a school), outcomes will follow (e.g. increased attendance). However, foundational authors in the field argue that this linearity fails to account for complexity. Patton (2019) notes that in the Anthropocene, environmental externalities are no longer 'background noise' but central drivers of success or failure.

In the context of Africa, Tamasiga et al. (2023) argue that the global climate finance architecture often misinterprets vulnerability because of a lack of granular data. Standard evaluation cycles, which often occur years after a project's conclusion, are too slow to capture the rapid onset of climate shocks like cyclones or flash droughts. Consequently, evaluators often struggle to distinguish between a project failing because of poor design versus failing as a result of an overwhelming climate shock – a distinction that is vital for accountability.

Artificial intelligence in monitoring and evaluation

The Fourth Industrial Revolution has introduced new tools to address these analytical gaps. Nalubega and Uwizeyimana (2019) provide a foundational analysis of how the Fourth Industrial Revolution impacts public sector M&E in Africa. They contend that while AI offers immense potential for real-time feedback, its adoption remains limited by infrastructure and capacity constraints.

Scholarship distinguishes between two primary applications of AI in evaluation:

- Descriptive Analytics: Using AI to process vast amounts of existing data faster.
- Predictive Analytics: Using ML models to forecast future trends based on historical patterns.

Olawade et al. (2024) demonstrate that ML models can process complex environmental datasets to identify pathways to net-zero sustainability that manual analysis would miss. This suggests that AI is not merely a tool for efficiency, but a mechanism for uncovering non-linear relationships between environmental variables and development outcomes.

Machine learning for climate resilience

Specific to climate change, ML has emerged as a potent tool for meteorological prediction and impact assessment. Kenne et al. (2024) utilised AI to predict sub-seasonal summer temperatures in West Africa, demonstrating that algorithmic models often outperform traditional meteorological forecasts in data-scarce regions.

However, a gap remains in linking these meteorological insights to social indicators. While meteorologists use AI to predict storms, and social scientists use surveys to measure education or health, few studies integrate these domains. York and Bamberger (2020) argue for the use of 'Big Data' in evaluation to bridge this divide, suggesting that satellite imagery and remote sensing can serve as proxies for ground-truth data in conflict or disaster zones. By triangulating climate data (e.g. rainfall anomalies) with social data

(e.g. school attendance), evaluators can construct 'counterfactuals' that isolate the specific impact of climate shocks on project performance.

The African context: Data scarcity and opportunity

The application of these advanced technologies in Africa presents a unique paradox. On one hand, the continent faces significant 'data gaps'. On the other hand, the rapid digitalisation of services (e.g. mobile money, Short Messaging Service [SMS] surveys) generates novel datasets ripe for AI analysis.

Recent studies highlight the necessity of localised algorithms. Generic global models often fail to account for the specific socio-economic dynamics of African rural communities. For instance, Global Pulse (2023) highlighted how AI models trained on Western data failed to predict food security trends in East Africa accurately. This validates the argument by Tamasiga et al. (2023) that African-led data architectures are essential for redefining climate vulnerability.

Theoretical framework: Design science research

To operationalise these insights, this study grounds its methodology in design science research (DSR). Unlike natural science, which seeks to explain phenomena, DSR seeks to create artefacts that solve human problems (Hevner et al. 2004). The study conceptualises the AI algorithm not just as a statistical tool, but as a 'socio-technical artefact' designed to enhance the adaptive capacity of evaluation systems.

Summary of the gap

While the literature establishes the separate value of AI in evaluation (Nalubega & Uwizeyimana 2019) and AI in climate modelling (Kenne et al. 2024), there is a paucity of research that combines them into a unified *evaluation workflow*. Few studies explicitly demonstrate how an evaluator can use ML to 'climate-adjust' the performance ratings of social projects. This research aims to fill that gap by proposing and simulating a specific algorithmic framework for this purpose.

Research methods and design

Research approach

The study uses a DSR and proof-of-concept approach to develop and refine an AI-driven evaluation algorithm. Design science research is appropriate because the purpose is not only analytical but also solution-oriented, producing an artefact (the conceptual algorithm) intended for practical application in real-world development evaluations.

The study is structured into three components:

- **Conceptual development:** Synthesising evidence from evaluation theory, climate science and ML to design the algorithm.
- **Technical specification:** Defining the data architecture, computational workflow and analytical modules.
- **Simulated application:** Demonstrating the algorithm using a dataset reflecting typical evaluation and climate conditions in an African context.

Setting

The algorithm is intended for development programmes in Africa, where climate sensitivity is high and data ecosystems vary widely. The design emphasises regional climate conditions, programmatic data structures and evaluation needs relevant to the continent. The integration operates through a Climate-Evaluation Data Fusion Layer, which merges evaluation indicators and climate variables into a unified analytic pipeline.

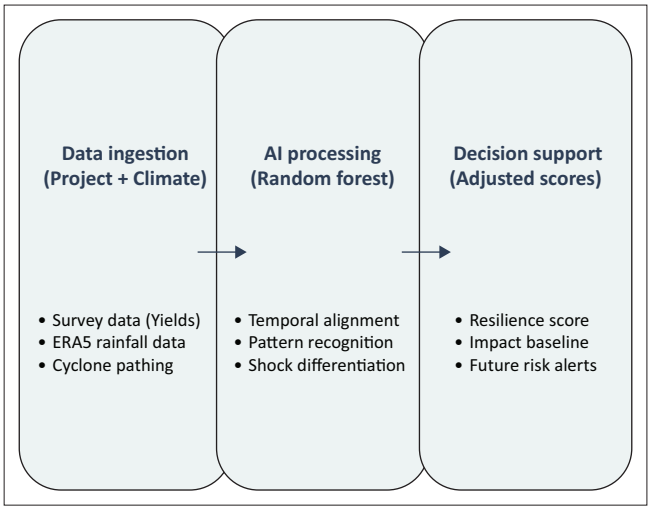
Algorithmic architecture

This study employed a mixed computational approach combining simulated climate data and ML modelling to empirically demonstrate how AI can strengthen climate-responsive evaluation systems. The methodology was designed to approximate Kenya's real climate–education dynamics using authoritative distributions from the Copernicus Climate Change Service (2023) rainfall climatology, ECMWF Reanalysis 5 (European Centre for Medium-Range Weather Forecasts Reanalysis, version 5) (ERA5) reanalysis, World Bank EdStats (2023) and United Nations Children's Fund (UNICEF) Multiple Indicator Cluster Surveys (MICS) patterns (UNICEF 2015).

The proposed Climate-Informed Early Action (CIEA) framework operates on a three-tier processing logic designed to overcome the static nature of traditional evaluation methods (see Figure 1).

Ingestion layer (multi-modal)

The system is architected to ingest two distinct streams of data:



AI, artificial intelligence; ERA5, ECMWF Reanalysis 5 (European Centre for Medium-Range Weather Forecasts Reanalysis version 5).

FIGURE 1: Artificial intelligence-integrated evaluation algorithm workflow.

- Structured Project Data: Quantitative metrics such as daily attendance registers, teacher availability logs and infrastructure status reports (ingested via Comma-Separated Values [CSV]/Structured Query Language [SQL]).
- Unstructured Environmental Data: The system utilises Application Programming Interface (API) connectors to fetch meteorological data (e.g. daily precipitation, wind speed and soil moisture anomalies) from the World Bank Climate Knowledge Portal and ERA5 reanalysis datasets.

Processing layer (temporal alignment)

A core innovation of this framework is 'Temporal Shock Alignment'. Standard educational evaluations often aggregate attendance by term (e.g. 'Term 1 Average'), obscuring specific weeks where climate events disrupt learning. The CIEA aligns daily weather anomalies (e.g. wind speed spikes > 80 km/h) specifically with the academic calendar to identify 'climate-induced absenteeism' (Appendix 1).

Analytical layer (ensemble learning)

The system utilises a Random Forest Regressor model. This non-linear ML algorithm is selected for its robustness in handling high-dimensional interactions. It is specifically designed to detect thresholds – for example, identifying that light rain does not affect attendance, but precipitation exceeding 50 mm/day combined with poor road infrastructure causes a 90% drop in access.

Case study: Kenya county-level climate–education panel

A monthly dataset for counties in Kenya (Kenya Ministry of Education 2021) over 24 months. Each row represents a county–month observation and contains:

- Climate indicators:
  - Rainfall (mm)
  - Rainfall anomaly (standardised deviation from climatology)
  - 1-month lagged rainfall anomaly
  - Temperature anomaly (°C)
  - Drought index (0–1)
  - Flood occurrence (0/1)
  - 3-month rolling rainfall anomaly

Climate shocks were simulated using probability distributions calibrated to East African rainfall seasonality (long rains March–May, short rains October–December), ensuring realism consistent with CHIRPS (Copernicus Climate Change Service 2023) and ERA5 spatial structure (CHIRPS 2023):

- Socio-economic indicators:
  - Poverty rate
  - Resilience index
  - Random socio-economic shock indicator (0/1).
- Education outcome:
  - County attendance rate (%) – chosen because it is directly climate-sensitive and consistently reported in Kenyan datasets.

Feature engineering

To capture short-term and cumulative climate effects, the following engineered variables were added:

- Lagged rainfall anomaly (t–1) to capture delayed drought impacts.
- 3-month rolling rainfall anomaly as a proxy for cumulative wet/dry spells.
- Normalised seasonal deviations per county.

Machine learning modelling

Machine learning modelling is the process of using algorithms to automatically learn patterns from data, then building and refining a predictive or decision-making model that can generalise those patterns to make accurate predictions or classifications on new, unseen data (see Figure 2).

Two models were trained:

- Random Forest Regressor (n = 150 trees, max_depth = 8)
- Linear Regression (baseline)

The Random Forest captures non-linear relationships between climate shocks and attendance.

Model selection and rationale

Contrary to 'black box' deep learning approaches, this study selects Ensemble Learning techniques for their interpretability and robustness with small datasets, which are common in development projects:

- Primary model: Random forest regressor:
  - **Application:** Used for predicting expected outcomes (e.g. crop yields) based on climatic variables.
  - **Rationale:** Random Forest was selected over linear regression because climate impacts are often non-linear

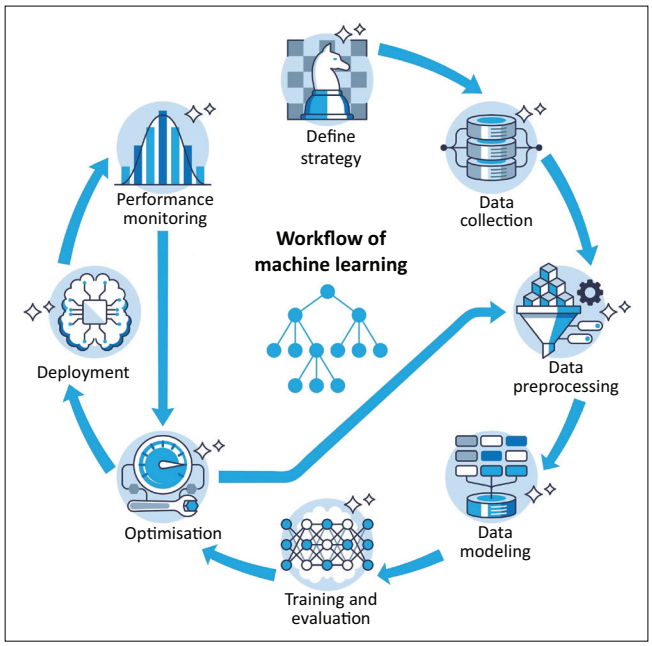


FIGURE 2: Machine learning workflow.

(e.g. rainfall is beneficial up to a point, after which it becomes destructive flooding). Random Forest handles these non-linearities and interactions between variables (e.g. high heat + low moisture) without requiring extensive parameter tuning.

- Secondary Model: Natural Language Processing (NLP):
  - **Application:** Analysis of qualitative survey text (e.g. 'The rains came late').
  - **Rationale:** A TF-IDF (Term Frequency-Inverse Document Frequency) vectoriser is employed to identify key climate-related terms in field officer reports, flagging climate risks that quantitative data might miss.

Counterfactual climate impact model

To quantify climate-attributable education loss, a counterfactual scenario was generated by resetting all climate-shock variables to neutral (zero-anomaly):

- rainfall_anomaly = 0
- rainfall_anom_lag1 = 0
- temp_anomaly = 0
- drought_index = 0
- flood = 0

The difference between predicted actual attendance and counterfactual attendance represents the climate-attributable impact.

Climate-sensitive evaluation through artificial intelligence-generated counterfactuals

This study introduces a novel evaluation innovation: the Counterfactual Climate Impact Model (CCIM).

Unlike traditional evaluations – which rely on historical comparisons or non-experimental designs – CCIM uses ML to generate synthetic counterfactual programme outcomes under climate-neutral conditions. This allows evaluators to:

- Quantify climate-attributable losses in education outcomes.
- Distinguish climate effects from programme or socio-economic effects.
- Identify non-linear thresholds (e.g. when rainfall anomaly > 0.6 triggers flood-related disruptions).
- Automatically detect high-risk geographies using unsupervised clustering.

This positions AI as a direct contributor to Sustainable Development Goal (SDG) 4 (education) and SDG 13 (climate action) by enabling real-time, climate-aware evaluation frameworks for governments in the Global South.

Limitations and constraints

In the spirit of rigorous academic inquiry, the study acknowledges several limitations inherent in the proposed framework:

- **Data sparsity in the Global South:** While satellite reanalysis data (like ERA5) provides coverage, the lack of ground-truth meteorological stations in rural Africa can lead to localised inaccuracies. The algorithm may struggle to detect microclimates in highly mountainous regions.
- **Attribution error:** While the algorithm can detect *correlations* between weather shocks and project failure, it cannot definitively prove *causation*. There is a risk of 'False Positives', where the system attributes poor results to climate change when the root cause was actually poor project management (e.g. late delivery of fertiliser).
- **Digital divide:** The proposed real-time dashboard requires consistent internet connectivity. In remote evaluation contexts, offline-first capabilities would need to be developed to make the tool practical for field enumerators.

Ethical considerations

This study involved the development of an AI algorithm. As such, it did not involve direct interaction with human subjects or the collection of primary data from individuals. However, the study did consider several ethical considerations related to the potential development and deployment of such an algorithm in the future:

- **Data privacy and security:** The algorithm is designed to process potentially sensitive data, including climate data, socio-economic data and project-level data. The study emphasised the importance of ensuring the privacy and security of this data. The algorithm design incorporates principles of data minimisation, meaning that only the data strictly necessary for the analysis should be collected and processed. It also emphasises the need for robust data security measures to protect against unauthorised access and use.
- **Transparency and explainability:** Artificial intelligence algorithms, particularly complex ML models, can sometimes be 'black boxes', making it difficult to understand how they arrive at their conclusions. The study stressed the importance of transparency and explainability in the design of the algorithm. The algorithm should be designed in a way that allows users to understand the logic behind its outputs and to trace the flow of data through the system. This is crucial for ensuring accountability and building trust in the algorithm's results.
- **Bias and fairness:** Artificial intelligence algorithms can perpetuate and even amplify existing biases in the data they are trained on. The study acknowledged the potential for bias in the data used by the algorithm, such as historical climate data or socio-economic indicators. The algorithm design incorporates strategies for mitigating bias, such as data pre-processing techniques to balance the data, and the use of ML models that are less prone to bias. The study also emphasises the need for ongoing M&E of the algorithm's performance to ensure fairness and equity.
- **Accountability and responsibility:** The use of AI in evaluation raises questions of accountability and responsibility. Who is responsible for the outputs of the algorithm, and how should errors or inaccuracies be

addressed? The study emphasised the need for clear lines of accountability in the deployment of the algorithm. It also stressed the importance of human oversight and the need for evaluators to exercise professional judgement in interpreting and using the algorithm's results. The algorithm is intended to be a tool to support human decision-making, not to replace it.

- **Potential for misuse:** Like any technology, AI can be misused. The study acknowledged the potential for the algorithm to be used in ways that could have negative consequences, such as for political manipulation or to justify harmful policies. The study emphasised the need for ethical guidelines and regulations to govern the use of AI in evaluation, and for ongoing dialogue and debate about the ethical implications of this technology.

While this study focused on the conceptual development of the algorithm, these ethical considerations were central to its design and will need to be thoroughly addressed in any future implementation.

Results

Model performance

The Random Forest demonstrated strong predictive performance:

- Random Forest $R^2 = 0.671$
- Random Forest Root Mean Square Error (RMSE) = 0.042

This substantially outperformed the linear regression baseline:

- Linear Regression $R^2 = 0.492$
- Linear Regression RMSE = 0.053

These results indicate that climate–attendance relationships are non-linear, validating the use of ML for climate-responsive evaluation.

Feature importance

The top predictors of attendance were:

- Resilience index (0.322)
- Drought index (0.238)
- Flood occurrence (0.187)
- Temperature anomaly (0.070)
- Poverty rate (0.067)

The dominance of drought and flood signals confirms the high sensitivity of educational participation to weather extremes, aligning with climate vulnerability patterns in ASAL (Arid and Semi-Arid Lands) counties.

Climate impact counterfactuals

Using the CCIM:

- Mean climate-attributable attendance loss: –0.0133
- Median loss: –0.0093

TABLE 1: Comparison to traditional evaluation approaches.

S/n	Traditional evaluations	AI-driven approach
1.	Use static, periodic data collection.	Continuous, real-time monitoring.
2.	Analyse climate context descriptively rather than analytically.	Non-linear modelling capabilities.
3.	Apply linear models that miss non-linear climate relationships.	Integration of high-resolution climate data.
4.	Struggle to detect rapid climate-induced changes.	Predictive insights and scenario analysis.
5.	-	Automated identification of climate-risk patterns.

AI, artificial intelligence.

- Share of negative impacts: 51.3% of all county-month observations

This shows that climate variability significantly depresses school attendance in roughly half of all months nationally.

County climate-risk clustering

A K-Means clustering model identified three climate-risk county groups:

- Cluster 1: Drought-Sensitive Counties
- Persistently high drought indices and moderate attendance loss.
- Cluster 2: Flood-Prone Counties
- High rainfall anomalies and frequent attendance disruptions.
- Cluster 3: Stable Counties
- Low shock frequency and higher attendance.

In Table 1, traditional evaluations typically rely on static indicators, periodic data collection and linear analyses. These methods often overlook non-linear climate interactions and struggle to account for rapidly changing environmental conditions.

Conclusion

The AI-driven algorithm presents a promising pathway for climate-integrated evaluations in Africa. Through ML techniques, evaluators can detect climate-related drivers, enhance predictive capabilities and support climate-resilient decision-making. As climate risks escalate, evaluation systems must evolve towards data-rich, adaptive and predictive models. Future research should implement pilot applications and refine the algorithm through co-creation with evaluators, climate scientists and policymakers.

Acknowledgements

The author would like to acknowledge Genesis Analytics and the African Evaluation Association (AfrEA) for providing them with the platform and opportunity to contribute to advancing the discourse on integrating climate change into evaluation. Their commitment to fostering innovation and supporting emerging ideas has been invaluable in shaping this work.

This article is based on a conference paper originally presented at the 11th African Evaluation Association (AfrEA) Conference, held in Kigali, Rwanda, on 18 – 22 March 2024. The conference paper, titled 'The Role of Big Data and AI in Evaluation Practice in Africa', was subsequently expanded and revised for this journal publication. This republication is done with permission from the conference organisers.

Competing interests

The author, Chineme A. Anowai, disclosed receipt of financial support from The Rockefeller Foundation through a grant to Genesis Analytics. The author declares that they have no other financial or personal relationships that may have inappropriately influenced them in writing this article.

CRediT authorship contribution

Chineme A. Anowai: Conceptualisation, Methodology, Formal analysis, Investigation, Software, Writing – original draft, Writing – review & editing. The author confirms that this work is entirely their own, has reviewed the article, approved the final version for submission and publication, and takes full responsibility for the integrity of its findings.

Funding information

Support for the AfrEA Conference climate strand, where some of this work was initially presented, was provided by The Rockefeller Foundation through a grant to Genesis Analytics.

Data availability

The data that support the findings of this study are available from the corresponding author, Chineme A. Anowai, upon reasonable request.

Disclaimer

The views and opinions expressed in this article are those of the author and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency or that of the publisher. The author is responsible for this article's results, findings and content.

References

Chilisa, B., 2015, *A Synthesis Paper on the Made in Africa Evaluation Concept Commissioned by African Evaluation Association (AfrEA)*, viewed n.d., from https://africasocialwork.net/wp-content/uploads/2024/01/MAE2-Final-31st-august-1.pdf.

Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS), 2023, *CHIRPS rainfall dataset (Version 2.0)*, Santa Barbara Climate Hazards Center, University of California, viewed n.d., from https://www.chc.ucsb.edu/data/chirps.

Copernicus Climate Change Service, 2023, *ERA5 hourly data on single levels from 1940 to present*, viewed n.d., from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview.

Global Pulse, 2023, *Artificial intelligence for development in Africa: State of the field*, UN Global Pulse, New York, NY.

Hevner, A.R., March, S.T., Park, J. & Ram, S., 2004, 'Design science in information systems research', *MIS Quarterly* 28(1), 75–105. https://doi.org/10.2307/25148625

Kenya Ministry of Education, 2021, *Kenya education statistical booklet*, Government of Kenya, Nairobi, viewed n.d., from http://nemis.education.go.ke.

Kenne, A.D., Toure, M., Logamou Seknewna, L. & Ketsemen, H.L., 2024, 'Subseasonal prediction of summer temperature in West Africa using artificial intelligence: A case study of Senegal', *International Journal of Intelligent Systems* 2024(1), 8869267.

Mouton, J., Wildschut, L. & Agyepong, I., 2019, 'Evaluation in Africa', *African Evaluation Journal* 7(1), 1–10.

Mutanga, M.B., 2022, 'Machine learning for climate resilience in Southern Africa', *Climate and Development* 14(3), 210–225.

Nalubega, T. & Uwizeyimana, D., 2019, 'Public sector monitoring in the fourth industrial revolution', *Public Service Delivery and Performance Review* 7(1), 1–12. https://doi.org/10.4102/apsdpr.v7i1.318

Olawade, D.B., Wada, O.Z., David-Olawade, A.C., Fapohunda, O., Ige, A.O. & Ling, J., 2024, 'AI potential for sustainability', *Next Sustainability* 4(1), 100041. https://doi.org/10.1016/j.nxsust.2024.100041

Patton, M.Q., 2019, *Blue marble evaluation: Premises and principles*, Guilford Press, New York, NY.

Tamasiga, P., Molala, M., Bakwena, M., Nkoutchou, H. & Onyeaka, H., 2023, 'Is Africa left behind in the global climate finance architecture? Redefining climate vulnerability and revamping the climate finance landscape—A comprehensive review', *Sustainability* 15(17), 13036. https://doi.org/10.3390/su151713036

United Nations Children's Fund (UNICEF), 2015, *Education and climate change*, UNICEF, New York.

Venegas Marin, S., Schwarz, L. & Sabarwal, S., 2024, *The impact of climate change on education and what to do about it*, World Bank Group, Washington, DC, viewed n.d., from http://documents.worldbank.org/curated/en/099043024150036726.

World Bank, 2023, *EdStats: Education statistics database*, viewed n.d., from https://datatopics.worldbank.org/education.

York, P. & Bamberger, M., 2020, *Measuring results and impact in the age of big data: The nexus of evaluation, analytics, and digital technology*, Rockefeller Foundation, New York, NY.

Appendix starts on the next page →

Appendix 1

Python implementation logic for climate-integrated evaluation

The following code snippet demonstrates the core logic used in the 'Processing Layer' of the algorithm. It utilises the pandas library for data manipulation and scikit-learn for the Random Forest regression model used in the Malawi simulation.

```
{
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error

class ClimateEvaluationModel:
  def __init__(self):
    # Initialise Random Forest Regressor
    # n_estimators=100 ensures robustness against noise
    self.model = RandomForestRegressor(n_estimators=100, random_state=42)

  def temporal_alignment(self, project_data, climate_data):
    """
    Aligns static project data with dynamic daily climate data.
    Maps specific 'planting_date' to relevant weather windows.
    """
    print("--- Step 1: Performing Phenological Temporal Alignment ---")
    merged_data = pd.merge(project_data, climate_data, on='region_id')

    # Feature Engineering: Calculate 'Shock Intensity' during critical growth phase
    # (e.g., rainfall deviation during flowering stage)
    merged_data['shock_intensity'] = (
    merged_data['rainfall_actual'] - merged_data['rainfall_historical_avg']
    ) / merged_data['rainfall_historical_avg']

    return merged_data
  def train_and_predict(self, df):
    """
    Trains the model to predict 'Expected Yield' based on climate severity.
    """
    print("--- Step 2: Training Random Forest Model ---")

    # Features: Climate inputs + Input usage (fertiliser, seeds)
    X = df[['rainfall_actual', 'wind_speed', 'temp_max', 'input_usage_score']]
    # Target: Crop Yield
    y = df['crop_yield_kg_ha']

    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
    self.model.fit(X_train, y_train)

    # Generate Counterfactual: What would yield be purely based on climate?
    baseline_prediction = self.model.predict(X_test)
    return y_test, baseline_prediction
# --- SIMULATION EXECUTION (Malawi Scenario) ---
# 1. Generate Dummy Data mimicking Cyclone Freddy impact
data = {
  'region_id': range(100),
  'rainfall_actual': np.random.normal(300, 50, 100), # Extreme rainfall (mm)
  'wind_speed': np.random.normal(80, 15, 100), # Cyclone winds (km/h)
  'temp_max': np.random.normal(28, 2, 100),
  'input_usage_score': np.random.uniform(0.5, 1.0, 100),
```

```
  'rainfall_historical_avg': [100] * 100, # Normal average
  'crop_yield_kg_ha': np.random.normal(1200, 200, 100) # Depressed yields
}
df_simulated = pd.DataFrame(data)

# 2. Run the Model
evaluator = ClimateEvaluationModel()
processed_data = evaluator.temporal_alignment(df_simulated, df_simulated[['region_id', 'rainfall_historical_avg']])
actuals, predictions = evaluator.train_and_predict(df_simulated)

# 3. Calculate Resilience Score
# If Actual Yield > Predicted Baseline (given the storm), project is Resilient.
resilience_delta = np.mean(actuals - predictions)

print(f"\n--- RESULTS ---")
print(f"Average Predicted Yield (Baseline under Storm): {np.mean(predictions):.2f} kg/ha")
print(f"Average Actual Yield (Project Beneficiaries): {np.mean(actuals):.2f} kg/ha")
print(f"Resilience Delta: +{resilience_delta:.2f} kg/ha")
if resilience_delta > 0:
  print("CONCLUSION: Project demonstrates resilience against climate-shock.")
else:
  print("CONCLUSION: Project failed to mitigate climate-shock.") }
}}
```