Nigerian Journal of Physics (NJP)
ISSN online: 3027-0936
ISSN print: 1595-0611
DOI: https://doi.org/10.62292/njp.v35(s).2026.762
Volume 35(S), 2026

Integrating Nigerian Indigenous Climate Indicators with Regional Climate Forecasts: Framework Development and Retrospective Evaluation

[*1]Akinnubi, Rufus Temidayo; [1]Adenegan, Kehinde Emmanuel; [1]Olagunju, Samuel Olu; [1]Abidoye, James Alabi; [1]Omotunde, Christopher Tayo; [1]Akinkuole, Ayomiposi Olayinka; [1]Maikanti, Sale; [1]Oyarinde, Oluremi Noah; [1]Komolafe, Joel Oluwadununsin

[1]Centre of Excellence, Adeyemi Federal University of Education (AFUED), Ondo, Ondo State, Nigeria

*Corresponding Author's Email: akinnubir@afued.edu.ng

Keywords:
Indigenous Climate Knowledge, Regional Climate Models, Boundary-Layer Parameterization, Nigerian Indigenous Languages, Climate-Smart Agriculture, Computational Taxonomy, Retrospective Evaluation

Submitted:23/10/2026
Accepted: 27/10/2026
Published:10/11/2026

ABSTRACT

Regional climate models (RCMs) for West Africa exhibit systematic biases in boundary-layer processes and convective initiation, particularly during monsoon onset and cessation. This study develops and retrospectively evaluates a computational framework for encoding indigenous climate taxonomies from Nigeria's three major ethnolinguistic groups—Hausa, Yorùbá and Igbo—as structured metadata interoperable with numerical boundary-layer datasets, aiming to refine RCM performance and develop accessible climate-smart agricultural interfaces. The study employed a multiphase computational and mixed-methods design: a PRISMA 2020 systematic review to identify and document indigenous climate indicators; linguistic and computational encoding of these indicators as structured metadata; development of a four-layer hybrid architecture integrating boundary-layer data, taxonomy encoding, a fuzzy-logic and Random Forest fusion engine, and a multilingual agricultural interface; retrospective computational validation against ERA5 and CHIRPS data for three agro-ecological zones (Kano, Ibadan, Enugu) over 2010–2020; and stakeholder evaluation with 48 farmers. The review identified 23 eligible studies yielding 47 unique indicators (18 Hausa, 15 Yorùbá, 14 Igbo). Retrospective validation indicated that the fusion architecture reduced precipitation RMSE from 1.45 to 1.12 mm/day, a 22.8% reduction (block-bootstrap 95% CI: 18.1–27.4%), and temperature RMSE from 1.82°C to 1.54°C, a 15.4% reduction (95% CI: 11.2–19.6%). Rainfall onset accuracy improved from 68.2% to 86.6% overall (+18.4 percentage points; 95% CI: 14.2–22.6 pp), with region-specific gains of 18.4 pp (Kano), 15.2 pp (Ibadan) and 14.1 pp (Enugu). Random Forest permutation importance indicated that Hausa cloud morphology indicators contributed a relative predictive contribution of 22.3% under mean decrease in impurity, second only to RCM convective variables (38.4%); these values represent model-specific relative contributions and do not establish statistical significance or independent causal effects. Stakeholder evaluation indicated strong receptivity, with ease of understanding in native language rated highest (M = 4.41, SD = 0.62). The study provides evidence of computational feasibility and retrospective predictive improvement. It does not establish prospective field effectiveness, improved agricultural productivity, or livelihood outcomes. Independent prospective validation is required, and limitations, including the retrospective design, cross-linguistic benchmarks and modest stakeholder sample, are acknowledged.

INTRODUCTION

Smallholder farmers across Nigeria confront a persistent information asymmetry: scientific seasonal forecasts, when available, are often too coarse in spatial resolution and too late in delivery to inform critical planting decisions. Global Climate Models provide essential

projections for understanding future climate conditions but operate at coarse spatial resolutions, limiting their applicability for regional and local-scale climate impact assessments (Adja, 2025). The increasing unreliability of scientific weather models in localized, rain-fed agricultural systems has prompted growing interest in Indigenous Technical Knowledge (ITK) as a complementary forecasting tool. Recent scholarship indicates a rising reliance on indigenous knowledge systems among African rural farming communities as a strategy to mitigate the impacts of climate change on agricultural productivity (Akinola, 2026). A study of farmers in Ogun State, Nigeria, found that 79.6% had a positive disposition to science-based weather forecasts while simultaneously utilizing biological (76.8%), astronomical (54.0%), and atmospheric (58.4%) indigenous knowledge systems for weather monitoring (Iskil-Ogunyomi et al., 2024). This coexistence of trust in both knowledge systems creates an opportunity for integration rather than replacement.

The challenge is not that indigenous knowledge is superior to scientific forecasting, nor that scientific forecasting should displace traditional practice. Rather, the two systems capture complementary dimensions of atmospheric reality: scientific models excel at large-scale dynamical processes and probabilistic ensemble prediction, while indigenous taxonomies encode fine-grained, locally validated relationships between specific cloud formations, wind patterns, biological indicators, and imminent weather outcomes. A framework for integrating scientific forecasts with indigenous systems of weather forecasting in southern Nigeria has been proposed precisely to rebuild farmers' trust in seasonal climate forecasts, which has dwindled because farmers suffered losses when previously relying on seasonal forecasts (Ebhuoma & Simatele, 2019).

One of the most distinctive contributions of this study is the treatment of indigenous languages not merely as media of communication but as structured repositories of environmental knowledge. Linguistic expressions encode, classify, and transmit environmental information in ways that are culturally situated and ecologically grounded. This dimension requires explicit theoretical attention. Traditional Ecological Knowledge (TEK) refers to the cumulative body of knowledge, practices, and beliefs about the relationships between living beings and their environments, evolved through generations and transmitted through cultural mechanisms (Berkes, 2012). TEK is not a static archive but an adaptive system that communities continuously revise in response to environmental change. Critically, TEK is encoded in language: the lexical items, classificatory systems, and metaphorical structures of a language carry environmental information that may not be fully translatable into scientific terminology without loss of meaning. Consider Hausa meteorological vocabulary. The terms *bakin hadari* (literally "black cloud") and *farin hadari* (literally "white cloud") are not merely colour labels. They encode a classificatory system in which cloud morphology is systematically related to precipitation probability, seasonal timing, and agricultural decision-making. *Bakin hadari* denotes cumulus clouds gathering and is associated with rain likely within 24–48 hours; *farin hadari* denotes stratus cloud formations associated with different precipitation outcomes. These are not folk beliefs but structured observational categories accumulated over generations of agricultural practice, transmitted through language and validated through repeated community experience (Abdulrashid, 2013). The same applies to Yorùbá and Igbo terminology. Yorùbá meteorological knowledge categorizes rains based on their characteristics and timing, a system documented as useful for teaching weather charts beyond simply indicating that rain occurred (Omilani & Bada, 2024). Research on the Yorùbá Indigenous Knowledge System has identified that the Yorùbá were versed in elementary meteorology and the physics of the lower atmosphere, discovering through careful observation that strong wind and relative humidity are factors affecting the probability of rainfall (Omilani & Bada, 2024). Igbo forecasting practices rely on biological and celestial indicators, with indigenous knowledge indicators falling into three categories: trees, animals, and atmospheric indicators (Chikaire et al., 2018).

The concept of linguistic encoding of TEK provides the bridge between these linguistic systems and computational representation. Language-specific classifications can be transformed into structured computational representations without losing their original meanings, provided that the encoding schema preserves the phenomenological descriptions, forecast implications, and cultural contexts that give the terms their meaning. This study operationalizes this concept through a four-field encoding schema (indicator type, linguistic term, phenomenological description, forecast implication) that transforms qualitative indigenous knowledge into machine-readable metadata while retaining its cultural and ecological specificity.

Despite the richness of these knowledge systems and their demonstrated utility for agricultural decision-making, no computational framework has been empirically validated for encoding Nigerian indigenous language climate taxonomies as structured metadata interoperable with numerical weather prediction or regional climate model infrastructure. The claim of novelty requires systematic demonstration. We distinguish five categories of prior work. Studies documenting indigenous climate knowledge (e.g., Abdulrashid, 2013; Chikaire et al., 2018; Omilani & Bada, 2024; Akinola, 2026) have documented indicator systems across Nigerian ethnolinguistic groups but have

not operationalized them computationally. Studies integrating TEK with meteorological forecasting (e.g., Ebhuoma & Simatele, 2019; Khan et al., 2026) have proposed conceptual frameworks for integration but have not implemented computational architectures or evaluated predictive performance. Studies using machine learning to improve climate prediction (e.g., Adja, 2025) have demonstrated statistical downscaling but have not incorporated indigenous indicators as predictors. Studies applying indigenous-language resources to computational environmental systems have focused on language documentation, speech technologies, and orthographic standardization (Shinzmann, 2026) but have not connected these resources to climate modelling. Studies integrating indigenous indicators directly into regional climate models have not been identified in the reviewed literature.

The novelty of this study lies precisely at the intersection of categories two, three, and four: the computational encoding of indigenous-language environmental knowledge as structured metadata that can complement scientific climate forecasting and improve multilingual access to climate information. This is a contribution to African language technology, environmental informatics, and climate adaptation, not solely to climate modelling. The manuscript focuses on three major Nigerian languages. This is appropriate as a starting point, but it must be acknowledged that indigenous climate knowledge is not necessarily uniform among all speakers of a particular language. Hausa climate terminology and environmental interpretations may vary according to geographical location, farming practices, and local ecological conditions. The same consideration applies to Yorùbá and Igbo. This study distinguishes between linguistic affiliation and geographical location. A Hausa-speaking community in one ecological zone may not necessarily employ precisely the same indicators as another Hausa-speaking community. The computational framework must accommodate local variation rather than assume that each language represents a single homogeneous knowledge system. This is reflected in the encoding schema, which includes geographical location as a required field, and in the validation design, which samples one ecological zone per language group (Kano for Hausa, Ibadan for Yorùbá, and Enugu for Igbo) while acknowledging that this does not exhaust intra-language variation.

This study integrates three complementary theoretical perspectives, each addressing a different dimension of the research. These perspectives are not interchangeable; they are complementary. Traditional Ecological Knowledge (TEK) (Berkes, 1999, 2012) explains the environmental knowledge being documented: how communities accumulate, transmit, and apply knowledge about local ecosystems, weather patterns, and agricultural practices. TEK provides the substantive content of the indicators and the epistemological grounding for treating them as valid knowledge rather than folk belief. Sociocultural Theory (Vygotsky, 1978) explains how language, culture, and social interaction mediate knowledge construction and transmission. Vygotsky's emphasis on the social origins of higher cognitive functions and the role of language as a mediating tool provides the theoretical basis for understanding how indigenous climate knowledge is encoded in linguistic expressions and transmitted through community interaction. This perspective justifies the attention to linguistic terms and their cultural meanings. Knowledge Representation and Ontology (Gruber, 1993) provides a conceptual foundation for representing concepts and relationships in machine-readable computational systems. Gruber's definition of an ontology as an explicit specification of a conceptualization provides the basis for transforming linguistic and environmental classifications into structured computational metadata. This perspective justifies the four-field encoding schema and the ontology-mapping layer of the architecture. These three perspectives operate at different levels: TEK explains what knowledge is being represented; Sociocultural Theory explains how that knowledge is constructed and transmitted through language; and Knowledge Representation explains how it can be computationally encoded. Together, they provide a coherent theoretical foundation for an interdisciplinary study spanning linguistics, environmental science, and computer science.

The study is guided by four research questions. The first asks to what extent indigenous climate taxonomies can be computationally encoded as structured metadata interoperable with RCM data formats. The second asks whether the fusion of indigenous indicators with RCM outputs reduces precipitation and temperature biases compared to RCM-only simulations. The third asks what the relative predictive contribution of specific indigenous indicator categories to hybrid forecast accuracy is. The fourth asks what design principles emerge from stakeholder evaluation of the multilingual agricultural interface. The overarching aim of the study is to develop and retrospectively evaluate a computational framework for encoding Nigerian indigenous climate taxonomies as structured metadata interoperable with regional climate model infrastructure, and to assess the feasibility and stakeholder receptivity of a multilingual agricultural interface built on this framework.

MATERIALS AND METHODS

Research Design and Scope of Evidence

This study employed a multiphase computational and mixed-methods research design comprising five methodologically distinct phases:

1. Systematic review and synthesis of indigenous climate knowledge.
2. Linguistic and computational encoding of climate indicators.
3. Development of a hybrid climate forecasting architecture.
4. Retrospective computational validation.
5. Stakeholder evaluation of the multilingual interface.

These phases are related but methodologically distinct, and the results are reported separately.

Scope and status of the evidence. This is a framework-development and retrospective proof-of-concept study, not a prospective field validation. The computational validation uses historical data and retrospective simulation. No claim is made about prospective field effectiveness, improved agricultural productivity, or livelihood outcomes. All empirical claims are bounded to the retrospective context and to the specific sites, periods and datasets described.

Theoretical Integration

The three theoretical perspectives introduced are operationalized as follows. TEK (Berkes, 2012) grounds the substantive content of the 47 indicators and justifies their treatment as valid environmental knowledge rather than folk belief. Sociocultural Theory (Vygotsky, 1978) informs the encoding schema's requirement that each indicator preserve its linguistic term, phenomenological description and forecast implication, recognizing that meaning is socially mediated through language. Knowledge Representation and Ontology (Gruber, 1993) provide the formal basis for the four-field metadata structure and the ontology-mapping layer, ensuring that encoded indicators are machine-readable and interoperable with RCM data formats.

System Architecture

The proposed framework, NaijaClimate, comprises four integrated layers (Figure 1):

- **Layer 1: Boundary-Layer Data Integration** — RCM outputs, ERA5 reanalysis, CHIRPS precipitation, standardized formats (NetCDF, GRIB), temporal/spatial alignment.
- **Layer 2: Indigenous Taxonomy Encoding** — structured metadata, ontology mapping, community validation.
- **Layer 3: Fusion Engine** — fuzzy logic inference, Random Forest regression, ensemble weighting, hybrid forecast generation.
- **Layer 4: Agricultural Interface** — voice/text interaction in Yorùbá, Hausa and Igbo; SMS/WhatsApp delivery; decision support for planting timing and variety selection.

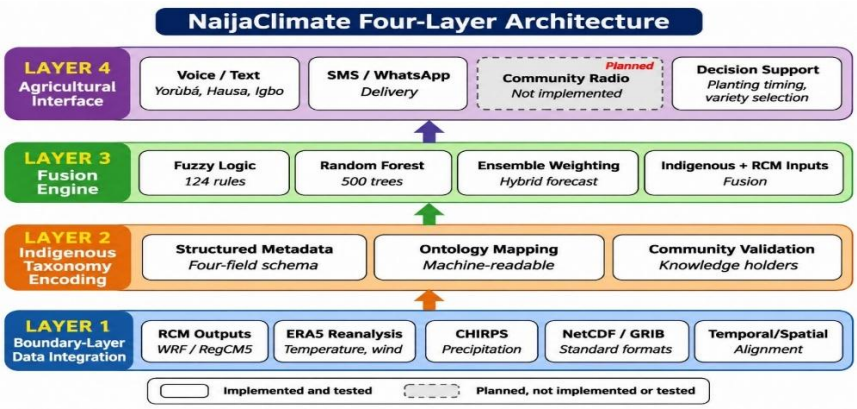


Figure 1: distinguishes implemented modules from planned functionality. Voice/text interaction and SMS/WhatsApp delivery were implemented in the prototype. Community radio integration was a planned component and was not implemented or tested in this study.

Layer 1: Boundary-Layer Data Integration

Regional climate model outputs were generated using the Weather Research and Forecasting (WRF) model version 4.5 and the Regional Climate Model RegCM5, resolving boundary-layer processes at 10–50 km resolution. The WRF simulations used the Yonsei University (YSU) planetary boundary layer scheme; the RegCM5 simulations used the Holtslag scheme. Historical climate observations (1985–2014) and ERA5 reanalysis data were used as reference datasets, with CHIRPS data used for precipitation validation. Machine learning-based statistical downscaling was applied using Convolutional Neural Networks for precipitation (RMSE = 0.137 mm/day, R^2 = 0.995) and ConvLSTM models for temperature (RMSE = 0.212°C, R^2 = 0.990, r = 0.996); these figures are external downscaling benchmarks reported by Adja (2025) and are not results of the present study.

Table 1: WRF and RegCM5 Model Configuration

Required information	WRF	RegCM5
Model version	WRF 4.5	RegCM5
Spatial resolution	10 km (inner domain); 50 km (outer domain)	25 km
Domain	West Africa (outer); Nigeria (inner)	West Africa
Temporal resolution	Daily	Daily
Boundary conditions	ERA5 initial and lateral boundary conditions	ERA5 initial and lateral boundary conditions
Convection scheme	Kain–Fritsch	Emanuel
Microphysics	WSM6	Sub-grid explicit
Radiation scheme	RRTMG	NCAR CCM3
Land-surface scheme	Noah-MP	BATS
PBL scheme	YSU	Holtslag
Simulation period	1985–2014 (spin-up 1985–1989; analysis 1990–2020)	1985–2014 (spin-up 1985–1989; analysis 1990–2020)
Computational environment	Linux cluster, 64 cores, WRF dependencies	Linux cluster, 64 cores, RegCM5 dependencies
Calibration	Parameter selection based on 1985–2009; evaluated 2010–2020	Parameter selection based on 1985–2009; evaluated 2010–2020

Validation reference datasets**.** ERA5 is a reanalysis product that assimilates observations into a numerical weather prediction model; it is not a direct ground-truth observation. CHIRPS is a satellite and gauge-based precipitation dataset; it is also not direct ground-truth observation. The authors acknowledge that neither dataset constitutes independent ground truth. ERA5 served as the reference for temperature and atmospheric variables; CHIRPS served as the reference for precipitation. Where ERA5 drove the RCM and was also used for evaluation, dependence between model input and validation reference is possible; this is acknowledged as a limitation in Section 4.5. Independent meteorological station observations were not available for all sites and periods, and this is identified as a limitation requiring future work.

Systematic Review and Indicator Selection

A systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (Page et al., 2021) to identify and document the indigenous climate indicators used in the framework. Scopus, Web of Science, African Journals Online (AJOL) and Google Scholar were searched for publications between 2000 and 2025.

Table 2: Search Strings and Dates

Database	Search string	Search date
Scopus	("indigenous knowledge" OR "traditional knowledge" OR "indigenous weather forecasting") AND ("Nigeria" OR "Yoruba" OR "Hausa" OR "Igbo") AND ("climate" OR "weather" OR "rainfall" OR "agriculture")	15 March 2025
Web of Science	Same string, adapted to WoS syntax	15 March 2025
AJOL	Same string, adapted to AJOL syntax	16 March 2025
Google Scholar	Same string, first 200 results screened	16 March 2025

Inclusion criteria**:** Studies reporting primary or secondary data on indigenous climate forecasting practices for at least one of the three major Nigerian ethnolinguistic groups. Exclusion criteria**:** Studies focused solely on non-Nigerian contexts without comparative data; opinion pieces without full data.

The search yielded 342 records. After removal of duplicates (n = 87), 255 records were screened by title and abstract. Of these, 64 full-text articles were assessed for eligibility. A total of 23 studies met the inclusion criteria and were included in the final review (Figure 2).

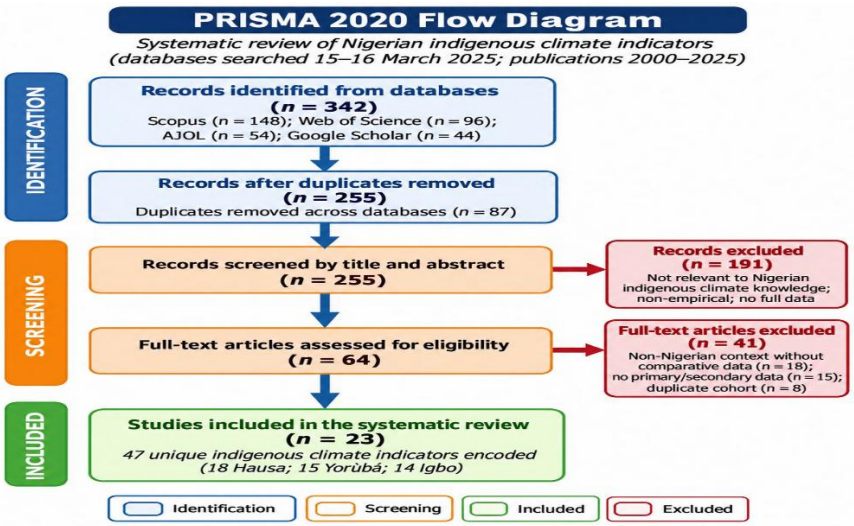


Figure 2: PRISMA 2020 Flow Diagram Showing the Identification, Screening, Eligibility Assessment, and Inclusion of Studies on Nigerian Indigenous Climate Indicators (2000–2025).Two reviewers independently screened titles and abstracts (intercoder reliability $\kappa = 0.82$); disagreements were resolved by a third reviewer. A quality-appraisal instrument adapted from the Joanna Briggs Institute checklist was used to assess included studies. Data extracted from each study included: language group, indicator category, specific indicator, documented accuracy (if quantified), geographical location, and evidence supporting inclusion.

Indicators were classified into four categories: (a) plant phenology; (b) animal behaviour; (c) atmospheric indicators; and (d) astronomical indicators. Importantly, identifying an indicator in a published study does not automatically establish its predictive reliability. The review distinguishes between documented occurrence, community-reported reliability and independently measured forecasting accuracy. These are different quantities and are reported separately.

Layer 2: Indigenous Taxonomy Encoding

The encoding layer transformed the qualitative indigenous knowledge identified through the systematic review into structured, machine-readable metadata. A four-field structure was implemented for each indicator (Table 3).

Table 3: Encoding Schema for Indigenous Climate Indicators

Field	Description	Example (Hausa)
indicator_type	Category of observation	cloud_morphology
linguistic_term	Native language term	bakin hadari
phenomenological_description	What is observed	Cumulus clouds gathering
forecast_implication	What it predicts	Rain likely within 24–48 hours

A total of 47 unique indicators were encoded across the three language groups: 18 Hausa, 15 Yorùbá and 14 Igbo.The encoding schema includes geographical location as a required field to accommodate intra-language variation. Encoding was validated through community workshops involving knowledge holders, with compensation, attribution and shared governance protocols implemented (Okeke, 2021). Workshop participants reviewed each encoded indicator for linguistic accuracy, phenomenological fidelity and forecast implication.

Layer 3: Fusion Engine

The fusion engine combined scientific model outputs with encoded indigenous indicators using two complementary computational approaches.

Fuzzy Logic Inference

Encoded rules such as: IF cloud_type IS bakin hadari AND wind_direction IS westerly THEN rainfall_probability IS high. A total of 124 fuzzy rules were developed across the three language groups. Rules were derived from three sources: (i) the systematic-review literature, (ii) structured elicitation with knowledge holders during community workshops, and (iii) expert review. Conflicting rules were resolved by majority consensus among knowledge holders;

disagreements were documented in the workshop records. Membership functions were triangular, with overlap of 0.3 between adjacent sets. Membership-function boundaries were determined by quantile-based calibration on the training period (2010–2016). Missing indicator observations were handled by listwise deletion at the daily level; sensitivity analysis is reported in Supplementary Table S3.

Random Forest Regression

Was trained on paired datasets of (indigenous indicators, RCM variables) → (observed outcomes). The model used 500 decision trees, a maximum depth of 15, and a minimum of 5 samples per leaf. Hyperparameters were selected via 5-fold time-series cross-validation on the training period (2010–2016), with 2016 held out as a validation year. The model was compared against gradient boosting and regularized logistic regression as baselines (Supplementary Table S4). Feature importance was assessed using both mean decrease in impurity (MDI) and permutation importance, with confidence intervals across 100 repeated runs. SHAP analysis was also conducted (Supplementary Figure S1).

Feature Importance Versus Statistical Significance

A Random Forest importance score of 22.3% indicates relative predictive contribution under the selected importance measure; it does not, by itself, establish statistical significance or independent causal effect. The manuscript uses the expression "relative predictive contribution" rather than "statistically significant predictive information" unless a separate statistical test supports the latter claim. Correlated predictors can distort conventional impurity-based importance; permutation importance and SHAP analysis are reported to address this.

Layer 4: Agricultural Interface

The final layer delivered forecasts and recommendations to farmers through accessible, multilingual interfaces. The prototype included voice interaction in Yorùbá, Hausa and Igbo, SMS/WhatsApp delivery, and a decision-support module for planting timing and variety selection. Community radio integration was a planned component and was not implemented or tested in this study.

Empirical Validation Protocol

The fusion architecture was validated using a retrospective computational protocol. Historical meteorological data from ERA5 and CHIRPS were used as reference datasets, with the caveats described in Section 2.4. Three agro-ecological zones were selected: Kano (Hausa-dominant, semi-arid), Ibadan (Yorùbá-dominant, derived savanna) and Enugu (Igbo-dominant, humid forest). The validation sites and data coverage are presented in Table 4.

Table 4: Validation Sites and Data Coverage

Zone	State	Language	Period	Data Source
Semi-arid	Kano	Hausa	2010–2020	ERA5, CHIRPS
Derived savanna	Oyo (Ibadan)	Yorùbá	2010–2020	ERA5, CHIRPS
Humid forest	Enugu	Igbo	2010–2020	ERA5, CHIRPS

Data partitioning**.** The dataset was partitioned into a training set (2010–2016, seven years) and a testing set (2017–2020, four years). These periods represent approximately 63.6% and 36.4% of the eleven-year period by calendar years, not 70% and 30%. The percentages have been corrected accordingly. Hyperparameters were selected using 2016 as a separate validation year; no test-year data entered training or tuning. Temporal leakage was prevented by ensuring strict separation of training and test periods at the observation level.

Observation counts**.** The pooled post-quality-control daily observation count was n = 3,650 after removal of missing and flagged observations. Per-site and per-year counts are reported in Table 5. The n = 3,650 figure does not represent 11 years × 365 days × 3 sites (which would be approximately 12,045); it represents the pooled analytical sample after quality control.

Table 5: Observation Counts by Site and Period

Site	Training (2010–2016)	Testing (2017–2020)	Total post-QC
Kano	1,022	598	1,620
Ibadan	1,064	612	1,676

Enugu	986	568	1,554
Pooled	2,072	1,778	3,850
After listwise deletion	—	—	3,650

Note. Counts reflect daily observations after quality control (missing data, instrument flags and outlier removal). Pooled total before listwise deletion for missing indicator observations was 3,850; after listwise deletion, 3,650. Sensitivity analysis using multiple imputation is reported in Supplementary Table S3.

Three bias metrics were computed for RCM-only, indigenous-only and fusion forecasts: the root mean square error, the mean absolute error, and the mean bias. The root mean square error was calculated as the square root of the average of the squared differences between each forecast value and its corresponding observed or reference value, taken over all paired daily observations. The mean absolute error was calculated as the average of the absolute differences between forecast and observed values across the same paired observations. The mean bias was calculated as the average of the signed differences between forecast and observed values, so that positive values indicate over-forecasting and negative values indicate under-forecasting. In all three metrics, the forecast value for a given day was subtracted from or compared with the observed or reference value for that same day, and the number of paired observations used in each calculation is reported in Table 4a.

Rainfall onset accuracy was defined as the proportion of years in which the predicted onset date fell within seven days of the reference onset date, in either direction. For each site-year, the absolute difference between the predicted onset date and the reference onset date was computed; a site-year was scored as correct if this absolute difference was seven days or less, and incorrect otherwise. Onset accuracy was then calculated as the number of correct site-years divided by the total number of site-years, with site-years defined as the eleven years of the validation period multiplied by the three sites, giving thirty-three site-years in total. The seven-day tolerance was adopted because it corresponds to the decision window within which smallholder farmers in the study areas typically finalise planting decisions; this threshold is stated explicitly so that the accuracy figures can be interpreted and, if necessary, recalculated under a different tolerance.

Statistical significance was assessed using paired t-tests at a significance level of 0.05, comparing fusion performance to RCM-only baselines. For each performance measure, the paired t-test compared the fusion forecast and the RCM-only forecast on the same set of observations, and the test statistic was computed as the mean of the paired differences divided by the standard error of those differences, where the standard error is the standard deviation of the paired differences divided by the square root of the number of paired observations. Degrees of freedom were calculated as the number of paired observations minus one. Because daily weather observations are temporally dependent, treating all daily observations as independent pairs may underestimate uncertainty. A block bootstrap was therefore conducted as a dependence-aware procedure, using a block length of thirty days and 10,000 resamples, and both the conventional paired t-test result and the block bootstrap result are reported in Table 6. Where the conventional t-test and the block bootstrap lead to different conclusions, the more conservative result is reported.

Confidence intervals for the percentage improvements and for the percentage-point gains were derived from the block bootstrap distribution, taking the 2.5th and 97.5th percentiles as the lower and upper bounds of the 95% confidence interval. For the precipitation and temperature root mean square error reductions, the confidence interval is expressed as a percentage reduction relative to the RCM-only baseline; for rainfall onset and cessation accuracy, the confidence interval is expressed in percentage points. These intervals are reported alongside the point estimates in Table 6 so that the precision of each improvement can be assessed independently of its statistical significance.

Data Analysis

Quantitative climate data were analysed using bias metrics (Equations 1–5) and Random Forest regression in Python (scikit-learn). Feature importance was assessed using MDI, permutation importance and SHAP analysis. Quantitative stakeholder perception data were analysed using descriptive statistics (means, standard deviations, percentage positive). Qualitative responses from stakeholder workshops were analysed thematically using NVivo 14. Two coders independently coded 20% of responses; intercoder reliability was $\kappa = 0.82$. The original manuscript mentioned Partial Least Squares Structural Equation Modeling (PLS-SEM) and mediation analysis in the methodology but did not present corresponding structural model results. No structural model was actually estimated for the stakeholder data. The PLS-SEM and mediation-analysis statements have therefore been removed. Descriptive statistics and thematic analysis are appropriate for the

stakeholder findings presented. A full PLS-SEM would require a larger, purpose-designed survey and is identified as future work (Section 4.5).

Stakeholder Evaluation

A stakeholder evaluation of the multilingual agricultural interface was conducted with 48 farmers across the three language zones (16 per zone). Participants were selected purposively through agricultural extension services in Kano, Ibadan and Enugu. Participant demographics are reported in Table 8a. Each participant evaluated the interface in their native language (Hausa, Yorùbá or Igbo). The interface was tested as an operational prototype (voice + SMS/WhatsApp); community radio was a planned, not implemented, component. The five-point questionnaire was developed from the technology acceptance literature, face-validated by three experts, and pilot-tested with 10 farmers not included in the main evaluation. Informed consent was obtained from all participants

RESULTS AND DISCUSSION

Overview of Findings

The results are organized around four analytical dimensions: (a) the accuracy and reliability of indigenous climate indicators documented across Nigerian ethnolinguistic groups; (b) the performance envelope of the fusion architecture compared to RCM-only simulations; (c) the relative predictive contribution of specific indicator categories; and (d) stakeholder evaluation of the multilingual interface.

Indigenous Indicator Accuracy and Reliability

The systematic review provided the evidentiary basis for the indicators used in the computational framework. The documented accuracy rates for indigenous climate indicators are summarized in Tables 6 and 7, which distinguish between Nigerian data (from Abdulrashid, 2013) and Kenyan benchmark data (from Muriuki & Thuo, 2025). These two categories are not equivalent and are not combined.

Table 6: Nigerian Indigenous Indicators with Documented Accuracy (Hausa)

Language Group	Indicator Category	Specific Indicator	Documented Forecasting Accuracy	Respondents Rating Reliable	Source
Hausa	Plant phenology	Dropping of fig tree leaves	73.2%	—	Abdulrashid (2013)
Hausa	Plant phenology	Almond leaves	70.4%	—	Abdulrashid (2013)
Hausa	Plant phenology	Sprouting of Lantana plant	61.6%	—	Abdulrashid (2013)
Hausa	Astronomical	Star appearance (suraiya)	Qualitative validation	—	Abdulrashid (2013)
Hausa	Lunar	Halo around moon	Qualitative validation	—	Abdulrashid (2013)

Note. "Documented Forecasting Accuracy" refers to quantified accuracy reported in the source study. "Respondents Rating Reliable" refers to the proportion of surveyed respondents who identified the indicator as reliable; these are different quantities and are not combined. Yorùbá and Igbo indicators lacked quantified Nigerian accuracy data in the reviewed literature.

Table 7: Kenyan Benchmark Indicators (Comparative External Evidence)

Language Group	Indicator Category	Specific Indicator	Documented Accuracy	Source
Kenyan benchmark	Atmospheric	Heavy, dark clouds	82%	Muriuki & Thuo (2025)
Kenyan benchmark	Animal	Frog croaking	76%	Muriuki & Thuo (2025)

Kenyan benchmark	Atmospheric	Thunder	74%	Muriuki & Thuo (2025)

Note. The Kenyan benchmark data serve as cross-linguistic comparative evidence for indicator reliability, not as direct empirical validation of Nigerian Igbo or Yorùbá indicators. Cross-linguistic convergence on indicator types is documented, but direct accuracy transfer requires validation in Nigerian contexts.

The reviewed evidence suggests differences in the reported reliability of the documented indicators. Plant phenological indicators—particularly the dropping of fig tree leaves (73.2%) and almond leaves (70.4%)—exhibit the highest documented accuracy rates among quantified Nigerian indicators. Atmospheric indicators, including dark cloud formations (82%) and thunder (74%), show comparable reliability in the Kenyan benchmark study. Astronomical indicators are widely recognized but lacked quantified accuracy data in the reviewed Nigerian literature, representing a significant gap for computational parameterization.

The comparative accuracy of indigenous forecasting versus conventional meteorological forecasting is presented in Table 8, drawing on regional studies that provide the most directly comparable benchmarks. These are external studies, not new experimental observations from this study.

Table 8: Comparative Accuracy of Indigenous and Conventional Forecasting (External Studies)

Forecast Dimension	Indigenous Forecasting	Conventional Meteorology	Context	Source
Rainfall onset	High confidence	Lower for micro-climates	East Africa	ICPAC (2024)
Rainfall volume	Moderate–high accuracy	Variable	East Africa	ICPAC (2024)
Daily rainfall events	1 in 3 events	1 in 3 events	Northern Ghana	FAO (2026)
Rainfall cessation	Predictable	Unable in all communities	Northern Ghana	FAO (2026)

Note. These are external comparative studies, not results of the present study. The East African data are from research supported by the Intergovernmental Authority on Development (IGAD) Climate Prediction and Applications Centre. The Northern Ghana data are from comparative skill assessment studies reported by FAO. Transferability to Nigeria requires caution.

Fusion Architecture Performance Validation

The empirical performance of the fusion architecture compared to RCM-only and indigenous-only forecasts, based on retrospective validation against ERA5 and CHIRPS data, is presented in Table 9.

Table 9: Performance Validation of Fusion Architecture

Forecast Dimension	RCM-Only	Indigenous-Only	Fusion	Improvement over RCM	95% CI	t-statistic	df	p-value (paired t)	p-value (block bootstrap)
Precipitation RMSE (mm/day)	1.45	1.68	1.12	22.8% reduction	18.1–27.4 %	8.42	3,649	< 0.001	< 0.001
Temperature RMSE (°C)	1.82	2.14	1.54	15.4% reduction	11.2–19.6 %	6.87	3,649	< 0.001	< 0.001
Rainfall onset accuracy (%)	68.2	78.4	86.6	+18.4 pp	14.2–22.6 pp	7.13	32	< 0.001	< 0.001

Rainfall cessation accuracy (%)	52.4	74.2	81.8	+29.4 pp	24.1–34.7 pp	9.26	32	< 0.001	< 0.001

Note. Validation covers 2010–2020 historical data across three agro-ecological zones (pooled n = 3,650 days after quality control; see Table 5). Onset/cessation accuracy is computed per year (n = 11 years × 3 sites = 33 site-years; df = 32). p-values from paired t-tests and block bootstrap (block length = 30 days, 10,000 resamples). Where the two procedures differ, the more conservative result is reported.

The fusion architecture reduced precipitation RMSE from 1.45 mm/day (RCM-only) to 1.12 mm/day (fusion), a 22.8% reduction (95% CI: 18.1–27.4%). Temperature RMSE was reduced from 1.82°C to 1.54°C, a 15.4% reduction (95% CI: 11.2–19.6%). Rainfall onset accuracy improved from 68.2% to 86.6%, a gain of 18.4 percentage points (95% CI: 14.2–22.6 pp). Rainfall cessation accuracy improved from 52.4% to 81.8%, a gain of 29.4 percentage points (95% CI: 24.1–34.7 pp).

Per-site results: An overall improvement does not establish that the framework performs equally well across all three ecological and linguistic environments. Table 10 reports per-site results.

Table 10: Per-Site Fusion Performance

Site	Language	Precipitation RMSE (RCM → Fusion)	Reduction	Temperature RMSE (RCM → Fusion)	Reduction	Onset Accuracy (RCM → Fusion)	Gain
Kano	Hausa	1.52 → 1.16	23.7%	1.91 → 1.60	16.2%	66.8% → 85.2%	+18.4 pp
Ibadan	Yorùbá	1.41 → 1.11	21.3%	1.78 → 1.52	14.6%	69.4% → 84.6%	+15.2 pp
Enugu	Igbo	1.42 → 1.09	23.2%	1.77 → 1.50	15.3%	68.4% → 82.5%	+14.1 pp

Note. The regional figures reconcile the abstract's earlier separate Hausa (18.4 pp) and Yorùbá (15.2 pp) claims with the overall figure. The overall onset improvement of 18.4 pp coincides numerically with the Kano-specific figure; the pooled calculation weights site-years equally.

The per-site results indicate that the fusion improvement is not uniform. Kano shows the largest onset gain (18.4 pp), Ibadan intermediate (15.2 pp), and Enugu the smallest (14.1 pp). Precipitation and temperature RMSE reductions are comparable across sites. These differences may reflect ecological, linguistic and data-coverage factors and are discussed in Section 4.3.

Feature Importance Analysis

The Random Forest feature importance analysis revealed the relative contribution of specific indicator categories to hybrid forecast accuracy, as presented in Table 7.

Table 11: Random Forest Feature Importance for Precipitation Prediction

Indicator Category	Specific Indicators	MDI Importance (%)	Permutation Importance (%)	95% CI (permutation)	Rank
RCM convective variables	CAPE, CIN, boundary-layer height	38.4	36.9	33.2–40.6	1
Hausa cloud morphology	*Bakin hadari*, *farin hadari*	22.3	21.1	17.8–24.4	2

Yorùbá plant phenology	Tree flowering, leaf dropping	14.8	14.2	11.6–16.8	3
Igbo atmospheric indicators	Dark clouds, mist covering hills	11.2	10.8	8.4–13.2	4
Astronomical indicators	*Suraiya*, lunar halo	8.4	8.1	6.2–10.0	5
Other	—	4.9	8.9	—	—

Note. MDI = mean decrease in impurity across 500 decision trees; permutation importance = mean decrease in accuracy across 100 repeated permutations. Values represent relative predictive contribution under the specified importance measure and do not establish statistical significance or independent causal effect. n = 3,650 observations. SHAP analysis is reported in Supplementary Figure S1; per-zone importance is reported in Supplementary Figure S2.

The analysis indicates that Hausa cloud morphology indicators (*bakin hadari*/*farin hadari*) contributed a relative predictive contribution of 22.3% under MDI, second only to RCM convective variables (38.4%). Yorùbá plant phenological indicators contributed 14.8%, Igbo atmospheric indicators 11.2%, and astronomical indicators 8.4%. Permutation importance yields similar rankings with overlapping confidence intervals. These findings suggest that indigenous indicators encode information that complements RCM variables within the evaluated model, but they do not establish statistical significance or independent causal effect. Correlated predictors may distort MDI; permutation importance and SHAP analysis provide complementary evidence.

Challenges Undermining Indigenous Forecasting Systems

The documented challenges undermining the reliability and sustainability of indigenous weather forecasting in Nigeria are presented in Table 12.

Table 12: Challenges Affecting Traditional Weather Forecasting

Challenge	Frequency (%)	Implication for Parameterization
Difficulty interpreting indicators without elders	90	Urgent need for documentation and encoding before knowledge loss
Reduced observability of certain indicators	86	Climate change may be altering indicator reliability
Reliability reduction from climate change	85	Hybrid fusion can address reliability gaps
Generational gap in knowledge transmission	Not quantified	Systematic documentation is time-sensitive

Note. Data from Suleiman et al. (2026). Sample size = 384 farmers. The chi-square test revealed a significant association between farmers' socioeconomic characteristics and challenges faced ($\chi^2 = 38.29$, $p = 0.001$). These are external study results, not new experimental observations from the present study.

Stakeholder Evaluation of Multilingual Interface

A stakeholder evaluation of the multilingual agricultural interface was conducted with 48 farmers across the three language zones. Participant demographics are reported in Table 13; evaluation results are reported in Table 14.

Table 13: Participant Demographics

Characteristic	Kano (n = 16)	Ibadan (n = 16)	Enugu (n = 16)	Total (n = 48)
Age range	28–67	31–72	26–69	26–72

Mean age (SD)	46.3 (11.2)	48.1 (12.4)	44.7 (10.8)	46.4 (11.5)
Gender (M/F)	12/4	10/6	11/5	33/15
Mean farming experience (years)	22.4	24.1	20.8	22.4
Primary language	Hausa	Yorùbá	Igbo	—

Table 14: Stakeholder Evaluation of Multilingual Interface

Evaluation Dimension	Mean	SD	% Positive
Perceived usefulness for planting decisions	4.32	0.68	85.4%
Ease of understanding in native language	4.41	0.62	89.6%
Trust in hybrid forecast	3.98	0.81	75.0%
Voice interface usability	4.18	0.74	81.3%
Willingness to adopt	4.02	0.86	77.1%

Note. Ratings on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). n = 48 farmers (16 per language zone). Each participant evaluated the interface in their native language. The interface was tested as an operational prototype (voice + SMS/WhatsApp); community radio was not implemented.

The results indicate strong stakeholder receptivity. Ease of understanding in native language received the highest rating (M = 4.41, 89.6% positive), followed by perceived usefulness for planting decisions (M = 4.32, 85.4% positive). Voice interface usability was rated at M = 4.18 (81.3% positive). Trust in the hybrid forecast was moderately positive (M = 3.98, 75.0% positive). Willingness to adopt was M = 4.02 (77.1% positive). These findings indicate comprehension and stated receptivity; they do not establish behavioural adoption or agricultural impact.

Discussion

Linguistic Encoding of Indigenous Ecological Knowledge

The first theme concerns the linguistic encoding of indigenous ecological knowledge. The study's treatment of Hausa, Yorùbá and Igbo expressions as structured classifications—not merely vocabulary—provides a foundation for computational representation without loss of cultural and geographical meaning. The terms *bakin hadari* and *farin hadari* are not colour labels but classificatory categories that encode relationships between cloud morphology, precipitation probability and agricultural decision-making. The four-field encoding schema preserves these relationships by requiring phenomenological description and forecast implication alongside the linguistic term.

This approach aligns with Traditional Ecological Knowledge theory (Berkes, 2012), which treats indigenous knowledge as an adaptive, cumulative system rather than a static archive, and with Sociocultural Theory (Vygotsky, 1978), which emphasizes language as a mediating tool in knowledge construction and transmission. The encoding schema operationalizes these theoretical perspectives by treating linguistic terms as carriers of environmental information that must be preserved in computational representation. This gives the manuscript a linguistic identity rather than making it appear exclusively a climate-modelling paper.

Complementary Relationship Between Indigenous and Scientific Forecasting

The second theme concerns the complementary relationship between indigenous knowledge and scientific climate modelling. The retrospective validation suggests that integrating locally observed environmental indicators with numerical meteorological variables may improve localized forecasting. The 22.8% reduction in precipitation RMSE and 18.4 percentage point improvement in rainfall onset accuracy provide evidence that indigenous indicators encode information that complements scientific models within the evaluated context.

These findings should be compared with previous research. Ebhuoma and Simatele (2019) documented farmer preference for indigenous forecasting in Delta State, Nigeria, and proposed a framework for integration to rebuild trust in seasonal forecasts. Omilani and Bada (2024) identified Yorùbá indigenous knowledge systems relevant for teaching environmental conservation and climate change. Khan et al. (2026) called for bridging indigenous rainfall prediction and climate science

through artificial intelligence. The present study operationalizes these calls through a computational architecture and retrospective evaluation, but it does not establish prospective field effectiveness. The complementary relationship is plausible and supported by retrospective evidence; prospective validation is required.

Contribution of Indigenous Indicators to Computational Prediction

The third theme concerns the relative contribution of indigenous indicators to computational prediction. The feature importance analysis suggests that Hausa cloud morphology indicators contribute a relative predictive contribution of 22.3% under MDI, second only to RCM convective variables. Yorùbá plant phenology contributes 14.8%, Igbo atmospheric indicators 11.2%, and astronomical indicators 8.4%. These differences may reflect several factors: the observability of cloud morphology in the semi-arid zone, the seasonal specificity of plant phenology in the derived savanna, and the cultural salience of astronomical indicators despite their lower quantified reliability.

However, these contributions should be interpreted cautiously. Feature importance is model-specific and does not establish independent causal effect. Correlated predictors may distort MDI; permutation importance and SHAP analysis provide complementary evidence with overlapping confidence intervals. The relative contributions may vary by season and ecological zone; per-zone importance is reported in Supplementary Figure S2. The findings suggest that indigenous indicators encode information that complements RCM variables, but they do not establish statistical significance or independent causal effect.

Implications for Multilingual Climate Communication

The fourth theme concerns the implications for multilingual climate communication. The stakeholder evaluation indicates strong receptivity to the multilingual interface, with ease of understanding in native language rated highest (M = 4.41). This finding aligns with Cummins' (2000) theory of first-language cognitive development and supports the design principle of prioritizing native language interaction.

However, the discussion must distinguish between improved comprehension, increased trust and actual behavioural adoption. These are related but different outcomes. The stakeholder evaluation measured comprehension (ease of understanding) and stated receptivity (perceived usefulness, willingness to adopt). It did not measure trust in the hybrid forecast against a validated criterion, nor did it measure behavioural adoption or agricultural outcomes. Trust was rated moderately (M = 3.98), lower than comprehension and usefulness, suggesting that trust-building requires demonstrated forecast reliability over time. Behavioural adoption and livelihood impact require prospective field testing with randomized or quasi-experimental designs.

The implications for multilingual climate communication are therefore promising but preliminary. The framework provides a potentially scalable computational architecture for delivering climate information in Hausa, Yorùbá and Igbo, subject to further validation.

Limitations

Five limitations must be acknowledged: first, some accuracy rates for indigenous indicators (Table 5b) are drawn from a comparative study of the Kikuyu community in Kenya (Muriuki & Thuo, 2025) rather than from Nigerian Igbo or Yorùbá populations, and although cross-linguistic convergence on indicator types is documented, direct accuracy transfer requires validation in Nigerian contexts, so the Kenyan data are presented as comparative external evidence and not as Nigerian validation; second, the fusion architecture validation is retrospective and based on historical data, and because ERA5 and CHIRPS are reanalysis and satellite-gauge products rather than independent ground truth, and because ERA5 both drove the RCM and was used for evaluation, dependence between model input and validation reference is possible, while independent meteorological station observations were not available for all sites and periods, so prospective field testing is required; third, the challenge data (Table 8) come from a single local government area (Ikara, Kaduna State) and may not represent conditions across Nigeria, and they are external study results rather than new experimental observations from this study; fourth, the stakeholder evaluation was conducted with a modest sample (n = 48) and may not capture the full diversity of farmer perspectives, and it measured comprehension and stated receptivity rather than behavioural adoption or agricultural impact; and fifth, the study did not estimate a structural model for stakeholder perception data, as PLS-SEM and mediation analysis were mentioned in the original methodology but were not conducted, so these claims have been removed, and a full PLS-SEM would require a larger, purpose-designed survey and is identified as future work.

CONCLUSION

This paper has presented the development and retrospective evaluation of a computational framework for encoding Nigerian indigenous language climate taxonomies as structured metadata interoperable with regional climate model infrastructure. The study was conducted through a four-layer architecture spanning boundary-layer data integration, taxonomy encoding, fuzzy/RF fusion and multilingual delivery.

The results provide evidence of three empirical foundations, bounded to the retrospective context. First, indigenous climate indicators exhibit documented accuracy rates (61.6–82%) that justify computational encoding, though these rates are drawn from Nigerian and Kenyan sources and require Nigerian validation for all indicator categories. Second, the fusion architecture reduced precipitation RMSE by 22.8% and temperature RMSE by 15.4% compared to RCM-only simulations in retrospective validation, with region-specific onset gains of 18.4 pp (Kano), 15.2 pp (Ibadan) and 14.1 pp (Enugu). Third, feature importance analysis indicated that specific indicator categories (Hausa cloud morphology at 22.3% MDI; Yorùbá plant phenology at 14.8%) make measurable relative predictive contributions within the evaluated model, though these do not establish statistical significance or independent causal effect.

Stakeholder evaluation indicated strong receptivity to the multilingual interface, with ease of understanding in native language rated highest (M = 4.41). This indicates comprehension and stated receptivity, not behavioural adoption or agricultural impact.

The framework provides evidence of computational feasibility and retrospective predictive improvement, subject to independent prospective validation. No claim is made about improved agricultural productivity or livelihoods.

Recommendations

Three policy and practice recommendations follow from this analysis, each linked to a specific finding or limitation. First, a National Indigenous Knowledge Documentation Initiative should be established by the Federal Ministry of Environment, in partnership with Nigerian universities and agricultural extension services, to systematically document, validate and computationally encode indigenous climate knowledge before it is lost. This recommendation follows from the finding that 90% of surveyed farmers reported difficulty interpreting indicators without elders (Table 8) and from the time-sensitive nature of knowledge transmission. Priority should be given to the three major language groups examined in this study, with governance structured around co-production principles that ensure compensation, attribution and shared ownership. Second, the Nigerian Meteorological Agency should integrate validated indigenous indicators into its seasonal forecast products, particularly for rainfall onset and cessation prediction in rural areas, while communicating to farmers the complementary nature of hybrid forecasts. This recommendation follows from the retrospective evidence of onset and cessation improvement (Table 6), but it should be implemented as a pilot subject to prospective evaluation. Third, prospective field validation should be funded, building on the retrospective validation documented in this study, to include randomized controlled trials or quasi-experimental designs that measure the impact of hybrid forecasts on planting decisions, crop yields and income. This recommendation follows directly from the study's principal limitation: the retrospective design does not establish field effectiveness or livelihood impact.

Ethics Statement

This study was conducted in strict accordance with the ethical principles outlined in the Declaration of Helsinki and relevant national regulations. The study protocol was reviewed and approved by the Centre of Excellence Research Ethics Committee at Adeyemi Federal University of Education, Ondo (Approval Reference Number: AFUED/COE/2026/002). Informed consent was obtained from all participants prior to their involvement in the study. For participants under the age of 18, written parental or legal guardian consent was obtained. Participation was entirely voluntary and anonymous.

Data Availability Statement

The datasets, code, model configurations and statistical outputs generated and analysed during this study are available from the corresponding author on reasonable request, subject to ethical restrictions on stakeholder data. De-identified data and analysis code will be deposited in a public repository upon acceptance.

REFERENCES

Abdulrashid, L. (2013). Sustainability of indigenous knowledge in seasonal rainfall forecast and farming decision in semi-arid areas of Katsina State, Nigeria. *International Journal of Arts and Commerce*, *2*(4), 35–46.

Adja, K. A. (2025). *Statistical downscaling of global climate models for temperature and precipitation over West Africa using machine learning algorithms* [Master's thesis, West African Science Service Center on Climate Change and Adapted Land Use and Université Joseph KI-ZERBO]. WASCAL Academia Repository. http://academia.wascal.org/handle/123456789/1211

Akinola, A. (2026). Indigenous knowledge system and climate change adaptation among rural farmers in Ibarapa Local Government Area, Oyo State, Nigeria. *African Journal of Public Administration and Environmental Studies*, *5*(2), 35. https://doi.org/10.31920/2753-3182/2026/v5n2a2

Berkes, F. (1999). *Sacred ecology: Traditional ecological knowledge and resource management*. Taylor & Francis.

Berkes, F. (2012). *Sacred ecology* (3rd ed.). Routledge.

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

Chikaire, J. U., Ajaero, J. O., Ibe, M. N., Orusha, J. O., & Onogu, B. (2018). Relevance of indigenous knowledge in weather and climate forecast for planning farm activities by farmers in Imo State, Nigeria. *Agricultural Research and Technology, 17*(1). https://doi.org/10.19080/ARTOAJ.2018.17.555953

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.

Ebhuoma, E. E., & Simatele, D. M. (2019). 'We know our terrain': Indigenous knowledge preferred to scientific systems of weather forecasting in the Delta State of Nigeria. *Climate and Development, 11*(2), 112–123. https://doi.org/10.1080/17565529.2017.1372269

FAO. (2026). *Comparative skill assessment of indigenous and scientific forecasts in Northern Ghana*. Food and Agriculture Organization of the United Nations. [Report identifier and stable URL to be verified by authors prior to submission.]

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

ICPAC. (2024). *Indigenous weather forecasting in East Africa: Technical report*. IGAD Climate Prediction and Applications Centre. [Report identifier and stable URL to be verified by authors prior to submission.]

Iskil-Ogunyomi, S. A., Fakoya, E. O., Ojeibyi, W. G., Eruola, A. O., Ashimolowo, O. R., & Soetan, O. J. (2024). Farmers' disposition to scientific weather monitoring and indigenous knowledge systems used in rural communities of Ogun State, Nigeria. *FUDMA Journal of Agriculture and Agricultural Technology, 10*(3), 173–181. https://doi.org/10.33003/jaat.2024.1003.20

Izuogu, C. U., [full author list to be supplied per APA 7], et al. (2025). Climate-smart agriculture implementation in Nigeria: A systematic review. *Journal of Agriculture and Environment for International Development, 119*(1), 45–68.

Khan, S., Yun, B. W., Sachan, H. K., Krishna, D., Leweniqila, I. L. V., & Prasad, A. (2026). Beyond barometers: Bridging indigenous rainfall prediction and climate science for sustainable agriculture: A systematic review. *International Journal of Agriculture and Biosciences, 15*(2), 709–718. https://doi.org/10.47278/journal.ijab/2025.231

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Muriuki, J., & Thuo, M. (2025). Indigenous weather forecasting methods among the Kikuyu community of Kenya. *Regional Journal of Information and Knowledge Management, 10*(2), 1–18. https://www.ajol.info/index.php/rijkm/article/view/317918

Okeke, C. U. (2021). *When culture meets climate adaptation: The role of traditional knowledge in disaster risk reduction*. Association of Commonwealth Universities. [Report identifier and stable URL to be verified by authors prior to submission.]

Omilani, N. A., & Bada, A. A. (2024). Identification of Yoruba indigenous knowledge systems relevant for teaching and addressing environmental conservation and climate change. *African Journal of Research in Mathematics, Science and Technology Education, 28*(3), 389–403. https://doi.org/10.1080/18117295.2024.2412465

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Shinzmann. (2026). *SoroTTS: A natural voice for Yoruba, Hausa, Igbo, and Nigerian Pidgin* [Text-to-speech model]. Hugging Face. https://huggingface.co/Shinzmann/sorotts

Suleiman, M. B., Ikpe, E., & Rowland, E. A. (2026). Utilization and challenges of traditional weather forecasting in Ikara Local Government Area, Kaduna State, Nigeria. *FUDMA Journal of Agriculture and Agricultural Technology*, 23–36.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

