InfoNote

# Integration of Artificial Intelligence in Seasonal to Sub-Seasonal Forecasting Systems in West Africa and the Sahel

*Mandela C. Houngnibo, Abdou Ali, Boubacar Toukal Assoumana, Bernard Minoungou, Alcade C. Segnon, Robert B. Zougmoré*

December 2024

**Key messages**

- This Info Note describes AGRHYMET's efforts to integrate Artificial Intelligence (AI) into seasonal and sub-seasonal forecasting systems.
- It discusses how AI can contribute to improve the accuracy and reliability of forecasts, and the outcomes of the Regional Climate Outlook Forum (RCOF).
- The report also highlights the different initiatives by AGRHYMET to integrate AI in forecasting systems and articulates AICCRA contributions.
- By integrating AI into its operations, AGRHYMET aims to address the unique challenges of forecasting in West Africa and the Sahel, regions characterized by complex and highly variable climatic conditions in addition to a poor ground-based data availability.

## Introduction

The integration of artificial intelligence (AI) into seasonal and sub-seasonal forecasting represents a revolutionary shift in climate science, enhancing the accuracy and depth of insights into atmospheric systems. AI's unparalleled ability to process vast and diverse datasets, including satellite imagery, in-situ measurements, and historical climate records, enables the detection of complex patterns and correlations often overlooked by traditional methods. By leveraging advanced machine learning models, such as neural networks and ensemble algorithms, AI has demonstrated its capacity to outperform conventional statistical and dynamical models, particularly in predicting temperature anomalies, precipitation variability, and extreme weather events. This is especially critical for sub-seasonal timescales, where the inherent uncertainty of atmospheric dynamics presents substantial challenges to traditional forecasting techniques.

Recent research highlights the broad applications of AI in climate prediction. Mouatadid et al. (2023) introduced a machine learning-based adaptive bias correction method that improves the reliability of sub-seasonal forecasts by systematically addressing and reducing biases inherent in climate models. Similarly, the "Outcomes of the WMO Prize Challenge to Improve Subseasonal to Seasonal Forecasts" (2022) showcased how AI-driven methodologies significantly enhance forecast accuracy, as evidenced by an international competition spearheaded by the World Meteorological Organization. Chen et al. (2024) highlighted the FuXi-S2S model, which leverages over seven decades of reanalysis data to outperform traditional ensemble forecasting tools, marking a significant leap in predictive capabilities.

In Europe, Van Straaten (2023) explored the application of machine learning techniques to refine sub-seasonal temperature forecasts, emphasizing the critical role of statistical post-processing techniques in achieving higher forecast accuracy. Elena et al. (2024) further demonstrated the utility of machine learning models as post-processing tools, integrating ensemble forecasts with observational data to enhance predictions of monthly precipitation and temperature. Research by Daniel et al. (2022) and Sijie et al. (2021) delves deeper into advanced AI methodologies, such as transformers and hybrid models, illustrating their potential to boost the efficiency and precision of traditional dynamical models. Additionally, Sijie et al. (2020) conducted a comprehensive analysis of machine learning approaches to sub-seasonal climate forecasting in the U.S., shedding light on the opportunities and challenges of AI adoption in this domain. Collectively, these studies illustrate AI's transformative potential in climate science, addressing the complexities of sub-seasonal and seasonal forecasting. By advancing prediction methodologies, AI is positioned to play a pivotal role in supporting climate resilience and informed decision-making in the face of escalating climate challenges.

Given the significant advancements in climate forecasting enabled by artificial intelligence (AI), AGRHYMET has expressed a strong interest in adopting this transformative technology to enhance its forecasting capabilities. This strategic initiative aligns with Decision 9 (EC-72) of the World Meteorological Organization (WMO), which emphasizes leveraging innovative technologies to improve climate predictions' accuracy, timeliness, and utility globally. By integrating AI into its operations, AGRHYMET aims to address the unique challenges of forecasting in West Africa and the Sahel, regions characterized by complex and highly variable climatic conditions in addition to a poor ground-based data availability.

## Enhancing climate forecasting in West Africa and the Sahel: integration of AI for accuracy and reliability

For decades, AGRHYMET has been developing seasonal climate forecasts for West Africa and the Sahel based on consensus approaches derived from regional meetings between experts and decision-makers. These meetings, organized within the framework of Regional Climate Outlook Forums (RCOFs), focused on qualitative analyses of climate trends, often guided by experience and human judgment. Although this method provided useful information, it had significant limitations. The subjectivity of individual interpretations, strongly influenced by expert intuition, introduced variability that sometimes compromised the reliability of results. In addition, the absence of a standardized methodological framework led to discrepancies in the climate scenarios proposed, which could cause confusion among decision-makers and end-users, such as farmers or government officials. Finally, these consensus forecasts had limited reliability, with recommendations that struggled to convince political decision-makers or meet the specific needs of local communities and economic actors. These shortcomings underscore the need to improve the forecasting approach, incorporating more robust tools based on measurable data and objective methodologies.

Developing advanced and reliable forecasting tools is imperative in the context of increasing extreme climatic events, including frequent floods and droughts. Therefore, AGRHYMET is engaging to transform its seasonal forecasting methods by integrating modern and innovative approaches. Artificial intelligence (AI) and machine learning (ML) are currently being explored to enhance the accuracy of seasonal and sub-seasonal forecasts. AI-based approaches offer the ability to analyze large datasets, recognize complex patterns, and make more adaptive forecasts for changing climate conditions. These technologies promise to complement existing methods and provide more robust and reliable forecasts, especially in a region as complex and variable as West Africa and the Sahel. Implementing AI and ML in seasonal forecasting does not seek to replace existing methods but to complement and enhance them. Traditional models often struggle with the inherent uncertainties of mid- to long-range atmospheric dynamics, particularly in regions with limited observational data. AI-based tools can bridge these gaps by integrating and optimizing diverse data sources, improving the reliability and granularity of forecasts. Moreover, these tools have the potential to be tailored to the unique climatic and socio-economic contexts of West Africa and the Sahel, providing actionable insights for policymakers, farmers, and disaster management teams.

## Capacitating AGRHYMET on AI-driven climate forecasting

To successfully transition to AI-based forecasting tools, AGRHYMET understands that human resources are the foundation of any successful technological transformation. While AI offers groundbreaking opportunities to enhance climate forecasting, these tools require skilled individuals who can effectively operate, interpret, and adapt them to the organization's specific needs. Recognizing this, AGRHYMET has identified resource persons with expertise in AI and machine learning (ML) to lead its capacity-building efforts. These experts are instrumental in equipping AGRHYMET staff with advanced knowledge and skills in modern data collection, analysis, and modeling techniques. Supported by the "Accelerating Impacts of CGIAR Climate Research for Africa" (AICCRA) project, AGRHYMET has implemented a robust capacity-building program to equip its staff with the knowledge and technical skills necessary to effectively leverage AI and machine learning (ML) technologies. These efforts are critical in fostering innovation, enabling AGRHYMET to harness the full potential of advanced forecasting tools tailored to the region's unique climatic and socio-economic contexts.

**The AGRHYMET team's capacity-building initiative focused on key modules as followed:**

1. **The first training module consisting of three sessions on machine learning methods applied to seasonal forecasting.**
   **Session 1:** Reminder on traditional methods, including multiple linear regression, multicollinearity (addressed through principal component regression and canonical correlation analysis), and regularization algorithms..

   Session 2 and 3: Support Vector Machines (SVM), decision trees, ensemble learning and random forests,

2. **The second training module consisting of three sessions:**
   **Session 1 and 2:** Neural network and Extreme Learning Machine.

   **Session 3:** Probabilistic Output of Extreme Learning Machine (PO-ELM), Extending Probabilistic Output of Extreme Learning Machine (EPO-ELM).

3. **The third training module consisting of three sessions:**
   **Session 1 and 2:** Extended Logistic Regression (ELR), Bayesian Model Averaging (BMA)

   **Session 3:** Introduction to deep learning

4. **An in-person session with experts to deepen the topics covered during the online sessions.**
   These modules aimed to equip participants with the essential skills for effective application and knowledge transfer. The training covered fundamental theory and methods with applications using appropriate software and tools. As a result of this training, participants acquired comprehensive expertise in all these methods allowing the development of AGRHYMET in-house tools for seasonal forecast embedding most of these methods.

## Potential Impact of AI-based approaches on Regional Climate Outlook Forums

AGRHYMET's RCOFs have historically played a crucial role in regional climate coordination, serving as a platform for experts to analyze climate trends and provide recommendations. These forums facilitated collaboration among meteorologists, hydrologists, and policymakers, ensuring a shared understanding of climate risks. However, they predominantly focused on qualitative discussions, heavily influenced by expert judgment and subjective (no-reproductible) interpretations. While these forums provided valuable insights, the reliance on intuition and experience introduced variability in forecasts and recommendations, occasionally undermining their reliability and effectiveness for decision-making. The integration of AI-based approaches and advanced tools will significantly transform the structure and outcomes of seasonal forecast forums, improving their scientific rigor and practical utility.

- **Discussions oriented on the forecast's quality and interpretation, rather than the forecast product development.** Traditional forums focused on qualitative debates to define the forecast product (manually and subjectively the delimitations of forecast boundaries), with the new approach the forecast product is obtained automatically from the outputs of the WAS S2S software application. Now the discussions will focus on the interpretation, analyses, and communication of the objective forecast product.
- **Localized recommendations:** One of the key advancements has been the ability to generate tailored forecasts for specific areas and contexts. AI's capacity to take into consideration (including localized data) has enhanced the relevance of recommendations, enabling actionable strategies for diverse stakeholders, from smallholder farmers to disaster management agencies. These personalized insights foster better preparedness and more effective responses to climate variability.

## AI-based seasonal to sub-seasonal forecasting systems implemented by AGRHYMET

### The role of AICCRA in supporting AGRHYMET in implementing the AI-based forecasting approach.

The project **Accelerating Impacts of CGIAR Climate Research for Africa (AICCRA)** has been instrumental in strengthening AGRHYMET's ability to deliver high-quality climate services. AICCRA provides the necessary support, resources, and collaborative networks that enable regional centers to develop and adopt innovative forecasting tools and methodologies. Through AICCRA, AGRHYMET has gained access to cutting-edge research, technical expertise, and peer partnerships—all of which contribute to more accurate and timely climate forecasts.

### Adaptation of PyCPT for Seasonal and Hydrological Forecasting

A significant outcome of AICCRA's Phase I is the development and deployment of PyCPT (Python Climate Predictability Tool) by the International Research Institute for Climate and Society (IRI). PyCPT automates Model Output Statistics (MOS), including techniques such as Multiple Linear Regression (MLR), Principal Component Regression (PCR), and Canonical Correlation Analysis (CCA), as well as ensemble methods like simple averaging and extended logistic regression.

Originally designed for climate forecasting, PyCPT has been successfully adapted by AGRHYMET for seasonal hydrological forecasts, demonstrating its versatility and robustness. Through automation, PyCPT improves the traceability and reproducibility of forecasts, ensuring that statistical methodologies are consistently applied and can be more easily evaluated and refined.

Despite these advancements, PyCPT does not yet integrate certain regional forecasting methods commonly used during West African Regional Climate Outlook Forums (RCOFs), such as analog methods and observation-based forecasts. In addition, the tool does not currently cover the full range of climate variables essential for comprehensive forecasting in West Africa and the Sahel. Recognizing these limitations, AGRHYMET is actively expanding PyCPT's capabilities to accommodate these methods and variables.

### AGRHYMET adopted an operational approach to implementing AI-based forecasts for West Africa and the Sahel

To address existing constraints and incorporate AI and machine learning techniques, AGRHYMET has introduced a methodology that is operationalized through a specialized Python tool named **WAS_S2S (West Africa and Sahel Seasonal to Sub-Seasonal)**, representing a significant leap forward in seasonal forecasting.

WAS_S2S synthesizes **traditional statistical methods** with **state-of-the-art ML algorithms**, enabling AGRHYMET to deliver more accurate, reliable, and adaptable forecasts. Below, we outline the main approaches and components that define the integration of AI in the new generation of seasonal forecasting:

i. Use of GCM Outputs
- Refers to predictions derived from **General Circulation Models (GCMs)**.
- GCM outputs are evaluated against observations (the "predictand") to identify the best-performing models.
- The three top-performing models are selected and then **weighted** based on performance metrics (e.g., Kling-Gupta Efficiency [KGE] or Pearson correlation).

ii. Application of Statistico-Dynamics method
- Bridges **statistical methods** and **dynamical outputs** (e.g., GCMs).
- Relies on relationships between GCM-derived predictors (e.g., Sea Surface Temperature, wind, specific humidity, rainfall) and the observed target variable (predictand).

iii. Observations-Based with Lag
- Uses observed historical data (e.g., rainfall) and reanalysis datasets (e.g., SST, wind, humidity) with **lagged relationships** to account for temporal dependencies in climate phenomena.

iv. Analog-Based
- Identifies historical analogs by examining key climate drivers (e.g., SST evolution).
- Employs **similarity/dissimilarity analysis** to predict future conditions.

### Primary functions used in the forecasting process

Within these approaches, three primary function types link predictors (X) to the target variable (Y):

i. f(•): Statistical learning techniques
- Linear methods (MLR, Ridge, Lasso, ElasticNet)
- Principal Component Regression (PCR)
- Canonical Correlation Analysis (CCA)
- Logistic Regression
- Quantile Regression
- Polynomial Regression
- Advanced ML models such as Support Vector Regression, MARS, Boosting, Random Forest, Neural Networks, etc.

ii. **g(•): Similarity/dissimilarity analyses**

- Techniques like Dynamic Time Warping (DTW), Self-Organizing Maps (SOM), Principal Component Analysis (PCA), and Convolutional Neural Networks (CNN) using Siamese Networks.

iii. **h(•): Combination methods**

- Extreme Learning Machines (ELM)
- Extended Logistic Regression (ELR)
- Weighted mean (based on correlation or skill scores such as ROC)
- Bayesian Model Averaging (BMA)

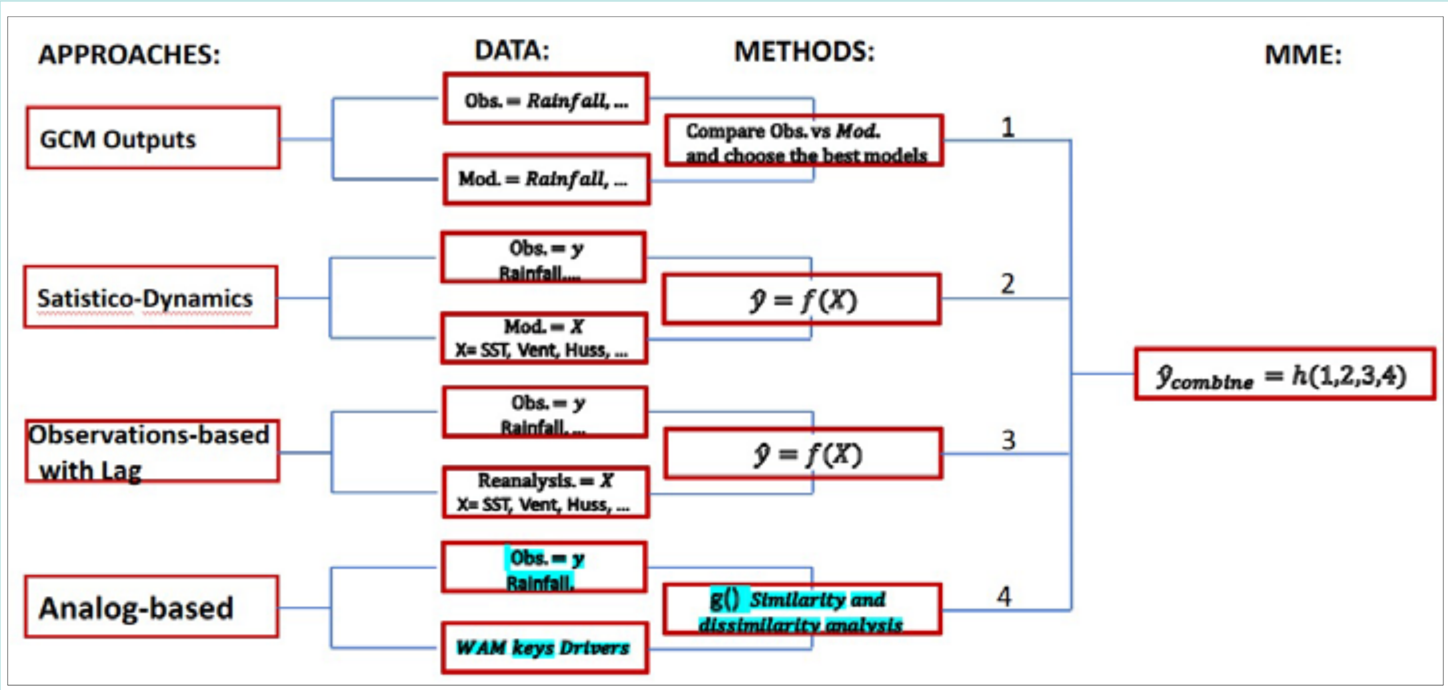


*Figure 1: Conceptual overview of the New Generation of Seasonal Forecast approaches in WAS_S2S*

## Key modules of WAS S2S new forecasting system

### i. WAS Download Module

This module offers utilities for downloading, processing, and combining a variety of meteorological and agro-meteorological datasets. It handles reanalysis data, seasonal forecast models, and indicators from sources like the Copernicus Climate Data Store (CDS) and the IRI Data Library. Key functionalities include:

- **Efficient Management of Large-Scale Climate Data**
- **Spatiotemporal Slicing** for targeted analysis
- **Data Preparation** for further modeling or analysis

### ii. WAS_Predictant Module

This module contains classes for analyzing **spatiotemporal rainfall and temperature data** and deriving indirect predictands such as onset and cessation dates for the rainy season. Key features include:

- **Onset Calculation**
  - Determines rainy season onset dates based on cumulative rainfall and dry-day thresholds.
- **Cessation Calculation**
  - Identifies rainy season cessation dates using **soil moisture balance** and **evapotranspiration** data.
- **Dry/Wet Spell Analysis**
  - Calculates the longest dry/wet spells, counts dry/wet spells, and tracks rainy days within the onset–cessation period.
- **Heatwave Index**
  - Computes **Heat Wave Severity Duration Index (HWSDI)** and counts seasonal heatwave events.

### iii. WAS_Merging Module

Designed to merge **observational and gridded** data. Main functionalities include:

- **Gridded & Point Data Integration**
  - Aligns and merges multiple datasets for improved forecast accuracy.
- **Flexible Interpolation Methods**
  - Supports regression, neural networks, and Kriging of residuals for spatial consistency.
- **Cross-Validation**
  - Implements metrics (e.g., RMSE) using a Leave-One-Out scheme.

- **Customizable Settings**
  - Permits variogram model selection, lag parameter tuning, and a range of Kriging configurations.

### iv. WAS_..._Models Module

Structured into submodules for different modeling approaches:

- **was_linear_models:** MLR and regularization algorithms (Ridge, Lasso, ElasticNet)
- **was_pcr:** Principal Component Regression (PCR)
- **was_cca:** Canonical Correlation Analysis (CCA)
- **was_machine_learning:** Logistic Regression, Quantile Regression, Polynomial Regression, Support Vector Regression, MARS, Random Forest, XGBoost, Neural Networks, etc.
- **was_analog:** Analog-based models using SOM, PCA, CNN, and Siamese Networks

### v. WAS_Cross Validate Module

Implements a custom time-series cross-validation scheme, splitting the data into training and test sets while preserving the temporal order. Users can optionally specify an omission period after the test index to ensure data independence.

### vi. WAS Verification Module

A **centralized hub** for calculating and visualizing both **deterministic** and **probabilistic** verification metrics. Core functionalities include:

- **Deterministic Metrics**
  - RMSE, MAE, Pearson Correlation, Nash-Sutcliffe Efficiency, Kling-Gupta Efficiency, and placeholders for advanced metrics (e.g., Taylor Diagrams).
- **Probabilistic Metrics**
  - Classification into terciles (below-normal, near-normal, above-normal).
  - Metrics such as GROC, RPSS, Ignorance Score, Resolution, and Reliability.
  - Emphasizes **grid-based** probabilistic verification essential for climate prediction.
- **Ensemble Metrics**
  - Continuous Ranked Probability Score (CRPS) and aggregation methods (mean, median) for ensemble forecasts.
- **Visualization**
  - Reliability diagrams, ROC curves (with bootstrapped CIs), and map-based scores for spatial interpretation.
- **GCM Validation**
  - Computes scores and probabilities for hindcasts.
  - Aligns model and observational data both spatially and temporally.
  - Supports ensemble mean or median for deterministic scores and handles ensemble outputs for probabilistic metrics.

### vii. WAS_MME Module

Implements **combination methods** (multimodel ensemble, or MME) to integrate outputs from different approaches (see Figure 1). Techniques include:

- Extreme Learning Machines (ELM)
- Extended Logistic Regression (ELR)
- Weighted Mean (**based on correlation or ROC skill scores**)
- Bayesian Model Averaging (BMA)

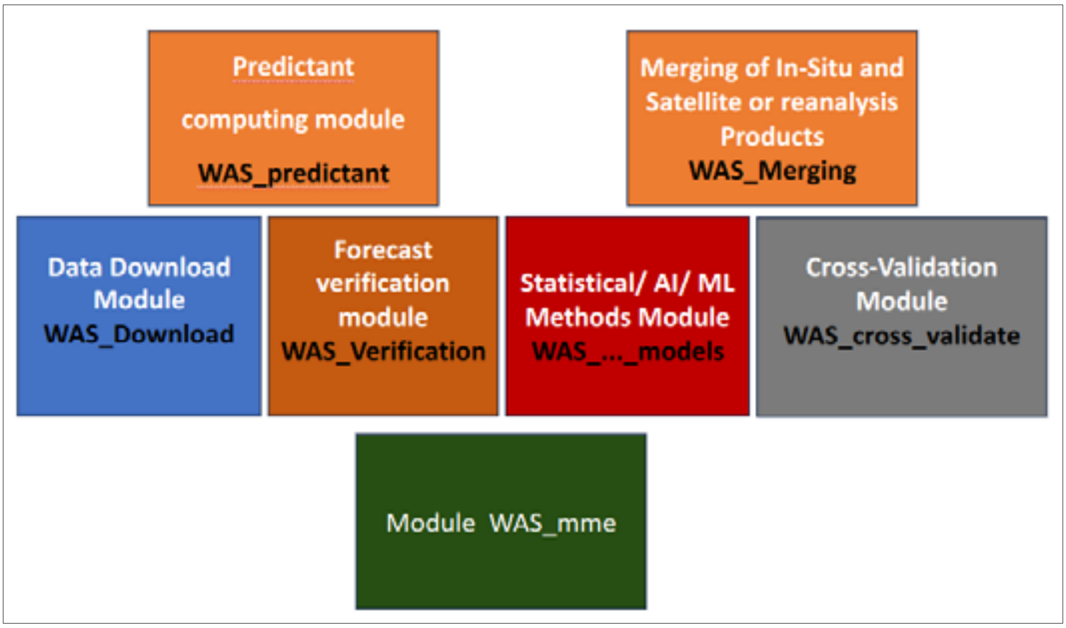


*Figure 2: Demonstration of the WAS_S2S modules and their use in combining multiple climate forecasting methods.)*

## Preliminary results of the use of AI-based forecasting system

WAS_S2S was used during a recent on-the-job training session on new generations of seasonal forecasts in West Africa and the Sahel (https://hmandela.github.io/WAS_S2S_Training). During this training, participants validated both global climate models and statistical methods integrated into WAS_S2S.

A key product from this session was a validation plot illustrating the performance of selected GCMs and statistical post-processing methods. The training underscored WAS_S2S's versatility in evaluating different forecast models, offering a user-friendly and efficient workflow for analyzing forecast skills in a structured manner.

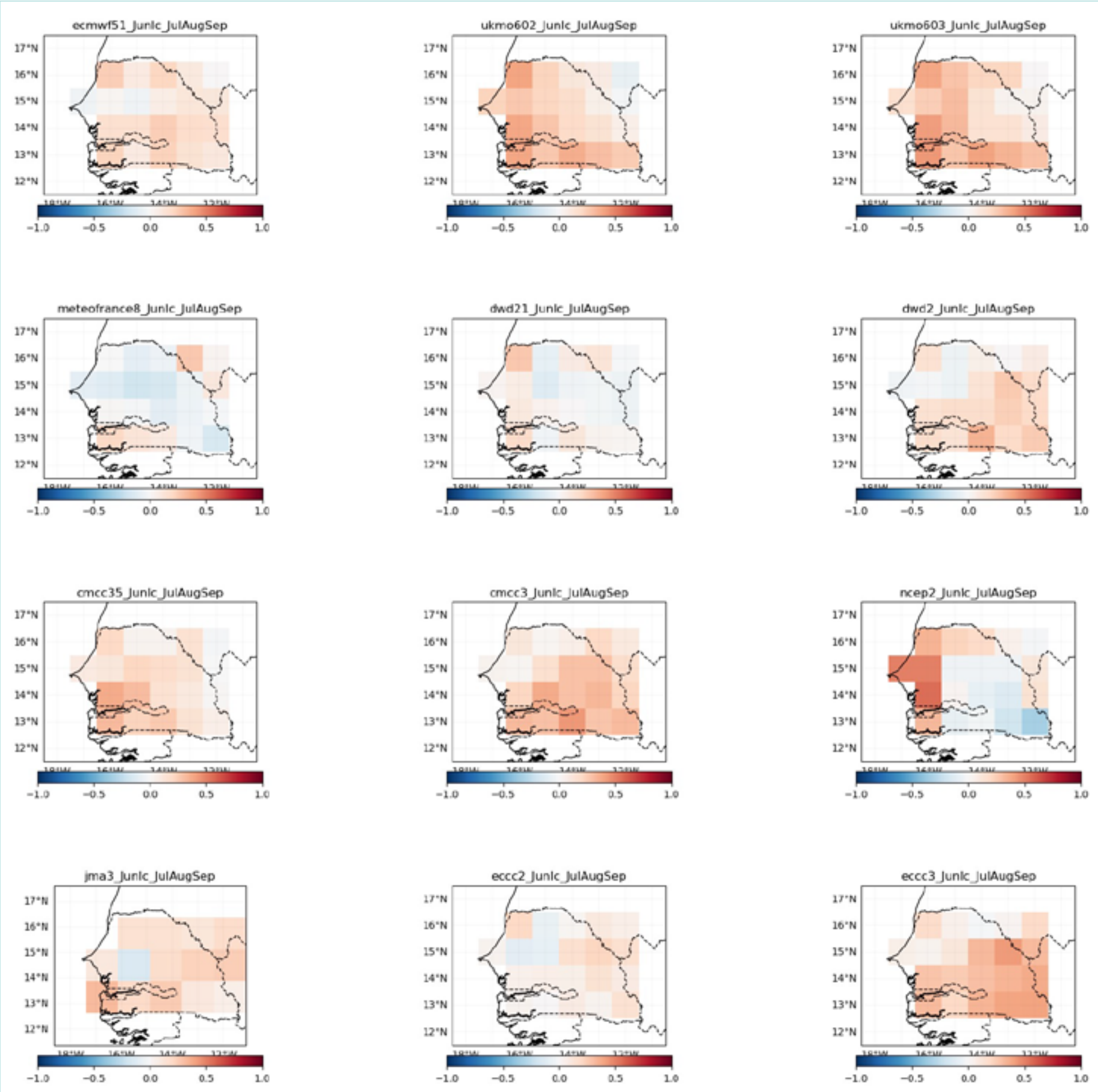


*Figure 3: Example output (Pearson correlation) from a training participant showcasing integrated GCM validation using WAS_S2S.)*

Notably, WAS_S2S leverages libraries such as xarray, dask, xeofs, scikit-learn, and xcast, ensuring robust data structures, modular statistical workflows, and comprehensive analysis. It's available through this link (https://github.com/hmandela/WAS_S2S).

## References

- Chen L., Xiaohui Z., Deliang Ch., Chensen L., Hao L., Shang-Ping X., Libo W., Jie W., Bo L., Qingchen C., Zixin H. and Yuan Q. (2024): A Machine Learning Model that Outperforms Conventional Global Subseasonal Forecasting. Nature Communications.
- Daniel S. C., Mohamed A. Z., Etienne E. V., Johannes S., Bianca Z. and Campbell W. (2022). Improving Sub-Seasonal to Seasonal Forecast Using a Time-Series Transformer. AGU Fall Meeting.
- Elena O., Haokun L., Raphael R., Benjamin A., And Rebecca W. (2024). Beyond Ensemble Averages: Leveraging Climate Model Ensembles for Subseasonal Forecasting with Machine Learning. Artificial Intelligence for the Earth Systems.
- He, S., Li, X., DelSole, T., Ravikumar, P., Banerjee, A. (2020). Sub-Seasonal Climate Forecasting via Machine Learning: Challenges, Analysis, and Advances. arXiv.
- He, S., Li, X., Trenary, L., Cash, B. A., DelSole, T., Banerjee, A. (2021). Learning and Dynamical Models for Sub-seasonal Climate Forecasting: Comparison and Collaboration. arXiv
- Mouatadid, S., Orenstein, P., Flaspohler, G., Cohen, J., Oprescu, M., Fraenkel, E., Mackey, L. (2023). Adaptive Bias Correction for Improved Subseasonal Forecasting. Nature Communications.
- Van Straaten, J. W. (2023). Machine Learning to Improve and Understand Sub-Seasonal Forecasts of European Temperature. PhD Thesis, Vrije Universiteit Amsterdam.
- WMO. (2022). Outcomes of the WMO Prize Challenge to Improve Subseasonal to Seasonal Forecasts. Bulletin of the American Meteorological Society.
- World Meteorological Organization (WMO) Decision 9 (EC-72)
- WMO (2020). Guidelines for Operational Practices on Objective Seasonal Forecasting.


This info note is an output of the partnership between AICCRA West Africa cluster and AGHRYMET Regional Climate Centre for West Africa and the Sahel, a specialized institution of the Permanent Interstates Committee for Drought Control in the Sahel (CILSS). Led by AGRHYMET, this info note describes regional efforts to integrate Artificial Intelligence (AI) into seasonal and sub-seasonal forecasting systems. It articulates how AI can contribute to improve the accuracy and reliability of forecasts, and the outcomes of the Regional Climate Outlook Forum (RCOF).




- **Mandela C. Houngnibo** is Agrometeorologist at AGRHYMET.
- **Abdou Ali** is Senior Hydroclimatologist, Head of AGRHYMET's Department of Climate, Water and Meteorology, and AICCRA Focal Point
- **Boubacar Toukal Assoumana** is User Engagement Specialist at AGRHYMET.
- **Bernard Minoungou** is Hydrological modeling Specialist at AGRHYMET.
- **Alcade C. Segnon** is Science Officer for AICCRA West Africa Cluster and Scientist at the Alliance of Bioversity and CIAT.
- **Robert B. Zougmoré** is Director of AICCRA program and Principal Scientist at the Alliance of Bioversity and CIAT



## ABOUT AICCRA

CGIAR | AICCRA Accelerating Impacts of CGIAR Climate Research for Africa

Accelerating Impacts of CGIAR Climate Research for Africa (AICCRA) is a project that helps deliver a climate-smart African future driven by science and innovation in agriculture. It is led by the Alliance of Bioversity International and CIAT and supported by a grant from the International Development Association (IDA) of the World Bank. Explore our work at aiccra.cgiar.org

aiccra.cgiar.org CGIARAfrica