Sequential Hybrid Approach for Reliable Detection of Rainfall Pauses at the Beginning of the Rainy Season in Senegal: Towards a Predictive Tool for False Starts

Pape El Hadji Abdoulaye Gueye[1*]; Cherif Bachir Deme[2]; Diery Ngom[3]; Adrien Basse[4]

Department of TIC, UFR SATIC University Alioune Diop of Bambey, Bambey Senegal
*Corresponding Author : https://orcid.org/0009-0000-5222-3474

Received:- 13 October 2025/ Revised:- 24 October 2025/ Accepted:- 29 October 2025/ Published: 05-11-2025

Abstract— *This work focuses on the detection of false onsets of the rainy sea son in Senegal, a critical factor that can lead farmers, particularly smallholders, to initiate agricultural activities prematurely. Such errors, caused by misleading early rainfall events, result in yield losses and increase farmers' vulnerability to climate variability. Unlike existing methods, our approach incorporates statistical tests (such as Pettitt, Kendall, and Lombard) to enrich the input dataset with relevant change points related to rainfall, soil moisture, and vegetation. This enrichment step, combined with a formal detection of false onsets based on climatic, phenological, and statistical criteria, enhances the relevance, robustness, and contextualization of detection compared to purely statistical or physical approaches. In this context, a deep learning methodology was developed to identify false onsets at an early stage using multivariate climatic data. We designed a hybrid model combining LSTM, GRU, and multi-head attention layers to extract complementary representations of the input sequence. Model hyperparameters were optimized through Bayesian search to enhance detection performance. Results show consistent improvements across all key metrics: accuracy increased from 0.84 to 0.88, F1-score from 0.833 to 0.86, recall remained perfect at 1.0, precision rose from 0.767 to 0.81, and AUC improved from 0.900 to 0.92. These gains demonstrate the overall robustness of the optimized model, ensuring more reliable detection of false onsets*.

Keywords— False onset, Deep learning, Statistical tests, LSTM, GRU, Attention, Bayesian optimization.

I. INTRODUCTION

In semi-arid regions, particularly in the Sahel, agricultural drought—defined as a water deficit affecting crops and their productivity—constitutes a major risk to food security [1]. This region, and notably Senegal, is characterized by strong interannual rainfall variability [2], frequent delays in the onset of the rainy season, an extended dry season, and a continuous increase in land surface temperature (LST) [3, 4]. These climatic changes disrupt traditional agricultural cycles and increase the vulnerability of rainfed production systems.

A critical yet often overlooked phenomenon in this context is the **false onset of the rainy season**. It occurs as a series of weak initial rains—sometimes totaling about 20 mm over 2–3 days—followed by a dry spell. This sequence can create a false sense of security for farmers, who may sow prematurely, exposing their crops to early-season water stress that can significantly reduce yields. Despite extensive research on drought detection, the false onset—a key factor in shifting the rainfall calendar—has received relatively little attention, limiting the predictive capabilities of existing models [5]. Early detection is therefore essential to adjust sowing schedules, mitigate water stress risks, and support farmers' decision-making.

Recent advances in **remote sensing** and **machine learning** have greatly enhanced drought monitoring and forecasting. Remote sensing provides continuous spatio-temporal information despite the scarcity of in situ measurements [4]. Vegetation and soil moisture indices such as NDVI, VHI, LSWI, and SIF have been widely used to assess vegetation water stress [6–9]. Moreover, composite indices integrating precipitation, temperature, and remote sensing data (e.g., CDI [10], SMADI [11], IDSI [12]) offer improved insights into drought dynamics.

In parallel, **machine learning** models have shown strong potential for drought and rainfall forecasting, particularly when combined with physical constraints that enhance predictive realism [9, 13]. Explainable Artificial Intelligence (XAI) techniques further highlight the relative importance of climatic variables, showing for example that temperature can be more influential than precipitation [1].

Within the **Sahelian context**, the timing of the rainy season is strongly influenced by large-scale climate phenomena such as ENSO. Some approaches, such as Kohonen maps, have been used to classify the onset and cessation of rainy seasons [5], while logistic regression models based on climate predictors (SST, precipitable water, dew point temperature, winds) have been proposed for seasonal onset forecasting [14]. However, these statistical models often show limited interregional transferability and struggle to capture complex intra-seasonal dynamics—particularly those linked to false starts.

To address these challenges, **recurrent neural networks (RNNs)** have been increasingly applied in hydrological and climatic studies. LSTM networks have proven effective for long-term temporal dependencies [15, 16], whereas GRU architectures better handle short-term fluctuations [20]. Transformer models, on the other hand, excel at capturing non-local relationships across long sequences. Hybrid architectures that integrate these models have recently emerged as a promising direction, leveraging heterogeneous data such as meteorological time series, soil properties, and remote sensing observations [18].

However, in West African contexts, the scarcity and heterogeneity of ground and satellite data remain a key limitation. Existing studies often overlook the **false onset phenomenon**, focusing instead on general drought indices or seasonal rainfall patterns. Statistical models alone lack the capacity for robust interregional generalization, while purely neural approaches may suffer from overfitting or reduced interpretability.

To overcome these limitations, this study proposes a **hybrid neural architecture combining LSTM, GRU, and Transformer components** for the detection and early prediction of false rainy season onsets in Senegal. The approach leverages multisource climatic and phenological data, enriched by statistical change-point detection tests (Pettitt, Kendall, Lombard) to identify structural shifts in rainfall, soil moisture, and vegetation dynamics. By integrating climatic, phenological, and statistical criteria, we define a **False Onset Index (FOI)** that enables more robust and realistic detection of false starts, improving the reliability of seasonal forecasts and supporting farmers in optimizing their sowing calendars.

The remainder of this paper is structured as follows: Section 2 describes the dataset, the LSTM–GRU–Transformer hybrid model, and hyperparameter optimization. Section 3 presents the results, including performance evaluation, robustness analysis, SHAP-based interpretability, and ablation studies. Section 4 also discusses the implications, limitations, and perspectives of the study, and Section 5 concludes by summarizing the main contributions.

II. MATERIAL AND METHODS

It is a The data used in this study were compiled from multiple climatic and remote sensing sources covering the period 2000–2016 and harmonized into a consistent dataset. They include precipitation from CHIRPS (5 km), temperature, humidity, radiation, and wind from POWER-DAV (0.5°), soil moisture, evapotranspiration, and net radiation from FLDAS (25 km), as well as NDVI and land surface temperature (LST) from MODIS (1 km). In addition, large-scale climate indicators such as the Standardized Precipitation–Evapotranspiration Index (SPEI) and the Oceanic Niño Index (ONI) were incorporated. All variables were aggregated over Senegal and over the rainy season period to obtain representative annual values.

False onsets of the rainy season were automatically identified using relevant climatic criteria. Specifically, the variable *faux_demarrage* was set to 1 when cumulative precipitation exceeded 600 mm and the start of the season was too early (sos_doy < 145) or too late (sos_doy > 180), or when the month of July showed significant water stress, defined as SPEI_1_July < –1. Otherwise, *faux_demarrage* was set to 0. To enhance the robustness of the dataset, several statistical tests were applied to precipitation and temperature time series. The Pettitt test was used to detect abrupt change points in precipitation, the Lombard test to evaluate the influence of extreme values, and Kendall's tau to assess the association between temperature trends and false onsets. The outputs of these tests were added as auxiliary predictors, providing additional features for the learning model.

A hybrid deep learning model was developed by combining Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Transformer components to capture the multi-scale temporal dynamics of the climatic data. The LSTM component was designed to model long-term dependencies, the GRU to represent short- and medium-term variability, and the Transformer to extract non-local relationships within sequences through multi-head attention. The outputs of these three branches were concatenated and passed through a dense layer with a sigmoid activation function to perform binary classification (false start or not). The model was trained using the Adam optimizer for stable convergence and balanced performance in terms of accuracy, recall, and F1-score.

Hyperparameter optimization was carried out using Bayesian search with *KerasTuner*, which iteratively updates a surrogate model (Gaussian Process) to balance exploration and exploitation during the search. Tuned parameters included the number of LSTM and GRU units, the number of Transformer heads and key dimensions, and the learning rate. The optimal configuration was obtained with 16 LSTM units, 48 GRU units, 2 attention heads, a key dimension of 4, and a learning rate of 2.04×10^{-4}. Cross-validation using 20% of the training set ensured model robustness and generalization. This optimized hybrid architecture leverages the complementarity of recurrent and attention mechanisms to improve early detection of false rainy season onsets, thereby supporting better agricultural decision-making and resilience to climate variability in the Sahel.

III. RESULT AND DISCUSSION

3.1 Data Preparation:

The data come from our tabular dataset for false start detection, with *faux_demarrage* as the target variable and physical variables (such as PRECTOTCORR_SUM and T2M) as predictors. The data were normalized using a StandardScaler and split into training (80%) and test (20%) sets in a stratified manner to preserve a balanced distribution of positive and negative cases.

3.2 Static Cross-Validation Training:

A stratified 5-fold cross-validation ($k = 5$) was used to evaluate the model's generalization performance. In each iteration, the model was trained on $k-1$ folds and validated on the remaining fold, ensuring that each observation was used once for validation. Results indicate that after hyperparameter optimization, the model becomes both more accurate on average and more stable across folds.

3.3 Performance Evaluation:

The metrics used to assess the model allow evaluation of its relevance from multiple perspectives. Model performance shows improvement after hyperparameter optimization: accuracy increases from 84% to 88%, precision rises from 0.767 to 0.810, and the F1-score reaches 0.86 compared to 0.833 before optimization. Recall remains maximal (1.0), ensuring detection of all false start cases, which is essential to avoid operational errors. The AUC also improves slightly to 0.92, indicating enhanced overall discriminative ability. These results demonstrate greater robustness of the model across key metrics, confirming the effectiveness of hyperparameter optimization.

TABLE 1
COMPARISON OF PERFORMANCE BEFORE AND AFTER HYPERPARAMETER OPTIMIZATION

Metric	Before Optimization	After Optimization
Accuracy	0.84 ± 0.196	0.88 ± 0.160
AUC	0.900 ± 0.200	0.920 ± 0.180
Precision	0.767 ± 0.291	0.810 ± 0.250
Recall	1.000 ± 0.000	1.000 ± 0.000
F1-score	0.833 ± 0.211	0.860 ± 0.196

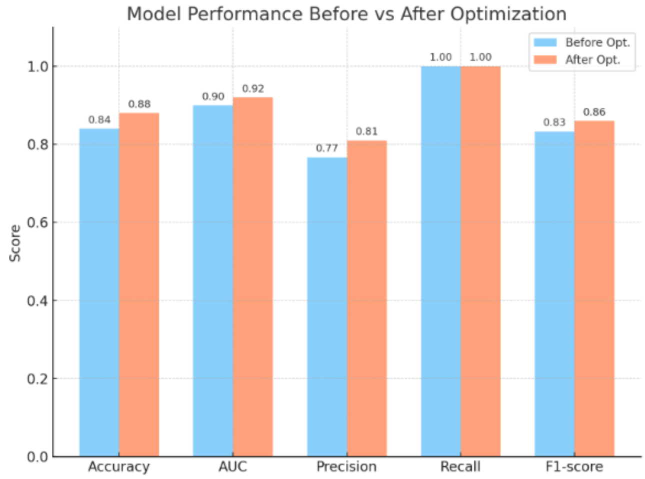

FIGURE 1: Performance before and after optimization

Performance improved after optimization: accuracy rose from 84% to 88%, precision from 0.767 to 0.810, F1-score from 0.833 to 0.86. Recall remains 1.0, ensuring all false start cases are detected. AUC also improved to 0.92.

3.4 Prediction Explanation (SHAP) and Robustness

To interpret the model, SHAP kernel values were computed on 100 test cases, with 30 bootstrap samples to ensure robustness. SHAP values identify the features with the greatest impact on predictions, while the mean and standard deviation highlight both their relevance and variability, which is crucial for model interpretation and improvement. Figure 2 shows that false starts are primarily influenced by recent hy 9 drological indicators. In particular, SPEI_1_July is the most important fea ture at the beginning of the season, with short-term indices (SPEI_1) having stronger effects than longer-term six-month indices. This highlights the pre dominance of immediate atmospheric conditions (June–July). Composite in dices such as SPEI are more relevant than raw precipitation (Rainf_f_tavg), confirming the effectiveness of synthetic water stress indicators. Global (oni) and phenological factors (sos_doy) have negligible influence, validating the predominance of local and recent causes

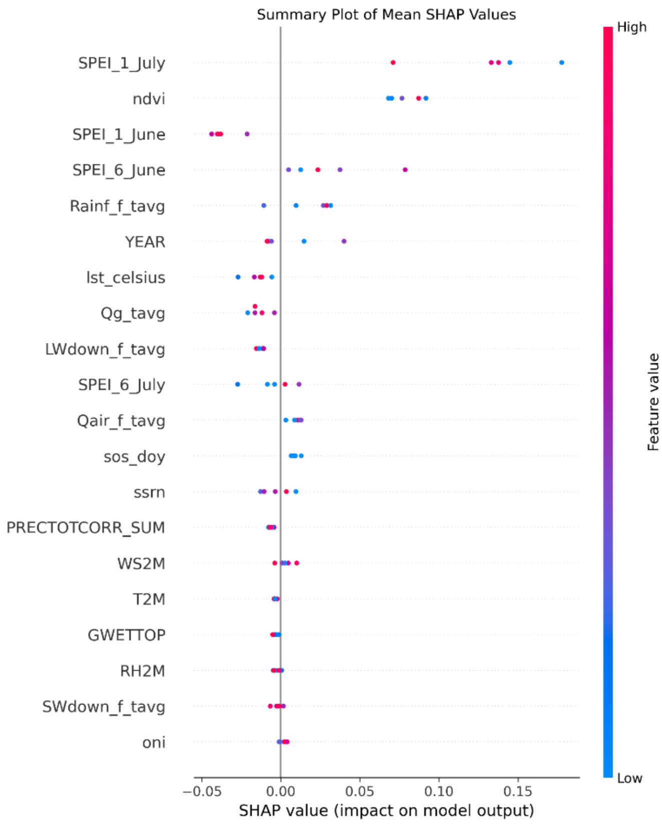

FIGURE 2: Variable importance according to SHAP summary

3.5 Robustness via Bootstrapping

To estimate the robustness or variability of SHAP values, 30 bootstrap samples were drawn from the test set, and SHAP values were computed for each sample. The mean and standard deviation of these values provide error bars and more reliable estimates. According to Figure 3, SPEI_1_July exhibits the highest mean SHAP value (+0.19), followed by PRECTOTCORR_SUM (+0.07). Both variables positively influence false starts during the rainy season. Bootstrapped SHAP confirms the robustness of these findings, while the relatively low standard deviation for the top features indicates consistent importance across resamples. This strengthens the understanding of underlying climatic mechanisms and provides confidence in model explanations.

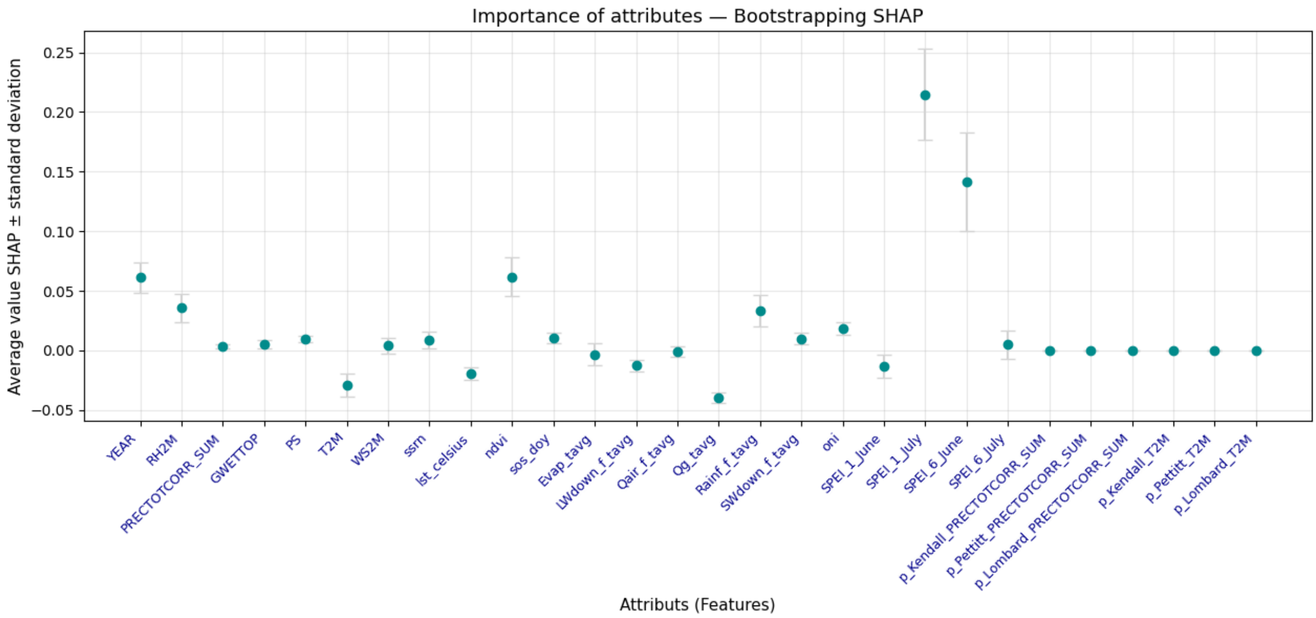

FIGURE 3: Variability of SHAP values estimated via bootstrapping

3.6 Performance Evaluation Based on Timestep

The time step determines the number of sequences used as model input to predict false starts.We varied this time step from 1 to 5 to assess its impact on model performance. Table 2 presents the mean accuracy, area under the ROC curve (AUC), precision, and F1-score (with standard deviation) across 5-fold cross-validation. According to Table 2, the model achieves the best performance in terms of Accuracy, AUC, Precision, andF1-score when the Timestep is 2. This highlights the importance of tuning this parameter to optimize the false start detection model.

TABLE 2
PERFORMANCE OF THE LSTM + GRU + TRANSFORMER MODEL AS A FUNCTION OF THE TIMESTEP

Timestep	Accuracy (mean ± std)	AUC (mean ± std)	Precision (mean ± std)	F1-score (mean ± std)
1	0.84 ± 0.196	0.900 ± 0.200	0.767 ± 0.291	0.833 ± 0.211
2	0.88 ± 0.160	0.920 ± 0.180	0.810 ± 0.250	0.860 ± 0.196
3	0.86 ± 0.120	0.910 ± 0.170	0.800 ± 0.240	0.850 ± 0.190
4	0.85 ± 0.140	0.905 ± 0.180	0.795 ± 0.230	0.845 ± 0.200
5	0.83 ± 0.150	0.890 ± 0.190	0.780 ± 0.250	0.830 ± 0.210

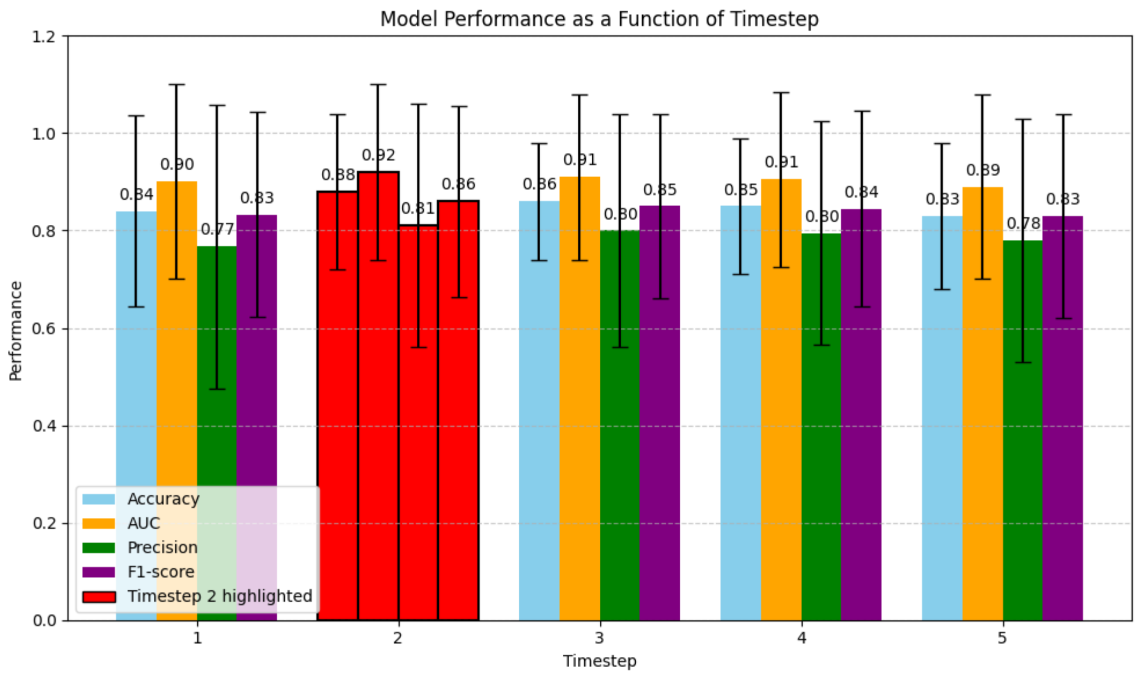

FIGURE 4: Selection of the optimal timestep for the hybrid model

Best performance occurs at Timestep = 2.

3.7 Ablation Study of the Hybrid Model:

To evaluate the individual contribution of the model components in predicting false starts of the rainy season, we conducted an ablation study analyzing the impact of each component on the achieved performance.

TABLE 3
ABLATION STUDY OF THE HYBRID MODEL. VALUES REPRESENT MEAN AND STANDARD DEVIATION OVER 5-FOLD CROSS-VALIDATION

Model	Accuracy	AUC	F1-score
LSTM only	0.84 ± 0.196	0.90 ± 0.20	0.767 ± 0.291
LSTM + GRU	0.88 ± 0.16	0.92 ± 0.18	0.81 ± 0.25
LSTM + Transformer	0.86 ± 0.12	0.91 ± 0.17	0.80 ± 0.24
GRU + Transformer	0.85 ± 0.14	0.905 ± 0.18	0.795 ± 0.23
Full (LSTM + GRU + Transformer)	0.88 ± 0.16	0.92 ± 0.18	0.86 ± 0.196

3.8 Justification of Performance

The full hybrid model (LSTM + GRU + Transformer) outperforms partial models (LSTM only, LSTM + GRU, LSTM + Transformer, GRU + Trans former) in terms of Accuracy, AUC, and F1-score. LSTM alone partially captures long-term dependencies. Adding GRU improves selective memory, enhancing recall and F1-score. Integrating the Transformer introduces an attention mechanism that helps identify critical time points, improving AUC. The combination LSTM + GRU + Transformer fully leverages the LSTM's sequential representation, the GRU's selective memory, and the Transformer's multi-head attention, providing a robust representation of the data for detecting false starts during the rainy season. We now compare our results with previous studies and present limitations and future perspectives in the following section.

IV. DISCUSSION

4.1 Main Results:

The study confirms the effectiveness of a hybrid model combining LSTM, GRU, and multi-head attention for detecting false starts of the rainy season. Bayesian optimization improved performance: Accuracy increased from 84% to 88%, F1-score from

0.79 to 0.86, and recall from 0.9 to 1.0. These results demonstrate that the model effectively detects all cases of false starts, which is crucial for preventing premature sowing. Across all experiments (see Table 3), the model achieved a mean accuracy of 0.88±0.16, a mean AUC of 0.92±0.18, and a mean F1-score of 0.86±0.196.

4.2 Comparison with Existing Methods:

Compared to traditional methods such as the Composite Drought Index (CDI) or SMADI, our hybrid approach integrates phenological, climatic, and statistical criteria (Pettitt and Kendall tests), as well as the temporal patterns of precipitation. This allows for a more precise and context-aware detection of false starts, outperforming physical or purely statistical models that struggle to handle climate variability and the multidimensionality of the data. In comparison with recent approaches:

• The IRD model [14], although effective, is limited by the lack of attention mechanisms and low scalability. Our approach integrates multi source sequences and provides a robust and interpretable solution adapted to Sahelian environments.

• The satellite-based approach by Balti et al. [16] focuses on global drought indices. Our model, by specifically targeting false rainfall starts, adds local value for agricultural planning.

• The model by Ichi et al. [15], based on CAMELS, addresses global hydrological dynamics. Our architecture, optimized for low-resource environments, stands out with fine temporal granularity suited to Sahelian dynamics.

• The Scisimple approach [17] focuses on global seasonal forecasts. Our model complements this work by anticipating short and localized events, which are critical for agricultural decision-making.

• The reference model by Agudelo et al. [18] demonstrates solid performance for predicting USDM categories. In contrast, our approach specifically targets false starts, events absent from these categories, achieving robust performance across 5-fold cross-validation. Overall, our hybrid model outperforms partial configurations (LSTM alone, LSTM+GRU, etc.) and stands out for its ability to handle the complexity and local variability of the Sahelian climate

4.3 Limitations:

Despite the overall improvement in all metrics, certain limitations remain. The complexity of the hybrid model (LSTM + GRU + Transformer) can lead to longer training times and sensitivity to hyperparameter choices. Furthermore, the reliance on the quality and resolution of climate data may affect the robustness of predictions under extreme or unprecedented conditions. These points highlight the need for continuous monitoring and field validation to ensure operational reliability.

4.4 Future Perspectives:

To further enhance the performance and robustness of the model, efforts should focus on increasing and diversifying the data (including field observations and high-resolution time series), performing field validation to compare predictions with real-world observations, adapting the model regionally to specific climatic and agronomic conditions, and enabling real-time deployment through early warning platforms for farmers and decision-makers. These strategies will maximize the operational applicability and agronomic relevance of false start detection.

V. CONCLUSION

This work demonstrated the effectiveness of a hybrid model combining LSTM, GRU, and multi-head attention for detecting false starts of the rainy season. Bayesian hyper-parameter optimization significantly improved all key metrics, including accuracy, F1-score, and AUC, ensuring reliable and comprehensive detection of critical events. SHAP value analysis enhanced the understanding of the respective contributions of climatic variables, reinforcing the model's transparency and robustness. Despite limitations related to the quality and spatial resolution of the data used, this hybrid approach outperforms traditional methods and shows promising potential for operational deployment as a decision-support tool for agricultural stakeholders. Finally, this work highlights the importance of a continuous improvement approach, based on integrating higher-resolution data and thorough field validation. These steps are essential to optimize the agronomic impact of the model and strengthen the resilience of agricultural systems in the face of increasing climate variability.

ACKNOWLEDGEMENT

This research did not receive any external funding. The corresponding author personally covered all related expenses.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data used and/or analyzed during this study are available from the corresponding author upon reasonable request. additionally, the source code associated with this work is hosted in a private github repository (papegu). interested researchers or students can contact the corresponding author to request access. Upon request, access will be granted by inviting the requester as a collaborator on the github repository, allowing them to download and use the code for their own research or study purposes.

AUTHOR CONTRIBUTIONS STATEMENT

The corresponding author, pape elhadji abdoulaye gueye, conducted all aspects of the research, from study design to manuscript writing. the co-authors, provided guidance, oversight, and final approval of the work.

REFERENCES

[1] Muhammad Rasool Al-Kilani, Jawad Al-Bakri, Michel Rahbeh, Cody Knutson, Tsegaye Tadesse, and Qasem Abdelal. Agricultural drought assessment in data-limited arid regions using opensource remotely sensed data: A case study from Jordan. 156(2):89.

[2] Pape ElHadji Abdoulaye Gueye, Cherif Bachir Deme, and Adrien Basse. Improved climate risk assessment tool: Integration of satellite data and machine learning models for climate prediction on farmland around the senegal river. Mathematical Modelling of Engineering Problems., 12(5):1513–1523, May 2025.

[3] Mouhamadou Sylla, Michel Nikiema, Peter Gibba, Ibourahima Kebe, and Nana Ama Browne Klutse. Climate Change over West Africa: Re cent Trends and Future Projections. pages 25–40, apr 2016.

[4] Nesrine Farhani. Contribution of Remote Sensing and Auxiliary Vari ables in the Study of the Evolution of Drought Periods. PhD thesis, Université Paul Sabatier- Toulouse III and Université de Carthage (Tunisia), February 2022.

[5] Faye Dioumacor, F Kaly, Abdou Lahat Dieng, Dahirou Wane, Cheikh Modou Noreyni Fall, and Amadou Gaye. Prediction of the Onset and Offset of the Rainy Season in Senegal Using Kohonen Self-Organizing Maps.

[6] Trisha Bhaga, Munyaradzi Shekede, and Cletah Shoko. Impacts of Climate Variability and Drought on Surface Water Resources in Sub Saharan Africa Using Remote Sensing: A Review. 12:4184.

[7] Kassahun Tenebo Alito, Mulu Sewinet Kerebih, and Dawit Asregedew Hailu. Characterization of Drought Detection With Remote Sensing Based Multiple Indices and SPEI in Northeastern Ethiopian Highland. 18:11786221251328833.

[8] Zhaoxu Zhang, Wei Xu, Zhenwei Shi, and Qiming Qin. Establishment of a comprehensive drought monitoring index based on multisource re mote sensing data and agricultural drought monitoring. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14:2113–2126, 2021.

[9] Hao Chen, Ni Yang, Xuanhua Song, Chunhua Lu, Menglan Lu, Tan Chen, and Shulin Deng. A novel agricultural drought index based on 15 multi-source remote sensing data and interpretable machine learning. 308:109303.

[10] Zakari Seybou Abdourahamane, Issa Garba, Aboubakr Gambo Boukary, and Alisher Mirzabaev. Spatiotemporal characterization of agricultural drought in the Sahel region using a composite drought in dex. 204:104789.

[11] Alzira Gabrielle Soares Saraiva Souza, Alfredo Ribeiro Neto, and Laio Lucas de Souza. Soil moisture-based index for agricultural drought assessment: SMADI application in pernambuco state-brazil. 252:112124, 2021.

[12] Rajkumar Guria, Manoranjan Mishra, Richarde Marques da Silva, and Carlos Antonio Costa dos Santos. Multisensor Integrated Drought Sever ity Index (IDSI) for assessing agricultural drought in Odisha, India. 37:101399.

[13] Mikhael G. Alemu and Fasikaw A. Zimale. Integration of remote sensing and machine learning algorithm for agricultural drought early warning over Genale Dawa river basin, Ethiopia. 197(3):243.

[14] Seyni Salack, Koufanou Hien, Namo K. Z. Lawson, Inoussa Abdou Saley, Jean-Emmanuel Paturel, and Moussa Waongo. Prévisibilité des faux départs de saison agricole au sahel. In Risques climatiques et agriculture en Afrique de l'Ouest, pages 31–43. IRD Éditions, 2025. Consulté en août 2025.

[15] 500daysofAI. Prédire les inondations et la sécheresse avec l'ia, 2025. Synthèse d'un article NeurIPS 2019 sur les LSTM appliqués aux bassins versants via le jeu de données CAMELS.

[16] Hanen Balti, Raja Inoubli, Ali Ben Abbes, and Riadh Farah. Prédiction de la sécheresse à partir des images satellitaires : une approche à base d'apprentissage profond. In Actes de la conférence TAIMA 2022, 2022.

[17] Simple Science. Exploiter l'apprentissage automatique pour prédire les pluies saisonnières en afrique de l'est, 2025. Consulté en août 2025. Résumé d'une étude sur l'utilisation de LASSO et Elastic Net pour la prévision pluviométrique saisonnière en Afrique de l'Est.

[18] Julian Agudelo, Vincent Guigue, Cristina Manfredotti, and Hadrien Piot. Prévision de sécheresse en utilisant une architecture neuronale 16 hybride intégrant des séries temporelles et des données statiques. In Conférence hébergée de l'AFIA, 2025.

[19] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. Neural Comput., 9(8):1735–1780, nov 1997. [20] Krzysztof Zarzycki and Maciej Ławryńczuk. Advanced predictive control for gru and lstm networks. Information Sciences, 616:229–254, 2022.