

RESEARCH

Open Access



Prediction of maize yield in Uganda using CNN-LSTM architecture on a multimodal climate and remote sensing dataset

Danison Taremwa^{1,5*}, Emmanuel Ahishakiye², Aggrey Obbo³, Paul Kategaya Kisozi⁴ and Fred Kagwa⁵

*Correspondence:

Danison Taremwa
taremwa.danison@gmail.com
¹Department of Computer Science,
Kyambogo University, Kampala,
Uganda

²Department of Networks, Data
Science and Artificial Intelligence,
Kyambogo University, Kampala,
Uganda

³Department of Software
Engineering, Mbarara University
of Science & Technology, Mbarara,
Uganda

⁴Department of Environmental
Science, Kyambogo University,
Kampala, Uganda

⁵Department of Computer Science,
Mbarara University of Science &
Technology, Mbarara, Uganda

Abstract

Maize is a staple crop in Uganda, underpinning both food security and rural livelihoods. Accurate forecasting of maize yields is therefore crucial for guiding agricultural planning, resource allocation, and policy design. Yet traditional statistical methods are often limited by low accuracy, poor scalability, and weak integration of diverse inputs, leaving them unable to capture complex, nonlinear, and spatiotemporal dynamics of crop growth. To overcome these constraints, we developed a hybrid convolutional neural network and long short-term memory (CNN-LSTM) model. This model integrates remotely sensed climatic variables and vegetation indices with biannual maize yield records from Uganda's Zonal Agricultural Research and Development Institute (ZARDI) zones for the period 2018–2020. Due to the scarcity of high-quality yield data, we applied the Synthetic Minority Oversampling Technique for Regression (SMOGR) alongside feature selection to balance the dataset and improve predictive robustness. The CNN-LSTM model's ability to select features and perform extensive hyperparameter tuning enabled it to outperform baseline models. It achieved a Mean Squared Error (MSE) of 0.107 tonnes², a Mean Absolute Error (MAE) of 0.267 tonnes, a Root Mean Squared Error (RMSE) of 0.327 tonnes, and an R^2 score of 0.783. A comparative analysis revealed that the CNN+Random Forest (RF) model achieved an MSE of 0.137 tonnes², a MAE of 0.281 tonnes, an RMSE of 0.370 tonnes, and an R^2 score of 0.722. These results outperformed the standalone CNN (MSE=0.216, R^2 =0.562) and RF (MSE=0.211, R^2 =0.573) models, underscoring the advantage of combining spatial-temporal learning for improved predictive accuracy. Residual analysis further confirmed the model's stability, showing minimal bias and close agreement between observed and predicted yields. These findings highlight the potential for integrating spatial-temporal deep learning and ensemble methods to deliver accurate crop yield forecasts in data-limited smallholder systems. By offering a scalable framework for evidence-based farm planning and food security policy, our study demonstrated that advanced machine learning can directly support sustainable development in sub-Saharan Africa. Future research will extend the framework to incorporate Transformer architectures, high-resolution satellite imagery, and explainable AI, further enhancing accuracy, interpretability, and decision-support capacity.



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Article highlights

- Developed a hybrid CNN-LSTM model that integrates remotely sensed climatic and vegetation indices to predict maize yields across Uganda's ZARDI zones (2018–2020).
- Achieved high predictive accuracy (MSE=0.107, explaining 78% of yield variation), outperforming standalone CNN, ensemble models such as RF, and CNN-RF.
- Introduced SMOGN-based data augmentation and feature selection techniques to overcome data sparsity, a novel approach for yield forecasting in smallholder, data-limited systems.
- Demonstrated that hybrid DL frameworks can inform scalable, data-driven agricultural planning, with potential to guide policymakers and strengthen food security strategies in SSA.
- Future work will focus on integrating Transformer architectures for improved sequence modelling alongside high-resolution imagery and explainable AI to enhance accuracy, interpretability, and practical decision support.

Keywords Maize yield prediction, Ensemble learning, Precision agriculture, CNN-LSTM, Vegetation indices

1 Introduction

Crop yield prediction (CYP) provides an estimate of yield per unit area [1]. Yield prediction is a crucial strategy for farmers and the agricultural sector to efficiently manage resources and make informed decisions during crop growth and harvesting [2]. Being informed of the yield before harvest aids resource planning, such as determining the optimal fertilizer use to achieve a high yield, scheduling labour, and preparing for packaging and manufacturing [3]. For instance, breweries rely on timely maize yield estimates to adjust procurement and production, demonstrating the industrial importance of accurate forecasts [4]. In Sub-Saharan African countries (SSA), such as Uganda, over 60% of the population relies on maize for food and as a source of household income. Therefore, timely yield information is vital to safeguard food security and strengthen rural resilience [5]. However, yield forecasts are often unavailable before harvest, leading to price volatility that heightens economic uncertainty for farmers and policymakers [6]. These vulnerabilities are likely to worsen as climate variability increases and food demand rises, underscoring the need for accurate, flexible, and scalable yield forecasting systems as essential tools for mitigating risk and enhancing resilience [7].

Yield prediction remains challenging because crop productivity depends on complex nonlinear interactions among genotype, environment, and management factors that vary spatially and temporally [8]. Traditionally, estimating maize yield in Uganda has relied on farmers' intuition, statistical regression methods, and crop simulation models [9, 10]. While informative and interpretable, these approaches are subjective, complex to transfer, and limited in their ability to capture these nonlinearities [11]. This may result in inaccuracies in maize yield information, which, in turn, could misinform decision-making along the supply chain and ultimately exacerbate food shortages [12]. Recent research emphasizes the need for improved accuracy, higher spatial resolution, and efficient algorithms that can learn from large, multimodal datasets to capture yield dynamics more reliably [13]. Satellite remote sensing and ground-based approaches have been found to complement each other in yield prediction. Together, they provide a scalable method for deriving timely and accessible observations of vegetation and climate over

large areas [9, 14]. However, extracting and analysing actionable insights from high-dimensional imagery remains technically challenging for farmers and agricultural professionals [15].

Globally, machine learning (ML) and deep learning (DL) techniques have advanced CYP by uncovering complex, nonlinear patterns in remotely sensed data, resulting in accurate outputs [16, 17]. Machine learning models, such as Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost), have shown strong performance but often rely on hand-engineered features. They can be prone to overfitting and may face efficiency and scalability issues as feature dimensionality and data complexity increase [18, 19]. These challenges become especially evident in agricultural applications, where crop growth dynamics are naturally represented as a time series. Each time point represents the crop's physiological status at a specific developmental stage across diverse environments. Hence, the spatial-temporal nature of crop development necessitates advanced DL architectures, such as CNNs and LSTM networks, which can jointly capture sequential dependencies and spatial variability, thereby enhancing predictive performance [20]. The CNN component extracts localized vegetation patterns from satellite images, reflecting spatial variation, while the LSTM network learns long-term temporal dependencies across the growing season to support phenological analysis [21]. This synergy enables more accurate modeling of spatial-temporal crop yield dynamics, enhancing the accuracy of CYP [17].

Several studies have demonstrated this integrated approach, particularly in maize yield prediction [16, 17, 22]. For instance, studies [23] and [22] applied a CNN-LSTM model to predict maize yields in the US and China. Study [22] utilized high-resolution Unmanned Aerial Vehicle (UAV)-based multispectral and Light Detection and Ranging (LIDAR) data, with an attention mechanism that enhanced both interpretability and model robustness. In contrast, [23] employed MODIS imagery, meteorological time series, and soil attributes. Both studies achieved precise results, with R^2 values of 0.73 and 0.78, respectively, demonstrating the effectiveness of the CNN-LSTM architecture in capturing complex spatiotemporal and nonlinear interactions across different agro-ecological contexts.

Together, these outcomes provide compelling evidence for the broader applicability of ensemble and hybrid DL models, which have already been successfully deployed in CYP [21, 24], underscoring the power of crop forecasting. While these successes motivate their application to the maize yield problem in Uganda, the use of such models utilizing remotely sensed data in smallholder intercropped farms in Uganda, where datasets are limited and spatial variability is high, remains largely underexplored. Moreover, models trained in high-input, homogeneous farming systems, such as the US Corn Belt, where abundant, reliable training data are available, often exhibit reduced performance in diverse agro-ecological zones, such as those found in Uganda. In these settings, domain shifts markedly diminish predictive accuracy, underscoring the need for model adaptation. This limitation is further exacerbated by the scarcity of high-quality long-term maize yield records across ZARDI zones [25, 26]. While acquiring large amounts of vegetation indices (VIs), soil characteristics, and climatic data can be relatively easy, collecting yield data in developing countries is costly and sporadic, relying heavily on smallholder farmers and typically producing only short-term time series for model development.

For example, datasets collected from the Uganda Bureau of Statistics (UBOS) between 2018 and 2020 were limited, incomplete, and spatially fragmented, resulting in imbalanced data that would constrain model training and evaluation [27, 28]. Such limitations hinder the learning process of DL models, often leading to biased predictions and reduced reliability. These challenges underscore the need for DL models that are locally calibrated, scalable, and accurate for maize yield prediction in Uganda's data-scarce, intercropped farming systems. In this study, we leverage the only consistent yield observation available and address this limitation by augmenting it with SMOGN, a regression-aware method that balances continuous outcomes. This approach enhances data diversity, mitigates class imbalance, and improves model robustness without distorting the underlying distributions. Consequently, the combined use of synthetic and original data broadens the training sample space, enhancing the efficiency, reliability, and generalizability of DL models across ZARDI zones with heterogeneous spatial and yield characteristics [29, 30].

To ensure robustness across dataset sizes, we deliberately employed a compact architecture, using a 1D-CNN and a single LSTM layer. The CNN extracts localized patterns from feature vectors at each time step, serving as adaptive feature detectors well-suited to tabular input. At the same time, the LSTM model captures temporal dependencies across seasons. This hybrid CNN-LSTM model was strengthened through synthetic data augmentation, significantly improving maize yield prediction in Uganda and providing a scalable framework for evidence-based decision-making in smallholders' agricultural systems. Hence, our specific objectives are: (i) to integrate multi-source data to compensate for the limited labels, (ii) apply SMOGN to mitigate imbalance in yield distribution, and (iii) develop and validate a CNN-LSTM model for maize yield prediction in Uganda. Accordingly, the study's contributions can be highlighted in the three aspects below:

- i. Multi-source data integration: we harmonized satellite VIs, climate variables, soil properties, and zonal yield records, covering Uganda's nine ZARDIs (2018–2020), creating a comprehensive dataset for model development.
- ii. Applied SMOGN data augmentation methods to overcome sparsity, incompleteness, and imbalances in maize yield datasets, reducing skewness while preserving data distributions and improving predictive accuracy across diverse ZARDI zones.
- iii. Developed and evaluated a CNN-LSTM model for data-scarce contexts, integrating synthetic augmentation with remote sensing and environmental data to capture intricate nonlinear spatial–temporal dependencies. The model, with hyperparameters optimized via grid search and cross-validation, effectively learns spatial and temporal patterns and outperforms CNN and CNN-Random Forest.

Together, these contributions deliver a resilient DL model for maize yield forecasting in Uganda's smallholder intercropped systems. The approach provides both methodological innovation and offers practical decision-support value for data-scarce agricultural regions, with strong potential for extension to other crops and diverse agro-ecological contexts across SSA.

2 Related literature

2.1 Transformative impact of artificial intelligence and remote sensing in agriculture

Artificial intelligence (AI) and remote sensing are transforming agriculture by providing scalable, timely, and cost-effective methods for monitoring crop growth under the dual pressures of climate change and rising food demand [31]. Although adoption has advanced rapidly in developed countries, implementation across the SSA remains limited, particularly among small-scale farmers who contribute 70% of the production. Enhancing crop production, monitoring diseases, managing fertilizer and irrigation, and optimizing harvesting are all challenging tasks that contribute to improving food security [3]. To meet these demands, remote sensing platforms, ranging from ground-based sensors (e.g., Internet of Things devices) to UAVs and satellites, provide vital information on crop growth, soil health, and climate variability [32]. Among these, satellite-based multispectral sensors such as Sentinel-2, Landsat, and MODIS are particularly valuable for large-scale monitoring, as they offer high resolution and continuous coverage at relatively low cost [33]. From these sensors, VIs including NDVI, EVI, LAI, NDWI, and CCI serve as robust indicators of crop vigor, water status, and chlorophyll content. At the same time, climatic and soil variables capture complementary environmental factors that influence yield. Integrating these multimodal datasets has consistently been demonstrated to improve yield prediction accuracy, as each source offers the limitations of the others [34, 35]. Consequently, predicting future crop yields from diverse data sources is a complex task that requires hybrid models capable of capturing nonlinear spatial patterns and temporal dependencies, thereby enhancing prediction accuracy [23].

2.2 Deep learning and multisource datasets for crop yield prediction

Various studies have used multimodal datasets to achieve accurate, reliable results by combining multiple sources of information. For example, a study by [36] developed a Bayesian optimization-based LSTM (BO-LSTM) model to predict winter wheat yields, which outperformed conventional ML models such as SVM and LASSO, achieving an RMSE of 177.84 kg/ha and an R^2 of 0.82. Their comparative experiments demonstrated that using single inputs, such as GPP ($R^2=0.72$, RMSE = 186.13 kg/ha), LAI ($R^2=0.67$, RMSE = 221.32 kg/ha), and VIs ($R^2=0.78$, RMSE = 190.96 kg/ha), yielded lower performance. Combining meteorological data with GPP significantly improved performance ($R^2=0.81$; RMSE = 180.66 kg/ha). This highlights the strength of multimodal fusion in capturing crop-environmental interactions.

While multimodal integration offers the highest accuracy, temporal models also show strong potential. Study [20] demonstrated that even small sets of NDVI time-series inputs can perform well when trained with LSTM, achieving an RMSE of 505.78 kg/ha (NRMSE = 0.0726). This underscores LSTM's strength in learning growth dynamics directly from sequential data, maintaining reasonable accuracy even under limited input conditions. Building on this, [23] developed a multilevel CNN-LSTM model that incorporated time-series remote sensing data and soil properties to predict corn yield across the U.S. Corn Belt states (2013 to 2016). The hybrid model achieved an RMSE of 1010.61, a MAPE of 7.97%, and an R^2 of 0.75, outperforming both traditional ML and standalone DL models, including the Deep Neural Network (DNN) model (RMSE of 1130.44, MAPE of 9.14%, and R^2 of 0.68). However, they observed that improvements

from additional modalities plateaued, with phenological features proving more influential than climate variables.

More studies have focused on an ensemble learning approach with DL models. For example, [21] developed a stacked CNN-DNN ensemble with LASSO meta-learning, achieving the best performance across seasons in the U.S. Corn Belt (RMSE = 874 kg/ha, $R^2 = 83.1\%$ in 2019, and RMSE = 999 kg/ha, $R^2 = 74.6\%$ in 2020). By comparison, the baseline CNN-RNN model without ensemble improvements exhibited high error rates (RMSE of 1,007 kg/ha in 2019 and 1,092 kg/ha in 2020). Similarly, [37] employed a concatenated-based 2D-CNN-BILSTM that fused Sentinel-1, Sentinel-2, and soil grids to predict corn yield in Iowa State (2018 to 2021). Their model achieved an RMSE of 0.698 tonnes per hectare, a MAPE of 4.4%, and an index of agreement (D) value of 84.67%, clearly surpassing the baseline model, 2D-CNN, which had a D of 14.77%. Both studies demonstrate the advantages of ensemble-hybrid pipelines in improving robustness and reliability. Finally, expanding the geographical focus, [17] employed a CNN-LSTM-Attention model to predict yields of maize, rice, and soybeans in Northeast China. The model achieved an R^2 of 0.80, an RMSE of 375.08, and a MAPE of 4.21%, outperforming both CNN (MAPE of 4.30%, R^2 of 0.77, and RMSE of 394.67) and LSTM. The inclusion of attention mechanisms enabled the extraction of higher-order features, addressing the complexity of high-dimensional agricultural data across time and space.

Taken together, these studies demonstrate that crop yield estimation has primarily been conducted in developed countries, where DL models such as CNNs and LSTMs have been applied to remote sensing data, achieving high accuracy [23, 36]. In these contexts, the accuracy, efficiency, and long-term usefulness of maize yield information have significantly improved, reflecting the advantages of homogeneous farming systems and well-documented agricultural practices. However, these conditions are difficult to replicate in the widely heterogeneous cropping patterns of Africa, which are characterized by smallholder-dominated systems [5]. Moreover, DL models often exhibit limited spatial transferability and remain highly location-specific due to domain shifts across diverse environments [25]. A further challenge DL faces is the need for extensive training datasets, which are difficult to acquire in developing countries, thereby constraining models' generalization in small data domains [37]. While remotely sensed data is widely available for maize yield prediction, access to high-quality ground-truth yield records remains limited in many developing countries due to the costs and time-intensive nature of the process involved.

To overcome these constraints, researchers have explored strategies such as a dimensionality reduction technique [23] and the generation of synthetic data to augment sparse records [30], thereby improving model robustness and reducing overfitting. These approaches have demonstrated promise in enhancing the reliability of training processes under data-limited conditions. However, there remains an urgent need for research in Africa, particularly in SSA, such as Uganda, that leverages remotely sensed data and DL models to develop scalable and context-specific maize yield prediction systems. The following section discusses the materials and methods for model development.

3 Materials and methods

3.1 Dataset description

Building on prior evidence that multimodal integration enhances spatial–temporal yield estimation [23, 36, 37], this study employs a dataset combining vegetation indices with soil and climatic data and historical maize records across nine ZARDI zones. Remote sensing features were obtained from publicly available sources, including MODIS imagery (frequent 16-day composite data at resolutions of 250–500 m) and Sentinel-2 (10 m). Climatic attributes collected from NASA POWER were first aggregated into daily data, then into seasonal values, aligned with Uganda’s two annual rain-fed maize production cycles (March–July and September–December), from 2018 to 2020 [38]. The Vegetation indices were derived from MODIS products (e.g., MOD13Q1 for NDVI/EVI/NDWI at 250 m resolution), LAI from MOD15A2H (500 m), CCI from Copernicus Sentinel-2 (10 m), and a cropland mask. Rainfall data were obtained from NASA GPM-IMERG (daily precipitation, aggregated by season). In contrast, solar radiation and soil moisture data came from NASA POWER/ERA5, and maximum and minimum temperature data were derived from NASA POWER.

Both NDVI and EVI capture crop vigor by reflecting biomass expansion and vegetation health. In parallel, LAI quantifies canopy structure, linking light interception, photosynthesis, and transpiration to biomass accumulation and yield formation. NDWI adds a water dimension by detecting canopy moisture and drought stress. At the same time, CCI reflects chlorophyll content and photosynthetic activity, both of which are directly tied to grain set and productivity. Management practices, such as fertilizer application, enhance canopy vigor, leaf area, and chlorophyll concentration, while pest and disease pressures reduce greenness and water content, leading to declines in NDVI, EVI, CCI, and NDWI [17, 39]. Together, these indices provide complementary insights into crop conditions. When integrated with climatic variables, such as rainfall, solar radiation, maximum and minimum temperatures, and soil moisture, they capture the key biophysical drivers of crop performance, regulating growth, phenology, and water availability. This integrated perspective forms a robust foundation for reliable yield prediction [10].

The maize yield datasets (measured in tons per hectare) used in this study were sourced from the Uganda Bureau of Statistics (UBOS¹). UBOS utilized standardized procedures through the Annual Agricultural Survey (AAS), thereby ensuring the data is nationally representative and quality-assured across all locations of the ZARDI. The use of UBOS, the national custodian of agricultural data, adds credibility, precision, and reliability to the empirically tested maize yield data used herein. These data were collected as part of the 50 × 2030 Initiative, an international effort led by the FAO and the World Bank to address agricultural data gaps in 50 countries by 2030. While the AAS provided standardized, nationally representative data, financial and logistical constraints prevented further follow-up localized surveys within the ZARDI zones after 2020 [27]. All 13 predictor variables, including Year, Rainfall, Solar Radiation, Max-Temp, Min-Temp, Humidity, Soil Moisture, NDVI, EVI, NDWI, CCI, LAI, and Cropland Fraction, were used in the study. These geographically diverse variables, spanning all ZARDI zones, enabled both spatial and temporal analyses of maize yield trends. The details of the data sources and variables employed are listed in Table 1.

¹ <https://www.ubos.org/>

Table 1 Sources of data

Data Type	Variables	Temporal Resolution	Spatial Resolution	Data Source	Link
Weather data	Temperature, solar irradiance, and humidity	Daily (varies)	~0.5°×0.5° (~50 km)	NASA POWER	https://power.larc.nasa.gov/
Rainfall	Precipitation (IMERG)	2014–present, daily	0.1°×0.1° (~10 km)	NASA GPM (IMERG)	https://pmm.nasa.gov/resources/documents/gpm-integrated-multi-satellite-retrievals-gpm-imerg-algorithm-theoretical-basis-https://ceres.larc.nasa.gov/data/
Solar radiation	All-Sky Surface Photosynthetically Active Radiation (PAR)	Daily, near real-time	~1° (CERES SYN1deg/FLASHFlux)	NASA CERES & FLASHFlux via POWER	https://ceres.larc.nasa.gov/data/
Soil moisture	Root zone soil wetness (0–100 cm depth)	2014–present, daily	~0.5° (~50 km)	NASA GMAO, MERRA-2 (GEOS DAS)	https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/
VIs	NDVI, EVI (MOD13Q1 V6.1); CCI (Copernicus S2 Harmonized)	2018–2020 (seasonal)	250 m (NDVI/EVI), 10 m (CCI)	MODIS (EOSDIS), Copernicus Sentinel-2	https://lpdaac.usgs.gov/products/mod13q1v061/ https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED
Photosynthetically active indices	LAI (Leaf Area Index, MOD15A2H V6.1)	2014–2023, 8-day	500 m	MODIS	https://doi.org/10.5067/MODIS/MOD15A2H.061
Land cover (masking)	Cropland areas (≥60% cultivated cropland, Band 12)	Annual	500 m	MODIS MCD12Q1 V6.1	https://lpdaac.usgs.gov/products/mcd12q1v061/

3.2 Data pre-processing

Data pre-processing was one of the most critical steps in preparing this data for efficient, reliable predictive modeling of maize yields. Processing of predictors was first conducted using Google Earth Engine (GEE), where variables were sampled at 500 m intervals, and VIs and environmental variables were standardized to monthly intervals. Within each ZARDI zone, a satellite-derived annual cropland mask derived from MODIS MCD12Q1 for each specific year from 2018 to 2020 was applied to delineate active cropland pixels. Within these masked cropland areas, sampling was further restricted to grid cells associated with maize production, as per UBOS Annual Agricultural Survey records and ZARDI agronomic reports. Vegetation indices (NDVI, EVI, LAI, NDWI, CCI) and environmental variables (rainfall, temperature, soil moisture, radiation) were then sampled at 500-m intervals only within these maize-specific pixels, and the resulting values were aggregated to seasonal means per ZARDI zone, aligning with Uganda's two annual maize seasons. This spatial filtering ensured that the aggregated inputs represented maize-specific conditions in each zone. To improve data reliability, the zonal-level maize yield records from AAS were first cross-checked against independent regional statistics for

consistency. Anomalies were corrected during data cleaning to provide a reliable ground truth, essential for model training. Subsequently, categorical variables, such as ZARDI zones, were encoded using one-hot encoding to ensure ML compatibility without introducing false ordinal relationships. At the same time, temporal variables (e.g., year) were converted to numerical values to retain sequential order by extracting the year component, thereby ensuring compatibility with predictive algorithms. Continuous features, such as NDVI, EVI, LAI, NDWI, rainfall, solar radiation, soil moisture, and temperature, were standardized using the StandardScaler to achieve uniform scaling and minimize potential bias arising from differences in magnitude across variables. They were then placed on a uniform scale with a mean of zero and a variance of one. This step eliminated disparities in variable magnitudes and prevented model bias toward features with larger numerical ranges. To address class imbalance, synthetic samples were generated using SMOGN [29], which expanded the dataset while preserving feature-target relationships, thereby reducing bias from underrepresented yield ranges.

Missing values, if present, were imputed with median values in numerical columns to minimize information loss and ensure robustness against outliers. Furthermore, interaction terms were engineered, such as those between climatic variables (e.g., NDVI and rainfall), to capture nonlinear relationships and enhance expressiveness in the feature space. To reduce dimensionality, a RF model was employed to rank features by importance. We retained the top 10 features (out of the initial 15) based on this ranking, including variables such as YEAR, MAX_TEMP, NDVI, and rainfall, which collectively accounted for the most variance in yield. Removing less influential features helped minimize noise and multicollinearity, ultimately improving model training efficiency. The final dataset was split into training and test sets at an 80:20 ratio to ensure good generalization and unbiased model performance. This split strategy, widely used in yield prediction studies, strikes a balance between the need for sufficient training data and the requirement for reliable test evaluation. Within the training set, we performed five-fold cross-validation to optimize model parameters and assess variability. This approach maximizes the use of the limited while guarding against overfitting. The resulting pre-processed dataset, enriched with synthesized data and feature engineering, captured spatial, temporal, and environmental variability. This robust foundation supports the development of cutting-edge ML models with high predictive accuracy for maize yields across diverse agro-ecological regions in Uganda. The section that follows describes SMOGN oversampling as a crucial pre-processing step to improve the model's performance.

3.2.1 SMOGN oversampling for imbalanced yield regression data

The lack of observations at the extremes of yield, such as very low or very high yields, in the target variable hampers the model's ability to learn those conditions [29]. SMOGN was applied as a pre-processing step to balance the yield distribution. It generates new samples in under-represented regions by interpolating between minority examples and their nearest neighbours, then adding Gaussian noise to increase variability. SMOGN also performs a mild under-sampling of the class to avoid biasing the model towards the middle of the distribution. We preferred SMOGN over simpler oversampling methods, such as SMOTER alone, because it better preserves continuous target relationships and has demonstrated superior performance on small, imbalanced regression datasets [40].

The process involves inputting the original dataset, applying SMOGN to create a balanced dataset, and then training the model on this new balanced data [29]. The pseudocode for the SMOGN technique is shown in Algorithm 1 [41].

Algorithm 1: Pseudocode for SMOGN.

```
1. Algorithm 1. Pseudocode for SMOGN.
2. Initialize SMOGN parameters (oversampling ratio, Gaussian noise parameters, k for k-
   nearest neighbours).
3. For iteration = 1 to the number of iterations, do
4.   Identify the minority samples in the training data based on the target variable.
5.   For each minority sample, do
6.     Use k-nearest neighbors to find the closest sample to the minority sample.
7.     Randomly select one of the k-nearest neighbours.
8.     Generate a synthetic sample by interpolating between the minority sample and
       the selected neighbour.
9.     Add Gaussian noise to the synthetic sample to make it more realistic.
10.  End For
11.  Combine the original training data with the synthetic samples to create a balanced
      training dataset.
12. End For
13. Return the balanced training dataset.
```

3.2.2 Application and limitations of SMOGN oversampling

After SMOGN, the training set increased to 295 samples, ensuring that extremely high and low yields were adequately represented. We verified that the augmented yields remained within realistic bounds, without introducing implausible extreme values, and oversampling preferentially added samples to previously sparse yield ranges, such as those below 1.0 t/ha and above 2.5 t/ha, resulting in a more uniform distribution. The synthetic data were generated with careful parameter tuning, using a neighbour count of $k=5$, a Gaussian noise level of 0.01, and 100% oversampling of minority instances, following the guidelines of Branco et al. [29]. To reduce the risk of unrealistic yield values being introduced during augmentation, we constrained the generated values within the biologically plausible range of maize yields observed in Uganda (0.5–4.5 t/ha). This ensured that synthetic samples remained consistent with known agronomic conditions. Additionally, we visually examined yield distribution plots before and after augmentation. The post-SMOGN distribution was more uniform, while still following the natural trends of the original dataset, thereby validating that no extreme outliers were created. As a result of this pre-processing step, the dataset's imbalance was significantly reduced while maintaining a realistic feature-target relationships approach validated in recent studies [41, 42]. Despite these improvements, SMOGN remains a heuristic method and may still introduce some artefactual values in edge cases; therefore, careful post-hoc validation and visualization remain essential.

Its assumption of local smoothness further limits the ability to capture complex, non-linear, high-dimensional interactions in a multimodal dataset [43]. As a result, while suitable for moderately structured data, SMOGN may underperform in contexts with discontinuities or latent heterogeneity, where hybrid approaches incorporating generative or manifold-based methods could offer more robust alternatives. Recent advances in deep generative models, particularly Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), have established themselves as the current state of the art for generating synthetic data for crop yield datasets. However, given our limited sample size, training a GAN or VAE reliably would be challenging. GANs typically require thousands of examples to capture the data distribution accurately and avoid model collapse [44].

3.3 The proposed model

Figure 1 illustrates the architecture of a CNN-LSTM model for maize yield prediction using multimodal climate and remote-sensing data. The process begins with multiple input streams for each maize-growing season, including satellite-derived indices (e.g., EVI and NDVI) and temporal climate variables (e.g., temperature and rainfall) spanning the season. After pre-processing these inputs (X_t), they are passed to the 1D-CNN, which extracts spatial features from the sequence (e.g., identifying patterns in NDVI that correlate with crop biomass). The CNN processes the inputs, yielding a spatially enriched feature vector summarizing patterns at each time step. This sequence of feature vectors is then fed into the LSTM module, enabling the system to learn how these features evolve from one time step to the next. The LSTM component facilitates an in-depth understanding of seasonality, the complexity of growth stages, and time-dependent behaviors, such as the impact of early-season rainfall on mid-season vegetation health, which are crucial for crop prosperity [45]. The building of the LSTM cell operates via forget (f_t), input (i_t), and output gates (o_t), each regulated by the sigmoid activations (σ), which regulate the movement of data through its cell state (c_t). The candidate memory cell (g_t) uses a tanh activation to propose updates to the cell state. The gates determine which parts of the past to keep or abandon, how much new input to take in at any given time, and what is essential to save and pass on. The LSTM's output at each time point, t , is denoted as H_t , and the hidden state (h_t) encodes the learned temporal relationships, serving as a distilled representation of all relevant historical data leading up to the current time step [46]. The output H_t is then passed onto a fully connected layer, acting as an advanced spatiotemporal integrator. During this step, the aggregated feature vector is transformed into a form suitable for prediction. Finally, the created representation is passed to the yield estimation layer, which is implemented as a dense output layer with a linear activation. This last one does the regression task and produces the predicted maize yield, measured in tonnes per hectare, as the final output [21, 23, 24].

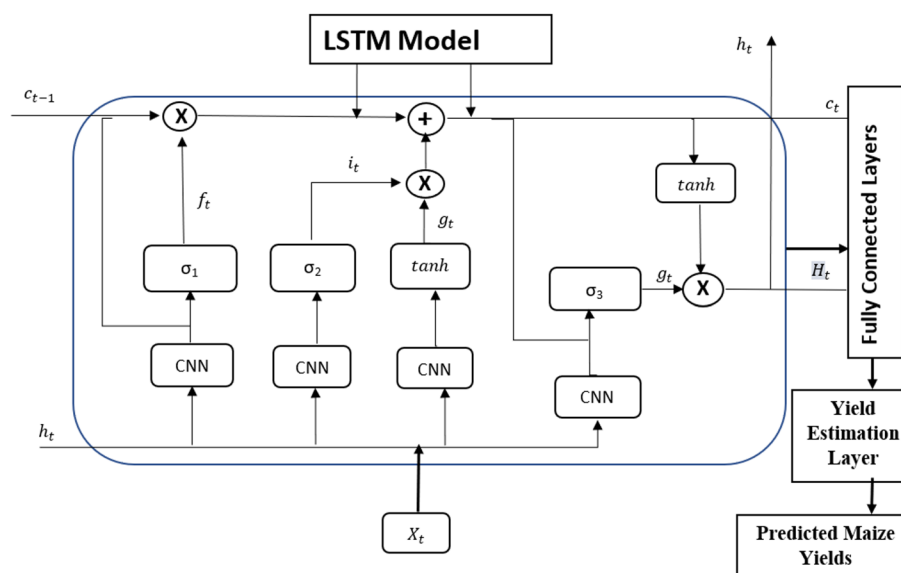


Fig. 1 The architecture of the proposed CNN-LSTM

3.3.1 Hyperparameter tuning strategy

This work used a suite of ML and DL models, each tuned via careful hyperparameter optimization, to evaluate the performance of various modeling paradigms for forecasting maize yield. The models were refined using a grid search and fivefold cross-validation on the training dataset, generated from 295 records using the SMOGN method. Independent, systematic hyperparameter optimization was performed separately for DL and baseline ML algorithms. The CNN-LSTM model was deliberately kept narrow (with only one CNN layer and one LSTM layer) to minimize the risk of overfitting on the limited data. A grid search was used to explore various combinations of convolutional filters (16, 32, 64), kernel sizes (3, 5, 7), LSTM units (32, 64, 128), learning rates (0.001, 0.0005), and batch sizes (16, 32). Model selection criteria were based on the mean validation loss across five-fold cross-validation, ensuring that the test set remained unseen during hyperparameter optimization. To prevent overfitting, early stopping (patience of 20 epochs) and dropout regularization (rate of 0.3) were implemented. The final architecture of the CNN-LSTM was intentionally kept narrow and straightforward, with just a single convolutional and a single recurrent layer, considering the dataset size. For the Random Forest algorithm, hyperparameters such as the number of estimators (100, 200, 500), the maximum depth of the tree (5, 10, 15), and the learning rate (0.05, 0.1, 0.2 for boosting) were tuned using grid searches with the assistance of five-fold cross-validation. This resulted in a balanced architecture that generalizes well while minimizing the risk of overfitting.

3.3.2 Model architectures and final configurations

RF was used as a baseline due to its robustness to high-dimensional data, fast training time, and ability to model nonlinear interactions between inputs. The optimal hyperparameter configuration consisted of 200 trees, a maximum depth of 20, and a minimum sample split of 2. The CNN model was used as another baseline. It comprised two 1-D convolutional layers with 32 and 64 filters, followed by a ReLU activation function for introducing non-linearity and a kernel size of 3. A max pooling layer followed this to reduce spatial dimensionality, and a dense regression output layer. Model training minimized mean squared error using the Adam optimizer with a learning rate of 0.001 and the MSE loss function for up to 50 epochs, with a batch size of 8. A hybrid CNN-RF model was also used as a comparative model. The CNN block consisted of one 1-dimensional convolutional layer with 64 filters of kernel size 3, followed by batch normalization, a ReLU activation, a max pooling layer, and a dropout layer with a rate of 0.3. The extracted feature maps from the CNN block were flattened and passed to an RF regressor with 100 decision trees. Each tree was trained on bootstrapped subsets of the data to enhance generalization. The final prediction was generated by averaging the outputs of all trees, allowing the model to perform nonlinear regression while maintaining interpretability and robustness to noise.

The proposed CNN-LSTM model was developed to learn the spatial features from climate inputs and the seasonal temporal dynamics of environment variables. The architecture began with a 1-D convolutional layer featuring 64 filters of kernel size 2, followed by a max pooling layer and a dropout layer with a dropout rate of 0.3 to reduce overfitting. The extracted spatial features were fed into an LSTM layer with 100 units, enabling the model to learn long-range sequential dependencies. The CNN and dense layers adopted

ReLU activation, while sigmoid and tanh activations were employed in the LSTM for gate operations and memory cell updates. A fully connected dense layer provided the final regression output with linear activation. The model was optimized using the Adam optimizer with a learning rate of 0.001 and MSE loss. Early stopping was implemented with a patience of 20 epochs to prevent overfitting, and a 20% validation split was applied. Optimization was performed with a batch size of 16 for up to 50 epochs (Table 2).

3.4 Performance metrics

The performance of the proposed model was evaluated using four key metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score. These metrics can be used comprehensively to assess the model's predictive accuracy. The MAE was chosen for its intuitive interpretation as the average magnitude of errors in the predictions, providing a simple measure of prediction accuracy that does not overweight larger errors [47, 48]. MSE was included because it squares the prediction errors, thereby emphasizing larger prediction errors and being especially useful for identifying those models that minimize significant deviations [18]. The MSE is derived into the RMSE, which presents error values in the same units as the target variable, maize yield, making them easily interpretable and practical [49]. Lastly, the R^2 score was selected to examine the proportion of variance in maize yield that the model explained. It provided a normalized measure of predictive power that accounted for dataset variability. It follows that these metrics provide a balanced assessment of the model's accuracy and reliability, thereby enhancing the robustness of the performance validation over Uganda's ZARDI zones. The MAE, MSE, RMSE, and R^2 Score are shown in Eqs. 1–4, respectively [50].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Table 2 Summary of optimal hyperparameters for each model

Model	Key Hyperparameters	Final Configuration
Random Forest (RF)	Number of estimators, maximum depth, minimum samples split	Estimators = 200; Max depth = 20; Min split = 2
CNN	Conv layers, filters, kernel size, activation, pooling, optimizer, training setup	2 Conv1D layers (32, 64 filters); Kernel size = 3; ReLU activation; Max pooling; Dense output; Adam optimizer (LR=0.001); MSE loss; 50 epochs; Batch size = 8
CNN–RF Hybrid	CNN block + RF regressor	Conv1D (64 filters, kernel = 3); Batch normalization; ReLU; Max pooling; Dropout = 0.3; Flatten → RF regressor (100 trees); Output = averaged predictions
CNN-LSTM	Conv filters, kernel size, LSTM units, dropout, activations, optimizer, training setup	Conv1D (64 filters, kernel = 2); Max pooling; Dropout = 0.3; LSTM (100 units); Dense output (linear); Adam optimizer (LR=0.001); MSE loss; 50 epochs; Batch size = 16; Early stopping (patience = 20)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (4)$$

where: y_i : Actual value for the i -th observation, \hat{y}_i : Predicted value for the i -th observation, \bar{y} : Mean of the actual values, n : Total number of observations.

3.5 Model implementation

The proposed model for predicting maize yield has been implemented in Google Colab, a cloud-based platform that enables easy collaboration in coding and powerful computing. The implementation of this model is written in Python 3.10; key libraries used include Pandas and NumPy for data manipulation and pre-processing, Scikit-learn for model training and evaluation, and Matplotlib for visualization. Pre-processing steps included scaling continuous features using StandardScaler to ensure consistent scaling across variables. The dataset was split 80–20 into training and test sets to ensure unbiased evaluation.

4 Experimental results

4.1 Performance of the proposed model

The performance metrics of the evaluated models provide insight into their predictive capability for maize yield, as shown in Table 3. The CNN-LSTM surpassed the CNN and Random Forest results in all comparisons. The R^2 of the CNN-LSTM was approximately 0.7833, and the RMSE was approximately 0.33 t/ha, with a MAE of roughly 0.27 t/ha. This indicated its effective ability to minimize prediction error and explained approximately 78.33% of the variance in the maize yield data, making it the most robust and accurate model among those tested. Notably, these results marked a substantial improvement over their counterparts from the CNN ($R^2 \approx 0.56$, RMSE ≈ 0.46 t/ha, MAE ≈ 0.37 t/ha) and the Random Forest ($R^2 \approx 0.57$, RMSE ≈ 0.46 t/ha, MAE ≈ 0.31 t/ha). The reference results from the CNN + RF ensemble were sensible ($R^2 \approx 0.72$, RMSE ≈ 0.37 t/ha, MAE ≈ 0.28 t/ha), demonstrating that combining spatial information extraction with a nonlinear regressor yields better results than using either model individually. This means that although Random Forest effectively captured the relationship in the data, it performed less accurately on its own than the ensemble and CNN-LSTM. Similarly, the CNN model may not adequately capture the full range of temporal dependencies in the data. The above findings also underscore the importance of combining complementary modeling techniques to improve predictive accuracy in agricultural yield forecasting. Notably, these performance gains were achieved without overfitting the data; the CNN-LSTM's hyperparameters were tuned via cross-validation, and its training was regularized with dropout and early stopping, ensuring that the model generalizes well to unseen data.

4.2 Comparison between the actual maize yields and the predictions of different models

Figure 2 presents the ground-truth real maize yields, represented by the blue line, against the predictions of different models: CNN (purple line), Random Forest (green line), CNN-LSTM (red line), and the Ensemble Model (CNN + Random Forest) (orange line). The performance of each model is analysed based on its ability to closely follow the trend in actual maize yields across different sample indices. The CNN-LSTM accurately

Table 3 Model performance results (calibration and validation)

Model	Calibration MSE	Calibration RMSE	Calibration R ²	Validation MSE	Validation MAE	Validation RMSE	Validation R ²
CNN	0.1985	0.4456	0.5894	0.2158	0.3735	0.4646	0.5622
CNN + Random Forest	0.1224	0.3499	0.7510	0.1368	0.2809	0.3699	0.7225
Random Forest	0.1942	0.4407	0.6018	0.2105	0.3054	0.4588	0.5730
Proposed CNN-LSTM	0.0956	0.3092	0.8125	0.1068	0.2667	0.3268	0.7833

tracked actual yields across the full range, including both high- and low-yield examples, whereas the CNN and RF often exhibited systematic biases. For instance, one of the highest observed yields at sample index 6 was 3.61 t/ha; the CNN-LSTM predicted approximately 3.13 t/ha, while the CNN underpredicted at around 2.99 t/ha, and the RF slightly overpredicted (~ 3.55 t/ha). In a low-yield case (actual 1.40 t/ha), CNN-LSTM's prediction was ~ 1.993 t/ha, compared to CNN's 2.291 t/ha and RF's 2.362 t/ha. These examples illustrate that CNN-LSTM errors (on the order of 0.1–0.5 t/ha) were minor and less biased than those of the other models at both extremes of the yield distribution. In general, the CNN tended to under-estimate peak yields, while the RF sometimes over-estimated lower yields; the hybrid CNN + RF ensemble mitigated some errors but still lagged CNN-LSTM's accuracy. It follows that the CNN-LSTM model is the most reliable predictor, as its trend closely follows the actual values of maize yield, particularly for both high and low yield extremes. This demonstrates its potential as a reliable model for accurately predicting maize yields over complex agricultural datasets.

4.3 Comparison of CNN, random forest, CNN-LSTM, and CNN-random forest

Figure 3 compares the four models, CNN, Random Forest, CNN-LSTM, and CNN-Random Forest, using MSE, MAE, RMSE, and R^2 Score to predict maize yield in tonnes per hectare. The CNN (with 1D convolutional layers) represented a naive DL approach that used the same input format; it achieved a moderate R^2 score of 0.56, indicating that temporal dependencies were not fully captured by the CNN alone. The RF represented a classic ML approach that performed similarly to a CNN, with an R^2 of 0.57, highlighting that a non-temporal model can capture some relationships but misses sequential effects. Building on these baselines, the CNN-Random Forest ensemble performed impressively, with an excellent MSE of 0.137, MAE of 0.281 tonnes, and RMSE of 0.370 tonnes, while achieving an R^2 score of 0.722. This validated the idea of integrating spatial feature extraction with a nonlinear regressor. This ensemble's success motivated the development of the combined CNN-LSTM model, which outperformed all other models with the lowest MSE of 0.107, MAE of 0.267 tonnes, and RMSE of 0.327 tonnes, with the highest R^2 score of 0.783. This reveals its strong ability to extract spatial and temporal patterns from yield data. It enables highly accurate maize yield predictions, for instance, 2.5 tonnes per hectare, with minimal deviations from the actual values. This proved to be an effective combination, leveraging the complementary strengths of both models while aligning with the characteristics of our dataset in our data-scarce setting. By contrast, single models such as CNNs and RFs showed higher errors and lower R^2 scores, suggesting their predictions are not particularly accurate.

These four models were selected as the baseline for predicting maize yield in Uganda. The chosen models are applicable and relevant to the available data. While gradient boosting algorithms, such as XGBoost and CatBoost, surpass RF on tabular datasets [45], they are not inherently designed for learning spatial–temporal sequences. Applying them in this study would have required extensive feature engineering to capture these dynamics, rendering them less practical in this context. We also recognize that more advanced architectures, such as CNN-Attention-LSTM, BiLSTM, 3D-CNN, and Transformer-based architectures, could further enrich the comparison given their ability to capture spatial and temporal dependencies [37, 51]. But they were excluded due to the limited size of the available training dataset and the substantial computational

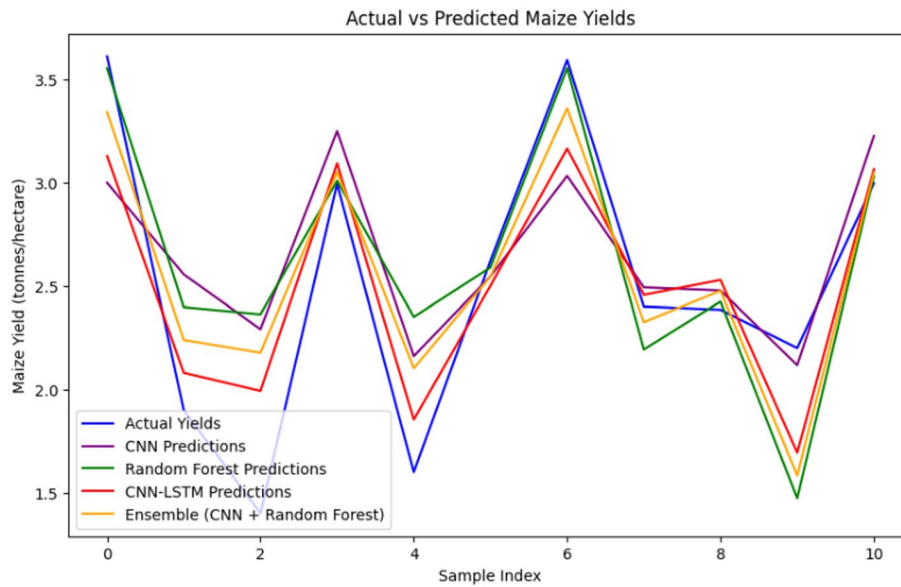


Fig. 2 Comparison between the actual maize yields and the predictions. The Sample index is an arbitrary sequential label assigned to each unique zone-season observation in the yield dataset (2018–2020). Each index corresponds to a single zone-season record, ordered chronologically

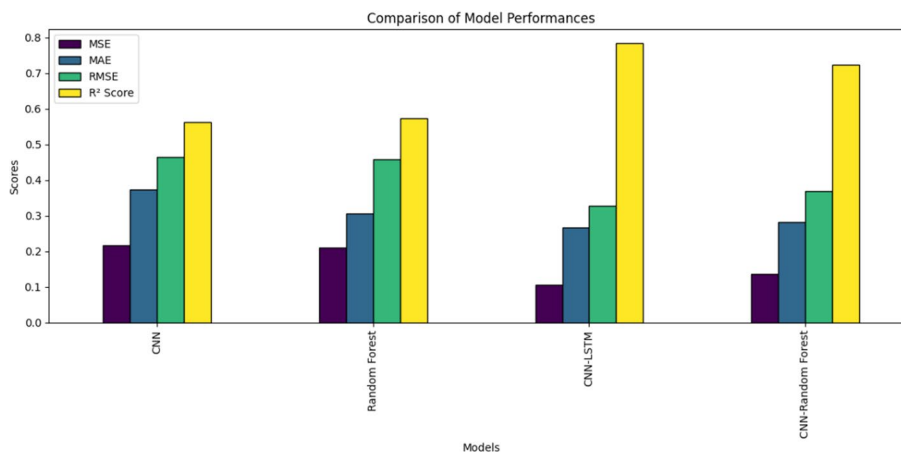


Fig. 3 The comparison of CNN, Random Forest, CNN-LSTM, and CNN-Random Forest

demands they entail. For instance, BiLSTM, which processes sequences both forward and backward, would roughly double the number of trainable parameters, increasing the risk of overfitting to our relatively limited time-series data. Likewise, a 3D-CNN was not employed, as it would require high-resolution spatial–temporal grids rather than our categorical zones and substantially larger datasets to support its parameter complexity. Similarly, transformer models have shown promise in capturing long-range dependencies in crop yield data; however, they typically require extensive training data and substantial computational resources [17, 52].

Consequently, the benchmarking was focused against top and realistic models rather than exhaustively examining all possible architectures. The models have been trained and evaluated using the same standardized data, which merged remotely sensed climatic and VIs with maize yield in Uganda's ZARDI zones for 2018–2020. The SMOGN technique was employed to address the natural imbalance in the yield data. The controlled

experiment setting adds weight to the conclusion and underscores the CNN-LSTM model's potential to handle spatiotemporal patterns. Overall, results indicate that the CNN-LSTM model provides the most accurate predictions, followed by the CNN-RF ensemble. Therefore, these models help improve maize yield forecasting in agricultural decision-making.

4.4 Predicted values of maize yield in tonnes per hectare for various models

Table 4 presents the predicted maize yields (tonnes per hectare) for the following models: CNN, CNN-LSTM (proposed), Random Forest, and CNN + Random Forest. These have been compared with the actual yield for each to assess their performance. The proposed CNN-LSTM produced the best performance in approximating actual values. For example, the proposed CNN-LSTM model estimated a yield of 3.610 tonnes/hectare, which was very close to the actual yield of 3.128 tonnes/hectare. For an actual value of 2.621 tonnes/hectare, the estimated value of CNN-LSTM was 2.501 tonnes/hectare, maintaining a minimal level of deviation. This demonstrates the ability to capture comprehensive spatial and temporal patterns in this dataset. While the random forest prediction was strong, there was some overprediction. For instance, an actual yield of 1.9 tonnes/hectare was predicted to be 2.397 tonnes/hectare, which is an overestimation of the value. Ensembles improved performance by combining the strengths of the different individual models. This provided much more refined predictions by the CNN + Random Forest ensemble, such as 3.341 tonnes/hectare for an actual yield of 3.610 tonnes/hectare and 2.239 tonnes/hectare for an actual yield of 1.9 tonnes/hectare, thereby reducing the prediction errors compared to standalone models. These findings raise the importance of integrating models to leverage their complementary strengths for reliable agricultural yield predictions.

4.5 The relationship between the predicted maize yields and the actual maize yields

Figure 4 shows a scatter plot of the relationship between the predicted and actual maize yields in tonnes per hectare, as forecasted by the CNN-LSTM model (blue dots) and the actual values from the dataset. The red dashed line represents the ideal fit, with the predicted values falling exactly on the actual values. The forecasted yields in the plot are typically very close to the perfect line, reflecting the model's ability to approximate the actual values accurately. For instance, a real yield of about 2.5 tonnes/hectare for maize

Table 4 Predictions of maize yields

Actual	CNN Predictions	CNN-LSTM Predictions	Random Forest Predictions	Ensemble (CNN- Random Forest)
3.61	2.999	3.128	3.553	3.341
1.9	2.556	2.08	2.397	2.239
1.4	2.291	1.993	2.362	2.178
2.994	3.249	3.093	3.008	3.051
1.6	2.161	1.854	2.35	2.102
2.622	2.546	2.501	2.589	2.545
3.593	3.032	3.164	3.554	3.359
2.4	2.493	2.457	2.193	2.325
2.385	2.478	2.529	2.424	2.477
2.2	2.117	1.694	1.473	1.583
2.996	3.225	3.064	3.029	3.047

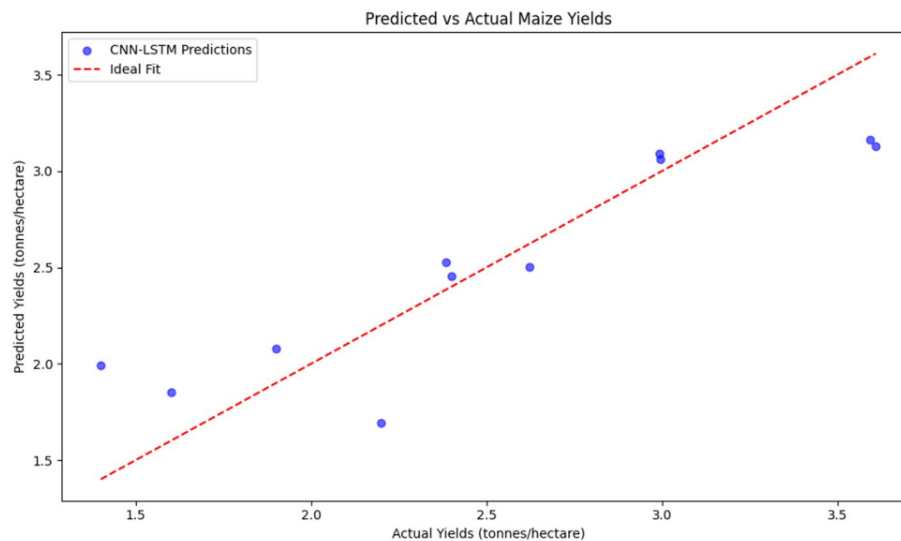


Fig. 4 Relationship between the predicted maize yields and actual maize yields

translates to approximately 2.6 tonnes/hectare in the CNN-LSTM model. This shows a little deviation from the actual. Similarly, predictions closely match actual values for a low yield of around 1.5 tonnes/hectare. However, there are minor deviations from the ideal fit. For example, at an actual yield of 2.0 tonnes/hectare, the predicted yield is slightly overestimated to about 2.2 tonnes/hectare. These minor discrepancies notwithstanding, the general scatter of the points around the red-dashed line indicates that the model captures the underlying structure of the maize yield data. From these, the resulting graphs show the predictive performance of the CNN-LSTM model: strong and with points closely clustered around the ideal fit line, further demonstrating that the CNN-LSTM is well-suited for this maize yield prediction problem.

4.6 The learning curves of the proposed CNN-LSTM model

Figure 5 shows the learning curves of the proposed CNN-LSTM model. It shows the trend in training and validation losses across 50 epochs. The training loss declines steeply within a few epochs (blue curve), indicating that the model has learned most of the underlying patterns in the training data remarkably quickly. Similarly, the validation loss also decreases drastically in the first few epochs, indicating the model's ability to generalize to unseen validation data. As the epochs progress, both curves converge and flatten, with the validation loss consistently lower than the training loss. That can be viewed as an indication that the model does not overfit the training data and has learned the patterns relevant to predicting the maize yield. This robustness and stability of the model are further supported by consistent convergence of both curves towards a low mean squared error. In the final epochs, the validation loss stabilizes, indicating that the model has reached its optimal learning capacity. This behaviour demonstrates the effectiveness of the CNN-LSTM architecture in modeling both spatial dependencies from the CNN and temporal dependencies from the LSTM in the dataset, thereby enabling accurate prediction of maize yield. The minimal gap between the training and validation losses further indicates that the model has a good balance between bias and variance, confirming its reliability for real-world applications.

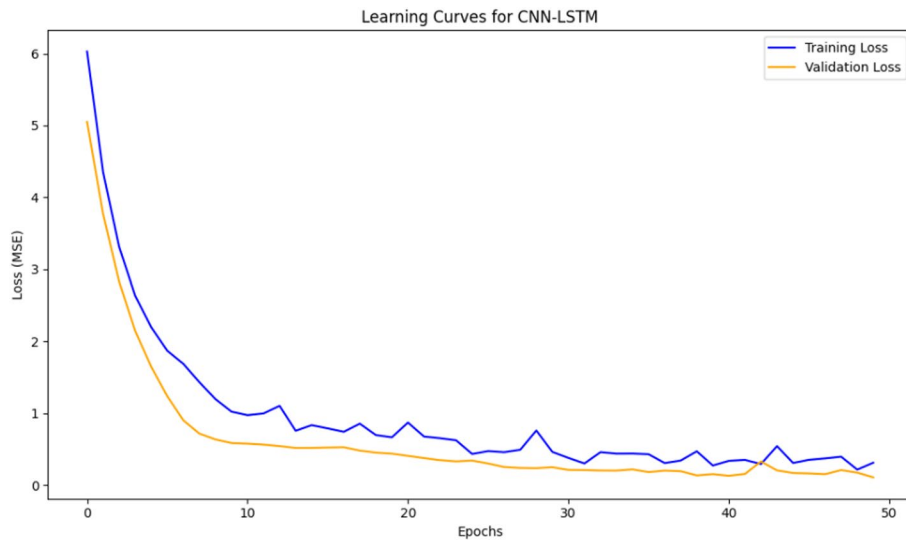


Fig. 5 The learning curves of the proposed CNN-LSTM model

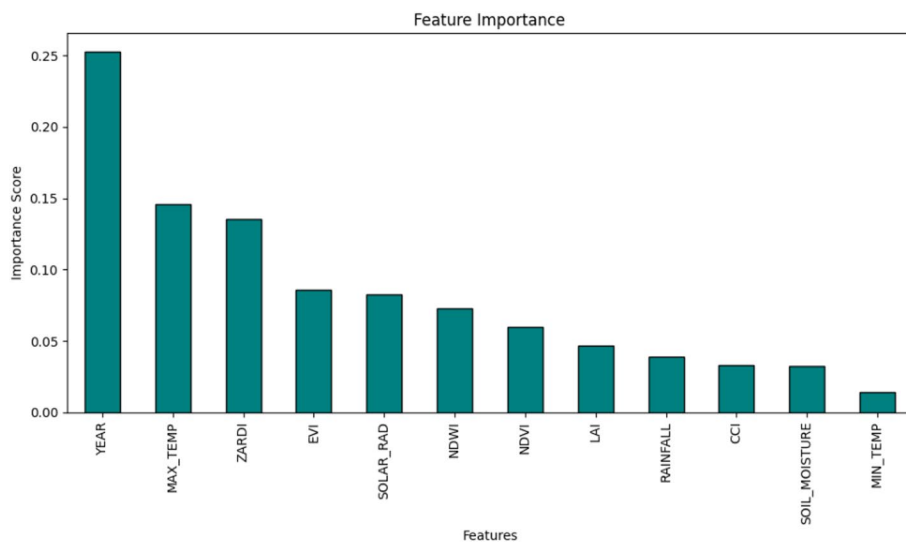


Fig. 6 The feature importance scores

4.7 The feature importance scores

Figure 6 shows the feature importance scores of the ML model in predicting maize yield. The prominence of 'Year' as the single most significant characteristic underscores the high interannual variation in maize yield in Uganda. 'Year' serves as a chronological indicator that integrates broader forces, including climatic changes, agronomic practices, and socio-economic variations that occur annually but are poorly captured through other covariates. For instance, changes in rainfall onset, distribution, and intensity from one year to the next significantly influence maize yields, and incorporating the 'Year' variable enables the model to capture these changes. Similar results have been reported in other studies, with 'Year' emerging as a significant predictor of crop yield variability because it captures systematic changes across growth seasons [53]. Thus, the high level of focus on 'Year' in this analysis does not indicate bias in the model, but rather reflects

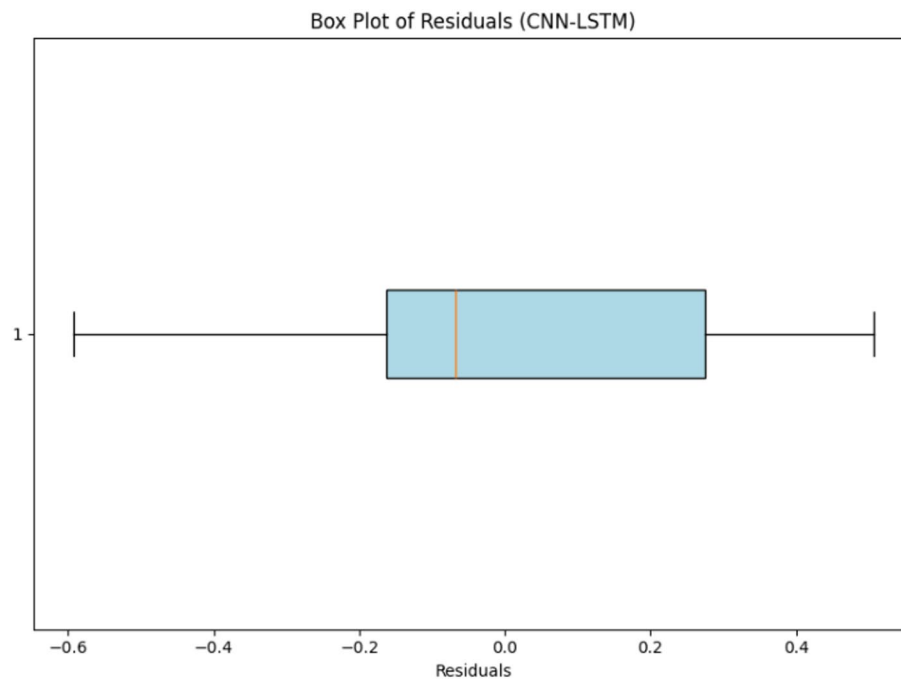


Fig. 7 The box plot of residuals for the CNN-LSTM model

its role in capturing inter-annual diversity, which has a significant impact on yield outcomes in smallholder farming enterprises.

This is followed by the feature `MAX_TEMP`, which reflects the critical role of maximum temperature in influencing maize growth and yield. Because very high temperatures can directly affect crop productivity, this feature is essential for accurate prediction. The `ZARDI` variable, representing geographical zones, ranks third, underscoring that yield varies across regions due to climatic and environmental conditions. Among these remote sensing variables, `EVI` and `NDWI` are the most significant contributors, indicating that vegetation health and water availability are key factors in yield prediction. Similarly, in photosynthesis and crop development, `SOLAR_RAD` and `NDVI` also contributed moderately. Features like `RAINFALL`, `CCI` (Canopy Chlorophyll Index), `SOIL_MOISTURE`, and `MIN_TEMP` contribute relatively low scores in this respect, yet they still add to the model's predictive potential. Although `RAINFALL` was not the highest-ranked feature overall, it remains agronomically critical. Its influence on yield is partly indirect; adequate rainfall improves VIs, such as `NDVI`, `NDWI`, and `EVI`, but extreme rainfall shortages or excesses directly impact yield and are reflected in the model's predictions. Together with soil moisture, it consistently emerged as one of the leading predictors, underscoring water availability as the principal constraint on maize in Uganda's semi-arid regions and reinforcing confidence in the model's robustness. Overall, this analysis highlights the interaction of temporal, climatic, spatial, and vegetation indices, where the model leverages these features to enhance the robustness and accuracy of maize yield predictions.

4.8 The box plot of residuals for the proposed CNN-LSTM model

Figure 7 presents the box plot of residuals for the CNN-LSTM model. It describes the distribution of errors, which is the difference between the predicted and actual maize

yields. The residuals are distributed symmetrically around the median, which is approximately zero. This suggests the model has no significant bias in underestimating or overestimating maize yield. The IQR is the width of the blue box, holding most of the residuals, which are small and fall within an acceptable range. The whiskers extend to the minimum and maximum values of the residuals; there are no extreme outliers, suggesting the model may be robust and consistent across the dataset. The compactness of the residual distribution demonstrates that the CNN-LSTM model effectively learns the underlying data pattern, thereby providing reliable predictions. Overall, this plot supports the claim that the CNN-LSTM minimizes prediction error and produces balanced, unbiased results for maize yield prediction.

5 Discussion of results

The results of this study indicate that the proposed CNN-LSTM model significantly enhances the accuracy of maize yield prediction compared to individual models such as CNN and Random Forest. With an MSE of 0.1068 tonnes², a RMSE of 0.3268 tonnes, and an R² score of 0.7833, the CNN-LSTM model demonstrated superior performance in learning both spatial and temporal dependencies in the data. These findings align with existing research that highlights the effectiveness of DL architectures for yield prediction [8, 35]. Random Forest (RF) performed well, achieving an MSE of 0.2105 tonnes², RMSE of 0.4588 tonnes, and R² score of 0.5730. While RF effectively captured complex relationships in the dataset, its reliance on static feature selection limited its ability to capture long-term dependencies in maize yield trends. This observation is consistent with previous studies that found ensemble tree-based methods, such as RF, to be strong predictors in structured agronomic data but limited in time-series forecasting [18]. In contrast, the CNN-LSTM model integrated sequential learning, enabling it to detect seasonal variations and long-term dependencies critical for yield prediction and surpassing RF in predictive accuracy.

Though effective in capturing spatial features, the CNN-only model performed relatively lower than CNN-LSTM, with an MSE of 0.2158 tonnes² and an R² score of 0.5622. This suggests that CNN alone cannot fully capture temporal patterns and sequence-based variations in yield trends. Similar findings were reported by Sun et al. [23], who observed that CNN models performed better when combined with sequence-learning models, such as LSTMs, in crop yield estimation. Furthermore, an ensemble of CNN and Random Forest showed substantial improvements, achieving an MSE of 0.1368 tonnes², an RMSE of 0.3699 tonnes, and an R² score of 0.7225. This improvement highlights that combining convolutional feature extraction from a CNN with decision-tree-based feature importance analysis from an RF enhances predictive accuracy. Similar trends were observed in studies where hybrid models outperformed standalone ML models by leveraging multiple feature extraction strategies [54]. Comparing the CNN-LSTM model's performance to prior studies, its R² score of 0.78 closely aligns with Zhou et al. [22], who achieved an R² of 0.78 using a CNN-Attention-LSTM model for maize yield prediction in China. Their model leveraged a similar multi-source dataset, incorporating remote sensing vegetation indices and climate variables. The proposed CNN-LSTM model also exhibited lower bias and variance in residual analysis compared to traditional ML models. This aligns with the findings of Muruganatham et al. [13], who noted that deep

learning models exhibit greater robustness when trained on multi-source datasets and with synthetic data augmentation techniques.

One key novelty in this study was the use of synthetic oversampling (SMOBN) to enhance model training. Acquiring high-quality ground-truth maize yield data across Uganda's ZARDI zones is challenging due to financial and logistical constraints. Consequently, the study utilized a dataset covering only three years, 2018 to 2020 (two seasons per year). By applying SMOBN, the training set size increased to 295 samples, ensuring that low- and high-yield cases were better represented. This argumentation significantly improved predictive accuracy, consistent with prior findings such as Ebrahimi et al. [30] and [42] that synthetic data can enhance crop yield models. Feature importance analysis identified YEAR as the most influential predictor. This likely reflects unmeasured temporal trends, such as gradual improvements in seed varieties and farming practices, that influenced yields from 2018 to 2020. In other words, the "Year" feature may serve as a proxy for factors not explicitly represented in the dataset. Among other top features are MAX_TEMP, NDVI, and EVI, which are sensible given maize's sensitivity to heat and the importance of canopy greens during growth. The prominence of NDWI and CCI among important variables further underscores that moisture status and chlorophyll levels are key factors in maize productivity. These findings are consistent with prior studies, which have shown that temperature and VIs are strongly linked to maize growth stages [17, 55, 56]. Thus, expanding irrigation systems in the north and northeast regions where rainfall deficits are most significant could secure yields and strengthen climate resilience. The findings highlight the capacity of hybrid DL methods to overcome data scarcity in smallholder farms, yielding more accurate and rapid yield estimates. Such advancements should aid in agricultural planning, enhance resource utilization, and bolster food security efforts in areas with limited data.

5.1 Limitations and future work

This study has several limitations, which also highlight directions for future research. The data set, covering the period from 2018 to 2020, was relatively small and may limit temporal generalizability, increasing the risk of overfitting. Although synthetic oversampling techniques, specifically SMOBN, were applied to address rare-yield situations, they assume local smoothness, which may produce unrealistic data in high-complexity attribute spaces. Moreover, the modest sample size prohibited explicit cross-zone validation. Splitting data would have further reduced sample sizes, risking instability during training and yielding unreliable estimates. Instead, we employed a fivefold stratified cross-validation, with SMOBN-based oversampling, dropout, and early stopping to improve robustness and reduce the risk of overfitting. While this approach achieved high predictive accuracy across the entire dataset, the lack of zone-specific holdout testing limits generalizability to novel agroecological zones. The exclusion of high-resolution imagery such as Sentinel-2 or high-resolution MODIS products, which better capture localized yield variability, also constrained spatial resolution. This limitation further restricted the use of data-intensive architectures, such as Transformers or 3D CNNs. Additionally, model explainability remains incomplete: Random Forest's use of importance scores provides only a global ranking of predictors without indicating the direction of their effects, i.e., whether a feature increases or decreases yield. Methods such as SHAP or LIME could yield more valid and locally interpretable insights. Prediction

uncertainty, such as confidence intervals, was not assessed; the analysis relied solely on deterministic performance metrics (e.g., R^2), which provide only point estimates on test data. Finally, although the CNN-LSTM model demonstrated improved predictive ability, the baseline ensemble models relied on equal-weight averaging, which may not fully leverage the CNN-LSTM's relative strengths under varying yield conditions. These limitations highlight the need for future research to enhance the CNN-LSTM framework through adaptive learning and rigorous uncertainty quantification.

- Adaptive Ensemble Learning. Moving beyond equal-weight averaging, weighted stacking can dynamically optimize each model's contribution, leading to further improvements in accuracy [57].
- Future work should also focus on quantifying prediction uncertainty. Generating confidence or prediction intervals, for example, through Bayesian neural networks, would convey valuable insights into the reliability of its forecasts. Such a certainty estimate is beneficial for decision-makers, as it indicates the level of confidence associated with each predicted yield and helps assess risk in planning [58].
- Integration of Transformer-based models. Future research can explore advanced architectures, such as BiLSTMs, 3D CNNs, and CNN-Transformer hybrids, to enhance attention mechanisms for maize yield prediction [52]. Recent studies [25] suggest that Transformers outperform LSTMs in capturing long-range dependencies.
- Incorporation of high-resolution satellite imagery. Expanding the dataset to include higher-resolution data, such as Sentinel-2 and fine-resolution MODIS products, UAV and IoT-derived field observations, could improve spatial detail and enhance predictive accuracy by capturing more localized crop variability [52].
- Explainable AI techniques. Incorporating interpretability methods, such as SHAP values, provides valuable insights into the influence of individual features on the model's predictions. Complementary tools, such as LIME for local interpretability and Grad-CAM, for visualizing CNN-LSTM decision processes, would further improve transparency and user trust in the model's outputs [59].
- Finally, future research should involve spatial cross-validation using larger and more geographically diverse datasets, enabling systematic evaluation of the model transferability across the highly heterogeneous maize-production environments in Uganda [60].

6 Conclusion

The study proposed a CNN-LSTM model to forecast maize yields across Uganda's ZARDI zones, integrating climate and remote-sensing data. A fully representative combined dataset capturing the complex interactions between climatic variables and maize yield was created through a comprehensive pre-processing pipeline that involved feature scaling, synthetic data augmentation, and dimensionality reduction. Among all these, the CNN-LSTM model has outperformed many single models, such as CNN and Random Forest, by significantly reducing prediction errors, demonstrating powerful predictive capability with an R^2 value of 0.78. The models further enhanced robustness and precision, with the CNN-Random Forest ensemble improving the R-squared score to 0.72, thereby underpinning the complementary strengths of deep learning and traditional machine learning. The residual analysis confirmed the reliability of the proposed model, with minimal bias in predictions and a strong fit to actual maize yield values. This

work underlines the importance of using multi-source datasets and state-of-the-art ML techniques to address the challenges of agricultural yield forecasting.

Thus, contributing both methodological innovation and regional insights, CNN-LSTM approaches emerge as a scalable tool for yield forecasting in data-scarce smallholder systems across SSA. They enable decision-makers and smallholder farmers to optimize resource allocation, mitigate risk, and plan farming practices effectively. While our model improved performance, limitations include a short data span, a lack of uncertainty quantification, the need for high-resolution satellite imagery, spatial crop validation, and interpretability. Future work will incorporate larger, multi-year datasets, higher-resolution imagery, uncertainty quantification, and employ explainable AI tools, along with advanced architectures such as Transformers, to yield more robust forecasts and stronger decision support for agricultural stakeholders. Hence, leveraging ML and DL technologies to predict crop yields would be crucial to advancing modern agriculture and alleviating global hunger.

Abbreviations

CNN	Convolutional neural network
LSTM	Long short-term memory
RF	Random forest
ANN	Artificial neural network
ML	Machine learning
DL	Deep learning
CYP	Crop yield prediction
SSA	Sub-Saharan Africa
ZARDI	Zonal Agricultural Research and Development Institute
NDVI	Normalized difference vegetation index
EVI	Enhanced vegetation index
NDWI	Normalized difference water index
LAI	Leaf area index
GPP	Gross primary productivity
SMOEN	Synthetic minority oversampling technique for regression
SMOTER	Synthetic minority oversampling technique (for regression)
MSE	Mean squared error
RMSE	Root mean squared error
MAE	Mean absolute error
R ²	Coefficient of determination
MAPE	Mean absolute percentage error
IQR	Interquartile range
UAV	Unmanned aerial vehicle
AE	Variational autoencoders
GAN	Generative adversarial networks
LIDAR	Light detection and ranging
CCI	Chlorophyll content index
AI	Artificial intelligence
SHAP	SHapley Additive exPlanations
LIME	Local interpretable model-agnostic explanations
Grad-CAM	Gradient-weighted class activation mapping

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s44163-026-00855-7>.

Supplementary Material 1

Acknowledgements

The authors gratefully acknowledge Kyambogo University for providing access to resources, a conducive research environment, and financial support.

Author contributions

Danison Taremwa: conceptualization, methodology, investigation, writing—original draft. Emmanuel Ahishakiye: supervision, guidance, review & editing. Aggrey Obbo: review & editing. Paul Kategaya Kisozi: review & editing. Fred Kagawa: supervision, guidance, review & editing. All authors approved the manuscript.

Funding

No organization, institution, or research centre funded this study.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations**Ethics approval and consent to participate**

Not applicable.

Clinical trial number

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 2 August 2025 / Accepted: 6 January 2026

Published online: 29 January 2026

References

1. Sambasivam G, Opiyo GD. A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolutional neural networks. *Egypt Inform J*. 2021;22(1):27–34. <https://doi.org/10.1016/j.eij.2020.02.007>.
2. Darra N, Anastasiou E, Kriezi O, Lazarou E, Kalivas D, Fountas S. Can yield prediction be fully digitized? A systematic review. *Agron*. 2023;13(9):1–53. <https://doi.org/10.3390/agronomy13092441>.
3. Tende IG, Aburada K, Yamaba H, Katayama T, Okazaki N. Development and evaluation of a deep learning based system to predict district-level maize yields in Tanzania. *Agric*. 2023;13(3):1–19. <https://doi.org/10.3390/agriculture13030627>.
4. Dabija A, Ciocan ME, Chettrariu A, Codină GG. Maize and sorghum as raw materials for brewing, a review. *Appl Sci*. 2021. <https://doi.org/10.3390/app11073139>.
5. Chivasa W, Mutanga O, Biradar C. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review. *Int J Remote Sens*. 2017;38(23):6816–45. <https://doi.org/10.1080/01431161.2017.1365390>.
6. Mahesh P, Soundrapandiyam R. Yield prediction for crops by gradient-based algorithms. *PLoS ONE*. 2024;19(8):1–20. <https://doi.org/10.1371/journal.pone.0291928>.
7. Yewle AD, Mirzayeva L, and Karakuş O. Multi-modal data fusion and deep ensemble learning for accurate crop yield prediction. 2025, [Online]. Available: <http://arxiv.org/abs/2502.06062>
8. Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. *Front Plant Sci*. 2020;10:1–14. <https://doi.org/10.3389/fpls.2019.01750>.
9. Lobell DB, et al. Eyes in the sky, boots on the ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis. *Am J Agric Econ*. 2020;102(1):202–19. <https://doi.org/10.1093/ajae/aaz051>.
10. Satpathi A, et al. Comparative analysis of statistical and machine learning techniques for rice yield forecasting for Chhattisgarh, India. *Sustain*. 2023;15(3):1–18. <https://doi.org/10.3390/su15032786>.
11. Aljahdali MO, Munawar S, Khan WR. Monitoring mangrove forest degradation and regeneration: landsat time series analysis of moisture and vegetation indices at Rabigh Lagoon, red sea. *Forests*. 2021;12(1):1–19. <https://doi.org/10.3390/f12010052>.
12. Mohammad N, Islam MA, Rahman MM, Ahmed I, Mahboob G. Yield forecasting model for maize using satellite multispectral imagery driven vegetation indices. *Qeios*. 2023. <https://doi.org/10.32388/coebosc>.
13. Muruganatham P, Wibowo S, Grandhi S, Samrat NH, Islam N. A systematic literature review on crop yield prediction with deep learning and remote sensing. *Remote Sens*. 2022. <https://doi.org/10.3390/rs14091990>.
14. Ali AM, et al. Integrated method for rice cultivation monitoring using Sentinel-2 data and leaf area index. *Egypt J Remote Sens Space Sci*. 2021;24(3):431–41. <https://doi.org/10.1016/j.ejrs.2020.06.007>.
15. Yang W, et al. Estimation of corn yield based on hyperspectral imagery and convolutional neural network. *Comput Electron Agric*. 2021. <https://doi.org/10.1016/j.compag.2021.106092>.
16. Nejad SMM, Abbasi-Moghadam D, Sharifi A, Farmonov N, Amankulova K, Laszlj M. Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches. *IEEE J Sel Top Appl Earth Observ Remote Sens*. 2023;16:254–66. <https://doi.org/10.1109/JSTARS.2022.3223423>.
17. Lu J, et al. Deep learning for multi-source data-driven crop yield prediction in northeast China. *Agriculture*. 2024;14(6):794.
18. Han Y, et al. Prediction of maize cultivar yield based on machine learning algorithms for precise promotion and planting. *Agric For Meteorol*. 2024. <https://doi.org/10.1016/j.agrformet.2024.110123>.
19. Kumar D. Biographical notes: Priyanka received her Bachelor of Technology in Computer Science and Engineering (CSE) and Master of Technology in CSE from GJUS&T. *Int J Inf Decis Sci*. 2020;12(3):246–69.
20. Wang Y, Feng K, Sun L, Xie Y, Song XP. Satellite-based soybean yield prediction in Argentina: a comparison between panel regression and deep learning methods. *Comput Electron Agric*. 2024. <https://doi.org/10.1016/j.compag.2024.108978>.
21. Shahhosseini M, Hu G, Khaki S, Archontoulis SV. Corn yield prediction with ensemble CNN-DNN. *Front Plant Sci*. 2021;12:1–13. <https://doi.org/10.3389/fpls.2021.709008>.

22. Zhou W, et al. A prediction model of maize field yield based on the fusion of multitemporal and multimodal UAV data: a case study in Northeast China. *Remote Sens.* 2023. <https://doi.org/10.3390/rs15143483>.
23. Sun J, Lai Z, Di L, Sun Z, Tao J, Shen Y. Multilevel deep learning network for county-level corn yield estimation in the U.S. corn belt. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2020;13:5048–60. <https://doi.org/10.1109/JSTARS.2020.3019046>.
24. Harsányi E, et al. Data mining and machine learning algorithms for optimizing maize yield forecasting in central Europe. *Agronomy.* 2023;13(5):1–22. <https://doi.org/10.3390/agronomy13051297>.
25. Ma Y, Yang Z, Huang Q, Zhang Z. Improving the transferability of deep learning models for crop yield prediction: a partial domain adaptation approach. *Remote Sens.* 2023;15(18):4562. <https://doi.org/10.3390/rs15184562>.
26. Hu X, Chen S, and Zhang D. Domain adaptation in agricultural image analysis: a comprehensive review from shallow models to deep learning. pp. 1–24, 2025, [Online]. Available: <http://arxiv.org/abs/2506.05972>
27. UBOS. "Annual agricultural survey," Report, no. 2, pp. 2–5, 2022, [Online]. Available: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679%26from=PT%0Ahttp://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52012PC0011:pt:NOT>
28. UBOS. Statistical Abstract Uganda Bureau of Statistics, Uganda Bur. Stat. Stat., pp. 1–336, 2022, [Online]. Available: [http://www.ubos.org/onlinefiles/uploads/ubos/pdf_documents/abstracts/Statistical Abstract 2013.pdf](http://www.ubos.org/onlinefiles/uploads/ubos/pdf_documents/abstracts/Statistical%20Abstract%202013.pdf)
29. Branco P, Ribeiro RP, Torgo L, Krawczyk B, Moniz N. SMOGN: a Pre-processing approach for imbalanced regression. *Proc Mach Learn Res.* 2017;74:36–50.
30. Ebrahimi H, Wang Y, Zhang Z. Utilization of synthetic minority oversampling technique for improving potato yield prediction using remote sensing data and machine learning algorithms with small sample size of yield data. *ISPRS J Photogram Remote Sens.* 2023;201:12–25. <https://doi.org/10.1016/j.isprsjprs.2023.05.015>.
31. Bali N, Singla A. Emerging trends in machine learning to predict crop yield and study its influential factors: a survey. *Arch Comput Methods Eng.* 2022;29(1):95–112. <https://doi.org/10.1007/s11831-021-09569-8>.
32. Jafarbiglu H, Pourreza A. A comprehensive review of remote sensing platforms, sensors, and applications in nut crops. *Comput Electron Agric.* 2022;197:106844. <https://doi.org/10.1016/j.compag.2022.106844>.
33. Bassine FZ, Epule TE, Kechchour A, and Chehbouni A. Recent applications of machine learning, remote sensing, and iot approaches in yield prediction: a critical review," 2023, [Online]. Available: <http://arxiv.org/abs/2306.04566>
34. Giovas R, Tassopoulos D, Kalivas D, Lougkos N, Priovolou A. Remote sensing vegetation indices in viticulture: a critical review. *Agriculture.* 2021. <https://doi.org/10.3390/agriculture11050457>.
35. Sun J, Di L, Sun Z, Shen Y, Lai Z. County-level soybean yield prediction using deep CNN-LSTM model. *Sensors (Switzerland).* 2019;19(20):1–21. <https://doi.org/10.3390/s19204363>.
36. Di Y, Gao M, Feng F, Li Q, Zhang H. A new framework for winter wheat yield prediction integrating deep learning and Bayesian optimization. *Agronomy.* 2022;12(12):1–15. <https://doi.org/10.3390/agronomy12123194>.
37. Fathi M, Shah-Hosseini R, Moghimi A. 3D-ResNet-BiLSTM model: a deep learning model for county-level soybean yield prediction with time-series Sentinel-1, Sentinel-2 imagery, and Daymet data. *Remote Sens.* 2023;15(23):1–20. <https://doi.org/10.3390/rs15235551>.
38. Epule TE, Dhiba D, Etongo D, Peng C, Lepage L. Identifying maize yield and precipitation gaps in Uganda. *SN Appl Sci.* 2021;3(5):1–12. <https://doi.org/10.1007/s42452-021-04532-5>.
39. Joshi A, et al. An explainable Bi-LSTM model for winter wheat yield prediction. *Front Plant Sci.* 2024;15(January):1–17. <https://doi.org/10.3389/fpls.2024.1491493>.
40. Li L, et al. Improving the estimation of alfalfa yield based on multi-source satellite data and the synthetic minority oversampling strategy. *Comput Electron Agric.* 2025;236:110497. <https://doi.org/10.1016/j.compag.2025.110497>.
41. Elabd E, Hamouda HM, Ali MAM, Fouad Y. Climate change prediction in Saudi Arabia using a CNN GRU LSTM hybrid deep learning model in al Qassim region. *Sci Rep.* 2025;15(1):1–19. <https://doi.org/10.1038/s41598-025-00607-0>.
42. Thihlum Z and Kiangte C. Impact of SMOGN on regression models for crop yield prediction in mizoram agriculture impact of SMOGN on regression models for crop yield prediction in Mizoram agriculture," no. May, 2025, <https://doi.org/10.1007/978-3-031-88039-1>
43. Li ZZ, Huang N, Yi LZ, Fu GH. Affine combination-based over-sampling for imbalanced regression. *J Chemom.* 2024;38(3):1–22. <https://doi.org/10.1002/cem.3537>.
44. Cao Z and Zhang Z. Corn yield prediction based on remotely sensed variables using variational autoencoder and multiple instance regression. *arXiv:2211.13286v1 [cs.CV]*. pp. 1–5
45. El-Kenawy ESM, Alhussan AA, Khodadadi N, Mirjalili S, and Eid MM. Predicting potato crop yield with machine learning and deep learning for sustainable agriculture, vol. 68, no. 1. Springer Netherlands, 2025. <https://doi.org/10.1007/s11540-024-09753-w>
46. Kumar R, Lad YA, Kumari P. Forecasting potato prices in Agra: comparison of linear time series statistical vs. neural network models. *Potato Res.* 2025. <https://doi.org/10.1007/s11540-024-09838-6>.
47. Jadon A, Patil A, and Jadon S. A comprehensive survey of regression based loss functions for time series forecasting. 2022, [Online]. Available: <http://arxiv.org/abs/2211.02989>
48. Terven JR, Cordova-esparza DM, Ramirez-pedraza A, Chavez-urbiola EA, and Romero-gonzalez JA. L f m d l. pp. 1–76
49. Steurer M, Hill RJ, Pfeifer N, Hill RJ, Pfeifer N. Metrics for evaluating the performance of machine learning based automated valuation models based automated valuation models. *J Prop Res.* 2021;38(2):99–129. <https://doi.org/10.1080/09599916.2020.1858937>.
50. Plevris V, Solorzano G, NP Bakas, and Ben Seghier MEA. Investigation of performance metrics in regression analysis and machine learning-based prediction models. *World Congr Comput Mech ECCOMAS Congr*, pp. 0–25, 2022, <https://doi.org/10.23967/eccomas.2022.155>
51. Eidele E, Ragab M, Chen Z, Wu M, Li X. TSLANet: rethinking transformers for time series representation learning. *Proc Mach Learn Res.* 2024;235:12409–28.
52. Challenges P. Applied deep learning-based crop yield prediction : a systematic analysis of current developments and potential challenges. *Technologies.* 2024;12(4):43.
53. Schumacher BL, Burchfield EK, Bean B, Yost MA. Leveraging important covariate groups for corn yield prediction. *Agric.* 2023;13(3):1–18. <https://doi.org/10.3390/agriculture13030618>.
54. Berveglieri A, et al. Remote prediction of Soybean yield using UAV-based hyperspectral imaging and machine learning models. *AgriEngineering.* 2024;6(3):3242–60. <https://doi.org/10.3390/agriengineering6030185>.

55. Yu L, et al. Near surface camera informed agricultural land monitoring for climate smart agriculture. *Clim Smart Agric.* 2024;1(1):100008. <https://doi.org/10.1016/j.csag.2024.100008>.
56. Kenduiywo BK, Miller S. Seasonal maize yield forecasting in South and East African countries using hybrid Earth observation models. *Heliyon.* 2024;10(13):e33449. <https://doi.org/10.1016/j.heliyon.2024.e33449>.
57. Tsang TK, Du Q, Cowling BJ, Viboud C. An adaptive weight ensemble approach to forecast influenza activity in an irregular seasonality context. *Nat Commun.* 2024;15(1):1–12. <https://doi.org/10.1038/s41467-024-52504-1>.
58. Ma Y, Zhang Z, Kang Y, Özdoğan M. Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach. *Remote Sens Environ.* 2021. <https://doi.org/10.1016/j.rse.2021.112408>.
59. Bouni M, Hssina B, Douzi K, Douzi S. Interpretable machine learning techniques for an advanced crop recommendation model. *J Electr Comput Eng.* 2024. <https://doi.org/10.1155/2024/7405217>.
60. Habibi LN, Matsui T, Tanaka TST. Critical evaluation of the effects of a cross-validation strategy and machine learning optimization on the prediction accuracy and transferability of a soybean yield prediction model using UAV-based remote sensing. *J Agric Food Res.* 2024;16:101096. <https://doi.org/10.1016/j.jafr.2024.101096>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.