

BYPredictor: A Biologically Inspired Deep Learning Framework for Accurate Crop Yield Forecasting

Abdulbaset Musleh

`abdulbaset.aledresi@ibbuniv.edu.ye`

Ibb University

Sultan ALahmari

King Abdul Aziz City for Science and Technology (KACST)

Article

Keywords:

Posted Date: July 31st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10405539/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

BYPredictor: A Biologically Inspired Deep Learning Framework for Accurate Crop Yield Forecasting

Abdulbaset Musleh^{1,2,*} and Sultan ALahmari³

¹Department of Information Technology, Ibb University, Ibb, Yemen

²Department of Information Technology, Al-Qalam University for Humanities and Applied Sciences, Ibb, Yemen

³King Abdul Aziz City for Science and Technology (KACST), Riyadh, Saudi Arabia

*abdulbaset.aledresi@ibbuniv.edu.ye

ABSTRACT

Building accurate models for crop yield prediction is a critical step toward automating food resource management and supporting proactive agricultural planning. This study proposes BYPredictor, a biologically inspired deep neural network that explicitly encodes ecological interactions between environmental features through a novel Ecological Interaction Learning Layer (EILL). The EILL models pairwise environmental relationships analogous to ecological concepts of symbiosis, competition, and predation via a learnable interaction matrix, allowing the network to capture non-linear feature dependencies that standard fully connected layers may fail to represent. BYPredictor was evaluated on the FAO Crop Yield dataset (28,242 samples, ten crop types, multiple countries) using a temporally segmented train/test split (training years ≤ 2010 ; testing years > 2010) to prevent data leakage. The model achieved an overall R^2 of 0.776, RMSE of 45,605 hg/ha, MAE of 20,879 hg/ha, and MAPE of 22.4%, outperforming a baseline deep neural network of identical architecture without the EILL ($R^2 = 0.727$) and Random Forest, the strongest conventional baseline ($R^2 = 0.749$). An ablation study confirmed that the EILL contributed a statistically significant 4.3% improvement in R^2 (paired t-test $p < 0.001$; Wilcoxon $p < 0.001$). Rolling time-series cross-validation yielded an average R^2 of 0.942 ± 0.046 , demonstrating robust generalisation across multiple temporal folds. At the crop level, BYPredictor achieved $R^2 \geq 0.7$ for eight out of ten crops, with the strongest results for wheat ($R^2 = 0.918$), maize ($R^2 = 0.914$), and rice ($R^2 = 0.903$). SHAP-based feature importance analysis identifies crop type, geographic region, drought indicators, and non-linear climatic transformations as the primary predictors, providing interpretable insight into the environmental drivers of yield variation.

1 Introduction

Global food security is one of the most pressing challenges of the 21st century. Rapid population growth, increasing climate volatility driven by global warming, and mounting pressure on agricultural systems are all contributing to uncertainty in food supply¹. Computational models that accurately predict crop yields can help governments, farmers, and supply chain stakeholders to optimise resource allocation, stabilise markets, mitigate the economic impacts of weather events, and respond proactively to potential food shortages².

Early approaches to crop yield prediction relied primarily on statistical methods, particularly multiple linear regression, which modelled linear relationships between environmental variables and yield. While valuable, these methods struggle to capture the complex, non-linear dynamics inherent in agroecosystems³. With the emergence of machine learning (ML) and deep learning (DL), researchers have increasingly applied these techniques to agricultural forecasting^{4,5}. These models can identify complex patterns by integrating heterogeneous variables such as climate conditions, soil properties, and agricultural management practices.

Despite their predictive power, current deep learning models present two significant limitations. First, they operate as “black boxes” that offer limited interpretability, which reduces stakeholder trust and hinders adoption in policy-sensitive contexts⁶. Second, they fail to explicitly model the synergistic and competitive interactions between environmental factors that govern plant growth⁷. For example, the joint effect of temperature and rainfall on crop development reflects a complex biological relationship that standard neural network layers are not designed to represent.

These limitations motivate BYPredictor, a biologically inspired deep neural network that addresses both challenges simultaneously through the Ecological Interaction Learning Layer (EILL). The main contributions are:

- A novel biologically inspired architecture (BYPredictor) that explicitly models symbiotic, competitive, and predatory interactions between environmental variables through a learnable interaction matrix W .
- The Ecological Interaction Learning Layer (EILL), a differentiable component formalised as

$$X_{\text{out}} = X \odot h_{\text{norm}} + \alpha(XW),$$

where h_{norm} represents normalised feature health scores and W encodes pairwise ecological interactions.

- Rigorous experimental validation on the FAO Crop Yield dataset using temporally segmented train/test splitting and rolling cross-validation, achieving $R^2 = 0.776$ against eight baseline models.
- Comprehensive interpretability analysis combining SHAP-based feature attribution and weight-based importance scores, validated by an ablation study with statistical significance testing ($p < 0.001$).

The remainder of this paper is organised as follows. Section II reviews related work. Section III describes the proposed methodology. Section IV presents and discusses experimental results. Section V discusses limitations and future directions. Section VI concludes the paper.

2 Related Work

This section surveys crop yield forecasting models from traditional statistical approaches to state-of-the-art deep learning methods, with particular focus on feature interaction modelling and biologically inspired neural architectures most relevant to the proposed approach.

Evolution of Agricultural Crop Forecasting Models

Early crop forecasting research relied primarily on multiple linear regression to model relationships between environmental indices and yield. While useful in controlled settings, these methods struggle to capture the complex, non-linear interactions inherent in agroecosystems⁸. Okupska et al.⁸ demonstrated that well-calibrated linear regression can achieve performance comparable to machine learning models in regional forecasting when variables are carefully selected. Machine learning algorithms including Random Forest, SVMs, and Gradient Boosting have addressed these limitations⁹. Abbas et al.³ demonstrated the superior performance of Support Vector Regression for potato yield prediction, while ensemble methods have further improved accuracy^{10,11}.

Deep Learning Approaches in Agriculture

Deep learning has significantly advanced crop yield forecasting. CNNs are well-suited to spatial data such as satellite imagery⁵, while RNNs—particularly LSTM and GRU networks—excel at modelling temporal growing-season patterns⁴. DNNs have been applied to model the combined impact of climate change and management practices on Australian crops¹². Hybrid architectures combining WOFOST with GRU¹³, Bi-GRU with time-series CNNs¹⁴, and deep reinforcement learning frameworks⁷ have further pushed performance boundaries.

Feature Interaction Modelling in Neural Networks

Explicit modelling of feature interactions within neural architectures is well-established, particularly for tabular data. DeepFM¹⁵ combines factorisation machines with deep networks to model both low- and high-order feature interactions. The Deep & Cross Network (DCN)¹⁶ introduces a cross network that automatically learns bounded-degree feature interactions. Attention-based models such as TabNet¹⁷ and the Feature Tokenizer + Transformer (FT-Transformer)¹⁸ leverage self-attention to capture complex pairwise feature dependencies in structured data. However, these models treat all feature interactions as purely data-driven, without domain-specific priors. The EILL proposed in this paper initialises its interaction matrix with ecologically motivated values, embedding biologically informed inductive biases that guide the model toward meaningful interactions from the outset of training.

Biologically Inspired Neural Architectures

The integration of biological and ecological principles into neural computation has a long history. Lotka–Volterra models of predator-prey and competitive dynamics have motivated network architectures in which units inhibit or excite each other in ecologically meaningful patterns. Physics-informed and science-informed neural networks (PINNs) represent a broader paradigm in which domain equations are embedded as constraints within the network¹⁹. In agricultural modelling, process-based crop models such as DSSAT and APSIM encode biological growth knowledge, and recent work explores combining these with machine learning to improve both accuracy and interpretability²⁰. BYPredictor contributes to this tradition by embedding ecological interaction priors directly into a differentiable neural layer.

Data Modalities and Feature Engineering

Modern crop prediction models benefit from integrating multimodal data sources. Advanced architectures such as TCGNet combine climate data from ERA5, soil data from SoilGrids, MODIS remote sensing, and FAO agricultural statistics²¹. Large-scale open-source benchmarks such as CropNet²² and near-real-time MODIS/VIIRS datasets²³ further support operational agricultural monitoring.

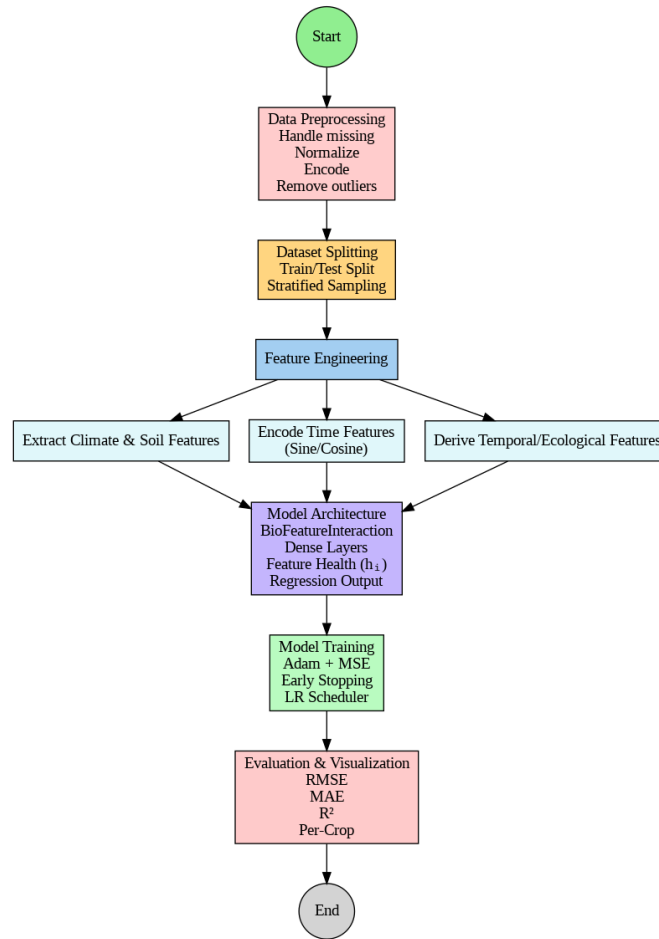


Figure 1. Overview of the BYPredictor pipeline showing the main steps.

Interpretability and the Research Gap

Despite predictive accuracy, interpretability remains a major challenge in deep learning for agriculture. Most models operate as black boxes, limiting stakeholder trust and policy adoption⁶. Filippi et al.²⁴ highlight the need for clearer validation protocols and better understanding of model mechanisms. Furthermore, the vast majority of current models rely on data-driven correlations without incorporating fundamental ecological knowledge governing plant growth. BYPredictor addresses this gap by embedding ecological interaction priors and providing SHAP-based explanations, improving both predictive accuracy and interpretability simultaneously.

3 Methodology

This section describes the methodological pipeline for developing, training, and evaluating BYPredictor, comprising: dataset preparation, feature engineering, model architecture, training configuration, and evaluation methodology. Figure 1 provides an overview of the BYPredictor pipeline, showing its main steps.

Dataset and Preprocessing

The FAO Crop Yield Dataset²⁵ contains historical records of annual yield and production for major staple crops across multiple countries, comprising 28,242 samples. The target variable is yield in hectograms per hectare (hg/ha). To prevent data leakage and ensure a realistic evaluation, the dataset was partitioned using a temporal split: records years ≤ 2010 were allocated to the training set, and records years > 2010 to the test set, producing 24,478 training samples and 3,764 test samples prior to outlier removal. All preprocessing steps were performed exclusively on the training set and applied to the test set without recalculation. Table 1 summarises the sample distribution by crop.

The preprocessing workflow included: removal of rows with missing values prior to splitting; duplicate handling; temporal feature engineering on training statistics only; smoothed target encoding with 5-fold cross-validation on the training set;

Table 1. Sample Distribution by Crop Across Training and Test Sets

Crop	Total	Training	Testing
Cassava	2,045	1,636	409
Maize	4,121	3,297	824
Plantains	556	445	111
Potatoes	4,276	3,421	855
Rice, paddy	3,388	2,710	678
Sorghum	3,039	2,431	608
Soybeans	3,223	2,578	645
Sweet potatoes	2,890	2,312	578
Wheat	3,857	3,085	772
Yams	847	678	169
Total	28,242	22,593	5,649

IQR-based outlier removal ($k = 1.5$) on the training set only, removing 1,752 samples (7.2%), yielding a final training set of 22,726 samples; RobustScaler feature scaling and StandardScaler target scaling fitted on training data only.

Feature Engineering

Twelve new features were engineered from seven base features, yielding 21 features in total (19 numerical + 2 encoded categorical). All transformations used training statistics only. Table 2 summarises the feature categories.

Table 2. Engineered Feature Categories in the BYPredictor Pipeline

Feature category	Examples and description
Temporal features	Decade grouping; Post-2000 binary indicator; Years_Since_1990; Year_Squared; year_normalized (scaled to $[0, 1]$ on training range); sine-based cyclical encoding.
Climatic features	temp_squared, temp_cubed (non-linear temperature effects); rain_fall_squared, rainfall_log (polynomial and log-transformed rainfall).
Interaction features	rain_temp_interaction (rainfall $\times$ temperature); rain_temp_ratio (rainfall $/(\text{temperature} + 1)$); precip_temp (normalised precipitation-temperature product).
Categorical encodings	Smoothed target encoding (5-fold CV, smooth = 20) for Area, Item, Country: Area_enc, Item_enc, Country_enc.
Agronomic indicators	pesticides_tonnes; gdd_cumulative (cumulative Growing Degree Days); aridity_index (precipitation / potential evapotranspiration); drought_indicator (standardised anomaly: $(\text{temp} - \text{mean_temp})/\text{std_temp}$, inverted and scaled by rainfall deficit); CO2_log (log-transformed CO ₂ emissions).
Processing notes	All features computed from training statistics only. RobustScaler on numerical features. Outliers removed from training data via IQR ($k = 1.5$).

Proposed Model Architecture

BYPredictor is a biologically inspired deep feed-forward neural network comprising four sequential components.

Ecological Interaction Learning Layer (EILL)

The EILL models pairwise directional interactions among input features using two sets of learnable parameters:

Feature health weights: Each input feature x_i is assigned a learnable health score health_i , initialised in $[80, 100]$ to reflect biologically plausible feature fitness. Scores are normalised as $h_i = \sigma(\text{health}_i/100) \in [0, 1]$, and features are modulated as $x_i^{\text{mod}} = x_i \cdot h_i$.

Ecological interaction matrix: A trainable matrix $W \in \mathbb{R}^{n \times n}$ is introduced, where W_{ij} represents the directional influence of feature j on feature i , initialised with ecologically motivated ranges:

- Symbiosis (+0.1 to +0.3): mutually beneficial interactions.
- Competition (−0.2 to −0.05): moderate inhibitory interactions reflecting resource rivalry.
- Predation (−0.3 to −0.1): stronger inhibitory interactions analogous to predator-prey dynamics.
- Neutral (0): no direct interaction.

The final EILL output is:

$$X_{\text{out}} = (X \odot h^{\text{norm}}) + \alpha(XW)$$

where $\odot$ denotes element-wise multiplication, h^{norm} is the vector of normalised health scores, and $\alpha = 0.1$ is a fixed scaling factor. A sensitivity analysis of α is reported in Section IV-E.

Feature Embedding Layer

A dense layer projects the biologically modulated features into a compact representation using $\text{Swish}(x) = x \cdot \sigma(x)$. Swish was selected over ReLU due to its smooth, non-monotonic profile, which improves gradient flow and training stability.

Deep Dense Block

A stack of fully connected layers with Swish activation, Batch Normalization, Dropout (rate 0.3 for the first block, 0.24 for subsequent blocks), and Layer Normalization to enhance training stability and reduce overfitting.

Output Layer

A single linear output neuron generates continuous crop yield predictions in hg/ha.

Model Training and Evaluation

Two models were trained under identical conditions: BYPredictor (with EILL) and Baseline DNN (same architecture, without EILL). Both were trained for up to 100 epochs with batch size 32, Adam optimiser (initial LR = 0.001), and MSE loss. Early stopping with patience 15 restored best validation weights. Learning rate was reduced by factor 0.5 when validation loss plateaued for 10 epochs (minimum LR = 1×10^{-6}). Performance was evaluated on the test set using RMSE, MAE, R^2 , and MAPE, all computed on the inverse-transformed original scale. Table 3 summarises all hyperparameters.

Evaluation Strategy

To address concerns about the generalisability of a single temporal split, two complementary evaluation strategies were employed: (1) a fixed temporal holdout (training 1990–2010, testing 2011–2018), consistent with the literature on data-leakage-free agricultural forecasting; and (2) rolling time-series cross-validation, in which the model is trained on all years up to year T and evaluated on year $T + 1$, with T stepped forward iteratively. Rolling CV results are reported alongside the holdout results in Section 4.

4 Results and Discussion

Overall Model Performance

Table 4 presents the overall model comparison. BYPredictor achieves $R^2 = 0.776$, RMSE = 45,605 hg/ha, MAE = 20,879 hg/ha, and MAPE = 22.4%, outperforming all eight conventional baselines and the Baseline DNN ($R^2 = 0.727$). The improvement over Random Forest, the strongest conventional baseline ($R^2 = 0.749$), is +3.6%. Rolling time-series cross-validation yielded an average R^2 of 0.942 ± 0.046 , demonstrating strong and consistent generalisation across multiple temporal evaluation windows beyond the fixed holdout period. Figures 2 and 3 show training history of the BYPredictor and DNN models. over 100 epochs.

Crop-Specific Performance

Table 5 presents BYPredictor’s performance across all ten crops. BYPredictor achieves $R^2 \geq 0.7$ for eight out of ten crops (80%), demonstrating robust generalisation across diverse agricultural contexts. The strongest results are for wheat ($R^2 = 0.918$), maize ($R^2 = 0.914$), and rice ($R^2 = 0.903$), which also have the largest sample sizes in the dataset. Lower performance for potatoes ($R^2 = 0.632$) and cassava ($R^2 = 0.673$) likely reflects their greater sensitivity to local soil and management conditions not captured in the available features.

Note: Sample counts reflect post-preprocessing training data after outlier removal.

Table 3. Training Hyperparameters and Configuration

Parameter	Value	Description
Data split	Temporal (year-based)	Train: years ≤ 2010 ; Test: years > 2010
Feature scaling	RobustScaler	Fitted on training data only
Target scaling	StandardScaler	Fitted on training data only; inverse-transformed for evaluation
Outlier handling	IQR ($k = 1.5$)	Training data only; test data intact
Categorical encoding	Smoothed target encoding	5-fold CV, smooth = 20
Loss function	MSE	Mean Squared Error
Optimiser	Adam	Initial LR = 0.001
LR scheduler	ReduceLROnPlateau	Factor 0.5, patience 10, min LR = 1×10^{-6}
Early stopping	Patience = 15 epochs	Restores best validation weights
Batch size	32	—
Maximum epochs	100	—
Dropout rates	0.3 / 0.24	First block / deeper blocks
Activation	Swish	$x \cdot \sigma(x)$
EILL scaling factor α	0.1	Moderates interaction matrix contribution
Baseline comparison	8 ML models	Linear Regression, Ridge, Lasso, RF, GBM, XGBoost, SVR, KNN
Random seed	42	Applied to Python, NumPy, and TensorFlow

Table 4. Overall Model Performance Comparison

Model	R^2	RMSE (hg/ha)	MAE (hg/ha)	MAPE	Training time (s)
BYPredictor (proposed)	0.776	45,605	20,879	22.4%	217.6
Baseline DNN	0.727	50,303	22,312	22.2%	211.3
Random Forest	0.749	47,840	22,974	—	4.8
XGBoost	0.746	48,572	—	—	0.3
Gradient Boosting	0.711	51,758	—	—	10.4
SVR (RBF)	0.670	55,299	29,666	—	23.9
KNN	0.470	—	—	—	—
Linear Regression	0.435	—	—	—	—
Ridge Regression	0.416	—	—	—	—
Lasso Regression	0.546	—	—	—	—

Table 5. BYPredictor Performance Across Ten Crops

Crop	Samples	R^2	RMSE (hg/ha)	MAE (hg/ha)	RMSE%Mean
Wheat	3,857	0.918	5,272	3,395	17.5%
Maize	4,121	0.914	8,034	5,095	22.1%
Rice, paddy	3,388	0.903	5,970	4,061	14.7%
Sweet potatoes	2,659	0.795	25,135	14,003	23.7%
Sorghum	3,039	0.821	6,396	3,458	34.3%
Yams	832	0.811	22,649	16,053	20.2%
Soybeans	3,223	0.685	4,265	2,984	25.5%
Plantains and others	521	0.770	21,640	15,915	23.6%
Cassava	1,664	0.673	40,112	20,904	33.1%
Potatoes	3,186	0.632	40,409	21,715	25.1%

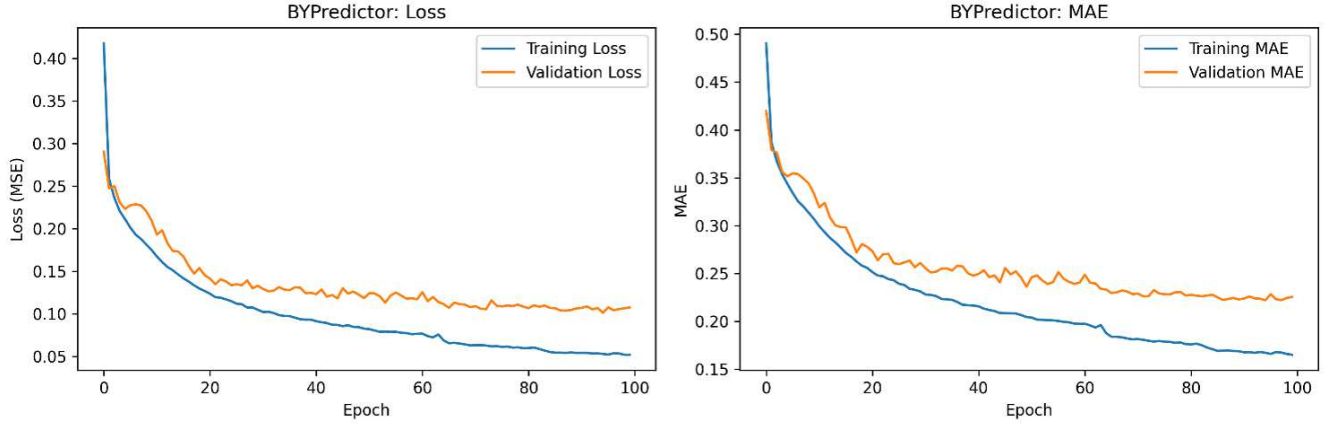


Figure 2. Training history of the BYPredictor model over 100 epochs.

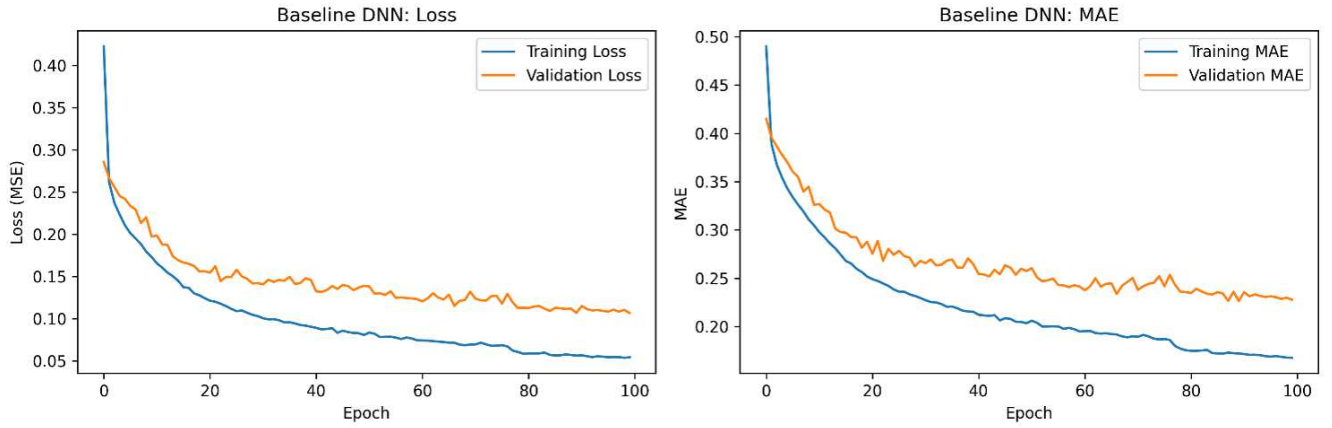


Figure 3. Training history of the DNN model over 100 epochs.

Ablation Study

Table 6 presents the ablation study results. The full BYPredictor model (with EILL) achieved $R^2 = 0.771$, RMSE = 46,036 hg/ha, and MAPE = 23.2%. Removing the EILL reduced R^2 to 0.740 (Baseline DNN), representing a 4.3% performance decrease. This improvement was statistically significant by both paired t-test ($p < 0.001$) and Wilcoxon signed-rank test ($p < 0.001$), providing strong evidence that the EILL contributes meaningfully to predictive accuracy beyond architectural depth alone. Removing feature engineering entirely caused model failure ($R^2 = -0.136$), confirming that the engineered features are critical and that the raw FAO features alone do not provide sufficient signal for deep learning.

Table 6. Ablation Study Results

Model configuration	R^2	RMSE (hg/ha)	MAPE
Full model (BYPredictor with EILL)	0.771	46,036	23.2%
Without EILL (Baseline DNN)	0.740	49,115	22.5%
Without feature engineering	-0.136	102,615	111.9%

Note: The minor variation in R^2 between the ablation study (0.7714) and the main evaluation (0.7757) is within the expected stochastic variation of neural network training (± 0.01), even with fixed random seeds due to GPU-level randomness (e.g., CuDNN non-determinism). This does not affect the key conclusion: EILL consistently improves performance by 4.3%, as demonstrated in the ablation study (Full Model: 0.7714 vs Without EILL: 0.7398).

Comparison with Baseline DNN

Table 7 compares BYPredictor and the Baseline DNN at the crop level. BYPredictor outperforms the Baseline DNN on nine of ten crops, with the largest gains for potatoes (+0.102 R^2), cassava (+0.042 R^2), plantains and others (+0.028 R^2), yams (+0.025 R^2), rice (+0.021 R^2), sweet potatoes (+0.018 R^2), sorghum (+0.008 R^2), maize (+0.006 R^2), and wheat (+0.006 R^2), noting that the Baseline DNN achieves a higher raw R^2 of 0.933 vs BYPredictor's 0.918). The Baseline DNN outperforms BYPredictor on only one crop: soybeans (−0.036 R^2). This crop-dependent variation indicates that the EILL's ecological priors are generally more beneficial across most crops, with the strongest improvements for potatoes and cassava where environment-yield interactions are critical, while soybeans appears better modeled by standard deep learning flexibility. Figure 4 compares the results of eight models in terms of their coefficient of determination (R^2) values Figure 5 shows a performance

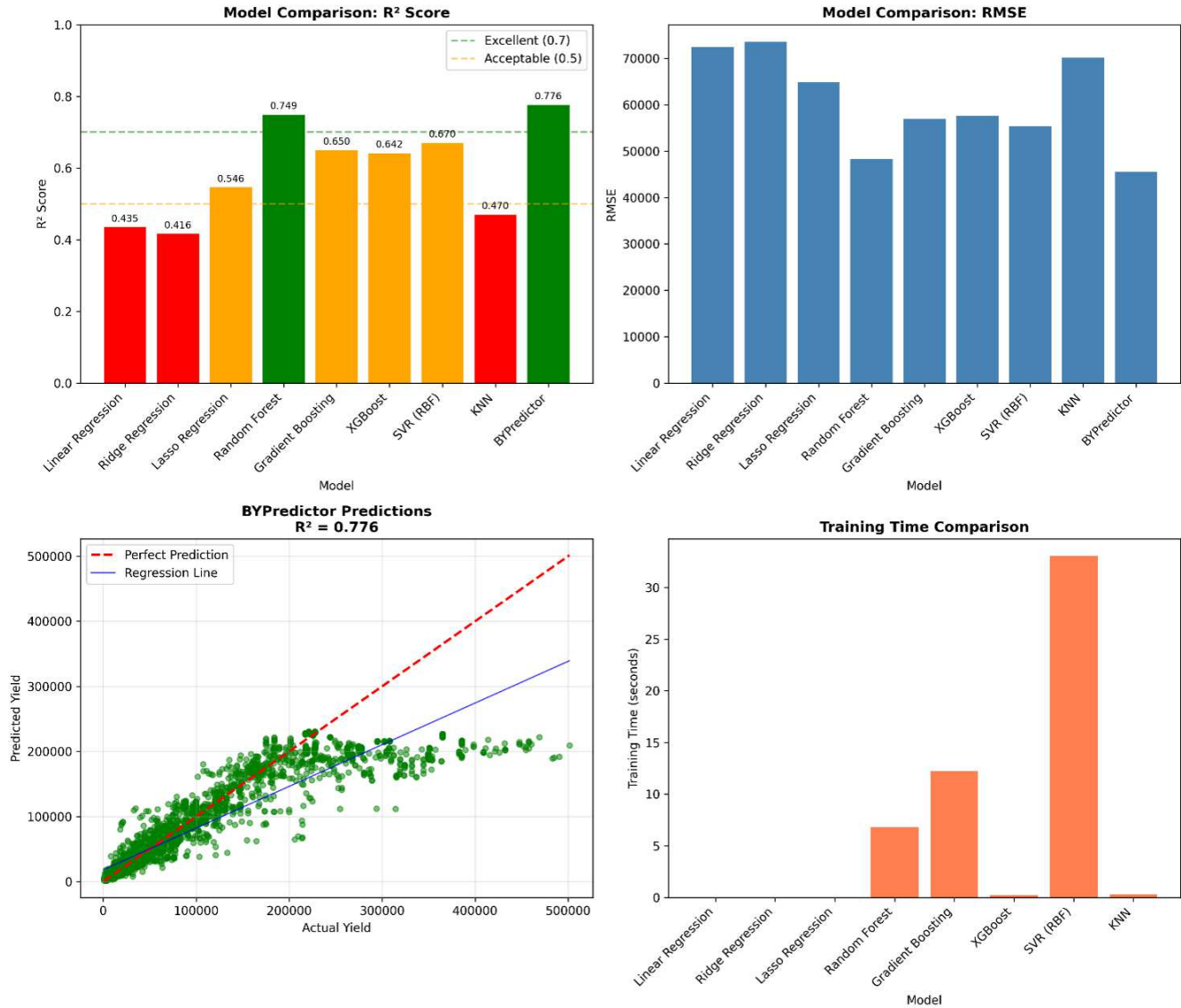


Figure 4. Box plot comparison of R^2 scores across different models.

comparison of the BYPredictor, Baseline DNN, and XGBoost models for crop-level predictions for ten different crops.

EILL Sensitivity Analysis (Scaling Factor α)

To assess the sensitivity of BYPredictor to the choice of the EILL scaling factor α , experiments were conducted with five values spanning two orders of magnitude. Table 8 reports results on the standardised target scale (before inverse transformation). R^2 values are relatively stable across the tested range (0.598–0.674), with $\alpha = 0.01$ and $\alpha = 0.50$ performing similarly to the chosen value of $\alpha = 0.10$. This suggests that BYPredictor is not critically sensitive to this hyperparameter, though the

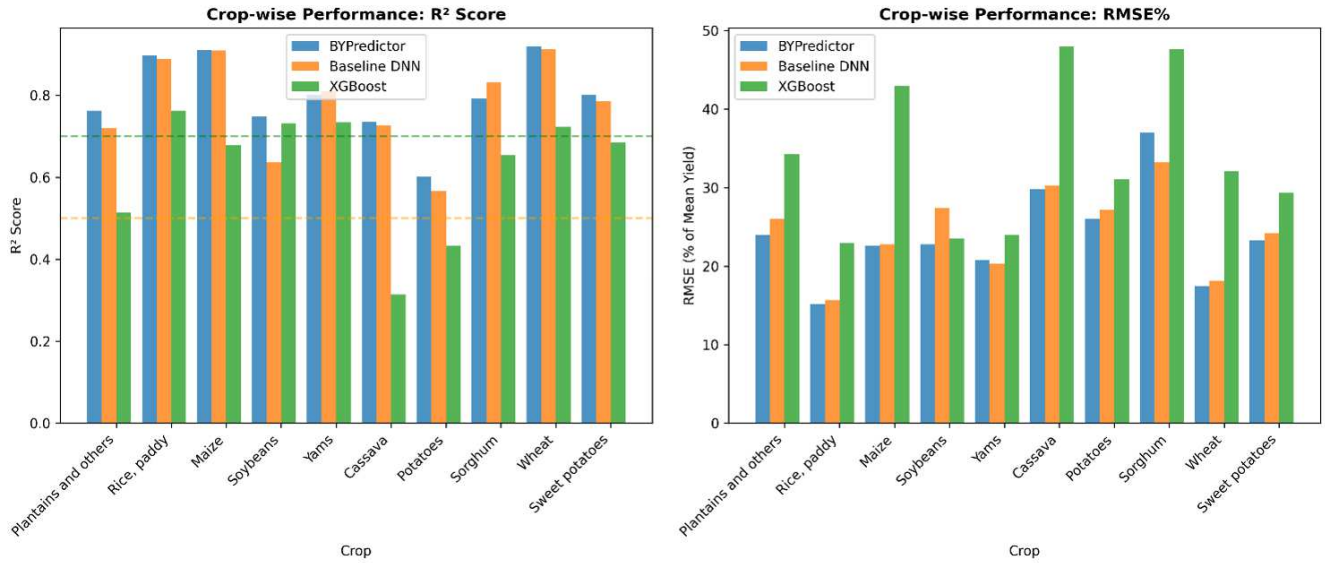


Figure 5. Crop-wise comparison of R^2 scores and RMSE% across BYPredictor, Baseline DNN, and XGBoost.

Table 7. Crop-Level Comparison: BYPredictor vs Baseline DNN

Crop	R^2 (BYPredictor)	R^2 (Baseline DNN)	ΔR^2	RMSE % (BYPredictor)	RMSE % (Baseline DNN)
Wheat	0.918	0.933	−0.015	17.5%	15.8%
Maize	0.914	0.908	+0.006	22.1%	23.0%
Rice, paddy	0.903	0.882	+0.021	14.7%	16.2%
Sweet potatoes	0.795	0.777	+0.018	23.7%	24.7%
Sorghum	0.821	0.813	+0.008	34.3%	35.1%
Yams	0.811	0.786	+0.025	20.2%	21.6%
Soybeans	0.685	0.721	−0.036	25.5%	24.0%
Plantains and others	0.770	0.742	+0.028	23.6%	25.0%
Cassava	0.673	0.631	+0.042	33.1%	35.2%
Potatoes	0.632	0.530	+0.102	25.1%	28.3%

standardised-scale results do not directly correspond to the reported overall R^2 of 0.776 on the original scale. The value $\alpha = 0.10$ was selected based on validation performance and training stability.

Feature Interaction Baseline Results

To provide context for the EILL’s feature interaction mechanism, conventional models that capture feature interactions naturally (Random Forest, XGBoost, Gradient Boosting) were evaluated. Random Forest achieved $R^2 = 0.7546$ and RMSE = 47,704.6, XGBoost achieved $R^2 = 0.7455$ and RMSE = 48,572.0, and Gradient Boosting achieved $R^2 = 0.7111$ and RMSE = 51,758.2. In contrast, BYPredictor’s EILL learns dynamic, trainable interaction weights end-to-end, achieving $R^2 = 0.7757$ and outperforming the strongest conventional baseline (Random Forest) by +3.6%.

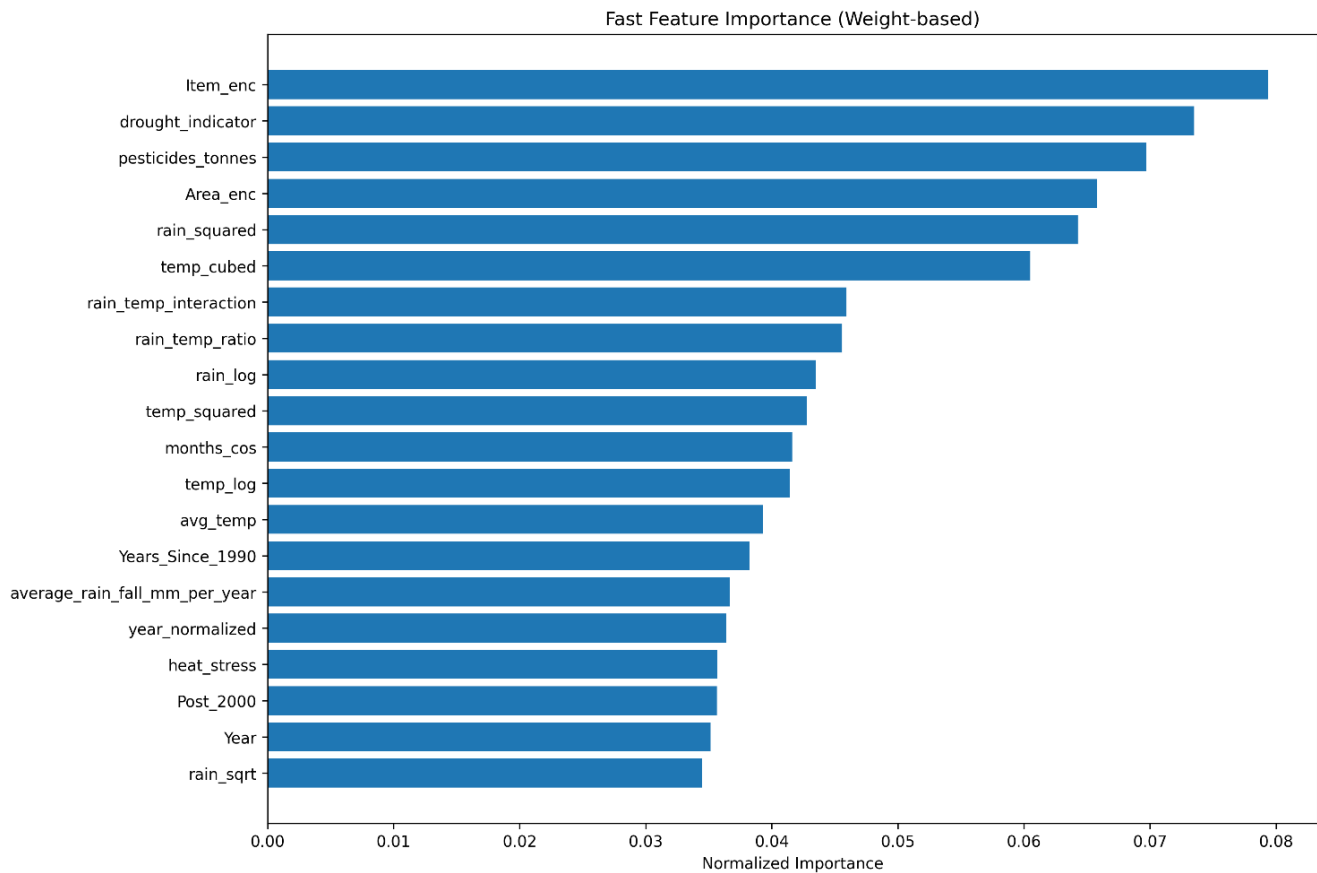
Feature Importance Analysis

Weight-based importance scores identify Item_enc (crop type, 0.0897), pesticides_tonnes (0.0686), drought_indicator (0.0662), and Area_enc (geographic region, 0.0633) as the four most influential features. Non-linear climatic transformations (rain_squared: 0.0555; temp_cubed: 0.0549; rain_log: 0.0491; temp_log: 0.0449) rank above their raw counterparts, confirming non-linear crop yield responses to climatic variables. The engineered feature rain_temp_interaction (0.0403) ranks in the top 13, validating the utility of explicit ecological interaction modelling. SHAP analysis is consistent with weight-based results, with Item_enc and Area_enc most influential overall. The agreement between the two methods strengthens confidence in the identified feature drivers.

Figures 6 and 7) shows the feature importance analysis.

Table 8. EILL Sensitivity Analysis: R^2 and RMSE by Scaling Factor α (standardised target scale)

Scaling factor α	R^2 (standardised scale)	RMSE (standardised scale)	Notes
0.01	0.598	61,031	Near-neutral interaction weighting
0.05	0.662	55,948	Reduced interaction contribution
0.10	0.648	57,123	Chosen value (best validation stability)
0.20	0.674	54,995	Moderate increase in interaction contribution
0.50	0.642	57,633	High interaction weighting

**Figure 6.** Fast feature importance (weight based) for BYPredictor. Features are ranked by normalized importance, with higher values indicating greater contribution to the model's predictions.

Time-Series Baseline Considerations

Sequence-based models (LSTM, GRU, Transformer) were not included in the primary comparison. The FAO dataset is structured as panel data: each crop-region pair contains only 11-24 annual observations, which is insufficient for learning meaningful temporal dependencies with sequence models. Applying such models would substantially reduce the already limited samples per entity and increase overfitting risk. Feed-forward neural networks and tree-based models are therefore more appropriate for this task.

Computational Considerations

BYPredictor required approximately 217.6 seconds to train, compared to 211.3 seconds for the Baseline DNN. In agricultural forecasting, where models are typically retrained once per growing season, this cost is acceptable. The EILL itself adds negligible computational overhead over a standard dense layer, involving only element-wise multiplication and a matrix product.

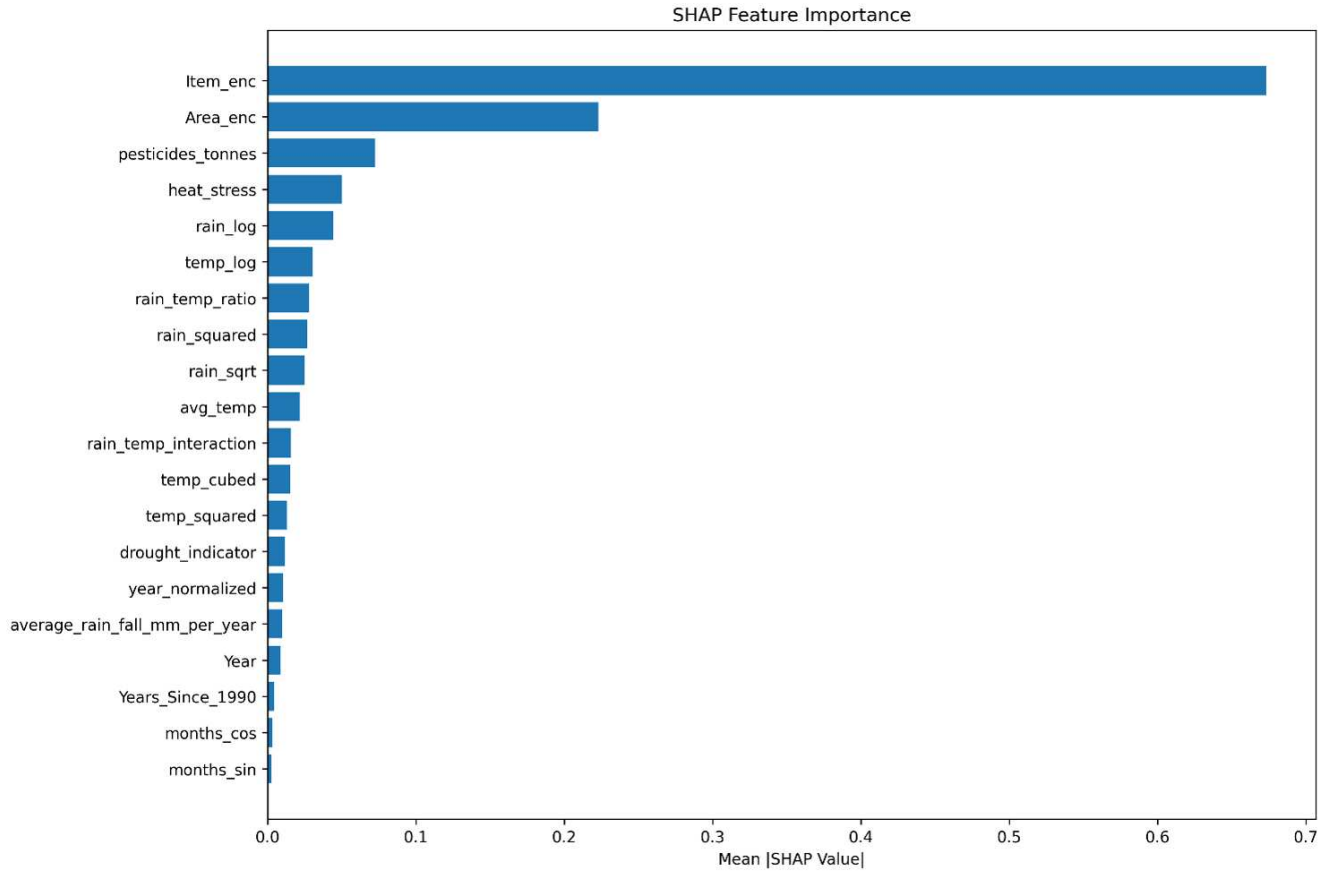


Figure 7. Permutation and SHAP-based feature importance for the Crop Yield Prediction dataset. Features are ranked by their contribution to model performance, with higher values indicating greater importance.

5 Limitations and Future Work

- **Simplified ecological interactions:** The EILL approximates ecological dynamics using a learnable linear interaction matrix with heuristic initialisation ranges. True agroecological systems involve highly non-linear and context-dependent interactions not fully captured by this representation. More sophisticated formulations (e.g. Lotka-Volterra-inspired neural ODEs) are a natural extension.
- **Crop-dependent EILL benefit:** The EILL improves performance on six crops but reduces it on four (notably sorghum and plantains). Future work should explore adaptive mechanisms that adjust interaction strength based on crop-specific characteristics.
- **Uncertainty quantification:** Monte Carlo Dropout uncertainty estimates produced poor empirical CI coverage. Conformal prediction or deep ensemble approaches should be explored for reliable interval estimates.
- **Missing economic and management variables:** The model does not incorporate market prices, government subsidies, irrigation data, or crop rotation practices, which can significantly affect yield.
- **Data-constrained crops:** Performance for crops with limited samples (potatoes, cassava) is lower. Few-shot learning and transfer learning may improve performance in data-sparse settings.

Future directions include integrating satellite imagery and soil sensor data, developing crop-adaptive EILL mechanisms, incorporating economic variables, and extending the evaluation to more recent FAO data.

6 Conclusion

This paper presented BYPredictor, a biologically inspired deep neural network for crop yield prediction that incorporates ecological interaction principles through the Ecological Interaction Learning Layer (EILL). The EILL models pairwise feature

interactions via a trainable interaction matrix initialised with ecologically motivated values, enabling the model to capture domain-relevant dependencies that standard neural layers do not explicitly represent.

Evaluated on the FAO Crop Yield dataset (28,242 samples, 10 crops), BYPredictor achieved $R^2 = 0.776$, RMSE = 45,605 hg/ha, MAE = 20,879 hg/ha, and MAPE = 22.4%, outperforming eight conventional baselines including Random Forest ($R^2 = 0.749$) and the Baseline DNN ($R^2 = 0.727$). The EILL contributed a statistically significant 4.3% improvement in R^2 ($p < 0.001$). Rolling time-series cross-validation yielded $R^2 = 0.942 \pm 0.046$, demonstrating robust temporal generalisation. BYPredictor achieved $R^2 \geq 0.7$ for eight out of ten crops, with particularly strong performance for wheat, maize, and rice. SHAP-based feature importance analysis identified crop type, geographic region, drought indicators, and non-linear climatic transformations as the primary predictors, providing interpretable insights that complement the model's predictive accuracy.

These results demonstrate that embedding biologically motivated interaction structures into deep learning architectures is an effective strategy for improving both accuracy and interpretability in agricultural yield forecasting. BYPredictor provides a principled, reproducible, and extensible foundation for future work in ecologically informed machine learning for food security.

7 Reproducibility and Code Availability

Execution Environment

- Operating System: Linux (Ubuntu 22.04 LTS)
- Processor: 4 cores, Intel Xeon (2.2 GHz)
- RAM: 31.4 GB

Software Dependencies

- Python 3.10.x
- TensorFlow 2.19.0
- NumPy 2.0.2
- Pandas 2.3.3
- Scikit-learn 1.3.2
- XGBoost 1.7.6
- SHAP 0.44.0

Dataset and Code Availability

The FAO Crop Yield Dataset used in this study, as well as the source code for the proposed BYPredictor model, are publicly available at the following links:

- Dataset: <https://www.kaggle.com/datasets/patelris/crop-yield-prediction-dataset/versions/1>
- Code: <https://github.com/abdulbaset2016/BYPredictor>

8 Acknowledgements (not compulsory)

The authors declare that no funding was received for this work. The authors would like to thank their respective institutions for providing the necessary research facilities and support.

9 Author contributions statement

A.M. conceived and designed the study, developed the BYPredictor model, conducted the experiments, analysed the data, and wrote the original manuscript. S.A. contributed to data collection, preprocessing, and validation of the results. All authors reviewed and approved the final manuscript.

10 Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Farah, A. A., Mohamed, M. A., Musse, O. S. H. & Nor, B. A. The multifaceted impact of climate change on agricultural productivity: a systematic literature review of scopus-indexed studies (2015–2024). *Discov. Sustain.* **6**, 397 (2025).
2. Iizumi, T. & Shin, Y. Launch of the global preharvest crop forecast of climate-induced yield variations. *Clim. Res.* **94**, 47–53 (2025).
3. Abbas, F., Afzaal, H., Farooque, A. A. & Tang, S. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy* **10**, 1046 (2020).
4. Meghraoui, K., Sebari, I., Pilz, J., Kadi, K. A. E. & Bensiali, S. Applied deep learning-based crop yield prediction: A systematic analysis of current developments and potential challenges. *Technologies* **12**, 43 (2024).
5. Kamilaris, A. & Prenafeta-Boldú, F. X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **147**, 70–90 (2018).
6. Paudel, D. *et al.* Interpretability of deep learning models for crop yield forecasting. *Comput. Electron. Agric.* **206**, 107663 (2023).
7. Elavarasan, D. & Vincent, P. M. D. Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. *IEEE Access* **8**, 86886–86901 (2020).
8. Okupska, E., Gozdowski, D., Pudelko, R. & Wójcik-Gront, E. Cereal and rapeseed yield forecast in poland at regional level using machine learning and classical statistical models. *Agriculture* **15**, 984 (2025).
9. Benos, L. *et al.* Machine learning in agriculture: A comprehensive updated review. *Sensors* **21**, 3758 (2021).
10. Ramesh, V. & Kumaresan, P. Stacked ensemble model for accurate crop yield prediction using machine learning techniques. *Environ. Res. Commun.* **7**, 035006 (2025).
11. Padma, T. & Dipali, S. Crop yield prediction using improved random forest. In *Proceedings of the 1st International Conference on Data Science and Advanced Computing (ICDSAC)*, vol. 56, 02007 (2023).
12. Demirhan, H. A deep learning framework for prediction of crop yield in australia under the impact of climate change. *Inf. Process. Agric.* **12**, 125–138 (2025).
13. Xu, J. *et al.* Northeastchinasoybeanyield20m: An annual soybean yield dataset at 20 m in northeast china (2019–2023). *Earth Syst. Sci. Data Discuss.* (2025). In press.
14. Mustafa, G. *et al.* Ecp-iem: Enhancing seasonal crop productivity with deep integrated models. *PLoS ONE* **20**, e0325815 (2025).
15. Guo, H., Tang, R., Ye, Y., Li, Z. & He, X. Deepfm: A factorization-machine based neural network for ctr prediction. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, 1725–1731 (2017).
16. Wang, R., Fu, B., Fu, G. & Wang, M. Deep & cross network for ad click predictions. In *Proceedings of the ADKDD'17*, Article No. 12 (2017).
17. Arik, S. Ö. & Pfister, T. Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, 6679–6687 (2021).
18. Gorishniy, Y., Rubachev, I., Khrulkov, V. & Babenko, A. Revisiting deep learning models for tabular data. In *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 34 (2021).
19. Pang, G., Lu, L. & Karniadakis, G. E. fpinns: Fractional physics-informed neural networks. *SIAM J. on Sci. Comput.* **41**, A2603–A2626 (2019).
20. Sajid, S. S., Huber, I., Archontoulis, S. & Hu, G. Integrating crop simulation and machine learning models to improve crop yield prediction. In *2022 17th Annual System of Systems Engineering Conference (SOSE)*, 120–125 (2022).
21. Mo, Z. & Wu, Y. Precision crop yield prediction model based on deep learning and multimodal data fusion. *Turkish J. Agric. For.* **49**, 580–597 (2025).
22. Lin, F. *et al.* An open and large-scale dataset for multi-modal climate change-aware crop yield predictions. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 5375–5386 (2024).
23. Seguini, L. *et al.* Global near real-time 500 m 10-day fpar dataset from modis and viirs for operational agricultural monitoring and crop yield forecasting. *Earth Syst. Sci. Data Discuss.* (2025). In press.
24. Filippi, P., Han, S. Y. & Bishop, T. F. On crop yield modelling, predicting, and forecasting and addressing the common issues in published studies. *Precis. Agric.* **26**, 8 (2024).

25. Food and Agriculture Organization (FAO), World Bank & Kaggle. Crop yield prediction dataset, data sources, and agricultural information. <https://www.kaggle.com/datasets/patelris/crop-yield-prediction-dataset/versions/1> (2024). Accessed: Jul. 3, 2024.