International Journal of Advances in Applied Computational Intelligence

♦ ISSN: 2833-5600 Vol: 08, No. 02, 2026 ♦ PP. 28–33

Article DOI: https://doi.org/10.54216/IJAACI.080204

FULL LENGTH ARTICLE

A Hybrid Computational Intelligence Framework Integrating XGBoost and Residual Neural Networks for Multi-Crop Yield Prediction

Yasser H. Elawady[1,*]

[1]Department of Artificial Intelligence, College of AI and Information, Horus University, Egypt

Email: yelawady@horus.edu.eg

Received: March 09, 2026 Revised: May 14, 2026 Accepted: July 14, 2026 ⋆ Corresponding author

ABSTRACT

Reliable crop-yield prediction is essential for strategic food planning, input allocation, and precision-agriculture analytics, yet a single predictive model must accommodate substantial heterogeneity across crop species, regions, seasons, and management conditions. This paper proposes a hybrid computational-intelligence framework that couples extreme gradient boosting with an ensemble of residual neural networks in order to model both dominant nonlinear structure and systematic residual error. A leakage-controlled benchmark was constructed from a recent public agricultural data release, yielding 2,698 observations spanning eight major crops, 30 Indian states and union territories, six agricultural seasons, and ten complete years from 2010 to 2019. The predictor space combines climatic and agronomic descriptors, including annual rainfall, cultivated area, fertilizer and pesticide intensities, crop identity, state, season, temporal trend, and rainfall–management interaction terms; the production variable was removed to prevent algebraic target leakage. The empirical study compares the proposed framework with nine competitive baselines under a strict chronological design in which 2010–2017 are used for training, 2018 for model selection, and 2019 for independent testing. On the holdout year, the proposed hybrid attains $R^2 = 0.974$, RMSE $= 3.146$ t ha^{-1}, and MAE $= 1.080$ t ha^{-1}, outperforming standalone XGBoost by 20.3% in RMSE and a direct deep neural network by 32.5%. Bootstrap analysis places the hybrid RMSE within 2.042–4.112 t ha^{-1} at the 95% confidence level. Grouped permutation analysis shows that crop identity, geographic location, and cultivated area contribute most strongly to predictive performance. Overall, the results support residual error correction as an effective strategy for strengthening tabular yield prediction while also underscoring the importance of leakage control, temporal validation, and competitive baseline design.

Keywords: Crop yield prediction ▪ Computational intelligence ▪ Extreme gradient boosting ▪ Deep neural networks ▪ Residual learning ▪ Precision agriculture ▪ Ensemble learning

1. INTRODUCTION

Crop yield estimates influence food-security assessment, market planning, insurance, irrigation, fertilizer allocation, and public policy. Their estimation is nevertheless challenging because the response arises from interactions among climate, soil, crop genetics, management practices, geography, and time. Machine-learning models can represent nonlinearities that are difficult to prescribe analytically, but their apparent accuracy is sensitive to data leakage, random rather than temporal validation, limited crop diversity, and weak baselines [1, 2, 3, 4].

Recent studies have demonstrated the value of convolutional, recurrent, boosted-tree, and ensemble architectures for re-

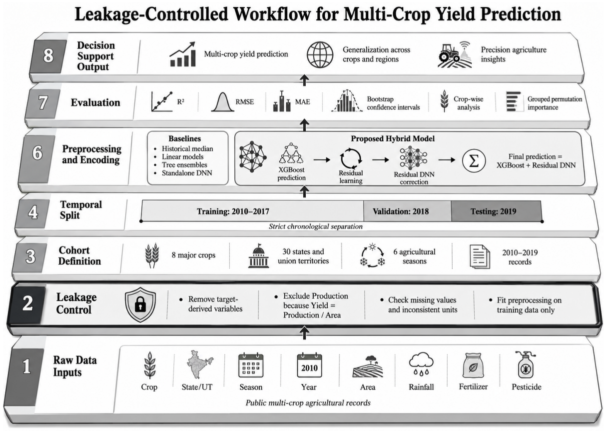

Figure 1. Three-dimensional leakage-controlled workflow for multi-crop yield prediction. The layered diagram summarizes raw agricultural inputs, leakage prevention, cohort construction, chronological partitioning, preprocessing, baseline comparison, residual-fusion modeling, evaluation, and decision-support outputs.

gional yield forecasting [5, 6, 7, 8, 9]. Deep networks can learn distributed representations, whereas gradient-boosted trees are effective on heterogeneous tabular predictors and limited sample sizes. A direct combination, however, may duplicate errors unless the components are assigned complementary roles. The present work therefore treats XGBoost as the primary learner and trains neural networks only on out-of-fold residuals. This design asks the neural component to model remaining structure rather than relearn the complete mapping.

The study makes five contributions. First, it constructs a leakage-controlled multi-crop benchmark from a recently released public corpus [10, 11]. Second, it introduces an out-of-fold XGBoost–residual-DNN architecture in which the neural component is restricted to systematic error correction rather than redundant end-to-end relearning. Third, it evaluates ten predictors under a strict chronological split and reports bootstrap uncertainty rather than relying on a single point estimate. Fourth, it compares the proposed model against a deliberately strong structure-aware historical baseline in addition to standard machine-learning competitors. Fifth, it interprets the final model through grouped permutation tests and a localized response-surface audit. Figure 1 summarizes the complete experimental sequence.

2. RELATED WORK

2.1 Machine learning for crop-yield estimation

Systematic reviews consistently identify weather, soil, vegetation, management, and historical yield as the most common input families [1, 12, 13, 14, 15]. Tree ensembles are widely used because they accommodate nonlinear thresholds and mixed predictors, while support-vector machines and neural networks remain competitive when sample construction and regularization are appropriate [16, 17, 18, 19, 20].

Studies in the U.S. Corn Belt have shown that ensemble methods and crop-model coupling improve corn and soybean forecasts [6, 7]. CNN–RNN and CNN–DNN systems have further demonstrated that deep architectures can exploit spatial and temporal dependencies [5, 8]. Nevertheless, these systems often target one or two crops and exploit dense time-series inputs. The present study addresses a different setting: heterogeneous annual tabular observations spanning eight

Table 1. Chronological data partitions and coverage.

Partition	Years	Records
Training	2010–2017	2,137
Validation	2018	282
Independent test	2019	279
Total	2010–2019	2,698
30 states/UTs; 8 crops; 6 seasons.		

crops and 30 administrative regions.

2.2 Climatic, agronomic, and interaction effects

Yield responses are not additive in general. Rainfall can alter the effectiveness or risk profile of fertilizer and pesticide applications, and geographic effects may proxy soil, cultivar, and management differences. Interaction regression has therefore been proposed for crop-yield modeling [21], while recent work explicitly combines meteorological and pesticide information [22]. Climate–vegetation fusion also improves crop-specific forecasts when remote-sensing data are available [23]. Because the selected corpus lacks within-season remote sensing, the current framework constructs rainfall–fertilizer and rainfall–pesticide interaction terms and interprets them as predictive, not causal, associations.

2.3 Evaluation gaps

Random cross-validation can mix future and past observations, leading to optimistic estimates when trends or repeated regional structures exist. Reviews have called for time-aware evaluation, explicit leakage checks, interpretable results, and comparisons against simple baselines [24, 9, 4]. We therefore reserve 2019 as an untouched test period and include a hierarchical historical-median predictor. This baseline is intentionally strong because crop–state–season strata exhibit persistent yield levels.

3. DATA AND FEATURE CONSTRUCTION

3.1 Public source and analytical subset

The source is Version 2 of the public dataset released by Ramesh and Kumaresan [10] in February 2025. The full corpus contains 19,689 observations from 1997–2020 and 55 crop categories. The analysis focuses on Rice, Wheat, Maize, Groundnut, Soyabean, Cotton (lint), Sugarcane, and Potato during 2010–2019. These crops provide contrasting yield scales and management regimes. After removing nonpositive yield or area values, 2,698 observations remain. Only ten selected-crop observations are present for 2020, so that incomplete tail is excluded rather than represented as a complete season.

Table 1 reports the chronological partitions. Figure 2 shows the annual and crop composition. The test period contains 279 observations distributed across all eight crops.

3.2 Leakage control and engineered variables

The source includes production, cultivated area, and yield. Because yield is computed from production and area, production is excluded from all predictors. Retaining it would permit the model to reconstruct the target algebraically and would invalidate the benchmark. For observation i, fertilizer and pesticide intensities are

$$f_i = \frac{F_i}{A_i}, \qquad p_i = \frac{P_i}{A_i}, \tag{1}$$

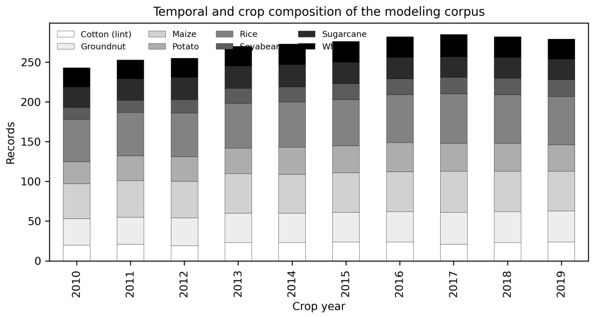

Figure 2. Temporal and crop composition of the modeling dataset. Counts remain comparatively stable across 2010–2019, whereas crop frequencies are unequal and motivate crop-stratified error reporting.

Table 2. Predictor groups used by all learning models.

Group	Variables and treatment
Identity	Crop, state/UT, and season; one-hot encoded
Climate	Annual rainfall in millimetres
Scale	$\log(1+\text{cultivated area})$
Management	Fertilizer/area and pesticide/area
Interactions	Rainfall $\times$ fertilizer rate; rainfall $\times$ pesticide rate
Time	Year index with 2010 as zero
Target	$z_i = \log(1+y_i)$; back-transformed for reporting

Equation (1) normalizes the two management inputs by cultivated area, where F_i, P_i, and A_i denote fertilizer, pesticide, and cultivated area. Area is transformed as $a_i = \log(1+A_i)$. Two cross-products, $r_i f_i$ and $r_i p_i$, encode first-order interactions with annual rainfall r_i. Table 2 defines the final inputs.

Missing numeric values, if encountered, are imputed with development-set medians; numeric variables are standardized and categorical variables are one-hot encoded. Every preprocessing parameter is estimated without the 2019 holdout.

4. PROPOSED HYBRID COMPUTATIONAL INTELLIGENCE FRAMEWORK

4.1 Gradient-boosted base learner

Let $x_i \in \mathbb{R}^d$ denote the transformed predictor vector and $z_i = \log(1+y_i)$. XGBoost represents the base prediction as an additive collection of regression trees,

$$\hat{z}_i^{(B)} = \sum_{m=1}^{M} h_m(x_i), \tag{2}$$

Equation (2) defines the additive tree predictor, where h_m is a tree. At boosting step m, the regularized objective is

$$\begin{aligned}\mathscr{L}^{(m)} &= \sum_i \ell\left(z_i, \hat{z}_i^{(m-1)} + h_m(x_i)\right) \\ &\quad + \gamma T_m + \frac{\lambda}{2}\|w_m\|_2^2 + \alpha\|w_m\|_1,\end{aligned} \tag{3}$$

Equation (3) balances prediction loss and tree complexity, with leaf count T_m and leaf weights w_m. The implementation uses 450 trees, depth five, learning rate 0.04, row and column subsampling of 0.85, and mixed L_1/L_2 regularization.

4.2 Out-of-fold residual learning

Training a residual network on in-sample XGBoost predictions would expose it to unrealistically small base errors. We instead partition the development data into five folds. For each fold, XGBoost is trained on the remaining folds and predicts the held-out fold. Concatenating these predictions gives $\hat{z}_i^{(B)}$, from which the residual target is

$$e_i = z_i - \hat{z}_i^{(B)}. \tag{4}$$

Equation (4) isolates errors produced under held-out prediction. The neural input augments the transformed features with the base prediction: $u_i = [x_i; \hat{z}_i^{(B)}]$. For layer ℓ,

$$q^{(\ell)} = \max\left(0, W^{(\ell)} q^{(\ell-1)} + b^{(\ell)}\right), \tag{5}$$

Equation (5) is applied using hidden widths 64, 32, and 16. The residual network is trained by minimizing

$$\mathscr{J}(\theta) = \frac{1}{N}\sum_{i=1}^{N}(e_i - g_\theta(u_i))^2 + \lambda_\theta \sum_{\ell}\|W^{(\ell)}\|_F^2, \tag{6}$$

where Equation (6) combines mean-squared residual fitting with weight decay to control variance. Five networks with different random seeds are trained with Adam, early stopping, and L_2 regularization. Their average residual estimate is $\hat{e}(x) = S^{-1}\sum_{s=1}^{S} g_s(u)$, where $S = 5$.

The final hybrid prediction is

$$\hat{y} = \max\left\{\exp\left[\hat{z}^{(B)} + \hat{e}(x)\right] - 1, 0\right\}. \tag{7}$$

Thus, the neural ensemble corrects only residual structure that is reproducible across out-of-fold predictions. Algorithm 1 states the training procedure.

Algorithm 1 Out-of-fold XGBoost–residual-DNN training

Require: Development data $\mathscr{D}$, folds $K = 5$, neural seeds $S = 5$
Fit preprocessing on $\mathscr{D}$ and obtain X, z
Initialize out-of-fold vector $\hat{z}^{(B)}$
for each fold k **do**
  Fit XGBoost on $\mathscr{D} \setminus \mathscr{D}_k$
  Predict $\mathscr{D}_k$ and fill $\hat{z}_k^{(B)}$
end for
Compute residuals $e = z - \hat{z}^{(B)}$
for $s = 1, \ldots, S$ **do**
  Fit DNN g_s to $[X; \hat{z}^{(B)}] \mapsto e$
end for
Refit XGBoost on all of $\mathscr{D}$
Predict by Equation (7)

5. EXPERIMENTAL DESIGN

5.1 Benchmark models and tuning

Table 3 summarizes the principal benchmark configurations.

Table 3. Benchmark configurations. All stochastic models use fixed seeds.

Model	Configuration
Historical median	Crop–state–season median with fallback to crop–state, crop, and global medians
Ridge / KNN / SVR	Ridge $\alpha = 10$; KNN $k = 10$ with distance weighting; RBF-SVR $C = 15$, $\varepsilon = 0.03$
Random forest	400 trees, minimum leaf size 2, feature fraction 0.85
AdaBoost	300 depth-4 regression trees, learning rate 0.03
Gradient boosting	300 trees, depth 3, Huber loss, learning rate 0.03
XGBoost	450 trees, depth 5, learning rate 0.04, row/column fractions 0.85, regularized squared-error objective
Direct DNN	Five 64–32–16 ReLU networks trained directly on log yield; averaged prediction
Proposed hybrid	Final XGBoost plus five neural residual correctors trained from five-fold out-of-fold errors

The benchmark contains a hierarchical historical median, ridge regression, k-nearest neighbors, support-vector regres-sion, random forest, AdaBoost, gradient boosting, XGBoost, and a direct DNN ensemble. Together with the proposed model, ten predictors are evaluated. Hyperparameters were selected using the 2018 validation season and then fixed before refitting on 2010–2018.

5.2 Metrics and statistical analysis

Performance is measured on the original yield scale. For n test observations,

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2}, \tag{8}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i-\widehat{y}_i|, \tag{9}$$

$$R^2 = 1-\frac{\sum_i(y_i-\widehat{y}_i)^2}{\sum_i(y_i-\bar{y})^2}. \tag{10}$$

The three principal criteria are defined jointly in Equation (10). Normalized RMSE is RMSE divided by the test-set mean. Percentile confidence intervals use 2,000 bootstrap resamples. A paired 5,000-resample bootstrap compares test RMSE of the hybrid and XGBoost; positive differences favor the hybrid. In addition, the median absolute error was inspected during model selection to avoid choosing a model whose performance depended excessively on a few extreme observations, although the main paper reports RMSE, MAE, and R^2 for comparability with prior studies. Grouped permutation importance jointly shuffles variables that share engineered interactions to avoid leaving an unpermuted proxy of the same information.

6. RESULTS

6.1 Overall predictive performance

Table 4 and Figure 3 show the independent 2019 results. The proposed hybrid ranks first with $R^2 = 0.974$, RMSE 3.146 t ha^{-1}, and MAE 1.080 t ha^{-1}. Standalone XGBoost obtains RMSE 3.950, so residual correction reduces RMSE by 20.3%. The direct DNN reaches RMSE 4.662, 32.5% above the hybrid. The hierarchical median is unexpectedly competitive (RMSE 3.389), indicating substantial persistence within crop–state–season strata.

Table 4. Independent 2019 test performance, ordered by RMSE.

Model	R^2	RMSE	MAE	NRMSE	Log-RMSE
Proposed XGB–DNN hybrid	**0.974**	**3.146**	**1.080**	**0.350**	0.192
Historical median	0.970	3.389	1.139	0.377	0.212
XGBoost	0.959	3.950	1.327	0.440	0.196
Random forest	0.946	4.548	1.319	0.507	**0.190**
Deep neural network	0.943	4.662	1.745	0.519	0.220
Gradient boosting	0.893	6.388	2.175	0.711	0.240
Support-vector regression	0.842	7.758	2.600	0.864	0.289
Ridge	0.727	10.183	3.367	1.134	0.329
AdaBoost	0.654	11.470	3.891	1.277	0.370
K-nearest neighbors	0.491	13.921	4.426	1.550	0.468

The hybrid's 95% bootstrap intervals are 0.957–0.987 for R^2, 2.042–4.112 t ha^{-1} for RMSE, and 0.750–1.438 t ha^{-1} for MAE. The observed RMSE difference relative to XGBoost is 0.803 t ha^{-1}. Across 5,000 paired bootstrap resamples, the mean reduction is 0.784 t ha^{-1}, with a two-sided percentile interval of -0.099 to 1.644 t ha^{-1} and a one-sided non-improvement tail probability of 0.044. Because the two-sided interval includes zero, this comparison provides suggestive directional evidence rather than definitive evidence of superiority. Even so, the direction of improvement is consistent across all three principal metrics and is accompanied by a visibly tighter alignment with the identity line in Figure 4, which indicates that residual correction improves accuracy

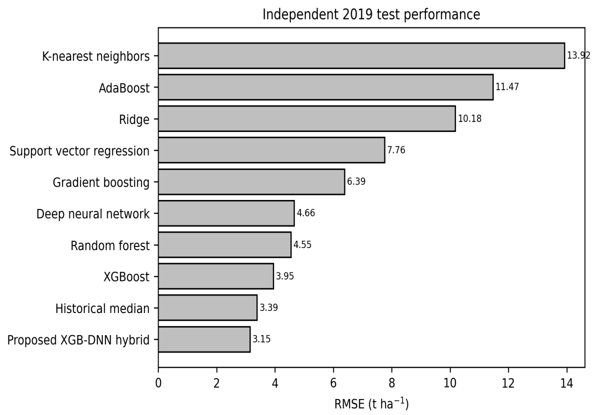

Figure 3. RMSE of the ten predictors on the independent 2019 holdout. The proposed residual-fusion model produces the lowest error; the structure-aware historical median is the second-best comparator.

without destabilizing the prediction range.

Figure 4 compares observations and hybrid predictions. Most values lie near the identity line. Larger absolute deviations occur at high sugarcane yields because that crop operates on a much wider numerical scale than grains and oilseeds.

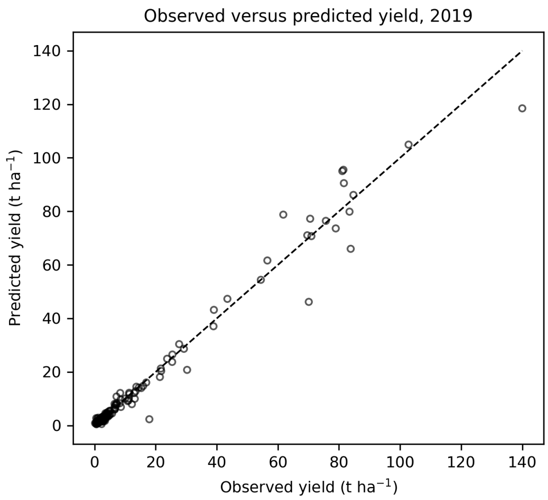

Figure 4. Observed and hybrid-predicted yield for the 2019 holdout. The dashed line represents perfect agreement.

6.2 Crop-specific generalization

Table 5 and Figure 5 report errors by crop. Wheat, potato, maize, and sugarcane achieve R^2 between 0.896 and 0.935. Cotton and soyabean have lower within-crop R^2, partly because their test subsets contain only 24 and 21 observations and have narrower yield variance. Sugarcane has the largest RMSE (10.013 t ha^{-1}) but also a strong $R^2 = 0.900$, reflecting its much larger absolute yield scale.

Table 5. Crop-wise performance of the proposed hybrid.

Crop	n	R^2	RMSE	MAE
Cotton (lint)	24	0.362	0.937	0.676
Groundnut	39	0.808	0.325	0.239
Maize	50	0.896	0.519	0.397
Potato	33	0.935	1.751	1.358
Rice	61	0.720	0.406	0.292
Soyabean	21	0.344	0.316	0.248
Sugarcane	26	0.900	10.013	6.962
Wheat	25	0.902	0.360	0.282

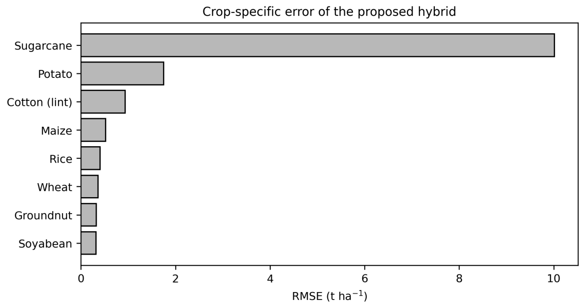

Figure 5. Crop-specific RMSE. Absolute error must be interpreted together with crop yield scale; sugarcane has the largest numerical range.

6.3 Feature-group interpretation

Grouped permutation results appear in Figure 6. Shuffling crop identity increases RMSE by 22.65 on average, followed by state (7.98) and cultivated area (3.76). Season, pesticide intensity, rainfall/interactions, and fertilizer intensity contribute smaller incremental signals. The year index has negligible holdout importance after crop, state, and season are known. These values quantify predictive reliance and must not be interpreted as causal effects.

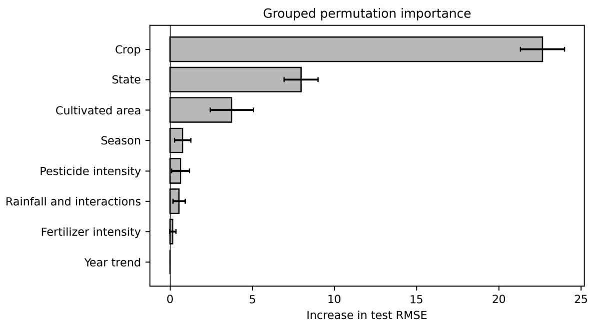

Figure 6. Grouped permutation importance for the hybrid model, expressed as the increase in 2019 test RMSE. Error bars show variation over 30 permutations.

Figure 7 provides a local model-response audit for a representative Uttar Pradesh Kharif rice case. The surface varies rainfall and fertilizer intensity while holding the remaining inputs fixed. It visualizes the model's learned association and should not be used to prescribe fertilizer because the observational corpus does not establish intervention effects.

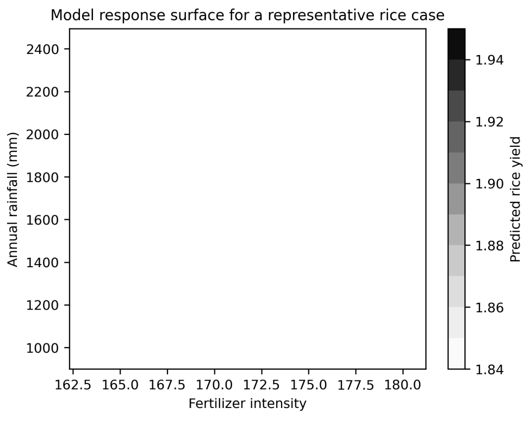

Figure 7. Associational response surface for a representative rice record. Grayscale levels denote predicted yield as rainfall and fertilizer intensity vary.

7. DISCUSSION

The empirical results support residual fusion rather than indiscriminate model stacking. XGBoost handles nonlinear partitions among crops, states, and management intensities, while the neural ensemble focuses on systematic residual patterns estimated from out-of-fold predictions. This division of labor produces a lower original-scale RMSE than either component alone. Related ensemble studies have also reported gains from complementary learners [6, 8, 11]; the present contribution differs by explicitly training the DNN on leakage-resistant residual targets.

The strong historical-median baseline is scientifically informative. A model that knows crop, state, and season already captures stable production regimes. Consequently, a random split could reward memorization of these strata. The chronological holdout reduces this risk, and the hybrid records a lower point-estimate error than the persistence baseline. At the same time, the paired uncertainty interval cautions against claiming universal superiority from one test year.

The dominance of crop and location is compatible with prior findings that crop type, agro-ecological environment, and management context govern yield variability [21, 24, 9]. Rainfall and input intensities contribute less after these identities are known, but their interactions remain relevant. The result does not imply that climate is agronomically unimportant; annual rainfall is a coarse summary and cannot represent within-season timing, heat stress, soil moisture, or extreme events.

From an applied computational-intelligence perspective, the framework has three practical advantages. It works with heterogeneous tabular data, accommodates unseen category levels through robust encoding, and produces a modular correction term that can be monitored separately from the base model. Updating can therefore be staged: the tree model may be retrained when structural conditions change, whereas the residual network may be recalibrated more frequently when modest systematic errors appear. A second practical advantage is methodological transparency. Because the hybrid is constructed as an additive correction to a well-understood gradient-boosted base model, performance gains can be traced to identifiable residual structure rather than to an opaque increase in model size alone. Third, the leakage-control workflow in Figure 1 makes the benchmark portable to other agricultural or environmental tabular datasets, thereby improving reproducibility and reducing the risk of inflated performance claims.

8. LIMITATIONS AND FUTURE WORK

The dataset is broad but not a field-level causal experiment. State-level aggregation can conceal local soil, cultivar, irrigation, and management differences. Fertilizer and pesticide rates are derived from total inputs and area and may contain reporting error. Annual rainfall does not preserve phenological timing. The 2019 holdout provides temporal separation but only one independent year, and the small crop-specific test groups lead to uncertain within-crop R^2 estimates. Moreover, the corpus ends before recent climate anomalies despite its 2025 publication date.

Future research should evaluate rolling-origin tests, district- or field-level measurements, remote-sensing time series, soil properties, temperature extremes, and calibrated predictive intervals. Domain adaptation across regions and multi-task losses that balance high- and low-yield crops are also promising. Physics- or crop-model-informed constraints could further improve extrapolation, provided they are evaluated against simple and structure-aware baselines rather than only against weak learners.

9. CONCLUSION

This paper presented a hybrid computational intelligence framework that combines XGBoost with an ensemble of residual neural networks for multi-crop yield prediction. Using 2,698 leakage-controlled observations across eight crops and 30 Indian states and union territories, the method was evaluated with 2018 validation and an untouched 2019 test period. It achieved $R^2 = 0.974$, RMSE 3.146 t ha^{-1}, and MAE 1.080 t ha^{-1}, recording the lowest point-estimate RMSE among the ten evaluated models and reducing XGBoost RMSE by 20.3%. Crop identity, state, and cultivated area were the principal predictive factors. Beyond the point estimates, the study shows that stronger publication-quality evidence in agricultural prediction requires three safeguards: leakage-resistant feature design, temporally separated validation, and comparison against both advanced and structure-aware baselines. Within this design, out-of-fold residual learning emerges as a credible and practically useful strategy for combining boosted trees and neural networks. The framework therefore provides a reproducible foundation for applied computational-intelligence research in precision agriculture and offers a defensible template for future extensions using richer weather, soil, remote-sensing, and field-scale management data.

REFERENCES

[1] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

[2] H. Burdett and C. Wellen, "Statistical and machine learning methods for crop yield prediction in the context of precision agriculture," *Precision Agriculture*, vol. 23, pp. 1553–1574, 2022.

[3] G. Lischeid, H. Webber, M. Sommer, C. Nendel, and F. Ewert, "Machine learning in crop yield modelling: A powerful tool, but no surrogate for science," *Agricultural and Forest Meteorology*, vol. 312, p. 108698, 2022.

[4] M. A. Jabed and M. A. A. Murad, "Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability," *Heliyon*, vol. 10, no. 24, p. e40836, 2024.

[5] S. Khaki, L. Wang, and S. V. Archontoulis, "A CNN-RNN framework for crop yield prediction," *Frontiers in Plant Science*, vol. 10, p. 1750, 2020.

[6] M. Shahhosseini, G. Hu, and S. V. Archontoulis, "Forecasting corn yield with machine learning ensembles," *Frontiers in Plant Science*, vol. 11, p. 1120, 2020.

[7] M. Shahhosseini, G. Hu, I. Huber, and S. V. Archontoulis, "Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt," *Scientific Reports*, vol. 11, p. 1606, 2021.

[8] M. Shahhosseini, G. Hu, S. Khaki, and S. V. Archontoulis, "Corn yield prediction with ensemble CNN-DNN," *Frontiers in Plant Science*, vol. 12, p. 709008, 2021.

[9] D. Paudel, A. de Wit, H. Boogaard, D. Marcos, S. Osinga, and I. N. Athanasiadis, "Interpretability of deep learning models for crop yield forecasting," *Computers and Electronics in Agriculture*, vol. 206, p. 107663, 2023.

[10] V. Ramesh and P. Kumaresan, "Stacked ensemble model for accurate crop yield prediction using machine learning techniques," Mendeley Data, 2025, version 2.

[11] V. Ramesh and P. Kumaresan, "Stacked ensemble model for accurate crop yield prediction using machine learning techniques," *Environmental Research Communications*, vol. 7, no. 3, p. 035006, 2025.

[12] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction," *IEEE Access*, vol. 9, pp. 63 406–63 439, 2021.

[13] P. Muruganantham, S. Wibowo, S. Grandhi, N. H. Samrat, and N. Islam, "A systematic literature review on crop yield prediction with deep learning and remote sensing," *Remote Sensing*, vol. 14, no. 9, p. 1990, 2022.

[14] S. M. Shawon, F. B. Ema, A. K. Mahi, F. L. Niha, and H. T. Zubair, "Crop yield prediction using machine learning: An extensive and systematic literature review," *Smart Agricultural Technology*, vol. 10, p. 100718, 2025.

[15] S. Saha, O. D. Kucher, A. O. Utkina, and N. Y. Rebouh, "Precision agriculture for improving crop yield predictions: A literature review," *Frontiers in Agronomy*, vol. 7, p. 1566201, 2025.

[16] F. Abbas, H. Afzaal, A. A. Farooque, and S. Tang, "Crop yield prediction through proximal sensing and machine learning algorithms," *Agronomy*, vol. 10, no. 7, p. 1046, 2020.

[17] D. Elavarasan and P. M. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, vol. 8, pp. 86 886–86 901, 2020.

[18] K. Jhajharia, P. Mathur, S. Jain, and S. Nijhawan, "Crop yield prediction using machine learning and deep learning techniques," *Procedia Computer Science*, vol. 218, pp. 406–417, 2023.

[19] M. Kuradusenge, E. Hitimana, D. Hanyurwimfura, P. Rukundo, K. Mtonga, A. Mukasine, C. Uwitonze, J. Ngabonziza, and A. Uwamahoro, "Crop yield prediction using machine learning models: Case of Irish potato and maize," *Agriculture*, vol. 13, no. 1, p. 225, 2023.

[20] P. Sharma, P. Dadheech, N. Aneja, and S. Aneja, "Predicting agriculture yields based on machine learning using regression and deep learning," *IEEE Access*, vol. 11, pp. 111 255–111 264, 2023.

[21] J. Ansarifar, L. Wang, and S. V. Archontoulis, "An interaction regression model for crop yield prediction," *Scientific Reports*, vol. 11, p. 17754, 2021.

[22] M. J. Hoque, M. S. Islam, J. Uddin, M. A. Samad, B. Sainz De Abajo, D. L. Ramirez Vargas, and I. Ashraf, "Incorporating meteorological data and pesticide information to forecast crop yields using machine learning," *IEEE Access*, vol. 12, pp. 47 768–47 786, 2024.

[23] M. Ashfaq, I. Khan, A. Alzahrani, M. U. Tariq, H. Khan, and A. Ghani, "Accurate wheat yield prediction using machine learning and climate-NDVI data fusion," *IEEE Access*, vol. 12, pp. 40 947–40 961, 2024.

[24] A. Morales and F. J. Villalobos, "Using machine learning for crop yield prediction in the past or the future," *Frontiers in Plant Science*, vol. 14, p. 1128388, 2023.