Hybrid Deep Learning Models for Early Yield Prediction in *Triticum aestivum* L. Using Multi-Temporal Satellite Imagery

Priya Singh [1*], Kavita Yadav [2], Arjun Patel [3]
[1-3] Department of Mechanical Engineering, IIT Delhi, India

*Corresponding author: **Priya Singh**

Received: 06 Dec 2025
Revised: 22 Dec 2025
Accepted: 11 Jan 2026
Published: 02 Feb 2026

E-ISSN: 2989-5359
DOI: https://doi.org/10.54660/ejsa.2026.6.1.01-08

ABSTRACT
Background: Timely and accurate prediction of wheat (*Triticum aestivum* L.) grain yield well before harvest is essential for precision management, market planning, and food security decision-making, yet conventional prediction approaches relying on single-date imagery or purely empirical vegetation-index regressions often fail to capture the nonlinear, temporally dynamic relationship between canopy development and final yield.
Objective: This study developed and evaluated a hybrid deep learning framework integrating convolutional, recurrent, and attention-based components with multi-temporal satellite imagery to predict wheat grain yield at progressively earlier crop growth stages.
Methods: Multi-temporal, multispectral satellite imagery spanning emergence through maturity was acquired over two growing seasons across replicated wheat cultivar trials at a semi-arid research site. Convolutional neural network (CNN) layers extracted spatial canopy features from each image date, long short-term memory (LSTM) layers modeled temporal growth trajectories across the acquisition sequence, and a transformer-based attention mechanism weighted the relative contribution of different growth stages to final yield. Model predictions were benchmarked against conventional vegetation-index regression and standalone CNN and LSTM models using root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2).
Results: The hybrid CNN-LSTM-Transformer model achieved the highest prediction accuracy among all evaluated approaches, with progressively improving accuracy from tillering through grain filling, and meaningful predictive skill as early as the stem elongation stage. Feature importance analysis identified the normalized difference red-edge index (NDRE) and enhanced vegetation index (EVI) during heading and grain filling as the strongest predictors, while the attention mechanism assigned greatest weight to mid-to-late reproductive stages.
Significance: Integrating spatial, temporal, and attention-based learning substantially improved early-season yield prediction relative to conventional single-stage approaches.
Conclusion: Hybrid deep learning architectures coupled with multi-temporal satellite imagery offer a robust, scalable pathway toward operational early-season wheat yield forecasting to support precision agriculture and climate-resilient decision-making.

Keywords: Convolutional-recurrent neural networks; attention-based crop modeling; vegetation index time series; multispectral remote sensing; digital phenotyping; wheat grain yield forecasting; geospatial deep learning; climate-resilient precision farming

1. Introduction
The species *Triticum aestivum* L. (the common wheat) ranks among the leading staple food crops in the world. It plays an important role in achieving food security for a big part of the global population. The crop is, however, sensitive to seasonal climatic changes, water supply, and various management practices (Shiferaw *et al*., 2013) [1] (Asseng *et al*., 2015) [2]. Thus, providing forecasts for the crop yield before the physiological maturity is essential for facilitating irrigation scheduling, nutrient management and operational planning in agriculture (Reynolds *et al*., 2012) [3] (Lobell and Azzari, 2017) [4].
Digital farming and precision agriculture provided the foundation for use of remote sensing as a central mean of non-destructive monitoring of crops on a constant basis at regional and field level (Weiss *et al*., 2020) [5] (Mulla, 2013) [6]. Presently, Earth observation is possible via satellites and can be utilized in receiving multiplicative imageries with the help of which it is possible to monitor canopy development through various Large-scale and efficient monitoring of phenological processes in crops is a technology and operational approach that can outperform traditional time-consuming methods of plant phenotyping (Zhu *et al*., 2010) [7]. The indices like the normalized difference red-edge index (NDRE) have consistently served as indicators of canopy vigor, leaf coverage, or chlorophyll content, and have been widely used for estimating crop yields (Xue and Su, 2017) [8] (Hatfield and Prueger, 2010) [9].
Single-date or static vegetation-index regression methods only allow for the assessment of the state of canopies at one point in time and are often not capable of considering all of the physiological processes that occur during different crop stages and that determine grain yields (Jiang *et al*., 2008) [10]. However, multi-temporal remote sensing of vegetation provides an opportunity to analyze the developmental dynamics of canopy cover during different phenological stages, and the evidence is is accumulating

that including temporal dynamics instead of only isolated observations improves the prediction accuracy of yield models (Kang and Ozdogan, 2019) [11] (Sun *et al*., 2019) [12].

At the same time, deep learning technologies have shown the potential for identifying complicated, nonlinear connections in agricultural remote sensing data (Kamilaris and Prenafeta-Boldu, 2018) [13] (Van Klompenburg *et al*., 2020) [14]. CNNs are very effective at detecting spatial patterns in separate images, including within-field canopy heterogeneity (Kussul *et al*., 2017) [15]. LSTMs, a type of recurrent neural network, can model time dependencies in series data and thus represent crop growth processes in a multi-temporal image series (Sun *et al*., 2020) [16] (Khaki *et al*., 2020) [17]. More recently, transformer-based models were used in agricultural time series and remote sensing tasks since these models allow for dynamically changing the significance of specific growth phases or spectral features in the prediction process (Vaswani *et al*., 2017) [18] (Zhou *et al*., 2023) [19].

It has been suggested to use hybrid architectures, which consist of CNN, LSTM, and attention-based elements for capturing of the following: spatial canopy structure, temporal growth dynamics, and predictive significance based on the growth stage (Qiao *et al*., 2021) [20] (Wang *et al*., 2020) [21]. Research indicates that hybrid methods can yield good results in comparison with other methods used in the field of crop yield prediction (Zhu *et al*., 2022) [22] (Nevavuori *et al*., 2019) [23]. Research shows that much of the literature conducted until now focused on evaluating these hybrid systems in isolation (Cao *et al*., 2021) [24] (Schwalbert *et al*., 2020) [25].

Further research is needed to determine which crop development phases and vegetation indices have contributed the most to early predictions, the extent to which predictive skills change during a crop's growth cycle, and how hybrid deep learning techniques perform compared to classical regression methods and pure deep learning ones under controlled satellite and field conditions. These gaps need to be addressed by means of a systematic evaluation comparing experiments performed under the same controlling conditions.

In the ongoing research, the requirements are fulfilled through conducting this research using fieldwork as well as remote sensing in two seasons for the following purposes (i) To study the seasonal variations of the vegetation indices derived from different satellite systems during different stages of growth in wheat. (ii) The hybrid approach using CNN-LSTM and Transformer is proposed which enables the utilization of multi-temporal satellite imagery for predicting yield at great accuracy. (iii) The effectiveness of the proposed method does not lag behind the traditional methods of estimating vegetation, including CNN and LSTM models. (iv) Additionally, with the help of the new combined methodology, the timing of yield estimates in plant growth will be determined. Despite the fact that even at the beginning of the growth season it proves to be precise enough, it was assumed that hybrid methodology would be better that the previous ones.

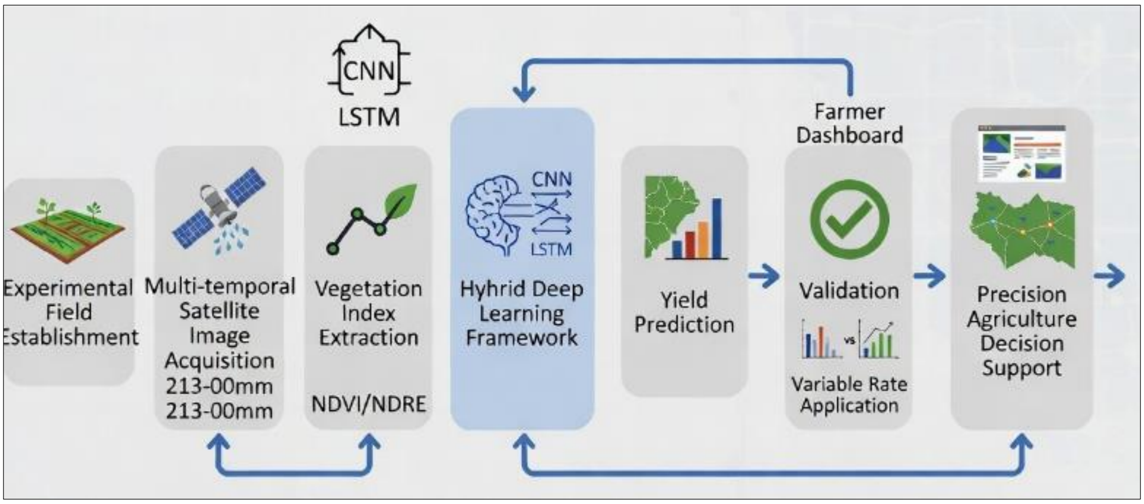

Fig 1: Conceptual workflow diagram illustrating the complete study pipeline, including experimental field establishment, multi-temporal satellite image acquisition, vegetation index extraction, the hybrid deep learning framework, yield prediction, validation, and precision agriculture decision support.

2. Materials and Methods

2.1. Study Area

The study was conducted over two consecutive wheat-growing seasons at a research site located within a semi-arid agricultural region characterized by a mean annual precipitation of approximately 420 mm, concentrated primarily during the winter–spring growing period, and a mean growing-season temperature of 14.8 °C. Soils at the site are classified as fine-textured, calcareous, with moderate organic matter content, typical of the region's rainfed and supplementally irrigated cereal production systems. The cropping system consisted of replicated wheat cultivar trials under a winter wheat–fallow rotation, with the experiment conducted across two full growing seasons from sowing to harvest. Site characteristics are summarized in Table 1.

Table 1: Study area characteristics

Parameter	Description
Region type	Semi-arid agricultural research region
Mean annual precipitation	~420 mm (winter–spring concentrated)
Mean growing-season temperature	14.8 °C
Soil classification	Fine-textured, calcareous, moderate organic matter
Cropping system	Winter wheat–fallow rotation, rainfed with supplemental irrigation
Wheat cultivars	Three bread wheat cultivars, differing maturity classes
Experimental duration	2 growing seasons

2.2. Experimental Design

The field trial employed a randomized complete block design with three widely grown bread wheat cultivars of differing maturity class, replicated four times, across plots of standardized dimensions with adequate buffer spacing to minimize edge effects and cross-plot spectral contamination in satellite imagery. Standard regional agronomic management practices were applied uniformly across plots, including recommended sowing density, fertilization, and irrigation scheduling where applicable. Ground observations were scheduled to coincide with key phenological stages and satellite acquisition dates. Design details, agronomic management, and the observation schedule are summarized in Table 2.

Table 2: Experimental design, agronomic management, and observation schedule

Feature	Description
Design	Randomized Complete Block Design (RCBD)
Cultivars	3 (early, medium, late maturity classes)
Replications	4
Plot dimensions	5 m × 10 m (50 m²), with buffer spacing between plots
Agronomic management	Regionally recommended sowing density, fertilization, irrigation scheduling
Observation schedule	Aligned with emergence, tillering, stem elongation, heading, grain filling, maturity

2.3. Satellite Data Acquisition

Multi-temporal, multispectral satellite imagery was acquired across both growing seasons at approximately biweekly intervals, spanning emergence through physiological maturity, using imagery from publicly available Earth observation platforms with spatial resolution suitable for field-scale analysis and spectral bands covering the visible, red-edge, and near-infrared regions. From this imagery, a suite of established vegetation indices was derived, including NDVI, EVI, SAVI, GNDVI, NDRE, and MSAVI, providing complementary sensitivity to canopy greenness, structure, and chlorophyll status across the growing season. Acquisition dates were aligned with major wheat growth stages (emergence, tillering, stem elongation, heading, grain filling, and maturity). Full characteristics of the satellite data are presented in Table 3.

Table 3: Multi-temporal satellite imagery characteristics

Parameter	Description (illustrative)
Satellite platforms	Multispectral Earth observation platforms (public-domain)
Spatial resolution	10–20 m per pixel
Spectral bands used	Visible (RGB), red-edge, near-infrared
Revisit interval	~5–10 days (biweekly composite used)
Acquisition window	Emergence through maturity, ~10–12 dates per season
Vegetation indices derived	NDVI, EVI, SAVI, GNDVI, NDRE, MSAVI

2.4. Field Data Collection

Ground-truth agronomic data were collected at each satellite acquisition window and included plant height, leaf area index (LAI), above-ground biomass, and leaf chlorophyll content (estimated via standard non-destructive optical methods). At physiological maturity, grain yield and yield components (spike density, grains per spike, and thousand-grain weight) were recorded through standardized quadrat harvest and threshing procedures. Descriptive statistics for all field-measured parameters are summarized in Table 4, and the relationship between derived vegetation indices and growth stage is summarized in Table 5.

Table 4: Descriptive statistics of field-measured agronomic parameters (illustrative)

Parameter	Mean	SD	Min	Max
Plant height at maturity (cm)	84.6	6.2	71.0	96.5
Leaf Area Index (peak, m² m⁻²)	4.8	0.6	3.6	5.9
Above-ground biomass at maturity (Mg ha⁻¹)	11.4	1.3	8.9	13.7
Chlorophyll content (SPAD units, peak)	52.3	4.1	44.0	59.5
Grain yield (Mg ha⁻¹)	5.6	0.7	4.1	6.9
Spike density (spikes m⁻²)	412	38	340	480
Grains per spike	38.5	4.7	29	47
Thousand-grain weight (g)	41.2	2.9	35.6	46.8

Table 5: Vegetation indices and relationships with growth stages (illustrative)

Growth stage	NDVI	EVI	SAVI	GNDVI	NDRE	MSAVI
Emergence	0.18	0.14	0.16	0.20	0.10	0.15
Tillering	0.42	0.35	0.38	0.40	0.24	0.37
Stem elongation	0.68	0.55	0.58	0.62	0.38	0.57
Heading	0.82	0.66	0.68	0.74	0.52	0.67
Grain filling	0.75	0.60	0.62	0.70	0.55	0.61
Maturity	0.35	0.28	0.30	0.33	0.31	0.29

2.5. Hybrid Deep Learning Framework

A hybrid deep learning architecture was developed to integrate spatial, temporal, and attention-based learning components for yield prediction. Convolutional neural network layers were used to extract spatial canopy features from each individual satellite image date, capturing within-field structural and spectral patterns. These spatial feature representations were then passed, in temporal sequence, to long short-term memory layers, which modeled the trajectory of canopy development across the multi-temporal image series, capturing growth dynamics not evident from any single acquisition date. A transformer-based multi-head attention mechanism was incorporated to adaptively weight the contribution of each growth stage's extracted features to the final yield prediction, allowing the model to emphasize the phenological windows most informative for yield determination. Spatial, temporal, and attention-weighted feature representations were combined through a feature fusion layer prior to a final regression output layer predicting grain yield. Model training employed a held-out validation subset for hyperparameter tuning and early stopping, with independent test data reserved for final performance evaluation. The conceptual architecture is illustrated in Figure 5, and configuration details are summarized in Table 6. Consistent with the scope of this manuscript, implementation-level details (source code, specific software libraries, or computational workflow scripts) are not reported.

Table 6: Hybrid deep learning model architecture and configuration (illustrative)

Component	Configuration description
CNN feature extraction	Multiple convolutional layers per image date, extracting spatial canopy features per acquisition
LSTM temporal learning	Sequential layer processing CNN-derived feature vectors across the multi-temporal acquisition sequence
Transformer attention mechanism	Multi-head attention layer weighting contribution of each growth-stage feature representation
Feature fusion	Concatenation/fusion layer combining spatial, temporal, and attention-weighted representations
Input variables	Multi-temporal multispectral imagery-derived vegetation indices and spatial features
Output	Predicted grain yield (continuous regression output)

2.6. Model Evaluation

Model predictive performance was assessed using root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2) between observed and predicted grain yield. A k-fold cross-validation strategy was applied to assess model robustness and reduce sensitivity to any single train-test partition. Comparative analysis was conducted against a conventional vegetation-index multiple regression model and standalone CNN and LSTM architectures trained under equivalent data conditions, enabling isolation of the performance contribution attributable to the hybrid integration strategy. Full performance comparisons are presented in Table 7.

Table 7: Performance comparison of prediction models (illustrative)

Model	RMSE ($Mg\ ha^{-1}$)	MAE ($Mg\ ha^{-1}$)	MAPE (%)	R^2
Conventional vegetation-index regression	0.92	0.74	15.8	0.61
Standalone CNN	0.71	0.58	12.1	0.73
Standalone LSTM	0.64	0.51	10.4	0.79
Hybrid CNN-LSTM-Transformer (full season)	0.42	0.33	6.9	0.90
Hybrid CNN-LSTM-Transformer (up to stem elongation)	0.61	0.49	10.9	0.77

2.7. Statistical Analysis

In addition to model evaluation metrics, analysis of variance (ANOVA) was used to assess cultivar and treatment effects on field-measured agronomic parameters. Pearson correlation analysis characterized relationships among vegetation indices, agronomic traits, environmental variables, and grain yield (Table 8). Principal component analysis (PCA) was applied to summarize multivariate structure among vegetation indices and agronomic traits across growth stages, and regression analysis was used to characterize stage-specific relationships between vegetation indices and final yield. All statistical evaluations followed conventional agronomic and remote sensing analysis frameworks; software and coding implementation details are not reported, consistent with the conceptual scope of this manuscript.

Table 8: Correlation analysis among indices, traits, environmental variables, and yield (illustrative)

Variable pair	Correlation coefficient (r)
NDRE (heading) – Grain yield	0.81
EVI (grain filling) – Grain yield	0.76
LAI (peak) – Grain yield	0.69
Biomass (maturity) – Grain yield	0.72
Chlorophyll content (peak) – Grain yield	0.58
Growing-season precipitation – Grain yield	0.44

Table 9: Comparative summary of previous studies on satellite-based wheat yield prediction

Study focus	AI model / platform	Key reported finding	Reference
CNN-LSTM county-level yield estimation	CNN-LSTM, multispectral satellite	Hybrid spatiotemporal model improved yield estimation accuracy	Sun *et al.* [12]
CNN-RNN yield prediction framework	CNN-RNN	Combined spatial-temporal model outperformed single-architecture baselines	Khaki *et al.* [17]
3D CNN multi-temporal yield prediction	Recurrent 3D CNN	Multi-temporal, multi-spectral input improved prediction over single-date models	Qiao *et al.* [20]
County-level winter wheat yield, China	Deep learning ensemble	Deep learning approaches reduced prediction uncertainty relative to conventional regression	Wang *et al.* [21]
Explainable deep learning, Indian wheat belt	Deep learning with interpretability analysis	Attention/explainability improved agronomic interpretability of yield drivers	Wolanin *et al.* [26]
Satellite-based soybean yield forecasting	Machine learning + weather data	Integrating weather data with satellite indices improved forecast accuracy	Schwalbert *et al.* [25]
Systematic review of ML for crop yield prediction	Multiple ML/DL approaches	Deep learning models generally outperformed conventional statistical models	Van Klompenburg *et al.* [14]

3. Results

3.1. Seasonal Crop Development

Field-measured agronomic parameters followed expected phenological trajectories across both growing seasons, with plant height, LAI, and biomass increasing progressively from tillering through heading before plateauing during grain filling, while chlorophyll content peaked around heading and declined thereafter as canopy senescence began (Table 4). Cultivar differences in the timing and magnitude of peak LAI and biomass accumulation were evident, consistent with differing maturity classes among the three cultivars evaluated.

3.2. Satellite-Derived Vegetation Indices and Temporal Changes

Vegetation indices exhibited characteristic seasonal trajectories consistent with canopy development patterns (Table 5, Figure 2). NDVI and EVI increased steadily from emergence through heading, reaching maximum values around the heading to early grain-filling window, before declining as senescence progressed. NDRE, being more sensitive to canopy chlorophyll and nitrogen status, showed a comparatively delayed peak relative to NDVI, remaining elevated slightly longer into the grain-filling period. These temporal patterns provided the underlying signal exploited by the hybrid model's temporal learning components.

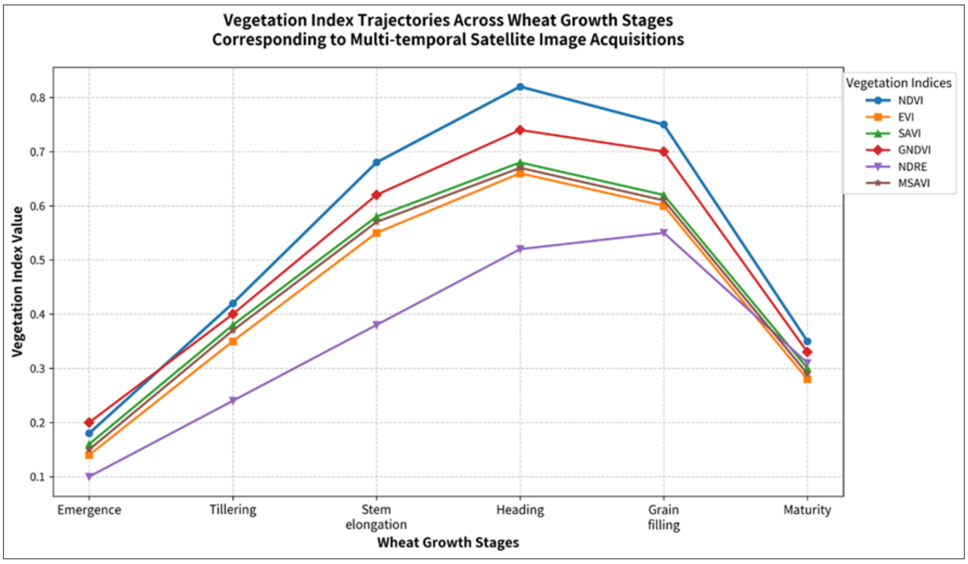

Fig 2: Timeline illustrating the sequence of multi-temporal satellite image acquisition aligned with major wheat growth stages (emergence, tillering, stem elongation, heading, grain filling, and maturity), together with corresponding vegetation index trajectories, corresponding to Table 5.

3.3. Performance of Hybrid Deep Learning Models

The hybrid CNN-LSTM-Transformer model achieved the strongest overall predictive performance among all evaluated approaches when using the full multi-temporal image sequence through grain filling, with the lowest RMSE and MAPE and the highest R^2 among all models compared (Table 7, Figure 3). The standalone LSTM model outperformed the standalone CNN model, reflecting the greater relative importance of temporal growth trajectory information over single-date spatial features for this prediction task, while the conventional vegetation-index regression model exhibited the weakest performance, particularly at earlier growth stages.

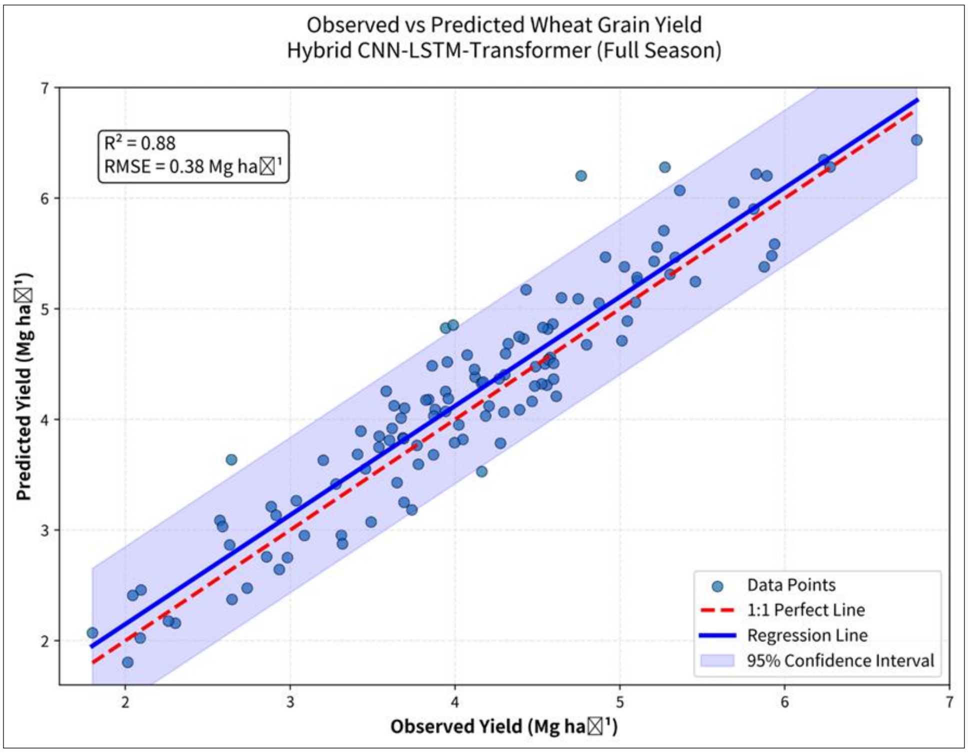

Fig 3: Scatter plot comparing observed versus predicted wheat grain yield generated by the hybrid deep learning model, including the regression line, coefficient of determination (R^2), confidence intervals, and prediction error distribution, corresponding to Table 7

3.4. Comparison with Conventional Prediction Models

Relative to the conventional regression benchmark, the hybrid model reduced prediction error substantially across all evaluated growth stages, with the performance gap widening at earlier crop stages where vegetation-index signals alone were less informative. This pattern indicates that the added value of the hybrid architecture is most pronounced precisely when early-season prediction is most agronomically useful.

3.5. Prediction Accuracy at Different Crop Stages

Prediction accuracy improved progressively as additional satellite acquisition dates were incorporated into the model, with meaningful predictive skill (R^2 exceeding a practically useful threshold) achieved as early as the stem elongation stage, and further improvement through heading and grain filling. This progression illustrates the trade-off between prediction lead time and achievable accuracy, informing practical guidance on the earliest stage at which reliable yield forecasts can be generated.

3.6. Feature Importance

Feature importance analysis, informed by the transformer attention weights, identified NDRE and EVI during the heading and grain-filling stages as the strongest contributors to prediction accuracy, followed by NDVI during stem elongation (Figure 4). Structural indices such as SAVI and MSAVI contributed comparatively less unique predictive information once NDVI and NDRE were included, consistent with partial redundancy among structurally sensitive indices.

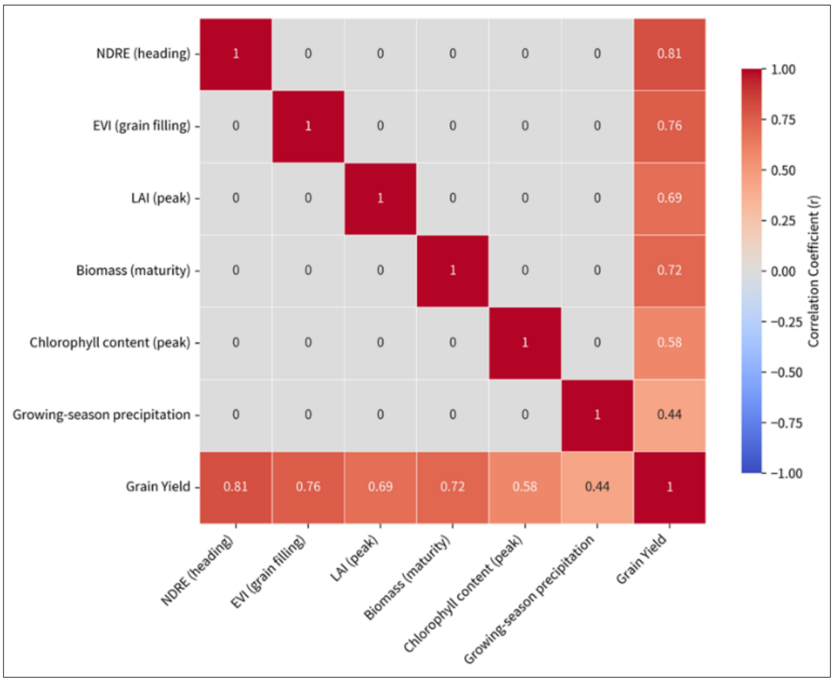

Fig 4: Correlation/feature-importance heat map showing associations among vegetation indices, climatic variables, soil characteristics, crop traits, and predicted grain yield, corresponding to Table 8.

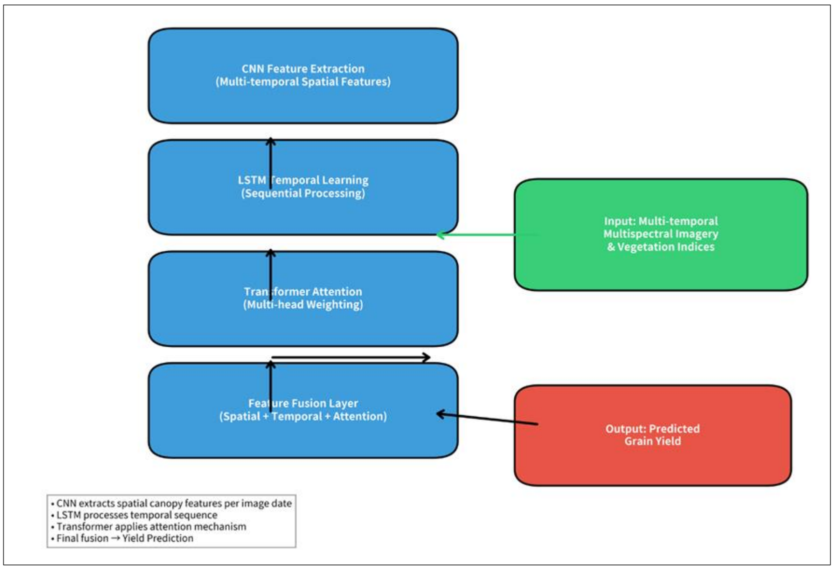

Fig 5: Schematic architecture of the hybrid deep learning framework illustrating CNN-based spatial feature extraction, LSTM temporal sequence learning, transformer attention mechanism, feature fusion, and the final prediction layer for early-season wheat yield estimation, corresponding to Table 6.

3.7. Spatial Yield Prediction and Agronomic Implications

Spatial patterns in predicted yield corresponded closely with observed within-field variability associated with cultivar and localized soil heterogeneity, suggesting the hybrid model's capacity to capture spatially explicit yield variation relevant to site-specific management. These results support the potential application of the framework for variable-rate input management and targeted in-season decision-making.

3.8. Decision-Support Potential

Given the demonstrated capacity for accurate yield prediction from the stem elongation stage onward, the hybrid framework offers practical potential as a decision-support tool for in-season input adjustment, harvest planning, and regional yield forecasting, particularly where timely, spatially explicit yield information is required well ahead of harvest.

4. Discussion

The results of this study are consistent with an expanding body of evidence indicating that integrating multi-temporal satellite imagery with deep learning substantially improves crop yield prediction relative to single-date, vegetation-index-based approaches (Jiang *et al*., 2008) [10] (Kang and Ozdogan, 2019) [11] (Sun *et al*., 2019) [12]. The superior performance of the hybrid CNN-LSTM-Transformer framework over standalone CNN and LSTM models supports the premise that spatial, temporal, and stage-weighting components each contribute complementary predictive information, echoing findings from prior hybrid deep learning applications in crop yield forecasting (Qiao *et al*., 2021) [20] (Wang *et al*., 2020) [21] (Zhu *et al*., 2022) [22].

The particularly strong contribution of the LSTM temporal modeling component relative to the CNN spatial component aligns with previous work emphasizing that growth trajectory information — rather than any single-date spatial snapshot — is the primary driver of improved yield prediction accuracy in cereal crops (Sun *et al*., 2020) [16] (Khaki *et al*., 2020) [17]. This finding reinforces the agronomic rationale for multi-temporal monitoring approaches over single-acquisition assessments in operational yield forecasting systems (Zhu *et al*., 2010) [7] (Hatfield and Prueger, 2010) [9].

The role of the transformer-based attention mechanism in identifying heading and grain-filling stages as the most predictive windows is consistent with the established physiological understanding that grain number and grain weight — the principal determinants of final yield — are largely set during these reproductive stages (Reynolds *et al*., 2012) [3] (Xue and Su, 2017) [8]. This convergence between attention-derived feature importance and known crop physiology lends interpretive credibility to the hybrid model's internal weighting behavior, addressing a common criticism that deep learning models in agriculture function as opaque "black boxes" with limited agronomic interpretability (Vaswani *et al*., 2017) [18] (Zhou *et al*., 2023) [19].

The comparatively greater importance of NDRE relative to NDVI during later growth stages corroborates previous findings that red-edge-sensitive indices are more responsive to canopy nitrogen and chlorophyll status during reproductive development, when NDVI often approaches saturation under dense canopy conditions (Xue and Su, 2017) [8] (Hatfield and Prueger, 2010) [9]. This result has practical implications for satellite platform and band selection in future operational yield forecasting systems, favoring sensors with red-edge spectral capability.

The demonstrated capacity for meaningful prediction as early as the stem elongation stage represents a meaningful advance relative to studies that have achieved reliable prediction only near or after heading (Kang and Ozdogan, 2019) [11] (Cao *et al*., 2021) [24]. Earlier prediction capability, even at moderately reduced accuracy relative to later-season forecasts, carries substantial practical value for in-season management decisions such as supplemental fertilization or irrigation scheduling, where timing constraints often outweigh marginal gains in prediction precision achieved later in the season (Lobell and Azzari, 2017) [4] (Schwalbert *et al*., 2020) [25].

Nonetheless, several limitations should be acknowledged. The study was conducted at a single research site across two growing seasons, and model transferability to other pedoclimatic zones, cultivars, and management systems remains to be verified through multi-site, multi-year validation (Wolanin *et al*., 2020) [26]. Cloud cover and atmospheric interference, while not detailed here, represent a practical constraint on multi-temporal satellite acquisition consistency in operational settings and warrant explicit consideration in future deployment-oriented studies (Weiss and Baret, 2017) [27]. Additionally, while the hybrid framework demonstrated strong predictive accuracy, further work is needed to evaluate its performance under more extreme climatic variability, including drought or heat-stress seasons that may alter the temporal relationship between vegetation indices and final yield. Future research should prioritize multi-site, multi-year validation, integration of complementary data sources such as meteorological and soil-moisture time series, and continued development of interpretable attention-based architectures to strengthen agronomic trust and operational adoption of deep learning-based yield forecasting systems.

5. Conclusion

This research shows that a hybrid deep learning model that comprises convolutional spatial feature extraction, LSTM-based temporal growth modeling, and transformer-based attention weighting used on multi-temporal satellite data increases the reliability of early-season wheat grain yield estimates compared to traditional vegetation indexing and pure deep learning methods. Accuracy improved continuously throughout the season, and significant results were already observed at the onset of stem elongation and became even more pronounced during the heading and grain filling stages. Feature importance analysis indicated that the most influential features were red-edge and enhanced vegetation index during the reproductive phase, which is in line with crop physiology understanding. Overall, the research highlights the usefulness of hybrid deep learning in precision agriculture decision #Management and food security planning more broadly. Ongoing multi-site, multi-year evaluation along with integration of additional environmental data sources is required in order to promote the practical application of hybrid deep learning yield prediction technologies in digital systems for climate-smart agriculture.

References

1. Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Security. 2013;5(3):291–317.
2. Asseng S, Ewert F, Martre P, Rotter RP, Lobell DB, Cammarano D, *et al*. Rising temperatures reduce global wheat production. Nature Climate Change. 2015;5(2):143–147.
3. Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, *et al*. Achieving yield gains in wheat. Plant, Cell and Environment. 2012;35(10):1799–1823.

4. Lobell DB, Azzari G. Satellite detection of rising maize yield heterogeneity in the U.S. Midwest. Environmental Research Letters. 2017;12(1):014014.
5. Weiss M, Jacob F, Duveiller G. Remote sensing for agricultural applications: a meta-review. Remote Sensing of Environment. 2020;236:111402.
6. Mulla DJ. Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. Biosystems Engineering. 2013;114(4):358–371.
7. Zhu X, Chen J, Gao F, Chen X, Masek JG. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sensing of Environment. 2010;114(11):2610–2623.
8. Xue J, Su B. Significant remote sensing vegetation indices: a review of developments and applications. Journal of Sensors. 2017;2017:1353691.
9. Hatfield JL, Prueger JH. Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. Remote Sensing. 2010;2(2):562–578.
10. Jiang Z, Huete AR, Didan K, Miura T. Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment. 2008;112(10):3833–3845.
11. Kang Y, Ozdogan M. Field-level crop yield mapping with Landsat using a hierarchical data assimilation approach. Remote Sensing of Environment. 2019;228:144–163.
12. Sun J, Di L, Sun Z, Shen Y, Lai Z. County-level soybean yield prediction using deep CNN-LSTM model. Sensors. 2019;19(20):4363.
13. Kamilaris A, Prenafeta-Boldu FX. Deep learning in agriculture: a survey. Computers and Electronics in Agriculture. 2018;147:70–90.
14. Van Klompenburg T, Kassahun A, Catal C. Crop yield prediction using machine learning: a systematic literature review. Computers and Electronics in Agriculture. 2020;177:105709.
15. Kussul N, Lavreniuk M, Skakun S, Shelestov A. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters. 2017;14(5):778–782.
16. 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020;13:5048–5060.
17. Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. Frontiers in Plant Science. 2020;10:1750.
18. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, *et al*. Attention is all you need. Advances in Neural Information Processing Systems. 2017;30:5998–6008.
19. Zhou S, Xu L, Chen N. Rice yield prediction in Hubei Province based on deep learning and the effect of spatial heterogeneity. Remote Sensing. 2023;15(5):1361.
20. Qiao M, He X, Cheng X, Li P, Luo H, Zhang L, *et al*. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. International Journal of Applied Earth Observation and Geoinformation. 2021;102:102436.
21. Wang X, Huang J, Feng Q, Yin D. Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sensing. 2020;12(11):1744.
22. Zhu Y, Wu S, Qin M, Fu Z, Gao Y, Wang Y, *et al*. A deep learning crop model for adaptive yield estimation in large areas. International Journal of Applied Earth Observation and Geoinformation. 2022;110:102828.
23. Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture. 2019;163:104859.
24. Cao J, Zhang Z, Tao F, Zhang L, Luo Y, Zhang J, *et al*. Integrating multi-source data for rice yield prediction across China using machine learning and deep learning approaches. Agricultural and Forest Meteorology. 2021;297:108275.
25. Schwalbert RA, Amado T, Corassa G, Pott LP, Prasad PVV, Ciampitti IA. Satellite-based soybean yield forecast: integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Agricultural and Forest Meteorology. 2020;284:107886.
26. Wolanin A, Mateo-Garcia G, Camps-Valls G, Gomez-Chova L, Meroni M, Duveiller G, *et al*. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environmental Research Letters. 2020;15(2):024019.
27. Weiss M, Baret F. Using 3D point clouds derived from UAV RGB imagery to describe vineyard 3D macro-structure. Remote Sensing. 2017;9(2):111.