Crop Yield Prediction Using Satellite Imagery, Weather Data, and Machine Learning: A Multi-Modal Framework for Precision Agriculture

Shravanakumari H J[1], Shinty P K[2]

[1]CMR University, School of Science and Computer Science, Bengaluru, Karnataka 560043, India
Email: *shravanakumari.hj[at]cmr.edu.in*

[2]CMR University, School of Science and Computer Science, Bengaluru, Karnataka 560043, India
Email: *shinty.p[at]cmr.edu.in*

Abstract: ***Accurate, early prediction of crop yield is one of the more stubborn problems in agricultural planning, food security policy, and commodity markets alike. Farmers, insurers, and governments all want the same thing at different scales: a reliable read on how much a field or a region is likely to produce, well before harvest. This paper presents a multi-modal machine learning framework that fuses multispectral satellite imagery with ground-level weather observations to forecast crop yield at the field level. Vegetation indices such as NDVI and EVI are extracted from Sentinel-2 and Landsat-8 imagery across the growing season and combined with temperature, rainfall, and soil-moisture time series. A hybrid architecture- a convolutional feature extractor for the imagery paired with a recurrent network for the temporal weather signal- is used to learn joint representations that a downstream regression head converts into yield estimates. On a multi-season wheat and maize dataset spanning three agro-climatic zones, the proposed model reaches an R^2 of 0.87 and reduces RMSE by roughly 18% compared to weather-only and imagery-only baselines. The results suggest that the two data sources are genuinely complementary rather than redundant, and that fusing them is worth the added engineering effort. We also discuss where the approach struggles - cloud cover, smallholder plot sizes, and limited ground-truth labels chief among them- and outline directions for making the system more field-ready.***

Keywords: crop yield prediction, remote sensing, satellite imagery, NDVI, weather data fusion, precision agriculture, convolutional neural network, LSTM, machine learning

1. Introduction

Predicting how much a crop will yield before it is harvested has always mattered- to a farmer deciding whether to book extra storage, to a bank underwriting a seasonal loan, to a ministry of agriculture trying to anticipate a shortfall months in advance. Traditionally this prediction has leaned on farmer-reported surveys, historical averages, and manual field surveys, all of which are slow, expensive to scale, and often arrive too late to be actionable.

Two developments have changed what is possible here. First, satellite constellations such as Sentinel-2, Landsat-8, and MODIS now provide frequent, free, multispectral imagery at resolutions fine enough to track the health of individual fields over a growing season. Second, machine learning models- particularly deep architectures capable of learning directly from raw imagery and time series- have matured to the point where they can extract useful signal from this data without heavy manual feature engineering. Weather data, meanwhile, has never really been optional: temperature stress, rainfall timing, and soil moisture are causally tied to yield, and any model that ignores them is working with one hand behind its back.

This paper works from a simple premise: imagery tells you what a crop looks like right now, and weather data tells you what conditions it has been growing under. Neither view is complete on its own. A field can look green and healthy in an image while sitting on the edge of a heat-stress event that will only show up in the yield a month later; conversely, favorable weather does not guarantee a good outcome if the crop is already stressed by disease or poor soil. We therefore propose a framework that fuses both signals rather than treating them as competing approaches, and we evaluate whether that fusion is actually worth the added complexity- a question the literature does not always test directly.

The remainder of this paper is organized as follows. Section II reviews related work in remote-sensing-based and weather-based yield prediction. Section III describes the proposed methodology, including data preprocessing and model architecture. Section IV details the dataset and experimental setup. Section V presents results and discussion. Section VI outlines limitations, and Section VII concludes with directions for future work.

2. Related Work

Early yield prediction models relied heavily on statistical regression over historical yield and weather records, an approach that works reasonably well at a coarse regional scale but tends to break down at the field level, where local variation in soil and management practices dominates. Vegetation indices derived from satellite imagery- most notably the Normalized Difference Vegetation Index (NDVI) - became a standard proxy for crop health once Landsat and later Sentinel-2 data became widely available, and a substantial body of work has shown correlation between season-long NDVI trajectories and final yield.

More recent work has moved toward deep learning. Convolutional neural networks have been applied directly to satellite image time series to predict yield without handcrafted indices, letting the network learn which spectral bands and spatial patterns matter. Separately, recurrent networks and, more recently, temporal attention models have been used to model the weather time series across a growing season, since yield outcomes depend on the sequence and timing of stress events, not just seasonal totals.

What is comparatively less explored is a systematic fusion of the two modalities within a single end-to-end model, evaluated against clean single-modality baselines on the same data. Several studies combine imagery and weather features but do so late, by concatenating hand-engineered summary statistics rather than learned representations, which limits how much the model can exploit interactions between the two signals- for instance, how a given NDVI trajectory should be interpreted differently under drought versus normal rainfall. This is the gap the present work targets.

3. Proposed Methodology

1) Overview

The proposed framework has three stages: (1) modality-specific preprocessing of satellite imagery and weather data, (2) a dual-branch neural network that learns separate representations for each modality, and (3) a fusion and regression head that combines the two representations to produce a final yield estimate. Fig. 1 (referenced conceptually here) summarizes the pipeline: imagery flows through a convolutional branch, weather time series flow through a recurrent branch, and the two output vectors are concatenated before the final regression layers.

2) Satellite Imagery Preprocessing

For each field boundary, Sentinel-2 Level-2A imagery is retrieved at roughly 10-day intervals across the growing season and clipped to the field polygon. Cloud-contaminated scenes are filtered using the scene classification layer, and remaining gaps are linearly interpolated across the time series. From each scene we compute NDVI, the Enhanced Vegetation Index (EVI), and the Normalized Difference Water Index (NDWI), stacked alongside the raw red, near-infrared, and short-wave infrared bands. The result is a per-field spatio-temporal tensor of shape ($T \times H \times W \times C$), where T is the number of time steps, H and W are spatial dimensions after resizing to a fixed patch size, and C is the number of channels.

3) Weather Data Preprocessing

Daily weather observations- minimum and maximum temperature, precipitation, relative humidity, and modeled soil moisture- are pulled from the nearest ground station or gridded reanalysis product and aligned to the same growing-season window as the imagery. Rather than only feeding raw daily values, we compute agronomically meaningful derived features: growing degree days, cumulative rainfall, and a simple heat-stress indicator counting days above a crop-specific temperature threshold. These derived features are concatenated to the raw daily series before being passed to the recurrent branch.

4) Model Architecture

The imagery branch uses a lightweight 3D convolutional stack (a small ResNet-style block applied across the temporal dimension) to extract a fixed-length embedding from the spatio-temporal tensor. The weather branch uses a two-layer bidirectional LSTM over the daily time series, with the final hidden state taken as the weather embedding. The two embeddings are concatenated and passed through two fully connected layers with dropout, ending in a single linear output for the predicted yield. The full network is trained end-to-end with mean squared error loss, using Adam with a learning rate of 1e-3 decayed on plateau, and early stopping based on validation RMSE.

We also trained two ablation variants- an imagery-only model (convolutional branch plus regression head) and a weather-only model (LSTM branch plus regression head)- to isolate how much the fusion itself contributes, rather than simply comparing against unrelated baselines from prior literature.

5) Data Set and Experimental Setup

The dataset covers three growing seasons of wheat and maize fields across three agro-climatic zones with differing rainfall and temperature regimes. Field-level yield labels were sourced from government agricultural statistics offices and cooperative harvest records, giving 4,200 labeled field-season observations after removing records with missing boundaries or incomplete imagery coverage. Fields were split 70/15/15 into training, validation, and test sets, stratified by zone and season so that the test set includes seasons and zones represented, but not overrepresented, in training.

All models were implemented in PyTorch and trained on a single GPU. Hyperparameters (learning rate, dropout rate, LSTM hidden size, convolutional channel widths) were tuned via grid search on the validation split. Performance is reported using root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2) on the held-out test set.

4. Results and Discussion

Table I summarizes test-set performance for the imagery-only, weather-only, and fused models. The fused model outperforms both single-modality baselines across every metric, with the largest relative improvement in RMSE. This supports the premise that imagery and weather data carry complementary rather than redundant information: imagery captures the realized state of crop health, while weather data provides context on the conditions producing that state, and the model appears to exploit both.

Table I: Test-Set Performance Comparison

Model	RMSE (t/ha)	MAE (t/ha)	R^2
Weather-only (LSTM)	0.71	0.55	0.74
Imagery-only (CNN)	0.64	0.49	0.79
Fused (proposed)	0.52	0.39	0.87

Breaking results down by growth stage, the advantage of fusion is largest mid-season, when vegetation indices alone are still ambiguous about whether the crop will finish strong, but recent rainfall and temperature trends already hint at the trajectory. Late-season predictions from all three models converge, which is expected: by then the imagery itself has largely absorbed the effect of the season's weather.

Performance also varies by zone. The semi-arid zone, where rainfall variability is the dominant driver of yield, benefits most from the weather branch; the more temperate, irrigated zone shows a smaller gap between imagery-only and fused models, since irrigation dampens the influence of rainfall timing. This zone-dependent pattern is a useful diagnostic in itself- it suggests that in practice, the relative weighting of the two modalities could reasonably be tuned per region rather than fixed globally.

5. Challenges and Limitations

A few limitations are worth stating plainly rather than glossing over. First, persistent cloud cover during the growing season- common in monsoon-affected regions- creates real gaps in the imagery time series that interpolation can only partially fix; performance degrades noticeably on fields with more than 30% cloud-affected scenes. Second, the dataset skews toward medium-to-large fields with clean boundary data; smallholder plots below roughly half a hectare are underrepresented, and satellite resolution becomes a genuine constraint at that scale. Third, ground-truth yield labels are themselves noisy, since cooperative harvest records are not always field-precise, which puts a ceiling on achievable accuracy regardless of model quality. Finally, the model has only been validated within the agro-climatic zones present in training; extending it to new regions or crop types would likely require additional fine-tuning rather than being a drop-in solution.

6. Conclusion and Future Work

This paper presented a fused deep learning framework for crop yield prediction that combines satellite-derived vegetation indices with weather time series, and showed on a multi-season, multi-zone dataset that the fusion meaningfully outperforms either data source used alone. The gain is not uniform- it is largest in rainfall-variable zones and mid-season, which is a useful, practically actionable finding rather than just an aggregate accuracy improvement.

Future work includes incorporating soil-type and management data (fertilizer application, irrigation schedules) as a third modality, testing transformer-based temporal encoders in place of the LSTM branch, and- most importantly for real-world deployment- evaluating the model on smallholder fields with lower-resolution imagery, where the practical need for accurate yield forecasting is arguably greatest. We also intend to explore uncertainty estimation alongside point predictions, since a calibrated confidence interval matters as much as the point estimate to a farmer or policymaker deciding how to act on it.

Acknowledgment
The authors thank the agricultural statistics offices and cooperative partners who provided field-level yield records used in this study.

References

[1] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," Comput. Electron. Agric., vol. 177, p. 105709, 2020.
[2] J. You, X. Li, M. Low, D. Lobell, and S. Ermon, "Deep Gaussian process for crop yield prediction based on remote sensing data," in Proc. AAAI Conf. Artif. Intell., 2017, pp. 4559–4565.
[3] A. Wolanin et al., "Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt," Environ. Res. Lett., vol. 15, no. 2, p. 024019, 2020.
[4] S. Khaki and L. Wang, "Crop yield prediction using deep neural networks," Front. Plant Sci., vol. 10, p. 621, 2019.
[5] M. Rußwurm and M. Körner, "Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images," in Proc. IEEE CVPR Workshops, 2017, pp. 1496–1504.
[6] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 5, pp. 778–782, 2017.
[7] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in Proc. ACM COMPASS, 2018, pp. 1–5.
[8] D. B. Lobell, "The use of satellite data for crop yield gap analysis," Field Crops Res., vol. 143, pp. 56–64, 2013.
[9] G. Azzari, M. Jain, and D. B. Lobell, "Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries," Remote Sens. Environ., vol. 202, pp. 129–141, 2017.
[10] J. Sun, L. Di, Z. Sun, Y. Shen, and Z. Lai, "County-level soybean yield prediction using deep CNN-LSTM model," Sensors, vol. 19, no. 20, p. 4363, 2019.
[11] E. Pekel, "Estimation of soil moisture using decision tree regression," Theor. Appl. Climatol., vol. 139, pp. 1111–1119, 2020.
[12] M. Shahhosseini, G. Hu, and S. V. Archontoulis, "Forecasting corn yield with machine learning ensembles," Front. Plant Sci., vol. 11, p. 1120, 2020.
[13] Copernicus / European Space Agency, "Sentinel-2 User Handbook," ESA Standard Document, 2015.
[14] USGS, "Landsat 8 Data Users Handbook," U.S. Geological Survey, 2019.
[15] A. Crane-Droesch, "Machine learning methods for crop yield prediction and climate change impact assessment in agriculture," Environ. Res. Lett., vol. 13, no. 11, p. 114003, 2018.