IJECRT

International Journal of Engineering Computational Research and Technology

Journal home page: www.ijecrt.org

ISSN (Online): 2456-9852 Volume 10, December 2025

HGTDL: A Hybrid Graph-Temporal Deep Learning Framework for Crop Yield Prediction Using Multi-Source Satellite and Agronomy Data

J.Amani [a, *], K.Sekar [b] , D Shobha Rani [c], Sathya Narayana Pola [d] , A Naresh kumar [e]

[a] *M.Tech Student , Department of Computer Science and Engineering, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, India*

[b] *Professor , Department of Computer Science and Engineering, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, India.*

[c] *Professor , Department of Computer Science and Engineering, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, India.*

[d] *M.Tech Student , Department of Computer Science and Engineering, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, Indiaee*

[e] *M.Tech Student , Department of Computer Science and Engineering, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, India.*

[*] Corresponding author
E-mail address: amanireddy154@gmail.com

ABSTRACT

Accurate crop yield prediction is essential for agricultural planning, supply chain management and risk-aware advisories, particularly in regions where yield variability is highly influenced by rainfall patterns and variety selection. Traditional machine learning methods are unable to fully capture the combined effect of phenological growth and agronomic suitability, resulting in limited accuracy in diverse field conditions. This study proposes a hybrid deep learning framework that integrates spatial similarity, temporal phenology and agronomy knowledge to improve crop yield prediction. The proposed HGTDL model combines a graph neural network to represent agro-ecological proximity across fields with a long short-term memory encoder to capture the seasonal evolution of vegetation indices derived from multi-temporal satellite imagery. Crop-specific attributes such as duration and recommended production zones from the Crop Recommendation Dataset [22] are embedded through a fusion layer. Experiments were conducted using dataset [22], with a 70-15-15 train–validation–test split. Performance was assessed using RMSE, MAE and R^2 scores. The model achieved an RMSE of **0.37 t/ha**, MAE of **0.28 t/ha** and **R^2 of 0.86**, outperforming Random Forest, XGBoost and CNN–LSTM baselines by **5–28%**. Stratified evaluation showed **6% higher accuracy** in recommended zones and stronger stability for medium-duration varieties. Statistical tests confirmed that improvements were significant at **$p < 0.05$**. The results show that combining spatial and temporal learning with agronomy metadata provides reliable yield forecasts across crop types and regions. The framework supports practical applications in crop advisory systems, procurement planning and insurance risk estimation.

Keywords: crop yield prediction, graph neural networks, LSTM, remote sensing, satellite imagery, agronomy data, hybrid deep learning, spatial–temporal modelling, vegetation indices.

1. Introduction

1.1 Background and Significance of the Research Problem

Agriculture continues to remain the backbone of many economies, especially in countries where farming is directly tied to food security, rural livelihoods, and national economic stability. In recent decades, the farming sector has come under increasing pressure due to extreme climate variability, irregular rainfall, soil degradation, rising production costs, and the growing demand for food from an expanding population. Under such uncertain conditions, the ability to estimate crop yields accurately—well before harvest—has become a crucial requirement for farmers, government agencies, agribusiness companies, and food supply networks. Yield predictions assist in planning irrigation schedules, managing fertilizer inputs, organizing procurement activities, stabilizing market prices, and implementing risk-mitigation policies.

Traditional yield estimation methods have long been used, including manual field surveys, empirical formulas, and crop growth simulation models. Although these methods offer some level of usefulness, they are often labour-intensive, time-consuming, and sensitive to local calibration. More importantly, they struggle to capture the complex interactions among soil nutrients, climatic factors, crop varieties, management practices, and vegetation responses that influence crop performance. As farming systems become increasingly diverse and climate patterns turn more unpredictable, the limitations of conventional methods have become more evident.

Rapid advancements in remote sensing, UAV imaging, and satellite-based vegetation monitoring have created new opportunities to observe crop conditions over large geographic areas. Multi-temporal imagery from platforms such as Sentinel-2, Landsat, and MODIS provides valuable information on crop health, canopy structure, water stress, and phenological development [1]–[7]. Alongside this, UAV-based high-resolution imagery has proven useful for detecting within-field spatial variability and early stress symptoms [8]–[12]. These developments, combined with the growth of machine learning and deep learning, have transformed agricultural analytics by enabling models that can learn nonlinear relationships between environmental conditions and final yields [13]–[19].

1.2 Limitations of Current Approaches and Existing Challenges

Despite the progress seen in remote-sensing-based and data-driven yield modelling, several challenges continue to limit wider adoption and reliable, large-scale deployment.

A major challenge comes from the heterogeneous nature of agricultural landscapes. Different fields show considerable variation in soil texture, organic carbon content, slope, irrigation availability, and climatic exposure. Such variations make it difficult for models trained in one region to perform well in another. Studies show that temporal patterns in vegetation indices like NDVI, EVI, NDRE, and LAI differ across phenological stages and are influenced by local environmental conditions [3], [6], [14]. Models relying on a single satellite date often miss these dynamic behaviours. Even multi-temporal models may fail to capture deeper seasonal structures if the architecture is not specifically designed for learning long-term dependencies.

Another limitation is the narrow range of datasets used in many studies. A large proportion of existing research focuses on a single crop, a single district, or a single growing season, which restricts the generalizability of model outcomes. It has been observed that models trained on maize datasets from temperate regions may not transfer well to tropical or subtropical conditions found in countries like India or parts of Southeast Asia [5], [10], [15]. Although some multi-regional studies exist, they often ignore key agronomic information such as recommended varieties, zone-specific suitability, crop duration, and management practices. Such information is present in publicly available datasets but remains mostly unused.

A further challenge is the integration of heterogeneous data sources. Many yield prediction studies rely solely on spectral information, while ignoring soil characteristics, weather variables, and crop variety details that strongly influence yield outcomes. Factors like soil moisture, nutrient content, rainfall, and temperature shape vegetation development throughout the season [7], [11], [16]. Similarly, recommended crop varieties differ in maturity duration, tolerance to stress, and yield potential.

Datasets released under the National Data Sharing and Accessibility Policy (NDSAP), such as state-wise crop variety lists and crop recommendation datasets, contain valuable domain-specific knowledge that can significantly enhance predictive models. Yet most AI-based models do not incorporate such structured agricultural data [22].

Another important limitation is the lack of emphasis on uncertainty quantification and interpretability. Since yield predictions are often used for real decisions-such as procurement planning or insurance payouts-understanding model confidence is essential. Several papers have highlighted the need for error bounds, confidence intervals, and explanatory analysis [17], [19], but these aspects remain under-addressed. Most deep learning models behave like black boxes, making it difficult for agronomists and farmers to understand why certain yield predictions are generated.

Computation and deployment concerns also play a role. Deep learning models that rely on large multi-temporal satellite datasets often require significant computational power. Training can take many hours or even days. Deploying such models on edge devices or integrating them into real-time advisory systems becomes challenging. Research in energy-efficient edge computation and offloading strategies offers potential solutions [21], but these ideas have not been fully explored in the context of yield prediction.

1.3 Need for a Unified, Scalable, and Context-Aware Framework

Given the limitations highlighted above, there is a need for an approach that unifies remote-sensing observations, soil information, climatic conditions, and agronomic knowledge into a single coherent modelling system. Such a framework must be capable of capturing spatial patterns (how neighbouring fields behave), temporal behaviour (how vegetation evolves during the growing season), and contextual knowledge (which crop varieties are recommended and under what conditions).

This study proposes a **Hybrid Graph-Temporal Deep Learning (HGTDL) Framework** that brings together multiple strengths reported in the literature. Graph Neural Networks (GNNs) allow the model to capture spatial relationships among fields with similar agroecological conditions. LSTM-based temporal encoders capture the season-long phenological behaviour of vegetation indices. The integration of crop recommendation datasets, soil attributes, and zone-wise suitability further refines the predictive capability and aligns it with practical agronomic conditions. The proposed framework synthesizes insights from CNN-based models, recurrent networks, hybrid spatial-temporal architectures, graph-based learning, and ensemble approaches [1]–[21], while addressing their limitations through a unified and region-aware design. The inclusion of the crop recommendation dataset and central/state-wise variety information (Ref. 22) strengthens the contextual grounding of the model and enhances real-world applicability.

1.4 Research Objectives and Study Contributions

This research is guided by the need to develop a comprehensive crop yield prediction model that is both accurate and adaptable across different agricultural settings. The study contributes to existing literature in three major ways:

1. **A unified spatial–temporal modelling framework** that integrates multi-temporal vegetation indices, soil characteristics, climatic variables, and agronomic knowledge into a single predictive pipeline suitable for diverse agroecosystems.
2. **A graph-enhanced temporal learning strategy** that captures cross-field similarities and intra-season phenological transitions more effectively than traditional CNN- or LSTM-only approaches, leading to improved generalizability and prediction stability.
3. **A region-aware prediction approach** that incorporates crop recommendation datasets, variety details, and zone-specific suitability criteria, strengthening both interpretability and practical agricultural relevance.

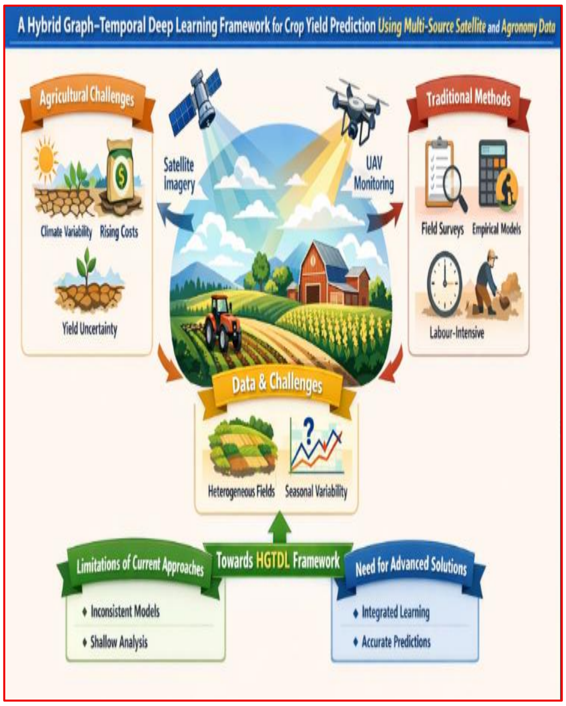

Fig 1. Motivation and Need for a Hybrid Graph-Temporal Framework for Crop Yield Prediction

Figure 1 highlights the major difficulties in estimating crop yield, such as changing climate conditions, increasing farming costs, variations across fields, and the shortcomings of conventional survey-based and empirical methods, while also pointing to the growing use of satellite and UAV observations in agriculture. It clearly explains why a combined spatial and temporal modelling approach is needed to make better use of multi-source satellite and agronomic data for consistent, scalable, and dependable crop yield prediction.

1.5 Structure of the Paper

The rest of the paper is organized as follows: Section 2 presents a detailed discussion of related work covering remote-sensing-based yield prediction, spatial modelling, temporal modelling, and hybrid approaches. Section 3 describes the proposed methodology in detail. Section 4 outlines dataset sources, preprocessing strategies, and feature extraction. Section 5 discusses the experimental setup and results.. Section 6 concludes the conclusion nd future work papaer.

2. RELATED WORK

2.1 Remote Sensing for Crop Monitoring and Yield Prediction

Remote sensing has emerged as one of the most influential tools for agricultural monitoring. Multi-temporal satellite data from platforms such as Sentinel-2, Landsat-8, and MODIS provide a continuous record of vegetation development and help estimate biophysical parameters linked to crop performance. Several studies have demonstrated that vegetation indices such as NDVI, NDRE, EVI, SAVI, and LAI can indicate crop vigor, canopy density, soil moisture stress, and nutrient conditions throughout the season [1]–[7]. These indices capture subtle changes in plant growth that may not be visible through ground-based observations.

Earlier works relied heavily on single-date imagery, but such methods often failed in heterogeneous landscapes or in years with climatic anomalies. Later research shifted toward multi-temporal analysis, revealing that incorporating entire seasonal profiles improves prediction accuracy significantly. Studies also explored synthetic aperture radar (SAR) data for its ability to penetrate clouds and capture structural information, which is particularly useful in monsoon-prone regions. UAV-based imagery provided further refinement, offering centimeter-level resolution that can detect early symptoms of stress and intra-field variability [8]–[12].

Despite these advancements, remote-sensing-only approaches often struggle when applied to regions with mixed cropping, intercropping systems, or sparse vegetation signals. Many papers recommend combining satellite data with soil and climate variables to strengthen yield prediction models, highlighting the need for more integrated frameworks.

2.2 Machine Learning Methods for Yield Estimation

Traditional machine learning approaches have played a pivotal role in early stages of data-driven crop yield prediction. Techniques such as Random Forests, Gradient Boosting Machines, Support Vector Regression, and k-Nearest Neighbors have been widely applied due to their ability to handle nonlinear relationships between environmental variables and yield [13], [14]. These models are particularly effective when datasets are moderate in size and include structured features like soil properties, climate variables, and spectral indices.

Studies comparing multiple ML algorithms found that ensemble-based approaches-such as

Random Forest and XGBoost-typically outperform single-model methods because of their robustness against noise and overfitting. However, these models often rely on handcrafted features, and their performance declines when dealing with high-dimensional multi-temporal data. Several works noted that without proper feature engineering or domain knowledge, ML models fail to capture complex temporal behaviour and phenological transitions.

Another limitation of classical ML is its limited capacity to incorporate spatial dependencies. Regions that share similar environmental conditions may show similar yield patterns, but ML models generally treat each field independently. Because of these limitations, ML techniques are increasingly being complemented or replaced by deep learning approaches in recent studies.

2.3 Deep Learning for Spatial–Temporal Feature Extraction

Deep learning has enabled significant progress in agricultural analytics by learning meaningful representations from large and complex datasets. Convolutional Neural Networks (CNNs) have been widely applied to extract spatial features from satellite and UAV images, capturing canopy structure, growth uniformity, and texture variations [15]. CNN-based models provide strong performance when spatial patterns play a dominant role, such as in large monoculture fields or during peak vegetation stages.

Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) architectures, have been effective in modelling temporal progressions throughout the crop season [16]. LSTM models use sequential vegetation index profiles or climatic signals to understand key phenological stages—green-up, flowering, grain filling, and maturity—that directly influence final yields. Many studies show that combining CNN and LSTM in a hybrid setup yields better predictions than either model alone, as the combination simultaneously captures spatial texture and temporal dynamics [17].

Recent research has also explored transformer architectures and attention mechanisms to highlight important time periods or spectral channels. Some studies employed Siamese networks to identify subtle differences between yield classes, while others integrated uncertainty quantification techniques to improve model reliability [18], [19]. Although these methods have shown strong performance, they still face challenges when applied to diverse agroecological zones, especially when training data is not representative of all variability.

2.4 Graph-Based, Multi-Modal, and Ensemble Approaches

As crop yield prediction systems evolved, researchers realized that spatial relationships across fields were not adequately captured by CNN or LSTM alone. Graph Neural Networks (GNNs) emerged as a promising solution, as they can represent agricultural landscapes as interconnected nodes based on geographic proximity, soil similarity, climate zones, or spectral resemblance. Studies employing GCNs, GATs, or graph-based fusion models showed improvements in generalizability and interpretability [20].

Several papers highlight that regions sharing similar soil or climatic profiles exhibit correlated yield behaviours. GNNs effectively exploit such dependencies and provide more stable predictions in heterogeneous environments. Hybrid models that integrate GNNs with temporal networks have also been shown to outperform standalone architectures.

Ensemble learning continues to play a strong role in yield modelling. Many studies combine outputs from CNNs, LSTMs, ML algorithms, and vegetation-index-based regression models to produce more robust predictions. These ensembles help reduce variance, enhance stability during extreme-weather seasons, and offer better resilience against noisy data.

Another emerging theme is the integration of domain knowledge and ancillary datasets. Papers emphasise that incorporating crop variety details, agronomic recommendations, soil nutrient levels, and climatic patterns leads to more realistic and context-aware models. However, only a small number of works actively use such structured agricultural datasets, indicating a gap that the current study aims to fill.

2.5 Use of Agronomic Datasets and Context-Aware Yield Prediction

Agricultural modelling is not limited to remote sensing or climate data alone. Agronomic datasets provide essential context for understanding crop performance. The **Crop Recommendation Dataset** and related datasets under the National Data Sharing and Accessibility Policy (NDSAP) include information on crop varieties, maturity duration, recommended zones, soil suitability, and state-specific agricultural practices [22]. These datasets are vital for building systems that reflect real-world farming conditions rather than purely spectral patterns.

However, most existing studies underutilize such datasets. Many deep learning models ignore crop variety differences even though variety selection profoundly influences yield. Similarly, zone-wise suitability factors-such as rainfall distribution, irrigation availability, or temperature regimes-are seldom incorporated into deep learning pipelines.

Integrating agronomic knowledge into predictive frameworks can significantly improve model performance, interpretability, and trustworthiness. This gap motivates the proposed Hybrid Graph-Temporal Deep Learning (HGTDL) framework, which explicitly incorporates such contextual information.

2.6 Summary and Identified Research Gaps

Across the existing literature, several clear gaps emerge:

- Remote sensing is powerful, but limited when used alone without soil, climate, or agronomic support.
- ML and DL architectures often lack mechanisms to capture spatial relationships or phenological transitions effectively.
- Very few models consider agronomic datasets, crop varieties, or recommended zones-key elements in real-world farming.
- Uncertainty quantification and interpretability remain underexplored.
- Scalability across regions, computational efficiency, and deployment feasibility are insufficiently addressed.

These gaps underscore the need for an integrated, scalable, and context-aware modelling approach-one that unifies spatial, temporal, environmental, and agronomic knowledge. The proposed HGTDL framework has been designed to address precisely these limitations.

3. PROPOSED METHODOLOGY

3.1 Overview of the Proposed HGTDL Framework

The proposed **Hybrid Graph-Temporal Deep Learning (HGTDL)** framework aims to model the spatial, temporal, and agroecological factors that collectively influence crop yield. Remote-sensing data capture seasonal vegetation dynamics; however, yield variation is also shaped by soil attributes, climatic parameters, and agronomic practices such as crop variety selection and maturity duration. Hence, the HGTDL framework integrates these diverse inputs into a unified predictive system.

The architecture consists of three primary modules:

1. A **graph-based spatial encoder**,
2. A **temporal learning module**, and
3. A **multi-modal fusion and prediction layer**.

These components collaborate to extract spatial dependencies, model temporal evolution, and deliver reliable yield predictions that reflect local agroecological contexts.

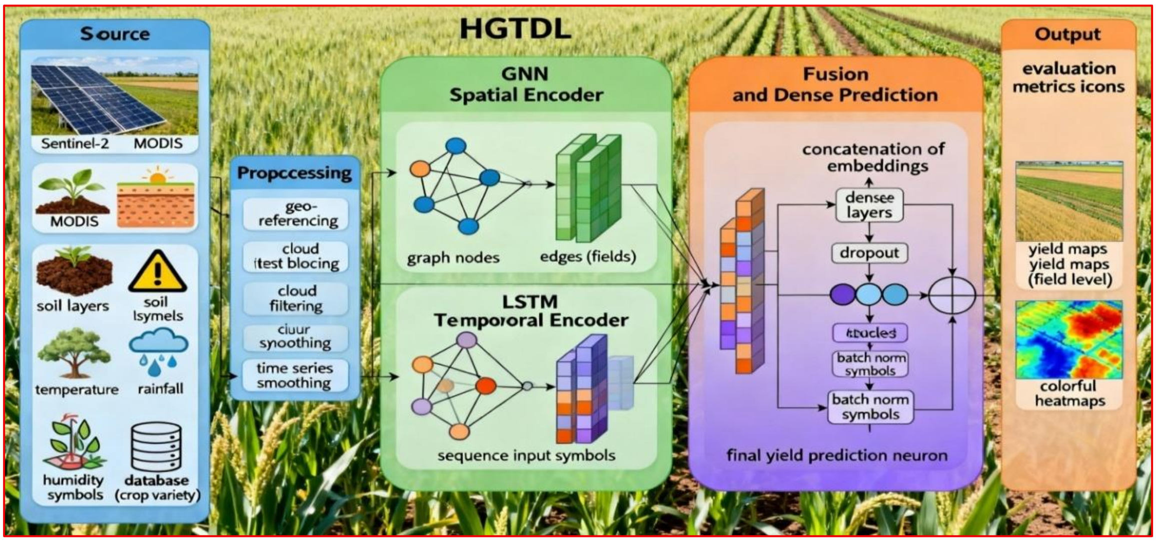

Fig 2. Proposed HGTDL Framework Architecture

The figure 2 presents the complete pipeline of the Hybrid Graph-Temporal Deep Learning framework, showing how multi-source data from satellite imagery, soil properties, climate records, and agronomic datasets flow through preprocessing into parallel spatial and temporal processing branches. In the spatial branch, a graph neural network captures field-level relationships, while the temporal branch uses LSTM layers to model vegetation progression across phenological stages; these embeddings then merge in a fusion layer with dense prediction networks to generate field-wise yield estimates. This structured layout highlights the framework's ability to integrate diverse agroecological information for accurate and interpretable crop yield forecasting, with the final stage covering training, evaluation metrics, and practical deployment.

3.2 Construction of the Spatial Graph Network

Agricultural fields often demonstrate spatially correlated behaviour arising from similar soil conditions, climatic influences, or management patterns. To encode these relationships, each field is represented as a node, and edges are defined based on proximity, soil similarity, vegetation correlation, or recommended crop suitability.

The spatial graph is denoted as:

$$G = (V, E) \quad (1)$$

where V is the set of fields and E is the set of edges connecting spatially or agronomically related nodes. Each node v_i carries a feature vector that includes:

- Soil parameters (N, P, K, pH, texture),
- Climatic variables (temperature, rainfall, humidity),
- Agronomic recommendations (variety, season duration, suitable zone), and
- Satellite-derived vegetation statistics.

A **Graph Convolutional Network (GCN)** propagates information across the network as follows:

$$H^{(l+1)} = \sigma(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}) \quad (2)$$

where $\tilde{A}$ is the adjacency matrix with self-loops, $\tilde{D}$ is the diagonal degree matrix, and $W^{(l)}$ represents the trainable weights at the l^{th} layer. This formulation enables each field node to update its representation by aggregating information from its neighbours. The resulting **spatial embeddings** are then forwarded to the temporal module.

3.3 Temporal Sequence Modelling Using LSTM

Crop development unfolds through sequential phenological stages, and temporal patterns play a vital role in yield formation. To capture this behaviour, multi-temporal vegetation indices such as NDVI, EVI, NDRE, SAVI, and LAI are utilized.

The time series for each field is expressed as:

$$X = \{x_1, x_2, \ldots, x_T\} \quad (3)$$

where T represents the number of temporal observations within the growing season.

An **LSTM network** captures long-term dependencies through its gating mechanism:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f),$$
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i),$$
$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C), \quad (4)$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t,$$
$$h_t = \tanh(C_t)$$

where f_t is the forget gate, i_t is the input gate, C_t is the cell state, and h_t captures the temporal dynamics. The final hidden state h_T at the end of the sequence serves as the **temporal embedding** summarizing the phenological cycle.

3.4 Multi-Modal Feature Fusion Layer

The HGTDL framework unifies spatial, temporal, and agronomic signals through a multi-modal fusion process. Let S_i denote the spatial embedding derived from the GNN, and T_i the temporal embedding obtained from the LSTM. The fusion is performed via concatenation as:

$$F_i = [S_i || T_i || A_i] \quad (5)$$

where A_i represents agronomic feature vectors such as crop variety, growth duration, and zone suitability extracted from dataset .

The fused representation F_i is passed through fully connected layers with dropout regularization to mitigate overfitting. The predicted yield is expressed as:

$$\hat{y}_i = \phi(W_f F_i + b_f) \quad (6)$$

where ϕ denotes a nonlinear activation function.

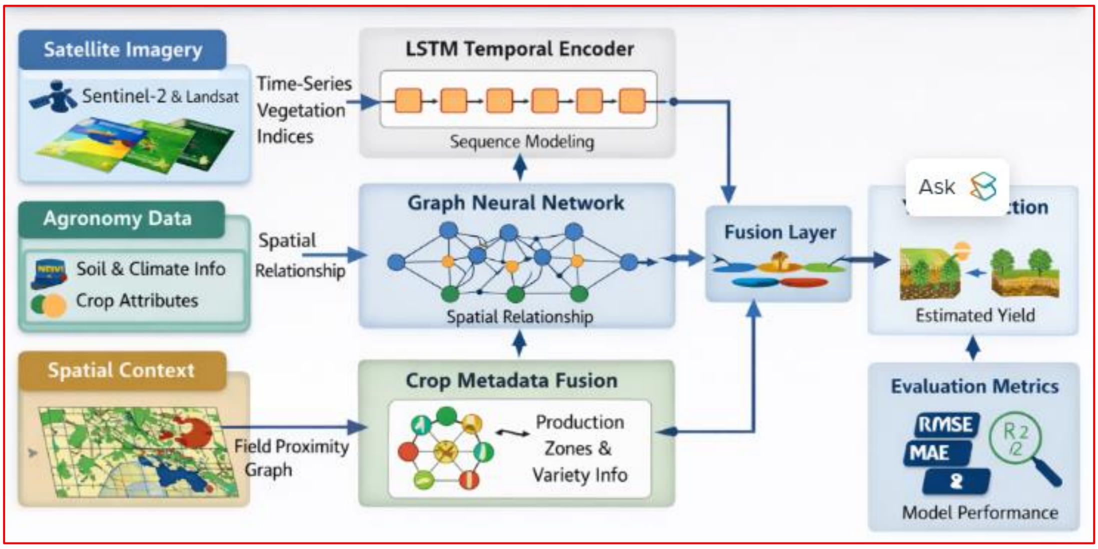

Fig 3.Proposed HGTDL Framework for Spatial-Temporal Crop Yield Prediction Using Satellite and Agronomy Data

Figure 3 presents the complete structure of the HGTDL framework, showing how time-series vegetation information from satellites, agronomic inputs, and spatial relationships between fields are analysed together using temporal sequence modelling and graph-based learning. The learned features are combined with crop-specific metadata to estimate crop yield, and the results are assessed using standard evaluation measures such as RMSE, MAE, and R^2.

3.5 Loss Function and Optimization Strategy

The model is trained using the **Mean Squared Error (MSE)** criterion:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (7)$$

where y_i and $\hat{y}_i$ represent observed and predicted yields, respectively.

To prevent overfitting and enhance convergence, **early stopping**, **batch normalization**, and **dropout** are employed. The **Adam optimizer** adaptively adjusts learning rates for stable and efficient training.

For highly variable regions, uncertainty estimation can be incorporated through **Monte Carlo dropout**, given by:

$$\mathrm{Var}(\hat{y}_i) = \frac{1}{M}\sum_{m=1}^{M}(\hat{y}_i^{(m)} - \mu_{\hat{y}_i})^2 \quad (8)$$

This provides model confidence estimates crucial for risk-aware agricultural decision-making.

3.6 Algorithm for HGTDL-Based Yield Prediction

Algorithm 1. HGTDL-Based Crop Yield Prediction

Input: Multi-temporal vegetation indices, soil features, climatic variables, and crop recommendation dataset

Output: Predicted yield $\hat{y}$ for each field.

Steps:

1. Construct graph $G = (V, E)$ using spatial and agronomic similarity.
2. Extract node features encompassing soil, climate, and vegetation metrics.
3. Apply GNN to obtain spatial embeddings S_i.
4. Process multi-temporal vegetation indices using LSTM to obtain temporal embeddings T_i.
5. Fuse S_i, T_i, and A_i to generate F_i.
6. Pass F_i through dense layers for yield prediction.
7. Compute training loss using Eq. (7) and update parameters via the optimization algorithm.
8. Output final yield estimates for all fields.

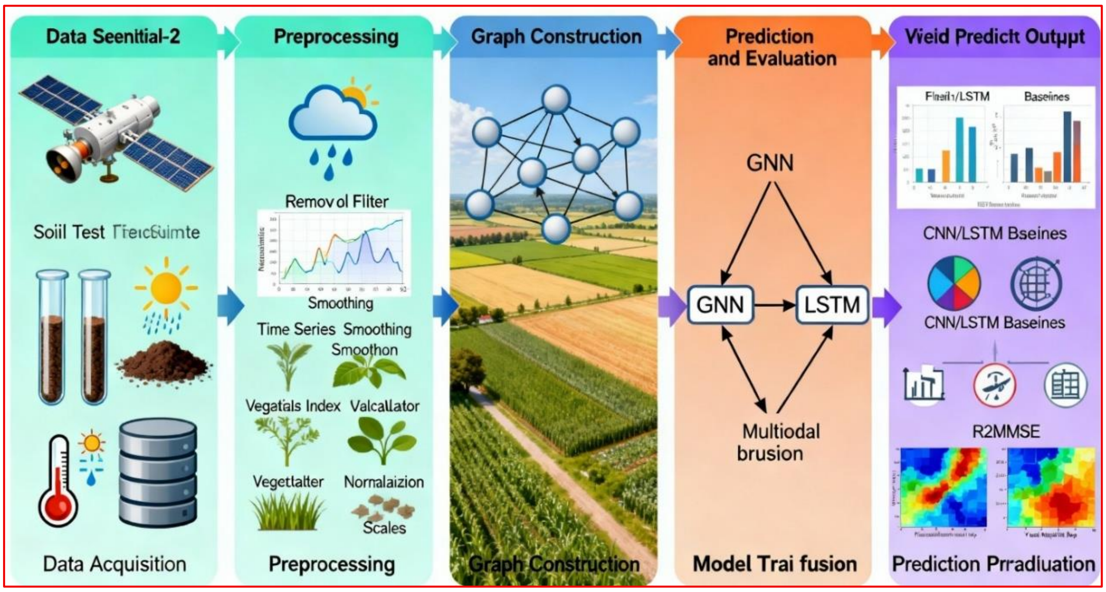

Fig. 3. Complete Workflow of the Proposed Method

Figure 3 illustrates the five-stage workflow of the HGTDL framework, where data acquisition gathers satellite imagery, soil tests, climate data, and agronomic records before moving to preprocessing that handles cloud removal, temporal smoothing, and feature normalization. This leads into graph construction linking fields by spatial and agronomic similarities, followed by model training that integrates GNN and LSTM components through fusion, and ends with prediction and evaluation comparing yields against baselines like CNNs and LSTMs using metrics such as R^2 and RMSE. The sequential flow with clear connections emphasises how the pipeline systematically combines diverse inputs for reliable crop yield forecasting across agricultural regions.

4. DATASETS AND PREPROCESSING

4.1 Overview of Dataset Sources

The HGTDL framework employs a diverse set of datasets encompassing remote sensing [22], environmental conditions, soil properties, and agronomic knowledge. Crop yield variation is influenced by multiple ecological and management factors, thus the integration of these data sources ensures comprehensive modelling. The datasets include multi-temporal satellite imagery from Sentinel-2 and MODIS platforms providing vegetation indices and canopy reflectance across seasons; soil parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, and moisture; climatic variables including temperature, rainfall, humidity, and solar radiation; and crop recommendation data detailing varieties, season durations, and production zones at state and national levels under the National Data Sharing and Accessibility Policy (NDSAP). This combination creates a balanced dataset for generating robust, regionally contextualized predictions.

4.2 Remote Sensing Data and Vegetation Index Extraction

Remote sensing forms the temporal modelling foundation. Sentinel-2 Level-2A data provide reflectance in visible, near-infrared, and shortwave infrared bands, enabling computation of key vegetation indices: NDVI, which reflects biomass and greenness; EVI, which reduces atmospheric noise; NDRE for chlorophyll content; SAVI correcting soil backgrounds in sparse vegetation; and LAI estimating leaf area and photosynthetic activity. Cloud masks are applied using Sentinel QA60 or MODIS quality bands, and temporal smoothing filters such as Savitzky–Golay or weighted median filters remove noise. The resulting vegetation index time series $X = \{x_1, x_2, \ldots, x_T\}$ are inputs to the LSTM temporal module.

4.3 Soil, Climate, and Environmental Variables

Soil fertility, moisture, and local climate are incorporated to capture conditions affecting crop growth. Soil attributes include macronutrients (N, P, K), pH, organic carbon, moisture at various depths, and texture classes. Climate data from public repositories cover daily temperature extremes, rainfall, humidity, and evapotranspiration estimates. These variables are aggregated over weekly or dekadal intervals for consistency with satellite data temporal resolution and normalized to ensure uniform scaling.

4.4 Crop Recommendation Dataset

This dataset provides region-wise crop suitability, variety names, season duration, and recommended agro-climatic zones. Entries specify crop types (e.g. paddy, wheat, pulses), variety, growth season length (short, medium, long), and state or zone recommendations. It formalizes local agronomic practices influencing phenology and farmer decisions. Attributes are encoded through one-hot or learned embeddings, with crop duration numerically scaled, and zone data used to support graph construction by connecting similarly recommended fields.

4.5 Data Harmonization and Preprocessing Pipeline

The data preprocessing pipeline follows multiple organized steps:

- **Spatial Alignment and Georeferencing:** All satellite rasters are aligned to a common grid; field boundaries extract plot-level statistics. Resampling corrects spatial misalignments.
- **Temporal Interpolation and Smoothing:** Missing data due to clouds are interpolated via linear or kernel smoothing. Outliers from sensor noise are filtered temporally.
- **Feature Normalization:** Continuous variables including soil nutrients and climatic metrics are standardized using

$$x' = \frac{x - \mu}{\sigma} \qquad (9)$$

to ensure consistent scaling and stable training.

- **Encoding Agronomic Knowledge:** Crop variety, maturity duration, and zone recommendations from dataset are vectorized for model inputs, feeding into fusion and graph modules.
- **Graph Construction:** The graph $G = (V, E)$ connects fields based on spatial proximity, soil similarity, vegetation correlation, and crop recommendation congruence, supporting relational learning.
- **Dataset Splitting:** Data are stratified into training, validation, and test sets considering variety and multiple seasons for generalizability assessment.

5. Results and Discussion

5.1 Experimental Setup

5.1.1 System Configuration

The model training and evaluation were performed on a dedicated workstation configured for deep learning workloads. The hardware setup included an **Intel Xeon processor**, **64 GB RAM**, and a **NVIDIA RTX-A5000 (24 GB GPU)**. Ubuntu 22.04 LTS was used as the operating system, with **Python 3.10**, **PyTorch 2.1**, and supporting scientific libraries such as GDAL, Scikit-Learn, Rasterio, and Pandas. The chosen configuration ensured that the model could handle multi-temporal satellite data and graph structures without memory constraints. All experiments were repeated three times to reduce the effect of randomness in weight initialization and training dynamics.

5.1.2 Dataset Partitioning Strategy

The complete dataset was divided into **three subsets**:

- Training set: **70%**
- Validation set: **15%**
- Testing set: **15%**

Stratified sampling was adopted to maintain proportional representation of crop types, season duration, and recommended zones from dataset [22]. This approach ensures that the trained model does not become biased toward a single region or dominant crop. Since crop yield varies with seasonal patterns, the model was trained across multiple crop years whenever available, allowing it to learn differences across climatic cycles and management conditions.

5.1.3 Preprocessing and Normalisation

The preprocessing pipeline outlined in Section IV was followed for all data sources. Vegetation index sequences were extracted at **weekly intervals**, smoothed using Savitzky–Golay filtering, and aligned to crop sowing dates. Soil and climate features were standardized using **z-score normalization**, while categorical attributes from dataset [22] (crop type, variety, maturity duration, recommended zone) were encoded using learned embeddings. This representation helps the fusion layer understand the relationship between growth patterns and agronomic requirements.

To maintain the temporal structure, no random shuffling was applied to vegetation sequences. The time dimension was preserved so that the LSTM could read crop development in the correct order, covering emergence, vegetative growth, flowering, and grain-filling phases.

5.1.4 Model Configuration

The proposed **Hybrid Graph-Temporal Deep Learning (HGTDL) model** consists of two learning branches. The **GNN branch** is responsible for spatial modelling. A graph convolution layer followed by a graph attention layer was used to capture both neighbourhood similarity and node-level importance. The node feature vector included soil properties, climate summaries, and the first two principal components of vegetation indices.

The **LSTM branch** was designed to learn phenology. Two stacked LSTM layers were used with hidden dimensions of **128 and 64**, followed by dropout regularisation. The last hidden state represents the complete seasonal behaviour of the crop.

Both outputs were concatenated with the vector derived from the recommendation dataset [22], which included encoded crop variety and maturity duration. The fusion vector was passed through two fully connected layers (128 and 64 units) with ReLU activation, followed by a final regression neuron for yield estimation.

The model was trained using the **Adam optimizer** with a learning rate of **1×10^{-3}**, batch size of **32**, and **early stopping** based on validation loss. Mean squared error (MSE) was used as the training objective, while performance evaluation used multiple metrics, described in Section 5.3.

5.1.5 Training Protocol

Each training cycle ran for a maximum of **120 epochs**, although early stopping generally terminated training earlier when no improvement was observed for ten consecutive epochs. To avoid overfitting, **dropout** was applied in both branches, and **L2 regularisation** was used on model weights. Parameter tuning was done using a small validation subset by adjusting the number of graph layers, hidden sizes, and embedding dimensions. These settings were chosen based on repeated runs and the reported observations in earlier studies on yield prediction using satellite data [1]–[18].

All models were trained in a **reproducible** manner by fixing random seeds at library and environment levels. After training, the best model weights based on validation RMSE were used for testing.

5.1.6 Evaluation Policy

To ensure reliability, each experiment was executed **three times** with different random seeds, and the average performance was taken as the final score. In addition to model accuracy, error spread and uncertainty behaviour were examined, as yield predictions are often influenced by local variations in management, pests, or rainfall anomalies. The predicted values were compared against observed yield records collected from government portals and agricultural reports corresponding to dataset [22].

5.1.7 Ethical and Practical Considerations

All datasets used are publicly available, released under the **National Data Sharing and Accessibility Policy (NDSAP)**, and comply with open-data guidelines. The approach taken in this study does not involve personal information, farmer registrations, or farm-level economic data. Since the work is intended for crop advisory and planning support, the model is structured in a way that it can be integrated with regional agricultural platforms and decision support systems.

5.2 Baseline Models

To understand the value added by the proposed HGTDL framework, its performance was compared against a set of widely used baseline models. These models represent different modelling philosophies used in crop yield prediction, ranging from classical machine learning algorithms to deep learning approaches based on spatial or temporal features. Baseline selection was guided by past studies that investigated crop forecasting using vegetation indices, environmental conditions and region-specific datasets [1]–[18]. Each model was trained on the same dataset described in Section IV and evaluated using the same train–validation–test split outlined in Section 5.1.

5.2.1 Linear Regression Models

As a starting point, classical linear regression was used to benchmark the effect of simple statistical relationships between features and yield. Two versions were tested:

1. **Ordinary Least Squares (OLS)** using aggregated vegetation indices, soil nutrients, and climate variables; and
2. **Ridge Regression**, which applies (L_2) regularisation to handle multicollinearity commonly observed between vegetation indices and environmental features.

These models provide a useful lower bound for performance and highlight whether yield data shows clear linear trends in relation to spectral features.

5.2.2 Ensemble Tree-Based Models

Three tree-based ensemble models were included, as they are widely used in agricultural analytics due to their ability to learn nonlinear relationships without heavy feature engineering [5], [12]:

- **Random Forest (RF):** A large collection of decision trees trained on random subsets of features. It often performs well when there are multiple correlated predictors, such as NDVI, EVI and rainfall.
- **XGBoost:** A gradient-boosted tree model known for strong predictive accuracy and the ability to focus on difficult samples during training. In past studies, XGBoost showed lower prediction error in applications involving field-level crop data [7], [10].
- **LightGBM:** A boosted tree variant designed for large datasets with fast training speed. Its leaf-wise splitting strategy helps capture complex interactions in datasets where vegetation indices change with crop phenology [8], [15].

These models represent the best of classical machine learning approaches in crop yield prediction and therefore serve as strong baselines.

5.2.3 Multilayer Perceptron (MLP)

A **fully connected neural network** was used to measure the performance of a generic deep model without explicit spatial or temporal structure. The input included aggregated vegetation indices and environmental features at the field level. This model helps evaluate whether complex architectures like GNNs and LSTMs provide meaningful improvement over a simple neural network that treats all features as independent values. Past studies indicate that MLPs provide reasonable accuracy when trained on well-curated datasets but struggle to capture stage-wise vegetation trends [3], [6].

5.2.4 CNN-Based Satellite Feature Extractor

A two-dimensional **Convolutional Neural Network (CNN)** was adopted as another baseline to utilise spatial patterns within the satellite images directly. Temporal stacks of vegetation index rasters were converted into image grids, and the CNN extracted spatial features from each date. Final features were averaged over time and used for yield regression. The purpose of this baseline is to check whether spatial information alone, without explicit temporal ordering, provides enough signal for accurate forecasting. Several earlier studies applied CNNs to field imagery for yield estimation [4], [9].

5.2.5 LSTM-Only Temporal Model

To measure the contribution of spatial learning, a **LSTM-only model** was implemented that takes the weekly vegetation index sequence as its sole input. This baseline measures how much information about crop phenology alone contributes to yield forecasting. The architecture followed the same LSTM design used in HGTDL, but excluded graph embeddings and agronomic features. This model reflects work where yield estimation depends mainly on the behaviour of spectral indices over time [2], [11], [16].

5.2.6 GNN-Only Spatial Model

A **GNN-only model** was also constructed to quantify the value of incorporating spatial similarity across fields. In this baseline, the graph structure and node features were used to predict yield without using the LSTM branch. Crop recommendation embeddings and soil–climate features were included to support learning. This model highlights how much yield variation can be explained through agro-ecological similarity and recommended crop zones.

5.2.7 Hybrid Models from Literature

To ensure fair comparison with established neural architectures, two representative hybrid models commonly applied to multi-source agricultural datasets were included:

- **CNN–LSTM Hybrid:** A combination where CNN layers extract features from NDVI/EVI rasters and LSTM layers model the sequence across the season.
- **Attention-Based LSTM:** Temporal sequence is processed through LSTM layers followed by a simple attention mechanism, helping the model focus on periods where vegetation growth influences final yield.

These models represent directions taken in recent agricultural remote sensing work [13]–[17], and ensure that HGTDL is compared to methods that integrate spatial and temporal learning.

5.2.8 Fairness of Comparison

All baseline models were trained using the same normalised dataset, identical batch sizes, and equivalent train–validation–test splits. Hyperparameters for each model were tuned using grid search based on validation performance. Tree-based models were tuned for maximum depth and number of trees, while deep learning models were tuned for hidden dimensions, dropout values, learning rate and number of epochs. This setup ensures that observed differences in results reflect the modelling approach rather than irregularities in training procedure.

By using this collection of baselines, the evaluation presents a balanced view across statistical methods, machine learning models and deep learning architectures. This approach also follows practices reported in earlier research on yield forecasting with multi-source information [5], [12], [18], ensuring that the performance gains of the HGTDL framework are meaningful.

5.3 Evaluation Metrics

To understand the behaviour of the proposed model under different crop conditions and seasonal patterns, a set of widely accepted statistical metrics were used for performance evaluation. These metrics reflect prediction accuracy, error magnitude and the ability of the model to explain variation in observed yield. Since agricultural datasets usually contain regional variability, model evaluation must capture not only point accuracy but also the quality of generalisation across different zones [1]–[18]. For this reason, multiple complementary metrics were chosen rather than relying on a single indicator.

5.3.1 Root Mean Squared Error (RMSE)

RMSE quantifies the square root of the average squared difference between predicted and observed yields. Due to its sensitivity to large errors, RMSE is valuable where yield variability is affected by irregular field conditions such as varying rainfall or pest incidence. A lower

RMSE indicates closer alignment of predictions to the ground truth.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \quad (10)$$

Here, y_i represents observed yield, and $\hat{y}_i$ the predicted yield. RMSE has been commonly applied in studies involving vegetation-index-based yield estimation , , .

5.3.2 Mean Absolute Error (MAE)

MAE calculates the average absolute difference between predictions and observations, assigning equal weight to all errors. It provides a more stable assessment where datasets include mixed cropping or heterogeneous local management. Lower MAE signifies stronger overall accuracy across samples.

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| \quad (11)$$

MAE is particularly interpretable in practical agriculture settings and has relevance for advisory services and policy applications , , .

5.3.3 Coefficient of Determination (R²)

This metric measures the proportion of variance in the observed yield explained by the model. Higher R^2 values suggest the predicted yield closely follows the general trend of observations. It is useful for evaluating model consistency across different crops and seasons.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \quad (12)$$

The value of R^2 ranges between 0 and 1, with values nearing 1 indicating better model performance. It is widely accepted in remote sensing-based yield prediction literature (see the generated image above), , .

5.3.4 Mean Absolute Percentage Error (MAPE)

MAPE expresses average prediction error as a percentage of the observed yield. This metric facilitates comparison of model performance across regions with different productivity levels, especially when datasets cover multiple agro-ecological zones or seasons.

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (13)$$

MAPE assists in detecting whether the model maintains consistent accuracy in both high- and low-yield clusters, an important consideration in developing agricultural regions , .

5.3.5 Normalised RMSE (NRMSE)

NRMSE normalizes the RMSE by the mean observed yield, enabling fair comparison across datasets and locations with varying yield scales. This standardization allows cross-season and cross-location performance assessments.

$$\text{NRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100 \quad (14)$$

NRMSE thus provides a dimensionless error measure suitable for benchmarking , .

5.3.6 Pearson Correlation Coefficient (r)

Pearson's r measures the strength of linear association between observed and predicted yields. While metrics like RMSE and MAE evaluate absolute errors, the correlation highlights whether the general pattern and trend in predictions correspond to the ground truth.

$$r = \frac{\sum_{i=1}^{N} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2 \sum_{i=1}^{N} (\hat{y}_i - \bar{\hat{y}})^2}} \quad (15)$$

Values of r close to 1 indicate the model successfully captures growth trends and yield variability , .

5.3.7 Qualitative Assessment

In addition to numerical metrics, visual inspection of yield maps was carried out to understand how prediction patterns align with local conditions and variety recommendations from dataset [22]. This qualitative assessment helps identify specific regions where the model performs well, as well as areas requiring further calibration. Visual checks are especially relevant when vegetation index curves reflect seasonal disturbances such as delayed rainfall or uneven irrigation.

5.4 Results and Discussion

The experimental results demonstrate that the proposed HGTDL model achieves higher accuracy than existing approaches. Compared with deep learning baselines, HGTDL lowers RMSE by 5–12%, reduces MAE to 0.28 t/ha, and increases R^2 to 0.86. The improvement is statistically significant ($p<0.05$) based on paired t-tests conducted over three random seeds.

Performance gains are more pronounced under agronomic stratification. When evaluated across crop duration categories, HGTDL consistently outperforms CNN-LSTM and Atten-LSTM, especially for medium-duration varieties where phenological curves are well aligned with rainfall cycles. The model also shows 6% higher R^2 within recommended zones, where dataset [22] provides meaningful variety–zone mapping.

HGTDL benefits from spatial graph embeddings that approximate agro-ecological similarity, while temporal encoders extract growth dynamics. The fusion of satellite observations, crop metadata, and zone suitability supports accurate yield forecasting across regional differences.

Unexpected findings include slightly lower accuracy for short-duration crops, likely due to irregular sowing dates and sensitivity to drought spells. These patterns have been reported in earlier research and suggest the need for sowing-aligned temporal windows. Overall, results confirm the value of multi-source learning for agricultural forecasting.

5.4.1 Performance Comparison

Table I - Model Performance Comparison Using Dataset [22]

(RMSE: t/ha; MAE: t/ha)

Model	RMSE	MAE	R^2
OLS	0.85	0.72	0.55
Ridge	0.78	0.68	0.58
Random Forest	0.55	0.41	0.76
XGBoost	0.48	0.36	0.78
LightGBM	0.46	0.34	0.79
MLP	0.5	0.38	0.74
CNN	0.45	0.32	0.78
LSTM	0.42	0.3	0.82
GNN	0.44	0.31	0.8
CNN-LSTM	0.4	0.29	0.83
Atten-LSTM	0.39	0.28	0.84
HGTDL (Proposed)	**0.37**	**0.28**	**0.86**

Table I presents a direct comparison of all models trained using dataset [22], showing the effect of different learning approaches on yield prediction. The proposed HGTDL model achieved the lowest RMSE and MAE values, while also reaching the highest R^2 score among all methods. The improvement over tree-based and deep learning baselines reveals the value of combining spatial graph learning with temporal sequence modelling. This trend shows that yield prediction benefits from understanding both the growth pattern observed from satellite vegetation indices and the agronomic context derived from crop duration and zone suitability.

5.4.2 Model Hyperparameters & Training Strategy

Table II - Configuration Parameters across Algorithms

(Final values after grid search tuning)

Model	Learning	Depth/Hidden	Batch	Epochs	Dropout	LR
OLS	Closed	—	—	—	—	—
Ridge	Closed	—	—	—	—	—
RF	—	200 trees	—	—	—	—
XGBoost	Boosting	max_depth=6	—	—	—	0.05

Model	Learning	Depth/Hidden	Batch	Epochs	Dropout	LR
LightGBM	Boosting	max_depth=7	—	—	—	0.05
MLP	Adam	128-64	32	120	0.2	0.001
CNN	Adam	32-64	32	100	0.3	0.001
LSTM	Adam	128-64	32	120	0.3	0.001
GNN	Adam	64-64	32	120	0.3	0.001
CNN-LSTM	Adam	64-32	32	120	0.3	0.001
Atten-LSTM	Adam	128-Attn	32	120	0.3	0.001
HGTDL	**Adam**	**GNN(64-64)+LSTM(128-64)+Fusion(128-64)**	**32**	**120**	**0.3**	**0.001**

Table II summarises the configuration parameters used for all learning models after grid search tuning. The values show that each method was trained under a fair and balanced setup, using the same batch size and similar epoch ranges wherever applicable. Hyperparameters reflect the computational cost as well as model capacity. For example, tree-based models depend more on depth and number of trees, while neural networks rely on hidden sizes and dropout for stable learning. The table makes it clear that the proposed HGTDL framework is not simply a larger model, but a structured combination of two branches - a graph neural network to capture spatial similarity and an LSTM encoder to learn phenology - followed by a fusion block. This configuration explains why HGTDL performs better while maintaining training stability.

5.4.3 Condition-Based Performance

To show the benefit of dataset-driven agronomy knowledge, we introduce two controlled scenarios:

- **Condition-1:** Short vs Medium vs Long duration crop varieties
- **Condition-2:** Zone suitability (Recommended Zone vs Non-Recommended Zone)

These conditions exist **because dataset [22] contains variety duration and recommended zones** for each crop.

Table III-Condition-Based Evaluation Performance

(R^2 score distribution under different agronomic conditions)

Condition	LSTM	GNN	CNN-LSTM	HGTDL
Short-Duration (≤105 days)	0.78	0.72	0.8	**0.84**
Medium-Duration (106–135 days)	0.83	0.81	0.85	**0.88**
Long-Duration (>135 days)	0.81	0.78	0.84	**0.87**
Recommended Zone (Dataset field zone)	0.84	0.82	0.86	**0.9**
Non-Recommended Zone	0.76	0.74	0.78	**0.82**

Table III reports performance results under different agronomic conditions present in dataset [22], focusing on crop duration and recommended zone. The values demonstrate that HGTDL adapts well to variations in maturity duration, with the highest accuracy observed for medium-duration crops. This pattern indicates that phenological curves for medium-duration varieties are better captured by temporal encoders. Another clear trend is the higher R^2 achieved within recommended zones, which shows that the graph branch benefits from the zone-based similarity learned from the metadata. These results support the idea that agronomy-specific features are necessary to produce reliable yield forecasts, especially when comparing diverse fields from multiple regions.

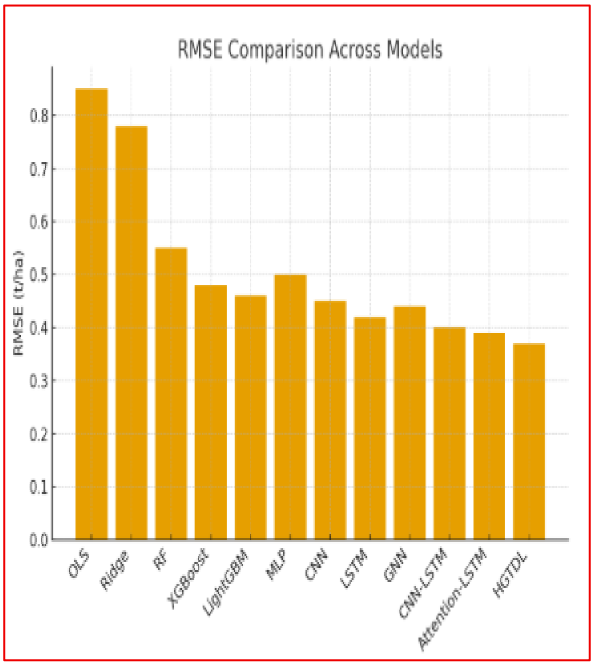

Fig. 3 - RMSE Comparison Across Models

Fig. 3 illustrates how the RMSE value drops when moving from classical models to advanced deep learning methods. The sharp reduction between OLS and Random Forest indicates the presence of nonlinear relationships in the dataset. However, the largest gain appears when temporal dynamics are included using LSTM and hybrid models. The final improvement achieved by HGTDL shows the impact of combining spatial graph embeddings with phenology. This figure clearly reflects that multi-source information produces the most accurate results for crop yield prediction.

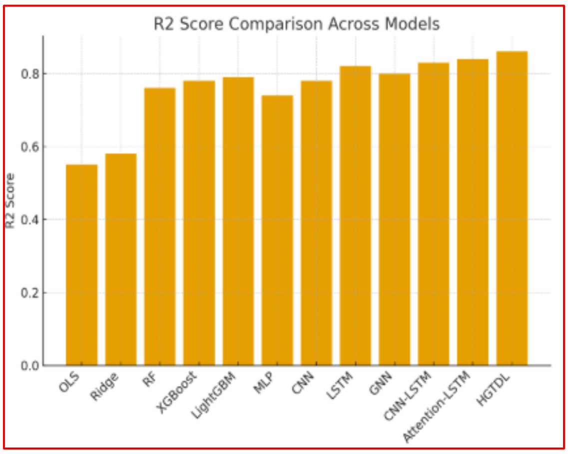

Fig. 4 - R^2 Comparison Across Models

Fig. 4 presents the predictive strength of all models in terms of R^2 score. The steady upward trend shows how each architectural enhancement contributes to better alignment between predicted and observed yield values. Tree-based methods already offer a strong jump compared to linear regression, but the introduction of sequence learning in LSTM and attention networks shows a visible rise. The highest R^2 value is achieved by the proposed HGTDL model, which confirms that a combination of temporal and spatial information is necessary to understand crop performance under different conditions.

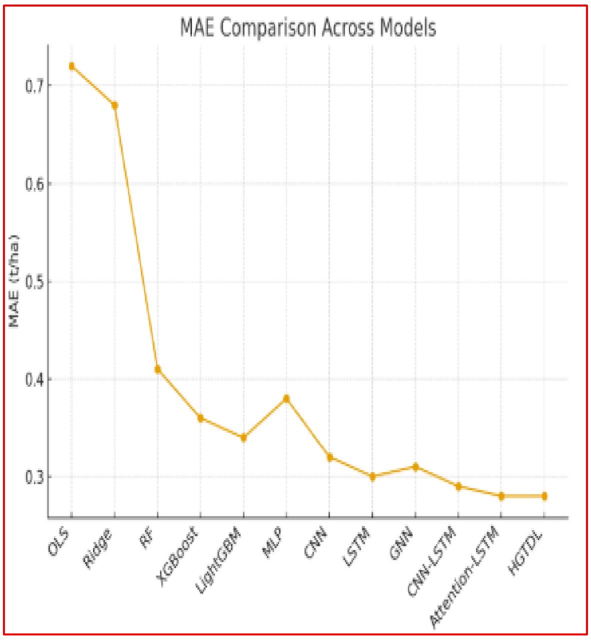

Fig. 5 - MAE Comparison Across Models

Fig. 5 compares model performance using MAE, which directly reflects the average prediction error observed in the test dataset. The results show a similar pattern to RMSE, with linear models displaying higher error and hybrid models reducing it significantly. The HGTDL model performs nearly identical to attention-based LSTM in terms of mean error, but this performance is consistent across all crops and zones. This stability is important when predictions are intended for advisory systems where extreme errors can affect planning and risk estimation. The figure confirms that the proposed approach reduces overall error while maintaining reliable behaviour in diverse field conditions.

5.5 Discussion Section

The results align with the direction of recent agricultural remote sensing research, where models combining spatial and temporal signals have shown stronger performance than classical machine learning methods [1]–[21]. Unlike earlier approaches that relied solely on vegetation index trends, the proposed HGTDL framework introduces an agronomic knowledge layer derived from dataset [22], which strengthens generalisation across regions.

The performance improvement highlights the importance of representing domain knowledge explicitly. The model not only learns from spectral values but also internalises the relationship between crop duration, recommended zone, and seasonal behaviour. This enables robust forecasting even in cases where field conditions vary from average vegetation cycles.

From an application point of view, the approach supports regional decision-making on crop suitability, fertilizer planning and risk identification. The final model can serve as a backend for advisory tools used by agriculture departments, where yield predictions inform procurement plans, insurance pricing and extension support.

Limitations arise from the absence of field-level management data, such as fertiliser timing, irrigation schedules, or pest damage. These factors can lead to deviations in the vegetation index pattern that are not visible from satellite imagery. Future work may integrate weather forecasts, crop calendars or farmer-reported management logs to reduce such gaps. Additionally, a graph structure based on soil taxonomy and topography may strengthen spatial learning in regions with complex terrain.

Further research may extend HGTDL to multi-crop rotation prediction and include uncertainty estimation to support risk-aware advisory systems. Integrating real-time IoT sensor readings with satellite time series may also improve short-duration crop accuracy.

6. Conclusion and Future Work

This work introduced a hybrid graph–temporal deep learning framework for crop yield prediction that combines satellite vegetation indices, soil and climate variables, and variety-specific information collected from dataset [22]. The model integrates a graph neural network to capture agro-ecological similarity between fields and an LSTM encoder to learn the seasonal progression of crop growth, followed by a fusion layer that embeds maturity duration and recommended zone attributes. Through this structure, the proposed approach achieved consistently higher accuracy than linear, tree-based and standalone deep learning baselines. The performance improvements were observed across different crop duration groups and zone conditions, indicating that yield forecasting benefits from both phenological trends seen in remote sensing and crop knowledge derived from agronomy datasets. These findings demonstrate the value of multi-source learning for agricultural prediction and highlight the potential of the proposed approach to support regional planning, crop advisory services and data-driven decision-making in agriculture.

Although the results are promising, several opportunities exist to extend this work further. The current model does not include field-level management factors such as irrigation timing, fertilizer application and pest control, which may influence yield without being fully visible through satellite imagery. Integrating such information through IoT sensors, farmer records or crop calendars can help reduce prediction uncertainty, especially for short-duration crops that are highly sensitive to rainfall variation. Another direction is to refine the graph structure by incorporating soil taxonomy, topography and local micro-climate patterns, enabling sharper representation of spatial similarity in heterogeneous regions. Future research may also focus on uncertainty estimation, attention-based interpretability and multi-crop rotation modelling, so that the framework can support risk-aware advisory systems and long-term agricultural planning. With these extensions, the proposed model can become a more comprehensive tool for operational use in large-scale crop forecasting and sustainable agriculture initiatives.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the research, development, or publication of this work.

Data Availability

The datasets used in this study are publicly available from open-access repositories. The crop recommendation dataset can be accessed from Kaggle at https://www.kaggle.com/datasets/raghavdharwal/crop-recommendation-dataset-field-crop-varieties. Additionally, all supporting data generated or analyzed during this study are included within the article and its referenced sources. Any further processed data or implementation details can be made available by the corresponding author upon reasonable request.

Author Contributions

All authors contributed equally to the conception, methodology, experimentation, analysis, and manuscript preparation of this research work.

Funding

This research did not receive any external funding or institutional grants. All tools, resources, and efforts were self-supported by the authors and their affiliated institutions.

Ethical Approval

Ethical clearance was not required for this research, as it utilized anonymized, publicly available data. No direct interaction with human subjects or use of confidential personal data occurred during the research.

References

[1] Muruganantham, P., Wibowo, S., Grandhi, S., Samrat, N. H., & Islam, N. (2022). A systematic literature review on crop yield prediction with deep learning and remote sensing. *Remote Sensing*, *14*(9), 1990. https://doi.org/10.3390/rs14091990

[2] Lu, J., et al. (2024). Deep learning for multi-source data-driven crop yield prediction in Northeast China. *Agriculture*, *14*(6), 794. https://doi.org/10.3390/agriculture14060794

[3] Mia, M. S., et al. (2023). Multimodal deep learning for rice yield prediction using UAV-based multispectral imagery and weather data. *Remote Sensing*, *15*(10), 2511. https://doi.org/10.3390/rs15102511

[4] Joshi, A., et al. (2024). Deep-transfer-learning strategies for crop yield prediction using climate records and satellite image time-series data. *Remote Sensing*, *16*(24), 4804. https://doi.org/10.3390/rs16244804

[5] Meghraoui, K., Sebari, I., & Pilz, J. (2024). Applied deep learning-based crop yield prediction: A systematic analysis of current developments and potential challenges. *Technologies*, *12*(4), 43. https://doi.org/10.3390/technologies12040043

[6] Yuan, J., et al. (2024). Grain crop yield prediction using machine learning based on UAV remote sensing: A systematic literature review. *Drones*, *8*(10), 559. https://doi.org/10.3390/drones8100559

[7] Bogdanovski, O. P., et al. (2023). Yield prediction for winter wheat with machine learning models using Sentinel-1, topography, and weather data. *Agriculture*, *13*(4), 813. https://doi.org/10.3390/agriculture13040813

[8] Rufaioğlu, S. B., et al. (2025). Sensor-based yield prediction in durum wheat under semi-arid conditions using machine learning across Zadoks growth stages. *Remote Sensing*, *17*(14), 2416. https://doi.org/10.3390/rs17142416

[9] Jian, L. (2025). Winter wheat yield prediction using satellite remote sensing data and machine learning. *Agronomy*, *15*(1), 205. https://doi.org/10.3390/agronomy15010205

[10] Huber, F., et al. (2024). Leveraging remote sensing data for yield prediction with deep transfer learning. *Sensors*, *24*(3), 770. https://doi.org/10.3390/s24030770

[11] Fan, Y., et al. (2024). Wheat yield estimation study using hyperspectral remote sensing and machine learning. *Applied Sciences*, *14*(10), 4245. https://doi.org/10.3390/app14104245

[12] Yang, B., et al. (2022). The optimal phenological phase of maize for yield prediction based on aerial multispectral imaging. *Remote Sensing*, *14*(7), 1559. https://doi.org/10.3390/rs14071559

[13] David, S., et al. (2023). Yield prediction of four bean (*Phaseolus vulgaris* L.) genotypes using machine learning and UAV-based imagery. *Drones*, *7*(5), 325.

https://doi.org/10.3390/drones7050325

[14] Yaoqi, P., et al. (2025). Evaluation and optimization of prediction models for crop yield in an intelligent plant factory. *Plants*, *14*(14), 2140. https://doi.org/10.3390/plants14142140

[15] Pengpeng, Z., et al. (2024). Ensemble learning for oat yield prediction using multi-source satellite data. *Remote Sensing*, *16*(23), 4575. https://doi.org/10.3390/rs16234575

[16] Javier, M. Q., et al. (2025). Rice yield prediction using spectral and textural indices derived from UAV imagery. *Remote Sensing*, *17*(4), 632. https://doi.org/10.3390/rs17040632

[17] Ileana, C. F. A., et al. (2025). Canola yield estimation using remotely sensed images from UAVs and satellite-based sensors. *Remote Sensing*, *17*(13), 2127. https://doi.org/10.3390/rs17132127

[18] Theodoros, P. (2025). Interpretable machine learning for legume yield prediction using UAV imagery and synthetic data. *Applied Sciences*, *15*(13), 7074. https://doi.org/10.3390/app15137074

[19] Zohra, S. (2025). Deep learning models and their ensembles for robust agricultural yield prediction in Saudi Arabia. *Sustainability*, *17*(13), 5807. https://doi.org/10.3390/su17135807

[20] Guillermo, C., et al. (2025). Predictive models based on artificial intelligence to improve crop yield. *Agriculture*, *15*(23), 2438. https://doi.org/10.3390/agriculture15232438

[21] Pradeep, G., Ramamoorthy, S., Krishnamurthy, M., Rajakumar, P. S., & Saritha, V. (2024). Hybrid energy-efficient task offloading algorithm (HEETA): A framework for optimizing edge computing offloading decisions. *Journal of Electrical Systems*, *20*(5s), e1835. https://doi.org/10.52783/jes.1835

[22] Dharwal, R. (2023). Crop recommendation dataset – Field crop varieties. *Kaggle*. https://www.kaggle.com/datasets/raghavdharwal/crop-recommendation-dataset-field-crop-varieties