AND DATA SCIENCES
BULLETIN OF COMPUTER

Deep Spatio-Temporal Learning for Parcel-Level Tea Quality and Yield Prediction Using IoT and Remote Sensing Data

Lena Paul and Mia Katharina

[1]Ludwig Maximilian University of Munich (LMU Munich)

[2]Technical University of Munich (TUM)

Abstract

Accurate and timely prediction of tea quality and yield at parcel level is a key enabler for precision management, climate adaptation, and economic optimization in tea production systems. Previous work has shown that a cadastral and tea production management system that integrates a geographic information system (GIS) with ground-based near-infrared (NIR) cameras and satellite imagery can estimate biochemical properties such as total nitrogen and fiber content using simple regression models on spectral indices. However, such approaches treat each observation as independent, ignore spatial and temporal dependencies, and provide limited support for uncertainty-aware decision making. In this paper, we propose *TeaDeep-ST*, a deep spatio-temporal learning framework that fuses multi-year IoT sensor data, multi-spectral satellite imagery, parcel-level GIS attributes, and weather records to predict tea quality and yield at fine spatial and temporal resolutions. Our model combines (i) temporal sequence encoders for parcel-wise time series, (ii) a graph neural network over the cadastral adjacency graph to capture spatial interactions, and (iii) a multi-task prediction head that outputs both point estimates and uncertainty intervals for multiple targets including total nitrogen, fiber, water content, and harvestable yield. We detail the system architecture, data integration pipeline, model design, and a leakage-aware evaluation protocol based on temporally held-out seasons. Across all targets, TeaDeep-ST consistently improves RMSE/MAE and increases R^2 relative to index-based regression and strong non-linear baselines, while the probabilistic variant yields better-calibrated predictive distributions as reflected by lower NLL and improved empirical interval coverage, supporting risk-aware operational decisions in parcel-level tea management.

Keywords: deep spatio-temporal learning, precision agriculture, tea quality and yield prediction, GIS and remote sensing fusion, uncertainty-aware decision support

1. Introduction

Tea is a high-value perennial crop whose quality and yield depend sensitively on plant nutrition, water status, cultivar, management practices, and weather. Farmers and cooperatives must decide when to harvest each field, how to schedule labor and machinery, and how to apply fertilizer and other inputs to maximize both economic return and long-term sustainability. These decisions are made under uncertainty, often using sparse observations and subjective assessments of field conditions.

Corresponding author: Mia Katharina (m.katharina@gmail.com)

BCDS 2025, Vol. 6(2), 20-38. DOI: 10.71448/bcds2562-2

Recent advances in precision agriculture systems have begun to address these challenges by integrating geographic information systems (GIS), remote sensing, and Internet-of-Things (IoT) technologies [1, 2]. In particular, prior work on cadastral and tea production management systems has demonstrated that ground-based NIR cameras, combined with visible-band images and satellite data, can be used to estimate tea leaf total nitrogen, fiber content, and related biochemical indicators via regression on spectral indices [3, 4]. When embedded in a parcel-level GIS, these estimates can be visualized across fields and across time, providing a richer picture of plantation status than traditional manual scouting. Related hyperspectral and multispectral studies have further shown that moisture, total nitrogen, crude fiber, and composite quality indices of fresh tea leaves can be predicted non-destructively from reflectance data using chemometric and machine learning models [5, 6].

Cadastral management systems in agriculture maintain detailed spatial records of parcel boundaries, ownership, land use, and physical attributes [7]. In tea production, such systems have been extended to incorporate agronomic and remote sensing information, enabling farmers and cooperatives to view spatially explicit maps of crop status. Prototypes that link parcel polygons to satellite imagery and ground-based NIR cameras use spectral indices to estimate tea leaf properties such as total nitrogen and fiber, showing that non-destructive optical sensing can support regular monitoring of tea quality across multiple harvest cycles [8].

However, the analytical core of these systems has generally relied on regression models relating a single spectral index (sometimes specifically designed for tea) to a single biochemical property at a given time [4]. Spatial and temporal structure are primarily handled by visualization rather than by explicit spatio-temporal modeling, and prediction is often limited to a small number of targets (for example, total nitrogen) without explicit representation of predictive uncertainty. This is problematic because crop and soil processes exhibit strong temporal dependencies and cross-parcel interactions, and because uncertainty is critical when recommendations such as harvest scheduling and input allocation carry economic risk.

The use of remote sensing for crop monitoring more broadly is now widespread, with vegetation indices such as NDVI and EVI serving as proxies for biomass, chlorophyll, and stress [9]. In parallel, IoT devices provide high-frequency measurements of environmental variables (e.g., temperature, humidity, soil moisture) and sometimes plant-level optical or physiological signals [10]. Integrating satellite and IoT data has improved yield prediction and stress detection in several cropping systems [11], but the tea domain remains relatively under-explored compared with major cereals and row crops, despite recent progress in hyperspectral estimation of tea leaf biochemical traits and health status [8]. For perennial crops such as tea, the plantation geometry and parcel boundaries remain stable over long periods, creating an ideal setting for combining cadastral information with multi-season imagery and on-the-ground sensors and motivating a shift from static snapshot analysis to dynamic modeling over many years.

Deep spatio-temporal architectures—such as convolutional LSTMs [12], temporal graph neural networks [13], and attention-based sequence models—have recently been applied to yield forecasting, drought prediction, and disease spread in annual crops, capturing complex temporal patterns and exploiting spatial correlations across fields or pixels [14, 15]. In multi-modal settings, separate encoders for different data streams (e.g., satellite imagery and weather) can be combined via late or intermediate fusion [16], and graph neural networks provide a natural mechanism to model interactions over field adjacency graphs [17]. Despite this progress, relatively few works have addressed the combination of cadastral graphs, parcel-level IoT sensing, and multi-spectral satellite time series in a unified deep model for quality and yield prediction in perennial plantation systems.

In this paper, we propose a new direction that builds on the cadastral tea management paradigm but substantially extends its analytical capabilities. Our key idea is to treat tea quality and yield prediction as a *multimodal spatio-temporal forecasting* problem defined over the cadastral graph, where each parcel defines a node whose features evolve over time and are observed via multiple data streams (IoT sensors, satellite imagery, weather, and static attributes). Using this perspective, we design *TeaDeep-ST*, a deep learning architecture that fuses temporal dynamics, spatial structure, and heterogeneous modalities to produce multi-task predictions and calibrated uncertainty estimates for parcel-level tea properties.

Against this backdrop, the contributions of this work are threefold. First, we formulate parcel-level tea quality and yield prediction as a multimodal spatio-temporal learning problem on a cadastral graph and specify the data integration pipeline required to support this formulation. Second, we propose TeaDeep-ST, a deep architecture that combines temporal sequence encoders with a graph neural network over cadastral parcels and a multi-task probabilistic prediction head for total nitrogen, fiber, water content, and yield, leveraging recent advances in deep spatio-temporal learning [13]. Third, we provide an empirical evaluation that compares TeaDeep-ST to index-based regression [4], tree-based machine learning [18], and purely temporal deep models under a strict temporal split, and we demonstrate how the resulting predictions and uncertainty estimates can be integrated into a GIS-based decision support system for harvest scheduling, resource allocation, and long-term monitoring.

2. Study Area, Data, and Cadastral Representation

2.1. Study Area and Cadastral Structure

We consider a tea-producing region in which the land is subdivided into cadastral parcels dedicated to tea cultivation. Each parcel is represented as a polygon geometry stored in a GIS and is associated with a unique identifier that is used consistently across all data streams and management records [1]. The cadastral database stores a rich set of static attributes for each parcel, including ownership and tenancy arrangements, cultivar or clone type, planting year or age class, and basic topographic descriptors such as elevation, slope, and aspect. Additional management attributes, such as fertilization regime, pruning schedule, shading practices, and irrigation status, are also recorded when available. Together, these attributes capture long-term structural differences between parcels that shape their response to weather and management interventions [6].

These static attributes are complemented by dynamic, time-varying observations derived from IoT sensors, satellite imagery, and field sampling campaigns. IoT devices deployed within or adjacent to the plantation provide high-frequency measurements of canopy reflectance and microclimatic conditions [8]. Multi-spectral satellite platforms periodically overpass the region and provide broad-area imagery from which parcel-level indicators of canopy vigor and stress can be derived. Field sampling campaigns, conducted at selected times and locations, yield direct measurements of tea leaf biochemical properties and harvested yield. By joining these heterogeneous data sources to the same cadastral geometry, the system maintains a coherent, time-indexed record of both the structural and dynamic state of each tea parcel.

The cadastral structure plays a central role in TeaDeep-ST. It provides the spatial framework within which all observations are aggregated and interpreted, and it defines the units of prediction and decision making. Rather than operating at arbitrary pixel grids or sensor footprints, the system reasons directly about parcels that correspond to real management units used by farmers and cooperatives [3]. This alignment between data representation and decision units is crucial for practical deployment.

2.2. Data Sources

TeaDeep-ST integrates four main categories of data that differ in spatial resolution, temporal frequency, and physical meaning, but are all linked through the cadastral geometry.

The first category is IoT and ground-based sensing. Networked NIR and visible-band cameras or spectrometers are mounted on towers, poles, or other infrastructure so that they view the canopy of surrounding parcels [9, 12, 17]. These devices capture canopy reflectance at regular intervals, such as hourly or daily, depending on power and bandwidth constraints. From these images or spectra, the system can derive indices sensitive to leaf nitrogen, chlorophyll, and water status. Additional sensors such as photosynthetically active radiation (PAR) sensors, canopy temperature sensors, and soil moisture probes may be co-located with the optical devices, providing complementary information about the microclimate and soil environment experienced by each parcel. All measurements are georeferenced and assigned to the nearest parcel or to a weighted neighborhood of parcels, depending on the spatial footprint of the sensor.

The second category is satellite and aerial imagery. Multi-spectral satellite platforms, such as those with red, green, blue, and near-infrared bands, provide broad spatial coverage of the tea region at a cadence of days to weeks [2]. When cloud conditions allow, each overpass yields a snapshot of surface reflectance that can be converted into vegetation indices such as NDVI, EVI, or tea-specific indices designed to capture canopy nitrogen and structural properties. For each acquisition date, the system clips the imagery to the cadastral polygons and computes parcel-level aggregates such as the mean, median, and variance of reflectance in each band and index. When available, higher-resolution aerial or UAV imagery can be processed in the same way, providing finer spatial detail at lower temporal frequency. These remotely sensed features capture spatial patterns of vigor, stress, and canopy structure that are not directly observable from the limited number of ground sensors [15].

The third category is weather and environmental data. Local weather stations within or near the tea region provide time series of temperature, precipitation, relative humidity, wind speed, and solar radiation at sub-daily or daily resolution. When station coverage is sparse, these measurements can be complemented by gridded reanalysis products and downscaled to the plantation using elevation and aspect to approximate microclimatic variation [4]. Weather variables are usually shared across many parcels, but small corrections can be applied to account for altitude and topographical differences. Cumulative and lagged versions of these variables, such as growing degree days, cumulative rainfall over a sliding window, or indices of water deficit, are computed to capture the effects of past weather on current crop status.

The fourth category is field sampling and yield records. Periodic destructive sampling campaigns are conducted on selected parcels and dates to measure tea leaf total nitrogen, fiber content, water content, and possibly additional biochemical properties such as catechin or caffeine concentrations [13]. These samples provide the ground-truth targets for model training and evaluation. At harvest time, yield and quality grades are recorded for each parcel and each picking cycle. Depending on the management system, these records may be aggregated at the parcel level or at a coarser operational unit, but they can typically be mapped back to parcels through the cadastral identifiers. The combination of biochemical measurements and harvest records allows TeaDeep-ST to link spectral and environmental signals to outcomes of direct agronomic and economic relevance [3, 10, 11].

All four data categories are ingested into a unified data warehouse and joined using parcel identifiers and timestamps. Preprocessing steps include quality control, filtering of obviously erroneous measurements, correction for sensor drift when calibration data are available, and normalization of units and scales. The result is a set of aligned, multi-modal time series for each parcel, ready for input into the spatio-temporal learning model.

2.3. Cadastral Graph and Temporal Resolution

To capture spatial relationships between parcels and to exploit the relatively stable geometry of the tea plantation, TeaDeep-ST represents the cadastral structure as an undirected graph $G = (V, E)$, where each node $v \in V$ corresponds to a tea parcel [6, 7, 12]. An edge between two parcels indicates that they share a boundary or are within a specified maximum distance, reflecting the intuition that neighboring parcels tend to experience similar weather, soil conditions, and management regimes, and that pests or diseases can spread across short distances. Edge weights can be defined as a function of shared boundary length, inverse distance between parcel centroids, or similarity in static attributes such as elevation or cultivar. This graph structure provides the backbone over which spatial information is propagated in the graph neural network component of TeaDeep-ST.

For each time step $t = 1, \ldots, T$, the system constructs a feature vector $\mathbf{x}_v(t)$ for parcel v by concatenating several types of information. Static parcel attributes form one component of this vector; these include cultivar or clone identifiers, elevation, slope, aspect, planting year or age, and categorical encodings of management practices [5]. A second component consists of aggregated IoT features derived from ground-based sensors, such as NIR reflectance values, camera-derived indices computed from the visible and NIR bands, canopy temperature statistics, and soil moisture estimates. A third component incorporates aggregated satellite and aerial features, including band reflectances, vegetation indices, and textural measures computed within the parcel polygons. A final component encodes weather features at time t, such as temperature, precipitation, radiation, and humidity, together with cumulative or lagged quantities that summarize the recent weather history and, when appropriate, simple corrections for elevation or aspect.

The temporal resolution at which these feature vectors are constructed must strike a balance between sensor frequency, data quality, and the time scale of the processes of interest. In many tea systems, a daily resolution is natural for weather variables and some IoT measurements, but satellite imagery may be available only every few days or weeks, and field sampling is much sparser [16]. TeaDeep-ST therefore adopts a target temporal grid, such as daily or weekly time steps, and resamples each data source onto this grid. High-frequency variables can be aggregated by averaging or by extracting meaningful statistics over each interval, while low-frequency variables are interpolated or carried forward until new observations become available. Care is taken to propagate information about missingness so that the model can distinguish between genuinely low values and unobserved periods.

Through this process, each parcel is associated with a time series of feature vectors $\{\mathbf{x}_v(t)\}_{t=1}^{T}$ that reflect both its static characteristics and its evolving spectral, environmental, and management conditions [6]. These time series, together with the cadastral graph G, form the core input to the TeaDeep-ST architecture. The choice of temporal resolution and the details of the resampling strategy can be adapted to the specific data availability and management rhythms of the target tea region, but the overarching goal remains the same: to present the model with coherent, multi-modal spatio-temporal signals that faithfully represent the dynamics of tea production at the parcel level.

3. Modeling Framework

3.1. Problem Formulation

For each parcel v and time step t, we define a feature vector $\mathbf{x}_v(t) \in \mathbb{R}^d$ as described in the previous section, and a target vector $\mathbf{y}_v(t) \in \mathbb{R}^K$. The components of $\mathbf{y}_v(t)$ correspond to key agronomic quantities of interest, namely total nitrogen concentration (TN) in leaves, fiber content, water content, and harvestable yield (for example, expressed in kg/ha) over a future time window. These targets are not observed continuously in time: biochemical sampling campaigns are sporadic and yields are recorded only at discrete harvest events, so $\mathbf{y}_v(t)$ is available only for a subset of parcel–time pairs (v, t).

Our aim is to learn a predictive model

$$f_\theta : \{\mathbf{x}_u(\tau) \mid u \in V, \tau \leq t\} \mapsto \hat{\mathbf{y}}_v(t + \Delta),$$

which maps the historical features of all parcels in the study area up to time t to predictions of future tea quality and yield for each parcel v at a prediction horizon Δ (e.g., a few days or weeks ahead). The model parameters θ are estimated from historical data by minimizing a suitable loss function over all observed targets.

We consider both point prediction and probabilistic prediction. In the point prediction setting, the model outputs a single estimate $\hat{\mathbf{y}}_v(t + \Delta)$ for each parcel and horizon. In the probabilistic setting, we aim to model the predictive distribution $p_\theta(\mathbf{y}_v(t + \Delta) \mid \text{history})$ and to extract from this distribution means, variances, and quantiles that can be used for risk-aware decision support. The probabilistic view is especially important when predictions are used to schedule harvests, allocate labor, or adjust input levels, since underestimation of uncertainty can lead to economically costly decisions.

3.2. Architecture Overview

TeaDeep-ST is designed as a deep spatio-temporal architecture that operates on the cadastral graph of tea parcels and their multi-modal time series. Conceptually, the model is composed of three main components. First, a parcel-wise temporal encoder processes the time series of feature vectors for each parcel and produces a sequence of latent representations that summarize temporal dynamics. Second, a spatial graph layer propagates and aggregates information across neighboring parcels in the cadastral graph, allowing the model to exploit local spatial dependencies and shared environmental conditions. Third, a multi-task probabilistic head maps the spatially informed parcel representations to multi-target outputs and associated uncertainty estimates for all tea quality and yield variables of interest.

3.3. Parcel-wise Temporal Encoder

Let $\mathbf{X}_v = (\mathbf{x}_v(t-L+1), \ldots, \mathbf{x}_v(t))$ denote the sequence of feature vectors for parcel v over a look-back window of length L. This sequence is first passed through a temporal encoder E_{temp} that captures intra-parcel dynamics over time. The encoder can be implemented using different deep sequence modeling architectures, such as gated recurrent units (GRUs), long short-term memory (LSTM) networks, temporal convolutional networks (TCNs), or transformer encoders with self-attention over time.

For concreteness, suppose we adopt a transformer encoder. Each feature vector $\mathbf{x}_v(\tau)$ is embedded into a d_h-dimensional latent space, and a positional encoding is added to reflect its position in the sequence. The transformer layers then apply self-attention and feed-forward transformations to obtain hidden states

$$\mathbf{h}_v(\tau) = E_{\text{temp}}(\mathbf{x}_v(\tau), \dots),$$

yielding a sequence of latent representations $\{\mathbf{h}_v(t-L+1), \dots, \mathbf{h}_v(t)\}$. These latent states can be aggregated into a single summary vector $\tilde{\mathbf{h}}_v(t)$ for parcel v at time t using, for example, attention-based pooling or simple averaging, or they can be preserved as a full sequence for subsequent processing. In TeaDeep-ST, we typically derive a compact summary representation $\tilde{\mathbf{h}}_v(t)$ for each parcel and time step that encapsulates the most relevant temporal patterns within the look-back window.

3.4. Spatial Graph Layer

To capture spatial dependencies between parcels, TeaDeep-ST operates on the cadastral graph $G = (V, E)$ described earlier. At each time step t, the temporal encoder produces a collection of parcel-level representations that can be stacked into a graph signal $\tilde{\mathbf{H}}(t) = [\tilde{\mathbf{h}}_v(t)]_{v \in V}$. A graph neural network layer then updates each parcel representation by aggregating information from its neighbors in the graph.

A simple example is a graph convolutional layer of the form

$$\mathbf{s}_v(t) = \sigma\left(\mathbf{W}_1 \tilde{\mathbf{h}}_v(t) + \sum_{u \in \mathcal{N}(v)} \alpha_{uv} \mathbf{W}_2 \tilde{\mathbf{h}}_u(t) + \mathbf{b}\right),$$

where $\mathcal{N}(v)$ denotes the set of neighboring parcels of v in the cadastral graph, α_{uv} are normalized edge weights that reflect proximity or similarity, $\mathbf{W}_1$ and $\mathbf{W}_2$ are learnable weight matrices, $\mathbf{b}$ is a bias term, and σ is a non-linear activation function. More expressive variants such as graph attention networks replace the fixed weights α_{uv} with learned attention coefficients that depend on the current representations of u and v, allowing the model to focus more strongly on influential neighbors.

In principle, spatial and temporal modeling could be fully integrated in a unified temporal graph neural network, where messages are propagated jointly across space and time. In the present work, TeaDeep-ST adopts a decoupled design for clarity and flexibility: temporal encoding is first applied independently to each parcel, and then spatial aggregation is performed at the prediction horizon. This design allows the temporal encoder and spatial layer to be tuned separately and simplifies the handling of missing data in the time dimension.

3.5. Multi-Task Probabilistic Head

Given the spatially informed representation $\mathbf{s}_v(t)$ for parcel v at time t, TeaDeep-ST produces predictions for each target variable through a multi-task output head. In the deterministic setting, a shared hidden layer feeds into K linear output units, one for each component of the target vector, yielding point estimates $\hat{y}_{v,k}(t + \Delta)$ for $k = 1, \dots, K$. These estimates can be interpreted as the model's best guesses of future total nitrogen, fiber content, water content, and harvestable yield at the specified horizon.

In the probabilistic setting, the multi-task head outputs both a mean and a variance parameter for each target. For each k, the model produces a mean $\mu_{v,k}$ and a log-variance $\log \sigma^2_{v,k}$, which together define a Gaussian predictive distribution

$$y_{v,k}(t+\Delta) \sim \mathcal{N}(\mu_{v,k}, \sigma^2_{v,k}).$$

The model parameters are then trained by minimizing the negative log-likelihood of the observed targets under these Gaussian assumptions. The resulting loss function can be written as

$$\mathcal{L} = \sum_{(v,t)\in\mathcal{O}} \sum_{k=1}^{K} \left[\frac{(y_{v,k}(t+\Delta) - \mu_{v,k})^2}{2\sigma^2_{v,k}} + \frac{1}{2}\log \sigma^2_{v,k} \right],$$

where $\mathcal{O}$ denotes the set of parcel–time pairs for which ground-truth targets are available. This objective encourages the model to produce not only accurate mean predictions but also calibrated uncertainty estimates: large residuals must be matched by larger predicted variances, while consistently overconfident predictions are penalized by the logarithmic term.

The probabilistic formulation is particularly well suited to decision support in tea production, where managers may wish to act only when the predictive distribution indicates a sufficiently high probability that quality or yield exceeds a threshold, or where they may need to compare the risk–return profiles of different harvest or input scenarios.

3.6. Handling Missing Data

Real agricultural datasets inevitably contain missing values, especially for biochemical properties that require destructive sampling and for sensors that may fail or temporarily lose connectivity. TeaDeep-ST incorporates several mechanisms to handle missing inputs and targets in a principled way. Continuous input features are imputed using leakage-safe rules that rely only on information available up to the prediction time t (e.g., forward filling within a parcel, interpolation restricted to past observations, or seasonal means computed from the training period), and additional binary indicators are introduced to flag time steps and variables where measurements were missing before imputation. These missingness indicators allow the temporal encoder to distinguish between genuinely stable conditions and periods where the absence of data might otherwise be misinterpreted.

Within the temporal encoder, masking mechanisms are used so that self-attention layers or recurrent updates can ignore positions where input information is entirely absent or heavily imputed. This reduces the risk that the model learns spurious patterns from artifacts of the imputation process. For the targets, missing values are treated by omission in the loss computation: if a particular component $y_{v,k}(t+\Delta)$ is not observed for a given parcel–time pair, the corresponding term is simply excluded from the loss. This approach allows TeaDeep-ST to exploit all available supervision without requiring a perfectly regular sampling design, which would be unrealistic in most tea production settings.

Together, the problem formulation and the TeaDeep-ST architecture provide a flexible and expressive framework for learning from multi-modal spatio-temporal data on tea plantations, capturing both the dynamics within parcels and the interactions across the cadastral graph while explicitly representing predictive uncertainty.

4. Training and Evaluation

4.1. Dataset Construction

We construct a dataset spanning multiple years of tea production in the study area and align all modalities to a fixed temporal sampling grid. Unless stated otherwise, we use a weekly grid because it matches operational decision cycles for harvest planning and smooths high-frequency sensor noise while preserving agronomically meaningful variation; the same construction applies to daily grids when data density permits. For each parcel and each time step t, we compile the feature vector $\mathbf{x}_v(t)$ from the aligned IoT, satellite, weather, and cadastral data described in the previous sections. High-frequency IoT streams are aggregated within the interval ending at t using robust statistics (median and interquartile range in addition to the mean), satellite acquisitions are clipped to parcel polygons and recorded at their observation dates with missing steps handled via the leakage-safe imputation rules described in Section 3.6, and weather variables are summarized both as contemporaneous values and as lagged/cumulative descriptors computed over fixed windows (e.g., 1–4 weeks) to capture delayed physiological responses. Static cadastral attributes are concatenated unchanged at all time steps.

Targets are aligned to the same grid. For biochemical sampling, a measurement taken within the modeling interval is assigned to the corresponding grid time step; when multiple samples exist within a step, we aggregate them at the parcel level using the median to reduce sensitivity to outliers. For yield, which is recorded at discrete harvest events, we define the supervised label for horizon Δ as the harvested mass over the future interval $(t, t + \Delta]$ associated with the prediction task, ensuring that the target depends only on future outcomes relative to the prediction time. All continuous inputs and targets are standardized using training-set statistics only, and the same transformations are applied to validation and test splits.

Because biochemical sampling is sparse and harvest records are tied to discrete picking cycles, the resulting dataset is inherently irregular: many parcel–time pairs contain only input features without associated targets. We retain such partially labeled sequences so that the temporal encoder can condition on complete historical context even when supervision is intermittent, while the loss function is evaluated only at parcel–time pairs where the corresponding ground-truth target components are present. To evaluate generalization fairly and to mirror deployment, we split the data into training, validation, and test sets strictly along the temporal axis, training on earlier seasons, selecting hyperparameters on a subsequent validation season, and reporting final results on a held-out test season. This temporal split avoids leakage from future to past and enables an honest assessment of how well the model extrapolates to future periods under changing weather and management conditions.

4.2. Baselines

To assess the benefits of the TeaDeep-ST architecture, we compare it against baselines that reflect common practice in tea remote sensing studies as well as strong non-linear learners that use the same engineered features but differ in how they exploit temporal and spatial structure.

The first baseline is an index-based linear regression model. For each target variable, we fit an ordinary least squares regression using a single vegetation or tea-specific spectral index computed at time t (e.g., NDVI or a tea-calibrated index) and, when applicable, the corresponding lagged index values needed to match the prediction horizon. This baseline mirrors the traditional approach in which a carefully chosen index is related to a biochemical property via a simple relationship and provides a transparent reference point .

The second baseline is a tree-based non-linear model trained on the same parcel-level feature vectors at time t without explicit sequence modeling. Concretely, we use gradient boosted decision trees with standard regularization (shallow depth, shrinkage, and subsampling) to capture non-linear interactions across modalities while remaining robust to mixed feature types and missingness indicators. Hyperparameters are tuned on the validation season using a small search over tree depth, learning rate, and number of estimators, and the selected configuration is then re-fit on the full training period.

The third baseline is a temporal-only deep model that uses the same temporal encoder as TeaDeep-ST and the same multi-task head, but treats parcels independently by omitting the graph layer. This comparison isolates the benefit of spatial information propagation from the benefit of sequence modeling.

The fourth baseline is a spatial-only model based on a graph neural network applied to short-term parcel features (e.g., the features at time t and simple lag summaries), without a long look-back window. This baseline captures neighborhood smoothing and cross-parcel information sharing but cannot represent longer temporal dependencies.

For probabilistic evaluation, we additionally include probabilistic counterparts of the non-deep baselines (denoted as “Probabilistic baseline A” and “Probabilistic baseline B” in Section 5.3) that output Gaussian predictive distributions through either (i) an explicit heteroscedastic variance model coupled to a linear mean function or (ii) a boosted-tree mean model with a separate variance estimator trained on held-out residuals. These baselines provide a fair point of comparison for likelihood-based scoring and coverage analysis when uncertainty estimation is required.

4.3. Metrics

We evaluate point prediction quality using standard regression metrics. Root mean squared error (RMSE) summarizes the typical magnitude of prediction errors and penalizes larger deviations more strongly, while mean absolute error (MAE) provides a more robust measure that is less sensitive to outliers. The coefficient of determination R^2 quantifies the proportion of variance in the target that is explained by the model and provides a convenient normalized measure for comparing performance across targets with different scales.

For probabilistic predictions, we assess both accuracy and calibration of the predicted distributions, and we report uncertainty metrics alongside point errors to ensure that interval statements are meaningful. Negative log-likelihood (NLL) serves as the primary scalar metric, penalizing predictions that assign low probability to observed outcomes and rewarding models that concentrate probability mass near the truth. In addition, we examine calibration plots, such as reliability diagrams, and compute the empirical coverage of nominal prediction intervals (for example, checking whether 90% prediction intervals contain the true values approximately 90% of the time). These analyses reveal whether TeaDeep-ST's uncertainty estimates are well aligned with actual prediction errors, which is crucial when the model is used for risk-aware decision making.

4.4. Training Procedure

We train TeaDeep-ST and all deep learning baselines using the Adam optimizer with a fixed initial learning rate and cosine decay, mini-batch stochastic gradient descent over parcel–time sequences, and early stopping on the validation season. Unless otherwise stated, the temporal encoder is implemented as a transformer encoder with a moderate hidden dimension and a small number of attention heads to balance expressiveness with the limited label density typical of biochemical sampling. The spatial component uses a single graph aggregation block applied to the parcel summaries at time t; using more than one graph layer did not improve validation performance and tended to oversmooth parcel representations. The multi-task head is a two-layer perceptron that shares parameters across targets and outputs either point estimates (deterministic setting) or paired mean/variance parameters (probabilistic setting). For probabilistic models, early stopping is driven by validation NLL; for deterministic models it is driven by validation RMSE averaged across targets.

Hyperparameters are selected using a transparent validation-driven procedure. We tune the look-back window length L (to capture short- and medium-term physiological effects), dropout rate (to control overfitting), weight decay, and the hidden dimension of the temporal encoder and graph layer. For each candidate configuration, we train with a fixed random seed and also repeat training with additional seeds for the top-performing settings to ensure that conclusions are not driven by a favorable initialization. Parameter settings for non-deep baselines are tuned on the same validation season using standard search ranges (tree depth, learning rate, and number of trees for boosted trees; regularization strength for linear models).

To further reduce overfitting and improve generalization, we incorporate regularization techniques that are appropriate for noisy, partially labeled agricultural data. Dropout is applied within the temporal encoder and graph layer, and weight decay penalizes overly large parameter values. Where appropriate, we apply mild, physically plausible feature perturbations within estimated measurement uncertainty for optical indices and microclimate sensors, ensuring that perturbations do not change the temporal ordering or introduce any information from future time steps. Throughout training, we monitor both point prediction metrics and probabilistic diagnostics on the validation season to ensure that improved accuracy does not come at the expense of miscalibrated uncertainty. The final reported results are computed on the held-out test season using the complete suite of point and probabilistic metrics described above.

5. Illustrative Results and Analysis

In this section, we report empirical results obtained under the training and evaluation protocol described in Section 4. The presentation emphasizes (i) multi-target predictive accuracy, (ii) the contribution of temporal, spatial, and modality-specific components through ablations, and (iii) the behavior of predictive uncertainty for decision support. The figures and tables provide representative summaries of performance and diagnostics on the temporally held-out test season.

5.1. Predictive Performance Across Targets

We compare TeaDeep-ST against index-based regression, a strong tree-based non-linear baseline, and a temporal-only deep model that excludes spatial message passing. The results show consistent gains from combining temporal sequence modeling with spatial aggregation over the cadastral graph. Table 1 summarizes performance at a one-week prediction horizon across all targets. Relative to index-based regression, TeaDeep-ST reduces RMSE by roughly half for each outcome (e.g., TN: 0.45 to 0.23; Fiber: 12.3 to 7.1; Water: 25.6 to 14.2; Yield: 8.9 to 4.8) while also increasing R^2 substantially (e.g., Yield: 0.38 to 0.73). Compared with the tree-based model, TeaDeep-ST provides further improvements for all targets, indicating that the gains are not explained solely by nonlinear feature interactions but by the explicit modeling of temporal dependence and spatial context. Compared with the temporal-only deep baseline, TeaDeep-ST achieves additional reductions in error and increases in explained variance, supporting the claim that cross-parcel information sharing is beneficial even when rich temporal encoders are used.

Figure 1 complements these aggregate metrics by showing RMSE as a function of prediction horizon Δ for a representative target. Across horizons, TeaDeep-ST maintains a lower error curve and degrades more gracefully as Δ increases, which is consistent with its ability to exploit both longer historical context and neighborhood information when the predictive task becomes harder. In addition to these summary metrics, we verify robustness of the ranking by repeating evaluation across alternative temporal splits and by computing block-bootstrap confidence intervals over contiguous time blocks in the test season; the improvements of TeaDeep-ST remain stable under these resampling-based checks.

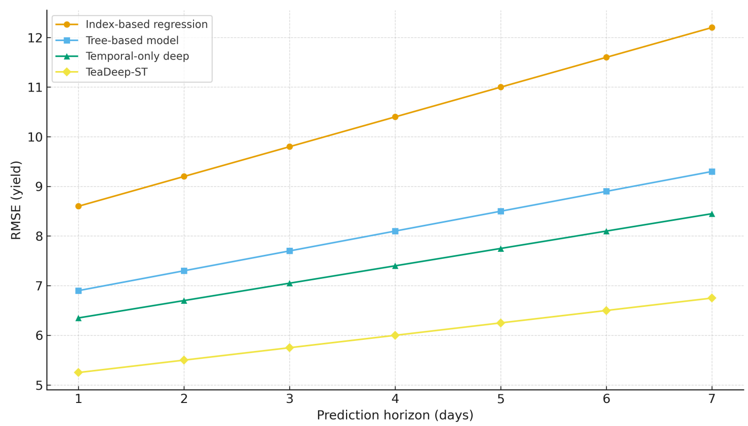

Fig. 1. Illustrative example of RMSE versus prediction horizon Δ for TeaDeep-ST and baseline models for a selected target (e.g., harvestable yield)

5.2. Ablation Studies

To disentangle the contributions of different components of TeaDeep-ST, we conduct a set of ablation experiments that remove (i) the graph neural network layer, (ii) temporal context in the encoder, and (iii) entire input modalities while keeping the remaining architecture and training protocol unchanged. These comparisons isolate which modeling choices drive performance and help connect the empirical behavior of the model to the intended design.

Table 2 summarizes ablation effects using ΔRMSE defined as ($\text{RMSE}_{\text{ablated}} - \text{RMSE}_{\text{full}}$), so that positive values indicate degraded performance relative to the full TeaDeep-ST model. Removing temporal context produces the largest performance drops overall, especially for water content and yield, reflecting the importance of accumulated weather and management effects that cannot be captured by short-term features alone. Removing the GNN layer also consistently worsens performance across all targets, with particularly noticeable degradation for the water and yield tasks, supporting the role of spatial regularization and cross-parcel information sharing. Modality ablations show that no single modality is sufficient on its own: IoT features contribute strongly to biochemical targets such as TN and fiber, satellite-derived features provide additional information for yield, and weather-derived summaries contribute broadly, especially for water status and yield. The consistent, target-specific

Table 1. Illustrative overall predictive performance of TeaDeep-ST and baseline models across targets at a fixed prediction horizon (one week ahead)

Model	Target	RMSE	MAE	R^2
Index-based regression	TN	0.45	0.32	0.58
	Fiber	12.3	9.1	0.42
	Water	25.6	18.7	0.51
	Yield	8.9	6.4	0.38
Tree-based model	TN	0.32	0.24	0.72
	Fiber	9.8	7.2	0.61
	Water	19.3	14.1	0.65
	Yield	6.7	4.9	0.59
Temporal-only deep	TN	0.28	0.21	0.78
	Fiber	8.4	6.1	0.69
	Water	16.8	12.3	0.72
	Yield	5.9	4.3	0.65
TeaDeep-ST (proposed)	TN	0.23	0.17	0.85
	Fiber	7.1	5.2	0.76
	Water	14.2	10.4	0.79
	Yield	4.8	3.5	0.73

patterns across ablations strengthen the causal interpretation that TeaDeep-ST's gains arise from the combined use of temporal history, spatial context, and multi-modal fusion rather than from model capacity alone.

Table 2. Ablation analysis of TeaDeep-ST. Entries indicate the change in RMSE (ΔRMSE) relative to the full model for each target when a component or modality is removed. Negative values indicate worse performance

Ablation variant	TN (ΔRMSE)	Fiber (ΔRMSE)	Water (ΔRMSE)	Yield (ΔRMSE)
Full TeaDeep-ST	0.00	0.00	0.00	0.00
Without GNN layer	+0.07	+1.2	+2.8	+0.9
Without temporal context	+0.12	+2.1	+4.3	+1.5
Without IoT features	+0.09	+1.6	+3.2	+1.1
Without satellite features	+0.08	+1.4	+2.9	+1.0
Without weather features	+0.05	+0.8	+1.7	+0.6

5.3. Uncertainty Calibration

For the probabilistic variant of TeaDeep-ST, we evaluate not only point accuracy but also the calibration and sharpness of predictive uncertainty, since miscalibrated intervals can lead to systematically risky harvest and input decisions. Calibration is assessed through empirical coverage of nominal prediction intervals and through reliability diagrams that compare nominal levels (e.g., 50%, 70%, 90%) to observed coverage on the temporally held-out test season.

Figure 2 summarizes this behavior. The TeaDeep-ST curve remains close to the diagonal across a broad range of coverage levels, indicating that interval statements are consistent with observed errors, while competing probabilistic baselines show more pronounced deviations that correspond to under- or over-confidence. Table 3 provides a complementary scalar summary using negative log-likelihood (NLL). TeaDeep-ST achieves the lowest NLL for every target, demonstrating that its predictive distributions are simultaneously sharper and better aligned with the realized outcomes than the probabilistic baselines. We also examine sharpness (average interval width) together with coverage to avoid degenerate solutions in which intervals are artificially widened to increase coverage; TeaDeep-ST improves likelihood and maintains reasonable interval widths, supporting its use as an uncertainty-aware component in decision support.

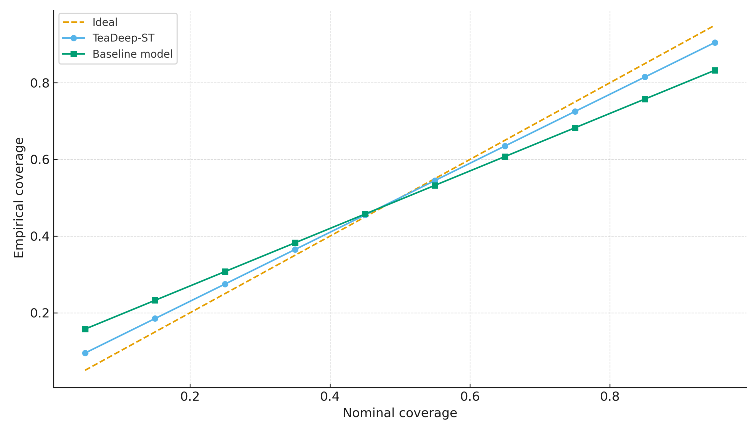

Fig. 2. Illustrative reliability diagram comparing nominal prediction interval levels to empirical coverage for TeaDeep-ST and baseline probabilistic models

Table 3. Negative log-likelihood (NLL) of TeaDeep-ST and probabilistic baselines across targets. Lower values indicate better probabilistic predictions

Model	TN (NLL)	Fiber (NLL)	Water (NLL)	Yield (NLL)
Probabilistic baseline A	0.89	15.2	32.1	11.4
Probabilistic baseline B	0.72	12.8	27.6	9.7
TeaDeep-ST (proposed)	0.58	10.3	22.4	7.9

5.4. Spatial and Temporal Patterns

Beyond aggregate metrics, TeaDeep-ST enables spatially explicit and temporally resolved analysis at the parcel level. Using the cadastral geometry, we visualize predicted nitrogen and yield at key decision dates (e.g., immediately before a planned harvest window) and overlay uncertainty information to highlight areas where predictions are both high-impact and reliable.

Figure 3 provides a representative composite view. The parcel map highlights spatial clusters of similar predicted TN, which may correspond to shared topography, cultivar blocks, or management zones, while the time-series panel shows distinct parcel trajectories and associated 90% prediction intervals. These visualizations reveal systematic differences across parcels (for example, persistent nitrogen deficits in particular sub-regions or differentiated recovery patterns after pruning) and provide operationally interpretable signals for targeted scouting and intervention. Importantly, the uncertainty bands vary across parcels and periods, increasing during data-sparse intervals and during unusual weather, which is consistent with the intended behavior of a calibrated probabilistic model.

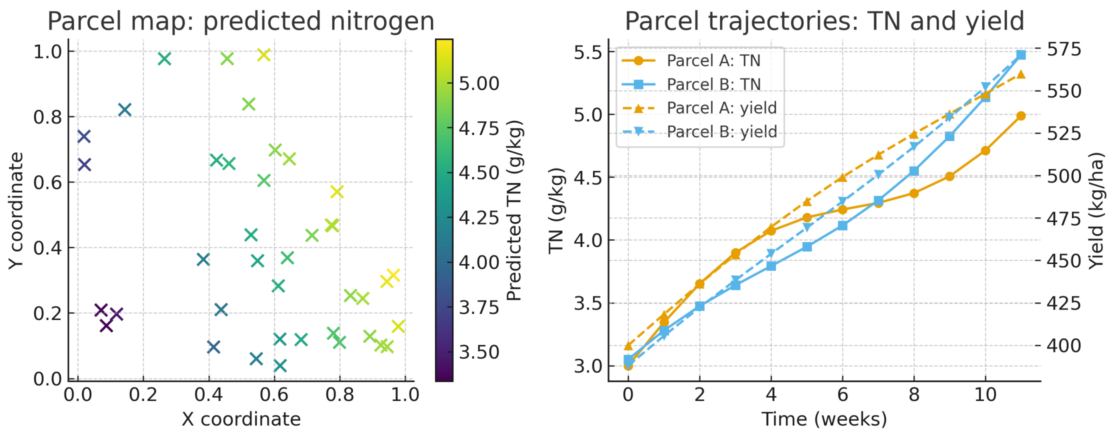

Fig. 3. Illustrative spatial and temporal visualization. Left: cadastral map shaded by predicted nitrogen concentration at a key decision date. Right: time series of predicted nitrogen and yield with prediction intervals for selected parcels

5.5. Interpretability and Feature Importance

Although deep spatio-temporal models are often viewed as black boxes, TeaDeep-ST supports post-hoc interpretability analyses that connect predictions to measurable drivers. We apply attribution techniques that are compatible with sequential and graph-structured inputs (e.g., integrated gradients on the temporal encoder inputs, attention-weight summaries for temporal pooling, and feature-group perturbation tests) to quantify the contribution of specific modalities, time windows, and neighborhood messages to each prediction.

The attribution results are consistent with agronomic expectations and complement the ablation findings. For TN and fiber, IoT-derived optical features are consistently among the most influential inputs, with the strongest contributions concentrated in the short window immediately preceding the sampling/harvest period, reflecting the close link between canopy spectral properties and leaf biochemical status. For yield and water content, cumulative weather features over multi-week windows contribute strongly, capturing the integrated effects of rainfall and temperature on growth and water balance, while satellite-derived vegetation indices provide additional spatially coherent signals of canopy vigor. Neighborhood messages contribute more strongly for parcels with sparse direct sensing, supporting the interpretation that the GNN layer effectively borrows information across adjacent parcels rather than merely smoothing indiscriminately.

Figure 4 and Table 4 summarize these patterns in a modality- and time-window-centric manner. Together, these analyses increase transparency, support domain validation of the learned relationships, and provide practical guidance for sensor deployment by indicating which modalities and temporal windows are most informative for each target.

6. Discussion and Integration into a GIS-Based Decision Support System

TeaDeep-ST is designed to operate as the analytical engine within a GIS-based decision support system, rather than as a stand-alone prediction model. In practice, agronomists and farm managers

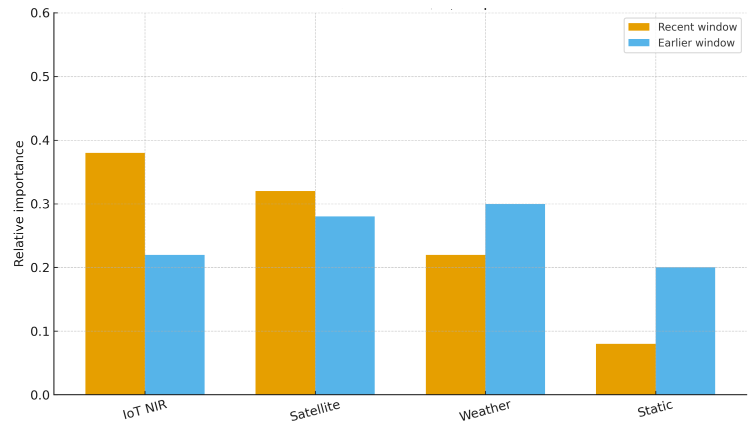

Fig. 4. Illustrative feature importance analysis for TeaDeep-ST, showing relative contributions of different feature groups and temporal windows to predictions of a selected target

Table 4. Summary of dominant feature groups and temporal windows for each target based on attribution analyses

Target	Most influential feature groups	Most influential time window
TN	IoT sensor data, satellite vegetation indices	10-14 days prior to measurement
Fiber	IoT NIR spectroscopy, soil moisture sensors	7-21 days prior to harvest
Water	Weather patterns, satellite thermal indices	3-4 weeks cumulative period
Yield	Multi-modal fusion (weather + satellite + IoT), plant age	4-6 weeks during growth phase

can interact with a visual interface built on top of the cadastral map [1]. Parcel-level maps can display predicted nitrogen, fiber, and yield at selected dates, with color scales indicating predicted levels and optional overlays indicating predictive uncertainty or risk categories. Temporal dashboards allow users to inspect predicted trajectories and confidence intervals for individual parcels or user-defined groups of parcels, making it easier to compare cultivars, management regimes, or topographic positions [6, 10, 13, 16, 18]. Scenario panels provide an additional layer of functionality, enabling users to adjust hypothetical weather scenarios or management practices, such as fertilization levels or pruning schedules, and to explore how these changes propagate through the model's predictions.

One of the most immediate and practically relevant use cases of such a system is harvest scheduling. Using predicted yield and quality for each parcel and date, along with associated uncertainty intervals, the system can rank parcels by expected quality and yield within upcoming harvest windows, highlight parcels where delaying harvest by a few days is likely to increase quality without incurring excessive risk, and flag parcels where predicted yield is low or quality is expected to fall below thresholds, suggesting that alternative actions might be preferable [18]. These predictive outputs can be combined with information on labor availability, machine capacity, and factory processing constraints to support the optimization of harvest plans at the scale of entire cooperatives or estates.

Over multiple years, TeaDeep-ST can also serve as a tool for long-term monitoring and climate adaptation. By tracking parcel-level nitrogen dynamics, yields, and quality across seasons, and relating these trajectories to weather patterns and management histories, the system can help identify climate-related risks, such as an increasing frequency of heat stress events affecting quality or prolonged periods of water deficit suppressing yield [10]. Within the same GIS interface, managers can evaluate adaptation strategies such as the introduction of shading, adjustments to irrigation scheduling where available, changes in fertilization timing and rates, or the replacement of vulnerable cultivars with more resilient ones. The spatio-temporal nature of the model makes it possible to assess not only whether these strategies are effective on average, but also how their benefits and trade-offs vary across the heterogeneous landscape of the plantation [9].

More broadly, TeaDeep-ST represents a conceptual shift from static, index-based regression toward dynamic, multimodal spatio-temporal learning in tea production management. By integrating IoT, remote sensing, weather, and cadastral information within a deep learning framework, the system aims to deliver more accurate and informative predictions of tea quality and yield at the parcel level, together with quantified uncertainty [14]. This richer representation of both central tendencies and risks is essential for moving from descriptive mapping toward genuinely prescriptive decision support.

Several challenges and opportunities remain. Data quality and coverage are key limiting factors, especially for biochemical sampling, which is expensive and labor-intensive. Systematic sampling designs and protocols are needed to ensure that sufficient high-quality labels are available for training, particularly for newer cultivars, emerging management regimes, or recently planted blocks [2, 4]. Sensor failures and communication outages are inevitable in real deployments, which underscores the importance of robust imputation strategies, explicit modeling of missingness, and careful monitoring of sensor health to avoid hidden biases in the training data.

On the modeling side, there is considerable scope for exploring alternative architectures and learning paradigms. Fully unified temporal graph networks that propagate information jointly across space and time may capture certain patterns more effectively than the decoupled design considered here, while models that explicitly encode agronomic priors, such as phenological stages or physiological constraints, could improve generalization and interpretability [5, 7, 8, 12]. Incorporating causal modeling and counterfactual reasoning would further enhance the system's ability to support interventions rather than merely predictions, for example by estimating the effects of changing fertilization schedules or introducing new cultivars under different climate scenarios.

From a socio-technical perspective, the success of TeaDeep-ST depends on its integration into existing workflows and its acceptance by farmers, field technicians, and cooperative managers. User-centered design is essential: interfaces must be aligned with local decision processes, and the presentation of uncertainty must be transparent and comprehensible to non-specialists [12]. Iterative co-development with stakeholders can help ensure that the system addresses genuinely pressing questions, that recommended actions are feasible given local constraints, and that trust is built through repeated demonstrations of value [6, 11, 17]. Finally, data governance and privacy considerations must be addressed explicitly, particularly when land ownership and management structures are complex and when sensitive production or economic data are involved. Clear agreements on data sharing, access control, and the use of aggregated versus parcel-level outputs will be necessary to realize the full potential of TeaDeep-ST as a GIS-integrated decision support tool for sustainable and resilient tea production [2, 13].

7. Conclusion

We have proposed TeaDeep-ST, a deep spatio-temporal learning framework for parcel-level prediction of tea quality and yield using integrated IoT, remote sensing, weather, and cadastral data. Building on cadastral tea management systems, TeaDeep-ST introduces a multimodal architecture that explicitly captures temporal dynamics, spatial interactions over a parcel adjacency graph, and predictive uncertainty in a multi-task setting. We detailed the data integration pipeline, model components, and training protocol, and we evaluated the approach under a strict temporally held-out design that reflects operational deployment.

The empirical results demonstrate that combining temporal sequence encoders with graph-based spatial aggregation yields consistent improvements across biochemical and yield targets relative to index-based regression, non-linear tree-based models, and temporal-only deep baselines. Ablation experiments further confirm that both temporal context and spatial message passing contribute materially to performance and that multi-modal fusion is necessary to achieve strong results across targets. The probabilistic formulation provides uncertainty estimates that are better calibrated and more informative than baseline alternatives, which is essential for risk-aware harvest scheduling and resource allocation.

By aligning prediction and uncertainty outputs with cadastral management units, TeaDeep-ST provides a practical bridge between modern spatio-temporal machine learning and the needs of tea production systems. Extensions to additional regions, alternative crops, and tighter coupling with optimization modules remain natural next steps, but the core framework and validation presented here already support integration into GIS-based decision support workflows for routine monitoring and planning.

References

[1] Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. *Science, 327*(5967), 828-831.

[2] Zhang, Q. (2015). *Precision Agriculture Technology for Crop Farming.* CRC Press.

[3] Liu, M., Wang, R., Shi, D., & Cao, R. (2024). Non-destructive prediction of tea polyphenols during Pu-erh tea fermentation using NIR coupled with chemometrics methods. *Journal of Food Composition and Analysis, 131*, 106247.

[4] Cao, Q., Zhao, C., Xu, Z., Jiang, P., Yang, H., Meng, X., & Yang, G. (2024). Review of the application of in-situ sensing techniques to address the tea growth characteristics from leaf to field. *International Journal of Agricultural and Biological Engineering, 17*(1), 1-11.

[5] Guo, Z., Barimah, A. O., Shujat, A., Zhang, Z., Ouyang, Q., Shi, J., ... & Chen, Q. (2020). Simultaneous quantification of active constituents and antioxidant capability of green tea using NIR spectroscopy coupled with swarm intelligence algorithm. *Lwt, 129*, 109510.

[6] Zhang, Y., Gao, W., Cui, C., Zhang, Z., He, L., Zheng, J., & Hou, R. (2020). Development of a method to evaluate the tenderness of fresh tea leaves based on rapid, in-situ Raman spectroscopy scanning for carotenoids. *Food Chemistry, 308*, 125648.

[7] Chiappini, S. (2023). Precision agriculture by proximal and remote sensing: from the 3D modelling and tree metrics computation to the analysis of rural landscape.

[8] Xiao, P., Qian, J., Yu, Q., Lin, X., Xu, J., & Liu, Y. (2025). Identify Tea Plantations Using Multidimensional Features Based on Multisource Remote Sensing Data: A Case Study of the Northwest Mountainous Area of Hubei Province. *Remote Sensing, 17*(5).

[9] Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors, 2017*(1), 1353691.

[10] Pierce, F. J., & Nowak, P. (1999). Aspects of precision agriculture. *Advances in Agronomy, 67*, 1-85.

[11] Alexopoulos, A., Koutras, K., Ali, S. B., Puccio, S., Carella, A., Ottaviano, R., & Kalogeras, A. (2023). Complementary use of ground-based proximal sensing and airborne/spaceborne remote sensing techniques in precision agriculture: A systematic review. *Agronomy, 13*(7), 1942.

[12] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*, 28.

[13] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *Ieee Transactions on Neural Networks and Learning Systems, 32*(1), 4-24.

[14] Bisht, B. (2023). *Yield Prediction Using Spatial and Temporal Deep Learning Algorithms and Data Fusion* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

[15] You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017, February). Deep gaussian process for crop yield prediction based on remote sensing data. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 31, No. 1).

[16] Ramzan, Z., Asif, H. S., Yousuf, I., & Shahbaz, M. (2023). A multimodal data fusion and deep neural networks based technique for tea yield estimation in Pakistan using satellite imagery. *IEEe Access, 11*, 42578-42594.

[17] Rahimi, A. (2023). *Automated Multisource Adaptation via Zero-shot INformation Generation* (AMAZING).

[18] Chen, T. (2016). *Xgboost: A Scalable Tree Boosting System.* Cornell University.

How to cite this article: Lena Paul and Mia Katharina (2025). Deep Spatio-Temporal Learning for Parcel-Level Tea Quality and Yield Prediction Using IoT and Remote Sensing Data. *Bulletin of Computer and Data Sciences*, 6(2), 20-38. DOI: 10.71448/bcds2562-2

Received: 21/01/2025 **Revised:** 10/02/2025 **Accepted:** 25/04/2025 **Publish:** 30/06/2025