International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(21s) 2026 pp. 27-49
https://doi.org/10.70917/ijcisim-2026-5277
Article

Kashmir-MungFormer: An Explainable Multimodal Transformer Framework for Green Gram Moisture-Stress and Yield Prediction under Temperate Conditions

Muneer Ahmad Shigan1*, Prince Arora[2], Ajaz Ahmad Lone[3]

[1]School of Computer Applications,Lovely Professional University, Punjab. Email: reenum03@gmail.com
[2]School of Computer Applications,Lovely Professional University, Punjab. Email: princearorabca@gmail.com
[3]Dryland Agriculture Research Station,Sher-e-Kashmir University of Agricultural Sciences and Technology Kashmir
Email: ajazlone@skuastkashmir.ac.in

Abstract: Green gram (Vigna radiata L.) is an important pulse crop whose productivity in the temperate Kashmir Valley is constrained by irregular rainfall, limited irrigation, cool-season variability, and episodic root-zone moisture deficits. Existing crop-yield models rarely integrate soil-water dynamics, weather, remote sensing, crop physiology, genotype, management, uncertainty, and decision support within a single leakage-aware framework. This study synthesizes the relevant evidence and proposes Kashmir-MungFormer, a physics-guided multimodal architecture for joint moisture-stress classification, phenological-state estimation, and yield prediction. The proposed system combines a temporal encoder for weather and soil-sensor sequences, an image encoder for satellite and unmanned aerial vehicle observations, genotype and management embeddings, reliability-aware cross-modal attention, and a differentiable soil-water-balance constraint. Multi-task learning, deep ensembles, conformal prediction, and explainability methods are specified for future multi-season field validation. An exploratory tabular benchmark was also conducted using 740 Moong (Green Gram) records from 26 Indian states with chronological model-fitting, calibration, and test partitions. Gradient Boosting produced the lowest test MAE (0.1625) and RMSE (0.2055) and the highest R^2 (0.3055), whereas Extra Trees produced the lowest MAPE (28.91%). Split-conformal intervals achieved 89.76% coverage at a nominal 90% level. The Jammu and Kashmir holdout yielded a negative R^2 (−0.1047), indicating substantial regional domain shift. Accordingly, the benchmark is reported as an exploratory baseline rather than a complete validation of the proposed multimodal architecture.
Keywords: soil-water balance; multimodal fusion; temporal attention; uncertainty calibration; precision irrigation; crop phenotyping

1. INTRODUCTION

Mungbean, or green gram is a short duration pulse crop and is not only an important source of dietary protein but also contributes to biological nitrogen fixation, relatively low water demand compared to cereals, diversifying the cropping systems. Even as its productivity is very sensitive to the timing, duration, and intensity of moisture deficit; its expansion into temperate production zones also creates opportunities for both nutritional security and sustainable rotations. Across locations and years, phenology, biomass accumulation and yield formation are highly sensitive to both measured environmental variability [1], [2] and the sequential occurrence of combined water and temperature stress that can appreciably impair pod set and seed filling [3]. Thus, the reproductive window is a vital management temporal scale and irrigation as well as residue studies highlight inter-relationships between soil moisture status, cultivar, irrigation timing and water productivity [4], [5], [6], [7]. Because zonation in root-zone water availability due to elevation gradients, cool nights, variable spring and summer rainfall, short growing seasons and heterogeneous soils creates strong spatial and temporal contrasts in the last-mentioned driver (root-zone water availability), Kashmir represents a unique prediction problem [18]. Climate-series analyses have revealed high temporal and spatial variability in temperature and precipitation patterns of Kashmir Himalaya, whereas vulnerability assessments suggest agriculture in Jammu, Kashmir and Ladakh are increasingly vulnerable to climate variability [8], [9]. One regional average is therefore insufficient; models need to represent district, elevation, soil and climate class, season and management context.

Conventional empirical regressions and mechanistic crop models have strengths but require high intensity calibration to untangle non-linear relationships between the timing of rainfall, potential evaporation, soil moisture and genotype [63]. Deep-learning systems are now routinely integrating weather, soil, remote-sensing, genotype and management data to improve crop-yield forecasting [10], [11], [12], [13], [14], [15], [16], [17], [18], [19],[20],[21]. While most studies have focused on maize, wheat, rice, soybean, cotton or potato, there are hardly any reports available investigating the same parameters in pulse crops grown under temperate environments. Only a handful of frameworks simultaneously integrate physiological diagnosis of water stress, yield prediction, physical consistency with water balance, uncertainty quantification and actionable irrigation advice. Thus, the empirical victory claimed by this paper for its approach is less than, or at least about as much as ideation, and hence defines a research direction rather than assuming superiority. We first compile existing biological and computational evidence on moisture stress in green gram, which is then used to develop Kashmir-MungFormer; a multimodal, physics-guided, uncertainty-aware framework. The framework enables leakage-safe validation, supports multi-season field evaluation, offers robust missing sensors handling, transparent explanations and calibrated decision support. The novelty of SMR is in merging crop-specific stress physiology, spatiotemporal multimodal learning, differentiable soil-water constraints, and uncertainty-aware irrigation recommendations—components that are typically investigated individually in the literature. Figure 1 provides a summary of (a) the biological response pathway and (b) the evidence streams needed for prediction.

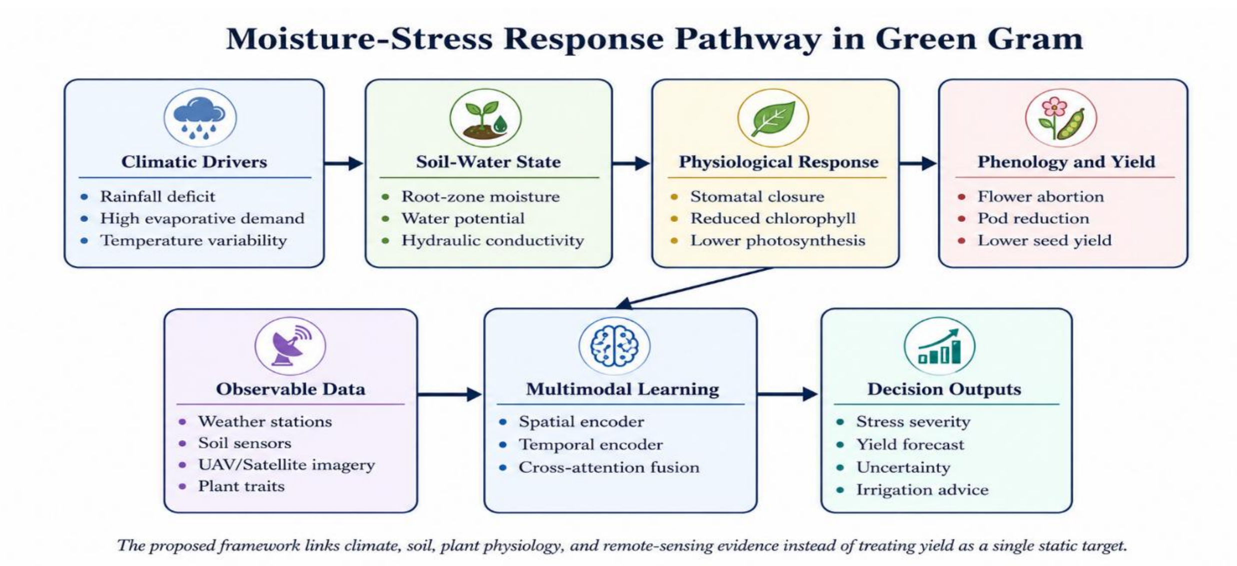

Figure 1. Moisture-stress response pathway and the corresponding multimodal evidence required for prediction.

2. Biological and Agroclimatic Basis of Moisture Stress

Coordinated hydraulic, biochemical and developmental responses of green gram (Vigna radiata L.) to moisture stress. Low root-zone water potential decreases stomatal conductance and carbon assimilation, disturbances in nutrient transport, increases foliar senescence, and diversions of resources between vegetative and reproductive organs. These studies suggest that key stages, such as flowering and pod initiation [2], [3] are sensitive to stress, leading to reduced floret-to-pod conversion and less seed filling. Data driven screening is possible for traits such as genotypic variation in osmotic adjustment, canopy cooling, root architecture, and recovery after dewatering [3], [5], [6], [7] provided these measures are interpreted with consideration of phonological stage.

Temperature modifies the moisture-stress response. Hot temperatures increase the evaporative demand during day and night, and can even exacerbate reproductive losses [2], [6], while cold temperatures can delay germination and hinder early-stage physiological development. The interaction is especially salient in Kashmir, where these cool early-season conditions may be followed by warm and dry spells. Thus, a model that was trained using only total seasonal rainfall might not correctly identify when stress occurs. Stage-aware features should include growing-degree days, days after emergence from crop establishment, flowering on set, pod initiation and canopy temp. Soil physical properties control how atmospheric elicitation influences soil water usability by vegetation. Texture, organic matter, bulk density, rooting-zone depth, field capacity and drainage control infiltration, storage and depletion. Soil-sensor data can be measured in a high frequency but are not immune to error that accompanied their calibration drift and installation depth together with missing measurements. Compared with in situ measurements, satellite and unmanned aerial vehicle (UAV) observations have a larger spatial extent but measure vegetation and surface properties indirectly. Instead of assuming reliability among all modalities, this framework uses these sources in a complementary manner and dynamically estimates the reliability of each modality. Management practices change the stress pathway as well by how they affect canopy development and water demand: sowing date, plant density, residue retention, nutrient application or irrigation interval, genotype [4], [22], [23]. Thus, the model should distinguish environmental stress from management-induced variation by considering management variables as first-class inputs and analyzing irrigation and agronomic treatments that have been varied purposely. The input blocks we propose are summarized in Table 1 along with their respective variables, temporal resolution, and model functionality.

Table 1. Proposed multimodal data specification for temperate green gram prediction.

Data block	**Representative variables**	**Resolution**	**Model role**
Weather	Rainfall; minimum/maximum temperature; relative humidity; solar radiation; wind; reference evapotranspiration	Daily or hourly	Atmospheric demand and stress timing
Soil	Volumetric water content at multiple depths; temperature; texture; EC; pH; field capacity; wilting point	15–60 min plus static profile	Root-zone water availability and retention
Remote sensing	Sentinel-1 backscatter; Sentinel-2 reflectance and indices; UAV RGB, multispectral and thermal imagery	5–10-day satellite; stage-specific UAV	Canopy structure, temperature, greenness and spatial heterogeneity
Crop physiology	Emergence; flowering; pod initiation; LAI; chlorophyll; canopy temperature; biomass	Growth-stage observations	Mechanistic response and intermediate supervision
Management	Genotype; sowing date; irrigation amount and timing; fertilizer; residue; plant density	Event-based	Separates controllable decisions from environment
Context	District; altitude; slope; soil class; season; device; sensor quality	Static and dynamic	Domain shift, calibration and reliability

3. Role of Deep Learning in Agricultural Prediction Systems

Deep learning is most suited for agricultural outcomes that respond nonlinearly to the combined effects of a diverse set of predictors, long temporal histories, spatial heterogeneity and high-dimensional sensor data. CNNs and vision Transformers learn spatial patterns from satellite and UAV imagery; recurrent networks and temporal Transformers capture weather, soil-moisture, and phenological sequences; and multimodal fusion fuses heterogeneous evidence. Deep models have been successful on several crops including soybean [12, 23], wheat [13, 24], maize [14, 15], safflower [16] and sunflower [17–19] when appropriate features and leakage-aware data splits were supported. Because no single sensor observes the entire causal pathway, multimodal data are particularly relevant in predicting moisture-stress. Soil sensors represent local water status structure; weather stations describe atmospheric forcing; radar and optical imagery measure field-sites conditions; thermal imagery indicates canopy cooling and management records represent interventions. Maimaitijiang et al. [17] intermediate-level fusion of UAV RGB, multispectral, and thermal features for soybean-yield estimation. Later studies came to use both satellite and non-satellite environmental, genotype and management variables to increase precision and generalizability of yield prediction [21], [24], [25], [31], [32], [33], [34].

Why interpretability is crucial : Predictive power alone does not guarantee agronomic validity. Specific sensitive growth windows can be detected using attention-based temporal models, and the quantification or visualization of feature contributions can be estimated through SHAP, LIME, integrated gradients, Grad-CAM and related methods [15], [16], [27], [35], [36], [37], [38], [39],[40]. In order for explanations to be useful they need to be verifiable, rather than accept an explanation simply because it seems plausible, the authors advocate that the explanation must pass a series of tests such as stability and faithfulness. This study thus evaluates explanations via feature deletion, temporal perturbation, and seasonal & model-seed consistency. One other route to reliability is physics-guided learning. Due to this failure of weather data alone to represent plant-available water [20], [41], [42], [43], [44], [45], the prediction can be improved by hybrid systems that include crop-model or hydrological variables. Static process knowledge are not replacing by Kashmir-MungFormer instead it uses a differentiable water-balance residual as regularizer, constraining the latent soil-water trojectory while optimizing a deep networks, to do so. Geospatial fusion and climate-impact modelling studies also corroborate heterogeneous-data integration and validation with data-aware models [47], [48].

4. Literature Review

4.1 Green Gram Response to Moisture and Temperature Stress

Moisture deficit is a major determinant of growth, phenology, and yield in recent green gram studies. Multi-location evaluation has shown interactions among rainfall, humidity, site characteristics, and crop duration [1]. Combined water and thermal stress can be more damaging than either stress alone, particularly during reproductive development [2], [3]. Drought-screening studies have also identified accessions and genotypes with improved phenology, protein yield, or recovery under deficit irrigation [5], [7]. Agronomic management remains equally important: residue retention and well-timed irrigation can improve water productivity, while integrated nutrient management can support canopy development and nodulation under temperate conditions [4], [22], [23]. These findings support a model that predicts both final yield and stage-specific stress severity, including the likely response to candidate irrigation actions. Cold tolerance is also relevant in Kashmir because low early-season temperatures can delay germination and alter subsequent canopy development [6]. A drought-only model may therefore misclassify cold-delayed growth as moisture stress. The proposed architecture includes phenological supervision and temperature-context embeddings to distinguish cold-limited establishment from moisture-limited growth [7].

4.2 Deep Learning for Yield, Soil Moisture, and Stress Prediction

The evolution of crop-yield prediction; from fully connected neural networks, to CNN-RNN hybrids, attention models and multimodal systems. Khaki and Wang [13] applied deep neural networks in modeling the interactive effects of genotype and environment, a feature that sets their methodology apart from methods such as machine learning (ML) models in this article, although at least 1 ML model was based on NN algorithms. Recently, Yang et al. [14] also proposed a CNN-RNN model for time-dependent environmental effect knowledge. Shook et al. We used a time aware attention on genotype and weather sequences and provided interpretable, growth-stage specific importance estimates [15]. Both the CNN-based winter-wheat model [16] and DeepCropNet [19] illustrated the importance of temporal feature extraction and spatial context, respectively.

While plot-level observation allows for prediction, remote sensing allows for repeated field- and regional-scale monitoring. UAV-based multimodal fusion integrates canopy structure, spectral reflectance and temperature; while multi-sensor fusion facilitates repeated observations from a joint array of satellite sources [17], [21], [24], [25]. Unlike soil-moisture retrieval, which connects directly hydrological status with crop response through deep-learning techniques [10], [49], [50]. These can lead to errors in crop predictions, as retrieval uncertainty, cloud contamination and temporal misalignment effects will propagate through the datasets. So fusion should include a knowledge of data quality and uncertain calibration.

Through this literature review, researchers define multiple a combination of multimodal inputs and hybrid architectures with optimization strategies as well as methods based on explain ability [26], [27], [28], [29], [30], [31], [32],[33],[34],[35],[36],[37],38,404244[46]{47}{48}[49]{50}51)[/53]. These studies provide evidence that attention and multi-source fusion will enhance predictive performance but the findings are still limited to dominance in major cereals and oilseeds. The fields of pulse crops, mountainous temperate region, and true out-of-location validation are still poorly represented. This gap justifies a shared view at the local level while providing an objective for transfer learning in data-poor regions. Table 2 summarizes some of the representative studies along with their strengths and remaining limitations.

Table 2. Representative studies and the unresolved gap addressed by the proposed framework.

Study	Crop/target	Inputs	Model	Strength	Limitation
Khaki and Wang [13]	Maize	Genotype + environment	DNN	Feature selection; nonlinear G×E modeling	No crop-physiology or uncertainty layer
Khaki et al. [14]	Corn/soybean	Weather + soil + management	CNN-RNN	Temporal dependence and unseen-environment prediction	Regional crop focus
Shook et al. [15]	Soybean	Genotype + weekly weather	Attention LSTM	Interpretable sensitive periods	No imagery or physical water balance
Maimaitijiang et al. [17]	Soybean	RGB + multispectral + thermal UAV	Multimodal DNN	Intermediate fusion	Single crop/site setting
Shahhosseini et al. [20]	Corn	Weather + APSIM outputs	Crop model + ML	Hydrological variables improve prediction	No end-to-end multimodal learning
Teshome et al. [10]	Soil moisture	Weather/soil observations	ML/DL	Direct moisture prediction	Not linked to crop outcome
Xing et al. [51]	Chinese fir	Image/time-series stress data	CNN-LSTM-attention	Drought/heat status modeling	Non-crop and non-yield setting

Khalifani et al. [26]	Sunflower	Morphological traits	CNN	Input-variable comparison under drought	No temporal remote sensing
Ali et al. [28]	Multiple crops	Physiology/environment	ML/DL	High stress-identification performance	Limited external validation
Sajid et al. [31]	Field crops	G×E×M multimodal data	Multimodal neural model	Interaction-aware prediction	Not specific to moisture-stressed pulses
Yadav et al. [32]	Potato	Time-series UAS multispectral	Conv1D-BiGRU-BiLSTM	Context-aware temporal modeling	Different crop and stress target
Demissie et al. [33]	Multiple crops	AI + remote sensing	Review/synthesis	Emerging technology roadmap	No region-specific implementation

5. Structured Evidence-Synthesis Method

Reproducible performance was ensured through a structured evidence-synthesis procedure. Search terms were "green gram," "mung bean," "soil moisture," "drought stress," "crop-yield forecasting," "deep learning" and several more like, the multimodal fusion of data sources are combined with explainable AI algorithm models to be applied to temperate agriculture. Eligible studies were peer-reviewed articles or conference papers with methodological and/or data detail that reported a target for either crop, soil, or stress; input data and model description; and outcome (validation design and quantitative metrics). Transferable data on important crops was incorporated when their architecture or evaluation approach illustrated the green gram framework proposed. For each study, it was coded by crop, region of application, sensing modality employed (if any), temporal resolution, model family used and the target measured, unit that data was split on in least a subset of cases (e.g., pixels or sites), chosen baseline for comparison to their approach and/or datasets evaluated against certain machine learning baselines (i.e., methods based on prediction only without trait experts knowledge added as an additional form way to learn about predictive accuracy), evaluation metric applied which is primarily concerned with signal-to-noise ratio expressed via understanding between predicted outputs relative how well they classify records using empirical judgement established by authors themselves so there may be discrepancies among evaluations carried out here. Some subsets of variables: quality assessment, leak control, independence at the level of the field or subject at risk, adequate baseline characterization and cross-sectional repeated evaluation over time along with preprocessing transparency. We did not compute a pooled performance estimate as the reported outcomes relate to different pairs of crops, scales, units, target definitions and data splits. Aggregating incompatible R^2 or RMSE values, therefore, is methodologically inappropriate: thematic synthesis provides a more appropriate selection of the literature.

The translation from evidence to model design is explicit. The physical-consistency layer is motivated by findings that soil-moisture and crop-model variables improve yield prediction [20], [21]. The stage-aware temporal encoder is supported by studies showing that temporal attention and sequence models can identify sensitive growth periods [15], [16], [54], [55], [56], [57], [58]. Intermediate fusion is motivated by multimodal remote-sensing evidence [17], [21], [24], [25]. Deep ensembles, conformal prediction, and multi-season holdouts are included because uncertainty and external validity remain weakly reported in agricultural prediction studies [59], [60], [61]. Figure 2 summarizes the structured workflow used to translate these evidence gaps into design requirements.

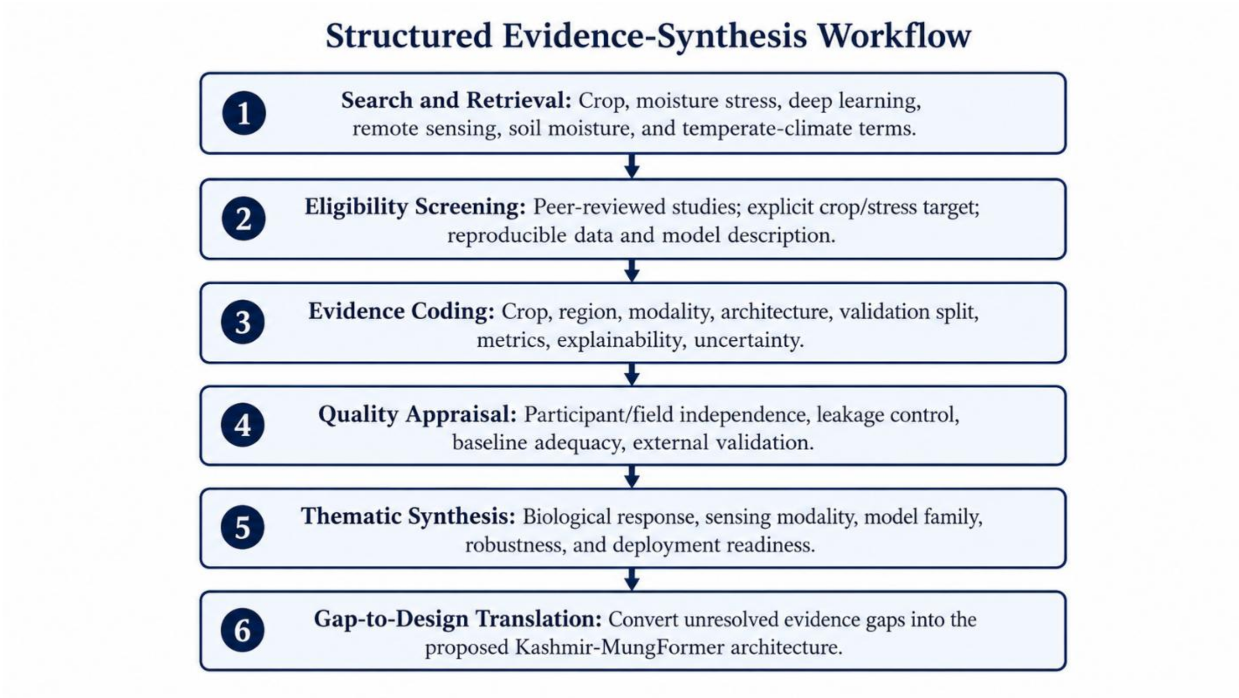

Figure 2. Structured evidence-synthesis workflow used to translate literature gaps into model-design requirements.

6. Evidence Synthesis and Research Gap

6.1 Dominant Evidence Patterns

Five dominant patterns emerge from the literature. First, temporal information is indispensable: seasonal averages conceal the timing of stress, whereas CNN-RNN, LSTM-attention, Transformer, and related sequence models can represent growth-stage-specific effects [14], [15], [16], [19], [51], [54], [55], [56], [57], [58]. Second, multimodal fusion is generally more informative than isolated imagery or weather streams when modalities are temporally and spatially aligned [17], [21], [24], [25], [34]. Third, soil-water variables often explain yield variation that meteorological inputs alone cannot capture [10], [20], [25], [49]. Fourth, explainability is increasingly reported but is rarely validated for stability or agronomic faithfulness [27], [35], [36], [37], [38], [39], [40]. Fifth, external validation and calibrated uncertainty remain uncommon, especially in small regional datasets.

Architecture complexity is not automatically beneficial. Larger models can overfit site identity, sensor characteristics, field boundaries, or season-specific artifacts. Crop studies often create many image patches or time windows from a small number of fields; random splitting can then leak field information across training and test sets. The proposed protocol therefore treats the field-season combination as the primary statistical unit and requires external district and season holdouts.

6.2 Specific Research Gap

No framework identified in the reviewed evidence jointly addresses green gram moisture-stress classification, final-yield regression, phenological supervision, physical water-balance consistency, multimodal reliability, calibrated uncertainty, and irrigation decision support for temperate Kashmir. Green gram and Kashmir-focused studies provide strong biological, agronomic, and regional evidence but limited end-to-end deep-learning implementation [1], [2], [3], [4], [5], [6], [7], [8], [9], [22], [23]. In contrast, advanced deep-learning architectures have largely been developed for other crops, targets, and regions [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53].

The central gap is therefore not merely the absence of a new neural architecture. It is the absence of a validation-ready system that links biological mechanism, regional context, multimodal sensing, physical consistency, uncertainty, and actionability. Kashmir-MungFormer is designed to fill this gap without claiming operational or agronomic readiness before prospective field evaluation. Table 3 maps the identified gaps to the corresponding novel design elements.

Table 3. Research gaps and corresponding novel design elements.

Research gap	**Risk**	**Proposed response**
Static seasonal prediction	Stress timing is hidden in seasonal averages	Stage-aware temporal Transformer with phenology tokens
Single-modality dependence	Cloud, sensor failure, or local measurement error	Reliability-aware multimodal fusion and modality dropout
Purely data-driven latent states	Predictions can violate plausible soil-water dynamics	Differentiable water-balance residual
Yield-only objective	Weak biological supervision and late feedback	Joint stress, phenology, and yield heads
Point prediction	Overconfident recommendations under domain shift	Deep ensembles and conformal intervals
Post hoc explanation only	Attributions may be unstable or unfaithful	SHAP + integrated gradients + perturbation audits
Random train/test split	Field and season leakage inflates accuracy	Field-, district-, and season-level holdouts
Prediction without action	Limited farmer and policy utility	Uncertainty-aware irrigation recommendation layer

7. Proposed Novel Framework: Kashmir-MungFormer

7.1 Multimodal Input Representation

Let $W \in R^{(T\times d_w)}$ denote a weather-soil sequence, $I \in R^{(T_i\times H\times W\times B)}$ a time-indexed remote-sensing tensor, $P \in R^{(T_p\times d_p)}$ crop-physiology observations, $M \in R^{(d_m)}$ management and genotype variables, and $C \in R^{(d_c)}$ static context. The model predicts stress severity s, phenological state g, yield $\hat{y}$, and calibrated uncertainty u.

$$z_w = E_{w(W)},\ \ z_i = E_{i(I)},\ \ z_p = E_{p(P)},\ \ z_m = E_{m(M)},\ \ z_c = E_{c(C)} \tag{1}$$

Equation (1) defines the modality-specific encoded representations. The weather-soil encoder E_w uses a temporal Transformer or selective state-space model with explicit missingness masks. The imagery encoder E_i uses a CNN or vision Transformer followed by phenology-aware temporal pooling. Crop-physiology, management, genotype, and contextual inputs are projected through separate encoders so that heterogeneous variables are not forced into a single undifferentiated feature vector [62], [63], [64].

7.2 Reliability-Aware Cross-Modal Fusion

For modality k, the model learns a predictive representation z_k and a reliability score r_k derived from sensor-quality indicators, cloud masks, missingness, interpolation distance, and out-of-distribution scores. Cross-attention exchanges information among modalities, while reliability gates suppress corrupted streams.

$$\alpha_k = \frac{exp\, exp\, (a_k + log\, log\, (r_k + \varepsilon))}{\Sigma_j\, exp\, exp\, (a_j + log\, log\, (r_j + \varepsilon))}, \quad h = \Sigma_k \alpha_k z_k \tag{2}$$

Equation (2) combines modality relevance and reliability into normalized fusion weights. The separation between relevance a_k and reliability r_k prevents a historically predictive but currently corrupted modality from dominating the fused representation. Modality dropout during training supports graceful degradation when imagery, soil sensors, or physiological measurements are unavailable.

7.3 Physics-Guided Soil-Water Consistency

A simplified root-zone water-balance residual is incorporated as a soft constraint. At time t, the expected change in root-zone storage is governed by rainfall, irrigation, evapotranspiration, runoff, and drainage. The neural model predicts latent storage Ŝ_t, and physically inconsistent trajectories are penalized.

$$R_t = \hat{S}_{t+1} - \hat{S}_t - (P_t + I_t - ET_t - Q_t - D_t) \tag{3}$$

$$L_{phys} = \left(\frac{1}{T}\right) \Sigma_t \rho(R_t) \tag{4}$$

Equations (3) and (4) define the water-balance residual and the corresponding physics-consistency loss, respectively. Here, ρ is a robust Huber penalty. This layer is not presented as a complete hydrological simulator; rather, it regularizes the latent trajectory using known conservation structure. Field capacity, wilting point, rooting depth, and drainage parameters can be estimated by soil class and updated through calibration. Physics-guided learning is supported by evidence that hydrological and crop-model variables improve machine-learning yield prediction [20], [41], [42], [43], [44], [45], [46].

7.4 Multi-Task Learning Objective

The model jointly optimizes ordinal or categorical stress severity, continuous yield, phenological stage, physical consistency, calibration, and representation regularization.

$$L_{total} = \lambda_s L_{stress} + \lambda_y L_{yield} + \lambda_g L_{stage} + \lambda_p L_{phys} + \lambda_c L_{cal} + \lambda_r L_{reg} \tag{5}$$

The complete multi-task objective is given in Equation (5). Stress can be represented as none, mild, moderate, or severe using ordinal cross-entropy; yield uses Huber or Gaussian negative log-likelihood; phenology uses cross-entropy or a time-to-event loss. The coefficients λ_s,λ_y,λ_g,λ_p,λ_c,and λ_r control the relative contributions of stress, yield, phenology, physical consistency, calibration, and regularization. Multi-task learning provides intermediate supervision and reduces the risk that the network learns a yield shortcut unrelated to moisture response.

7.5 Uncertainty and Selective Decision Support

Deep ensembles estimate epistemic uncertainty, while the predictive head can learn heteroscedastic aleatoric variance. Conformal calibration on a held-out set provides distribution-free prediction intervals under exchangeability assumptions [59], [60], [61]. Equation (6) defines the ensemble mean, epistemic variance, and conformal interval, where K is the ensemble size and q(1−α) is the finite-sample calibration-residual quantile.

$$\begin{gathered} \hat{y}(x) = \frac{1}{K}\sum_{k=1}^{K} \hat{y}_k(x). \\ U_{epi}(x) = Var_k[\hat{y}_k(x)]. \\ PI_{1-\alpha}(x) = [\hat{y}(x) - q_{1-\alpha},\ \hat{y}(x) + q_{1-\alpha}]. \end{gathered} \tag{6}$$

The system abstains from irrigation recommendations when uncertainty, missingness, or domain-shift indicators exceed prespecified limits. This converts uncertainty from a descriptive statistic into a safety mechanism.

7.6 Explainability and Agronomic Audit

Instead of using a single attribution method, an explain ability suite is utilized. SHAP provides both global and local explanations for tabular variables [36], LIME yields local surrogate explanations [37], Grad-CAM highlights relevant spatial parts of images [38] and integrated gradients quantify contributions along path-ways of differentiable input dimensions [39]. The higher-level description fits well with existing principles of explainable-AI research [35], [40]. Attention analysis is used to analyze sensitive time frames. Deletion and insertion tests, seed stability, and agronomic review are all used to measure explication quality. The findings need not be causal but should identify plausible drivers (e.g., root-zone moisture, evaporative demand, canopy temperature, flowering-stage stress, and rainfall timing) to which the model associates.

7.7 Irrigation Recommendation Layer

A candidate irrigation action a is evaluated by simulating the short-horizon change in predicted stress and yield under the calibrated model. Equation (7) selects the feasible action that maximizes expected yield benefit after penalties for water use and predictive risk; β and γ are non-negative decision weights, and A is the feasible action set.

$$a^* = arg\, max_{a \in A} \{E[\Delta Y \mid a] - \beta\, Water(a) - \gamma\, Risk(a)\}. \tag{7}$$

The recommendation layer is constrained by agronomic rules, irrigation capacity, and phenological stage. It should be deployed initially in silent mode, where recommendations are compared with expert decisions without directly controlling irrigation. The complete proposed architecture is illustrated in Figure 3.

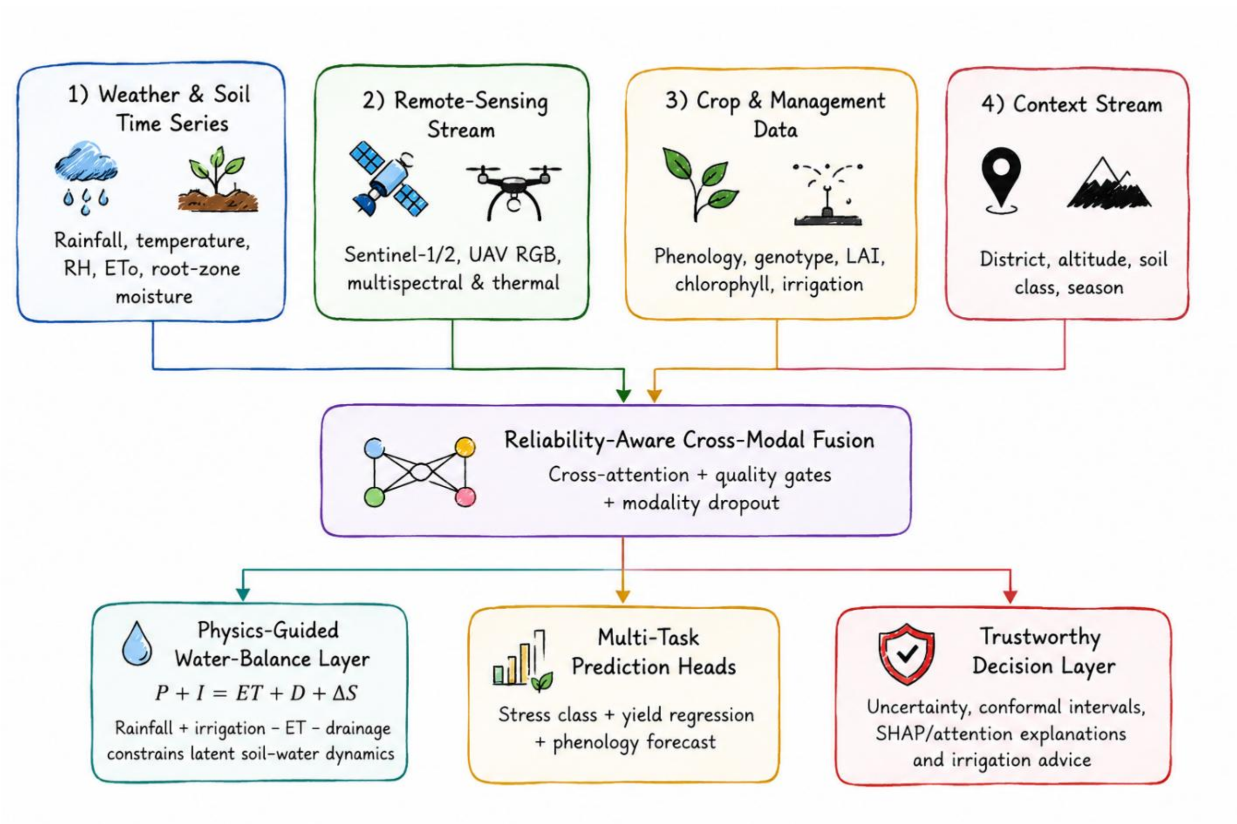

Figure 3. Proposed Kashmir-MungFormer architecture integrating temporal, spatial, physiological, management, and contextual evidence.

8. Proposed Experimental Design and Validation

8.1 Field Sampling Strategy

The field trial should be implemented in a range of districts or agro ecological zones within the Kashmir Valley, for a minimum of two growing seasons, across green gram genotypes and with controlled irrigation regimes. Herbaceous or woody perennial species should be exposed to well-watered control, mild-deficit, moderate-deficit and severe or stage-specific-deficit under a factorial or partially balanced design. Soil sensors should be placed at multiple depths and along with weather, management, phenology, chlorophyll, biomass and final yield measurements made simultaneously. The harmonization of field boundaries with satellite observations and cloud masks. Detect emergence, vegetative growth, flowering, pod initiation and seed filling through UAV campaigns. The parameters for all preprocessing steps (radiometric correction, resampling, image mosaicking, sensor calibration, missing-value imputation and temporal alignment) must be derived without reference to test fields.

8.2 Leakage-Safe Validation

The main split unit is the combination of field and season. It is misleading to just randomly split pixels, image patches or sensor observations as the location and crop identity leak across partitions of your training and test. Hyperparameter selection must use nested cross-validation. The test model should be validated on (i) unseen fields in seen districts, (ii) an unseen district, and (iii) an unseen season. This hierarchy distinguishes between interpolation, spatial transfer and temporal transfer. An overview of the end-to-end validation process is shown in Figure 4.

8.3 Baselines and Ablations

Our baselines will be a linear or mixed-effects regression, Random Forest, XGBoost, LSTM, temporal CNN transformer and also CNN-LSTM (e.g. We should have reporting of parameter numbers and tuning budgets. Deep-learning baselines need to be trained with the Adam optimizer [65] in an established framework, such as Tensorflow [66] or any other reproducible platform. Ablation studies that remove for each modality, the physics constraint, reliability gating, phenolological supervision, uncertainty calibration and context embeddings. The minimum experimental design is summarized in Table 4, the evaluation metrics are given in Table 5 and the necessary ablations are listed in Table 6.

Table 4. Minimum experimental design for credible validation.

Component	Recommended design	Purpose
Sites and seasons	≥3 agroecological zones; ≥2 seasons	Tests spatial and temporal generalization
Genotypes	Locally relevant and contrasting drought responses	Measures genotype transfer and interaction
Moisture treatments	Control, mild, moderate, severe or stage-specific deficit	Creates supervised stress levels
Primary split	Complete field-season groups	Prevents patch/window leakage
External tests	Unseen district and unseen season	Tests transportability
Robustness	Sensor dropout, image clouding, temporal shifts, noise	Tests deployment failure modes
Decision phase	Silent recommendation comparison before field control	Protects against unsafe automation

Table 5. Evaluation metrics for prediction, calibration, robustness, and decision utility.

Evaluation block	**Metrics**	**Interpretation**
Stress classification	Macro-F1, balanced accuracy, class-wise sensitivity/specificity, AUPRC	Performance under imbalance
Yield regression	MAE, RMSE, R^2, concordance correlation, bias by stress level	Accuracy and systematic error
Calibration	Brier score, expected calibration error, reliability curves	Probability trustworthiness
Uncertainty	Prediction-interval coverage, width, risk-coverage curve	Selective prediction
Robustness	Relative performance under missing modalities and corruption	Operational resilience
Explainability	Deletion/insertion, stability, expert agreement	Faithfulness and agronomic utility
Decision utility	Water saved, severe-stress events avoided, yield-risk trade-off	Practical impact

Table 6. Required ablation studies.

Ablation	**Scientific question**
Without imagery	Quantifies dependence on remote sensing
Without soil sensors	Measures satellite/weather-only viability
Without physiology	Tests value of intermediate crop observations
Without context token	Measures district/season information
Without reliability gates	Tests corrupted-modality handling
Without physics loss	Tests physical consistency contribution
Without phenology head	Tests biological supervision
Point prediction only	Quantifies benefit of uncertainty and conformal intervals

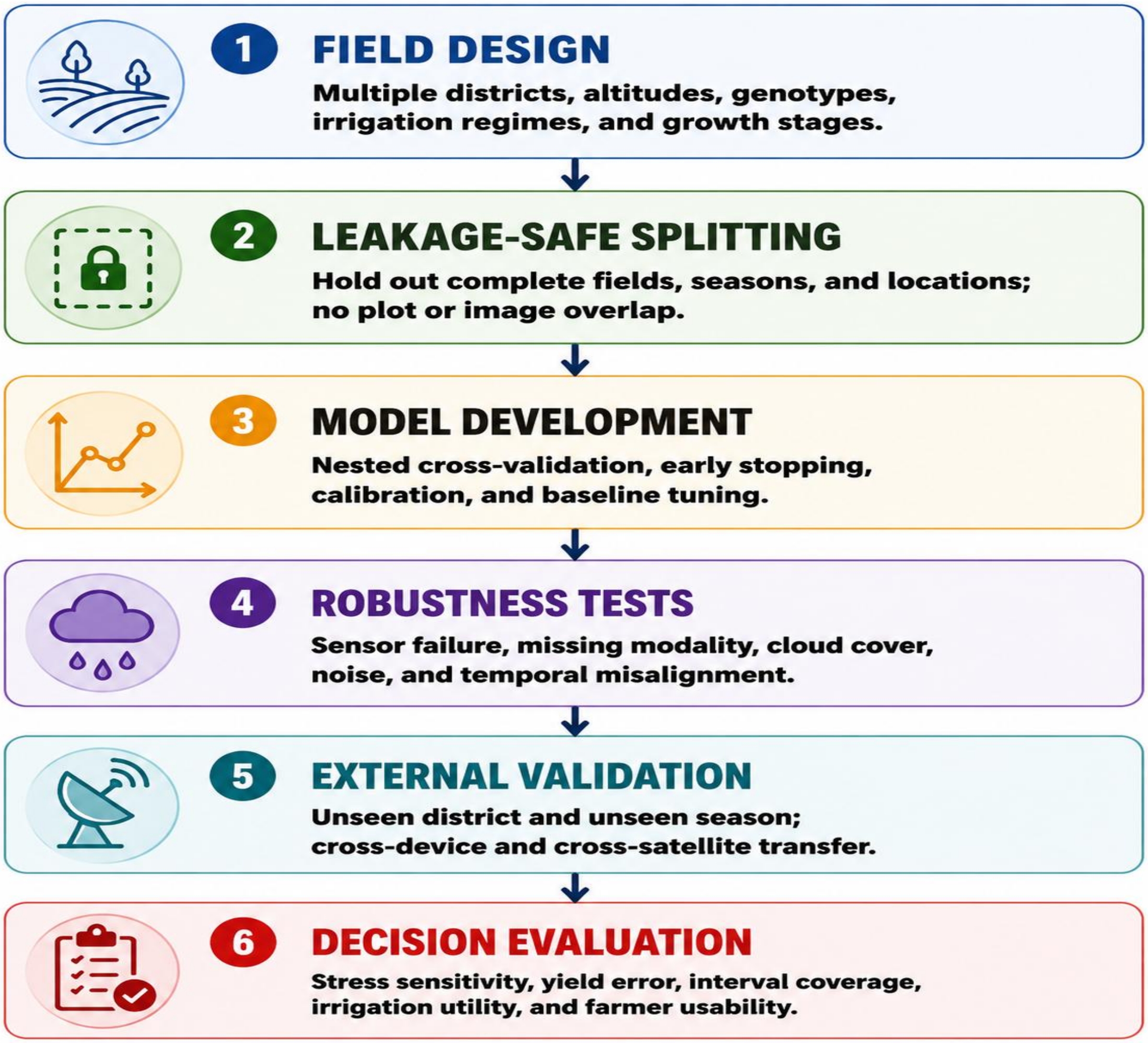

Figure 4. Multi-season and multi-location validation protocol for journal-quality empirical assessment.

8.4 Empirical Dataset and Tabular Benchmark Methodology

For a preliminary quantitative baseline, we obtained the Crop Yield in Indian States data set through KaggleHub [67]. It consists of 19,689 records and has 10 variables which includes crop, crop year, season, state area production annual rainfall fertilizer use pesticide use yield. The analysis contained 740 Moong (Green Gram) observations across six seasons, 26 states and crop years: 1997–2020. Production was excluded from the predictor set due to its direct influence on yield and the potential for direct target leakage. The streams that were not implemented for this benchmark include the satellite, UAV (drone), soil sensor, genotype and crop physiology components shown in Fig 3; only the weather, management, temporal and regional-context components of Kashmir-MungFormer are evaluated here.

8.4.1 Data Preparation and Feature Construction

Standardized column names and categorical labels, removed duplicate records in fields where they could occur, converted numerical fields with error-tolerant parsing, excluded observations without a validated yield target. The set of predictors included crop year, area (size), annual rainfall, fertilizer, pesticide, season and state. To reduce skewness, first-order and square root log-transformed area, rainfall, fertilizer and pesticide were included. For validity within a large dataset, upon determining the counts of how many selected fields were seen prior to imputation each time, we calculated a reliability score at a row-level (i.e.) A four-level moisture-stress proxy was derived from annual-rainfall quartiles: severe (≤Q1), moderate (Q1–Q2), mild (Q2–Q3), and low deficit (>Q3). These proxy categories support exploratory stress analysis only and must not be interpreted as measured physiological drought labels. The source dataset's yield field and continuous predictors were retained in their native reported units; therefore, MAE, RMSE, bias, and interval width are expressed in the same native yield units.

8.4.2 Leakage-Safe Temporal Validation

A chronological split was used instead of random sampling. Crop years 1997–2012 formed the model-fitting set (n = 457), 2013–2015 formed the calibration set (n = 117), and 2016–2020 formed the unseen temporal test set (n = 166). Numerical variables were median-imputed and standardized, whereas season and state were most-frequent-imputed and one-hot encoded. All preprocessing operations were fitted only on the model-fitting partition. A second transportability test trained the leading exploratory Gradient Boosting model on all non-Kashmir records and evaluated it on the 21 Jammu and Kashmir observations.

8.4.3 Models, Evaluation and Uncertainty Calibration

Random Forest, Extra Trees, and Gradient Boosting regressors were implemented using scikit-learn [68]. All three pre-specified benchmark models are reported on the 2016–2020 temporal test period. Because the same test period is used to identify the lowest-RMSE benchmark, the ranking is interpreted as exploratory rather than as unbiased confirmatory model selection. Future work should use nested temporal cross-validation or a separate validation period for model and hyperparameter selection. MAE, RMSE, R^2, MAPE, and mean bias are reported as complementary measures. Split-conformal calibration [61] uses absolute residuals from 2013–2015 to construct nominal 90% prediction intervals. Permutation importance quantifies the increase in MAE after repeatedly shuffling each feature, providing a model-agnostic explanation consistent with the paper's explainability principles [35], [36], [37], [38], [39].

$$MAE = \left(\frac{1}{n}\right) \Sigma_i \, |y_i - \hat{y}_i| \quad (8)$$

$$RMSE = \sqrt{\left[\left(\frac{1}{n}\right) \Sigma_i \, (y_i - \hat{y}_i)^2\right]} \quad (9)$$

$$R\hat{}2 = 1 - [\Sigma_i \, (y_i - \hat{y}_i)\hat{}2 / \Sigma_i \, (y_i - \bar{y})\hat{}2\,] \quad (10)$$

$$PI_{0.90}(x) = [\hat{y}(x) - q_{0.90},\ \hat{y}(x) + q_{0.90}]. \quad (11)$$

Equations (8)–(10) define MAE, RMSE, and R^2, respectively, while Equation (11) defines the split-conformal prediction interval. In Equation (11), $q_{0.90}$ is the finite-sample conformal quantile of the absolute calibration residuals. The resulting interval is interpreted through observed coverage and mean interval width rather than through an assumption of normally distributed errors.

8.5 Algorithmic Procedure

Algorithm 1. Leakage-Safe Tabular Kashmir-MungFormer Benchmark

Input: Indian crop-yield dataset (D), green gram records, predictor variables (X), yield target (y), temporal split years, and significance level (α = 0.10).

Step 1: Load the crop-yield dataset through KaggleHub and standardize column names, data types, categorical labels, and missing-value representations.

Step 2: Retain Moong (Green Gram) observations; remove duplicate records, records with missing targets, and invalid yield values (yield < 0).

Step 3: Exclude production from the predictor set to prevent direct leakage into the yield target.

Step 4: Construct log-transformed numerical variables, a row-level data-reliability score, and rainfall-quartile classes as exploratory moisture-stress proxies.

Step 5: Partition complete crop years into model-fitting (1997–2012), calibration (2013–2015), and unseen test (2016–2020) sets.

Step 6: Fit the preprocessing pipeline on the model-fitting partition only: numerical imputation, scaling, categorical imputation, and one-hot encoding.

Step 7: Train Random Forest, Extra Trees, and Gradient Boosting regressors.

Step 8: Report MAE, RMSE, R^2, MAPE, and mean bias for every pre-specified model; identify the lowest-RMSE model as the leading exploratory benchmark.

Step 9: Calibrate absolute residuals on the calibration partition and construct nominal 90% conformal prediction intervals for unseen test observations.

Step 10: Estimate permutation feature importance and conduct external validation by withholding all Jammu and Kashmir observations from model training.

Output: Comparative yield predictions and evaluation metrics, calibrated uncertainty intervals, a feature-importance profile, moisture-stress-proxy sensitivity, and external transfer-test results.

9. Empirical Results and Discussion

9.1 Dataset Profile and Validation Split

The available dataset was comprised of 19,689 observations in total, which included 740 Moong (Green Gram) traits. The data, filtered, covered 26 states across six seasons and 24 crop years. Table 7 provides the sample size and confirms that the final assessment employed unseen years subsequent to November 2023 rather than a random row-wise split. This design is able to decrease temporal leakage and offer a more realistic assessment of generalization in the face of varying climatic, input, and state-level production conditions.

Table 7. Dataset profile and leakage-safe temporal partition.

Element	**Value**	**Role in evaluation**
Original dataset	19,689 records × 10 variables	National crop-yield source data
Green gram subset	740 records	Crop-specific analytical sample
Coverage	26 states; 6 seasons; 1997–2020	Spatial, seasonal, and temporal context
Model-fitting partition	457 records; 1997–2012	Parameter estimation
Calibration partition	117 records; 2013–2015	Conformal residual calibration
Unseen temporal test	166 records; 2016–2020	Final model comparison
External Jammu and Kashmir test	21 records	State-level transfer assessment

9.2 Yield-Prediction Performance

Gradient Boosting reached the lowest temporal-test MAE (0.1625) and RMSE (0.2055) correspondingly, R^2 = 0.3055. In RMSE second place was reached by Extra Trees (0.2106), and third Random Forest (0.2154), meanwhile, the lowest MAPE was also achieved by Extra Trees (28.91 %). Because the RMSE differences were small (Table 8 and Figure 5(a)), this ranking should be viewed as a modest exploratory advantage rather than clear superiority. However, all three models displayed a negative mean bias (i.e., underprediction of the higher-yield observations). Figure 5(b): The predicted yield still reflects the overall positive tendencies as reflected by observed vs. predicted yield from the leading Gradient Boosting model but predictions become compressed towards the center of distribution for observed data. High-yield cases are consistently underestimated and low-to-moderate yields are closer represented. This is consistent with the negative mean bias, moderate R^2, and suggests that calibrated intervals may be more informative than point estimates only.

Table 8. Unseen-year yield-regression performance (2016–2020).

Model	MAE	RMSE	R^2	MAPE (%)	Mean bias
Gradient Boosting	0.1625	0.2055	0.3055	29.33	−0.1047
Extra Trees	0.1635	0.2106	0.2707	28.91	−0.1090
Random Forest	0.1703	0.2154	0.2370	31.26	−0.1070

9.3 Calibrated Uncertainty and External Transfer

Split-Conformal Residual quantile: 0.3361 For the nominal 90% interval on temporal test set, the coverage achieved was 89.76%, which is only 0.24 percentage points off of what we targeted, and has a mean width of 0.6723 native yield units. As seen in Figure 6(b), most observed yields fell between the intervals, and misses clustered near the upper and lower extremes. The close agreement of nominal and empirical coverage supports the calibration procedure, while relatively wide intervals reveal persistent residual variance in a dataset that lacks soil moisture, genotype, phenology and remote-sensing measurements. For the external evaluation for Jammu and Kashmir, MAE: 0.1560, MAPE: 23.11%, RMSE: increase to 0.2644, R^2 = negative −0.1047. Therefore, although mean absolute error stayed moderate, the Kashmir subset was modelled to not do better than a mean-based predictor. This result validates the concept of regional domain shift and supports our previous recommendation for incorporating key elements including region, altitude, soil, phenology and sensor streams in the complete Kashmir-MungFormer pipeline.

Table 9. Calibration and Jammu-and-Kashmir transfer results.

Evaluation	Metric	Result
Temporal uncertainty	Conformal residual quantile	0.3361
Temporal uncertainty	Nominal interval level	90.00%
Temporal uncertainty	Observed interval coverage	89.76%
Temporal uncertainty	Mean interval width	0.6723
External Jammu and Kashmir (n=21)	MAE	0.1560
External Jammu and Kashmir (n=21)	RMSE	0.2644
External Jammu and Kashmir (n=21)	R^2	−0.1047
External Jammu and Kashmir (n=21)	MAPE	23.11%

External Jammu and Kashmir (n=21) MAPE 23.11%

9.4 Explainability and Moisture-Stress Proxy Analysis

Permutation analysis identified state to be the most important variable, with shuffling it resulting in an increase in MAE of 0.07898. What were the largest quarterly increases in factor level (from October 2023) season (0.00325), fertilizer (0.00254), annual rainfall (0.00103) and log area's smallest increase (0.00085). As can be seen, thus, from Table 10 & Figure 6(a), the model is very dependent upon regional context. This dependence may have roots in actual agroclimatic variation, but it might also capture state-level practices or reporting effects that were not measured. The importance profile facilitates substitution of broad administrative proxies with explicit soil, genotype, management and remote-sensing inputs. Thresholding on rainfall quartiles, as indicated in Figure 7, yielded nearly equal numbers of proxy group members: severe-deficit (185 observations), moderate-deficit (186 observations), mild-deficit (184 observations) and low-deficit (185 observations). This equilibrium emerges mechanically from quartile partitioning and is relevant only for exploratory stratification; it does not validate physiological stress severity. Canonical stress labels must also take into account root-zone soil moisture and crop-stage observations, as well as canopy temperature and controlled field-treatment measurements.

Table 10. Leading permutation-importance results for Gradient Boosting.

Feature	**Increase in MAE**	**Standard deviation**
State	0.07898	0.00898
Season	0.00325	0.00156
Fertilizer	0.00254	0.00120
Annual rainfall	0.00103	0.00045
Log area	0.00085	0.00048
Log fertilizer	0.00058	0.00102
Pesticide	0.00054	0.00026

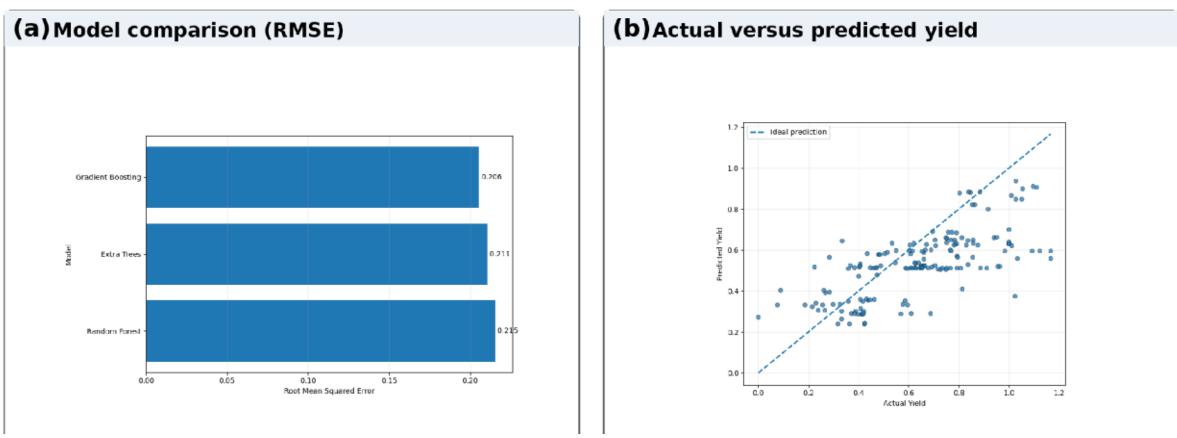

Figure 5. Yield-model results: (a) temporal-test RMSE comparison and (b) actual versus predicted yield for Gradient Boosting.

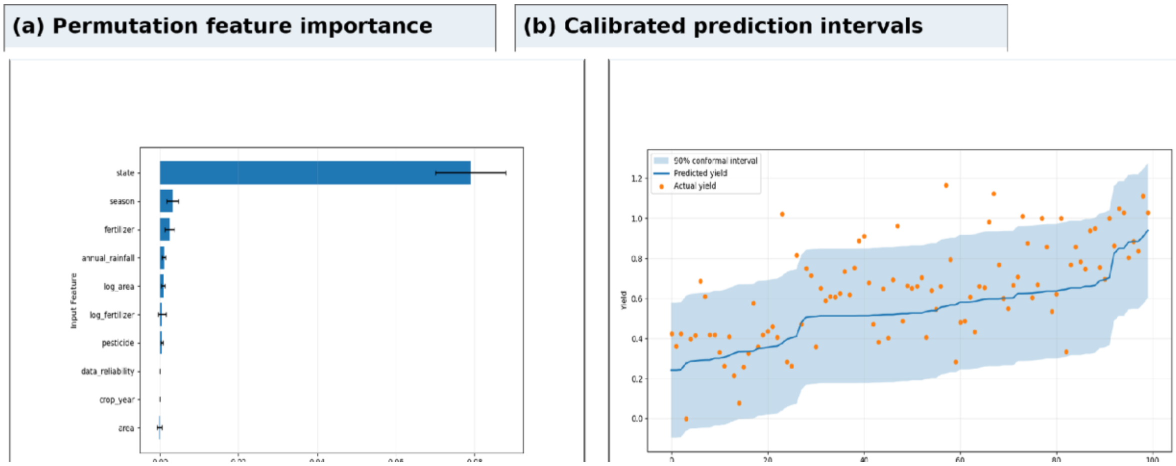

Figure 6. Model interpretation and uncertainty: (a) permutation feature importance and (b) calibrated 90% split-conformal prediction intervals.

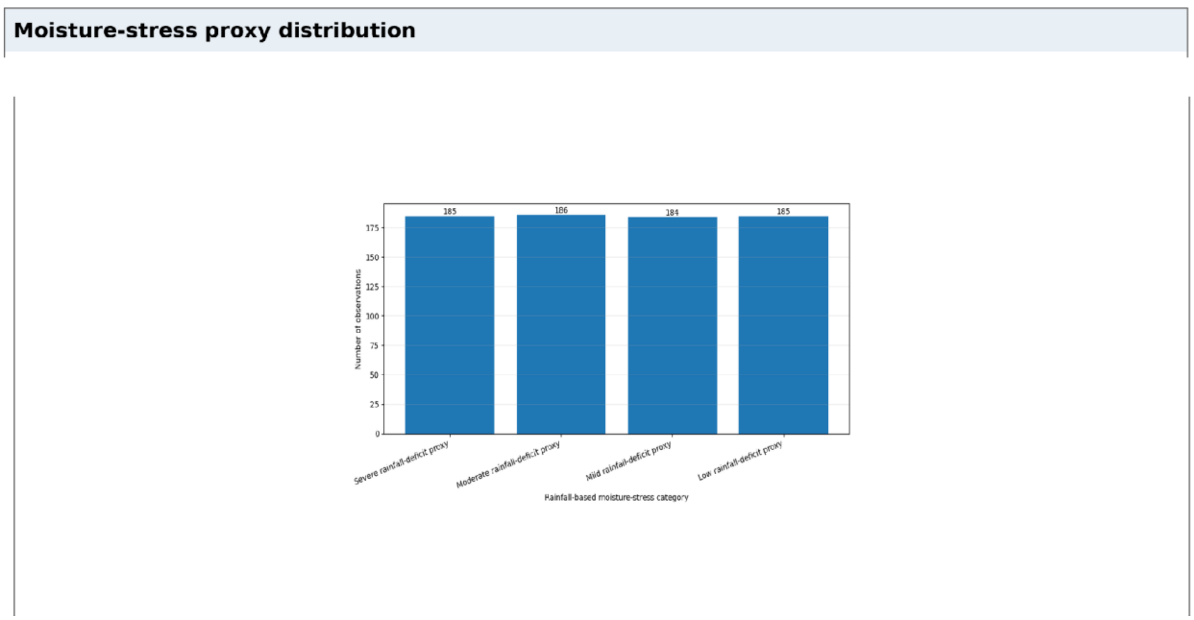

Figure 7. Rainfall-quartile moisture-stress proxy distribution for the 740-record green-gram subset.

9.5 Overall Interpretation of the Benchmark

Collectively, Figures 5–7 and Tables 7–10 establish a reproducible exploratory tabular baseline for the proposed research programme. The benchmark confirms that regional context and seasonal management signals contain useful information and that uncertainty calibration can provide nearly nominal coverage. At the same time, the moderate R^2, underprediction of high-yield cases, wide prediction intervals, exploratory test-set ranking, and negative external Kashmir R^2 show that the available Kaggle variables cannot support the full claims of a physics-guided multimodal model. The results therefore strengthen, rather than replace, the planned multi-season field protocol. The next empirical stage must collect soil-water trajectories, genotype, phenology, canopy temperature, and Sentinel/UAV measurements and evaluate the complete multi-task architecture through the ablations defined in Table 6.

10. DISCUSSION

Kashmir-MungThe Kashmir-MungFormer is deliberately more expansive than your typical yield predictor. Moisture stress is a auto to physiologic process, final yield is an ending out come with influences from weather, soil, genotype and management. Thus, joint stress and phenology supervision should lead to a more biologically interpretable representation compared to training with only the yield loss. The explicit part of the water-balance residual discourages implausible latent trajectories, and reliability gates respond to a class of errors common in agricultural sensing due to ground truth incompleteness or measurement. It also decouples predictive performance from decision readiness. A model with a good R^2 value is still unsafe if it is mis-calibrated or unstable to missing imagery, or overconfident in an out-of-sample district. Thus the evaluation includes calibration, prediction intervals, abstention tests, corruption tests and external holdout. These requirements are in line with calls to environmentally-relevant machine learning that is physically informed, explainable, and uncertainty-aware [35], [36], [37], [38], [39], [40], [41],[59],[60],[61]. And in building on this, ablation studies have to take care to prove that each proposed piece of novelty actually contributes additional value and is not simply an architectural chi mage. A rigorous evaluation should measure whether the physical constraint improves out-of-distribution performance or just regularizes training; compare multimodal fusion with a strong weather-soil baseline; assess explanation stability; and test whether uncertainty grows before large errors. Negative results teach just as much scientifically— if UAV imagery adds minimal incremental value beyond soil and weather data, perhaps a lower-cost sensor configuration is chosen. Discovered that green gram datasets are small in volume and beneficial for transfer learning due to Kashmir specific nature. SQLite 3.9.2 Full Unlocked Note that while encoders are pretrained on larger crop, remote-sensing, and weather datasets fine-tuning needs to avoid domain mismatch. Unlabeled imagery and sensor streams can also be exploited with the use of self-supervised and contrastive methods [69, 70, 71]. Because pretrained representations may retain crop, region, or sensor biases, external validation and subgroup analysis remain necessary. The irrigation recommendation layer should not be interpreted as autonomous control. A responsible deployment pathway is retrospective validation, prospective silent-mode testing, comparison with expert decisions, supervised pilot evaluation, and only then limited decision support. Trade-offs among false alarms, missed severe-stress events, farmer workload, accessibility, water-use efficiency, and yield protection must be evaluated explicitly.

11. Limitations and Risk Controls

This study has several limitations. First of all, field data collection is expensive, even more so when thinking of the independent fields relatively large number of images or time windows. Thus, the unit of statistical independence must therefore stay at the level of field-season combination. Second, irrigation treatments may not accurately reflect plant water status; physiological measurements need to be used in conjunction with soil-water thresholds to improve the quality of stress classifications. Third, in cases where the overpass lies within a cloudy period, remote-sensing observations may not be available; therefore missing-modality performance should be directly assessed. Fourth, the water-balance constraint is a simplified representation of runoff, deep drainage, root growth and capillary rise. Instead of treating its parameters as constant, they should be calibrated and their uncertainty propagated. Fifth, it is possible that recommendations do not transfer across irrigation systems, resource constraints, or soil types. Sixth, the standard benchmark compares a restricted number of pre-specified models against each other on the identical and temporally ordered test period that generates reported ranking; thus exploratory ranking requires confirmation through nested temporal validation or an independent model-selection period. Moreso, the explain ability methods provide associations from the model but not a biological cause, so agronomic interpretation is cautious and should be confirmed by intervention experiments.

12. Future Scope

Future work can extend the framework to longitudinal genotype screening, conduct of federated learning across agricultural research stations, and transfer it to other pulse crops. The interactions between the neighboring fields, and irrigation networks, as well as landscape position may be represented through graph neural network. Earth-observation foundation models and agricultural language models are unlikely to lessen data demand without domain bias audits and computational costs. A potential extension of this project could use the learned model in combination with crop simulation allowing for analysis of hypothetical irrigation schedules, sowing dates and genotype selections. Then reinforcement learning could be applied to optimize water allocation with agronomic safety constraints at a later stage.

The focus for mobile and edge deployment should be low-cost configurations like weather / soil sensing combined with sporadic satellite updates as opposed to continuous UAV access.

13. CONCLUSION

Prediction systems providing information on why, when and where moisture stress occurs in temperate Kashmir is the need for greater productivity particularly with respect to highly susceptible crops like green gram. While growth components based on recent deep-learning and remote-sensing studies possess huge potential, they are so far scattered across crops, modalities and regions. Kashmir-MungFormer, a validation-ready framework integrating spatiotemporal multimodal learning, soil-water consistency constraint, multi-task physiological supervision, calibrated uncertainty and explainability with irrigation decision support. Only architectural novelty is not enough evidence for a journal; the framework also needs multi-season field experiments, strong baselines, leakage-safe splitting, comprehensive ablations, external validation and decision-level evaluation. A first (multi-model) exploratory tabular benchmark in fact appears here: Gradient Boosting developed by Chennupati et al. on unseen temporal test uncovered RMSE = 0.2055, R^2 = 0.3055 at the level of nominal 90% conformal intervals coverage = 89.76%. Negative external Kashmir R^2 and exploratory model ranking illustrate the need for local multimodal field observations, nested temporal validation, and regional recalibration prior to operational use.

REFERENCES

1. A. K. Parihar, S. Gupta, K. K. Hazra, A. Lamichaney, D. Sen Gupta, D. Singh, et al., "Multi-location evaluation of mungbean (Vigna radiata L.) in Indian climates: Ecophenological dynamics, yield relations, and characterization of locations," Frontiers in Plant Science, vol. 13, Art. no. 984912, 2022, doi: 10.3389/fpls.2022.984912.
2. G. Mahajan, K. Wenham, and B. S. Chauhan, "Mungbean (Vigna radiata) growth and yield response in relation to water stress and elevated day/night temperature conditions," Agronomy, vol. 13, no. 10, Art. no. 2546, 2023, doi: 10.3390/agronomy13102546.
3. M. Jincy, V. Prasad, A. Senthil, P. Jeyakumar, and N. Manivannan, "Physiological divergence in green gram genotypes for drought and high-temperature stress tolerance during flowering," Legume Research, vol. 45, no. 8, pp. 960-967, 2022, doi: 10.18805/LR-4314.
4. S. Tripathi, A. Kaur, A. S. Brar, K. S. Sekhon, S. Singh, A. Malik, and O. Kisi, "Optimizing yield and water productivity in summer mung bean through crop-residue management and irrigation strategies," BMC Plant Biology, vol. 24, Art. no. 1012, 2024, doi: 10.1186/s12870-024-05640-1.
5. M. R. Islam, J. Hossain, M. S. Huda, M. G. Azam, M. F. Hossain, M. Aktar-Uz-Zaman, et al., "Assessment of drought tolerance in mung bean through phenology, growth, protein yield, and cluster heatmap analysis," Philippine Agricultural Scientist, vol. 107, no. 4, pp. 366-378, 2024.
6. L. S. Manasa, M. Panigrahy, K. C. Panigrahi, G. Mishra, S. K. Panda, and G. R. Rout, "Cold tolerance mechanisms in mungbean genotypes during germination," Agriculture, vol. 13, no. 2, Art. no. 315, 2023, doi: 10.3390/agriculture13020315.
7. U. N. Ukwu, B. N. Oburu, D. P. Udochukwu, S. O. Adewuyi, S. O. Muojiama, V. O. Osadebe, et al., "Performance of mungbean accessions under intermittent water-deficit stress in a tropical environment," Journal of Applied Biology & Biotechnology, vol. 12, no. 3, pp. 129-137, 2024, doi: 10.7324/JABB.2024.158000.
8. J. M. Dad, M. Muslim, I. Rashid, and Z. A. Reshi, "Time-series analysis of climate variability and trends in the Kashmir Himalaya," Ecological Indicators, vol. 126, Art. no. 107690, 2021, doi: 10.1016/j.ecolind.2021.107690.
9. I. Malik, M. Ahmed, Y. Gulzar, S. H. Baba, M. S. Mir, A. B. Soomro, et al., "Estimation of agricultural vulnerability to climate change using analytical and deep-learning methods: A case study in Jammu, Kashmir, and Ladakh," Sustainability, vol. 15, no. 14, Art. no. 11465, 2023, doi: 10.3390/su151411465.
10. F. T. Teshome, H. K. Bayabil, B. Schaffer, Y. Ampatzidis, and G. Hoogenboom, "Improving soil-moisture prediction with deep-learning and machine-learning models," Computers and Electronics in Agriculture, vol. 226, Art. no. 109414, 2024, doi: 10.1016/j.compag.2024.109414.
11. M. A. Jabed and M. A. A. Murad, "Crop-yield prediction in agriculture: A comprehensive review of machine-learning and deep-learning approaches, with insights for sustainability," Heliyon, vol. 10, no. 24, Art. no. e40836, 2024, doi: 10.1016/j.heliyon.2024.e40836.
12. A. Oikonomidis, C. Catal, and A. Kassahun, "Hybrid deep-learning-based models for crop-yield prediction," Applied Artificial Intelligence, vol. 36, no. 1, pp. 1-18, 2022, doi: 10.1080/08839514.2022.2031823.
13. S. Khaki and L. Wang, "Crop-yield prediction using deep neural networks," Frontiers in Plant Science, vol. 10, Art. no. 621, 2019, doi: 10.3389/fpls.2019.00621.
14. S. Khaki, L. Wang, and S. V. Archontoulis, "A CNN-RNN framework for crop-yield prediction," Frontiers in Plant Science, vol. 10, Art. no. 1750, 2019, doi: 10.3389/fpls.2019.01750.

15. J. Shook, T. Gangopadhyay, L. Wu, B. Ganapathysubramanian, S. Sarkar, and A. K. Singh, "Crop-yield prediction integrating genotype and weather variables using deep learning," PLOS ONE, vol. 16, no. 6, Art. no. e0252402, 2021, doi: 10.1371/journal.pone.0252402.
16. A. K. Srivastava, N. Safaei, S. Khaki, G. Lopez, W. Zeng, F. Ewert, T. Gaiser, and J. Rahimi, "Winter-wheat yield prediction using convolutional neural networks from environmental and phenological data," Scientific Reports, vol. 12, Art. no. 3215, 2022, doi: 10.1038/s41598-022-06249-w.
17. M. Maimaitijiang, V. Sagan, P. Sidike, S. Hartling, F. Esposito, and F. B. Fritschi, "Soybean-yield prediction from UAV using multimodal data fusion and deep learning," Remote Sensing of Environment, vol. 237, Art. no. 111599, 2020, doi: 10.1016/j.rse.2019.111599.
18. P. Nevavuori, N. Narra, and T. Lipping, "Crop-yield prediction with deep convolutional neural networks," Computers and Electronics in Agriculture, vol. 163, Art. no. 104859, 2019, doi: 10.1016/j.compag.2019.104859.
19. T. Lin, R. Zhong, Y. Wang, J. Xu, H. Jiang, J. Xu, et al., "DeepCropNet: A deep spatial-temporal learning framework for county-level corn-yield estimation," Environmental Research Letters, vol. 15, no. 3, Art. no. 034016, 2020, doi: 10.1088/1748-9326/ab66cb.
20. M. Shahhosseini, G. Hu, I. Huber, and S. V. Archontoulis, "Coupling machine learning and crop modeling improves crop-yield prediction in the US Corn Belt," Scientific Reports, vol. 11, Art. no. 1606, 2021, doi: 10.1038/s41598-020-80820-1.
21. L. Li, B. Wang, P. Feng, D. L. Liu, Q. He, Y. Zhang, et al., "Developing machine-learning models with multi-source environmental data to predict wheat yield in China," Computers and Electronics in Agriculture, vol. 194, Art. no. 106790, 2022, doi: 10.1016/j.compag.2022.106790.
22. Q. A. Aga, M. Attri, F. Aga, and K. K. Bansal, "Production performance of mungbean (Vigna radiata) as influenced by nutrient management under temperate conditions of Kashmir Valley," Journal of Crop and Weed, vol. 19, no. 2, pp. 250-255, 2023, doi: 10.22271/09746315.2023.v19.i2.1728.
23. Qurat-ul-Ain-Aga, M. Attri, F. Aga, K. K. Bansal, and V. Kumar, "Response of green gram to comprehensive nutrient management under temperate conditions of Kashmir Valley," Biological Forum - An International Journal, vol. 15, no. 5a, pp. 92-97, 2023.
24. J. Lu, J. Li, H. Fu, X. Tang, Z. Liu, H. Chen, Y. Sun, and X. Ning, "Deep learning for multi-source data-driven crop-yield prediction in Northeast China," Agriculture, vol. 14, no. 6, Art. no. 794, 2024, doi: 10.3390/agriculture14060794.
25. J. Lu, H. Fu, X. Tang, Z. Liu, J. Huang, W. Zou, et al., "GOA-optimized deep learning for soybean-yield estimation using multi-source remote-sensing data," Scientific Reports, vol. 14, Art. no. 7097, 2024, doi: 10.1038/s41598-024-57278-6.
26. S. Khalifani, R. Darvishzadeh, N. Azad, M. G. Shayesteh, H. Kalbkhani, and N. Akbari, "Comparison of influential input variables in deep-learning modeling of sunflower grain yield under normal and drought-stress conditions," Field Crops Research, vol. 303, Art. no. 109145, 2023, doi: 10.1016/j.fcr.2023.109145.
27. I. Malashin, V. Tynchenko, A. Gantimurov, V. Nelyub, A. Borodulin, and Y. Tynchenko, "Predicting sustainable crop yields: Deep-learning and explainable-AI tools," Sustainability, vol. 16, no. 21, Art. no. 9437, 2024, doi: 10.3390/su16219437.
28. T. Ali, S. U. Rehman, S. Ali, K. Mahmood, S. A. Obregon, R. C. Iglesias, et al., "Smart agriculture: Utilizing machine learning and deep learning for drought-stress identification in crops," Scientific Reports, vol. 14, Art. no. 30062, 2024, doi: 10.1038/s41598-024-74127-8.
29. S. Godara, R. Parsad, R. S. Bana, D. Singh, G. Avinash, and S. Marwaha, "DL-RSM: Deep-learning-integrated response-surface methodology for environmental-condition estimation for crop yield," Journal of Cleaner Production, vol. 456, Art. no. 142381, 2024, doi: 10.1016/j.jclepro.2024.142381.
30. R. Zhong, X. Xiong, Q. Tian, J. Huang, L. Zhu, Y. Yang, and T. Lin, "PSNet: A hierarchical deep-learning framework for crop-yield estimation," Computers and Electronics in Agriculture, vol. 239, Art. no. 110917, 2025, doi: 10.1016/j.compag.2025.110917.
31. S. S. Sajid, Z. Khalilzadeh, L. Wang, and G. Hu, "Analyses of crop-yield dynamics and a multimodal neural-network model with genotype-environment-management interactions," Frontiers in Plant Science, vol. 16, Art. no. 1537990, 2025, doi: 10.3389/fpls.2025.1537990.
32. S. A. Yadav, X. Zhang, N. K. Wijewardane, M. Feldman, R. Qin, Y. Huang, et al., "Context-aware deep learning model for yield prediction in potato using time-series UAS multispectral data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 18, pp. 6096-6115, 2025, doi: 10.1109/JSTARS.2025.3539217.
33. W. A. Demissie, L. Sebastiani, and R. Rossetto, "Integration of artificial intelligence and remote sensing for crop-yield prediction and crop-growth parameter estimation in Mediterranean agroecosystems," European Journal of Agronomy, vol. 173, Art. no. 127894, 2026, doi: 10.1016/j.eja.2025.127894.
34. M. U. Shahzad, S. Tahir, J. Rashid, O. A. Khashan, R. Ahmad, S. Mansoor, and A. Ghani, "Machine-learning-based cotton-yield forecasting under climate change for precision agriculture," Smart Agricultural Technology, vol. 12, Art. no. 101117, 2025, doi: 10.1016/j.atech.2025.101117.
35. A. B. Arrieta, N. Diaz-Rodriguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, et al., "Explainable artificial intelligence: Concepts, taxonomies, opportunities and challenges," Information Fusion, vol. 58, pp. 82-115, 2020, doi: 10.1016/j.inffus.2019.12.012.
36. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems 30, pp. 4765-4774, 2017, doi: 10.5555/3295222.3295230.

37. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135-1144, 2016, doi: 10.1145/2939672.2939778.
38. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE International Conference on Computer Vision, pp. 618-626, 2017, doi: 10.1109/ICCV.2017.74.
39. M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in Proc. 34th International Conference on Machine Learning, PMLR, vol. 70, pp. 3319-3328, 2017, doi: 10.48550/arXiv.1703.01365.
40. A. Wolanin, G. Mateo-Garcia, G. Camps-Valls, L. Gomez-Chova, M. Meroni, G. Duveiller, et al., "Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt," Environmental Research Letters, vol. 15, no. 2, Art. no. 024019, 2020, doi: 10.1088/1748-9326/ab68ac.
41. M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, "Deep learning and process understanding for data-driven Earth-system science," Nature, vol. 566, no. 7743, pp. 195-204, 2019, doi: 10.1038/s41586-019-0912-1.
42. A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, et al., "Theory-guided data science: A new paradigm for scientific discovery from data," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 10, pp. 2318-2331, 2017, doi: 10.1109/TKDE.2017.2720168.
43. J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating scientific knowledge with machine learning for engineering and environmental systems," ACM Computing Surveys, vol. 55, no. 4, Art. no. 66, pp. 1-37, 2022, doi: 10.1145/3514228.
44. C. Shen, "A transdisciplinary review of deep-learning research and its relevance for water resources scientists," Water Resources Research, vol. 54, no. 11, pp. 8558-8593, 2018, doi: 10.1029/2018WR022643.
45. F. Kratzert, D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger, "Rainfall-runoff modelling using long short-term memory networks," Hydrology and Earth System Sciences, vol. 22, no. 11, pp. 6005-6022, 2018, doi: 10.5194/hess-22-6005-2018.
46. G. S. Nearing, F. Kratzert, A. K. Sampson, C. S. Pelissier, D. Klotz, J. M. Frame, et al., "What role does hydrological science play in the age of machine learning?" Water Resources Research, vol. 57, no. 3, Art. no. e2020WR028091, 2021, doi: 10.1029/2020WR028091.
47. H. Jiang, H. Hu, R. Zhong, J. Xu, J. Xu, J. Huang, et al., "A deep-learning approach to conflating heterogeneous geospatial data for corn-yield estimation in the US Corn Belt," Global Change Biology, vol. 26, no. 3, pp. 1754-1766, 2020, doi: 10.1111/gcb.14885.
48. A. Crane-Droesch, "Machine learning methods for crop-yield prediction and climate-change impact assessment in agriculture," Environmental Research Letters, vol. 13, no. 11, Art. no. 114003, 2018, doi: 10.1088/1748-9326/aae159.
49. A. Singh and K. Gaurav, "Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images," Scientific Reports, vol. 13, Art. no. 2251, 2023, doi: 10.1038/s41598-023-28939-9.
50. M. Dabboor, G. Atteia, S. Meshoul, and W. Alayed, "Deep-learning-based framework for soil-moisture retrieval of bare soil from satellite data," Remote Sensing, vol. 15, no. 7, Art. no. 1916, 2023, doi: 10.3390/rs15071916.
51. D. Xing, Y. Wang, P. Sun, H. Huang, and E. Lin, "A CNN-LSTM-att hybrid model for classification and evaluation of growth status under drought and heat stress in Chinese fir (Cunninghamia lanceolata)," Plant Methods, vol. 19, Art. no. 66, 2023, doi: 10.1186/s13007-023-01044-8.
52. J. Logeshwaran, D. Srivastava, K. S. Kumar, M. J. Rex, A. Al-Rasheed, M. Getahun, and B. O. Soufiene, "Improving crop production using an agro-deep-learning framework in precision agriculture," BMC Bioinformatics, vol. 25, Art. no. 341, 2024, doi: 10.1186/s12859-024-05970-9.
53. V. S. Yakkala, K. V. Nusimala, B. Gayathri, S. Kanamarlapudi, S. S. Aravinth, A. O. Salau, and S. Srithar, "Deep-learning-based crop-health enhancement through early disease prediction," Cogent Food & Agriculture, vol. 11, no. 1, Art. no. 2423244, 2025, doi: 10.1080/23311932.2024.2423244.
54. B. Lim, S. O. Arik, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for interpretable multi-horizon time-series forecasting," International Journal of Forecasting, vol. 37, no. 4, pp. 1748-1764, 2021, doi: 10.1016/j.ijforecast.2019.12.005.
55. B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, "N-BEATS: Neural basis expansion analysis for interpretable time-series forecasting," in Proc. International Conference on Learning Representations, unpaginated, 2020, doi: 10.48550/arXiv.1905.10437.
56. H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient Transformer for long sequence time-series forecasting," in Proc. AAAI Conference on Artificial Intelligence, vol. 35, no. 12, pp. 11106-11115, 2021, doi: 10.1609/aaai.v35i12.17325.
57. H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition Transformers with autocorrelation for long-term series forecasting," in Advances in Neural Information Processing Systems 34, pp. 22419-22430, 2021, doi: 10.48550/arXiv.2106.13008.
58. Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with Transformers," in Proc. International Conference on Learning Representations, unpaginated, 2023, doi: 10.48550/arXiv.2211.14730.

59. Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in Proc. 33rd International Conference on Machine Learning, PMLR, vol. 48, pp. 1050-1059, 2016, doi: 10.48550/arXiv.1506.02142.
60. B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in Advances in Neural Information Processing Systems 30, pp. 6402-6413, 2017, doi: 10.48550/arXiv.1612.01474.
61. A. N. Angelopoulos and S. Bates, "Conformal prediction: A gentle introduction," Foundations and Trends in Machine Learning, vol. 16, no. 4, pp. 494-591, 2023, doi: 10.1561/2200000101.
62. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., "Attention is all you need," in Advances in Neural Information Processing Systems 30, pp. 5998-6008, 2017, doi: 10.48550/arXiv.1706.03762.
63. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, et al., "An image is worth 16 x 16 words: Transformers for image recognition at scale," in Proc. International Conference on Learning Representations, unpaginated, 2021, doi: 10.48550/arXiv.2010.11929.
64. A. Gu and T. Dao, "Mamba: Linear-time sequence modeling with selective state spaces," arXiv:2312.00752, 2023, doi: 10.48550/arXiv.2312.00752.
65. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. International Conference on Learning Representations, unpaginated, 2015, doi: 10.48550/arXiv.1412.6980.
66. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, et al., "TensorFlow: A system for large-scale machine learning," in Proc. 12th USENIX Symposium on Operating Systems Design and Implementation, pp. 265-283, 2016. [Online]. Available: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.
67. A. Gupta, "Crop Yield in Indian States Dataset," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/datasets/akshatgupta7/crop-yield-in-indian-states-dataset. Accessed: Jun. 24, 2026.
68. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011. [Online]. Available: https://jmlr.org/papers/v12/pedregosa11a.html.
69. A. van den Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," arXiv:1807.03748, 2018, doi: 10.48550/arXiv.1807.03748.
70. J.-B. Grill, F. Strub, F. Altche, C. Tallec, P. H. Richemond, E. Buchatskaya, et al., "Bootstrap your own latent: A new approach to self-supervised learning," in Advances in Neural Information Processing Systems 33, pp. 21271-21284, 2020, doi: 10.48550/arXiv.2006.07733.
71. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in Proc. 37th International Conference on Machine Learning, PMLR, vol. 119, pp. 1597-1607, 2020, doi: 10.48550/arXiv.2002.05709.