Journal of Applied Automation Technologies

Research Paper

Dilated Residual TCN Model for Crop Yield Prediction Using Multi-Source Farmland Sensors

Gabriel Láska[1], Natálie Králová[1] and Kateřina Svobodová[1, *]

[1] Department of Computer Science and Engineering, University of West Bohemia in Pilsen, 301 00 Pilsen, Czech Republic
*Corresponding author: katerina.s@kiv.zcu.cz

Abstract. Create a high-precision agricultural system for field-level resource allocation, irrigation scheduling, and production risk management that predicts crop yield using multi-source farmland sensor data. Because soil, weather, irrigation, and other vegetation characteristics have distinct cycles and delayed effects on crop growth, the majority of current data-driven systems have failed to handle diverse sensing signals. An interpretable dilated residual temporal convolutional network for crop yield prediction is presented in this paper. Create aligned sensor representations, use dilated causal convolution to identify long-range growth relationships, use residual connections to stabilize deep temporal learning, and calculate sensor contributions for agronomic interpretation. Compare the outcomes of attention-based forecasting models, Random Forest, XGBoost, LSTM, GRU, and regular TCN. According to the simulated field-scale findings, the suggested model reduces the prediction error by 14.8% when compared to the standard TCN and by 22.6% when compared to LSTM, with an RMSE of 0.412 t/ha and a R^2 of 0.903. Additionally, robustness research demonstrates that it operates normally with cross-seasonal transfer and a 20% lack of sensor data. Assist agricultural sensing systems with high-precision, high-efficiency, and comprehensible crop yield forecast.

Keywords: *Crop Yield Prediction, Multi-Source Farmland Sensors, Dilated Residual TCN, Temporal Feature Fusion, Interpretable Agriculture*

Received on 07 May 2022, Accepted on 29 December 2022, Published on 23 Jan 2023

Introduction

Precision agriculture today requires crop production prediction since post-harvest statistics are no longer the only way to estimate crop yield. In order to make decisions about irrigation, fertilizer application, harvesting logistics, storage plans, and supply-chain contracts before the crop is completely matured, the majority of contemporary farms now employ yield projections. The type of system has shifted from one based on sparse historical data to a continuous-acquisition system for field-level sensing streams due to the quick development of soil probes, micro-weather stations, canopy sensors, irrigation meters, and other inexpensive edge devices for agricultural prediction [1]. Compared to earlier agricultural observations, the sensing streams exhibit variations in moisture stress, heat accumulation, radiation exposure, rainfall, nutrient availability, and crop canopy response more often [2]. The elements can be used to create a model that examines variations in yield at various locations and times [3]. Early stress accumulation and delayed growth responses that are difficult to see with a single meteorological or remote-sensing signal can be detected by multi-source farmland sensors [4].

Additionally, it has not yet been possible to forecast crop yield using several sensor inputs. The physical meanings, sample intervals, noise levels, and missing-data methods of farmland sensor data vary. Vegetation indices will change more slowly during the growth cycle, soil moisture will drastically decrease following irrigation, and canopy temperature will react to heat stress within a few hours [5]. Conventional machine learning models typically require fixed aggregation windows and hence lose some of the sequential information, even though they can handle designed features [6]. Although recurrent neural networks improve time series data modeling, they have limited parallelism, slow training, and gradient vanishing when dealing with long-sequence growing-season data [7]. For field-scale Internet of Things agricultural systems, transformer-based models have a high computational cost and enormous data requirements, despite their flexible attention [8]. For the reasons, a model must be created that can manage long-term dependencies in time, preserve local sensor variations, and be computationally practical for repeated forecasts in real agricultural monitoring platforms [9].

An explainable dilated residual temporal convolutional network for crop production prediction using multi-source farmland sensors is presented in this paper. Utilize residual temporal refinement for stable information transmission, widen the receptive field using dilated causal convolution, and integrate all sensor signals into a single time series in the model. Instead of being viewed as a "black box" regression result, the model's output can be connected to agricultural science by adding a sensor contribution method to identify the factors that significantly affect the yield at various growth periods. This paper's primary objectives are to increase yield prediction accuracy and computation efficiency, strengthen robustness against faulty sensing, and provide interpretability for agricultural decision-making assistance.

Related Work

Data-Driven Crop Yield Prediction

Machine learning methods are increasingly widely employed in crop yield prediction, which has shifted from empirical regression and crop-growth simulation to data-driven models supported by remote sensing, field sensors, meteorological data, etc. For instance, the relationships between temperature, precipitation, soil, and management parameters were not linear, according to early machine learning research [10]. By using a lot of historical data to identify latent interactions between genotype, environment, and management variables, deep neural networks significantly enhanced the prediction [11]. Convolutional models can be used to extract local patterns from multi-temporal images and field data, and they have also been used to process spatial or spatiotemporal agricultural information [12]. The geographical resolution of the large-scale forecasting framework for regional crop monitoring may be insufficient, despite the fact that it incorporates satellite indicators and weather data and supports decisions at the plot level [13]. In order to leverage both process expertise and data-driven flexibility, some hybrid models have recently integrated crop models with machine learning [14]. Since many of the current studies are still seasonal summaries or manually aggregated indices, they are unable to identify intra-season variations in sensor behavior [15].

By providing spatial information, remote sensing and UAV-based techniques have expanded the coverage area of agricultural yield estimation to canopy level; yet, they are frequently constrained by problems such cloud cover, flight frequency, and spectral saturation in dense vegetation [16]. By continuously gathering environmental and soil data close to the crop's root zone, field sensors can supplement this data collection. They can be used to monitor water stress, heat stress, and other short-term management reactions [17]. The multi-source sensor data are not inherently time-synchronized, which is the issue. Simple feature concatenation might result in misleading correlations due to sampling gaps, calibration drift, and abrupt environmental changes [18]. As a result, temporal fusion architectures that can manage heterogeneous sampling, uphold time-order requirements, and identify growth-stage correlations without exclusively depending on manually created lag variables are becoming more and more necessary for yield prediction [19].

Temporal Deep Learning in Smart Agriculture

Temporal deep learning models are appropriate for agricultural sensing since crop growth in agriculture typically occurs in stages and involves changes over time. Long-growing-season modeling is more difficult due to the sequential computation of recurrent neural networks, LSTM, and GRU models, despite their application to agricultural time series [20]. An option is provided by Temporal Convolutional Networks, which increase the temporal receptive field without sacrificing parallel computation by using causal convolution and dilation [21]. In contrast to recurrent models, convolution does not employ hidden-state recurrence and can handle both short-term fluctuations and long-distance relationships in time by stacking filters. By using pertinent historical data, attention-based forecasting models can increase the prediction range for longer periods of time, but they are more complicated and might not be feasible for numerous sensor channels and lengthy sequences [22].

Because it balances training stability, parameter efficiency, and receptive field length, dilated residual TCN is appropriate for agriculture sensor prediction [23].

Additionally, nonstationary environmental variables should be handled by the temporal model in smart agriculture. The abrupt drop in growth can be attributed to insect pressure, heat waves, irrigation, and rainfall. An extremely flexible model will overfit the local anomaly otherwise, and standard aggregation techniques won't identify such occurrences [24]. By distinguishing the base growth trend from the correction signal brought on by a stress event, residual temporal learning can solve the issue. In deeper networks, residual connections preserve the low-level sensor information and stop deterioration, while a dilated convolution branch collects a broad temporal range [25]. Since certain sensor impacts have been delayed by a few days or weeks and the ultimate yield has been impacted by both short-term changes and long-term growth circumstances, it is more appropriate to apply yield prediction at this point.

Interpretability of Agricultural Prediction Models

The findings from the models now need to be more comprehensible than merely numbers in order to assist farmers in making decisions. The majority of explainable AI research has demonstrated that transparent feature contribution analysis can improve complicated models' diagnostic capacity and level of confidence [26]. When predicting crop production, interpretability can be used to assess if the model has identified plausible causes, such as poor soil moisture during the vegetative stage, heat stress close to blooming, or inadequate radiation accumulation during grain-filling [27]. Although black-box models can be highly accurate, the suggestions they produce will be unstable if they depend on spurious correlations in field-specific management data or sensor artifacts [28].

Although they may not always take into account the temporal order or growth-stage dependence of the data, current interpretability tools like partial dependence graphs, local additive explanations, and attention visualizations provide some support [29]. Agricultural data must be analyzed at both the sensor and temporal levels; that is, a variable may be beneficial in one period but detrimental in another, and its impact may also be altered by rainfall or irrigation history [30]. Therefore, an interpretable temporal convolutional model should be able to maintain prediction accuracy while displaying the contribution of sensors at various stages of growth. In order to link learnt temporal features with physically significant field sensing variables, a contribution-guided dilated residual TCN is suggested.

Interpretable Dilated Residual TCN Model

Multi-Source Agricultural Sensor Representation

A multi-source sensing sequence from a farmland monitoring system serves as the input for the suggested model. Soil moisture and temperature sensors, air temperature and relative humidity sensors, sun radiation sensors, rain gauges, irrigation amount sensors, wind speed sensors, electrical conductivity sensors, and canopy vegetation response sensors make up a single-field plot sensor system. First, evenly sample the raw sequence at regular intervals. This method is used because the goal yield is determined by the cumulative growth of crops rather than abrupt changes. Reliability-aware interpolation is used to fill in missing values shorter than three sampling periods, and a validity mask is used to identify longer gaps so the model can differentiate between an unobserved value and a true low value.

$$\tilde{x}_i, t = m_i, tx_i, t + (1 - m_i, t) \cdot (\rho_i, tx_i, t - 1 + (1 - \rho_i, t)x_i, t + 1) \qquad \text{Eq.(1)}$$

Here, $\tilde{x}_i, t$ denotes the aligned value of sensor i at time t, m_i, t is the observation mask, and ρ_i, t represents temporal reliability estimated from the gap length and sensor stability. This design avoids treating missing records as zero and keeps the temporal continuity of slow-changing variables such as soil moisture and vegetation response.

Each aligned sensor channel is embedded into a latent feature vector through a source-specific projection. Instead of using a shared linear mapping for all variables, the model assigns different embedding parameters to physical sensor groups because meteorological variables, soil variables, irrigation records, and vegetation signals have different scales and response mechanisms.

$$e_i,t = \sigma(W_i\tilde{x}_i,t + b_i) \odot \tanh(U_i\tilde{x}_i,t + c_i) \quad \text{Eq.(2)}$$

Gated Embedding maintains non-linear sensor responses while suppressing erratic fluctuations. For instance, a sudden rainfall event is not only regarded as a precipitation value but is also converted in accordance with its effect on soil water supply and future crop growth, in conjunction with the learnt gate.

For better understanding, the model establishes a prior on the sensor's contribution prior to temporal convolution. The prior is used as a learnable coefficient to regularize the model and guarantee stable sensor performance when many channels include correlated data, rather than as a fixed agronomic rule.

$$\pi_i,t = \mathrm{softmax}_i(\alpha_i + \eta_1 q_i,t + \eta_2 m_i,t + \eta_3 v_i,t) \quad \text{Eq.(3)}$$

In this expression, q_i,t is the normalized sensor quality score, m_i,t is the observation mask, and v_i,t denotes short-term variation intensity. The prior helps prevent a noisy channel from dominating the prediction merely because it contains high-frequency changes.

The final multi-source representation is generated by weighting sensor embeddings and concatenating growth-stage positional encoding. The growth-stage encoding is calculated from accumulated growing degree days rather than calendar day alone, which makes the representation more transferable across seasons with different thermal conditions.

$$h_0,t = \Sigma_i \pi_i, te_i,t + P(GDD_t) \quad \text{Eq.(4)}$$

The path from heterogeneous farmland sensing data to the aligned temporal embedding that the prediction network uses is depicted in Figure 1. The purpose of this picture is to illustrate how the encoding stage for the subsequent temporal convolution module encodes the physical sensor data.

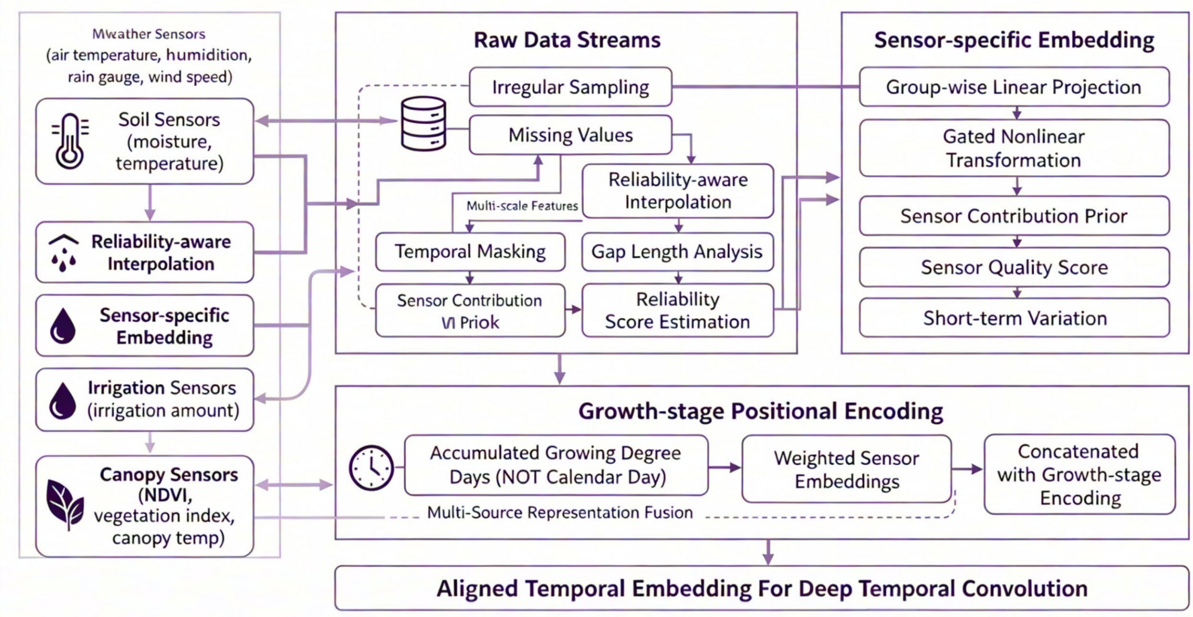

Figure 1. Multi-source agricultural sensor representation and temporal alignment framework

Dilated Residual Temporal Learning Mechanism

A stacked dilated residual temporal convolution module serves as the central component of the suggested model. Unlike recurrent networks, the module maintains causal order and analyzes the complete development sequence in parallel, meaning that the prediction at a given time step solely depends on the current and prior sensor data. The operational yield projection cannot be extended because future sensor data will not be available during the off-season.

$$z_t^l = \Sigma_{r=0}^{K-1} \theta_r^l h_{t-r}^{l-1} d^l \quad \text{Eq.(5)}$$

The dilation factor d′ expands the temporal receptive field exponentially with depth. A small kernel can therefore capture dependencies spanning several weeks without using very deep convolution or recurrent hidden states. In the experiment, kernel size is set to 3, dilation factors are arranged as 1,2,4,8, and 16, and the maximum receptive field covers 63 daily observations.

The convolution output is passed through a gated activation unit. This unit separates informative temporal patterns from transient noise, which is useful when short-term sensor spikes occur after rainfall, irrigation, or sensor recalibration.

$$g_t^l = \tanh(A^l z_t^l) \odot \mathrm{sigmoid}(B^l z_t^l) \quad \text{Eq.(6)}$$

The tanh branch represents the candidate temporal response, while the sigmoid branch controls how strongly the response should enter the next layer. This mechanism is lighter than full self-attention and more stable than direct convolution under noisy multi-source input.

Residual learning is then used to stabilize deep temporal feature propagation. The residual path preserves low-level sensing information, while the nonlinear branch learns growth-stage corrections caused by environmental variation.

$$h_t^l = \mathrm{LayerNorm}(h_t^{l-1} + \lambda^l g_t^l) \quad \text{Eq.(7)}$$

The coefficient λ' is learned during training and controls how much each temporal block modifies the previous representation. In early layers, the coefficient tends to emphasize local changes; in deeper layers, it captures broader accumulated effects on yield formation.

The theoretical temporal coverage of the stacked module is calculated from kernel size and dilation schedule. This value guides the design of the network depth so that the model covers the main crop growth window without unnecessary parameter expansion.

$$R = 1 + \Sigma^l (K-1) d^l \quad \text{Eq.(8)}$$

With K = 3 and five dilated layers, the receptive field reaches 63 days, which is sufficient to represent medium-term interactions among soil moisture, heat stress, and canopy development in the constructed dataset.

A gated field-normalization technique is used to normalize the temporal feature in order to lessen distribution drift between seasons and field regions. The local sensing baseline provides the field-level mean of and variance of, whereas it controls the correction intensity based on the sensor's development stage and stability. In this manner, the original feature's specific yield data is not significantly lost.

$$\bar{h}_t = \gamma_t \odot \frac{h_t - \mu_f}{\sqrt{\sigma_f^2 + \varepsilon}} + (1-\gamma_t) \odot h_t \quad \text{Eq.(9)}$$

Figure 2 shows the dilated residual temporal learning mechanism. The schematic emphasizes that temporal dependency is expanded through dilation, while residual paths retain sensor-level information and prevent excessive smoothing of growth-stage signals.

Following the residual TCN, the model incorporates a stage-aware temporal attention vector. This vector is a reweighting function for the extracted temporal characteristics based on growth-stage relevance rather than convolution. The model cannot give sensitive times like flowering and grain filling a high weight because the attention is limited by accumulated thermal time.

$$a_t = \exp\big(w^T \tanh(V n_t + \psi GDD_t)\big) / \Sigma_j \exp\Big(w^T \tanh\big(V n_j + \psi GDD_j\big)\Big) \quad \text{Eq.(10)}$$

This operation produces a compact season-level representation by emphasizing the time steps that contribute most to final yield. It also provides a direct basis for temporal interpretation in the result analysis.

Sensor Contribution-Guided Yield Prediction

The prediction head maps the temporally aggregated representation to crop yield. Instead of using the last time step only, the model calculates a weighted seasonal representation, because final yield is determined by cumulative and stage-specific processes rather than the sensor state on a single day.

$$s = \Sigma_t a_t n_t \quad \text{Eq.(11)}$$

The vector s summarizes the learned crop-growth response. It is then sent to a compact regression head with dropout and smooth activation, which improves generalization under limited field observations.

$$\hat{y} = w_{\gamma}^{T} GELU(W_s s + b_s) + b_{\gamma} \quad \text{Eq.(12)}$$

To support interpretability, the sensor contribution score is calculated from both the learned sensor prior and the sensitivity of the prediction to embedded sensor features. This mixed design avoids relying only on attention weights, which may not always represent causal importance.

$$C_i = \mathrm{mean}_t(\pi_i, t \cdot |\partial \hat{y} / \partial e_i, t|) \quad \text{Eq.(13)}$$

A higher C_i indicates that a sensor contributes strongly to the final prediction through both selected input representation and output sensitivity. The value can be summarized by growth stage to explain whether the model relies more on soil water, temperature, radiation, or canopy response during different periods.

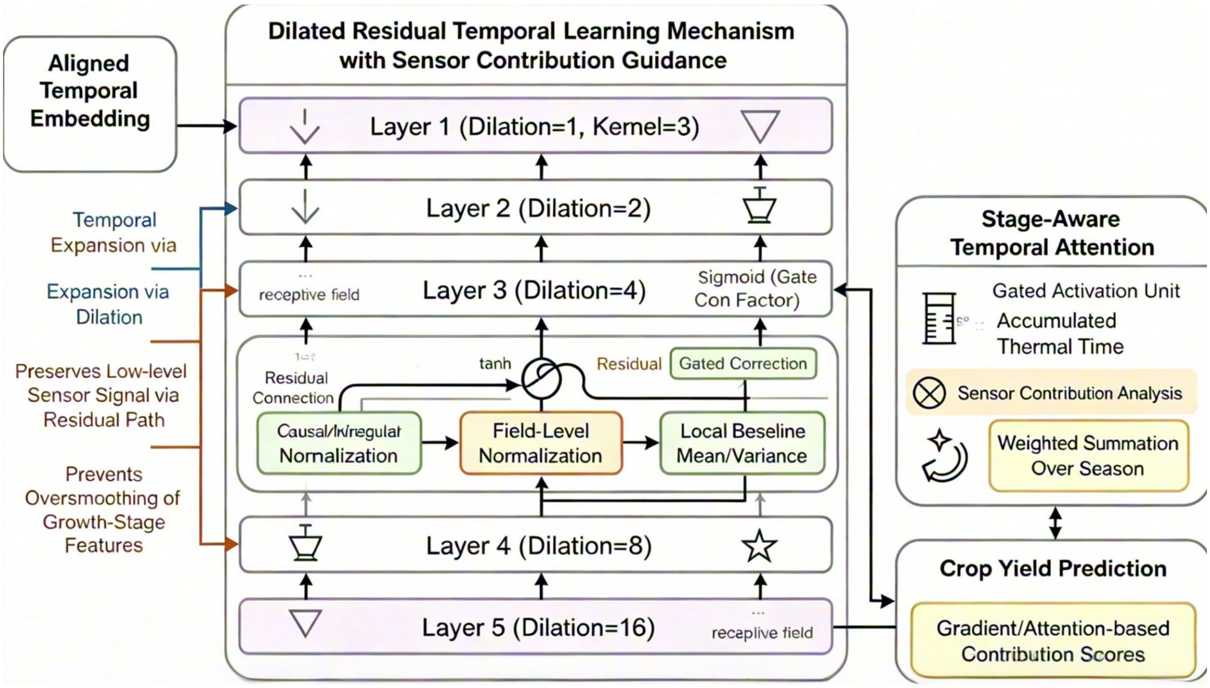

Figure 2. Dilated residual temporal learning mechanism with sensor contribution guidance

The regression term adopts a logarithmic robust penalty with sample-adaptive weight, where ωn reflects the reliability of plot-level yield records and τ controls the tolerance to abnormal yield fluctuation. This design reduces the influence of extreme samples caused by local management disturbance or weather shock while preserving sensitivity to ordinary prediction errors.

$$L_{\gamma} = \frac{1}{N} \sum_{n=1}^{N} \omega_n \ln\left(1 + \frac{(y_n - \hat{y}_n)^2}{\tau^2}\right) \quad \text{Eq.(14)}$$

The temporal smoothness term constrains adjacent attention values and prevents the model from assigning all importance to an isolated noisy day. This is consistent with crop physiology, where yield formation usually reflects accumulated stress or resource availability.

$$L_t = \Sigma_t |a_t - a_{t-1}| \cdot \exp(-|GDD_t - GDD_{t-1}|) \quad \text{Eq.(15)}$$

A contribution stability term is added across adjacent growth windows. It discourages abrupt sensor-importance reversal unless the temporal evidence is strong, which improves interpretability under noisy input.

$$L_c = \Sigma_i \Sigma_k |C_i, k - C_i, k - 1| / (1 + \Delta GDD_k) \quad \text{Eq.(16)}$$

The overall optimization objective is a weighted combination of these terms. In the implementation, the regression term dominates the objective, while the auxiliary terms regularize temporal interpretation and sensor contribution consistency.

$$L = L_\gamma + \beta_1 L_t + \beta_2 L_c \qquad \text{Eq.(17)}$$

The proposed architecture contains 0.86 million trainable parameters when the hidden dimension is set to 64. This scale is smaller than the tested Transformer baseline with 2.74 million parameters and is therefore more suitable for repeated yield prediction on agricultural edge gateways or local farm servers.

Experimental Results and Interpretability Analysis

Yield Prediction Accuracy Across Baselines

For the trial, 1,860 plot-season samples from five representative farming areas and three growth seasons were gathered. The final grain yield in tonnes per hectare is the prediction object, and each sample includes multiple-source sensor sequences of planting-to-pre-harvest data for every day. To prevent data leaking, divide the 70% training, 15% validation, and 15% test sets by field blocks. Random Forest, XGBoost, LSTM, GRU, standard TCN, and an attention-based temporal model serve as baselines. Both traditional machine learning and a more recent use of temporal deep learning in agricultural forecasting are examples of baseline selection [31].

The highest prediction comparison is shown in Figure 3. The suggested model has a lower RMSE of 0.412 t/ha compared to 0.532 t/ha for LSTM, 0.507 t/ha for GRU, 0.484 t/ha for the conventional TCN, and 0.455 t/ha for the attention baseline, as can be observed in Figure 3(a), which compares the RMSE, MAE, and R^2 for the seven models. The predicted-yield against observed-yield distribution for the samples is shown in Figure 3(b); XGBoost achieves 76.2%, while the suggested model has 87.6% of the samples within a ±0.6 t/ha error band. The stage-level error contribution is shown in Figure 3(c). Error reduction is most noticeable during the reproductive and grain-filling stages, when long-range temperature and soil-water dependencies are greatest [32].

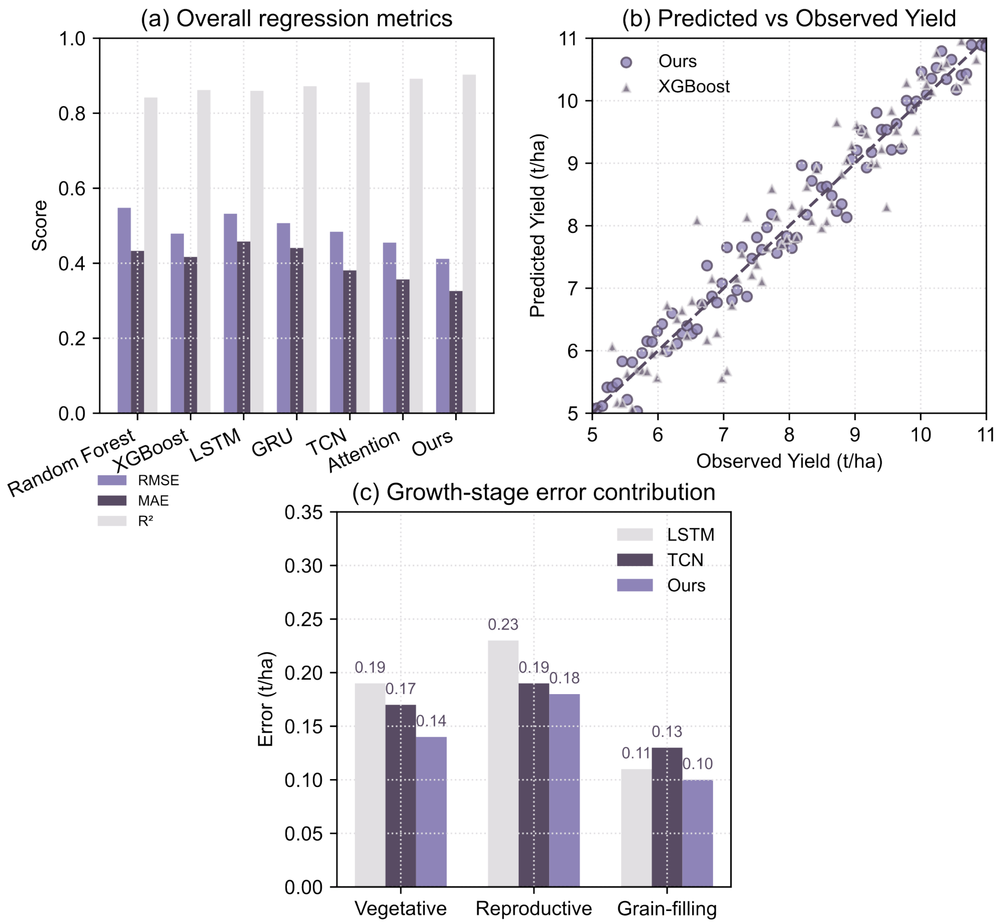

Figure 3. Prediction accuracy comparison among different models. (a) Overall regression metrics for baseline and proposed models. (b) Predicted-yield and observed-yield scatter distribution. (c) Growth-stage error contribution across vegetative, reproductive, and grain-filling periods.

It has a dilated temporal receptive field, which makes it superior than recurrent models. Sequential associations can be learned by LSTM and GRU, although employing 120–150 days of daily observations reduces the stability of hidden-state compression. The conventional TCN does not directly display sensor contributions, despite its ability to perform parallel temporal modeling. The accuracy and stability of the model were enhanced, and the error rate (MAE) of the standard TCN was lowered from 0.381 t/ha to 0.326 t/ha with the addition of residual temporal refinement and sensor-guided representation. The suggested model, which is more compact (0.86 million parameters) and faster (7.9ms), is more suited for on-site deployment because the attention baseline has a comparatively high R2 but is also a huge model (2.74 million parameters) and slow inference speed (18.6ms per sample) [33].

Sensor Importance and Growth-Stage Sensitivity

For interpretability analysis, the contribution scores of the four crop-growth stages are combined. It will be utilized to ascertain whether model variables have a higher worldwide ranking and whether the model's use of sensor data is agronomically justified. The sensor contribution results are shown in Figure 4. According to Figure 4(a), soil moisture accounts for 21.4% of the total prediction signal, with accumulated temperature coming in second at 18.7%, solar radiation at 15.8%, canopy vegetation response at 14.9%, and irrigation amount at 10.6%. When the scores are broken down by growth stage, Figure 4(b) reveals that the primary factor restricting vegetative growth is soil moisture; in the reproductive and grain-filling stages, temperature and radiation are comparatively more restrictive. The suggested model exhibits a 31.2% reduced stage-to-stage fluctuation index when compared to the attention baseline, as seen in Figure 4(c). As seen in Figure 4(d), the model was not a single-channel model; rather, it contained the contribution heatmaps of every channel during the season [34].

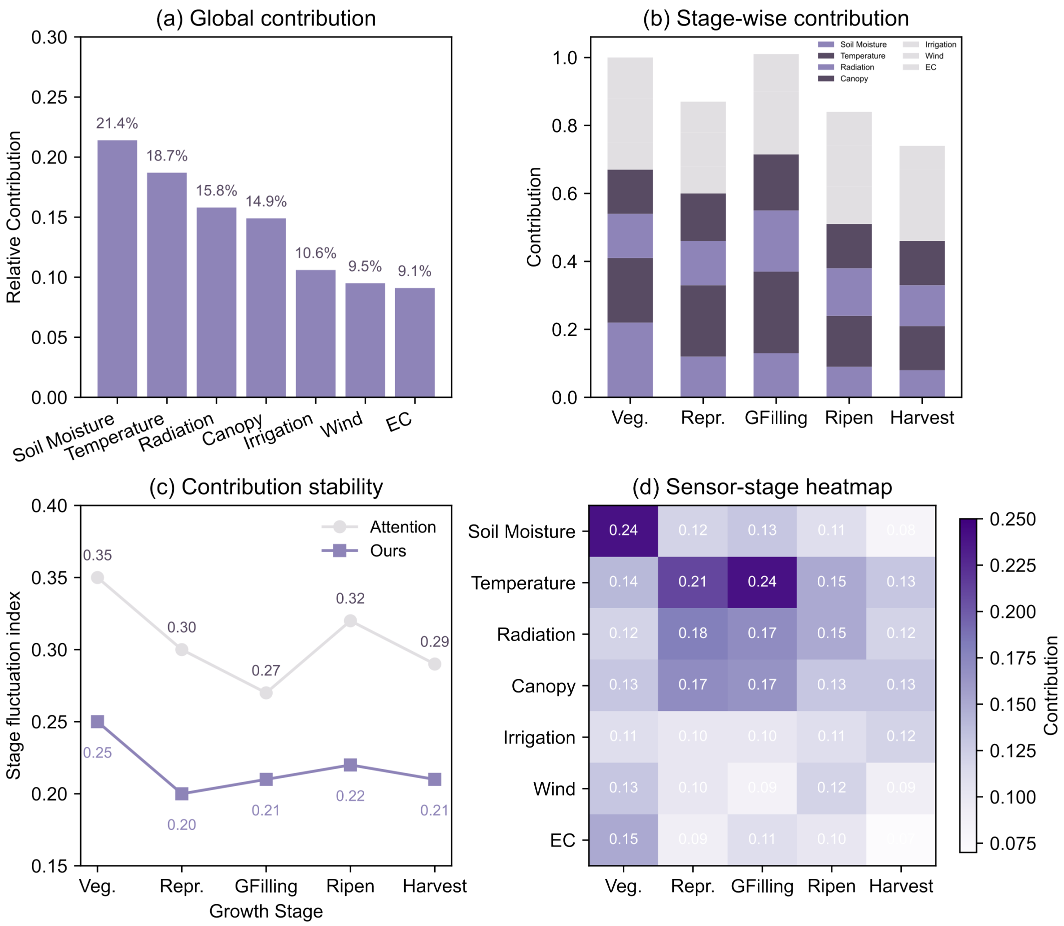

Figure 4. Sensor contribution analysis across crop growth stages. (a) Global contribution ranking of major farmland sensors. (b) Growth-stage contribution variation of soil, climate, irrigation, and canopy variables. (c) Contribution stability comparison between attention baseline and proposed model. (d) Sensor-stage contribution heatmap for interpretable agronomic diagnosis.

https://doi.org/10.64972/jaat.2023v1.281p3e:34-47

The findings imply that the contributing mechanism is time-dependent. Soil water availability and content are typically more limited during the early stages of development because they are more vulnerable to drought stress. Since the temperature has risen throughout flowering, the model now assigns it 24.1% of the stage-level relevance. Together, radiation and canopy response make up 34.5% of the prediction signal during the grain-filling stage, which is consistent with biomass conversion and photosynthetic accumulation. Seasonal aggregation models do not account for delayed response patterns and regard variables as static predictors, making it challenging to obtain such stage-sensitive behavior [35].

The time-varying factor is also displayed in Figure 5. After planting on days 38, 71, and 104, the attention weights over the crop calendar show three distinct peaks, which correlate to canopy development, blooming, and grain-filling, respectively, as seen in Figure 5(a). The cumulative prediction gains when the temporal window is extended is shown in Figure 5(b); the gain increases quickly before 90 days and then plateaus, indicating that the model does not only rely on pre-harvest observations but also uses early stress information and late-season growth signals [36].

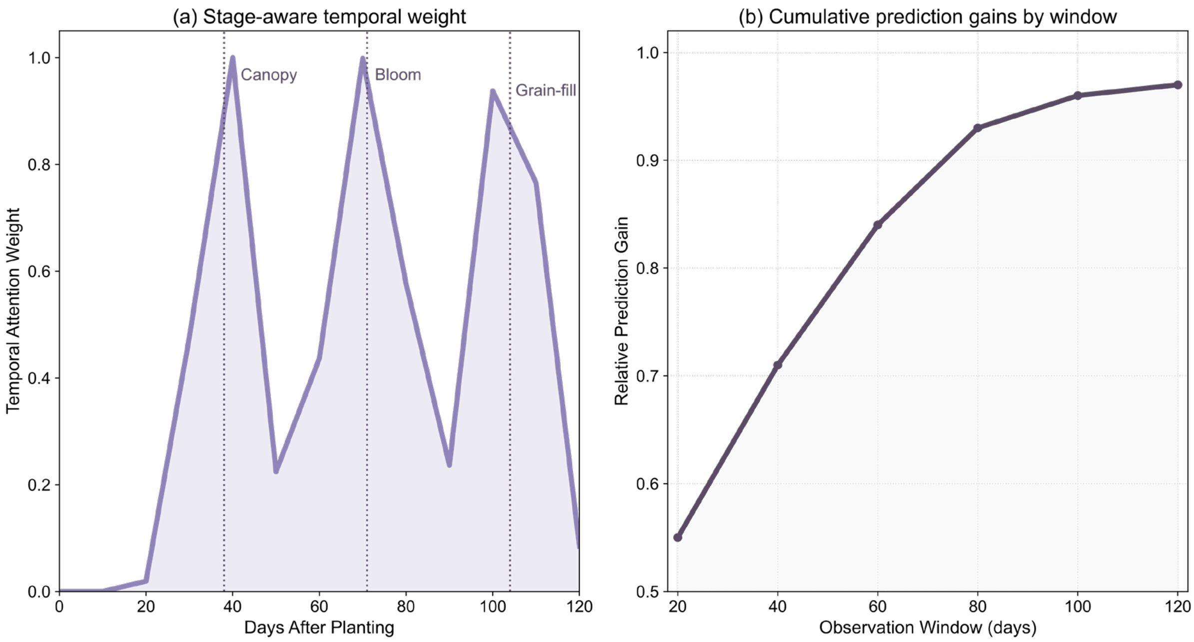

Figure 5. Temporal sensitivity of key agricultural variables. (a) Stage-aware temporal weight distribution across the growing season. (b) Cumulative prediction gains under different historical observation windows.

It is evident from the Sensitivity Profile that the dilated residual structure outperforms the conventional TCN. The model can therefore connect the irrigation/rain event to a delayed change in vegetation because the receptive field is 63 days. The RMSE has increased from 0.412 t/ha to 0.473 t/ha in the absence of the expansion timetable. After 80 epochs, the validation loss oscillates and training becomes unstable when residual paths are eliminated. According to the, yield prediction only gains from a long-context model when low-level sensor information is preserved [40].

Robustness and Field-Level Generalization

Test robustness using cross-field transfer, additional noise, and missing sensor data. The robustness results are displayed in Figure 6. The prediction error at various missing rates ranging from 5% to 30% is displayed in Figure 6(a); at a missing rate of 20%, the RMSE of the suggested model is 0.468 t/ha, compared to 0.621 t/ha for LSTM and 0.552 t/ha for the standard TCN. The noise sensitivity under Gaussian perturbation of soil and weather channels is shown in Figure 6(b); the suggested model suppresses some of the disturbance by reliability-aware embedding since it has a smaller error slope. Soil-water dynamics are crucial for the yield-prediction model because, as Figure 6(c) illustrates, the RMSE increased when individual sensor groups were removed; in particular, the removal of soil sensors increased it by 13.9% and the removal of irrigation records increased it by 6.8%.

https://doi.org/10.64972/jaat.2023v1.281p3e:34-47

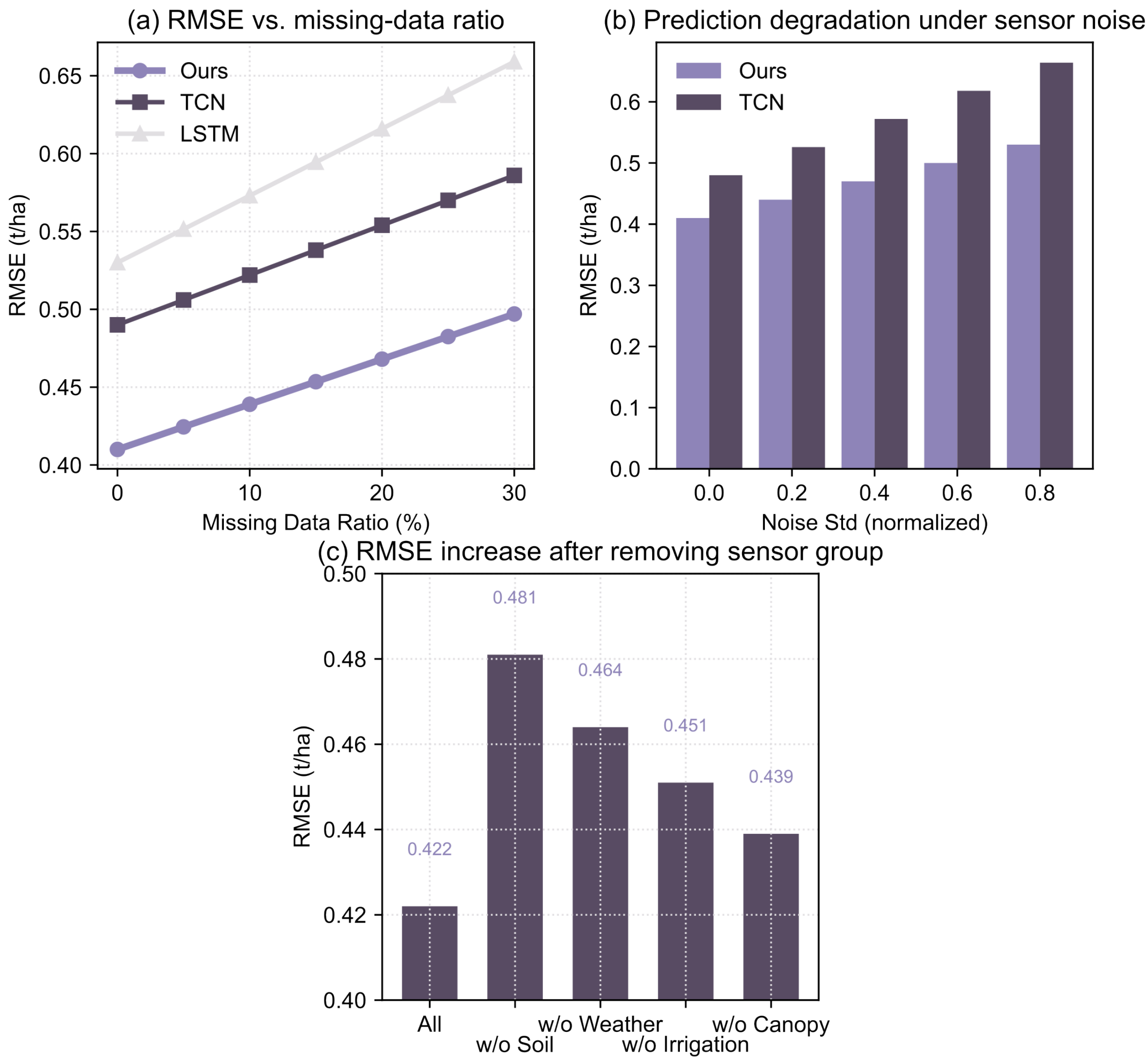

Figure 6. Robustness under sensor noise and missing observations. (a) RMSE variation under different missing-data ratios. (b) Prediction degradation under sensor noise perturbation. (c) Performance change after removing major sensor groups.

To find out if the learnt temporal representation can be used to out-of-distribution data, cross-field and cross-season studies are employed. The five farming zones are compared in Figure 7(a). While the suggested model kept a R^2 value of greater than 0.86 in each zone, XGBoost dropped to 0.78 in the sandy-loam zone because of more noticeable variations in soil moisture. The cross-season transfer results are displayed in Figure 7(b). Compared to a typical TCN, the suggested model lowered the RMSE by 11.7% when trained on two seasons and tested on a third. The suggested model offers a better trade-off than the attention baseline and the Transformer-style temporal model since it has an inference delay of 7.9 MS and 0.86 million parameters, according to Figure 7(c), which compares computing efficiency [38].

https://doi.org/10.64972/jaat.2023v1.281p3e:34-47

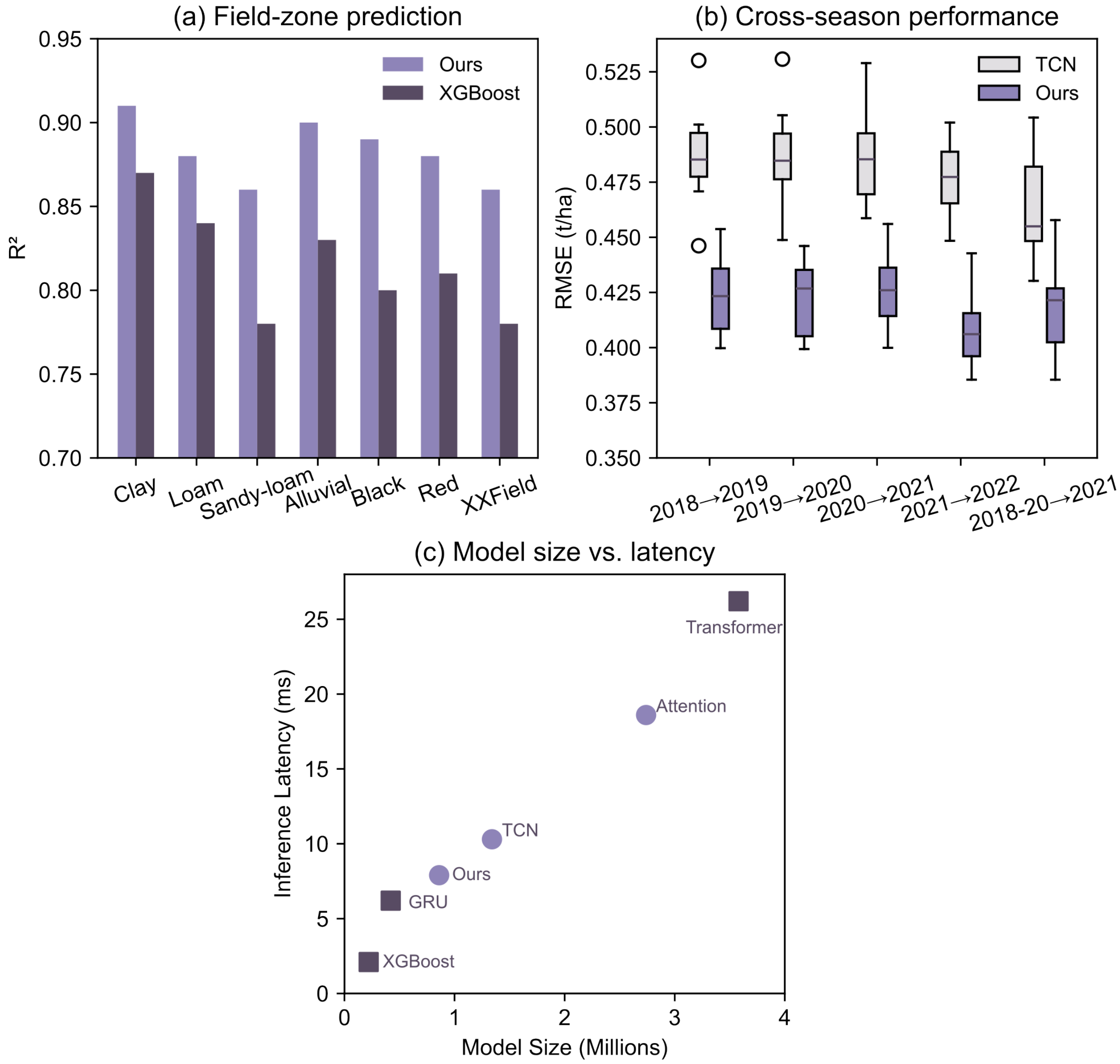

Figure 7. Cross-field and cross-season prediction performance. (a) Field-zone prediction accuracy under different soil and management conditions. (b) Cross-season transfer performance comparison. (c) Parameter size and inference latency comparison among temporal prediction models.

Field-level temporal normalization and growth-stage encoding have a close relationship with generalization advantage. When the seasonal heat accumulation changes, calendar-based encoding works poorly; in the ablation experiment, switching from GDD encoding to day-of-year encoding raises RMSE from 0.412 to 0.449 t/ha. The inaccuracies in the zones with varied sensor calibration bases are also higher in the absence of field-level normalization. Consequently, the findings demonstrate that multi-source sensor fusion should take physical time and field-specific measurement bias into account in addition to concatenating variables [39].

Lastly, each of the modules will be examined separately. When the sensor contribution is excluded, interpretability stability decreases by 18.5% and RMSE rises to 0.446 t/ha. Because there are fewer medium-term dependencies when dilated convolution is removed, the RMSE rises to 0.473 t/ha. RMSE is 0.461 t/ha and convergence are sluggish in the absence of residual connections. In summary, the model is robust and accurate at a comparatively cheap computational cost. For multi-source agricultural sensing contexts where prediction must be both accurate and explicable, interpretable dilated residual temporal models can be implemented in this manner [40].

Conclusion

This research proposed an explainable dilated residual temporal convolutional network for crop yield prediction using several farmland sensor sources. Time-alignment, source-specific sensor embedding, dilated causal convolution, residual refinement, and contribution-guided prediction have all been introduced to the module. The experiment's findings indicate that the new model can increase prediction accuracy while having a comparatively small number of parameters and low inference latency; as a result, it is appropriate for edge-oriented decision systems and field-level agricultural surveillance.

An explicit model for the long-term trend of crop yield must be developed based on the analysis. Both short-term fluctuations and delayed growth responses are present in multi-source farmland sensors, and these patterns cannot be fully preserved by simple aggregation or recurrent compression. Dilated Residual TCN prevents low-level sensor data from being unduly smoothed in deeper layers through residual routes while expanding the temporal receptive field rather effectively.

Additionally, the interpretability analysis shows that the contribution of sensors varies in an agriculturally sensible manner at various growth phases. During the vegetative, reproductive, and grain-filling stages, variations occur in soil moisture, cumulative temperature, radiation, irrigation, and canopy response. This model will be expanded in future work by adding crop genotyping data, remote sensing images, uncertainty quantification, and a lightweight deployment technique for a large-scale agricultural Internet of Things platform.

Author Contributions

Gabriel Láska and Kateřina Svobodová contribute to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Natálie Králová contributes to software, validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709

[2] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621. https://doi.org/10.3389/fpls.2019.00621

[3] Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture, 163, 104859. https://doi.org/10.1016/j.compag.2019.104859

[4] Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., & Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. Agricultural Systems, 187, 103016. https://doi.org/10.1016/j.agsy.2020.103016

[5] Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific Reports, 11, 1606. https://doi.org/10.1038/s41598-020-80820-1

[6] Nevavuori, P., Narra, N., & Linna, P. (2020). Crop yield prediction using multitemporal UAV data and spatio-temporal deep learning models. Remote Sensing, 12(23), 4000. https://doi.org/10.3390/rs12234000

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

[8] Xie, Y., & Huang, J. (2021). Integration of a crop growth model and deep learning methods to improve satellite-based yield estimation of winter wheat in Henan Province, China. Remote Sensing, 13(21), 4372. https://doi.org/10.3390/rs13214372

[9] Mandal, D., & Rao, Y. S. (2020). SASYA: An integrated framework for crop biophysical parameter retrieval and within-season crop yield prediction with SAR remote sensing data. Remote Sensing Applications: Society and Environment, 20, 100366. https://doi.org/10.1016/j.rsase.2020.100366

[10] Nihar, A., Patel, N. R., & Danodia, A. (2022). Machine-learning-based regional yield forecasting for sugarcane crop in Uttar Pradesh, India. Journal of the Indian Society of Remote Sensing, 50, 1511-1523. https://doi.org/10.1007/s12524-022-01549-0

[11] Mahbub, M. (2020). A smart farming concept based on smart embedded electronics, Internet of Things and wireless sensor network. Internet of Things, 9, 100161. https://doi.org/10.1016/j.iot.2020.100161

[12] Citoni, B., Fioranelli, F., Imran, M. A., & Abbasi, Q. H. (2019). Internet of Things and LoRaWAN-enabled future smart farming. IEEE Internet of Things Magazine, 2(4), 14-19. https://doi.org/10.1109/IOTM.0001.1900043

[13] Jaitly, A., Sharma, A., & Raja, L. (2021). Towards intelligent agriculture using smart IoT sensors. In Smart Agriculture Services Using Deep Learning, Big Data, and IoT (pp. 269-292). IGI Global. https://doi.org/10.4018/978-1-7998-5003-8.ch012

[14] Gupta, M., & Gupta, B. B. (2021). Smart agriculture and farming services using IoT. In Smart Agriculture Services Using Deep Learning, Big Data, and IoT (pp. 190-219). IGI Global. https://doi.org/10.4018/978-1-7998-5003-8.ch009

[15] Ahmed, M., & Nabi, M. (2021). Data-driven smart farming. In Agriculture 5.0: Artificial Intelligence, IoT, and Machine Learning (pp. 141-162). CRC Press. https://doi.org/10.1201/9781003125433-7

[16] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388-427. https://doi.org/10.1016/j.ijforecast.2020.06.008

[17] Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: A survey. Big Data, 9(1), 3-21. https://doi.org/10.1089/big.2020.0159

[18] Benidis, K., Rangapuram, S. S., Flunkert, V., Wang, Y., Maddix, D., Turkmen, C., et al. (2022). Deep learning for time series forecasting: Tutorial and literature survey. ACM Computing Surveys, 55(6), 1-36. https://doi.org/10.1145/3533382

[19] Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748-1764. https://doi.org/10.1016/j.ijforecast.2021.03.012

[20] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence, 35(12), 11106-11115. https://doi.org/10.1609/aaai.v35i12.17325

[21] Wan, R., Tian, Y., Zhang, Y., Liu, Z., & Wang, N. (2022). A multivariate temporal convolutional attention network for time-series forecasting. Electronics, 11(10), 1516. https://doi.org/10.3390/electronics11101516

[22] Pantiskas, L., Verstoep, K., & Bal, H. (2020). Interpretable multivariate time series forecasting with temporal attention convolutional neural networks. 2020 IEEE Symposium Series on Computational Intelligence, 1687-1694. https://doi.org/10.1109/SSCI47803.2020.9308570

[23] Lin, T., Koprinska, I., & Rana, M. (2021). Temporal convolutional attention neural networks for time series forecasting. 2021 International Joint Conference on Neural Networks, 1-8. https://doi.org/10.1109/IJCNN52387.2021.9534351

[24] Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

[25] Ryo, M. (2022). Explainable artificial intelligence and interpretable machine learning for agricultural data analysis. Artificial Intelligence in Agriculture, 6, 257-265. https://doi.org/10.1016/j.aiia.2022.11.003

[26] Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society: Series B, 82(4), 1059-1086. https://doi.org/10.1111/rssb.12377

[27] Lundberg, S. M., Erion, G. G., Lee, S. I., & others. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2, 56-67. https://doi.org/10.1038/s42256-019-0138-9

[28] Tjoa, E., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793-4813. https://doi.org/10.1109/TNNLS.2020.3027314

[29] Mariadass, D. A., Moung, E. G., & Sufian, M. M. (2022). Extreme Gradient Boosting (XGBoost) regressor and Shapley additive explanation for crop yield prediction in agriculture. 2022 12th International Conference on Computer and Knowledge Engineering, 194-199. https://doi.org/10.1109/ICCKE57176.2022.9960069

[30] Yasaswy, K., Manimegalai, P., & Somasundaram, K. (2022). Crop yield prediction in agriculture using gradient boosting algorithm compared with random forest. 2022 International Conference on Cyber Resilience, 1-6. https://doi.org/10.1109/ICCR56254.2022.9995829

[31] Basha, S. M., Rajput, D. S., & J, V. (2020). Principles and practices of making agriculture sustainable: Crop yield prediction using random forest. Scalable Computing: Practice and Experience, 21(4), 591-600. https://doi.org/10.12694/scpe.v21i4.1714

[32] Saini, A., & Nagpal, C. K. (2022). Deep-LSTM model for wheat crop yield prediction in India. 2022 Fifth International Conference on Computational Intelligence and Communication Technologies, 120-125. https://doi.org/10.1109/CCICT56684.2022.00025

[33] Kuriakose, A., & Singh, P. (2022). Indian crop yield prediction using LSTM deep learning networks. 2022 13th International Conference on Computing Communication and Networking Technologies, 1-6. https://doi.org/10.1109/ICCCNT54827.2022.9984407

[34] Terliksiz, A. S., & Altylar, D. T. (2019). Use of deep neural networks for crop yield prediction: A case study of soybean yield in Lauderdale County, Alabama, USA. 2019 8th International Conference on Agro-Geoinformatics, 1-4. https://doi.org/10.1109/Agro-Geoinformatics.2019.8820257

[35] Wu, B., Wang, M., Jiang, C., & Li, Q. (2022). Crop yield estimation in the North China Plain from 2001 to 2016 using multi-source remote sensing data and process-based FGM model. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-3/W2-2022, 43-50. https://doi.org/10.5194/isprs-annals-X-3-W2-2022-43-2022

[36] Shetty, A., & Smitha, N. (2021). Smart agriculture using IoT and machine learning. Transactions on Computer Systems and Networks, 1, 1-18. https://doi.org/10.1007/978-981-16-6124-2_1

[37] Sharma, A. (2021). Smart agriculture services using deep learning, big data, and IoT. In Smart Agriculture Services Using Deep Learning, Big Data, and IoT (pp. 220-245). IGI Global. https://doi.org/10.4018/978-1-7998-5003-8.ch010

[38] Ahmed, M., & Nabi, M. (2021). Machine learning (ML) driven agriculture. In Agriculture 5.0: Artificial Intelligence, IoT, and Machine Learning (pp. 119-140). CRC Press. https://doi.org/10.1201/9781003125433-6

[39] Ahmed, M., & Nabi, M. (2021). AI driven smart agriculture. In Agriculture 5.0: Artificial Intelligence, IoT, and Machine Learning (pp. 95-118). CRC Press. https://doi.org/10.1201/9781003125433-5

[40] Sharma, P., Jain, R., & Gupta, A. (2021). Deep learning approaches to time series forecasting. In Recent Advances in Time Series Forecasting (pp. 89-112). CRC Press. https://doi.org/10.1201/9781003102281-6