IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

Transformer Architectures for Multi-Crop Yield Forecasting: A Systematic Review of Multi-Source Remote Sensing Fusion and Interpretable Deep Learning Approaches (2020–2025)

Seema Sharma[1] and Dr. Praveen Gupta[2]
ARNI University, Tanda, Himachal Pradesh[1,2]
seemasharma1974@gmail.com

Abstract: *Accurate crop yield forecasting is fundamental to global food security, agricultural policy, and supply chain management. Over the past decade, deep learning approaches—particularly convolutional neural networks and recurrent architectures—have substantially advanced yield prediction accuracy beyond traditional statistical and machine learning methods. More recently, transformer architectures, originally developed for natural language processing, have emerged as a promising paradigm for agricultural forecasting due to their capacity to model long-range temporal dependencies and fuse heterogeneous data streams through self-attention mechanisms. This systematic review synthesises the peer-reviewed literature published between 2020 and 2025 on transformer-based crop yield forecasting, multi-source remote sensing data fusion, and explainable artificial intelligence in agricultural deep learning. Following a structured search across Scopus, Web of Science, IEEE Xplore, and Google Scholar, we identify and critically analyse studies spanning convolutional, recurrent, hybrid, and transformer architectures applied to yield prediction at field, county, and regional scales. Our synthesis reveals that while transformers such as the Multi-Modal Spatial-Temporal Vision Transformer (MMST-ViT), Informer, and Temporal Fusion Transformer have demonstrated competitive or superior performance compared to CNN-LSTM baselines, most studies remain confined to single crops in data-rich regions of the United States and China. Multi-crop forecasting frameworks, interpretable model designs, and uncertainty-aware architectures remain underexplored. A dedicated analysis of Canadian crop yield forecasting literature reveals that no published study has applied transformer architectures to multi-crop regional yield prediction across the Canadian Prairies, representing a significant and defensible research gap. We conclude with a future research agenda addressing multimodal transformer design, operational deployment, physics-informed learning, and the specific needs of Canadian Prairie agriculture.*

Keywords: transformer; crop yield forecasting; remote sensing; deep learning; explainable artificial intelligence; multi-crop prediction; data fusion; attention mechanism.

I. INTRODUCTION

Global food systems face unprecedented pressure from population growth, climate variability, and shifting land-use patterns. Reliable pre-harvest crop yield forecasts are essential for informing trade policy, managing commodity markets, triggering early-warning food security mechanisms, and guiding on-farm decision-making. Historically, yield forecasting relied on process-based crop simulation models and statistical regression approaches that linked remotely sensed vegetation indices or agroclimatic variables to observed yields. While these methods remain operationally valuable—Canada's Integrated Canadian Crop Yield Forecaster (ICCYF) employs Bayesian regression with MODIS-

DOI: 10.48175/IJARSCT-37840

derived indices (Chipanshi et al., 2015)—they struggle to capture the nonlinear, spatiotemporally heterogeneous dynamics that characterise modern agricultural landscapes.

The advent of machine learning offered a partial remedy. Random forests, support vector machines, and gradient-boosted ensembles improved predictive accuracy over linear regression by accommodating nonlinear feature interactions (Chlingaryan et al., 2018; Van Klompenburg et al., 2020). Yet these models typically require extensive manual feature engineering and do not natively exploit the spatial or temporal structure inherent in remote sensing time series. The deep learning revolution, beginning with the seminal work of You et al. (2017), who combined convolutional neural networks (CNNs) with long short-term memory (LSTM) networks and Gaussian processes for soybean yield prediction from MODIS imagery, demonstrated that end-to-end representation learning could bypass manual feature construction while improving accuracy by approximately 15% in mean absolute percentage error relative to prior approaches.

Subsequent work rapidly expanded the deep learning toolkit for yield forecasting. Khaki et al. (2020) developed a CNN-RNN framework achieving relative root mean square errors (RRMSE) of 8–9% for corn and soybean in the U.S. Corn Belt. Sun et al. (2019) and Wang et al. (2020) introduced hybrid CNN-LSTM architectures that jointly modelled spatial features from satellite imagery and temporal dynamics across the growing season. Gavahi et al. (2021) proposed DeepYield, which coupled ConvLSTM with three-dimensional CNNs for county-level soybean prediction across 1,836 U.S. counties. These architectures consistently outperformed classical machine learning baselines, establishing deep learning as the methodological frontier for yield forecasting.

The introduction of the transformer architecture by Vaswani et al. (2017) catalysed a paradigm shift across machine learning. Unlike recurrent networks, transformers process entire sequences in parallel through self-attention mechanisms, enabling efficient capture of long-range dependencies without the vanishing-gradient limitations of LSTMs. The Vision Transformer (ViT) extended this paradigm to image recognition by treating image patches as token sequences (Dosovitskiy et al., 2021), while the Temporal Fusion Transformer (TFT) introduced interpretable multi-horizon time-series forecasting with built-in variable selection and temporal attention (Lim et al., 2021). These architectural innovations present compelling opportunities for agricultural yield forecasting, where predictions depend on integrating spatially distributed satellite imagery, temporally varying weather observations, and static soil properties across multi-month growing seasons.

Despite growing interest, the application of transformers to crop yield forecasting remains nascent. Recent reviews have surveyed deep learning for yield prediction broadly (Muruganantham et al., 2022; Oikonomidis et al., 2023; Van Klompenburg et al., 2020), but none has systematically examined the emerging transformer-specific literature alongside the critical questions of multi-source data fusion and model interpretability. Three interconnected gaps motivate the present review. First, no comprehensive synthesis exists of transformer architectures designed specifically for yield forecasting. Second, the question of how multi-source remote sensing fusion—combining optical, synthetic aperture radar (SAR), thermal, and meteorological data—interacts with transformer design choices remains largely unaddressed in the review literature. Third, explainable artificial intelligence (XAI) methods such as SHapley Additive exPlanations (SHAP), attention visualisation, and integrated gradients have seen limited integration with transformer-based yield models, despite the agronomic imperative for interpretable predictions that build stakeholder trust and inform management decisions.

This review addresses these gaps by providing the first systematic synthesis of transformer-based crop yield forecasting, contextualised within the broader deep learning and remote sensing fusion literature (2020–2025). We pay particular attention to multi-crop forecasting—predicting yields for several crop types within a unified modelling framework—which remains rare despite its operational importance in mixed-cropping regions such as the Canadian Prairies. A dedicated section analyses the Canadian yield forecasting literature to establish whether a transformer-based multi-crop framework exists for this agriculturally significant region, thereby grounding the research gap claim that underpins future work.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

II. REVIEW METHODOLOGY

This review follows the principles of systematic literature review as described in the PRISMA framework, adapted for a narrative synthesis appropriate to the heterogeneous modelling approaches under consideration. We queried Scopus, Web of Science, IEEE Xplore, ScienceDirect, MDPI journals, Google Scholar, and Semantic Scholar using Boolean combinations of transformer-, yield-, and remote sensing-related terms, with complementary queries targeting explainability and the Canadian context.

Studies were included if they (a) applied machine learning, deep learning, or transformer architectures to crop yield prediction or forecasting; (b) used remotely sensed, weather, or management data; (c) were published in peer-reviewed journals, conferences, or established preprint repositories between 2020 and 2025; and (d) reported quantitative evaluation metrics. Seminal pre-2020 works providing foundational context (Chipanshi et al., 2015; Vaswani et al., 2017; You et al., 2017) were retained where necessary. The final review draws on 49 verified primary studies coded along eight dimensions: architecture type, crop(s), region, data sources, spatial scale, temporal horizon, metrics, and explainability methods.

III. CONCEPTUAL AND TECHNICAL BACKGROUND

Crop yield forecasting seeks to predict harvested output per unit area before harvest occurs. The problem is multivariate, spatiotemporal, and scale-dependent, with inputs comprising time-series remote sensing observations, meteorological records, static soil properties, and management data. Modern forecasting leverages MODIS (daily, 250–500 m), Landsat 8/9 (30 m, 16-day), Sentinel-2 (10–20 m, 5-day), and Sentinel-1 SAR (cloud-independent C-band), complemented by UAV platforms at centimetre resolution (Fei et al., 2023; Muruganantham et al., 2022; Toledo et al., 2024).

Early machine learning approaches—random forests, support vector regression, and gradient-boosted trees—improved on linear regression but required hand-crafted features (Chlingaryan et al., 2018; Van Klompenburg et al., 2020). CNNs automated spatial feature extraction while recurrent networks captured temporal dynamics; hybrid CNN-LSTM architectures combining both strengths became dominant by 2020 (Khaki et al., 2020; Sun et al., 2019).

The transformer architecture (Vaswani et al., 2017) replaces recurrence with self-attention, computing pairwise relevance scores across all sequence positions simultaneously as Attention(Q, K, V) = $Softmax(\frac{QK^T}{\sqrt{d_k}})$V. Multi-head attention projects queries, keys, and values into multiple subspaces while positional encoding injects sequence order. Three variants are particularly relevant: the Vision Transformer partitions images into patches for transformer encoding (Dosovitskiy et al., 2021); the Temporal Fusion Transformer adds variable selection and interpretable multi-head attention for multi-horizon forecasting (Lim et al., 2021); and the Informer employs ProbSparse attention to handle long sequences efficiently.

Explainable AI methods relevant to yield forecasting include SHAP (Lundberg & Lee, 2017), regression activation mapping (Wolanin et al., 2020), attention weight visualisation, integrated gradients, and Monte Carlo dropout for approximate Bayesian uncertainty estimation (Ma et al., 2021).

IV. SYSTEMATIC REVIEW AND THEMATIC SYNTHESIS

Before deep learning transformed yield prediction, machine learning methods set the performance baselines. Van Klompenburg et al. (2020) reviewed 567 studies, identifying neural networks, random forests, and SVMs as most frequent. Chlingaryan et al. (2018) noted that random forests offered the best accuracy-to-complexity trade-off for tabular data. Shahhosseini et al. (2021a) showed that integrating APSIM crop simulation outputs with gradient-boosted ensembles reduced corn RMSE by 7–20% in the U.S. Corn Belt, foreshadowing physics-informed deep learning.

The transition to deep learning began with You et al. (2017), who converted MODIS reflectance into histograms processed through CNN-LSTM-Gaussian Process branches for soybean. Khaki and Wang (2019) achieved RMSE near 12% of average yield for maize, later extending to a CNN-RNN framework with RRMSE of 9% for corn and 8% for

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

soybean across 13 U.S. states (Khaki et al., 2020). Sun et al. (2019) reduced RMSE by 8.2% over CNN-only and 9.3% over LSTM-only baselines for county-level soybean. Wang et al. (2020) achieved $R^2 = 0.77$ for Chinese winter wheat with a dual-branch design. DeepYield (Gavahi et al., 2021) combined ConvLSTM with 3D CNNs across 1,836 U.S. counties, while Shahhosseini et al. (2021b) reported RRMSE of 8.5% with CNN-DNN ensembles.

Transformer-based yield forecasting has emerged rapidly since 2022. MMST-ViT (Lin et al., 2023), presented at ICCV, processes Sentinel-2 imagery and meteorological data through separate spatial and temporal transformer encoders, achieving soybean $R^2 = 0.843$ across U.S. counties for corn, soybean, cotton, and winter wheat. The associated CropNet dataset (Lin et al., 2024) provides the first terabyte-scale multimodal benchmark. Liu et al. (2022) applied the Informer architecture to rice yield prediction in the Indo-Gangetic Plains, achieving $R^2 = 0.81$ and RMSE = 0.41 t/ha. Guo et al. (2024) developed a Spatio-Temporal Learning Framework transformer for Chinese winter wheat, identifying FPAR and LAI at heading-filling stage as most influential. Bi et al. (2023) applied a ViT architecture to soybean in Ontario, Canada, achieving a 40% RMSE reduction versus CNN-LSTM baselines—the only transformer study with a Canadian setting. Kamangir et al. (2024) proposed CMAViT for California vineyard yield, reporting $R^2 = 0.84$ and MAPE = 8.22%.

The complementarity of optical and radar observations makes multi-source fusion a natural strategy. Kalecinski et al. (2024) showed that combined SAR-optical models reduced prediction error by 50% relative to optical alone across 258 Arkansas fields. Alebele et al. (2021) raised rice R^2 from 0.65 to 0.81 by combining Sentinel-1 coherence with Sentinel-2 indices. Liao et al. (2019) blended Landsat-8 and MODIS for Ontario corn and soybean, while He et al. (2018) achieved county-level production correlations of $r = 0.96$ across seven Montana crops. At sub-field resolution, Mena et al. (2025) introduced an adaptive Multi-Modal Gated Fusion model achieving $R^2 \approx 0.80$ across soybean, wheat, and rapeseed in Argentina, Uruguay, and Germany.

Effective prediction demands more than spectral reflectance. Fei et al. (2023) demonstrated that three-sensor UAV fusion consistently outperformed any single sensor for wheat ($R^2 > 0.70$ with DNN and RF). Toledo et al. (2024) integrated UAV hyperspectral and LiDAR time-series with attention mechanisms for maize ($R^2 > 0.80$), producing attention weights aligned with phenological stages. Kamangir et al. (2024) showed that removing management data from CMAViT reduced R^2 from 0.84 to 0.73, quantifying each modality's contribution.

A striking gap is the dominance of single-crop studies (Oikonomidis et al., 2023). Khaki et al. (2021) developed YieldNet, the first CNN predicting corn and soybean simultaneously (MAE ≈ 8.7% of average yield). Lu et al. (2024) extended this with a CNN-LSTM-Attention model for maize, rice, and soybean in Northeast China, achieving relative errors below 10% for maize and rice but above 20% for soybean in some subregions. Lin et al. (2023) demonstrated MMST-ViT across four crops, but treats each independently at inference time rather than learning shared cross-crop representations.

Spatial transferability remains a critical bottleneck. Wang et al. (2018) demonstrated deep transfer learning from U.S. soybean data to Argentina, while Joshi et al. (2024) compared three transfer strategies for winter wheat, achieving MAE reductions of 9–28%. Transformer-based transfer learning for yield forecasting remains essentially unexplored.

Regarding explainability, Wolanin et al. (2020) pioneered regression activation mapping for Indian wheat, revealing downward shortwave radiation as most influential. Srivastava et al. (2022) applied SHAP to a CNN for winter wheat across 271 German counties. Paudel et al. (2023) evaluated gradient-based attribution for LSTM and CNN models in the European MARS system, finding that neural networks learned features comparable to expert-designed ones but struggled under extreme temperature and moisture. Hu et al. (2023) delivered a sobering finding: an intrinsically interpretable Bayesian ensemble was the only method among seven tested that correctly uncovered true climate-crop relationships on synthetic data, while black-box models predicted well for wrong reasons. Ryo (2022) showed that different ML algorithms can produce divergent interpretability outputs for identical datasets. For transformers specifically, attention visualisation has been demonstrated (Guo et al., 2024; Tian et al., 2025), but its agronomic faithfulness remains debated.

Ma et al. (2021) applied a Bayesian neural network to corn yield, finding that over 84% of observed yields fell within 95% confidence intervals. Abbaszadeh et al. (2022) extended this via copula-embedded Bayesian model averaging. Tian et al. (2025) combined attention with integrated gradients and Monte Carlo dropout for wheat yield estimation. However, no published study has yet integrated transformer-based forecasting with full Bayesian uncertainty quantification. Operationally, the ICCYF (Chipanshi et al., 2015) remains one of the few documented systems, employing relatively simple statistical methods that prioritise robustness over peak accuracy.

V. COMPARATIVE CRITICAL ANALYSIS

The literature reveals three patterns. First, transformers consistently match or exceed CNN-LSTM performance, with the largest gains appearing when models must integrate heterogeneous data modalities or capture long-range temporal dependencies. Second, most transformer studies remain geographically concentrated in the United States and China. Third, multi-crop transformer forecasting is represented primarily by MMST-ViT; the remaining studies focus on single crops. Transformers offer the strongest theoretical capacity for modelling spatial-temporal dependencies, but their data requirements and computational costs are substantially higher than CNN-LSTM alternatives. The multi-crop advantage remains more theoretical than demonstrated, and native attention-based interpretability has not yet been rigorously validated against ground-truth agronomic mechanisms. Tree-based methods retain the most mature XAI tooling via SHAP (Lundberg & Lee, 2017).

VI. THE CANADIAN MULTI-CROP TRANSFORMER GAP

Canada is the world's fifth-largest agricultural exporter, with the Prairie provinces of Alberta, Saskatchewan, and Manitoba producing the majority of the nation's spring wheat, canola, barley, oats, and pulse crops. Accurate yield forecasting is essential for the Canadian Grain Commission, Agriculture and Agri-Food Canada, and international markets. The landscape presents distinctive challenges: a compressed May–September growing season, extreme heterogeneity across sub-humid, semi-arid, and arid zones, sparse meteorological coverage, and concurrent cultivation of five or more crops.

Canada's operational infrastructure relies primarily on the ICCYF (Chipanshi et al., 2015; Newlands et al., 2014), which combines Bayesian regression, MCMC algorithms, and robust least angle regression with MODIS-derived indices and a soil water budget. Evaluated for spring wheat, barley, and canola, the ICCYF achieves R^2 values of 66%, 51%, and 67% respectively at the Census Agricultural Region scale, with national-level MAPE of 8%, 5%, and 9%.

The Canadian literature has explored several ML and remote sensing pathways. Mkhabela et al. (2011) demonstrated that MODIS NDVI explained 32–90% of yield variability for Prairie crops. Johnson et al. (2016) compared Bayesian neural networks, recursive partitioning, and linear regression, finding that nonlinear models did not consistently outperform linear approaches. Zhang et al. (2019) showed that crop-specific masks from the AAFC Annual Crop Inventory improved yield predictions. White et al. (2020) demonstrated that SMOS-based soil moisture improved canola forecasts, with soil moisture selected in 74.2% of models. Kross et al. (2020) applied artificial neural networks for within-field corn and soybean in Ontario. Newlands et al. (2023) assessed satellite-based models across the Prairies using Landsat, Sentinel-2, and MODIS, finding Landsat-NDWI the best single predictor.

Critically, the only transformer-based study with any Canadian setting is Bi et al. (2023), confined to a single crop (soybean), a single province (Ontario), and field-scale camera imagery rather than satellite-based regional forecasting. No published study has applied transformer architectures to multi-crop regional yield forecasting across the Canadian Prairies—the scale relevant to AAFC and national food security monitoring. The Prairie multi-crop system represents precisely the scenario where a unified transformer could offer efficiency advantages, and the increasing availability of Sentinel-1 SAR alongside Sentinel-2 optical data over the Prairies since 2017 offers a fusion opportunity transformers are uniquely positioned to exploit.

VII. RESEARCH GAPS AND FUTURE RESEARCH AGENDA

Five interconnected frontiers emerge. First, multimodal transformer architectures must natively fuse optical, SAR, meteorological, soil, and management data through modality-specific encoders with cross-modal attention; MMST-ViT (Lin et al., 2023) and CMAViT (Kamangir et al., 2024) show feasibility but neither incorporates SAR or operational soil information. Second, multi-crop representation learning through conditional attention, crop-type embedding tokens, or modular decoders could enable single models to predict across diversified portfolios while respecting crop-specific phenology. Third, uncertainty-aware transformer forecasting via variational attention, ensemble transformers, or evidential deep learning would address the probabilistic gap. Fourth, domain adaptation through self-supervised pretraining on satellite archives followed by regional fine-tuning offers a scalable path to data-sparse regions. Fifth, physics-informed transformer learning—embedding radiation-use efficiency, water balance, and thermal time constraints—could address the concern raised by Hu et al. (2023) that purely data-driven models may predict accurately for wrong reasons. Finally, operational XAI pipelines must translate interpretability from research curiosity to production requirement, with the Canadian CCYF system offering an ideal testbed. These priorities converge on a specific proposal: a multimodal, multi-crop, interpretable transformer framework for Canadian Prairie yield forecasting fusing Sentinel-1/2, gridded climate, and soil data; predicting wheat, canola, barley, oat, and pulse yields within a unified architecture; quantifying predictive uncertainty; and providing agronomically meaningful explanations.

VIII. CONCLUSION

This review traced the evolution of deep learning for crop yield forecasting from CNN-LSTM hybrids to the transformer paradigm, synthesising 49 verified studies (2020–2025). Three findings stand out. First, transformer architectures have demonstrated competitive or superior performance relative to CNN-LSTM baselines, with MMST-ViT achieving $R^2 = 0.84$ for soybean and the Informer $R^2 = 0.81$ for rice. Second, multi-source SAR-optical fusion reduces error by up to 50% relative to single-sensor approaches, yet current transformers have not fully exploited this opportunity. Third, interpretability and uncertainty quantification remain underdeveloped in transformer-based forecasting. The dedicated Canadian analysis confirms that no published study has applied transformers to multi-crop regional yield forecasting across the Canadian Prairies—a scientifically and operationally important gap. Addressing it through multimodal fusion, multi-crop representation learning, Bayesian uncertainty, physics-informed priors, and operational XAI would advance both the methodological frontier and practical food security monitoring in one of the world's major grain-producing regions.

REFERENCES

1. Abbaszadeh, P., Gavahi, K., Alipour, A., Deb, P., & Moradkhani, H. (2022). Bayesian multi-modeling of deep neural nets for probabilistic crop yield prediction. Agricultural and Forest Meteorology, 314, 108773. https://doi.org/10.1016/j.agrformet.2021.108773
2. Alebele, Y., Fang, H., Zhang, X., Wang, H., Liang, D., Wei, S., Zheng, G., Zhao, L., & Wang, L. (2021). Estimation of crop yield from combined optical and SAR imagery using Gaussian kernel regression. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 10520–10534. https://doi.org/10.1109/JSTARS.2021.3118707
3. Bi, L., Wally, O., Hu, G., Tenuta, A. U., Kandel, Y. R., & Mueller, D. S. (2023). A transformer-based approach for early prediction of soybean yield using time-series images. Frontiers in Plant Science, 14, 1173036. https://doi.org/10.3389/fpls.2023.1173036
4. Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian, B., Daneshfar, B., Bedard, F., & Reichert, G. (2015). Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agricultural and Forest Meteorology, 206, 137–150. https://doi.org/10.1016/j.agrformet.2015.03.007

IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

5. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61–69. https://doi.org/10.1016/j.compag.2018.05.012
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. In Proceedings of the 9th International Conference on Learning Representations (ICLR 2021). https://arxiv.org/abs/2010.11929
7. Fei, S., Hassan, M. A., He, Z., Chen, Z., Shu, M., Wang, J., Li, C., & Xiao, Y. (2023). UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat. Precision Agriculture, 24, 187–212. https://doi.org/10.1007/s11119-022-09938-8
8. Gavahi, K., Abbaszadeh, P., & Moradkhani, H. (2021). DeepYield: A combined convolutional neural network with long short-term memory for crop yield forecasting. Expert Systems with Applications, 184, 115511. https://doi.org/10.1016/j.eswa.2021.115511
9. Guo, F., Wang, P., Tansey, K., Zhang, Y., Li, M., Liu, J., & Zhang, S. (2024). A novel transformer-based neural network under model interpretability for improving wheat yield estimation using remotely sensed multi-variables. Computers and Electronics in Agriculture, 223, 109606.
10. He, M., Kimball, J. S., Maneta, M. P., Maxwell, B. D., Moreno, A., Beguería, S., & Wu, X. (2018). Regional crop gross primary productivity and yield estimation using fused Landsat-MODIS data. Remote Sensing, 10(3), 372. https://doi.org/10.3390/rs10030372
11. Hu, T., Zhang, X., Bohrer, G., Liu, Y., Zhou, Y., Martin, J., Li, Y., & Zhao, K. (2023). Crop yield prediction via explainable AI and interpretable machine learning: Dangers of black box models for evaluating climate change impacts on crop yield. Agricultural and Forest Meteorology, 340, 109555. https://doi.org/10.1016/j.agrformet.2023.109555
12. Johnson, M. D., Hsieh, W. W., Cannon, A. J., Davidson, A., & Bédard, F. (2016). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agricultural and Forest Meteorology, 218–219, 74–84. https://doi.org/10.1016/j.agrformet.2015.11.003
13. Joshi, A., Pradhan, B., Chakraborty, S., Varatharajoo, R., Gite, S., & Alamri, A. M. (2024). Deep-transfer-learning strategies for crop yield prediction using climate records and satellite image time-series data. Remote Sensing, 16(24), 4804. https://doi.org/10.3390/rs16244804
14. Kalecinski, N. I., Skakun, S., Torbick, N., Huang, X., Franch, B., Roger, J.-C., & Vermote, E. (2024). Crop yield estimation at different growing stages using a synergy of SAR and optical remote sensing data. Science of Remote Sensing, 10, 100153. https://doi.org/10.1016/j.srs.2024.100153
15. Kamangir, H., Sams, B. S., Dokoozlian, N., Sanchez, L., & Earles, J. M. (2024). CMAViT: Integrating climate, management, and remote sensing data for crop yield estimation with multimodel vision transformers. arXiv. https://doi.org/10.48550/arXiv.2411.16989
16. Khaki, S., Pham, H., & Wang, L. (2021). Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning. Scientific Reports, 11, 11132. https://doi.org/10.1038/s41598-021-89779-z
17. 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
18. Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN framework for crop yield prediction. Frontiers in Plant Science, 10, 1750. https://doi.org/10.3389/fpls.2019.01750
19. Kross, A., Znoj, E., Callegari, D., Kaur, G., Sunohara, M., Lapen, D. R., & McNairn, H. (2020). Using artificial neural networks and remotely sensed data to evaluate the relative importance of variables for prediction of within-field corn and soybean yields. Remote Sensing, 12(14), 2230. https://doi.org/10.3390/rs12142230

20. Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., & Song, Y. (2019). Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. Science of the Total Environment, 650, 1707–1721. https://doi.org/10.1016/j.scitotenv.2018.09.308
21. 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
22. Lin, F., Crawford, S., Guillot, K., Zhang, Y., Chen, Y., Yuan, X., Chen, L., Williams, S., Minvielle, R., Xiao, X., Gholson, D., Ashwell, N., Setiyono, T., Tubana, B., Peng, L., Bayoumi, M., & Tzeng, N.-F. (2023). MMST-ViT: Climate change-aware crop yield prediction via multi-modal spatial-temporal vision transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 5774–5784). https://doi.org/10.1109/ICCV51070.2023.00531
23. Lin, F., Guillot, K., Crawford, S., Zhang, Y., Yuan, X., & Tzeng, N.-F. (2024). An open and large-scale dataset for multi-modal climate change-aware crop yield predictions. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 5375–5386). https://doi.org/10.1145/3637528.3671536
24. Liu, Y., Wang, S., Chen, J., Chen, B., Wang, X., Hao, D., & Sun, L. (2022). Rice yield prediction and model interpretation based on satellite and climatic indicators using a transformer method. Remote Sensing, 14(19), 5045. https://doi.org/10.3390/rs14195045
25. Lu, J., & Fu, Y. (2024). Deep learning for multi-source data-driven crop yield prediction in Northeast China. Agriculture, 14(6), 794. https://doi.org/10.3390/agriculture14060794
26. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (pp. 4765–4774).
27. Ma, Y., Zhang, Z., Kang, Y., & Özdoğan, M. (2021). Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach. Remote Sensing of Environment, 259, 112408. https://doi.org/10.1016/j.rse.2021.112408
28. Mena, F., Pathak, D. K., Najjar, H., Sanchez, C., Helber, P., Bischke, B., Habelitz, P., Lorenz, M. M., Arenas, D., Vollmer, M., & Dengel, A. (2025). Adaptive fusion of multi-modal remote sensing data for optimal sub-field crop yield prediction. Remote Sensing of Environment, 318, 114547. https://doi.org/10.1016/j.rse.2024.114547
29. Mkhabela, M. S., Bullock, P., Raj, S., Wang, S., & Yang, Y. (2011). Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. Agricultural and Forest Meteorology, 151(3), 385–393. https://doi.org/10.1016/j.agrformet.2010.11.012
30. 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
31. Newlands, N. K., Davidson, A., Daneshfar, B., & Champagne, C. (2023). Assessing the performance of satellite-based models for crop yield estimation in the Canadian Prairies. Canadian Journal of Remote Sensing, 49(1), 2252926. https://doi.org/10.1080/07038992.2023.2252926
32. Newlands, N. K., Zamar, D. S., Kouadio, L. A., Zhang, Y., Chipanshi, A., Potgieter, A., Toure, S., & Hill, H. S. J. (2014). An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. Frontiers in Environmental Science, 2, 17. https://doi.org/10.3389/fenvs.2014.00017
33. Oikonomidis, A., Catal, C., & Kassahun, A. (2023). Deep learning for crop yield prediction: A systematic literature review. New Zealand Journal of Crop and Horticultural Science, 51(1), 1–26. https://doi.org/10.1080/01140671.2022.2032213

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, July 2026
Impact Factor: 8.2

34. Paudel, D., de Wit, A., Boogaard, H., Marcos, D., Osinga, S., & Athanasiadis, I. N. (2023). Interpretability of deep learning models for crop yield forecasting. Computers and Electronics in Agriculture, 206, 107663. https://doi.org/10.1016/j.compag.2023.107663
35. 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
36. Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021a). 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
37. Shahhosseini, M., Hu, G., Khaki, S., & Archontoulis, S. V. (2021b). Corn yield prediction with ensemble CNN-DNN. Frontiers in Plant Science, 12, 709008. https://doi.org/10.3389/fpls.2021.709008
38. Srivastava, A. K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F., Gaiser, T., & Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. Scientific Reports, 12, 3215. https://doi.org/10.1038/s41598-022-06249-w
39. Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-level soybean yield prediction using deep CNN-LSTM model. Sensors, 19(20), 4363. https://doi.org/10.3390/s19204363
40. Tian, H. (2025). An interpretable wheat yield estimation model using an attention mechanism-based deep learning framework with multiple remotely sensed variables. International Journal of Applied Earth Observation and Geoinformation, 136, 104355.
41. Toledo, J. D., Crawford, M. M., & Tuinstra, M. R. (2024). Integrating multi-modal remote sensing, deep learning, and attention mechanisms for yield prediction in plant breeding experiments. Frontiers in Plant Science, 15, 1408047. https://doi.org/10.3389/fpls.2024.1408047
42. 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
43. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems 30 (pp. 5998–6008).
44. Wang, A. X., Tran, C., Desai, N., Lobell, D., & Ermon, S. (2018). Deep transfer learning for crop yield prediction with remote sensing data. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '18) (pp. 1–5). https://doi.org/10.1145/3209811.3212707
45. Wang, X., Huang, J., Feng, Q., & Yin, D. (2020). Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sensing, 12(11), 1744. https://doi.org/10.3390/rs12111744
46. White, J., Berg, A. A., Champagne, C., Zhang, Y., Chipanshi, A., & Daneshfar, B. (2020). Improving crop yield forecasts with satellite-based soil moisture estimates: An example for township level canola yield forecasts over the Canadian Prairies. International Journal of Applied Earth Observation and Geoinformation, 89, 102092. https://doi.org/10.1016/j.jag.2020.102092
47. Wolanin, A., Mateo-García, G., Camps-Valls, G., Gómez-Chova, L., Meroni, M., Duveiller, G., You, L., & Guanter, L. (2020). Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environmental Research Letters, 15(2), 024019. https://doi.org/10.1088/1748-9326/ab68ac
48. You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian Process for crop yield prediction based on remote sensing data. In Proceedings of the AAAI Conference on Artificial Intelligence, 31(1), 4559–4565. https://doi.org/10.1609/aaai.v31i1.11172
49. Zhang, Y., Chipanshi, A., Daneshfar, B., Koiter, L., Champagne, C., Davidson, A., Reichert, G., & Bédard, F. (2019). Effect of using crop specific masks on earth observation based crop yield forecasting across Canada. Remote Sensing Applications: Society and Environment, 13, 121–137. https://doi.org/10.1016/j.rsase.2018.10.002

DOI: 10.48175/IJARSCT-37840