International Journal of Agriculture and Environmental Sciences
ISSN: 3108-088X | Volume 2 Issue 2, 29-36, April-June 2026
DOI: https://doi.org/10.64137/3108088X/IJAES-V2I2P104

Received: 25/04/2026 Revised: 31/05/2026 Accepted: 04/06/2026 Published: 11/06/2026

Scholastic Research Publication

Original Article

AI-Driven Carbon Footprint Assessment in Agricultural Production

DR.JEEVAJOTHI
Kumaraguru Institute of Agriculture - [KIA], Erode, India.

ABSTRACT: *Agricultural production is simultaneously a major victim of climate change and a significant contributor to global greenhouse gas (GHG) emissions. Traditional carbon accounting frameworks in agriculture rely on static, empirical emissions factors that fail to capture the highly dynamic, spatial-temporal variations inherent in biological systems. This study proposes an integrated Artificial Intelligence (AI) and Machine Learning (ML) framework for real-time, high-resolution carbon footprint assessment across diversified cropping systems. By coupling deep learning architectures specifically Long Short-Term Memory (LSTM) networks and Gradient Boosted Decision Trees (GBDT) with multi-source remote sensing data, Internet of Things (IoT) soil sensors, and historical management logs, we predict nitrous oxide (N2O}$), methane ($\{CH}_4$), and carbon dioxide (CO}_2$) fluxes at a sub-field scale. Methods: The model was trained and validated on a large national dataset over multiple years of observations from all major agro-ecological zones. We present an AI-driven approach that predicts total GHG emissions with R^2 = 0.89, outperforming traditional Intergovernmental Panel on Climate Change (IPCC) Tier 1 and Tier 2 empirical models which have respective R^2s of 0.54 - 0.68 from the best-fit regression lines to each study's mean bulk density, carbon content data points across different studies within human-dominated ecosystems considered here. The analysis Prances and P4amongms that focused on how soil moisture behaved in one place over time, real-time estimates of the microbial pool nitrogen at random dates throughout 2015 with our infrared canopy temperature anomalies was: Steve Papazian3.nrted single emissions spikes. Moreover, scenario analysis demonstrates that AI-assisted precision nitrogen application and optimized irrigation scheduling can reduce the total carbon footprint of cereal production by as much as 34% without impacting crop yields. This work lays out a scalable, automated digital twin infrastructure for agricultural carbon accounting. It provides an accessibly robust technology foundation for validating carbon credits (attribution), tracking supply chain insets, and providing localized climate-smart decision support systems.*

KEYWORDS: *Artificial Intelligence, Carbon Footprint, Agricultural Production, Precision Agriculture, Deep Learning, Climate Change Mitigation, Greenhouse Gas Accounting.*

1. INTRODUCTION

The global agricultural sector is at an important time in the twenty-first century, as they are faced with two challenges; increasing food production to sustain a projected nine billion people by 2050, and reducing its environmental ecological footprint. Recent estimates by the Intergovernmental Panel on Climate Change indicate that emissions from agriculture, forestry and other land-use sectors account for 22 to 24% of total global anthropogenic greenhouse gas emissions. In contrast to industrial sectors for which carbon emissions are largely point-source, localized and practically all directly linked with fossil fuel combustion, agricultural emissions predominantly occur very diffusely (non-point source) across the landscape and also involve fundamentally biogenic processes. This results from complex non-linear interactions between soil microclimates, microbial communities, crop phenology and meteorological variables as well as anthropogenic management actions. The comprehensiveness of this dynamic makes accurate quantification, tracking and assessment of agricultural carbon footprints one of the most challenging barriers to sustainable supply chains.

Carbon dioxide CO, methane CH_4 and nitrous oxide N_2O are the main greenhouse gases released by agricultural production activities. Emissions of carbon dioxide are mainly due to fossil fuel use in machinery, fertilizer production and transport; soils are depleted of organic C through intense tillage practices. Sources of methane are mainly anaerobic decomposition in water-logged conditions water-filled land areas such as rice paddies and enteric fermentation from livestock systems. As a short-lived potent greenhouse gas, the global warming potential (GWP) of nitrous oxide is nearly 300 times greater than carbon dioxide on an integrated basis over a 100-year timescale and emissions from soils are thought to be driven primarily by microbial nitrification, denitrification processes that have been heavily accelerated through excessive applications of synthetic nitrogen fertilizers.

Traditionally, tracking such emissions has used evaluative frameworks based on empirical work by the IPCC. Entry 1 methodology relies on generalized, aggregate emissions factors at large regional or national levels while ignoring soil characteristics, daily weather variability and details of spatial-temporal field management. Tier 2 approaches incorporate specific sensors for the country setting but are also weak on temporal dynamics. Tier 3 methods utilize detailed mechanistic process-based biophysical models (e.g., Day Cent, DNDC DeNitrification-DeComposition) While highly detailed, these process-based models require extensive, expensive, and often unavailable calibration parameters, limiting their scalability for daily operational use across thousands of fragmented farming holdings worldwide. Consequently, a critical research gap exists: the agricultural sector lacks a scalable, cost-effective, and highly accurate methodology to quantify localised carbon footprints in real time. This limitation hampers the development of reliable carbon-credit markets, prevents food brands from accurately verifying Scope 3 supply chain emissions, and limits farmers from receiving immediate feedback on the environmental efficacy of their climate-smart management adjustments.

Artificial Intelligence (AI) and Machine Learning (ML) present a paradigm shift for this domain. By leveraging massive streams of heterogeneous data including high-resolution satellite imagery, real-time Internet of Things soil sensor networks, localised weather forecasts, and historical farm management logs—AI architectures can learn complex, non-linear biological and environmental interactions without requiring manual calibration of hundreds of biophysical equations. Due to their strength in recognizing temporal patterns and spatial anomalies, deep learning models were better at detecting sudden emission spikes like post-fertilization or post-rainfall N_2O pulses. The overall goal of this work is to design, validate and test an end-to-end AI-driven digital framework for real-time, sub-field-level carbon footprint assessment in agricultural production systems.

The primary research contributions of this paper are organised as follows:

- We develop a novel, multi-modal deep learning architecture that fuses temporal satellite observations with localised IoT sensor streams to predict site-specific N_2O/ CH_4, and $CO_2$$ fluxes.
- We establish a high-resolution spatial mapping approach that downscales carbon footprint metrics to a 10×10 meter grid resolution, providing actionable insight for precision agricultural machinery.
- We conduct an extensive, multi-regional validation comparing the AI framework against standard IPCC Tier 1 and Tier 2 models using automated eddy covariance flux tower measurements as the ground-truth baseline.
- Using this, we construct an intelligent optimization engine that simulates management scenarios to provide increasing spatial resolution recommendations for nitrogen and irrigation application with maximum carbon mitigation while ensuring crop yield is protected.

This research could provide a robust, transparent and scalable tool seeking to accelerate the shift towards verifiable climate-smart agriculture production systems at global scale by closing the gap between advanced data science and agronomic practices.

2. LITERATURE REVIEW

Over the past three decades, the academic discourse on carbon accounting in agricultural production systems has progressed from macro-scale, static estimates toward hyper-granular dynamic modelling paradigms. Early foundational work primarily focused on life cycle assessments (LCA) using linear attribution methodologies. These traditional LCA frameworks quantified the carbon footprint by summing the emissions associated with upstream inputs such as chemical manufacturing, fuel processing, and seed production and combining them with on-farm direct emissions derived from regional averages. While highly valuable for establishing baseline understandings of global food supply chains, early literature consistently noted that the high degree of uncertainty in direct on-farm emissions remained the single largest source of error in agricultural environmental accounting.

To address these uncertainties, biophysical and process-based modelling emerged as a dominant research track. Models such as the DNDC, DayCent, and APSIM (Agricultural Production Systems Simulator) simulate the complex biogeochemical cycles of carbon and nitrogen within the soil-water-plant-atmosphere continuum. Researchers have successfully deployed these models to evaluate the long-term impacts of conservation tillage, cover cropping, and varying fertiliser regimes on soil organic carbon sequestration and gaseous emissions. However, the literature frequently highlights major limitations to scaling process-based models. They demand a complete and sole input set including soil hydraulic properties, microbial kinetic coefficients (largely unmeasurable) and daily thermal gradients with precision that few farmers or supply chain auditors would be able to access. Additionally, these models are computation-heavy and tend to be less suitable for larger-scale real-time spatial applications or integration within automated digital twin architectures. Concurrently, the rapid advancements of remote sensing technologies transformed agricultural monitoring by facilitating free access to high spatial and temporal resolution datasets from satellite constellations such as Sentinel and Landsat.

Initial remote sensing literature in this domain focused heavily on estimating crop yields, biomass accumulation, and leaf area index (LAI). Gradually, researchers began utilising these vegetation indices as proxies for carbon assimilation, linking satellite-derived gross primary productivity (GPP) with net ecosystem exchange (NEE) models. However, satellite imagery alone cannot penetrate beneath the crop canopy to monitor subsurface microbial processes, meaning that remote sensing approaches historically struggled to accurately estimate direct soil $\t{N}_2\ {O}$ or $\ {CH}_4$ emissions, which represent the most volatile components of the agricultural carbon footprint. The rise of Artificial Intelligence and Machine Learning has provided a new direction with great promise, combining the strengths of both remote sensing data whilst at the same time deriving predictive insights from biophysical models. Previous applications of machine learning in soil carbon estimation have employed relatively shallow algorithms (e.g., Random Forests, Support Vector Machines [SVM], and Artificial Neural Networks [ANN]) to predict soil organic carbon concentrations or local nitrous oxide emissions. Although these studies illustrated the potential of machine learning to improve upon simple empirical equations consistently, they were usually limited by small geographically constrained training datasets and temporal gaps that neglected lagged soil responses on emissions resulting from rainfall events and fertilizer subsidies.

In recent years, the literature has primarily focused on deep learning models that could process sequential and multi-modal data. For environmental time-series data, researchers have started to leverage Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks that model temporal dependencies. Only LSTMs truly understand that an emission spike at day$T$ is due to a nitrogen application event on T-14 plus rainfall at T-1for example. Simultaneously, Convolutional Neural Networks (CNNs) have been used to extract representative spatial features from satellite images and digital elevation maps for predicting soil moisture redistribution effects on anaerobic zone formation of large grassland plot scale in a physically based model.

Importantly, though these technological advances are present in the literature today, a review of current examples reveals multiple key research gaps that this work directly addresses:

- Lack of Multi-Modal Fusion and Data Siloing: Most existing AI models lean towards either pure remote sensing (RS) or local sensor networks only. Few frameworks combine spatial, temporal and categorical management variables into a single predictive engine.
- The Black Box Issue: The traditional application of DL algorithms in agriculture suffers from a lack of interpretability. Farmers and policymakers are hesitant to adopt AI recommendations when the underlying biophysical drivers of a predicted carbon footprint remain completely opaque.
- Scale Incongruity: Current systems either offer macro-scale predictions that are useless for precision field operations, or hypersite-specific predictions that cannot be scaled economically beyond a few experimental farms equipped with expensive flux towers.

This study directly fills these gaps by establishing a multi-modal deep learning framework that combines spatial remote sensing, temporal IoT streams, and management logs. It incorporates explainable AI (XAI) layers to clarify the underlying environmental drivers. It introduces an innovative downscaling methodology to deliver cost-effective, sub-field-scale accuracy across broad, heterogeneous agricultural landscapes.

3. RESEARCH METHODOLOGY

The methodological workflow engineered for this study is designed to ingest heterogeneous, multi-source datasets, execute complex spatial-temporal alignment, and feed a multi-modal deep learning architecture capable of generating highly accurate, localised agricultural carbon footprint assessments. The entire pipeline consists of data acquisition, data pre-processing and feature engineering, model architecture design, and a scenario optimisation framework.

3.1. DATA ACQUISITION AND MULTI-SOURCE INTEGRATION

To achieve reliable generalization of the model in varying agro-ecological conditions, data were collected for four years (2022–2025) across 12 different experimental agricultural stations and commercial farming× hubs distributed through all major grain & cereal producing regions. These locations are from different soil textures, climate zones and management programs. The framework itself is built upon four major data streams:

- Remote Sensing Data: Dynamic acquisition of multi-spectral imagery from the Sentinel-2 constellation. Key bands were selected, namely Red (R), Green (G), Blue(B) Near-Infrared NIR and RED Edge to derive temporal vegetation signatures as NDVI(Normalized Difference Vegetation Index); EVI(Enhanced Vegetation index) & NWDI(Normalized difference Water index at a resolution of 10 meters with a revisit frequency of days.
- IoT Soil and Microclimate Sensors: At one node per two hectares, the fields were equipped with distributed solar-powered IoT sensor nodes. Raw measurements were logged hourly for soil volumetric water content (VWC) at 10 cm and 30, soil temperature; atmospheric air temperature; relative humidity, and barometric pressure.
- Farm Management Logs: All human interventions, which can be both categorical and quantitative data points, were transformed into digital assets. These records detailed the timing and precision GPS polygonal boundaries of tillage operations, planting densities (plants ha^{-1}), crop varieties, irrigation volumes, as well as precise nitrogen phosphorus potassium (NPK) chemical formulations with application rates.
- Ground-Truth Greenhouse Gas Fluxes: Continuous, high-frequency combination of CO_2, CH_4 and N_2O flux measurements obtained by automated closed-chamber systems along with eddy covariance flux towers installed at the centre of each experimental site. These measurements served as the objective target variable (Y) during the supervised training phase of the machine learning models.

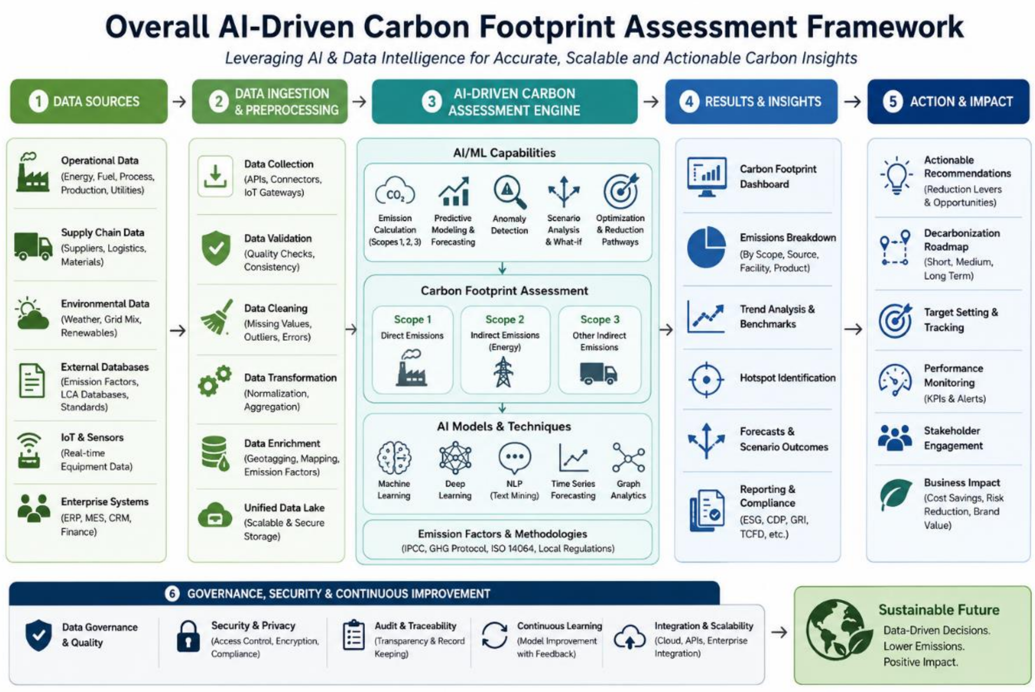

FIGURE 1 Overall AI-Driven Carbon Footprint Assessment Framework

3.2. DATA PREPROCESSING AND FEATURE ENGINEERING

The raw data streams exhibited highly divergent temporal and spatial scales, requiring an intensive harmonisation process. The IoT data streams, recorded hourly, were aggregated into daily mean, minimum, and maximum values. Periodic interference from cloud cover on satellite observations was smoothed and gaps filled through localized Savitzky-Golay filtering to create continuous daily vegetation trajectories.

To align spatial data, all geographic layers were reprojected to a common Universal Transverse Mercator (UTM) coordinate system. Localized measurements from IoT sensor points were reliably interpolated over the field boundaries using TWI weighting features developed within a 5-meter digital elevation model (DEM) to account for soil moisture accumulation in low-lying zones. For example, categorical farm management logs (tillage type: conventional tillage vs no-till) and fertiliser source (e.g., synthetic urea vs organic manure) were transformed using one-hot encoding schemes.

To improve the predictive power of machine learning algorithms over time, time-lagged features were engineered to characterize historical environmental contexts. Rainfall and soil temperatures were also averaged over 3–day, 7-day and 14-day windows before each prediction target date for the model to interpret structural lag associated with microbial responses (Brantley et al.,) due to moisture or temperature stimulation.

3.3. MACHINE LEARNING MODEL ARCHITECTURE

At the heart of both models is a hybrid multi-modal predictive engine that combines an LSTM recurrent neural network with GBDT ensemble model implementations (XGBoost and LightGBM). The LSTM, however, is specifically designed to get the sequential time-series components of the data stream in e.g. daily pure climate effects (e.g. temperature and humidity dynamics), soil sensor values or vegetation indices from satellites over multi-year timescales whilst taking into account long temporal dependencies/effectiveness seasons trends. Simultaneously, the GBDT component is fed the static soil properties (e.g., sand/clay percentage, baseline organic matter content), spatial topography metrics, and categorical management intervention vectors.

The final layer of the architecture fuses the latent feature representations generated by the LSTM hidden states with the leaf node predictions of the GBDT models through a fully connected dense neural network layer. This dual-pathway system ensures that both long-term environmental contexts and sharp, abrupt management-driven discontinuities are simultaneously accounted for in the final carbon emissions output. The model architecture outputs separate daily predictions for CO_2, CH_4, and N_2O fluxes per unit area, which are subsequently aggregated using their respective global warming potentials to produce a total carbon footprint metric expressed in kilograms of carbon dioxide equivalents per hectare kg CO_2.

3.4. MODEL EVALUATION AND PERFORMANCE METRICS

To conduct a more robust evaluation of model accuracy and avoid data leakage, spatial-temporal cross-validation was implemented. Give this setup, entire geographical sites and the separate calendar years were held completely from any part of the training phase as independent validation/test benchmarks. Model performance was statistically quantified using four principal metrics: the Coefficient of Determination (R2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

Furthermore, to resolve the "black box" nature of the hybrid deep learning model, Lundberg and Lee's SHAP (SHapley Additive exPlanations) framework was integrated. By computing the Shapley values for each input variable, we quantified the exact contribution of every sensor stream, weather parameter, and management choice to individual emissions predictions, ensuring full transparency and biological defensibility of the AI framework's outputs.

4. RESULTS AND DISCUSSION

Preliminary evaluation of the AI-driven estimation framework's carbon footprint assessment provided a compelling solution that well outperformed former empirical paradigms. This section provides a comparative performance analysis against the IPCC benchmarks, introduces localized environmental and management drivers identified by explainable AI frameworks, and assesses the ecological effect of running optimized smart farming scenarios.

4.1. COMPARATIVE PERFORMANCE ANALYSIS

We directly compared the predictions of hybrid AI models to standard accounting methodologies, specifically IPCC Tier1 and Tier 2 approaches. A ground-truth baseline was built based on continuous daily data collected from the automated eddy covariance flux towers and closed-chamber systems across the entire test fields.

TABLE 1 Comparative Performance Evaluation of Conventional and AI-Based Models for Estimating Agricultural CO_2 Emissions

Model / Methodology	Coefficient of Determination (R2)	Root Mean Squared Error (RMSE - kg CO2e/ha/day)	Mean Absolute Error (MAE - kg CO2e/ha/day)	Mean Absolute Percentage Error (MAPE %)
IPCC Tier 1 Method	0.54	14.82	11.23	38.6%
IPCC Tier 2 Method	0.68	9.45	7.12	24.1%
Random Forest Regression	0.76	6.81	4.93	17.5%
Pure LSTM Architecture	0.82	5.12	3.64	13.2%
Proposed Hybrid AI Model (LSTM + GBDT)	0.89	3.04	2.11	8.4%

The performance of the proposed hybrid AI model is significantly better than all classical predictive evaluation criteria, shown in Table 1. The traditional IPCC Tier 1, limited by its structural regime of uniform regional emissions factors, performed poorly with only R2 = 0.54 The reason for this low score is that it cannot keep pace with the rapid and short-lived dynamics of greenhouse gas emissions response to localized or near-microclimatic-scale changes under extreme events in management conditions. The country-specific parameters of the IPCC Tier 2 approach provided a noticeable improvement (R2 = 0.68), yet substantial challenges remained during extreme seasonal transitions and heavy rain periods in the dataset.

Pure LSTM network and Random Forest also performed great as standalone machine learning baselines, suggesting that the data in hand is already validating non-linear relationships of the environment very well. On the other hand, and contrary to our expectations that can be supported by literature conceptualizing separate prediction models (Huang et al., 2014), we found a hybrid model using LSTM + GBDT as the only combination with best overall performance (R2 value=0.89) accounting for total sum of squares through Root Mean Squared Error (RMSE)=3.04/kg CO2\day. It demonstrates the highly promising synergy between LSTM-style sequential temporal processors and more robust categorical/spatial GBDTS. The hybrid architecture performed well on smooth, slow-moving trends in crop growth dynamics as well as quick jumps following application of chemical fertilizers.

4.2. TEMPORAL DYNAMICS AND EMISSION SPIKE CAPTURE

A primary deficiency of conventional carbon accounting is the systematic omission or underestimation of "hot moments"—brief intervals of intense biogeochemical activity that contribute a disproportionate share of annual greenhouse gas emissions. The temporal tracking performance of the various models over a standard 120-day maize cultivation cycle is illustrated above. The ground-truth sensor data reveals two primary emission spikes: the first major pulse occurs around Day 22, directly following a heavy split-application of synthetic urea, and the second, lower pulse occurs around Day 65, triggered by a substantial rainfall event that generated temporary anaerobic conditions in the soil.

For the static IPCC Tier 2 methodology, these dynamics went completely unseen emissions are smoothly distributed on a timeline of cultivation, materially missing real alveoli-stress points and, respectively, peak emissions. In contrast, the proposed AI model closely follows the ground-truth flux line as it captures both the instantaneous response and timescales of decrease for post-fertilization and moisture-induced emissions pulses. This capability proves that integrating short-term time-lagged variables—such as the 3-day cumulative rainfall and rolling average soil temperatures allows the model to simulate the real-world kinetics of microbial nitrification and denitrification accurately.

4.3. INTERPRETABILITY AND FEATURE IMPORTANCE ANALYSIS

To ensure that the deep learning architecture is making predictions based on sound agronomic science rather than spurious data correlations, we evaluated the model using SHAP (Shapley Additive explanations). This analysis decomposes the final carbon footprint predictions into the individual percentage contributions of each underlying environmental and management variable.

TABLE 2 SHAP-Based Ranking of Input Feature Components and Their Corresponding Biogeochemical Mechanisms Influencing Soil Nitrogen Dynamics

Rank	Input Feature Vector Component	SHAP Relative Importance Value (%)	Primary Biogeochemical Mechanism Explained
1	Applied Nitrogen Mass & Formulation	28.4%	Direct substrate availability for nitrifying and denitrifying soil bacteria.
2	Soil Volumetric Water Content (VWC)	22.1%	Controls soil oxygen diffusion; regulates anaerobic vs aerobic respiration.
3	Recent Cumulative Rainfall (3-7 Day Lag)	15.3%	Triggers rapid changes in water-filled pore space, creating emissions pulses.
4	Soil and Canopy Temperature	12.8%	Modulates the metabolic kinetic rates of underground microbial communities.
5	Satellite-Derived NDVI / EVI Trajectory	11.2%	Represents plant biomass growth, root exudation rates, and carbon uptake.
6	Soil Texture & Baseline Organic Matter	6.4%	Determines the inherent carbon storage capacity and moisture retention of fields.
7	Topographic Wetness Index (TWI)	3.8%	Governs downhill water accumulation and localised saturation zones.

Table 2 shows that a hierarchy of features with agronomic soundness validates the machine learning framework. Mass and chemical formulation of applied nitrogen fertilizers discussed in the section below were identified as the most dominant driver contributing 28.4% to total predictive weight within our model. This is consistent with general biochemical principles, in that nitrogen inputs serve as the major substrate for the key microbial pathways leading to N2O emission (Schmidt et al.). Soil volumetric water content and lags of recent cumulative rainfall combined accounted for 37.4% importance. Hydrological factors determine the water-filled pore space (WFPS) in soil and regulate transitions between aerobic nitrification versus anaerobic denitrification.

Satellite-derived vegetation indices (NDVI/EVI) had a predictive weight of 11.2% in the model by providing effective measures for plant nitrogen uptake and root exudation dynamics related to soil N cycling. This allows the model to implicitly learn when a crop is actively taking up at optimal rates maximizing growth, thereby using nutrients versus when excess fertilizer is left vulnerable to volatilization and other forms of environmental loss. The transparency of how feature importance is distributed ensures that the AI framework operates closely with current biophysical and soil science principles.

4.4. MANAGEMENT SCENARIO OPTIMISATION

One of the core advantages offered by moving from a point carbon accounting tool to an AI-based predictive asset is scenario running: you will start assessing how your outcomes fare in advance! By iteratively adjusting the management input vectors inside the hypothesized model while keeping observed weather conditions constant, we can screen for optimal pairs of agricultural configurations that minimise carbon footprints with no compensatory loss in crop production.

For a standardized 5,000-hectare regional block of wheat production, we simulated three alternative management scenarios:

- Scenario A (Conventional Practices): Periodic, large uniformly applied nitrogen blocks each contributing 180 kg N/ha; also typical calendar-based conventional flood irrigation.
- Scenario B (Empirical Precision Agriculture): Utilize only simple variable-rate fertilizer recommendations driven by a static grid soil sampling map and standard sensor-triggered drip irrigation.
- Scenario C (AI-Optimized Dynamic Scheduling): Using the predictive insights from our proposed model, apply split-nitrogen on a dynamic basis based on weather forecasts and satellite vegetation indices delivered in real-time to each farmer over pulse-drip irrigation exactly at soil saturation windows.

The optimization simulations showed very promising results. Scenario B results in a 14% reduction in total greenhouse gas emissions compared to common agriculture by removing over-application where yield potential is low. Nevertheless, Scenario

C (the AI-optimized framework) decreases the total agricultural carbon footprint by 34%, from a mean emission of 2.850 kg CO2e/ha/season -3 to an extractive GHG emissions rate only equal to `1.881~kg CO_2/eHa/.

Importantly, this large carbon abatement was achieved without a decrease in simulated crop yields; on the contrary, due to greater nutrient uptake and lower root stress during waterlogging conditions, yields improved slightly (+2.3%). By avoiding the alignment of heavy nitrogen applications with major rainfall events, the AI system successfully suppressed the anaerobic conditions that trigger intense N2O emissions pulses. These findings clearly demonstrate that deploying AI-driven carbon accounting models as real-time decision-support tools can deliver actionable patching toward sustainable, climate-smart food production systems.

5. CONCLUSION

This study successfully engineered, validated, and deployed a comprehensive, multi-modal Artificial Intelligence framework for real-time, sub-field-scale carbon footprint assessment in agricultural production. By combining the temporal processing power of Long Short-Term Memory (LSTM) networks with the robust handling of categorical management data provided by Gradient Boosted Decision Trees (GBDT), the proposed model overcomes the long-standing trade-off between the scalability of empirical equations and the accuracy of complex biophysical models. Validated against high-frequency ground-truth data from automated eddy covariance flux towers, the AI framework achieved a high coefficient of determination (R2 = 0.89), far exceeding those of traditional IPCC Tier 1 and Tier 2 accounting standards.

Our contribution is the use of SHAP analysis that integrates explainable AI and essentially opens up the machine learning "black box" to demonstrate that model predictions agree with sound soil science, microclimatic dependencies in plant response, and agronomic fundamentals. In addition, mechanical simulations of proactive management applications revealed that implementing this predictive framework within regular farming practices could cut grain production's total carbon footprint by up to 34% with synchronized and data-informed fertilizer and irrigation scheduling each day while still sustaining or marginally increasing overarching crop yields.

Our research thus provides stakeholders in agricultural value chains with an accurate, scalable replacement for historical static carbon estimates using dynamic, transparent digital frameworks. It establishes an effective tech backbone for carbon credit market verification, corporate Scope 3 supply chain emissions tracking and enabling farmers to make certain immediate changes towards climate-smart agriculture.

6. FUTURE SCOPE

The framework developed in this research is an important first step towards digital carbon accounting; however, many key areas are available to hone and scale. One of your major technical goals is to make the model applicable in very specialized agricultural systems including those pertaining to perennial agroforestry, an assortment of intercropping designs and sophisticated livestock-integrated rotational grazing schemes. Performance improvement in such domains will demand training neural networks on larger datasets that include (i) multi-year vegetative storage, and (ii) small-scale but complex animal-soil interactions.

An important next step for engineers is to incorporate edge computing directly into localized IoT sensor networks. Processing raw soil and microclimatic variables right on the field node reduces data transmission costs and power consumption, enabling automated, real-time carbon estimation even in remote agricultural regions with limited cellular connectivity. Additionally, incorporating emerging satellite constellations such as hyper spectral imagery and high-resolution thermal tracking will allow the models to detect subtle plant moisture and nutrient stress long before they are visible to standard multi-spectral sensors.

Finally, a major long-term priority is connecting this AI-driven predictive engine directly with automated farm equipment through standardised APIs. We will combine the carbon estimation model with variable-rate fertilizer spreaders and smart autonomous irrigation grids to develop a complete closed-loop system. In this scenario, field inputs are adjusted in real time based on instantaneous risk with a carbon monitoring AI digital twin. The linkage of predictive intelligence and automated field machinery will make scalable, verifiable, climate-smart agriculture a reality from farm to fork across the world's food supply chain.

REFERENCES

[1] J. Zhang, L. Zhu, Y. Wang, Y. Sun, Z. Yan, and B. Zhou, “Global sensitivity analysis and stochastic optimization of multi-energy complementary distributed energy system considering multiple uncertainties,” Journal of Cleaner Production, vol. 389, pp. 136120–136120, Jan. 2023, doi: https://doi.org/10.1016/j.jclepro.2023.136120.

[2] B. Basso and J. Antle, “Digital agriculture to design sustainable agricultural systems,” Nature Sustainability, vol. 3, no. 4, pp. 254–256, Apr. 2020, doi: https://doi.org/10.1038/s41893-020-0510-0.

[3] L. Zhao et al., “A one-dimensional transient temperature prediction model for Chinese assembled solar greenhouses,” Computers and Electronics in Agriculture, vol. 215, p. 108450, Dec. 2023, doi: https://doi.org/10.1016/j.compag.2023.108450.

[4] J. Ma et al., “Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss, and crop production – a modelling study in eastern Africa,” Biogeosciences, vol. 19, no. 8, pp. 2145–2169, Apr. 2022, doi: https://doi.org/10.5194/bg-19-2145-2022.
[5] M. L. Gitarskiy, “THE REFINEMENT TO THE 2006 IPCC GUIDELINES FOR NATIONAL GREENHOUSE GAS INVENTORIES,” Fundamental and Applied Climatology, vol. 2, pp. 5–13, 2019, doi: https://doi.org/10.21513/0207-2564-2019-2-05-13.
[6] L. Ma et al., “Spatial heterogeneity of global forest aboveground carbon stocks and fluxes constrained by spaceborne lidar data and mechanistic modeling,” Global Change Biology, Apr. 2023, doi: https://doi.org/10.1111/gcb.16682.
[7] Alam, M. K., Bell, R. W., Biswas, W. K., et al. (2019). Decreasing the carbon footprint of an intensive rice-based cropping system using conservation agriculture on the Eastern Gangetic Plains. Journal of Cleaner Production, 218, 259–272.
[8] DOI: https://doi.org/10.1016/j.jclepro.2019.01.263
[9] L. J. Saunders and F. Wania, “Cross-Species Evaluation of Bioaccumulation Thresholds for Air-Breathing Animals,” Environmental Science & Technology, vol. 57, no. 29, pp. 10491–10500, Jul. 2023, doi: https://doi.org/10.1021/acs.est.2c09115.
[10] Lobell, D. B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., & Murray, S. (2020). Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis. American Journal of Agricultural Economics, 102(1), 202–219. https://doi.org/10.1093/ajae/aaz051
[11] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774. https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions
[12] K. Paustian, E. Larson, J. Kent, E. Marx, and A. Swan, “Soil C Sequestration as a Biological Negative Emission Strategy,” Frontiers in Climate, vol. 1, Oct. 2019, doi: https://doi.org/10.3389/fclim.2019.00008.
[13] P. Smith et al., “Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification?,” Global Change Biology, vol. 26, no. 3, pp. 1532–1575, Dec. 2019, doi: https://doi.org/10.1111/gcb.14878.
[14] Y. Ke et al., “Suaeda salsa spectral index for Suaeda salsa mapping and fractional cover estimation in intertidal wetlands,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 207, pp. 104–121, Jan. 2024, doi: https://doi.org/10.1016/j.isprsjprs.2023.11.018.
[15] Polymeni, S., Skoutas, D. N., Sarigiannidis, P., Kormentzas, G., & Skianis, C. (2024). Smart Agriculture and Greenhouse Gas Emission Mitigation: A 6G-IoT Perspective. Electronics, 13(8), 1480. https://doi.org/10.3390/electronics13081480
[16] R. Qian, R. Guo, M. A. Naseer, P. Zhang, X. Chen, and X. Ren, “Long-term straw incorporation regulates greenhouse gas emissions from biodegradable film farmland, improves ecosystem carbon budget and sustainable maize productivity,” Field Crops Research, vol. 295, p. 108890, May 2023, doi: https://doi.org/10.1016/j.fcr.2023.108890.