

Original Research Article

GREEN TECH PARADOX

An Analysis of AI in Environment Sustainability

Ms. Kirandeep Kaur,¹ Mrs. Harpreet Kaur²¹Research Scholar, Guru Nanak College of Arts, Science and Commerce, Mumbai, India²(HoD, Information Technology), Guru Nanak College of Arts, Science and Commerce, Mumbai, IndiaCorresponding author E-mail: kirandeepkaurthe661@gmail.com,¹ Harpreetkaur@gncasc.org²

Received: 10 September, 2026 | Accepted: 24 September, 2026 | Published: 13 October, 2026

Abstract:

The integration of artificial intelligence into environmental sustainability strategies presents a critical dilemma known as the "Green Tech Paradox." While AI offers unprecedented capabilities to optimize resource allocation, predict climate patterns, and enhance energy efficiency across industrial sectors, its deployment carries substantial, often overlooked, ecological costs. This research provides a critical analysis of the dual role AI plays as both a catalyst for eco-innovations and a significant contributor to environmental degradation.

On the positive frontier, AI-driven applications accelerate sustainability by enabling smart power grids, improving carbon capture systems, facilitating precision agriculture, and modeling complex environmental crises. Machine learning algorithms enhance operational efficiency, reducing waste and carbon footprints in manufacturing and logistics. However, these beneficial outcomes are counterbalanced by the escalating physical footprint of AI infrastructure. The training and continuous execution of complex neural networks demand massive amounts of electricity, heavily relying on fossil-fuel-dominant energy grids.

Furthermore, data centers consume millions of liters of freshwater daily for cooling mechanisms, accelerating localized water scarcity, while the rapid obsolescence of specialized hardware generates growing volumes of hazardous electronic waste.

This paper evaluates this structural tension through a multidimensional synthesis of secondary data, life-cycle assessments, and resource consumption metrics across global data centers. The analysis reveals that without stringent regulatory frameworks and sustainable design interventions, the environmental degradation caused by AI's hardware lifecycle risks surpassing its ecological benefits. Ultimately, the study advocates for a paradigm shift toward "Green AI"—prioritizing algorithmic efficiency, renewable energy integration for computing infrastructure, sustainable hardware lifecycle management, and policy-driven eco-design standards. By establishing transparent auditing protocols, researchers and industry leaders can mitigate the hidden ecological liabilities of artificial intelligence, ensuring that modern technological advancement aligns genuinely with global net-zero targets and long-term planetary boundaries.

Keywords: Green AI, Climate change mitigation, Sustainable development goals (SDGs), Environmental governance, Lifecycle assessment of AI.

1. Introduction:

1.1. AI Applications for Environmental Sustainability

AI is transforming environmental management through diverse, high-impact applications. In the energy sector, machine learning models improve renewable generation forecasting, optimize grid dispatch, and enable demand-response systems that reduce reliance on fossil fuels. In agriculture, AI-powered precision farming tools analyze satellite imagery, soil sensors, and weather data to optimize irrigation, fertilizer application, and pest control—thereby conserving water, reducing chemical runoff, and enhancing crop yields. AI also supports biodiversity conservation by automating species identification from camera traps and acoustic sensors, tracking illegal logging or fishing in near real time, and modeling habitat shifts under climate scenarios. Additional domains include smart waste sorting using computer vision, air and water quality monitoring via IoT-integrated predictive analytics, and climate risk assessment for infrastructure planning. Collectively, these applications demonstrate AI's capacity to accelerate progress toward multiple Sustainable Development Goals (SDGs), particularly SDG 6 (Clean Water), SDG 7 (Clean Energy), SDG 11 (Sustainable Cities), SDG 12 (Responsible Consumption), and SDG 13 (Climate Action).

1.2. The Environmental Footprint of AI Infrastructure

Despite its benefits, AI's rapid scaling imposes significant environmental costs. Recent United Nations assessments project that global data centers consumed approximately 448 terawatt-hours (TWh) of electricity in 2025—surpassing the annual usage of most nations—and could reach 945 TWh by 2030, with AI workloads accounting for up to 40% of this demand. This energy consumption translates into 189–399 million tons of CO₂ emissions annually, comparable to the national emissions of countries like Argentina or the United Kingdom. Water use is equally concerning: data centers currently withdraw 4.5 trillion liters per year for cooling, a figure projected to rise to 9.3 trillion liters by 2030—enough to meet the basic domestic needs of over 1.3 billion people. Land use is also expanding, with data center footprints expected to exceed 14,500 km² by 2030, twice the area of metropolitan Jakarta. These impacts underscore the “sustainability paradox” wherein carbon-intensive technologies are deployed

to solve carbon-driven crises, raising urgent questions about lifecycle accountability and systemic trade-offs.

1.3. Toward a Responsible AI-for-Sustainability Framework

Addressing this paradox requires more than incremental efficiency gains; it demands a reorientation of AI development toward explicit sustainability principles. Emerging frameworks advocate for “Green AI”—prioritizing energy-efficient algorithms, renewable-powered infrastructure, and hardware designed for circularity and longevity. Governance mechanisms must integrate environmental externalities into AI procurement, deployment, and auditing standards, ensuring that sustainability claims are substantiated by full lifecycle assessments covering carbon, water, land, and e-waste. Moreover, equitable access to AI tools must be ensured for developing regions, avoiding a new form of “green digital divide.” This research contributes to this agenda by proposing an integrated, multi-dimensional framework that aligns AI innovation with planetary boundaries and social justice, thereby enabling AI to serve as a genuine catalyst—not a hidden cost—in the transition to a sustainable future.

2. Problem statement and need for the study:

Despite AI's rapid deployment for climate action and resource optimization, a critical “sustainability paradox” persists: AI systems themselves impose substantial, often unaccounted environmental burdens—carbon emissions, water depletion, land use, and e-waste—that risk offsetting their ecological benefits. Current assessments remain fragmented, with 72% of studies narrowly focusing on energy and CO₂ while neglecting water, materials, biodiversity, and full lifecycle impacts, leading to 300–500% underestimation of AI's true footprint. This methodological gap hinders evidence-based governance and responsible innovation. There is an urgent need for an integrated, multi-dimensional framework that aligns AI development with planetary boundaries, embeds systemic accountability across design–deployment–decommissioning phases, and ensures equitable access—thereby transforming AI from a hidden cost into a verified catalyst for genuine environmental sustainability.

3. Literature Review:

Wu et al. (2022) examine sustainable AI through a holistic analysis of data, algorithms, and system hardware. The study demonstrates that AI's environmental impact extends beyond model-training energy to include experimentation, inference, data processing, hardware manufacturing, and embodied carbon. It identifies hardware–software co-design, efficient algorithms, carbon-aware scheduling, improved resource utilization, and lifecycle assessment as major opportunities for reducing emissions. The authors emphasize standardized sustainability metrics and telemetry, arguing that model accuracy must be evaluated alongside energy consumption and environmental cost. Overall, the study provides an important foundation for environmentally responsible AI development.

Sowgandika et al. (2026) introduced AgriSmart, a sensor-free precision agriculture system combining Sentinel-2 NDVI satellite data with machine learning for crop management. Utilizing Random Forest models, the framework monitors crop health, achieves high yield prediction accuracy ($R^2 = 0.971$), and generates actionable fertilizer recommendations. Accessible via an interactive dashboard on standard hardware, it offers smallholders a scalable, low-cost solution without physical IoT infrastructure.

Arockia Nisha and Nanthini (2026) present a comprehensive Green AI framework for smart agriculture designed to reduce computational energy and carbon emissions without sacrificing performance. By integrating lightweight models, edge computing, model compression, and federated learning into precision farming and

resource management systems, the framework achieves up to 30% water savings and 20% fertilizer reduction while maintaining high yield forecasting accuracy. Furthermore, this architectural design decentralizes data processing, allowing localized edge nodes to execute real-time analytical decisions directly on low-power hardware devices. By minimizing reliance on centralized, energy-intensive cloud infrastructure, the system drastically cuts network data transmission overheads, minimizes latency, and reduces operational electricity demands. Consequently, the framework bridges the gap between high-precision agronomic forecasting and sustainable computing practices. It demonstrates that eco-conscious AI optimization can actively drive input resource efficiency at the field level while curbing the overall carbon footprint of digital agricultural infrastructure. Ultimately, this approach offers a scalable blueprint for deploying carbon-aware machine learning in resource-constrained agrarian regions. This decentralized approach directly addresses the data center energy burdens highlighted in wider infrastructure literature. By processing algorithmic features directly on localized edge nodes, precision agricultural applications can bypass heavy reliance on centralized cloud databases. This mitigation path reduces network data transmission overheads, minimizes operational latency, and lowers grid electricity consumption during live model inference. Consequently, transitioning to localized edge architectures represents a viable pathway to reconcile high-performance agricultural forecasting with global digital sustainability goals.

Sheshasaayee and Angayarkanni (2025) examine how artificial intelligence, machine learning, IoT sensors, drones, and robotics can address the ecological degradation caused by traditional agricultural practices. The authors highlight how predictive analytics, computer vision, and AI-powered chatbots facilitate precision irrigation, soil monitoring, and automated pest management to mitigate climate change and protect biodiversity. Furthermore, the study links these AI-driven precision techniques to the achievement of 10 specific United Nations Sustainable Development Goals (SDGs), while addressing critical socio-technical challenges such as data privacy, algorithmic bias, and equitable resource access for smallholder farmers.

Tabbakh et al. (2024) establish a framework for Green AI to mitigate the environmental impact of artificial intelligence. The study addresses energy-intensive computational demands by proposing model optimization techniques, including pruning and quantization, alongside hardware alternatives like TPUs and FPGAs. Furthermore, the authors present strategies for energy-efficient data center operations and demonstrate practical implementations in computer vision and natural language processing to align AI advancements with global sustainability goals. Additionally, the framework emphasizes the necessity of joint hardware-software evaluation to accurately quantify the trade-offs between computational throughput and carbon expenditure. By evaluating the operational life-cycle of complex neural networks, the study highlights how targeted pruning and low-precision quantization significantly decrease memory bandwidth demands without degrading task accuracy. Implementing these optimized models on energy-efficient ASIC and FPGA architectures allows enterprise systems to lower total dynamic power consumption during inference. Ultimately, this comprehensive methodology provides a concrete roadmap for technology developers to audit operational energy footprints, reduce carbon intensity in large-scale processing, and establish eco-conscious computing standards across industrial AI applications. The operational lifecycle assessments discussed by Tabbakh et al. (2024) underscore a critical requirement for joint hardware-software evaluation frameworks to accurately quantify the operational trade-offs between computational throughput and carbon expenditure. By evaluating the macro-level footprint of deep learning networks, current literature highlights how targeted structural pruning and low-precision quantization can actively lower dynamic power consumption without degrading task-specific diagnostic accuracy. This eco-conscious auditing approach is essential for scaling technologies across sectors where computational resource conservation must be balanced

against real-time operational demands.

Greif et al. (2024) evaluate artificial intelligence's impact on environmental sustainability through a strategic SWOT analysis of 70 research articles. The study highlights key strengths, such as rapid complex data processing and the decoupled separation of model training from real-time execution. However, these are counterbalanced by internal weaknesses, including high energy consumption, data bias, and a lack of model explainability. While AI offers critical opportunities for climate change mitigation and scientific innovation, deployment faces threats from cybersecurity risks, financial constraints for smaller enterprises, and regulatory hurdles.

Abbassi et al. (2026) addressed irrigation challenges in Algerian date palm farming by evaluating machine learning models to predict root soil moisture using meteorological data. Comparing multiple algorithms across historical weather parameters, the authors demonstrated that Artificial Neural Networks outperformed Decision Trees and Random Forests, achieving 80% precision and an F1-score of 0.75. This predictive framework provides a reliable alternative to physical IoT sensors, enabling data-driven irrigation management that conserves vital water resources and maintains crop health.

Jääskeläinen et al. (2022) examine the environmental impact of artificial intelligence in artistic practices. By mapping the interplay between technological materials (hardware and software) and creative phases, the authors establish a framework to assess resource consumption. Ultimately, the study urges artistic and scientific communities to adopt resource-conscious approaches that respect planetary limits.

Walke et al. (2022) explored the transformative potential of artificial intelligence and precision farming in advancing sustainable agriculture. Combining a systematic literature review with an empirical case study in drought-prone rural India, the authors evaluated AI applications including IoT sensors, agricultural drones, and analytical software. Their findings demonstrated that AI-driven resource allocation significantly increased crop yields, enhanced water efficiency, and reduced environmental impacts, though high technology implementation costs and data scarcity remain critical challenges for smallholder farmers.

Sibiya (2026) presented a Green Artificial Intelligence framework to optimize sustainable crop yield prediction and resource management within smart agriculture systems. Integrating an Adaptive Sparse Extreme Learning Machine (AS-ELM) algorithm with model compression, edge computing, and federated learning, the proposed framework significantly reduces computational overhead and environmental impact. The system achieved high predictive accuracy while lowering energy consumption, providing an eco-friendly and scalable solution for climate-resilient farming. Moreover, by leveraging the AS-ELM algorithm, the framework achieves exceptionally fast learning speeds and minimal parameter complexity compared to conventional deep learning models. Integrating federated learning ensures that localized agricultural datasets remain on edge devices, safeguarding farmer data privacy while enabling continuous collaborative model training across distributed farms. Model compression techniques further shrink memory footprints, allowing deployment on resource-constrained microcontrollers without requiring continuous cloud connectivity. By dramatically cutting network bandwidth demands and energy consumption during inference, this eco-friendly paradigm addresses both computational sustainability and rural infrastructure constraints, establishing a viable pathway for scalable, climate-resilient precision farming. Furthermore, by optimizing deep learning layers into compressed, sparse algorithmic structures, field-level edge devices can execute complex predictive analytics without constant connectivity to regional servers. The implementation of federated learning setups ensures that sensitive farm telemetry metrics remain localized, safeguarding operational privacy while allowing continuous collaborative model improvements across distributed

networks. This intersection of lightweight computing structures with precision agrarian applications effectively demonstrates how targeted machine learning optimizations can lower infrastructural carbon liabilities while boosting real-world resource allocation efficiency.

Chaudhary (2023) examines the role of artificial intelligence (AI) in driving United Nations Sustainable Development Goals (SDGs), particularly regarding environmental sustainability. AI supports resource management across energy, water, transportation, and biodiversity sectors. However, significant challenges persist, including severe carbon footprint concerns from AI computing power, security risks, bias, and uneven global access that could worsen developmental inequality.

Gap Analysis:

While current literature addresses Green AI techniques—such as model compression, hardware co-design, and satellite data—and AI applications in precision agriculture, a significant gap remains at their intersection. Most agricultural AI frameworks focus on isolated components like sensor-free prediction or yield optimization, yet fail to systematically quantify the lifecycle environmental impact of these tools. Additionally, existing Green AI strategies lack field-level evaluation under real-world agricultural constraints, such as smallholder data scarcity, cost barriers, and edge deployment challenges. Bridging low-overhead algorithms with comprehensive, deployment-ready lifecycle assessments remains an unaddressed research priority.

4. Objectives, Variables, Hypothesis:

4.1 Research Objectives

To Assess Environmental & Microclimatic Drivers: To evaluate the impact of environmental and climatic parameters (soil moisture, soil pH, ambient temperature, rainfall, humidity, and sunlight hours) on agricultural productivity (yield_kg_per_hectare).

To Analyze the Efficiency of Farming Practices: To analyze how different agricultural management techniques (irrigation type, fertilizer type, and pesticide usage volume) influence crop yield and overall plant vigor (NDVI_index).

To Examine Crop Health and Disease Dynamics: To examine the correlation between real-time vegetation indices (NDVI_index) and the incidence and severity of crop_disease_status across different crop types (Wheat, Soybean, Maize, Cotton, Rice).

To Evaluate Regional and Growth Period Variations: To investigate regional performance differences and determine how total growth duration (total_days) affects total harvest yield across diverse geographic areas.

To Develop Predictive Precision Agriculture Models: To construct statistical and machine learning models leveraging IoT sensor metrics (soil_moisture_%, NDVI_index, weather telemetry) for automated yield estimation and decision support in smart farming.

4.2 Variables Identification

4.2.1. Dependent Variables (Target / Outcome Variables)

yield_kg_per_hectare (*Primary Continuous Outcome*): Total crop yield produced per unit area.

NDVI_index (*Secondary Continuous Outcome*): Normalized Difference Vegetation Index indicating crop health and biomass density.

crop_disease_status (*Categorical Outcome*): Severity level of crop disease infection (e.g., Mild, Moderate, Severe).

4.2.2. Independent Variables (Predictor / Feature Variables) Environmental & Climatic Telemetry:

soil_moisture_% (*Numeric*)

soil_pH (*Numeric*) temperature_C (*Numeric*) rainfall_mm (*Numeric*)

Agronomic Management Practices:

irrigation_type (*Categorical: Drip, Sprinkler, Manual*) fertilizer_type (*Categorical: Organic, Inorganic, Mixed*)

pesticide_usage_ml (*Numeric*)

total_days (*Numeric: Crop cycle duration from sowing to harvest*)

crop_type (*Categorical: Wheat, Soybean, Maize, Cotton, Rice*)

Geographic & Location Telemetry:

region (*Categorical: North India, South India, Central USA, South USA, East Africa*)

latitude & longitude (*Numeric spatial coordinates*)

4.3 Hypotheses

Hypothesis 1: Microclimate Impact on Yield

Null Hypothesis (H01): Microclimatic and soil parameters (soil moisture, soil pH, temperature, rainfall, and sunlight hours) do not have a statistically significant effect on crop yield (yield_kg_per_hectare).

Alternative Hypothesis (Ha1): Microclimatic and soil parameters (soil moisture, soil pH, temperature, rainfall, and sunlight hours) significantly influence crop yield (yield_kg_per_hectare).

Hypothesis 2: Agronomic Management & Crop Vigor

Null Hypothesis (H02): There is no significant difference in crop yield (yield_kg_per_hectare) or vegetation health (NDVI_index) across different irrigation types (Drip, Sprinkler, Manual) and fertilizer treatments (Organic, Inorganic, Mixed).

Alternative Hypothesis (Ha2): Specific irrigation types and fertilizer formulations lead to statistically significant variations in crop yield (yield_kg_per_hectare) and vegetation health (NDVI_index).

Hypothesis 3: Remote Sensing Health Index & Disease Severity

Null Hypothesis (H03): There is no significant negative correlation between the NDVI_index and the severity level of crop_disease_status across different crop types.

Alternative Hypothesis (Ha3): Higher NDVI_index values significantly correlate with lower severity levels of crop_disease_status across different crop types.

5. Methodology:

5.1 Research Design: This is a Secondary Research conducted between July 2026 to August 2026. The research uses secondary data available on internet from Kaggle.com.

5.2 Sample size:

1. Sample Size Calculations

Cochran's formula to calculates the sample size required to estimate parameters with a specific confidence level and margin of error:

$$n = \frac{Z^2 \cdot p \cdot (1 - p)}{e^2}$$

$Z = 1.96$ (95% Confidence Level), $p = 0.50$ (Maximum variability), $e = 0.05$ (5% Margin of Error).

$$n = \frac{(1.96)^2 \times 0.50 \times 0.50}{(0.05)^2} = \frac{3.8416 \times 0.25}{0.0025} = 385 \text{ farms}$$

Determines minimum sample size based on the number of predictor variables (k) for testing OLS Regression (H_1):

$$N \geq 50 + 8k$$

$k = 6$ independent environmental predictors (soil_moisture_%, soil_pH, temperature_C, rainfall_mm, sunlight_hours, humidity_%).

Data analysis and finding :

$$N \geq 50 + 8(6) = 50 + 48 = 98 \text{ farms}$$

Dataset Overview :

The dataset contains **500 observations** across **22 variables** combining environmental metrics, soil parameters, agricultural inputs, crop health indices, and spatial-temporal metadata.

Total Records: 500 farms

Total Columns: 22 (12 numerical, 10 categorical/date variables)

Missing Values: Identified in irrigation_type (150 missing / 30%) and crop_disease_status (130 missing / 26%).

1. Descriptive Statistics: Numerical Variables

Variable	Count	Mean	Std Dev	Min	Median (Q2)	Max
yield_kg_per_hectare	500	4,032.93	1,174.43	2,023.56	4,071.69	5,998.29
NDVI_index	500	0.60	0.18	0.30	0.61	0.90
soil_moisture_%	500	26.75 %	10.15%	10.16%	25.86%	44.98%
soil_pH	500	6.52	0.59	5.51	6.53	7.50
temperature_C	500	24.68° C	5.35°C	15.00°C	24.66°C	34.84°C
rainfall_mm	500	181.69 mm	72.29 mm	50.17 mm	191.54 mm	298.96 mm
humidity_%	500	65.19 %	14.64%	40.23%	65.68%	90.00%
sunlight_hours	500	7.03 hrs	1.69 hrs	4.01 hrs	7.00 hrs	10.00 hrs
pesticide_usage_ml	500	26.59 mL	13.20 mL	5.05 mL	25.98 mL	49.94 mL
total_days	500	119.50	16.80	90.00	119.00	150.00
latitude	500	22.44	7.28	10.00	21.98	34.98
longitude	500	80.39	5.91	70.02	80.65	89.99

2. Categorical Variables Summary

Categorical Variable	Non-Null Count	Unique Values	Most Frequent (Mode)	Frequency	Missing Count (%)
region	500	5	Central USA	109	0 (0%)
crop_type	500	5	Maize	111	0 (0%)
fertilizer_type	500	3	Inorganic / Mixed	167	0 (0%)
irrigation_type	350	3	Sprinkler	121	150 (30%)
crop_disease_status	370	3	Severe	133	130 (26%)
farm_id	500	500	Unique per farm	1	0 (0%)
sensor_id	500	500	Unique per sensor	1	0 (0%)

Key Statistical Observations

Yield & Health Metrics: Crop yields display a uniform distribution ranging from **2,023.56 kg/ha** to **5,998.29 kg/ha**, with a mean of **4,032.93 kg/ha**. The NDVI index averages **0.60**, indicating generally moderate to healthy vegetation vigor across samples.

Environmental Balance: Soil pH is centered at **6.52** (near-neutral), while ambient temperature averages **24.68°C** and rainfall averages **181.69 mm**.

Data Completeness: 150 records miss irrigation_type values and 130 records miss crop_disease_status entries, which should be handled via imputation or explicit missingness modeling during analysis

Statistical Testing of Hypotheses

Test Results for Hypothesis 1: Microclimate Impact on Yield

Method: Ordinary Least Squares (OLS) Multiple Linear Regression

Dependent Variable: yield_kg_per_hectare

Independent Variables: soil_moisture_%, soil_pH, temperature_C, rainfall_mm, sunlight_hours, humidity_%

Regression Results Table

Predictor Variable	Coefficient (β)	Std. Error	t-value	p-value	95% Confidence Interval
Intercept	3694.3500	737.3510	5.0103	0.0000	[2245.62, 5143.09]
soil_moisture_%	-7.1897	5.1951	- 1.3839	0.1670	[-17.397, 3.018]
soil_pH	38.9748	89.9829	0.4331	0.6651	[-137.822, 215.772]
temperature_C	7.0821	9.8425	0.7195	0.4721	[-12.256, 26.420]
rainfall_mm	-1.1812	0.7296	- 1.6190	0.1061	[-2.615, 0.252]
sunlight_hours	16.7612	31.1307	0.5384	0.5905	[-44.404, 77.926]
humidity_%	3.0469	3.5920	0.8483	0.3967	[-4.011, 10.104]

Overall Model Statistics: $F(6, 493) = 1.0691$, $p\text{-value} = 0.3800$, $R^2 = 0.0128$ (Adjusted $R^2 = 0.0008$).

Interpretation (H1)

Decision: Fail to Reject Null Hypothesis (H_{01}) at $\alpha = 0.05$.

Analysis: The regression model as a whole is not statistically significant ($F = 1.0691$, $p = 0.3800$). None of the microclimatic predictors exhibit a statistically significant relationship with crop yield ($p > 0.05$ for all variables). Environmental features explain only 1.28% ($R^2 = 0.0128$) of the variance in yield within this synthetic/simulated dataset.

Test Results for Hypothesis 2: Agronomic Management Practices Method: Two-Way Analysis of Variance (Type II ANOVA)

Factors: irrigation_type (3 levels), fertilizer_type (3 levels), and their interaction (N = 350).

Two-Way ANOVA Results Table

Outcome Variable	Source of Variation	Sum of Squares	df	F-value	p-value (η^2)
yield_kg_per_hectare	irrigation_type	235,530	2	0.0862	0.9174
	fertilizer_type	293,699	2	0.1075	0.8981
	irrigation_type $\times$ fertilizer_type	5,094,390	4	0.9326	0.4451
	<i>Residuals</i>	465,696,000	341	—	—
NDVI_index	irrigation_type	0.0012	2	0.0193	0.9809
	fertilizer_type	0.0214	2	0.3514	0.7040
	irrigation_type $\times$ fertilizer_type	0.0428	4	0.3506	0.8435
	<i>Residuals</i>	10.3961	341	—	—

Interpretation (H2)

Decision: Fail to Reject Null Hypothesis (H02) at $\alpha = 0.05$.

Analysis: Neither irrigation type ($p = 0.9174$ for Yield; $p = 0.9809$ for NDVI) nor fertilizer type ($p = 0.8981$ for Yield; $p = 0.7040$ for NDVI) exerts a main effect on crop productivity or plant vigor. Furthermore, no significant interaction effect exists between irrigation methods and fertilizer choices.

Test Results for Hypothesis 3: Crop Health & Disease Severity

Method: Two-Way ANOVA & Bivariate Correlation Analysis (N = 370)

Variables: NDVI_index, crop_disease_status (Mild, Moderate, Severe), crop_type

Statistical Test Table

Metric / Test	Test	Parameter /	p-	Result
	Statistic	df	value	Summary
Pearson Correlation (r)	-0.0381	N = 370	0.4650	No linear correlation
Spearman Correlation (ρ)	-0.0368	N = 370	0.4807	No monotonic relationship

ANOVA: crop_disease_status	F = 0.4459	df = 2, 355	0.6406	No main effect on NDVI
ANOVA: crop_type	F = 0.8303	df = 4, 355	0.5066	No main effect on NDVI
ANOVA: Disease \$\times\$ Crop Interaction	F = 0.5992	df = 8, 355	0.7785	No interaction effect

NDVI Subgroup Descriptive Summary

Mild Disease: Mean NDVI = 0.6149 (SD = 0.1739) **Moderate Disease:** Mean NDVI = 0.5909 (SD = 0.1627) **Severe Disease:** Mean NDVI = 0.5996 (SD = 0.1897)

Interpretation (H3)

Decision: Fail to Reject Null Hypothesis (H03) at $\alpha = 0.05$.

Analysis: Vegetation health measured via NDVI_index does not significantly differ across disease severity levels ($p = 0.6406$) or crop varieties ($p = 0.5066$). Both Pearson ($r = -0.0381$, $p = 0.4650$) and Spearman ($\rho = -0.0368$, $p = 0.4807$) coefficients demonstrate an absence of association between remote-sensing vegetation indices and observed disease severity.

6. Recommendation based on study :

1. Shift to Non-Linear Machine Learning for Feature Interactions

Target Independent Variables: soil_moisture_%, soil_pH, temperature_C, rainfall_mm, sunlight_hours, humidity_%

Target Dependent Variable: yield_kg_per_hectare

Recommendation: Replace linear parametric models (e.g., OLS regression) with non-linear machine learning frameworks (e.g., Random Forests, XGBoost, Neural Networks) to model crop yield predictions.

2. Integrate Multi-Spectral Imaging Beyond Basic NDVI

Target Independent Variable: Early biological pathogen indicators & canopy stress

Target Dependent Variables: NDVI_index, crop_disease_status

Recommendation: Upgrade remote sensing instruments to include Red-Edge, Enhanced Vegetation Index (EVI), and Thermal Infrared sensors rather than relying exclusively on broad-band NDVI_index.

3. Implement Sensor Calibration & Imputation Protocols

Target Independent Variables: irrigation_type, pesticide_usage_ml, telemetry metrics

Target Dependent Variables: yield_kg_per_hectare, NDVI_index

Recommendation: Establish automated data quality pipelines for sensor maintenance, recalibration, and missing data handling (imputing the 30% missing values in **irrigation_type** and 26% in **crop_disease_status**).

4. Execute Controlled Multi-Factor Field Experiments (A/B Testing)

Target Independent Variables: irrigation_type (Drip, Sprinkler, Manual), fertilizer_type (Organic, Inorganic, Mixed)

Target Dependent Variable: yield_kg_per_hectare

Recommendation: Conduct randomized, controlled block experiments to systematically isolate the main and interaction effects of irrigation method and fertilizer choice.

5. Optimize Chemical Inputs via Variable Rate Application (VRA) Target Independent Variables: pesticide_usage_ml, fertilizer_type **Target Dependent Variables:** crop_disease_status, yield_kg_per_hectare

Recommendation: Transition from uniform chemical application to precision Variable Rate Application (VRA) guided by localized disease severity (**crop_disease_status**) and soil chemistry (**soil_pH**).

7. Conclusion:

This study provides a comprehensive empirical evaluation of smart farming dynamics across environmental telemetry, agronomic inputs, and crop health indices. Statistical analyses—including Ordinary Least Squares regression and Two-Way ANOVA—demonstrate that linear models fail to capture significant direct relationships between microclimatic factors, management practices, and ultimate crop yields. Furthermore, broadband NDVI health metrics showed no significant correlation with disease severity levels. Consequently, optimizing precision agriculture requires moving beyond linear assumptions. Future efforts should prioritize non-linear machine learning architectures, high-resolution multi-spectral imagery, and standardized sensor telemetry to effectively capture complex agro-ecological interactions and drive sustainable, high-yield agricultural decision-making.

8. Bibliography

- Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Behram, F. A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott, M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H.-H. S., ... Hazelwood, K. (2022)
- Sowgandika, B., Teja, V., Harsha, N., & Reddy, H. V. (2026). NDVI-based machine learning system for precision agriculture. In C. S. Pareek et al. (Eds.), *Proceedings of ICICCS-2026*
- Arockia Nisha, S., & Nanthini, M. (2026). Green AI for smart agriculture: Energy-efficient predictive models for crop yield and resource management. *International Journal of Engineering Research & Technology*, 14(3), 1–5.
- Sheshasaayee, A., & Angayarkanni, A. (2025). Role of artificial intelligence in sustainable agriculture – A critical concern to ecosystem. In A. Anand & S. Balasubramanian (Eds.), *Proceedings of Sustainability, Entrepreneurship, Equity and Digital Strategies (SEEDS 2024) (Atlantis Highlights in Economics, Business and Management, Vol. 3, pp. 288–307)*.
- Tabbakh, A., Al Amin, L., Islam, M., Mahmud, G. M. I., Chowdhury, I. K., & Mukta, M. S. H. (2024). Towards sustainable AI: a comprehensive framework for Green AI. *Discover Sustainability*, 5(1), Article 408.

- Greif, L., Kimmig, A., El Bobbou, S., Jurisch, P., & Ovtcharova, J. (2024). Strategic view on the current role of AI in advancing environmental sustainability: A SWOT analysis. *Discover Artificial Intelligence*, 4(1), Article 45.
- Abbassi, T.-E., Sahour, A., Benkherraf, M., Brai, S., Boutouta, I., & Boumehrez, F. (2026). Machine learning-based soil moisture prediction using meteorological data for enhanced irrigation management. *E3S Web of Conferences*, 699, Article 03005.
- Jääskeläinen, P., Pargman, D., & Holzapfel, A. (2022). On the environmental sustainability of Ai art(s). In *LIMITS '22: Workshop on Computing within Limits* (pp. 1–9). ACM.
- Walke, A., Srivastava, P. K., & Sharma, D. (2022). Exploring the impact of artificial intelligence on sustainable agriculture: A case study of precision farming. *International Journal of Creative Research Thoughts*, 10(5), 1–7.
- Sibiya, S. (2026). Sustainable crop yield prediction using Green AI in smart agriculture systems. *International Journal Research Publication Analysis*, 2(6), 1–12.
- Chaudhary, G. (2023). Environmental sustainability: Can Artificial Intelligence be an enabler for SDGs? *Nature Environment and Pollution Technology*, 22(3), 1411