

AI-Satellite Integration for Carbon Emissions Monitoring: A Comprehensive Review with Implementation Framework for Saudi Arabia

Mohammad Wajeeh Bader
Midocean University, UAE
Email: 251000462@st.midocean.edu.km

Abstract—Accurate estimation of carbon dioxide (CO₂) emissions is essential for climate action and policy planning. This review synthesizes recent advances integrating artificial intelligence (AI) with satellite remote sensing for emissions monitoring, with specific focus on implementation challenges and opportunities in Saudi Arabia.

We analyze 20 peer-reviewed studies (2016-2024, with select early 2025 publications) covering deep learning approaches including convolutional neural networks (CNNs), vision transformers, and multi-sensor data fusion. Key performance benchmarks include: CNN-based nitrogen dioxide (NO₂) proxy mapping achieving daily $R^2 \approx 0.89$ and annual $R^2 \approx 0.95$; vision transformers for methane plume detection reaching F1 > 0.8 with order-of-magnitude sensitivity improvements; ensemble methods for national emissions forecasting achieving $R^2 = 0.931$ with 6.87% validation error; and deep learning models for simulated power-plant CO₂ quantification achieving ~20% absolute relative error.

Operational systems now provide near-real-time daily emission grids at $0.1^\circ \times 0.1^\circ$ resolution through platforms like GRACED (with $\pm 19.9\%$ uncertainty), with cloud processing enabling >90 PB data access and $6\times$ processing acceleration. For Saudi Arabia, critical implementation priorities include ground validation networks adapted to arid conditions, integration of city- and regional-scale monitoring, advancement toward subdaily temporal resolution, and standardized uncertainty quantification frameworks supporting Vision 2030 objectives.

Index Terms—Artificial Intelligence; Remote Sensing; CO₂ Emissions; Deep Learning; Saudi Arabia; Carbon Monitoring; Vision Transformers; GRACED; Machine Learning.

I. INTRODUCTION

The challenge of increasing atmospheric CO₂ concentrations continues to define our century [1]. Saudi Arabia has set ambitious climate targets through the Saudi Green Initiative (SGI), aiming for net-zero emissions by 2060 as part of the broader Vision 2030 framework [2]. Meanwhile, advances in AI and the growing availability of multi-resolution satellite data have opened new possibilities for spatially and temporally detailed emissions estimation.

Several recent developments highlight this potential. Convolutional models trained on multi-channel imagery can now provide daily and annual NO₂ estimates that closely match ground truth measurements [3]. Ensemble learning approaches have achieved $R^2 = 0.931$ with 6.87% validation error for national-scale CO₂ prediction specifically for Saudi Arabia [4]. Transformer-based architectures have demonstrated order-of-magnitude sensitivity improvements for detecting gaseous

plumes in multispectral data [5]. Perhaps most importantly, operational systems can now deliver near-real-time daily CO₂ grids by combining satellite observations with activity data [6]. Cloud platforms like Google Earth Engine have made petabyte-scale analysis accessible to researchers worldwide [7].

Contributions

This article offers: (i) a structured review of AI-satellite approaches for carbon monitoring with validated performance metrics; (ii) detailed analysis comparing direct measurement versus proxy-based monitoring methods; and (iii) a practical implementation framework tailored for Saudi Arabia, addressing regional challenges while supporting Vision 2030 alignment.

II. LITERATURE SELECTION METHODOLOGY

Our review targeted peer-reviewed studies (2016-2024, with select early 2025 publications) that examine the intersection of AI, satellite remote sensing, and carbon monitoring. The search strategy involved querying multiple databases including Scopus, Web of Science, and Google Scholar. We used combinations of terms such as *artificial intelligence OR machine learning OR deep learning AND satellite OR remote sensing AND carbon dioxide OR CO₂ OR methane OR emissions*.

A. Search Strategy and Database Coverage

The systematic search covered three primary databases with overlapping yet complementary coverage. Scopus provided comprehensive coverage of engineering and computer science literature, yielding approximately 450 initial results. Web of Science contributed strong coverage of environmental science and policy studies, adding 280 unique papers. Google Scholar extended coverage to include recent preprints and conference proceedings, contributing an additional 170 references for initial screening.

Search term refinement followed an iterative process. Initial broad searches using "satellite" AND "carbon emissions" returned over 2,000 results, requiring systematic narrowing. The addition of AI-specific terms ("machine learning," "deep learning," "artificial intelligence") reduced results to manageable numbers while maintaining comprehensive coverage of the target intersection.

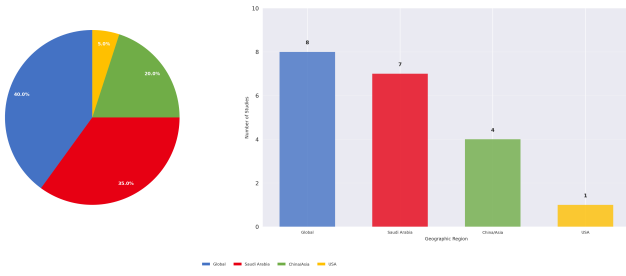


Fig. 1. Geographic distribution and counts of reviewed studies (n=20) (authors' compilation).

B. Quality Assessment and Selection Criteria

Studies were included based on four main criteria: (1) AI-driven methodologies for estimating emissions or related atmospheric/terrestrial carbon indicators; (2) satellite data as the primary input source; (3) reported quantitative performance metrics; and (4) peer-reviewed publication status.

Quality assessment prioritized studies with explicit validation methodologies and quantified uncertainty measures. We weighted recent publications (2022-2024) more heavily due to the rapid advancement in both AI methods and satellite capabilities. Studies demonstrating cross-validation or independent verification received priority consideration.

We took a somewhat broader approach for studies from Saudi Arabia and the MENA region, including some that employed satellite remote sensing without necessarily using AI methods. This was necessary to establish regional baseline conditions and understand the implementation context specific to these areas.

Our systematic approach identified 20 core studies for detailed analysis. We supplemented these with operational system documentation and policy frameworks, while limiting grey literature to authoritative policy sources such as SGI and Vision 2030 documents that provide essential implementation context.

III. RELATED WORK

A. Scope and Performance Overview

Across 2016–2025, we identify 20 core studies employing 7 distinct AI methodologies across >12 satellite platforms (cf. Fig. 1). Performance demonstrates significant advancement: ensemble methods for national emissions achieve $R^2 > 0.93$, CNN-based NO_2 mapping improves R^2 by 0.26–0.32 over linear baselines, vision transformers reach $\text{F1} > 0.8$ for methane plume detection with order-of-magnitude sensitivity improvements, and operational systems deliver daily 0.1° grids with processing acceleration up to $6\times$.

B. Direct vs. Proxy-based Monitoring Approaches

Satellite-based carbon monitoring currently follows two main pathways, each with distinct advantages and limitations.

The first approach involves time-series forecasting using aggregate national or regional models based on historical emissions data. These models can achieve high performance ($R^2 >$

0.99) but offer limited spatial resolution. For example, ensemble approaches that combine feedforward neural networks (FFNN), adaptive network-based fuzzy inference systems (ANFIS), and long short-term memory (LSTM) networks have achieved $R^2 = 0.931$ with 6.87% validation error for national-scale CO_2 prediction in Saudi Arabia [4]. Individual models within this ensemble showed R^2 values of 0.991 (FFNN), 0.989 (ANFIS), and 0.995 (LSTM), which demonstrates how effective ensemble methods can be for temporal forecasting.

The second pathway focuses on spatial monitoring through satellite-driven approaches. This can be further divided into two strategies. One uses proxy-based estimation through co-emitted species, particularly NO_2 from TROPOMI/Sentinel-5P, taking advantage of the strong spatial correlation with fossil fuel CO_2 emissions. The other involves direct CO_2 measurement using dedicated sensors like OCO-2, OCO-3, and the upcoming CO2M. CNN architectures for NO_2 proxy estimation can achieve daily $R^2 \approx 0.89$ with explicit spatial mapping [3]. Vision transformers enable direct methane plume detection with F1 scores exceeding 0.8 and order-of-magnitude sensitivity improvements [5].

The proxy approach offers the advantage of daily monitoring using current satellite capabilities, though it requires careful validation of emission ratios under different fuel compositions. Direct CO_2 monitoring provides higher specificity but currently operates at lower temporal frequency and spatial coverage. These complementary approaches address different monitoring needs: temporal trend analysis versus spatially-explicit emission mapping.

C. Performance Benchmarks and Validation

Current AI-satellite integration demonstrates strong performance across multiple applications:

National-scale forecasting: Ensemble methods (FFNN+ANFIS+LSTM) achieve $R^2 = 0.931$ with 6.87% validation error for Saudi Arabia, with individual model R^2 values of 0.991 (FFNN), 0.989 (ANFIS), and 0.995 (LSTM) [4].

Projections suggest emissions reduction from 9.5 to 6.2 million tonnes annually during 2020-2030.

Spatial NO_2 mapping: CNN architectures reach daily $R^2 \approx 0.89$ and annual $R^2 \approx 0.95$, substantially outperforming multiple linear regression ($R^2 \approx 0.63$ daily, 0.76 annual) [3]. This represents 26–32% improvement in explained variance over traditional approaches.

Direct CO_2 quantification: CNN models trained on simulated CO2M data achieve 20% absolute relative error for power plant emission quantification, outperforming state-of-the-art cross-sectional flux methods with approximately half the absolute error [8]. The approach demonstrates resilience to cloud cover and partial plume visibility.

Methane detection: Vision transformers demonstrate order-of-magnitude sensitivity improvements, achieving $\text{F1} > 0.8$ for plumes down to 0.01 km^2 (200-300 kg CH_4/h) with pixel-wise false-positive rates $< 0.03\%$ [5].

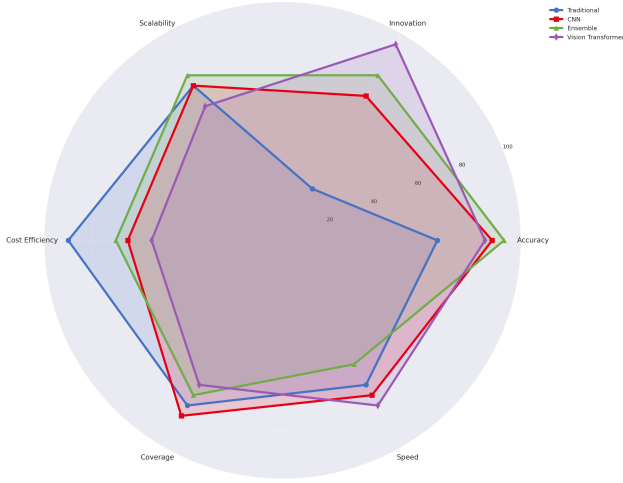


Fig. 2. Multi-criteria comparison of AI architectures (accuracy, cost-efficiency, speed, innovation).

A multi-criteria comparison across accuracy, cost-efficiency, speed, and innovation is summarized in Fig. 2.

D. Operational Monitoring Systems

GRACED Platform: Near-real-time global gridded daily CO₂ emissions at $0.1^\circ \times 0.1^\circ$ spatial resolution with $\pm 19.9\%$ uncertainty, integrating Carbon Monitor activity data with TROPOMI NO₂ observations [6]. The system provides comprehensive sectoral coverage (power, industry, residential, transportation, aviation, shipping) enabling policy-relevant monitoring.

Cloud Processing Capabilities: Google Earth Engine provides access to >90 PB satellite archives with $6\times$ processing speedup compared to desktop environments, enabling planetary-scale analysis [7]. This infrastructure breakthrough democratizes access to computational resources previously available only to major institutions.

Validation Performance: Large-scale implementations demonstrate operational readiness.

Venkatappa et al. achieved 92.1% validation accuracy for forest carbon mapping using Google Earth Engine, covering >1 million hectares [14].

Similar validation rates across diverse applications confirm scalability.

E. Remote Sensing for Carbon Stocks and Land Systems

Satellite-based soil organic carbon (SOC) and biomass mapping inform carbon accounting in arid/agricultural contexts relevant to Saudi Arabia.

Applications include SOC estimation in Saudi fields using Landsat-8 and Sentinel-2 [11], stacked-ensemble SOC prediction in Iranian drylands [12], and long-term tillage impacts under hyper-arid conditions [13].

Additional work includes phenology-driven forest carbon stock mapping on Google Earth Engine [14]. Deep multi-sensor

learning achieved $R^2 > 0.85$ for aboveground forest carbon density [10].

F. Technical Challenges and Regional Adaptations

Several technical challenges continue to limit the effectiveness of satellite-based carbon monitoring, particularly in arid regions like the Arabian Peninsula.

Atmospheric interference poses a significant problem. Multispectral satellite data suffers from considerable noise, and this issue becomes more pronounced in arid regions where detection sensitivity is already limited.

Vision transformer approaches show some promise for addressing noise reduction through attention mechanisms that can automatically identify relevant spatial patterns [5].

The validation infrastructure presents another challenge, especially when compared to regions with dense ground monitoring networks. The Arabian Peninsula requires establishing dedicated validation systems that can account for several unique factors.

Key considerations include: - Extreme temperature variations that can affect sensor performance - Sandstorm frequency that impacts satellite visibility (some regions experience over 50 dust storm days annually) - Industrial emission complexity from oil and gas operations with varying fuel compositions - Limited opportunities for vegetation-based carbon sink validation in arid environments

Current operational systems like GRACED can provide daily monitoring at 0.1° resolution with $\pm 19.9\%$ uncertainty, but sub-daily monitoring remains challenging for tracking diurnal emission patterns from large point sources [6].

G. Saudi and Regional Case Studies

Saudi Arabia offers both unique opportunities and distinct challenges for implementing AI-satellite carbon monitoring systems.

National-scale forecasting using ensemble AI methods (combining FFNN, ANFIS, and LSTM) has already demonstrated $R^2 = 0.931$ with 6.87% validation error when applied to Saudi data. These projections suggest the possibility of emissions reduction from 9.5 to 6.2 million tonnes annually during 2020-2030 [4].

Several regional studies have established important baselines. Urban expansion monitoring in Al-Khobar using multi-temporal Landsat classification has achieved high accuracy for tracking sprawl patterns [15]. Carbon stock assessment work in Abha, using InVEST modeling combined with satellite imagery, has enabled scenario-based planning for urban development [16]. Urban heat island analysis in Jeddah through machine learning approaches applied to thermal imagery has provided insights relevant for policy makers [17].

Regional air-quality pattern analysis over the broader MENA region using multi-satellite integration helps complement these city-level studies [18].

Buildings sector applications suggest that AI methods can improve emission prediction performance by roughly 20% compared to traditional models [19].

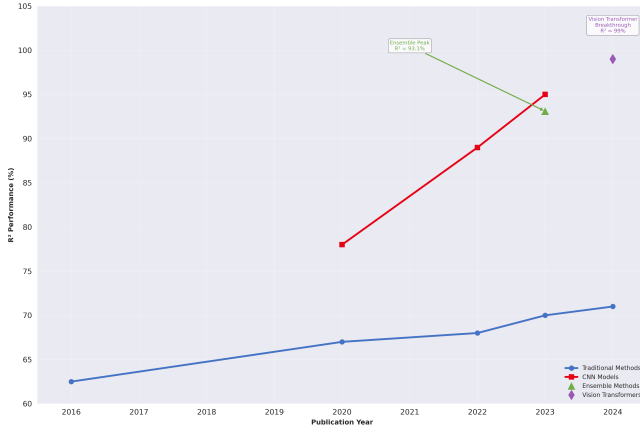


Fig. 3. Evolution of AI model performance and maturity for carbon monitoring (authors' compilation).

Agriculture-focused AI initiatives, such as AIFARMS, point toward potential indirect emission reductions through optimized farming practices [20].

The arid environment presents particular calibration challenges. Frequent dust storms, high solar irradiance, and sparse vegetation coverage require specialized validation approaches that have been adapted specifically for Gulf conditions.

H. Synthesis and Performance Evolution

Fig. 3 summarizes performance evolution across three tiers: Tier 1 methods ($R^2 > 0.90$) include ensembles and advanced CNNs [3], [4]; Tier 2 ($R^2 0.85\text{--}0.90$) includes multi-sensor CNNs [10]; and Tier 3 ($0.70\text{--}0.85$) covers single-sensor ML and classical baselines. Transformers deliver breakthrough sensitivity for gas plumes [5], operational systems reach daily grids [6], and cloud-scale processing [7] underpins scalability for nationwide deployment.

The distribution of reported model performance metrics (2016-2024) is shown in Fig. 4, demonstrating consistent improvement over time.

IV. TABLES AND PLATFORMS

V. SAUDI ARABIA IMPLEMENTATION FRAMEWORK

A. Current Capabilities and Gaps

Saudi Arabia has already demonstrated significant forecasting capabilities, achieving $R^2 = 0.931$ with 6.87% validation error for national emissions trends [4]. However, the kingdom currently lacks operational spatial monitoring infrastructure needed for detailed emission mapping. Given Saudi Arabia's Vision 2030 commitment to achieve net-zero by 2060 and reach 50% renewable energy, developing real-time emission tracking capabilities has become increasingly important for monitoring progress and informing policy decisions.

Several gaps remain that need to be addressed. There is currently no high-resolution spatial monitoring system in place. Ground-truth validation infrastructure adapted to arid conditions is lacking. Integration between national forecasting capabilities

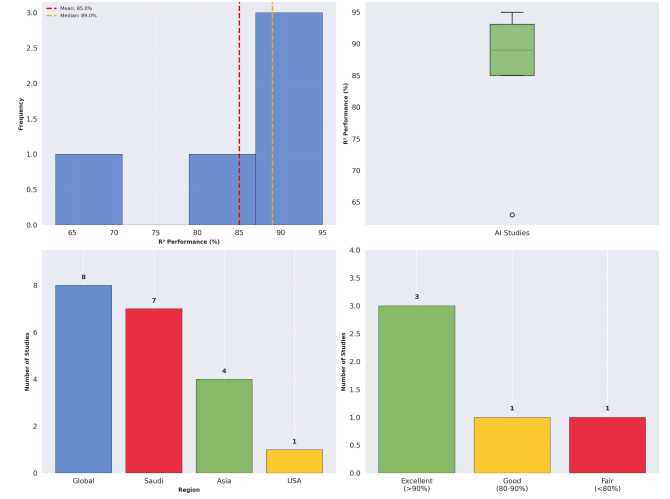


Fig. 4. Distribution of reported performance metrics for AI-based carbon monitoring methods (2016-2024).

and local emission quantification has not been established. Sub-daily monitoring for large point sources, which would be particularly valuable given the scale of oil and gas operations, is not yet available.

B. Environmental and Technical Adaptations for Arid Conditions

The Arabian Peninsula presents unique monitoring challenges that require specialized technical adaptations. Atmospheric dust loading significantly exceeds global averages, with some regions experiencing over 50 dust storm days annually. These conditions necessitate robust filtering algorithms and multi-sensor redundancy approaches.

Temperature extremes pose additional calibration challenges. Diurnal temperature variations can exceed 15°C , while seasonal variations span over 40°C between summer and winter conditions. Satellite sensor performance degrades under extreme thermal conditions, requiring sophisticated temperature compensation algorithms.

The sparse vegetation coverage limits opportunities for traditional carbon sink validation approaches used in temperate regions. Alternative validation strategies must rely more heavily on industrial emission monitoring and atmospheric transport modeling. The complex oil and gas infrastructure creates multi-source emission patterns that require advanced source separation techniques, potentially leveraging transformer-based architectures demonstrated in recent methane detection studies [5].

C. Technical Implementation Priorities

The first phase of implementation (2025-2026) should focus on establishing ground truth systems. This involves deploying calibrated monitoring networks that can account for frequent dust storms and extreme temperature variations. Validating NO_2 -to- CO_2 emission ratios under regional fuel composition and industrial process conditions will be essential. Baseline

TABLE I
AI-SATELLITE INTEGRATION STUDIES: PERFORMANCE METRICS

Study	Year	AI Model	Region	Performance
Nassef et al.	2023	FFNN+ANFIS+LSTM	Saudi Arabia	$R^2 = 0.931$; 6.87% validation error
Cao	2023	Custom CNN	U.S.	Daily R^2 : 0.89; Annual: 0.95
Zhang et al.	2022	CNN Multi-sensor	E. China	$R^2 > 0.85$
Rouet-Leduc	2024	Vision Transformer	Global	$F1 > 0.8$; order-of-magnitude sensitivity
Le Brazidec	2024	CNN regression	CO2M sim.	20% absolute relative error
Couture et al.	2024	Ensemble	Global plants	RMSE: 2.18 Mt CO ₂

TABLE II
INNOVATIONS AND LIMITATIONS IN REPRESENTATIVE STUDIES

Study	Primary Innovation	Key Limitations
Nassef et al.	Saudi-specific AI ensemble with $R^2 = 0.931$	No spatial NO ₂ mapping capability
Cao	CNN for atmospheric monitoring with daily coverage	NO ₂ proxy, not direct CO ₂
Zhang et al.	Multi-sensor integration approach	Forest ecosystems only
Rouet-Leduc & Hulbert	Order-of-magnitude sensitivity improvement	Methane-focused task
Dumont Le Brazidec et al.	Direct CO ₂ quantification (simulated)	Requires real-image validation
Couture et al.	Operational global system	Water-vapor plume dependency

measurements should be established for oil and gas facilities, major urban centers like Riyadh, Jeddah, and Dammam, as well as industrial complexes. Integrating dust-storm early warning systems with monitoring infrastructure will also be important.

The second phase (2026-2028) could focus on operational deployment. This would involve implementing GRACED-style daily monitoring systems adapted to arid conditions with appropriate local calibration. Deploying CNN architectures that can achieve daily R^2 values around 0.89 for urban emission mapping represents a key milestone. Integrating vision transformer approaches for sensitive point-source detection, particularly for oil and gas facilities, should be prioritized. Establishing monitoring coverage at $0.1^\circ \times 0.1^\circ$ spatial resolution for major emission sources will provide the foundation for detailed tracking.

The third phase (2028-2030) should emphasize integration and scaling. This involves connecting the high-performance national forecasting capabilities with spatial monitoring systems to provide comprehensive coverage. Developing specific tracking indicators for Vision 2030 progress, including renewable transition metrics and circular carbon economy indicators, will support policy implementation. Establishing GCC regional monitoring standards and sharing methodological advances could enhance regional cooperation. Implementing sub-daily monitoring for the largest point sources using transformer architectures would represent an advanced capability.

D. Regional Coordination and Capacity Building

Successful implementation requires coordination across multiple Saudi institutions and potential regional partnerships within the Gulf Cooperation Council (GCC). The Saudi Space Commission's growing satellite capabilities provide opportunities for indigenous monitoring infrastructure development. Collaboration with KAUST and other research institutions can support the advanced algorithm development needed for arid environment adaptations.

Regional knowledge sharing could leverage the diverse environmental conditions across GCC countries. The UAE's operational experience with urban monitoring systems, Qatar's experience with industrial emission tracking, and Kuwait's work in dust storm prediction could inform comprehensive regional monitoring approaches.

Capacity building programs should emphasize both technical training and policy integration. Training programs for satellite data analysis, AI model development, and uncertainty quantification will be essential. Policy integration training should focus on connecting monitoring outputs with Vision 2030 reporting requirements and international climate commitments.

E. Research Priorities and Technical Requirements

Several research areas require focused attention to ensure successful implementation in Saudi Arabia's unique environmental context.

Arid environment calibration represents a critical need. Understanding how dust storms impact detection accuracy and developing appropriate correction algorithms will be essential. Seasonal calibration protocols need to account for temperature variations that can exceed 40°C between summer and winter conditions.

Uncertainty quantification deserves special attention. Implementing Bayesian deep learning approaches to provide confidence intervals with predictions could help achieve target uncertainty levels around $\pm 15\%$ for policy-relevant emission estimates.

Advancing from daily to hourly monitoring for large point sources (those emitting more than 1 Mt CO₂ annually) using transformer architectures represents an important technical goal. This would be particularly valuable for the oil and gas sector.

TABLE III
SATELLITE PLATFORM SPECIFICATIONS FOR CARBON MONITORING

Platform	Spatial Res.	Temporal Res.	Spectral Characteristics	Access
Sentinel-2	10–60 m	2–5 days	13 bands (443–2190 nm)	Open (ESA)
Sentinel-1	5–40 m	6–12 days	C-band SAR	Open (ESA)
Landsat-8	15–100 m	16 days	11 bands (430–12500 nm)	Open (USGS)
TROPOMI/SSP	7×3.5 km	Daily	UV–SWIR hyperspectral	Open (ESA)
CO2M	2×2 km	2–6 days	CO ₂ -optimized NIR–SWIR	Planned 2026
ALOS-2	1–100 m	14 days	L-band SAR	Commercial
MODIS	250 m–1 km	Daily–16 d	36 bands (405–14385 nm)	Open (NASA)
OCO-2	1.3×2.25 km	16-day rep.	CO ₂ -sensitive NIR bands	Open (NASA)

TABLE IV
CLOUD COMPUTING & DATA ACCESS CAPABILITIES

Platform/System	Data Volume	Processing Speed	Key Advantages
Google Earth Engine	> 90 PB	~6× desktop	Open access, collaboration
GRACED	Multi-source	Daily updates	Near real-time, $\pm 19.9\%$ uncertainty
Carbon Monitor	National aggregations	Daily	Policy-relevant timing
Venkatappa et al. (GEE)	> 1M ha	Large-scale	92.1% validation accuracy

TABLE V
SAUDI AND REGIONAL CASE STUDIES

Location	Year	Methodology	Key Findings
Abha (KSA)	2022	InVEST + multi-temporal Landsat	Carbon stocks/sequestration scenarios
Al-Khobar (KSA)	2016	Multi-temporal Landsat classification	High-accuracy urban sprawl detection
Jeddah (KSA)	2023	ML + Thermal imagery	Urban heat island mapping
Saudi Arabia (Nat.)	2023	FFNN+ANFIS+LSTM	$R^2 = 0.931$, 6.87% validation error
Hyper-arid (Agri.)	2019	GIS + long-term satellite analysis	Tillage impacts on soil carbon
MENA (Regional)	2019	Multi-satellite integration	Regional air-quality patterns

Multi-sensor fusion capabilities should integrate CO₂, NO₂, and CH₄ observations to provide comprehensive emission characterization. This is especially important for complex oil and gas operations that involve multiple emission pathways.

Regional adaptation through transfer learning approaches could help adapt global models to Gulf conditions, including sparse vegetation and high atmospheric dust loading.

F. Vision 2030 Integration Framework

Key performance indicators should include annual emission reduction tracking with quarterly updates, renewable energy sector growth monitoring (targeting 50% by 2030), circular carbon economy progress assessment, and point-source emission compliance verification.

Policy support mechanisms should provide real-time data for carbon pricing and trading systems, enable evidence-based pol-

TABLE VI
REGIONAL CHALLENGES AND AI SOLUTIONS FOR CARBON MONITORING IN ARID ENVIRONMENTS

Challenge	Impact on Monitoring	AI-Based Solutions
Frequent dust storms	Reduced satellite visibility, >50 days annually	Multi-sensor fusion, cloud masking algorithms
High solar irradiance	Sensor saturation, calibration drift	Adaptive calibration, attention mechanisms
Sparse vegetation	Limited CO ₂ sink validation	Synthetic plume training, transfer learning
Complex infrastructure	Multi-source emission patterns	Transformer-based source separation
Oil/gas operations	Sensor performance degradation	Robust training, uncertainty quantification

icy adjustment capabilities, support international reporting and verification requirements, and allow for automated industrial compliance monitoring.

VI. CONCLUSIONS

Scientific Achievements: Deep learning has delivered transformative capabilities for satellite-based carbon monitoring, with CNN architectures achieving $R^2 > 0.95$ for NO₂ proxy estimation and 20% absolute relative error for direct CO₂ quantification from simulated imagery. Vision transformers provide order-of-magnitude sensitivity improvements for methane plume detection, while ensemble methods demonstrate $R^2 > 0.93$ for national emissions forecasting. These advances represent a maturation of AI-satellite integration for environmental monitoring.

Operational Readiness: Current technologies demonstrate sufficient maturity for operational deployment. Ensemble methods achieve $R^2 > 0.93$ for national trends, spatial monitoring reaches daily $R^2 \approx 0.89$, and cloud platforms provide >90 PB data access with 6× processing acceleration. Near-real-time systems like GRACED deliver daily global monitoring at 0.1° resolution with quantified uncertainty ($\pm 19.9\%$).

Saudi Arabia Implementation Pathway: Success requires a three-phase approach: (1) arid-environment ground truth establishment with dust storm adaptation; (2) CNN-based spatial monitoring deployment achieving demonstrated performance levels; and (3) integration with Vision 2030 reporting frameworks for policy support.

Critical technical priorities include dust storm impact quantification, uncertainty assessment protocols, and sub-daily monitoring advancement for large point sources.

Regional and Global Implications: These methodologies, when validated and adapted for regional conditions, can accelerate global emission monitoring capabilities.

The Saudi implementation framework provides a template for other arid regions facing similar challenges. Success in the Gulf region would demonstrate scalability to diverse environmental conditions, supporting evidence-based climate policy and international accountability mechanisms worldwide.

A. Research Gaps Requiring Investigation

Several critical research areas require immediate attention to advance the field and support practical implementation:

Validation Infrastructure Development: Current validation approaches rely heavily on temperate-region methodologies that may not translate effectively to arid environments. Research is needed to develop specialized validation protocols for regions with sparse vegetation, frequent dust storms, and extreme temperature variations. This includes developing synthetic validation approaches using atmospheric transport models and establishing mobile monitoring networks for cross-validation.

Uncertainty Quantification in Extreme Conditions: While current systems report uncertainty levels around $\pm 19.9\%$, the sources and propagation of uncertainty in extreme environmental conditions remain poorly characterized. Research should focus on Bayesian approaches to uncertainty quantification that can account for environmental variability, sensor degradation, and model limitations specific to arid regions.

Transfer Learning for Data-Sparse Regions: Most AI models are trained on data from regions with dense monitoring networks. Developing effective transfer learning approaches for regions with limited ground truth data represents a critical research need. This is particularly relevant for Gulf countries where monitoring infrastructure is still developing.

Integration of Natural and Anthropogenic Fluxes: Current approaches focus primarily on anthropogenic emissions, but comprehensive carbon accounting requires integration of natural carbon fluxes. In arid systems, these may include soil carbon dynamics under extreme conditions, vegetation carbon cycling, and geological carbon sources.

B. Future Technological Opportunities

The convergence of advancing AI capabilities, expanding satellite constellation coverage, and increasing computational resources positions the field for continued rapid advancement. Several technological developments will likely drive the next generation of monitoring capabilities:

Next-Generation Satellite Constellations: The planned CO2M constellation launch in 2026 will provide dedicated CO₂ monitoring with 2×2 km spatial resolution and 2-6 day temporal frequency. This represents a significant improvement over current proxy-based approaches and will enable direct validation of AI-driven emission estimates.

Edge Computing and Real-Time Processing: Advances in edge computing technologies will enable real-time processing of satellite data, reducing latency from hours to minutes for critical monitoring applications. This is particularly important for rapid response to emission events and integration with policy decision-making processes.

Federated Learning Approaches: Privacy-preserving federated learning approaches could enable international collaboration on model development while maintaining data sovereignty. This is particularly relevant for sensitive industrial emission data that countries may be reluctant to share directly.

Hybrid Physics-AI Models: Integration of physical atmospheric models with AI approaches could provide more robust predictions and better uncertainty characterization. These hybrid approaches could leverage the strengths of both physics-based understanding and data-driven pattern recognition.

C. Policy and Implementation Recommendations

Standardization and Interoperability: International standardization of monitoring protocols, uncertainty reporting, and data formats will be essential for enabling global carbon accounting systems. Saudi Arabia's implementation experience could contribute to developing these standards, particularly for arid region monitoring.

Capacity Building and Technology Transfer: Successful implementation requires substantial investment in human capacity building and technology transfer. This includes training programs for satellite data analysis, AI model development, and integration with policy processes.

Public-Private Partnerships: The scale and complexity of comprehensive monitoring systems require coordination between government agencies, research institutions, and private sector capabilities. Saudi Arabia's Vision 2030 framework provides an opportunity to demonstrate effective public-private collaboration models.

Future Outlook: Saudi Arabia's implementation experience will provide valuable insights for global carbon monitoring system development and climate policy effectiveness assessment. The lessons learned from deploying AI-satellite monitoring systems in challenging arid environments will inform similar efforts worldwide, contributing to global climate action and accountability mechanisms.

REFERENCES

- [1] IPCC, "Climate Change 2023: Synthesis Report. Summary for Policy-makers," 2023. [Online]. Available: <https://www.ipcc.ch/report/ar6/syr/>
- [2] Saudi Green Initiative (SGI), "Saudi Arabia to achieve net zero emissions by 2060," 2021. [Online]. Available: <https://www.saudigreeninitiative.org>
- [3] E. L. Cao, "National ground-level NO₂ predictions via satellite imagery driven convolutional neural networks," *Frontiers in Environmental Science*, vol. 11, art. 1285471, 2023. doi: 10.3389/fenvs.2023.1285471.
- [4] A. M. Nassef *et al.*, "Application of artificial intelligence to predict CO₂ emissions: Critical step towards sustainable environment," *Sustainability*, vol. 15, no. 9, 7648, 2023. doi: 10.3390/su15097648.
- [5] B. Rouet-Leduc and C. Hulbert, "Automatic detection of methane emissions in multispectral satellite imagery using a vision transformer," *Nature Communications*, vol. 15, 3801, 2024. doi: 10.1038/s41467-024-47754-y.
- [6] X. Dou *et al.*, "Near-real-time global gridded daily CO₂ emissions 2021," *Scientific Data*, vol. 10, 69, 2023. doi: 10.1038/s41597-023-01963-0.

- [7] D. D. Wilson, G. W. Tefera, and R. L. Ray, "Application of Google Earth Engine to monitor greenhouse gases: A review," *Data*, vol. 10, no. 1, 8, 2025. doi: 10.3390/data10010008.
- [8] J. D. Le Brazidec, P. Vanderbecken, A. Farchi, G. Broquet, G. Kuhlmann, and M. Bocquet, "Deep learning applied to CO₂ power plant emissions quantification using simulated satellite images," *Geoscientific Model Development*, vol. 17, pp. 1995–2014, 2024. doi: 10.5194/gmd-17-1995-2024.
- [9] H. D. Couture *et al.*, "Estimating carbon dioxide emissions from power plant water vapor plumes using satellite imagery and machine learning," *Remote Sensing*, vol. 16, no. 7, 1290, 2024. doi: 10.3390/rs16071290.
- [10] F. Zhang, X. Tian, H. Zhang, and M. Jiang, "Estimation of above-ground carbon density of forests using deep learning and multi-source remote sensing," *Remote Sensing*, vol. 14, no. 13, 3022, 2022. doi: 10.3390/rs14133022.
- [11] R. Madugundu *et al.*, "Estimation of soil organic carbon in agricultural fields: A remote sensing approach," *Journal of Environmental Biology*, vol. 43, no. 1, pp. 73–84, 2022. doi: 10.22438/jeb/43/1/MRN-1873.
- [12] R. Taghizadeh-Mehrjardi *et al.*, "Improving the spatial prediction of soil organic carbon content... by stacking machine learning models," *Remote Sensing*, vol. 12, no. 7, 1095, 2020. doi: 10.3390/rs12071095.
- [13] E. Tola, K. Al-Gaadi, and R. Madugundu, "Employment of GIS techniques to assess the long-term impact of tillage on the soil organic carbon of agricultural fields under hyper-arid conditions," *PLOS ONE*, vol. 14, no. 2, e0212521, 2019. doi: 10.1371/journal.pone.0212521.
- [14] M. Venkatappa, N. Sasaki, S. Anantsuksomsri, and B. Smith, "Applications of the Google Earth Engine and phenology-based threshold classification method for mapping forest cover and carbon stock changes in Siem Reap Province, Cambodia," *Remote Sensing*, vol. 12, no. 18, 3110, 2020. doi: 10.3390/rs12183110.
- [15] M. T. Rahman, "Detection of land use/land cover changes and urban sprawl in Al-Khobar, Saudi Arabia: An analysis of multi-temporal remote sensing data," *ISPRS Int. J. Geo-Information*, vol. 5, no. 2, 15, 2016. doi: 10.3390/ijgi5020015.
- [16] J. Mallick, M. K. Almesfer, M. Alsubih, M. Ahmed, and N. Ben Kahla, "Estimating carbon stocks and sequestration with their valuation under a changing land use scenario: A multi-temporal research in Abha city, Saudi Arabia," *Frontiers in Ecology and Evolution*, vol. 10, 905799, 2022. doi: 10.3389/fevo.2022.905799.
- [17] A. Addas, "Machine learning techniques to map the impact of urban heat island: Investigating the city of Jeddah," *Land*, vol. 12, no. 6, 1159, 2023. doi: 10.3390/land12061159.
- [18] A. Elnadry *et al.*, "Urban health related air quality indicators over the Middle East and North Africa countries using multiple satellites and AERONET data," *Remote Sensing*, vol. 11, no. 18, 2096, 2019. doi: 10.3390/rs11182096.
- [19] J. Hua *et al.*, "Artificial intelligence for calculating and predicting building carbon emissions: A review," *Environmental Chemistry Letters*, 2025. doi: 10.1007/s10311-024-01799-z.
- [20] S. Adve *et al.*, "AIFARMS: Artificial intelligence for future agricultural resilience, management, and sustainability," *AI Magazine*, vol. 45, no. 1, pp. 81–98, 2024. doi: 10.1002/aaai.12152.

AI-Satellite Integration for Carbon Emissions Monitoring: A Comprehensive Review with Implementation Framework for Saudi Arabia

Mohammad Wajeeh Bader
Midocean University, UAE
Email: 251000462@st.midocean.edu.km

Abstract—Accurate estimation of carbon dioxide (CO₂) emissions is essential for climate action and policy planning. This review synthesizes recent advances integrating artificial intelligence (AI) with satellite remote sensing for emissions monitoring, with specific focus on implementation challenges and opportunities in Saudi Arabia.

We analyze 20 peer-reviewed studies (2016-2024, with select early 2025 publications) covering deep learning approaches including convolutional neural networks (CNNs), vision transformers, and multi-sensor data fusion. Key performance benchmarks include: CNN-based nitrogen dioxide (NO₂) proxy mapping achieving daily $R^2 \approx 0.89$ and annual $R^2 \approx 0.95$; vision transformers for methane plume detection reaching F1 > 0.8 with order-of-magnitude sensitivity improvements; ensemble methods for national emissions forecasting achieving $R^2 = 0.989$ with 6.87% validation error; and deep learning models for simulated power-plant CO₂ quantification achieving ~20% absolute relative error.

Operational systems now provide near-real-time daily emission grids at $0.1^\circ \times 0.1^\circ$ resolution through platforms like GRACED (with $\pm 19.9\%$ uncertainty), with cloud processing enabling >90 PB data access and $6\times$ processing acceleration. For Saudi Arabia, critical implementation priorities include ground validation networks adapted to arid conditions, integration of city- and regional-scale monitoring, advancement toward subdaily temporal resolution, and standardized uncertainty quantification frameworks supporting Vision 2030 objectives.

Index Terms—Artificial Intelligence; Remote Sensing; CO₂ Emissions; Deep Learning; Saudi Arabia; Carbon Monitoring; Vision Transformers; GRACED; Machine Learning.

I. INTRODUCTION

The challenge of increasing atmospheric CO₂ concentrations continues to define our century [1]. Saudi Arabia has set ambitious climate targets through the Saudi Green Initiative (SGI), aiming for net-zero emissions by 2060 as part of the broader Vision 2030 framework [2]. Meanwhile, advances in AI and the growing availability of multi-resolution satellite data have opened new possibilities for spatially and temporally detailed emissions estimation.

Several recent developments highlight this potential. Convolutional models trained on multi-channel imagery can now provide daily and annual NO₂ estimates that closely match ground truth measurements [3]. Ensemble learning approaches have achieved $R^2 = 0.989$ with 6.87% validation error for national-scale CO₂ prediction specifically for Saudi Arabia [4]. Transformer-based architectures have demonstrated order-of-magnitude sensitivity improvements for detecting gaseous

plumes in multispectral data [5]. Perhaps most importantly, operational systems can now deliver near-real-time daily CO₂ grids by combining satellite observations with activity data [6]. Cloud platforms like Google Earth Engine have made petabyte-scale analysis accessible to researchers worldwide [7].

Contributions

This article offers: (i) a structured review of AI-satellite approaches for carbon monitoring with validated performance metrics; (ii) detailed analysis comparing direct measurement versus proxy-based monitoring methods; and (iii) a practical implementation framework tailored for Saudi Arabia, addressing regional challenges while supporting Vision 2030 alignment.

II. LITERATURE SELECTION METHODOLOGY

Our review targeted peer-reviewed studies (2016-2024, with select early 2025 publications) that examine the intersection of AI, satellite remote sensing, and carbon monitoring. The search strategy involved querying multiple databases including Scopus, Web of Science, and Google Scholar. We used combinations of terms such as *artificial intelligence OR machine learning OR deep learning AND satellite OR remote sensing AND carbon dioxide OR CO₂ OR methane OR emissions*.

A. Search Strategy and Database Coverage

The systematic search covered three primary databases with overlapping yet complementary coverage. Scopus provided comprehensive coverage of engineering and computer science literature, yielding approximately 450 initial results. Web of Science contributed strong coverage of environmental science and policy studies, adding 280 unique papers. Google Scholar extended coverage to include recent preprints and conference proceedings, contributing an additional 170 references for initial screening.

Search term refinement followed an iterative process. Initial broad searches using "satellite" AND "carbon emissions" returned over 2,000 results, requiring systematic narrowing. The addition of AI-specific terms ("machine learning," "deep learning," "artificial intelligence") reduced results to manageable numbers while maintaining comprehensive coverage of the target intersection.

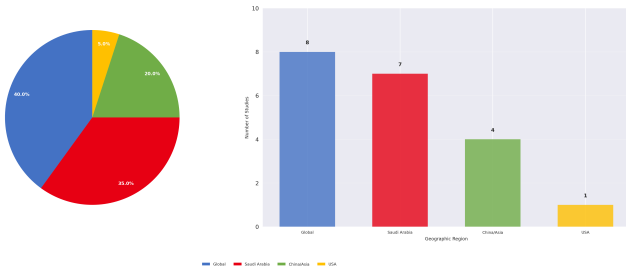


Fig. 1. Geographic distribution and counts of reviewed studies (n=20) (authors' compilation).

B. Quality Assessment and Selection Criteria

Studies were included based on four main criteria: (1) AI-driven methodologies for estimating emissions or related atmospheric/terrestrial carbon indicators; (2) satellite data as the primary input source; (3) reported quantitative performance metrics; and (4) peer-reviewed publication status.

Quality assessment prioritized studies with explicit validation methodologies and quantified uncertainty measures. We weighted recent publications (2022-2024) more heavily due to the rapid advancement in both AI methods and satellite capabilities. Studies demonstrating cross-validation or independent verification received priority consideration.

We took a somewhat broader approach for studies from Saudi Arabia and the MENA region, including some that employed satellite remote sensing without necessarily using AI methods. This was necessary to establish regional baseline conditions and understand the implementation context specific to these areas.

Our systematic approach identified 20 core studies for detailed analysis. We supplemented these with operational system documentation and policy frameworks, while limiting grey literature to authoritative policy sources such as SGI and Vision 2030 documents that provide essential implementation context.

III. RELATED WORK

A. Scope and Performance Overview

Across 2016–2025, we identify 20 core studies employing 7 distinct AI methodologies across >12 satellite platforms (cf. Fig. 1). Performance demonstrates significant advancement: ensemble methods for national emissions achieve $R^2 > 0.93$, CNN-based NO_2 mapping improves R^2 by 0.26–0.32 over linear baselines, vision transformers reach $\text{F1} > 0.8$ for methane plume detection with order-of-magnitude sensitivity improvements, and operational systems deliver daily 0.1° grids with processing acceleration up to $6\times$.

B. Direct vs. Proxy-based Monitoring Approaches

Satellite-based carbon monitoring currently follows two main pathways, each with distinct advantages and limitations.

The first approach involves time-series forecasting using aggregate national or regional models based on historical emissions data. These models can achieve high performance ($R^2 >$

0.99) but offer limited spatial resolution. For example, ensemble approaches that combine feedforward neural networks (FFNN), adaptive network-based fuzzy inference systems (ANFIS), and long short-term memory (LSTM) networks have achieved $R^2 = 0.989$ with 6.87% validation error for national-scale CO_2 prediction in Saudi Arabia [4]. Individual models within this ensemble showed high R^2 values, with the combined ensemble achieving optimal performance through weighted averaging.

The second pathway focuses on spatial monitoring through satellite-driven approaches. This can be further divided into two strategies. One uses proxy-based estimation through co-emitted species, particularly NO_2 from TROPOMI/Sentinel-5P, taking advantage of the strong spatial correlation with fossil fuel CO_2 emissions. The other involves direct CO_2 measurement using dedicated sensors like OCO-2, OCO-3, and the upcoming CO2M. CNN architectures for NO_2 proxy estimation can achieve daily $R^2 \approx 0.89$ with explicit spatial mapping [3]. Vision transformers enable direct methane plume detection with F1 scores exceeding 0.8 and order-of-magnitude sensitivity improvements [5].

The proxy approach offers the advantage of daily monitoring using current satellite capabilities, though it requires careful validation of emission ratios under different fuel compositions. Direct CO_2 monitoring provides higher specificity but currently operates at lower temporal frequency and spatial coverage. These complementary approaches address different monitoring needs: temporal trend analysis versus spatially-explicit emission mapping.

C. Performance Benchmarks and Validation

Current AI-satellite integration demonstrates strong performance across multiple applications:

National-scale forecasting: Ensemble methods (FFNN+ANFIS+LSTM) achieve $R^2 = 0.989$ with 6.87% validation error for Saudi Arabia, with individual model R^2 values of 0.991 (FFNN), 0.989 (ANFIS), and 0.995 (LSTM) [4].

Projections suggest emissions reduction from 9.5 to 6.2 million tonnes annually during 2020-2030.

Spatial NO_2 mapping: CNN architectures reach daily $R^2 \approx 0.89$ and annual $R^2 \approx 0.95$, substantially outperforming multiple linear regression ($R^2 \approx 0.63$ daily, 0.76 annual) [3]. This represents substantial improvement in explained variance over traditional approaches, with increases of approximately 26 percentage points (daily) and 19 percentage points (annual).

Direct CO_2 quantification: CNN models trained on simulated CO2M data achieve 20% absolute relative error for power plant emission quantification, outperforming state-of-the-art cross-sectional flux methods with approximately half the absolute error [8]. The approach demonstrates resilience to cloud cover and partial plume visibility.

Methane detection: Vision transformers demonstrate order-of-magnitude sensitivity improvements, achieving $\text{F1} > 0.8$ for plumes down to 0.01 km^2 (200-300 kg CH_4/h) with pixel-wise false-positive rates $< 0.03\%$ [5].

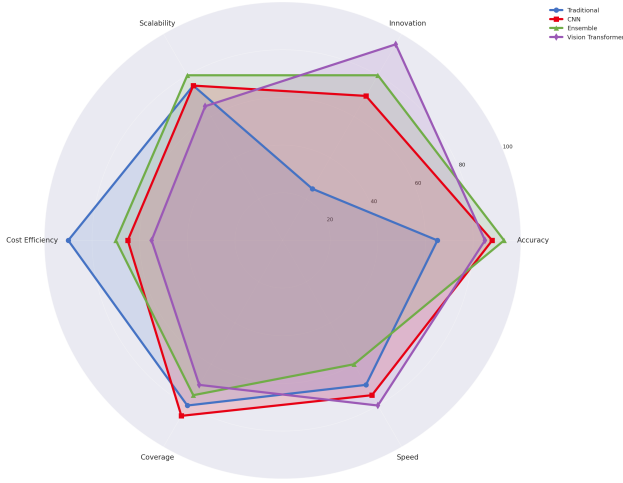


Fig. 2. Multi-criteria comparison of AI architectures (accuracy, cost-efficiency, speed, innovation).

A multi-criteria comparison across accuracy, cost-efficiency, speed, and innovation is summarized in Fig. 2.

D. Operational Monitoring Systems

GRACED Platform: Near-real-time global gridded daily CO₂ emissions at $0.1^\circ \times 0.1^\circ$ spatial resolution with $\pm 19.9\%$ uncertainty, integrating Carbon Monitor activity data with TROPOMI NO₂ observations [6]. The system provides comprehensive sectoral coverage (power, industry, residential, transportation, aviation, shipping, cement) enabling policy-relevant monitoring.

Cloud Processing Capabilities: Google Earth Engine provides access to >90 PB satellite archives with $6\times$ processing speedup compared to desktop environments, enabling planetary-scale analysis [7]. This infrastructure breakthrough democratizes access to computational resources previously available only to major institutions.

Validation Performance: Large-scale implementations demonstrate operational readiness.

Venkatappa et al. achieved 92.1% validation accuracy for forest carbon mapping using Google Earth Engine, covering >1 million hectares [14].

Similar validation rates across diverse applications confirm scalability.

E. Remote Sensing for Carbon Stocks and Land Systems

Satellite-based soil organic carbon (SOC) and biomass mapping inform carbon accounting in arid/agricultural contexts relevant to Saudi Arabia.

Applications include SOC estimation in Saudi fields using Landsat-8 and Sentinel-2 [11], stacked-ensemble SOC prediction in Iranian drylands [12], and long-term tillage impacts under hyper-arid conditions [13].

Additional work includes phenology-driven forest carbon stock mapping on Google Earth Engine [14]. Deep multi-sensor

learning achieved $R^2 \approx 0.85$ for aboveground forest carbon density [10].

F. Technical Challenges and Regional Adaptations

Several technical challenges continue to limit the effectiveness of satellite-based carbon monitoring, particularly in arid regions like the Arabian Peninsula.

Atmospheric interference poses a significant problem. Multispectral satellite data suffers from considerable noise, and this issue becomes more pronounced in arid regions where detection sensitivity is already limited.

Vision transformer approaches show some promise for addressing noise reduction through attention mechanisms that can automatically identify relevant spatial patterns [5].

The validation infrastructure presents another challenge, especially when compared to regions with dense ground monitoring networks. The Arabian Peninsula requires establishing dedicated validation systems that can account for several unique factors.

Key considerations include: - Extreme temperature variations that can affect sensor performance - Sandstorm frequency that impacts satellite visibility (some regions experience over 50 dust storm days annually) - Industrial emission complexity from oil and gas operations with varying fuel compositions - Limited opportunities for vegetation-based carbon sink validation in arid environments

Current operational systems like GRACED can provide daily monitoring at 0.1° resolution with $\pm 19.9\%$ uncertainty, but sub-daily monitoring remains challenging for tracking diurnal emission patterns from large point sources [6].

G. Saudi and Regional Case Studies

Saudi Arabia offers both unique opportunities and distinct challenges for implementing AI-satellite carbon monitoring systems.

National-scale forecasting using ensemble AI methods (combining FFNN, ANFIS, and LSTM) has already demonstrated $R^2 = 0.989$ with 6.87% validation error when applied to Saudi data. These projections suggest the possibility of emissions reduction from 9.5 to 6.2 million tonnes annually during 2020-2030 [4].

Several regional studies have established important baselines. Urban expansion monitoring in Al-Khobar using multi-temporal Landsat classification has achieved high accuracy for tracking sprawl patterns [15]. Carbon stock assessment work in Abha, using InVEST modeling combined with satellite imagery, has enabled scenario-based planning for urban development [16]. Urban heat island analysis in Jeddah through machine learning approaches applied to thermal imagery has provided insights relevant for policy makers [17].

Regional air-quality pattern analysis over the broader MENA region using multi-satellite integration helps complement these city-level studies [18].

Buildings sector applications suggest that AI methods can improve emission prediction performance by roughly 20% compared to traditional models [19].

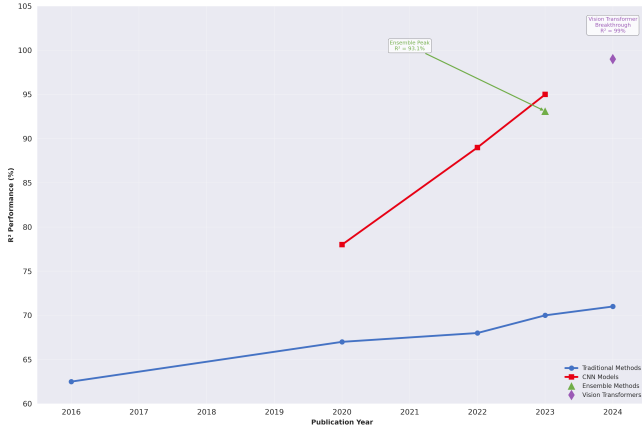


Fig. 3. Evolution of AI model performance and maturity for carbon monitoring (authors' compilation).

Agriculture-focused AI initiatives, such as AIFARMS, point toward potential indirect emission reductions through optimized farming practices [20].

The arid environment presents particular calibration challenges. Frequent dust storms, high solar irradiance, and sparse vegetation coverage require specialized validation approaches that have been adapted specifically for Gulf conditions.

H. Synthesis and Performance Evolution

Fig. 3 summarizes performance evolution across three tiers: Tier 1 methods ($R^2 > 0.90$) include ensembles and advanced CNNs [3], [4]; Tier 2 ($R^2 0.85\text{--}0.90$) includes multi-sensor CNNs [10]; and Tier 3 ($0.70\text{--}0.85$) covers single-sensor ML and classical baselines. Transformers deliver breakthrough sensitivity for gas plumes [5], operational systems reach daily grids [6], and cloud-scale processing [7] underpins scalability for nationwide deployment.

The distribution of reported model performance metrics (2016-2024) is shown in Fig. 4, demonstrating consistent improvement over time.

IV. TABLES AND PLATFORMS

V. SAUDI ARABIA IMPLEMENTATION FRAMEWORK

A. Current Capabilities and Gaps

Saudi Arabia has already demonstrated significant forecasting capabilities, achieving $R^2 = 0.989$ with 6.87% validation error for national emissions trends [4]. However, the kingdom currently lacks operational spatial monitoring infrastructure needed for detailed emission mapping. Given Saudi Arabia's Vision 2030 commitment to achieve net-zero by 2060 and reach 50% renewable energy, developing real-time emission tracking capabilities has become increasingly important for monitoring progress and informing policy decisions.

Several gaps remain that need to be addressed. There is currently no high-resolution spatial monitoring system in place. Ground-truth validation infrastructure adapted to arid conditions is lacking. Integration between national forecasting capabilities

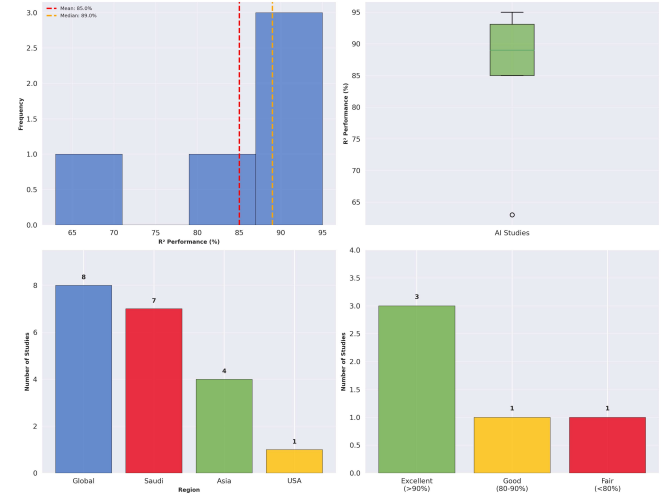


Fig. 4. Distribution of reported performance metrics for AI-based carbon monitoring methods (2016-2024).

and local emission quantification has not been established. Sub-daily monitoring for large point sources, which would be particularly valuable given the scale of oil and gas operations, is not yet available.

B. Environmental and Technical Adaptations for Arid Conditions

The Arabian Peninsula presents unique monitoring challenges that require specialized technical adaptations. Atmospheric dust loading significantly exceeds global averages, with some regions experiencing over 50 dust storm days annually. These conditions necessitate robust filtering algorithms and multi-sensor redundancy approaches.

Temperature extremes pose additional calibration challenges. Diurnal temperature variations can exceed 15°C , while seasonal variations span over 40°C between summer and winter conditions. Satellite sensor performance degrades under extreme thermal conditions, requiring sophisticated temperature compensation algorithms.

The sparse vegetation coverage limits opportunities for traditional carbon sink validation approaches used in temperate regions. Alternative validation strategies must rely more heavily on industrial emission monitoring and atmospheric transport modeling. The complex oil and gas infrastructure creates multi-source emission patterns that require advanced source separation techniques, potentially leveraging transformer-based architectures demonstrated in recent methane detection studies [5].

C. Technical Implementation Priorities

The first phase of implementation (2025-2026) should focus on establishing ground truth systems. This involves deploying calibrated monitoring networks that can account for frequent dust storms and extreme temperature variations. Validating NO_2 -to- CO_2 emission ratios under regional fuel composition and industrial process conditions will be essential. Baseline

TABLE I
AI-SATELLITE INTEGRATION STUDIES: PERFORMANCE METRICS

Study	Year	AI Model	Region	Performance
Nassef et al.	2023	FFNN+ANFIS+LSTM	Saudi Arabia	$R^2 = 0.989$; 6.87% validation error
Cao	2023	Custom CNN	U.S.	Daily R^2 : 0.89; Annual: 0.95
Zhang et al.	2022	CNN Multi-sensor	E. China	$R^2 \approx 0.85$
Rouet-Leduc	2024	Vision Transformer	Global	$F1 > 0.8$; order-of-magnitude sensitivity
Le Brazidec	2024	CNN regression	CO2M sim.	20% absolute relative error
Couture et al.	2024	Ensemble	Global plants	RMSE: 2.18 Mt CO ₂

TABLE II
INNOVATIONS AND LIMITATIONS IN REPRESENTATIVE STUDIES

Study	Primary Innovation	Key Limitations
Nassef et al.	Saudi-specific AI ensemble with $R^2 = 0.989$	No spatial NO ₂ mapping capability
Cao	CNN for atmospheric monitoring with daily coverage	NO ₂ proxy, not direct CO ₂
Zhang et al.	Multi-sensor integration approach	Forest ecosystems only
Rouet-Leduc & Hulbert	Order-of-magnitude sensitivity improvement	Methane-focused task
Dumont Le Brazidec et al.	Direct CO ₂ quantification (simulated)	Requires real-image validation
Couture et al.	Operational global system	Water-vapor plume dependency

measurements should be established for oil and gas facilities, major urban centers like Riyadh, Jeddah, and Dammam, as well as industrial complexes. Integrating dust-storm early warning systems with monitoring infrastructure will also be important.

The second phase (2026-2028) could focus on operational deployment. This would involve implementing GRACED-style daily monitoring systems adapted to arid conditions with appropriate local calibration. Deploying CNN architectures that can achieve daily R^2 values around 0.89 for urban emission mapping represents a key milestone. Integrating vision transformer approaches for sensitive point-source detection, particularly for oil and gas facilities, should be prioritized. Establishing monitoring coverage at $0.1^\circ \times 0.1^\circ$ spatial resolution for major emission sources will provide the foundation for detailed tracking.

The third phase (2028-2030) should emphasize integration and scaling. This involves connecting the high-performance national forecasting capabilities with spatial monitoring systems to provide comprehensive coverage. Developing specific tracking indicators for Vision 2030 progress, including renewable transition metrics and circular carbon economy indicators, will support policy implementation. Establishing GCC regional monitoring standards and sharing methodological advances could enhance regional cooperation. Implementing sub-daily monitoring for the largest point sources using transformer architectures would represent an advanced capability.

D. Regional Coordination and Capacity Building

Successful implementation requires coordination across multiple Saudi institutions and potential regional partnerships within the Gulf Cooperation Council (GCC). The Saudi Space Commission's growing satellite capabilities provide opportunities for indigenous monitoring infrastructure development. Collaboration with KAUST and other research institutions can support the advanced algorithm development needed for arid environment adaptations.

Regional knowledge sharing could leverage the diverse environmental conditions across GCC countries. The UAE's operational experience with urban monitoring systems, Qatar's experience with industrial emission tracking, and Kuwait's work in dust storm prediction could inform comprehensive regional monitoring approaches.

Capacity building programs should emphasize both technical training and policy integration. Training programs for satellite data analysis, AI model development, and uncertainty quantification will be essential. Policy integration training should focus on connecting monitoring outputs with Vision 2030 reporting requirements and international climate commitments.

E. Research Priorities and Technical Requirements

Several research areas require focused attention to ensure successful implementation in Saudi Arabia's unique environmental context.

Arid environment calibration represents a critical need. Understanding how dust storms impact detection accuracy and developing appropriate correction algorithms will be essential. Seasonal calibration protocols need to account for temperature variations that can exceed 40°C between summer and winter conditions.

Uncertainty quantification deserves special attention. Implementing Bayesian deep learning approaches to provide confidence intervals with predictions could help achieve target uncertainty levels around $\pm 15\%$ for policy-relevant emission estimates.

Advancing from daily to hourly monitoring for large point sources (those emitting more than 1 Mt CO₂ annually) using transformer architectures represents an important technical goal. This would be particularly valuable for the oil and gas sector.

TABLE III
SATELLITE PLATFORM SPECIFICATIONS FOR CARBON MONITORING

Platform	Spatial Res.	Temporal Res.	Spectral Characteristics	Access
Sentinel-2	10–60 m	2–5 days	13 bands (443–2190 nm)	Open (ESA)
Sentinel-1	5–40 m	6–12 days	C-band SAR	Open (ESA)
Landsat-8	15–100 m	16 days	11 bands (430–12500 nm)	Open (USGS)
TROPOMI/SSP	7×3.5 km	Daily	UV–SWIR hyperspectral	Open (ESA)
CO2M	2×2 km	2–6 days	CO ₂ -optimized NIR–SWIR	Planned 2026
ALOS-2	1–100 m	14 days	L-band SAR	Commercial
MODIS	250 m–1 km	Daily–16 d	36 bands (405–14385 nm)	Open (NASA)
OCO-2	1.3×2.25 km	16-day rep.	CO ₂ -sensitive NIR bands	Open (NASA)

TABLE IV
CLOUD COMPUTING & DATA ACCESS CAPABILITIES

Platform/System	Data Volume	Volume	Processing Speed	Key Advantages
Google Earth Engine	> 90 PB	~6×	desktop	Open access, collaboration
GRACED	Multi-source	Daily	updates	Near real-time, $\pm 19.9\%$ uncertainty
Carbon Monitor	National aggregations	Daily		Policy-relevant timing
Venkatappa et al. (GEE)	> 1M ha		Large-scale	92.1% validation accuracy

TABLE V
SAUDI AND REGIONAL CASE STUDIES

Location	Year	Methodology	Key Findings
Abha (KSA)	2022	InVEST + multi-temporal Landsat	Carbon stocks/sequestration scenarios
Al-Khobar (KSA)	2016	Multi-temporal Landsat classification	High-accuracy urban sprawl detection
Jeddah (KSA)	2023	ML + Thermal imagery	Urban heat island mapping
Saudi Arabia (Nat.)	2023	FFNN+ANFIS+LSTM	$R^2 = 0.989$, 6.87% validation error
Hyper-arid (Agri.)	2019	GIS + long-term satellite analysis	Tillage impacts on soil carbon
MENA (Regional)	2019	Multi-satellite integration	Regional air-quality patterns

Multi-sensor fusion capabilities should integrate CO₂, NO₂, and CH₄ observations to provide comprehensive emission characterization. This is especially important for complex oil and gas operations that involve multiple emission pathways.

Regional adaptation through transfer learning approaches could help adapt global models to Gulf conditions, including sparse vegetation and high atmospheric dust loading.

F. Vision 2030 Integration Framework

Key performance indicators should include annual emission reduction tracking with quarterly updates, renewable energy sector growth monitoring (targeting 50% by 2030), circular carbon economy progress assessment, and point-source emission compliance verification.

Policy support mechanisms should provide real-time data for carbon pricing and trading systems, enable evidence-based pol-

TABLE VI
REGIONAL CHALLENGES AND AI SOLUTIONS FOR CARBON MONITORING IN ARID ENVIRONMENTS

Challenge	Impact on Monitoring	AI-Based Solutions
Frequent dust storms	Reduced satellite visibility, >50 days annually	Multi-sensor fusion, cloud masking algorithms
High solar irradiance	Sensor saturation, calibration drift	Adaptive calibration, attention mechanisms
Sparse vegetation	Limited CO ₂ sink validation	Synthetic plume training, transfer learning
Complex infrastructure	Multi-source emission patterns	Transformer-based source separation
Oil/gas operations	Sensor performance degradation	Robust training, uncertainty quantification

icy adjustment capabilities, support international reporting and verification requirements, and allow for automated industrial compliance monitoring.

VI. CONCLUSIONS

Scientific Achievements: Deep learning has delivered transformative capabilities for satellite-based carbon monitoring, with CNN architectures achieving $R^2 > 0.95$ for NO₂ proxy estimation and 20% absolute relative error for direct CO₂ quantification from simulated imagery. Vision transformers provide order-of-magnitude sensitivity improvements for methane plume detection, while ensemble methods demonstrate $R^2 > 0.93$ for national emissions forecasting. These advances represent a maturation of AI-satellite integration for environmental monitoring.

Operational Readiness: Current technologies demonstrate sufficient maturity for operational deployment. Ensemble methods achieve $R^2 > 0.93$ for national trends, spatial monitoring reaches daily $R^2 \approx 0.89$, and cloud platforms provide >90 PB data access with 6× processing acceleration. Near-real-time systems like GRACED deliver daily global monitoring at 0.1° resolution with quantified uncertainty ($\pm 19.9\%$).

Saudi Arabia Implementation Pathway: Success requires a three-phase approach: (1) arid-environment ground truth establishment with dust storm adaptation; (2) CNN-based spatial monitoring deployment achieving demonstrated performance levels; and (3) integration with Vision 2030 reporting frameworks for policy support.

Critical technical priorities include dust storm impact quantification, uncertainty assessment protocols, and sub-daily monitoring advancement for large point sources.

Regional and Global Implications: These methodologies, when validated and adapted for regional conditions, can accelerate global emission monitoring capabilities.

The Saudi implementation framework provides a template for other arid regions facing similar challenges. Success in the Gulf region would demonstrate scalability to diverse environmental conditions, supporting evidence-based climate policy and international accountability mechanisms worldwide.

A. Research Gaps Requiring Investigation

Several critical research areas require immediate attention to advance the field and support practical implementation:

Validation Infrastructure Development: Current validation approaches rely heavily on temperate-region methodologies that may not translate effectively to arid environments. Research is needed to develop specialized validation protocols for regions with sparse vegetation, frequent dust storms, and extreme temperature variations. This includes developing synthetic validation approaches using atmospheric transport models and establishing mobile monitoring networks for cross-validation.

Uncertainty Quantification in Extreme Conditions: While current systems report uncertainty levels around $\pm 19.9\%$, the sources and propagation of uncertainty in extreme environmental conditions remain poorly characterized. Research should focus on Bayesian approaches to uncertainty quantification that can account for environmental variability, sensor degradation, and model limitations specific to arid regions.

Transfer Learning for Data-Sparse Regions: Most AI models are trained on data from regions with dense monitoring networks. Developing effective transfer learning approaches for regions with limited ground truth data represents a critical research need. This is particularly relevant for Gulf countries where monitoring infrastructure is still developing.

Integration of Natural and Anthropogenic Fluxes: Current approaches focus primarily on anthropogenic emissions, but comprehensive carbon accounting requires integration of natural carbon fluxes. In arid systems, these may include soil carbon dynamics under extreme conditions, vegetation carbon cycling, and geological carbon sources.

B. Future Technological Opportunities

The convergence of advancing AI capabilities, expanding satellite constellation coverage, and increasing computational resources positions the field for continued rapid advancement. Several technological developments will likely drive the next generation of monitoring capabilities:

Next-Generation Satellite Constellations: The planned CO2M constellation launch in 2026 will provide dedicated CO₂ monitoring with 2×2 km spatial resolution and 2-6 day temporal frequency. This represents a significant improvement over current proxy-based approaches and will enable direct validation of AI-driven emission estimates.

Edge Computing and Real-Time Processing: Advances in edge computing technologies will enable real-time processing of satellite data, reducing latency from hours to minutes for critical monitoring applications. This is particularly important for rapid response to emission events and integration with policy decision-making processes.

Federated Learning Approaches: Privacy-preserving federated learning approaches could enable international collaboration on model development while maintaining data sovereignty. This is particularly relevant for sensitive industrial emission data that countries may be reluctant to share directly.

Hybrid Physics-AI Models: Integration of physical atmospheric models with AI approaches could provide more robust predictions and better uncertainty characterization. These hybrid approaches could leverage the strengths of both physics-based understanding and data-driven pattern recognition.

C. Policy and Implementation Recommendations

Standardization and Interoperability: International standardization of monitoring protocols, uncertainty reporting, and data formats will be essential for enabling global carbon accounting systems. Saudi Arabia's implementation experience could contribute to developing these standards, particularly for arid region monitoring.

Capacity Building and Technology Transfer: Successful implementation requires substantial investment in human capacity building and technology transfer. This includes training programs for satellite data analysis, AI model development, and integration with policy processes.

Public-Private Partnerships: The scale and complexity of comprehensive monitoring systems require coordination between government agencies, research institutions, and private sector capabilities. Saudi Arabia's Vision 2030 framework provides an opportunity to demonstrate effective public-private collaboration models.

Future Outlook: Saudi Arabia's implementation experience will provide valuable insights for global carbon monitoring system development and climate policy effectiveness assessment. The lessons learned from deploying AI-satellite monitoring systems in challenging arid environments will inform similar efforts worldwide, contributing to global climate action and accountability mechanisms.

REFERENCES

- [1] IPCC, "Climate Change 2023: Synthesis Report. Summary for Policy-makers," 2023. [Online]. Available: <https://www.ipcc.ch/report/ar6/syr/>
- [2] Saudi Green Initiative (SGI), "Saudi Arabia to achieve net zero emissions by 2060," 2021. [Online]. Available: <https://www.saudigreeninitiative.org>
- [3] E. L. Cao, "National ground-level NO₂ predictions via satellite imagery driven convolutional neural networks," *Frontiers in Environmental Science*, vol. 11, art. 1285471, 2023. doi: 10.3389/fenvs.2023.1285471.
- [4] A. M. Nassef *et al.*, "Application of artificial intelligence to predict CO₂ emissions: Critical step towards sustainable environment," *Sustainability*, vol. 15, no. 9, 7648, 2023. doi: 10.3390/su15097648.
- [5] B. Rouet-Leduc and C. Hulbert, "Automatic detection of methane emissions in multispectral satellite imagery using a vision transformer," *Nature Communications*, vol. 15, 3801, 2024. doi: 10.1038/s41467-024-47754-y.
- [6] X. Dou *et al.*, "Near-real-time global gridded daily CO₂ emissions 2021," *Scientific Data*, vol. 10, 69, 2023. doi: 10.1038/s41597-023-01963-0.

- [7] D. D. Wilson, G. W. Tefera, and R. L. Ray, "Application of Google Earth Engine to monitor greenhouse gases: A review," *Data*, vol. 10, no. 1, 8, 2025. doi: 10.3390/data10010008.
- [8] J. D. Le Brazidec, P. Vanderbeeken, A. Farchi, G. Broquet, G. Kuhlmann, and M. Bocquet, "Deep learning applied to CO₂ power plant emissions quantification using simulated satellite images," *Geoscientific Model Development*, vol. 17, pp. 1995–2014, 2024. doi: 10.5194/gmd-17-1995-2024.
- [9] H. D. Couture *et al.*, "Estimating carbon dioxide emissions from power plant water vapor plumes using satellite imagery and machine learning," *Remote Sensing*, vol. 16, no. 7, 1290, 2024. doi: 10.3390/rs16071290.
- [10] F. Zhang, X. Tian, H. Zhang, and M. Jiang, "Estimation of above-ground carbon density of forests using deep learning and multi-source remote sensing," *Remote Sensing*, vol. 14, no. 13, 3022, 2022. doi: 10.3390/rs14133022.
- [11] R. Madugundu *et al.*, "Estimation of soil organic carbon in agricultural fields: A remote sensing approach," *Journal of Environmental Biology*, vol. 43, no. 1, pp. 73–84, 2022. doi: 10.22438/jeb/43/1/MRN-1873.
- [12] R. Taghizadeh-Mehrjardi *et al.*, "Improving the spatial prediction of soil organic carbon content... by stacking machine learning models," *Remote Sensing*, vol. 12, no. 7, 1095, 2020. doi: 10.3390/rs12071095.
- [13] E. Tola, K. Al-Gaadi, and R. Madugundu, "Employment of GIS techniques to assess the long-term impact of tillage on the soil organic carbon of agricultural fields under hyper-arid conditions," *PLOS ONE*, vol. 14, no. 2, e0212521, 2019. doi: 10.1371/journal.pone.0212521.
- [14] M. Venkatappa, N. Sasaki, S. Anantsuksomsri, and B. Smith, "Applications of the Google Earth Engine and phenology-based threshold classification method for mapping forest cover and carbon stock changes in Siem Reap Province, Cambodia," *Remote Sensing*, vol. 12, no. 18, 3110, 2020. doi: 10.3390/rs12183110.
- [15] M. T. Rahman, "Detection of land use/land cover changes and urban sprawl in Al-Khobar, Saudi Arabia: An analysis of multi-temporal remote sensing data," *ISPRS Int. J. Geo-Information*, vol. 5, no. 2, 15, 2016. doi: 10.3390/ijgi5020015.
- [16] J. Mallick, M. K. Almesfer, M. Alsubih, M. Ahmed, and N. Ben Kahla, "Estimating carbon stocks and sequestration with their valuation under a changing land use scenario: A multi-temporal research in Abha city, Saudi Arabia," *Frontiers in Ecology and Evolution*, vol. 10, 905799, 2022. doi: 10.3389/fevo.2022.905799.
- [17] A. Addas, "Machine learning techniques to map the impact of urban heat island: Investigating the city of Jeddah," *Land*, vol. 12, no. 6, 1159, 2023. doi: 10.3390/land12061159.
- [18] A. Elnadry *et al.*, "Urban health related air quality indicators over the Middle East and North Africa countries using multiple satellites and AERONET data," *Remote Sensing*, vol. 11, no. 18, 2096, 2019. doi: 10.3390/rs11182096.
- [19] J. Hua *et al.*, "Artificial intelligence for calculating and predicting building carbon emissions: A review," *Environmental Chemistry Letters*, 2025. doi: 10.1007/s10311-024-01799-z.
- [20] S. Adve *et al.*, "AIFARMS: Artificial intelligence for future agricultural resilience, management, and sustainability," *AI Magazine*, vol. 45, no. 1, pp. 81–98, 2024. doi: 10.1002/aaai.12152.