ISSN
2822-4450

AINTELIA SCIENCE NOTES

Volume 4, Issue 1
June 2025

2025
Volume: 4
Issue: 1

AIntelia

AIntelia Sciences Notes
Volume 4, Issue 1
June , 2025

Editor
İlyas Özer
Bandırma Onyedi Eylül University
Bandırma, Balıkesir, Turkey

Publisher's Note

AINTELIA Science Notes Journal (ASNJ) has been published under the institutional umbrella of AINTELIA Yapay Zeka Teknolojileri since its inception. As of Volume 4, Issue 1, the journal continues its publication activities under BAUDER — Scientific Research and Applications Association (Bilimsel Araştırmalar ve Uygulamalar Derneği), a non-profit research association based in Türkiye.

The journal's editorial standards, peer review processes, and open-access policy remain unchanged. ASNJ will continue to be indexed and archived through its established channels, with its digital platform now hosted at www.bauderpress.org.tr.

The editorial board thanks all authors, reviewers, and readers for their continued support.

Table of Contents

Microplastic Contamination Assessment in Urban Stormwater Runoff Using Hyperspectral Imaging and Machine Learning: A Case Study of the İzmir Metropolitan Drainage Network, Turkey

Zeynep Çelik[1], Irmak Yıldırım [2]
[1] Dokuz Eylül University, İzmir, Turkey, zeynep.celik@ogr.deu.edu.tr, ORCID: 0009-0000-6569-9762
[2] Dokuz Eylül University, İzmir, Turkey, irmak.yildirim@ogr.deu.edu.tr, ORCID: 0000-0003-3347-5195

Abstract

Microplastic (MP) pollution in urban stormwater systems represents a critical but undercharacterised environmental pathway, with stormwater runoff constituting a primary vector transporting terrestrially originating MPs to receiving water bodies. This paper presents a hyperspectral imaging and machine learning framework for rapid, non-destructive quantification and polymer-type classification of microplastics in stormwater samples collected from the İzmir Metropolitan Municipality drainage network, Turkey. A shortwave infrared (SWIR, 900–1700 nm) hyperspectral camera coupled to a conveyor-belt scanning platform acquires spectral data cubes from filtered stormwater membrane samples. A random forest (RF) classifier trained on spectral signatures of seven reference polymer types (polyethylene, polypropylene, polystyrene, polyethylene terephthalate, polyamide, polyvinyl chloride, and polycarbonate) achieves a mean classification accuracy of 91.3% and a mean F1-score of 0.893 on held-out test samples. Gradient-boosted tree (XGBoost) regression models trained on environmental and catchment-scale predictor variables — including impervious surface fraction, traffic density, and antecedent dry period — explain 76% of variance in MP particle concentration across 38 drainage sub-catchments ($R^2 = 0.76$, RMSE = 412 particles/L). SHAP analysis identifies impervious surface fraction and commercial land-use proportion as the dominant drivers of MP concentration. The framework enables basin-scale MP load mapping at operational throughput (80 samples/day) and provides decision-support outputs for targeted source-control intervention.

Keywords: *microplastics; hyperspectral imaging; random forest; polymer classification; water quality*

1. Introduction

Microplastics (≤5 mm) have been identified as pervasive contaminants in freshwater, marine, and terrestrial environments, with documented adverse effects on aquatic biota through physical ingestion, chemical toxicity, and trophic transfer [1]. Urban environments are primary sources of microplastic generation, with stormwater drainage networks serving as critical transport pathways that mobilise MPs from streets, car parks, and waste infrastructure to receiving water bodies during precipitation events [2]. The concentration and polymer composition of MPs in urban stormwater are highly variable, reflecting heterogeneous land-use patterns, traffic intensity, and the diversity of plastic-containing products prevalent in modern urban areas.

Conventional microplastic characterisation methods — Fourier-transform infrared spectroscopy (FT-IR) and Raman spectroscopy — provide definitive polymer identification but are labour-intensive, low-throughput, and incompatible with the sample volumes required for catchment-scale monitoring programmes. Hyperspectral imaging (HSI) in the shortwave infrared (SWIR) spectral range has emerged as a promising high-throughput alternative, exploiting the characteristic C-H, C=O, and C-C bond absorption features that distinguish polymer types while enabling spatial mapping of particle distribution on filter membranes [3]. Machine learning classifiers applied to HSI spectral signatures have demonstrated accuracies exceeding 90% for reference polymer libraries under controlled conditions [3].

The İzmir Metropolitan Municipality operates a combined sewer and stormwater drainage network serving 4.4 million residents across a catchment of 11,973 km². Coastal receiving water bodies — including İzmir Bay, classified as a sensitive receiving water under the EU Water Framework Directive — are subject to significant stormwater-derived MP loading, but systematic catchment-scale characterisation has not been conducted. This paper presents the deployment of a SWIR-HSI and ML framework for MP assessment across 38 sub-catchments of the İzmir drainage network, quantifying spatial MP concentration patterns and identifying the catchment-level drivers of MP loading using SHAP-interpreted XGBoost models [4].

2. Literature Review

2.1 Microplastics in Urban Stormwater

The global dataset of microplastics in urban stormwater (DURUM, 2018–2024) compiled by Ziajahromi et al. [2] from 24 published studies across four continents documents median MP concentrations of 80–1,200 particles/L in stormwater runoff, with peak concentrations following the first flush of antecedent dry period events exceeding 5,000 particles/L. Tire wear particles and fibres are the dominant morphological categories across all land-use types. The machine learning assessment of Serry et al. [4] applied gradient-boosted decision trees (GBDT) models to the DURUM dataset, identifying impervious surface fraction, catchment area, and vehicle traffic as the three principal drivers of MP concentration, and demonstrating that XGBoost models achieve $R^2 = 0.71$ for inter-catchment concentration prediction.

2.2 Hyperspectral Imaging for Microplastic Classification

Ai et al. [3] demonstrated the application of SWIR hyperspectral imaging (900–1700 nm) to rapid identification of microplastics in farmland soil, combining principal component analysis (PCA) for dimensionality reduction with a support vector machine (SVM) classifier to achieve 92.4% polymer identification accuracy for six reference polymers. Valls-Conesa et al. combined FT-IR hyperspectral imaging with random forest algorithms for microplastic classification in environmental matrices, establishing that spectral ranges spanning the 1,700–1,800 nm and 2,200–2,300 nm windows provide the highest inter-polymer discriminability. The automated deep learning classification of scanning electron micrograph images of microplastics reported by Shi et al. [5] achieved 88% classification accuracy across polymer types, establishing a benchmark for comparison with HSI-based approaches.

2.3 Research Gaps

Integration of high-throughput HSI polymer classification with catchment-scale ML concentration modelling in a single operational monitoring framework has not been reported for the Mediterranean urban stormwater context. The present work addresses this gap, providing both a technology demonstration and catchment-scale driver analysis applicable to coastal Turkish and broader Mediterranean cities facing stormwater quality management obligations under the EU Water Framework Directive alignment process.

3. Methodology

3.1 Study Area and Sampling Design

Stormwater samples were collected from 38 selected sub-catchments of the İzmir metropolitan drainage network during four storm events in the autumn 2023 wet season (October–December), representing a range of antecedent dry periods (2–12 days), precipitation intensities (3.2–28.7 mm/h), and land-use types (residential, commercial, industrial, mixed). Samples were collected at sub-catchment outfalls using stainless-steel bucket samplers during the rising limb of the hydrograph. Aliquots of 10 L were transported to the laboratory on ice within 4 hours of collection and vacuum-filtered through 125 µm stainless-steel sieves followed by 20 µm polycarbonate membrane filters.

3.2 SWIR Hyperspectral Imaging System

Filtered membrane samples were air-dried and scanned using a Specim FX17 SWIR hyperspectral line-scan camera (900–1700 nm, 224 spectral bands, 2.7 nm spectral resolution) mounted on a motorised conveyor-belt scanning stage at a spatial resolution of 0.5 mm/pixel. Spectral pre-processing comprised dark/bright reference correction, standard normal variate (SNV) normalisation, and Savitzky-Golay first-derivative smoothing (window width 9 bands). A reference spectral library comprising 350 spectra per polymer type (7 polymers) was assembled from characterised reference materials to train and validate the RF classifier.

3.3 Random Forest Classifier and XGBoost Concentration Model

The RF classifier (500 trees, max depth 25, Gini impurity) was trained on the reference spectral library with 5-fold stratified cross-validation. Environmental concentration models used XGBoost regression with 18 catchment-scale predictor variables extracted from GIS layers (land use, traffic counts, imperviousness, population density, antecedent dry period, rainfall intensity, catchment area and slope). SHAP (SHapley Additive exPlanations) analysis quantified the marginal contribution of each predictor to model output, providing interpretable driver rankings consistent with the approach of Serry et al. [4].

4. Results and Discussion

4.1 Polymer Classification Performance

The RF classifier achieves mean accuracy of 91.3% and mean F1-score of 0.893 on the held-out test set across all seven polymer types. Per-polymer F1-scores range from 0.943 (polyethylene, distinctive 1,730 nm absorption) to 0.821 (polyamide, overlapping spectral features with polypropylene in the 1,560–1,600 nm range). Classification throughput is 80 samples/day at the conveyor-belt scanning speed, representing a 25× throughput improvement over manual FT-IR analysis. The most abundant polymer types in İzmir stormwater are polyethylene (31.4%), polypropylene (22.7%), and polystyrene (15.2%), consistent with global urban stormwater polymer distributions reported in the DURUM dataset [2].

4.2 Catchment-Scale Concentration Modelling

The XGBoost concentration model achieves $R^2 = 0.76$ and RMSE = 412 particles/L on the test set across 38 sub-catchments. SHAP analysis identifies impervious surface fraction (mean $|SHAP| = 0.38$) and commercial land-use proportion (0.29) as the two dominant predictors of MP concentration, followed by antecedent dry period (0.21) and heavy vehicle traffic density (0.17). Sub-catchments dominated by

commercial and transport land use in the central İzmir metropolitan area exhibit mean MP concentrations of 2,840 particles/L, approximately 4.7× the mean across residential sub-catchments (607 particles/L). The first-flush coefficient of variation is 2.3 across all catchments, confirming that antecedent dry period concentrations substantially exceed mean event concentrations.

4.3 Source Control Implications

The spatial MP load map derived from the XGBoost predictions identifies three high-priority sub-catchments (A3, B7, C12) in the Karsıyaka–Bornova commercial corridor as contributing 41% of total estimated annual MP load to İzmir Bay despite comprising only 18% of total drainage area. Implementation of gross pollutant traps (GPTs) with 100 µm mesh screens at the five highest-load outfalls is estimated, using the XGBoost load predictions, to reduce total İzmir Bay MP input by 28–34%, providing a cost-effective prioritisation basis for infrastructure investment within the İzmir Municipal Water and Sewerage Administration (\.IZSU) capital programme.

5. Discussion

5.1 Limitations and Uncertainty

The SWIR-RF classification system is trained on reference polymer spectra acquired under controlled laboratory conditions; field-weathered particles with surface oxidation or biofilm coating may exhibit spectral shifts that reduce classification accuracy. A validation campaign using FT-IR confirmation of 200 field-collected particles yielded a field classification accuracy of 84.6% — a reduction of 6.7 percentage points relative to laboratory performance, consistent with the weathering effect documented by Ai et al. [3]. The XGBoost concentration model explains 76% of variance, with the residual attributable primarily to small-scale spatial heterogeneity in local MP sources and intra-event concentration variability not captured by the event-integrated catchment predictors. Future work will incorporate event-scale dynamic predictors (antecedent rainfall pattern, runoff velocity) to improve model fidelity.

5.2 Scalability and Policy Application

The HSI + ML framework is designed for laboratory deployment at IZSU's existing water quality monitoring centre with modest infrastructure requirements: a Specim FX17 camera system (EUR 35,000–45,000), a conveyor scanning stage, and a workstation-class GPU for inference. The framework's spatial driver analysis output — XGBoost SHAP maps overlaid on municipal GIS — directly supports the source control prioritisation required under Turkey's Water Framework Directive alignment process, providing a replicable model for other Turkish coastal metropolitan areas facing similar receiving water quality obligations.

6. Conclusions

A SWIR hyperspectral imaging and machine learning framework for microplastic assessment in urban stormwater has been presented and validated across 38 sub-catchments of the İzmir metropolitan drainage network. Key findings are: (i) RF polymer classification achieves 91.3% accuracy and 0.893 F1-score in laboratory conditions, reducing to 84.6% under field-weathered particle conditions; (ii) XGBoost catchment concentration modelling explains 76% of inter-catchment variance with impervious surface fraction and commercial land-use as dominant SHAP predictors; (iii) spatial MP load mapping identifies three priority sub-catchments contributing 41% of estimated annual İzmir Bay MP load; and (iv) targeted GPT installation at five high-load outfalls is projected to reduce total bay MP input by 28–34%. Future work will extend the polymer library, incorporate event-scale predictors, and evaluate automated field-deployable HSI units for real-time stormwater monitoring.

REFERENCES

[1] Y. Xu, Q. He, Z. Li, C. Xu, X. Zhao, Z. Mao, and X. Shi, "Advances in machine learning for the detection and characterization of microplastics in the environment," Front. Environ. Sci., vol. 13, pp. 1573579, Mar. 2025.

[2] S. Ziajahromi, P. Kumar Neale, and F. D. Leusch, "Microplastics and tire wear particles in urban stormwater: Abundance, characteristics, and potential mitigation strategies," Environ. Sci. Technol., 2023.

[3] W. Ai, S. Liu, H. Liao, J. Du, Y. Cai, C. Liao, H. Shi, Y. Lin, M. Junaid, X. Yue, and J. Wang, "Application of hyperspectral imaging technology in the rapid identification of microplastics in farmland soil," Sci. Total Environ., vol. 807, pp. 151030, Feb. 2022.

[4] M. Serry, A. Dujardin, and C. Schwarz, "Assessment of environmental and socioeconomic drivers of urban stormwater microplastics using machine learning," npj Clean Water, 2025.

[5] B. Shi, M. Patel, D. Yu, J. Yan, Z. Li, D. Petriw, T. Pruyn, K. Smyth, E. Passeport, R. J. D. Miller, and J. Y. Howe, "Automatic quantification and classification of microplastics in scanning electron micrographs via deep learning," Sci. Total Environ., vol. 825, pp. 153903, Jun. 2022.

[6] M. Stuart, M. Davies, M. Hobbs, T. Pering, A. McGonigle, and J. Willmott, "High-resolution hyperspectral imaging using low-cost components: Application within environmental monitoring scenarios," Sensors, vol. 22, no. 12, pp. 4652, Jun. 2022.

[7] N. B. Beni, S. Karimifard, J. Gilley, T. Messer, A. Schmidt, and S. Bartelt-Hunt, "Higher concentrations of microplastics in runoff from biosolid-amended croplands than manure-amended croplands," Commun. Earth Environ., vol. 4, no. 1, pp. 42, 2023.

[8] X. Hu, C. Xie, Z. Fan, Q. Duan, D. Zhang, L. Jiang, X. Wei, D. Hong, G. Li, and X. Zeng, "Hyperspectral anomaly detection using deep learning: A review," Remote Sens., vol. 14, no. 8, pp. 1973, Apr. 2022.

[9] International Commission for the Protection of the Danube River, Danube River Basin Management Plan 2021 Update. Vienna: ICPDR, 2021.

[10] A. Turkyilmaz, M. Ozkan, and H. Engin, "Assessment of plastic pollution in Turkish coastal waters using Raman spectroscopy," Environ. Pollut., vol. 287, pp. 117629, Oct. 2021.

[11] B. B. Sezer, H. Turkmen, and U. Nuriyev, "PPFchain: A novel framework privacy-preserving blockchain-based federated learning method for sensor networks," Internet of Things, vol. 22, pp. 100781, Jul. 2023.

[12] K. Luo, M. Tang, C. Luo, and J. Guo, "Identification of microplastics in environmental matrices using hyperspectral imaging: Optimizing pre-processing and spectral wavelength selection," J. Hazard. Mater., vol. 421, pp. 126684, Jan. 2022.

[13] European Parliament, Directive 2000/60/EC Establishing a Framework for Community Action in the Field of Water Policy. Luxembourg: Official Journal of the European Communities, 2000.

[14] H. Arp, C. Bogdal, and K. Oetken, "Microplastic levels in stormwater show high variability between events and catchments," Sci. Total Environ., vol. 834, pp. 155166, Aug. 2022.

[15] M. Schwarzer, J. Brehm, M. Vollmer, J. Jasinski, C. Xu, S. Zainuddin, T. Fröhlich, M. Schott, A. Greiner, T. Scheibel, and C. Laforsch, "Shape, size, and polymer dependent effects of microplastics on Daphnia magna," J. Hazard. Mater., vol. 426, pp. 128136, 2022.

[16] P. Corcoran, C. Biesinger, and M. Grifi, "Plastics and beaches: A degrading relationship," Mar. Pollut. Bull., vol. 58, no. 1, pp. 80–84, 2009.

[17] R. Coyle, M. Service, U. Witte, G. Hardiman, and J. McKinley, "Modeling microplastic transport in the marine environment," ACS ES&T Water, vol. 3, no. 4, pp. 984–995, Apr. 2023.

Federated Learning for Privacy-Preserving Predictive Maintenance in Multi-Site Industrial IoT Networks: Architecture, Convergence Analysis, and Benchmark Evaluation

Dian Rezki Wijaya[1], Bambang Setiawan[2], Khairul Anuar[3]

[1] Department of Computer Science and Engineering, Universitas Bengkulu, Bengkulu 38371, Indonesia dian.w@eng.unib.ac.id, ORCID: 0000-0002-7064-275X

[2] Department of Computer Science and Engineering, Universitas Bengkulu, Bengkulu 38371, Indonesia, bambang@eng.unib.ac.id, ORCID: 0000-0003-1094-1114

[3] Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, 26300 Gambang, Kuantan, Pahang, Malaysia, khairulanuar@umpsa.edu.my, ORCID: 0000-0001-7006-6601

Abstract

Predictive maintenance (PdM) in multi-site industrial settings requires training machine learning models on condition monitoring data that is distributed across facilities operated by competing organisations, making centralised data aggregation infeasible due to privacy, intellectual property, and regulatory constraints. Federated learning (FL) offers a paradigm-shift solution by enabling collaborative model training without raw data exchange: each site trains locally and shares only model gradient updates with a central aggregation server. This paper presents FedPdM, a federated learning architecture tailored for industrial IoT predictive maintenance, incorporating: (i) FedAvg-based global aggregation with adaptive client weighting proportional to local dataset quality; (ii) a non-IID robustness mechanism using Scaffold variance reduction to address heterogeneous failure mode distributions across sites; and (iii) differential privacy (DP) noise injection at the gradient level to provide formal ε-DP guarantees. FedPdM is evaluated on the CMAPSS turbofan engine degradation benchmark and a proprietary 12-site industrial compressor dataset, achieving Remaining Useful Life (RUL) prediction RMSE of 14.3 cycles (CMAPSS FD001 subset) — comparable to centralised baseline RMSE of 13.8 cycles — while providing (ε=2.1, δ=10^{-5}) differential privacy at a communication overhead of 2.3 MB per round. Convergence is achieved in 38 rounds across 12 heterogeneous clients, confirming practical viability for real industrial multi-site deployment.

Keywords: *federated learning, predictive maintenance, industrial IoT,RUL prediction, Industry 4.0*

1. Introduction

Predictive maintenance — the use of condition monitoring data and machine learning models to forecast equipment failures before they occur — is among the highest-value applications of Artificial Intelligence in manufacturing and process industry [1]. Large-scale industrial datasets, spanning multiple machines, sites, and operating regimes, are essential for training robust PdM models that generalise across the variability of real operational conditions. However, in practice, condition monitoring data is fragmented across facilities operated by competing organisations or across business units with strict data governance policies, rendering centralised model training infeasible without compromising data sovereignty.

Federated learning (FL), introduced by McMahan et al. in 2017, addresses this challenge by enabling collaborative model training across distributed clients (sites) without requiring raw data to leave the local facility: each client trains a local model update on its own data and transmits only the gradient or model parameter update to a central aggregation server, which combines updates into an improved global model [2]. FL has been demonstrated in healthcare, mobile applications, and smart grid settings, but its application to multi-site industrial PdM — where data heterogeneity (non-IID distributions), communication constraints, and strong privacy requirements coexist — requires specialised architectural adaptations [3].

This paper presents FedPdM, a federated learning framework designed for industrial IoT predictive maintenance in multi-site settings. The framework integrates three key innovations: adaptive client weighting in the FedAvg aggregation, Scaffold variance reduction for non-IID robustness, and differential privacy gradient perturbation. Evaluation on the widely-used CMAPSS turbofan benchmark and a real 12-site industrial compressor dataset demonstrates that FedPdM achieves near-centralised RUL prediction accuracy while providing formal privacy guarantees and remaining within practical communication budgets.

2. Literature Review

The FedAvg algorithm [2] alternates between local gradient descent steps at each client and weighted averaging of model parameters at the server, achieving communication efficiency by reducing round trips relative to distributed SGD. Boobalan et al. [4] surveyed the fusion of federated learning and industrial IoT, documenting 47 application domains and identifying non-IID data distribution and heterogeneous client compute capacity as the two principal technical barriers to industrial deployment. Ahn et al. [5] applied FL to predictive maintenance with time-series IIoT data, demonstrating that federated models achieve predictive accuracy within 3% of centralised baselines across 12 heterogeneous manufacturing machines while maintaining data locality.

Non-identically distributed (non-IID) data across clients — arising from different machine types, operating profiles, and failure modes at each industrial site — causes FedAvg to diverge or converge to a suboptimal global model [6]. The Scaffold algorithm addresses client drift by introducing variance reduction control variates that correct the bias introduced by heterogeneous local updates, achieving convergence rates comparable to IID settings under mild assumptions. Differential privacy (DP) in federated learning, implemented by adding Gaussian noise to client gradient updates before transmission, provides (ε, δ)-DP guarantees that bound the information extractable about individual training samples from the shared model updates [3].

Existing FL-PdM studies have typically evaluated on single-domain benchmark datasets (CMAPSS, PHM08) under IID assumptions, without formal privacy analysis. The combination of non-IID robustness (Scaffold), adaptive weighting, and DP guarantees in a single framework evaluated on both benchmark and real multi-site industrial data has not been reported. FedPdM addresses this gap.

3. Methodology

FedPdM implements a three-tier hierarchy: edge computing nodes at each client site run local model training; a site aggregation tier performs intra-site gradient compression; and a central FL server performs FedAvg-based global aggregation. Each client maintains a local LSTM-based RUL prediction model (2 layers, hidden dim 128, dropout 0.2) trained on its local vibration, temperature, and pressure time-series data. Local training runs for $E = 5$ epochs per FL round using the Adam optimiser with learning rate 1×10^{-3}.

Client aggregation weights in FedAvg are set proportional to the estimated local data quality score $Q_k = n_k \times H_k$, where n_k is local dataset size and H_k is a data quality index (0–1) derived from label completeness and sensor availability metrics at site k. Scaffold control variates c_k are maintained per client and updated after each local training phase, correcting the gradient direction to compensate for client-specific distributional bias and reducing the inter-client gradient variance by 34–58% relative to vanilla FedAvg in heterogeneous conditions [6].

Gradient clipping (clip norm $L2 = 1.0$) followed by Gaussian noise addition ($\sigma = 0.8$) is applied to each client gradient update before transmission, implementing the DP-SGD mechanism. The Rényi differential privacy (RDP) accountant tracks privacy budget consumption across rounds, converting to (ε, δ)-DP via the moments accountant method. Gradient sparsification (top-k, k = 20% of parameters) and 8-bit quantisation reduce per-round communication volume from 11.4 MB (dense FP32) to 2.3 MB per client, enabling deployment over standard industrial 4G/LTE backhaul [4].

4. Results and Discussion

On the CMAPSS FD001 test subset (100 degradation trajectories, single fault mode), FedPdM achieves RUL prediction RMSE of 14.3 cycles and RMSE score of 312 (NASA scoring function weighting late predictions more severely). The centralised baseline achieves RMSE = 13.8 cycles. The performance gap of 0.5 cycles (3.6%) is attributable to the DP noise injection and the non-IID data splitting across 12 simulated clients; Scaffold variance reduction recovers 71% of the accuracy lost relative to non-private FedAvg. On the more heterogeneous FD003 subset (multiple fault modes), FedPdM achieves RMSE = 19.7 cycles vs. centralised RMSE = 18.9 cycles — a 4.2% gap — confirming robustness to multi-modal failure distributions.

On the proprietary 12-site industrial compressor dataset (bearing and seal degradation, 18-month operational records), FedPdM achieves mean absolute error (MAE) of 6.8 days for failure horizon prediction, compared to 6.1 days for the centralised model. Convergence is achieved in 38 FL rounds (approximately 76 minutes wall-clock time at 2-minute round cadence), with the Scaffold-enabled FedPdM converging 28% faster than vanilla FedAvg (53 rounds) in the non-IID client configuration. Privacy budget at convergence is $\varepsilon = 2.1$ with $\delta = 10^{-5}$, meeting the DP requirement specified by the industry consortium.

Total communication cost per client to convergence is 87.4 MB (38 rounds × 2.3 MB), achievable in approximately 12 minutes over a 1 Mbps uplink — within practical industrial 4G budgets. Local training compute per round requires 3.2 GPU-minutes on the NVIDIA Jetson Xavier NX (21 TOPS) edge unit at each client, confirming compatibility with edge computing hardware within current industrial deployments [4]. The privacy-utility-efficiency three-way trade-off analysis demonstrates that the FedPdM parameter configuration ($\sigma = 0.8$, k = 20%, E = 5) lies at the Pareto frontier for this problem class.

5. Discussion

The most significant challenge in the 12-site industrial deployment is the heterogeneity of failure mode distributions: four sites operate air-cooled compressors with predominant bearing failures, while the remaining eight operate water-cooled units with seal degradation as the primary failure mode. Without Scaffold, FedAvg converges to a model that performs adequately for the majority class (seal failures) but shows 23% higher MAE for bearing failure prediction. Scaffold reduces this class imbalance effect to 8%

through variance reduction, confirming the mechanism's value in real multi-fault-mode industrial settings [6].

Deployment of FedPdM in production industrial environments requires integration with existing SCADA and MES systems for data ingestion, and with IT security frameworks for gradient transmission authentication. The emerging FATE (Federated AI Technology Enabler) open-source platform provides a production-grade infrastructure compatible with the FedPdM aggregation protocol. Regulatory alignment with the EU AI Act (2024) — which classifies industrial PdM AI as limited-risk — and with GDPR Article 25 (data protection by design) is supported by the DP guarantees and by the data minimisation achieved through local training [7].

6. Conclusions

FedPdM, a federated learning framework for privacy-preserving predictive maintenance in multi-site industrial IoT networks, has been presented and evaluated. Key findings: (i) FedPdM achieves CMAPSS FD001 RUL prediction RMSE of 14.3 cycles (vs. 13.8 centralised), a 3.6% accuracy gap attributable to DP noise; (ii) Scaffold variance reduction enables convergence in 38 rounds for a 12-client non-IID configuration, 28% faster than vanilla FedAvg; (iii) (ε=2.1, δ=10^{-5})-DP guarantees are achieved within 2.3 MB/round communication budget; and (iv) real 12-site industrial compressor deployment confirms practical viability with 6.8-day MAE for failure horizon prediction. Future work will extend FedPdM to asynchronous client participation, personalised federated learning for site-specific model fine-tuning, and adversarial robustness against gradient poisoning attacks.

REFERENCES

[1] P. Boobalan, S. P. Ramu, Q. V. Pham, K. Dev, S. Pandya, P. K. R. Maddikunta, T. R. Gadekallu, and T. Huynh-The, "Fusion of federated learning and industrial Internet of Things: A survey," Comput. Netw., vol. 212, pp. 109048, Jul. 2022.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. AISTATS, pp. 1273–1282, 2017.

[3] W. Yang, Y. Yang, Z. Zhou, C. Ye, J. Han, Q. Liu, and L. Hanzo, "Adaptive optimization federated learning enabled digital twins in industrial IoT," J. Ind. Inf. Integr., vol. 41, pp. 100645, Sep. 2024.

[4] B. Farahani and A. K. Monsefi, "Smart and collaborative industrial IoT: A federated learning and data space approach," Digit. Commun. Netw., vol. 9, no. 1, pp. 23–41, Feb. 2023.

[5] J. Ahn, Y. Lee, N. Kim, C. Park, and J. Jeong, "Federated learning for predictive maintenance and anomaly detection using time series data distribution shifts in manufacturing processes," Sensors, vol. 23, no. 17, pp. 7331, Aug. 2023.

[6] X. Ma, J. Zhu, Z. Lin, S. Chen, and Y. Qin, "A state-of-the-art survey on solving non-IID data in federated learning," Future Gener. Comput. Syst., vol. 135, pp. 244–258, Oct. 2022.

[7] B. B. Sezer, H. Turkmen, and U. Nuriyev, "PPFchain: A novel framework privacy-preserving blockchain-based federated learning method for sensor networks," Internet of Things, vol. 22, pp. 100781, Jul. 2023.

[8] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, "Efficient and privacy-enhanced federated learning for industrial artificial intelligence," IEEE Trans. Ind. Inf., vol. 16, no. 10, pp. 6532–6542, Oct. 2020.

[9] A. R. Javed, M. Shahzad, S. ur Rehman, Y. B. Zikria, I. Razzak, Z. Jalil, and G. Xu, "Integration of blockchain technology and federated learning in vehicular (IoT) networks: A comprehensive survey," Sensors, vol. 22, no. 12, pp. 4394, Jun. 2022.

[10] K. Hazra, M. Adhikari, S. Nandy, K. Doulani, and V. G. Menon, "Federated-learning-aided next-generation edge networks for intelligent services," IEEE Netw., vol. 36, no. 3, pp. 56–64, May/Jun. 2022.

[11] M. Al-Quraan, L. Mohjazi, L. Bariah, A. Centeno, A. Zoha, K. Arshad, and M. A. Imran, "Edge-native intelligence for 6G communications driven by federated learning: A survey of trends and challenges," IEEE Trans. Emerg. Top. Comput. Intell., vol. 7, no. 3, pp. 957–979, Jun. 2023.

[12] P. Chhikara, R. Tekchandani, N. Kumar, M. Guizani, and M. M. Hassan, "Federated learning and autonomous UAVs for hazardous zone detection and AQI prediction in IoT environment," IEEE Internet Things J., vol. 8, no. 20, pp. 15456–15467, Jul. 2021.

[13] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: Fundamentals, enabling technologies, and future applications," Inf. Process. Manag., vol. 59, no. 6, pp. 103061, Nov. 2022.

[14] L. Shanmugam, R. Tillu, and M. Tomar, "Federated learning architecture: Design, implementation, and challenges in distributed AI systems," J. Knowl. Learn. Sci. Technol., vol. 2, no. 2, pp. 371–384, 2023.

[15] G. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "End-to-end privacy preserving deep learning on multi-institutional medical imaging," Nat. Mach. Intell., vol. 3, no. 6, pp. 473–484, Jun. 2021.

[16] Z. Liu, J. Guo, W. Yang, J. Fan, K. Y. Lam, and J. Zhao, "Privacy-preserving aggregation in federated learning: A survey," IEEE Trans. Big Data, vol. 14, no. 8, pp. 1–21, 2022.

[17] N. Rieke, J. Hancox, W. Li, F. Milletarì, H. R. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. A. Landman, K. Maier-Hein, S. Ourselin, M. J. Sheller, R. M. Summers, A. Trask, D. Xu, M. Baust, and M. J. Cardoso, "The future of digital health with federated learning," npj Digit. Med., vol. 3, no. 1, pp. 119, Sep. 2020.

AIntelia

Grid-Forming Inverter Control for Black-Start and Islanded Operation in Renewable-Dominated Rural Microgrids: Application to Off-Grid Communities in Sub-Saharan Africa

Amara Diallo[1], Nneka Okafor[2]

[1] Department of Electrical Engineering, Université Cheikh Anta Diop, Dakar, Senegal, a.diallo@ucad.edu.sn, ORCID: 0009-0015-4417-2963

[2] Department of Electrical and Electronic Engineering, University of Lagos, Lagos, Nigeria, n.okafor@unilag.edu.ng, ORCID: 0009-0016-8823-6174

Abstract

Sub-Saharan Africa (SSA) hosts approximately 600 million people without electricity access, and solar PV-battery mini-grids are the most cost-effective pathway for rural electrification in low-density communities beyond economic reach of grid extension. However, the overwhelmingly inverter-based generation composition of PV-battery mini-grids creates power system stability challenges: conventional grid-following (GFL) inverters require an external voltage and frequency reference that is absent during grid-off conditions, precluding black-start and stable islanded operation. Grid-forming (GFM) inverter control architectures — which independently establish and regulate voltage and frequency — are essential for reliable standalone mini-grid operation. This paper presents a droop-based GFM inverter control implementation with black-start capability, virtual impedance, and load-sharing coordination for a 48 kWp PV/100 kWh lithium-ion battery mini-grid supplying an off-grid community of 320 households in Tambacounda, Senegal. MATLAB/Simulink simulation and hardware-in-the-loop (HIL) validation on a CHIL test bench confirm: voltage magnitude regulation within ±1.5% (1σ), frequency deviation below ±0.3 Hz during cold-load pickup, black-start sequence completion in 380 ms, and reactive load sharing error below 8% between parallel inverters. The architecture provides a replicable technical baseline for GFM mini-grid deployments across SSA.

Keywords: *grid-forming inverter; islanded microgrid; droop control; PV-battery; virtual impedance*

1. Introduction

Access to electricity remains one of the most persistent development challenges in Sub-Saharan Africa, with an estimated 600 million people — approximately 80% of the global population without electricity access — residing in the region as of 2022 [1]. National grid extension remains economically infeasible for dispersed rural communities with low consumption density, and diesel generator-based mini-grids impose high and volatile fuel costs. Solar PV-battery mini-grids have emerged as the technically and economically preferred solution, with approximately 2,200 operational mini-grids in SSA as of 2020 and projections of 111 million household connections by 2030 [1].

PV-battery mini-grids are by nature dominated by inverter-based resources (IBRs), which present fundamentally different control and stability characteristics compared to synchronous generator-based systems. Conventional grid-following inverter control, which requires an external voltage and frequency reference for synchronisation via phase-locked loop (PLL), cannot operate in the absence of a grid — creating a bootstrapping problem for black-start and islanded operation. Grid-forming (GFM) inverter control addresses this limitation by generating an internal voltage and frequency reference independently, effectively emulating the behaviour of a voltage source rather than a current source [2].

Despite the theoretical advantages of GFM control for islanded mini-grid operation, practical implementation challenges — including load-sharing between parallel inverters, cold-load pickup voltage stability, and the specific constraints of low-cost embedded controllers prevalent in SSA mini-grid hardware — have limited published case studies of GFM deployment in real SSA community mini-grids. This paper addresses this gap by presenting a complete GFM control architecture implemented and validated for the Tambacounda, Senegal community mini-grid, providing a documented technical and implementation baseline for GFM deployment in similar SSA contexts.

2. Literature Review

The review of renewable off-grid mini-grids in SSA by Domegni and Azouma [1] documented the rapid growth of solar PV mini-grids, noting that lead-acid batteries remain dominant (66% of installed base) despite lower cycle life than lithium-ion alternatives, driven by lower upfront cost. Operational reliability — defined as the fraction of demand hours with voltage and frequency within acceptable limits — is the primary technical metric driving community acceptance of mini-grids, with field data from Senegal, Kenya, and Tanzania showing that voltage sag events during motor starting are the most frequent reliability complaint. The International Growth Centre [2] and IRENA identified GFM inverter capability as a key enabling technology for improving mini-grid reliability at high solar penetration.

GFM inverter control encompasses several architectures, of which droop control, virtual synchronous generator (VSG), and matching control are the most practically implemented. Droop control, which adjusts output frequency proportional to active power deviation (P-ω droop) and output voltage proportional to reactive power deviation (Q-V droop), is the most widely adopted architecture for multi-inverter mini-grids due to its simplicity and communication-free load sharing [3]. The analysis of GFM inverter controls for grid-connected and islanded microgrid integration confirmed that droop-based GFM controllers achieve black-start sequence completion within 0.2–0.5 s without voltage overshoot, and maintain stable operation during cold-load pickup of up to 120% rated load [3]. NREL's multi-microgrid black-start study [4] demonstrated coordinated GFM black start across four microgrids using GFM-enabled PV-battery inverters, confirming scalability beyond single-site application.

Published GFM mini-grid studies primarily address European, North American, or East Asian installations with reliable grid backup, sophisticated SCADA infrastructure, and high-capacity battery banks. The specific constraints of SSA rural deployments — limited battery capacity (1–3 h autonomy), high motor starting fractions due to agricultural loads, and the need for robust embedded controller implementation on low-cost hardware — have received limited systematic treatment. The present work addresses these constraints in the Tambacounda deployment context.

3. Methodology

The Tambacounda mini-grid supplies 320 households plus a community health centre, primary school, and grain mill in a rural community of approximately 1,850 residents in eastern Senegal. The installed generation and storage capacity is: 48 kWp monocrystalline PV (96 × 500 W modules in 12-string configuration); 100 kWh lithium-iron-phosphate (LFP) battery bank (50 kWh usable at 80% DoD); and three 15 kVA GFM bidirectional inverters connected in parallel on a 240 V single-phase AC bus. Peak demand is 22.4 kW (residential evening peak), with the grain mill motor (5.5 kW, single-phase induction) constituting the largest single load step at 155% starting current.

Each GFM inverter implements P-ω droop (droop coefficient $m_p = 2\times10^{-5}$ rad/s/W) and Q-V droop ($m_q = 1.5\times10^{-4}$ V/VAr) with a reference frequency of 50 Hz and reference voltage of 240 V. Virtual impedance ($R_v = 0.2\ \Omega$, $L_v = 0.5$ mH) is inserted in the output voltage reference calculation to emulate a resistive-inductive coupling impedance, improving reactive power sharing accuracy between parallel inverters with mismatched cable resistances. The inner voltage and current control loops are implemented as PI controllers in the synchronous rotating frame (d-q) with bandwidth 600 Hz (current) and 100 Hz (voltage), implemented on a Texas Instruments C2000 microcontroller at 10 kHz sampling rate.

The black-start sequence initiates from fully de-energised conditions: (i) Battery management system (BMS) powers on and confirms State of Charge (SoC) > 20%; (ii) Inverter 1 (master) initialises voltage reference at 240 V, 50 Hz and ramps output over 200 ms to avoid inrush; (iii) Critical loads (health centre, communications) are connected via automatic transfer switch (ATS) priority scheme; (iv) Inverters 2 and 3 (followers) synchronise to the master bus voltage via PLL and transition to GFM droop mode; (v) remaining loads connected sequentially with 50 ms inter-step delays to limit voltage dip accumulation. The entire sequence is designed for completion within 500 ms of BMS power-on [4].

4. Results and Discussion

HIL simulation confirms black-start sequence completion in 380 ms from BMS power-on to full community load connection, within the 500 ms design target. The critical load health centre ATS connects at t = 210 ms with a peak voltage transient of +8.3% (26.4 V above nominal) damped to within ±1% in 62 ms. The grain mill motor starting at t = 340 ms causes a voltage dip of 11.2% (26.9 V below nominal) recovering to nominal within 95 ms — within the EN 50160 voltage sag immunity range for Class 2 equipment. Frequency nadir during motor starting is 49.62 Hz, recovering to 50.0 ± 0.05 Hz within 180 ms.

Under full evening load (22.4 kW, PF 0.87 lagging), steady-state voltage magnitude is 239.6 V (99.8% nominal) at the AC bus with ±1.5% (1σ) spatial variation at household metering points across the 1.2 km LV distribution network. Reactive power sharing error between the three parallel inverters is 6.8% — reduced from 18.3% without virtual impedance, confirming that virtual impedance compensation substantially improves reactive load sharing for inverters with differing cable resistance paths [3]. Frequency deviation under normal load variations is within ±0.15 Hz, well within the ±0.5 Hz IEC 61727 mini-grid tolerance.

The droop-based GFM control is augmented with a secondary SoC-proportional frequency offset that adjusts the P-ω droop reference frequency between 49.8 Hz (SoC = 20%) and 50.2 Hz (SoC = 95%), communicating battery state to PV inverters that implement frequency-responsive curtailment above 50.15 Hz. This eliminates the need for explicit communication between the PV generation and battery inverter controllers for charge management, simplifying the control architecture and improving reliability in communication-degraded conditions common in rural SSA infrastructure environments.

5. Discussion

The GFM control modifications described in this paper require no additional hardware beyond that present in modern bidirectional inverters — only firmware updates to the embedded microcontroller

implementing the droop and virtual impedance algorithms. The C2000-based embedded implementation adds approximately EUR 12,000–18,000 in engineering cost per mini-grid deployment for firmware integration and commissioning, relative to a baseline GFL-controlled system. This cost is justified by the documented improvement in voltage reliability that drives community acceptance and productive-use load growth. Three additional community mini-grids in the Tambacounda region are planned for GFM deployment in 2025–2026, building on the architecture validated in this paper.

The HIL test bench uses single-phase 240 V representation; the extension to three-phase 400 V rural distribution networks prevalent in West Africa requires adaptation of the droop control for unbalanced load conditions. The SoC-frequency droop communication-free approach relies on all inverters observing the same bus frequency, which may be compromised by harmonic distortion in the presence of unfiltered non-linear loads. Future work will address three-phase unbalanced operation, evaluate virtual synchronous generator control as an alternative to droop for improved transient performance, and conduct long-term field monitoring of the Tambacounda installation.

6. Conclusions

A droop-based grid-forming inverter control architecture with black-start, virtual impedance, and SoC-frequency droop has been presented and validated for the 48 kWp PV/100 kWh battery Tambacounda community mini-grid. Key results: (i) black-start sequence completes in 380 ms with voltage transients within EN 50160 Class 2 immunity limits; (ii) steady-state voltage regulation is ±1.5% (1σ) and frequency deviation is ±0.15 Hz under nominal loading; (iii) reactive power sharing error is 6.8% with virtual impedance, reduced from 18.3% without; and (iv) SoC-frequency droop eliminates explicit communication requirements for charge management. The architecture provides a documented, replicable technical baseline for GFM mini-grid deployment across SSA, directly applicable to the 110 million household connections target by 2030.

REFERENCES

[1] T. Domegni and T. Azouma, "A review of renewable off-grid mini-grids in Sub-Saharan Africa," Front. Energy Res., vol. 10, pp. 1089025, Jan. 2023.

[2] International Growth Centre, "How microgrids can facilitate energy access and electrify rural Africa," IGC Blog, Oct. 2023.

[3] B. Ounissa and A. Berboucha, "Analysis of grid-forming inverter controls for grid-connected and islanded microgrid integration," Sustainability, vol. 16, no. 5, pp. 2148, 2024.

[4] E. Fix, A. Banerjee, U. Muenz, and G.-S. Seo, "Investigating multi-microgrid black start methods using grid-forming inverters," in Proc. IEEE ISGT NA, pp. 1–5, Washington, DC, 2023.

[5] B. Bahrani, M. H. Ravanji, B. Kroposki, D. Ramasubramanian, and X. Guillaud, "Grid-forming inverter-based resource research landscape," IEEE Power Energy Mag., vol. 22, no. 2, pp. 48–62, Mar./Apr. 2024.

[6] S. Ansari, A. Chandel, and M. Tariq, "A comprehensive review on power converters control and control strategies of AC/DC microgrid," IEEE Access, vol. 9, pp. 17998–18015, 2021.

[7] International Energy Agency, Africa Energy Outlook 2022. Paris: IEA, 2022.

[8] International Renewable Energy Agency, Renewable Power Generation Costs in 2022. Abu Dhabi: IRENA, 2023.

[9] D. Omoniwa, R. Hussain, M. A. Shah, S. A. Bhatti, S. Awan, and A. Ain, "Fog/edge computing-based IoT (FECIoT): Architecture, applications, and research issues," IEEE Internet Things J., vol. 6, no. 3, pp. 4118–4149, Jun. 2019.

[10] A. Akinyele, R. Munda, and A. Msunzu, "Grid-forming inverter for islanded and grid-connected modes in South African mini-grids," Energies, vol. 16, no. 4, pp. 1912, Feb. 2023.

[11] A. Ndiaye, A. Charki, A. Kobi, C. M. F. Kebe, P. A. Ndiaye, and V. Sambou, "Degradations of silicon photovoltaic modules: A literature review," Sol. Energy, vol. 96, pp. 140–151, Oct. 2013.

[12] S. Koohi-Fayegh and M. A. Rosen, "A review of energy storage types, applications and recent developments," J. Energy Storage, vol. 27, pp. 101047, Feb. 2020.

[13] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. de Vicuña, and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids — A general approach toward standardization," IEEE Trans. Ind. Electron., vol. 58, no. 1, pp. 158–172, Jan. 2011.

[14] Q. C. Zhong and G. Weiss, "Synchronverters: Inverters that mimic synchronous generators," IEEE Trans. Ind. Electron., vol. 58, no. 4, pp. 1259–1267, Apr. 2011.

[15] A. D. Paquette and D. M. Divan, "Virtual impedance current limiting for inverters in microgrids with synchronous generators," IEEE Trans. Ind. Appl., vol. 51, no. 2, pp. 1630–1638, Mar./Apr. 2015.

[16] IRENA, Grid Codes for Renewable Powered Systems. Abu Dhabi: International Renewable Energy Agency, 2022.

[17] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," IEEE Power Energy Mag., vol. 5, no. 4, pp. 78–94, Jul./Aug. 2007.

Life Cycle Assessment and Carbon Footprint Optimisation of Agrivoltaic Systems Combining Bifacial Solar Panels with Drip-Irrigated Vegetable Production in Semi-Arid Mediterranean Climates

Fatima Ben Salah[1], Fadl Labeeb Halabi[2]

[1] *Department of Agronomy and Plant Biotechnology, Institut National Agronomique de Tunisie (INAT), Tunis, Tunisia, f.bensalah@inat.agrinet.tn, ORCID: 0009-0017-3341-9915*

[2] *Department of Information Engineering, Università degli Studi di Brescia, Brescia, Italy, halabi.f@unibs.it, ORCID: 0009-0006-4555-5538*

Abstract

Agrivoltaic (AV) systems — the co-location of solar photovoltaic generation with agricultural crop production on the same land — are emerging as a high-impact strategy for addressing the food-energy-water (FEW) nexus in semi-arid Mediterranean climates, where high irradiance, water scarcity, and land competition are simultaneously acute. This paper presents a cradle-to-gate Life Cycle Assessment (LCA) of a 2 ha agrivoltaic system combining elevated bifacial PERC photovoltaic panels (1.8 m ground clearance, 264 kWp total capacity) with drip-irrigated tomato and pepper production in the Beja governorate, northern Tunisia. Three land-use configurations are systematically compared: (C1) conventional open-field cropping with grid-sourced irrigation pumping; (C2) agrivoltaic with fixed-rate drip irrigation; and (C3) agrivoltaic with soil-sensor-guided precision drip irrigation. LCA is performed using the ReCiPe 2016 midpoint (H) methodology with a 25-year system lifetime horizon and the ecoinvent 3.9 background database. The C3 agrivoltaic configuration achieves a Global Warming Potential (GWP) of 42.3 g CO_2-eq per kWh of delivered electricity, which is 92% lower than the Tunisian STEG grid emission factor of 527 g CO_2-eq/kWh, and 18% below a comparable ground-mount utility PV system. The precision irrigation variant reduces seasonal water consumption by 33.4% relative to C1 (from 4,820 to 3,210 m^3/ha/year for tomato), driven equally by shading-induced evapotranspiration reduction and sensor-triggered scheduling. The Land Equivalent Ratio (LER) of C3 is 1.42, confirming that 42% more land would be required in separated monocultures to achieve the same combined food and energy output. Bifacial rear-face irradiance gain under the elevated AV array geometry is 8.3%, contributing a measurable 4.1% system energy yield improvement over monofacial configurations. Sensitivity analysis identifies module manufacturing emission factor and avoided grid emission credit as the dominant uncertainty drivers. The results provide a quantitative, regionally calibrated evidence base for agrivoltaic policy and investment decisions in Tunisia and the broader semi-arid Mediterranean agricultural zone.

Keywords: *agrivoltaic; life cycle assessment; carbon footprint; bifacial solar panel; drip irrigation*

1. Introduction

The food-energy-water (FEW) nexus the recognition that food production, energy generation, and water availability are interdependent and must be managed jointly is acutely relevant in the semi-arid Mediterranean basin. Climate projections under the RCP 4.5 and RCP 8.5 scenarios consistently indicate declining summer precipitation across the southern and eastern Mediterranean, with Tunisia projected to experience 15-25% reductions in annual precipitation by 2050 relative to the 1990-2020 baseline, combined with increases in the frequency and severity of droughts [1]. In parallel, the region faces growing competition for land between expanding solar energy infrastructure and agriculture, with utility-scale PV deployment on agricultural soils generating significant social and political resistance in both North African and Southern European contexts.

Agrivoltaic systems address this nexus challenge by enabling simultaneous production of solar electricity and food crops on the same land parcel, exploiting physical complementarities that enhance the productivity of both systems relative to their separated monocultures. The partial shading provided by PV panels reduces crop heat stress and evapotranspiration demand, particularly valuable in semi-arid conditions where photoinhibition and heat stress are the primary yield-limiting factors during summer months. Conversely, the microclimate modification induced by crop canopies — lower ambient temperature, higher relative humidity, reduced wind speed improves PV module temperature and thus conversion efficiency relative to bare-ground installations [2]. Bifacial PV modules, which generate electricity from both front- and rear-face irradiance, offer an additional synergy in agrivoltaic configurations: the elevated mounting height required for under-panel agriculture increases effective albedo and rear-face irradiance relative to standard ground-mount configurations, amplifying the bifacial energy yield gain.

Despite rapid growth in agrivoltaic installations globally with installations now documented in over 40 countries and cumulative capacity exceeding 14 GW as of 2023 [3] life cycle assessment studies comparing integrated agrivoltaic configurations against conventional land-use baselines remain limited, and regional calibration for semi-arid North African conditions is particularly sparse. Tunisia presents an archetypal deployment context: the national electricity generation mix is predominantly fossil fuel-based, resulting in a grid emission factor of 527 g CO_2-eq/kWh that is among the highest in the Mediterranean region; approximately 65% of agricultural area is classified as arid or semi-arid; and the agricultural sector consumes 80% of national freshwater withdrawals through surface and groundwater irrigation. These conditions create a powerful alignment between the decarbonisation, water saving, and food security objectives that agrivoltaic systems can simultaneously address.

The present paper fills the regional LCA evidence gap by presenting a rigorous comparative LCA of three land-use configurations for a representative 2 ha vegetable production unit in the Beja governorate of northern Tunisia, using the ReCiPe 2016 midpoint methodology with the ecoinvent 3.9 background database. The three configurations conventional open-field cropping (C1), agrivoltaic with standard drip irrigation (C2), and agrivoltaic with soil-sensor-guided precision drip irrigation (C3) are compared across five midpoint impact categories: Global Warming Potential, Water Consumption, Land Use, Freshwater Eutrophication, and Human Toxicity. Energy yield modelling explicitly accounts for the bifacial gain under AV panel geometry. The paper is structured as follows: Section 2 provides the literature review; Section 3 details the system description and LCA methodology; Section 4 presents the results; Section 5 provides sensitivity analysis and discussion; and Section 6 concludes.

2. Literature Review

The agrivoltaic concept was formally proposed by Dupraz et al. [4] who introduced the Land Equivalent Ratio as the key metric demonstrating that co-located solar and crop production achieves higher combined productivity than separated systems on the same land area. The seminal Montpellier field experiment documented LER values of 1.28-1.35 for shade-tolerant crops under fixed-tilt monofacial PV systems, establishing the quantitative foundation for subsequent AV research. A 2022 Fraunhofer ISE guideline for Germany [3] documented LER values of 1.1-1.5 across 18 agrivoltaic pilot sites, confirming that the productivity surplus is robust across different climates, crop types, and PV configurations.

The systematic review by Gorjian et al. [2] covering progress in agrivoltaic technology across 22 countries established that crop response to partial shading is strongly species-dependent: shade-tolerant crops including tomato, pepper, lettuce, herbs, soft fruit, and vineyard crops maintain 90-105% of open-field yields under 30-45% shading, while shade-sensitive crops including maize, sunflower, and cereals exhibit yield reductions of 15-30% at equivalent shading intensities. In hot semi-arid environments, the yield penalty is frequently reversed for shade-tolerant crops: the reduction in heat stress and vapour pressure deficit under panel shade can increase yield relative to unshaded conditions during peak summer months, with tomato yield improvements of 5-18% reported in trials in the Middle East and North Africa.

Fixed-tilt agrivoltaic configurations with ground cover ratios (GCR) in the range 0.25-0.35 are identified as the optimal range for dual-use productivity in semi-arid Mediterranean climates, providing sufficient shading for crop benefit without excessive yield penalty for shade-sensitive crops in rotations [5]. East-west oriented panel rows at moderate tilt angles (30-35° at 36°N latitude) provide more uniform daily light distribution beneath the panels than north-south oriented rows, reducing the spatial heterogeneity of crop growth across the field and simplifying agronomic management.

Bifacial PERC (Passivated Emitter and Rear Cell) modules have become the dominant commercial technology for new PV installations, with market share exceeding 70% as of 2023 due to their superior energy yield per installed watt peak. In standard ground-mount configurations at 0.3-0.5 m above grade, bifacial gains of 3-7% relative to monofacial equivalents are typical, depending on ground albedo and row pitch. In elevated agrivoltaic mounting configurations with 1.5-2.0 m ground clearance, the effective bifacial gain increases to 7-14% due to: (i) reduced ground shading of the module rear face; (ii) higher apparent ground albedo from crop canopy reflectance (typically 0.22-0.30) relative to bare soil (0.10-0.18); and (iii) improved sky view factor of the rear surface [6].

The Middle Eastern agrivoltaics review [6] documented bifacial gain values of 8.1-11.4% in elevated AV configurations in arid and semi-arid climates with similar irradiance profiles to northern Tunisia, confirming that the bifacial advantage is particularly pronounced in high-irradiance, high-albedo environments. Module temperature coefficients (−0.35 to −0.45%/°C power output) interact with the AV microclimate benefit: the lower module temperature under crop canopy conditions (typically 4-8°C lower than bare-ground installations in summer) provides an additional 1.4-3.6% energy yield improvement beyond the bifacial gain, partially captured in the system-level energy modelling of the present study.

Life cycle assessment of solar PV systems has been comprehensively reviewed by the IEA PVPS Task 12 group, establishing that the GWP of crystalline silicon PV electricity ranges from 20 to 80 g CO_2-eq/kWh depending on technology, irradiance, and grid emission factor for manufacturing energy. The cradle-to-gate LCA of a utility-scale ground-mount system in a Mediterranean irradiance location (1,800-2,000 kWh/m²/year) typically yields 40-60 g CO_2-eq/kWh, with module manufacturing accounting for 55-65% of GWP [7]. Agostini et al. [7] assessed innovative agrivoltaic system configurations in the Po Valley, Italy, reporting energy payback times of 2.1-2.7 years and GWP of 33-48 g CO_2-eq/kWh, confirming that the functional unit sharing between food and energy outputs reduces per-kWh GWP relative to standalone PV systems.

The comprehensive LCA framework for agrivoltaic systems proposed at the AgriVoltaics Conference 2023 addressed the challenge of defining appropriate functional units for AV LCA, recommending a dual-output functional unit (energy + food per unit land area per year) as the most transparent approach for comparative studies. The LCA of potato-based AV systems by Busch and Wydra [8] reported GWP of 38.6 g CO_2-eq/kWh (electricity output basis), with module manufacturing dominating

at 58% of total GWP. Pascaris et al. [9] extended AV LCA to integrated livestock-pasture systems, demonstrating that land-sharing impacts across multiple output categories must be handled through system expansion rather than allocation to avoid methodological bias.

Precision drip irrigation, wherein irrigation events are triggered by real-time soil moisture sensor readings rather than by crop calendar schedules, has been documented to reduce water consumption by 15-35% in Mediterranean vegetable production relative to schedule-based drip irrigation at equivalent or improved yields [10]. The interaction between AV partial shading and irrigation demand operates through two mechanisms: (i) direct reduction of soil evaporation (documented at 30-65% under AV panels relative to open field in arid conditions); and (ii) reduction of crop transpiration through lower vapour pressure deficit and air temperature under panel shade. Barron-Gafford et al. [11] measured a 65% reduction in total evapotranspiration under AV arrays in Arizona desert conditions, attributable primarily to soil evaporation reduction. The combination of AV shading and precision irrigation thus provides multiplicative water savings that substantially exceed those achievable by either intervention alone.

The sensor-guided irrigation approach implemented in the C3 configuration of the present study uses capacitance-based soil moisture sensors (Decagon 5TM) at 20 and 40 cm depths to maintain soil water content within the 60-75% field capacity range, triggering drip irrigation events with a minimum inter-event interval of 2 hours to allow infiltration front development. Real-time weather data (temperature, humidity, wind speed, solar radiation) are used to update daily ET_0 estimates via the FAO-56 Penman-Monteith algorithm, providing an anticipatory irrigation demand signal that prevents stress-event onset under conditions of rapidly increasing evaporative demand.

Tunisia's agricultural sector contributes approximately 10-12% of GDP and employs 14% of the workforce, with irrigated agriculture covering approximately 470,000 ha and accounting for 80% of national freshwater withdrawals [12]. The principal irrigated crops in the northern plains — where Beja governorate is located — include cereals, legumes, vegetables, and arboriculture (olive, citrus). The Beja region benefits from the highest precipitation in Tunisia (500-700 mm/year), but summer precipitation is negligible and irrigation is required for all summer vegetable crops from June through September.

Tunisia's electricity generation is 97% fossil-fuel based (natural gas and oil), resulting in a STEG grid emission factor of 527 g CO_2-eq/kWh in 2022 — one of the highest in the Euro-Mediterranean area. The National Renewable Energy Plan (PANER) targets 35% renewable electricity capacity by 2030, with solar PV as the primary growth vector. The Prosumer Decree (2016, updated 2021) enables agricultural enterprises to install and operate solar PV for self-consumption with net-metering, creating a regulatory framework for agrivoltaic installations at farm scale.

Despite the growing international AV literature, three specific gaps motivate the present work: (i) LCA studies explicitly calibrated for the Tunisian electricity grid emission factor and North African solar resource remain absent; (ii) the combined effect of bifacial panel configuration and precision soil-moisture-based irrigation optimisation on LCA outcomes has not been jointly quantified; and (iii) comparative analysis across multiple ReCiPe midpoint impact categories — beyond GWP alone — for AV systems in semi-arid Mediterranean vegetables is lacking. The present work addresses all three gaps through a fully documented, ecoinvent-based comparative LCA.

3. Methodology

The study system is located in the municipality of Slouguia, Beja governorate, northern Tunisia (36.73°N, 9.18°E, elevation 84 m). The site has deep calcareous vertisol soils with a field capacity of 0.38 m³/m³ and a permanent wilting point of 0.18 m³/m³. Annual horizontal irradiation is 1,820 kWh/m²/year (Meteonorm TMY); the site is connected to the STEG 30 kV distribution network with a grid emission factor of 527 g CO_2-eq/kWh [12]. The three configurations occupy the same 2 ha land parcel and are compared over a 25-year study horizon.

C1 (conventional open-field): tomato (Year 1-2) / pepper (Year 3-4) rotation, furrow and drip irrigation from a farm well supplemented by SONEDE municipal supply, pumping energy from STEG grid. C2 (agrivoltaic, standard drip): identical PV system to C3 with fixed-schedule drip irrigation at $1.0 \times ET_0$ crop water requirement without soil moisture feedback. C3 (agrivoltaic, precision drip): 264 kWp bifacial PERC PV system with soil-moisture-sensor-triggered drip irrigation. The system boundary is cradle-to-gate for the PV system and includes agricultural inputs (seeds, fertilisers, pesticides, packaging) and pumped irrigation water for all configurations. End-of-life treatment applies 85% module material recovery credit following ecoinvent module recycling datasets.

The AV PV system comprises 480 bifacial PERC modules (manufacturer: LONGi Solar, model LR5-72HBD 550 Wp, bifaciality factor 0.70, temperature coefficient −0.35%/°C) mounted on galvanised steel racking structures at 32° fixed tilt, south-facing, with 8 m east-west row pitch providing a GCR of 0.28. Minimum ground clearance is 1.8 m (portal-frame structure), enabling passage of standard agricultural machinery (tractor + equipment to 1.7 m width). The inverter system comprises six 44 kW string inverters (SMA Tripower 50000TL-JP) connected to a 264 kWp string array. Generation is consumed on-site for pumping, lighting, and crop processing; surplus is exported to STEG under the net-metering prosumer framework. Annual net energy generation is modelled at 412 MWh/year (year 1), degrading at 0.5%/year per the module manufacturer's linear performance warranty.

Mounting structure material quantities are estimated from structural engineering drawings for the portal-frame AV racking system: 42 kg/m of galvanised steel tube per portal frame at 8 m pitch yields approximately 96 frames for the 2 ha array, for a total structural steel mass of 38.4 tonnes. Concrete foundation volume for driven pile footings is 1.2 m^3 per frame, totalling 115 m^3. These material quantities are incorporated in the LCA foreground inventory using ecoinvent 3.9 datasets for galvanised steel tube and ready-mix concrete.

Bifacial energy yield is simulated using the pvfactors v1.5.2 open-source bifacial irradiance model with hourly TMY meteorological data from Meteonorm 8.2 for the Beja site. The rear-face irradiance at each row position is computed as a function of direct, diffuse, and ground-reflected irradiance components, with the effective ground albedo modelled as a time-varying weighted average of crop canopy albedo (0.22-0.28, phenologically varying) and bare soil albedo (0.18) according to the seasonal crop cover fraction. The bifacial energy yield simulation yields a system-level bifacial gain BG = 8.3% over a monofacial equivalent at the same nameplate capacity, corresponding to an annual energy production increment of 34.2 MWh/year. The energy yield is converted to GWP impact using the avoided STEG grid emission factor of 527 g CO_2-eq/kWh as the system expansion credit for all AV configurations.

Crop yield under partial shading is estimated from published experimental dose-response relationships for tomato (Solanum lycopersicum L.) and pepper (Capsicum annuum L.) calibrated to semi-arid Mediterranean conditions. At the modelled GCR of 0.28 (corresponding to approximately 28% panel-shadow coverage at solar noon during the summer season), the mean light transmittance under the AV array is 0.72 relative to open-sky conditions. Tomato yield response at this shading level is modelled as 94% of open-field yield (based on Marrou et al. [13] and confirmatory Tunisian field data from INAT experimental plots), and pepper at 97%. These yield responses are incorporated in the functional unit calculation and the LER.

Irrigation demand under each configuration is estimated using the FAO-56 dual crop coefficient approach. Reference evapotranspiration ET_0 is computed from the Beja TMY meteorological data using the Penman-Monteith equation. Crop evapotranspiration under open-field conditions (C1) is $ET_c = K_c \times ET_0$, with crop coefficients K_c = 0.85 (initial), 1.15 (mid-season), and 0.80 (late-season) for tomato. For AV configurations, a shading correction factor f_{shade} = 0.78 is applied to the crop transpiration component of ET_c, following the empirical relationship established by Marrou et al. [13], yielding a net crop ET reduction of 19% under AV shading. The additional precision irrigation saving in C3 (18%) is estimated from published sensor-triggered drip irrigation efficiency comparisons for Mediterranean vegetables [10].

The LCA is performed in openLCA 2.1 software using the ReCiPe 2016 midpoint (H) characterisation method. The five impact categories assessed are: (i) Global Warming Potential (GWP, kg CO_2-eq, 100-year time horizon); (ii) Water Consumption (m^3 water-eq, scarcity-weighted); (iii) Land Use (m^2□year of agricultural land, annual crop); (iv) Freshwater Eutrophication (kg P-eq); and (v) Human Toxicity, non-cancer (kg 1,4-DCB-eq). Background data for PV module manufacturing, inverter production, steel, concrete, and agricultural inputs (seeds, NPK fertiliser, pesticides, packaging) are drawn from ecoinvent 3.9 with Consequential system model. The functional unit is 1 ha□year of combined crop and electricity production for the AV configurations, and 1 ha□year of crop production for C1 (with an equivalent electricity allocation from the STEG grid).

Parameter uncertainty is assessed through a Monte Carlo simulation with N = 1,000 iterations, sampling from triangular distributions around the key foreground parameters: bifacial gain (±2% absolute), crop yield response under shading (±5% absolute), irrigation water saving (±5% absolute), and module manufacturing GWP (±15% relative). Results are reported as mean ± one standard deviation. A scenario analysis supplements the Monte Carlo assessment by evaluating the sensitivity of GWP to the Tunisian grid emission factor (low scenario: 420 g CO_2-eq/kWh, reflecting partial renewable grid integration; high scenario: 570 g CO_2-eq/kWh, reflecting grid degradation) and to module lifetime (20-year vs. 30-year scenarios).

4. Results

The GWP results per kWh of delivered electricity for the 25-year system lifetime are: C2 (agrivoltaic, standard drip) 44.8 g CO_2-eq/kWh; C3 (agrivoltaic, precision drip) 42.3 g CO_2-eq/kWh (Monte Carlo 95% CI: 38.7-46.2 g CO_2-eq/kWh). The 5.5% GWP reduction from C2 to C3 reflects the larger avoided grid electricity credit attributable to increased crop productivity per unit water input in C3, reducing the allocation of irrigation pumping electricity to the functional unit. Both AV configurations achieve GWP reductions of 91-92% relative to the STEG grid emission factor of 527 g CO_2-eq/kWh, and 14-19% below a co-located utility-scale ground-mount PV system (51.7 g CO_2-eq/kWh, estimated using the same site and module data without AV structural premium).

The contribution analysis reveals that PV module manufacturing accounts for 56.3% of total GWP in C3, followed by mounting structure fabrication (14.8%), inverter and BOS manufacturing (8.4%), agricultural inputs (fertiliser 7.1%, pesticides 4.2%, seeds 1.8%), and irrigation system infrastructure (3.2%). End-of-life aluminium and silicon recovery credits offset 4.8% of total GWP. The dominance of module manufacturing in the GWP profile confirms that the choice of module manufacturing location and energy mix is the most impactful variable for LCA-based comparison of competing PV technologies — a finding consistent with the broader PV LCA literature [7].

The bifacial gain of 8.3% reduces GWP by approximately 3.5 g CO_2-eq/kWh relative to a monofacial equivalent through two mechanisms: increased energy yield (increasing the avoided grid emission credit) and reduced proportional structural and manufacturing burden per kWh of output. This bifacial GWP benefit is in addition to the direct energy performance advantage and reinforces the case for bifacial module selection in AV systems at high-irradiance, elevated-mount configurations.

Seasonal irrigation water consumption in C3 is 3,210 m^3/ha/year for tomato (C1: 4,820 m^3/ha/year), representing a 33.4% reduction. For pepper in the rotation years, C3 consumption is 2,180 m^3/ha/year compared to 3,090 m^3/ha/year in C1 (29.5% reduction). In the scarcity-weighted ReCiPe water consumption metric, the Beja region carries a water scarcity characterisation factor of 1.43 m^3 water-eq/m^3 consumed (reflecting moderate-to-high stress classification for the Medjerda basin), yielding a lifecycle water consumption impact of 4,590 m^3 water-eq/ha/year for C1 and 3,025 m^3 water-eq/ha/year for C3 — a 34.1% reduction that is among the most compelling environmental benefits of the agrivoltaic plus precision irrigation approach.

The water saving mechanism deserves disaggregation: shading-induced ET reduction accounts for 57% of the total saving (equivalent to 924 m^3/ha/year for tomato), and precision irrigation scheduling accounts for the remaining 43% (688 m^3/ha/year). This disaggregation confirms that the AV structural shading provides the larger water saving contribution, but that the sensor-guided irrigation layer is a substantial incremental improvement that cannot be attributed to the PV system alone. From a water governance perspective, the 1,610 m^3/ha/year net water saving in C3 relative to C1 corresponds to a reduction in annual Medjerda aquifer abstraction volume of 3,220 m^3 for the 2 ha study farm — a concrete contribution to the Beja regional water balance.

The ReCiPe land use impact category (annual crop land, m²□year) is identical for C1, C2, and C3 by construction of the functional unit (same 2 ha land parcel). However, the Land Equivalent Ratio explicitly quantifies the land productivity gain of the agrivoltaic approach over separated monocultures. The C3 LER of 1.42 (tomato rotation mean) is computed as: LER_crop = 0.94 (tomato yield fraction) + LER_energy = 0.47 (fraction of equivalent PV land for same energy yield at standard GCR 0.55 system). The LER of 1.42 indicates that 42% more total land would be required in separated monocultures (conventional tomato farm + utility PV plant) to deliver the same combined output as the C3 AV system. This land productivity premium has direct policy relevance in the Tunisian context where competition for fertile agricultural land in the northern plains limits large-scale PV deployment.

The LER sensitivity to the crop yield response assumption is the most significant uncertainty: at the pessimistic yield assumption (tomato 85% of open-field at GCR = 0.28), LER decreases to 1.31 — still a substantial land productivity advantage. At the optimistic assumption (tomato yield parity at 100% due to heat stress relief), LER increases to 1.47. The range 1.31-1.47 across the full uncertainty envelope confirms that the agrivoltaic land productivity advantage is robust to reasonable variation in the crop yield response parameter.

Freshwater eutrophication GWP (kg P-eq/ha/year) is dominated by phosphate fertiliser application and leaching in all three configurations. C3 shows a 12.8% reduction relative to C1 (0.187 vs. 0.214 kg P-eq/ha/year), primarily attributable to reduced irrigation volume and thus reduced phosphate leaching flux per unit crop output. The reduction in irrigation-driven leaching in C3 partially offsets the small eutrophication impact associated with PV system manufacturing chemicals. Human Toxicity (non-cancer, kg 1,4-DCB-eq/ha/year) is dominated by pesticide application in all configurations. C3 shows a marginal 3.1% reduction relative to C1, reflecting the modest reduction in pesticide application rate achievable under AV conditions due to lower canopy temperature and humidity-induced fungal disease pressure. These impact category improvements are secondary to the GWP and water savings but contribute to the overall environmental superiority of the C3 configuration.

5. Discussion

The Monte Carlo sensitivity analysis (N = 1,000 iterations) confirms that the GWP result for C3 (42.3 g CO_2-eq/kWh) is robust within a 95% confidence interval of 38.7-46.2 g CO_2-eq/kWh. The dominant uncertainty source is the Tunisian grid emission factor, which drives the avoided emission credit: a ±10% variation in the grid emission factor propagates to a ±8% variation in the net GWP result. Module manufacturing GWP uncertainty (±15%) propagates to ±6.4% on the total GWP. Bifacial gain uncertainty (±2% absolute) contributes only ±0.8 g CO_2-eq/kWh — confirming that bifacial yield modelling precision is not a critical uncertainty driver for LCA conclusions. The crop yield response to shading (±5% absolute) propagates primarily to the LER metric (±0.05-0.08) rather than to GWP, confirming methodological robustness of the carbon results.

Scenario analysis on system lifetime reveals that a 20-year lifetime (conservative scenario) increases GWP to 51.2 g CO_2-eq/kWh (C3), while a 30-year lifetime (optimistic, conditional on module performance warranty extension) reduces GWP to 37.1 g CO_2-eq/kWh. These results confirm that the choice of system lifetime is a significant LCA parameter for long-lived infrastructure assets and highlight

the importance of module warranty duration and actual degradation rate as key drivers of LCA outcomes in semi-arid high-temperature deployment environments.

The C3 GWP of 42.3 g CO_2-eq/kWh compares favourably with the 38.6 g CO_2-eq/kWh reported by Busch and Wydra [8] for a Central European AV system, noting that the lower result in the German study reflects the lower module manufacturing emission factor (European grid mix) rather than superior system design. The GWP of the present study is 15-22% below values reported for conventional ground-mount utility PV in comparable North African irradiance conditions (51-65 g CO_2-eq/kWh), confirming that the functional unit sharing between food and energy outputs reduces the per-kWh LCA burden. The LER of 1.42 in the present study is at the upper end of the published range for Mediterranean AV systems (1.28-1.47), reflecting the specific advantage of bifacial elevated panels combined with precision irrigation for tomato and pepper crops in a high-irradiance, water-stressed environment.

The quantified environmental benefits of the C3 agrivoltaic configuration — 92% GWP reduction relative to grid, 34% water saving, LER of 1.42 — provide a rigorous evidence base for several policy instruments. First, the GWP reduction supports inclusion of agrivoltaic electricity generation in Tunisia's NDC (Nationally Determined Contribution) renewable energy accounting, with the 412 MWh/year energy yield of the 2 ha C3 system avoiding 217 tCO_2/year relative to STEG grid electricity. Second, the water saving of 1,610 m^3/ha/year is directly quantifiable against the Ministry of Agriculture's water efficiency targets for irrigated agriculture, potentially qualifying AV irrigation systems for water efficiency subsidy programmes. Third, the LER of 1.42 provides a quantitative basis for valuing the land footprint efficiency of AV systems in the STEG renewable energy procurement framework, where land availability for large-scale PV is a binding constraint in agriculturally productive regions.

The agrivoltaic prosumer regulatory framework, established by Tunisian Decree No. 2016-1123 and updated in 2021, currently permits self-consumption installations up to 200 kWp without grid export tariff. The 264 kWp C3 system slightly exceeds this threshold, requiring a tailored utility agreement for the surplus export portion (approximately 55 kWp equivalent at average load factor). Regulatory clarification for AV systems in the 200-500 kWp range — explicitly recognising the agricultural primary land use — would remove the principal regulatory barrier to AV deployment in the Beja and Bizerte governorates where the combination of high irradiance, fertile soils, and water stress makes agrivoltaic systems particularly compelling.

The LCA presented is a prospective assessment based on modelled crop yields, irrigation demands, and energy generation rather than operational field data. Field validation from at least two complete crop rotation cycles (four growing seasons) is required to confirm the modelled water savings, crop yield responses, and bifacial energy yield predictions. Key parameters requiring field validation include: the effective crop canopy albedo contribution to bifacial gain (currently estimated from literature values); the spatially integrated ET reduction under the specific panel geometry of the C3 system; and the fertiliser leaching reduction attributable to precision irrigation under AV conditions. A companion agronomic trial at the INAT Beja experimental station is planned for 2025-2026 to provide these validation data.

Future research will extend the LCA scope to include: (i) the social LCA dimension, including impacts on agricultural labour employment and seasonal worker income; (ii) techno-economic analysis under Tunisian STEG prosumer tariff structures to determine internal rate of return and payback period as functions of tomato and pepper market prices; (iii) multi-crop rotation scenarios including drought-tolerant cereals and legumes in the shaded zone; and (iv) the floating agrivoltaic configuration potential for irrigation reservoir surfaces in the Sidi Salem dam catchment area. The ReCiPe characterisation factors used in the present study do not include a specific characterisation factor for the avoided freshwater withdrawl from stressed aquifers, representing a methodological gap that future water-scarcity-aware LCA methods should address.

6. Conclusions

This paper has presented a cradle-to-gate comparative LCA of three land-use configurations for a 2 ha vegetable production unit in Beja governorate, northern Tunisia, demonstrating that the integrated agrivoltaic plus precision drip irrigation configuration (C3) achieves substantial environmental improvements across multiple impact categories. The principal quantitative findings are: (i) GWP of 42.3 g CO_2-eq/kWh (Monte Carlo 95% CI: 38.7-46.2), representing a 92% reduction relative to the Tunisian STEG grid emission factor and an 18% improvement over co-located utility PV; (ii) seasonal irrigation water consumption reduction of 33.4% for tomato (from 4,820 to 3,210 m^3/ha/year), with 57% attributable to AV shading-induced ET reduction and 43% to precision sensor-triggered scheduling; (iii) Land Equivalent Ratio of 1.42 (95% scenario range: 1.31-1.47), confirming 42% superior land productivity relative to separated monocultures; (iv) bifacial gain of 8.3%, contributing 34.2 MWh/year incremental energy yield and 3.5 g CO_2-eq/kWh GWP improvement over monofacial equivalents; and (v) freshwater eutrophication reduction of 12.8% through reduced phosphate leaching under precision irrigation.

These results provide the most comprehensive region-specific LCA evidence to date for agrivoltaic systems in the semi-arid Mediterranean agricultural context, directly supporting the case for regulatory enablement of AV deployment under Tunisia’s renewable energy prosumer framework. The quantified synergies between bifacial elevated panel configuration and precision drip irrigation — in water savings, GWP reduction, and land use efficiency — confirm that the integrated C3 approach represents a superior deployment strategy over agrivoltaic systems without precision irrigation optimisation. Field validation across two complete crop rotation cycles is the immediate next step required to transition these findings from modelling-based evidence to operational evidence for policy application.

REFERENCES

[1] IEA, Renewables 2022: Analysis and Forecast to 2027. Paris: International Energy Agency, 2022.

[2] S. Gorjian, E. Bousi, O. E. Özdemir, M. Trommsdorff, N. M. Kumar, A. Anand, K. Kant, and S. S. Chopra, "Progress and challenges of crop production and electricity generation in agrivoltaic systems using semi-transparent photovoltaic technology," Renew. Sustain. Energy Rev., vol. 158, pp. 112126, Apr. 2022.

[3] Fraunhofer ISE, Agrivoltaics: Opportunities for Agriculture and the Energy Transition - A Guideline for Germany. Freiburg: Fraunhofer Institute for Solar Energy Systems, 2022.

[4] C. Dupraz, H. Marrou, G. Talbot, L. Dufour, A. Nogier, and Y. Ferard, "Combining solar photovoltaic panels and food crops for optimising land use: Towards new agrivoltaic schemes," Renew. Energy, vol. 36, no. 10, pp. 2725-2732, Oct. 2011.

[5] M. Trommsdorff, J. Kang, C. Reise, S. Schindele, G. Bopp, A. Ehmann, A. Weselek, P. Häring, and T. Obergfäll, "Combining food and energy production: Design of an agrivoltaic system applied in arable and vegetable farming in Germany," Renew. Sustain. Energy Rev., vol. 140, pp. 110694, Apr. 2021.

[6] K. Al Shalabi et al., "Middle Eastern agrivoltaics: Technologies, sustainability, and economic effects," Sustainability, vol. 18, no. 3, pp. 1596, 2025.

[7] A. Agostini, M. Colauzzi, and S. Amaducci, "Innovative agrivoltaic systems to produce sustainable energy: An economic and environmental assessment," Appl. Energy, vol. 281, pp. 116102, Jan. 2021.

[8] C. Busch and K. Wydra, "Life cycle assessment of an agrivoltaic system with conventional potato production," J. Renew. Sustain. Energy, vol. 15, no. 4, pp. 043501, Jul. 2023.

[9] A. S. Pascaris, R. Handler, C. Schelly, and J. M. Pearce, "Life cycle assessment of pasture-based agrivoltaic systems: Emissions and energy use of integrated rabbit production," Clean. Responsible Consum., vol. 3, pp. 100030, 2021.

[10] E. A. Abioye, O. Hensel, T. J. Esau, O. Elijah, M. S. Z. Abidin, A. S. Ayobami, O. Yerima, and A. Nasirahmadi, "Precision irrigation management using machine learning and digital farming solutions," AgriEngineering, vol. 4, no. 1, pp. 70-103, Feb. 2022.

[11] G. A. Barron-Gafford, M. A. Pavao-Zuckerman, R. L. Minor, L. F. Sutter, I. Barnett-Moreno, D. T. Blackett, M. Thompson, K. Dimond, A. K. Gerlak, G. P. Nabhan, and J. E. Macknick, "Agrivoltaics provide mutual benefits across the food-energy-water nexus in drylands," Nat. Sustain., vol. 2, no. 9, pp. 848-855, Sep. 2019.

[12] GIZ, Study on the Opportunities of Power-to-X in Tunisia. Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit, 2021.

[13] H. Marrou, L. Guilioni, L. Dufour, C. Dupraz, and J. Wery, "Microclimate under agrivoltaic systems: Is crop growth rate affected in the partial shade of solar panels?" Agric. For. Meteorol., vol. 177, pp. 117-132, Sep. 2013.

[14] GEF, Agrivoltaics: Technical and Policy Overview. Washington DC: Global Environment Facility, 2024.

[15] A. Leon and K. N. Ishihara, "Assessment of new functional units for agrivoltaic systems," J. Environ. Manag., vol. 226, pp. 493-498, Nov. 2018.

[16] A. Goyal et al., "Saving Southeast Asia’s crops: Four key steps toward food security," McKinsey & Company, Jun. 2023.

[17] European Parliament, Directive 2009/28/EC on the Promotion of the Use of Energy from Renewable Sources. Luxembourg: Official Journal of the European Union, 2009.

Real-Time Forest Fire Detection and Spread Prediction Using Transformer-Based Vision Models on Edge Computing Nodes: A Case Study of the Aegean Pine Forest Zone, Greece

Elif Yıldız [1]

[1] *Department of Forest Engineering, Faculty of Forestry, Kahramanmaraş Sütçü İmam University, 46040 Kahramanmaraş, Türkiye, 241103045@ogr.ksu.edu.tr, ORCID: 0000-0002-6942-1474*

Abstract

Forest fires in the Eastern Mediterranean, particularly in the Aegean pine (Pinus brutia) forest zone of Greece and Turkey, are increasing in frequency and intensity under projected climate change conditions, with the 2021 and 2023 Greek fire seasons setting historical records for annual burned area within the European Union. Early automated detection and real-time fire spread prediction are critical capabilities for enabling effective suppression resource pre-positioning and evacuation decision-making within the narrow window of fire suppression feasibility. This paper presents EgeoFireNet, an integrated end-to-end system combining a Swin Transformer-based fire and smoke detection model with a physics-informed cellular automaton spread prediction module, deployed on NVIDIA Jetson Orin NX edge computing nodes integrated into a 43-camera fixed-mount dual-channel (visible + NIR) monitoring network covering 12,400 ha of Aegean pine forest in the Chios Regional Unit, Greece. The Swin Transformer detection model (Swin-T backbone, 28M parameters, pretrained ImageNet-22K) is fine-tuned on a domain-specific dataset of 18,600 annotated Aegean pine fire images augmented with 4,200 frames from the FLAME2 multispectral benchmark, achieving mean Average Precision at IoU 0.5 (mAP@0.5) of 0.889 and F1-score of 0.912 for fire and smoke detection at 22.4 FPS on the Jetson Orin NX (TensorRT INT8 quantisation). The cellular automaton spread prediction model, initialised from real-time Swin Transformer fire perimeter segmentation masks and driven by LPWAN-networked wind and humidity sensor data, generates geo-referenced 15-minute fire spread forecasts in 4.2 seconds on the edge node CPU, with a Sorensen-Dice similarity coefficient of 0.76 relative to observed burned perimeters in six validation test fire events. End-to-end system alert latency from fire ignition to confirmed multi-camera detection and SMS/API alert dispatch to the Chios Regional Emergency Operations Centre is 87 seconds (median) across 14 controlled test fire events, a 12-24-fold reduction relative to conventional helicopter patrol detection. False positive rate is 0.023 alerts per camera per hour, within operational service tolerance.

Keywords: *forest fire detection; spread prediction; cellular automaton; early warning*

1. Introduction

Wildfire is among the most damaging natural hazards in the Mediterranean basin, causing annual economic losses estimated at EUR 1.0-3.5 billion in Greece alone during severe fire seasons [1]. The 2021 Greek fire season burned approximately 125,000 ha, the worst since the catastrophic 2007 season; the 2023 Alexandroupolis wildfire complex in northeastern Greece burned over 81,000 ha in a single event the largest recorded wildfire in European Union history by area, resulting in 20 fatalities and displacement of 70,000 residents. Climate projections under EURO-CORDEX regional scenarios consistently indicate that fire weather conditions in the Eastern Mediterranean will intensify through mid-century: the number of days per year with fire weather index (FWI) exceeding the “very high” threshold is projected to increase by 40-60% in Greece and Turkey by 2050 relative to the 1990-2020 baseline, driven by higher summer temperatures, longer drought periods, and increased wind frequency [2].

The Aegean pine (Pinus brutia Ten.) is the dominant forest species of the eastern Aegean islands, coastal Anatolia, and the eastern Greek mainland, covering approximately 3.2 million ha across Greece and 5.8 million ha across Turkey. Pinus brutia forest is characterised by high flammability its resinous needle and bark chemistry, open canopy structure, and tendency to accumulate dry litter fuel make it one of the most fire-prone forest types in Europe and by high ecological and economic value, including watershed protection, timber production, and tourism landscape services. The combination of high ignition probability under summer drought conditions and high consequence-per-hectare burned makes the Aegean pine zone the highest fire risk stratum in the Greek National Fire Risk Management System, receiving the largest share of aerial and ground suppression resource deployment during the fire season.

Early detection of nascent fires specifically, detection and confirmed alert dispatch within the first 2-10 minutes of ignition before a fire transition from the “suppression-feasible” phase to the “suppression-infeasible” phase has been identified by fire management authorities as the highest-leverage technical intervention for improving suppression outcomes. Analysis of 340 Greek fire events between 2010 and 2022 confirms that fires detected within 5 minutes of ignition were suppressed within 1 ha in 71% of cases, while fires with detection delays exceeding 30 minutes resulted in suppression failure and large fire development in 64% of cases under High fire danger weather conditions [1]. The critical implication is that detection system latency is a primary determinant of fire outcome, motivating investment in automated early detection technologies capable of sub-10-minute detection under all weather and visibility conditions.

Computer vision approaches to automated fire detection have advanced substantially with the adoption of deep learning, progressing from CNN-based classifiers to the current generation of vision transformer (ViT) architectures that demonstrate superior performance on fine-grained visual recognition tasks including early-stage smoke detection in forested backgrounds [3]. The Swin Transformer, which introduces a hierarchical feature representation with shifted window self-attention, is particularly well suited to fire and smoke detection tasks that require multi-scale feature extraction: early smoke plumes at the scale of a few pixels require high-resolution fine features, while the classification of a scene as a fire event requires integration of evidence across a wide spatial context [4]. The deployment of transformer inference on embedded GPU hardware specifically the NVIDIA Jetson platform family enables each camera node to perform real-time detection independently, eliminating the cloud connectivity dependency that represents a critical failure mode during active fire events when power and communication infrastructure is vulnerable.

Fire spread prediction forecasting the trajectory of fire perimeter expansion over the next 15-60 minutes is complementary to detection in enabling effective fire management: while detection triggers initial alert and resource dispatch, spread prediction enables pre-positioning of resources at projected fire head locations, advance evacuation of threatened communities, and real-time resource reallocation as fire behaviour evolves. Physics-based spread models (FARSITE, Prometheus) provide high-fidelity predictions but require minutes to hours of computation time, limiting their operational utility for short-horizon real-time decision support. Cellular automaton (CA) models, which propagate fire states on a regular grid using spatially local rules encoding fuel, topography, and weather conditions, provide a computationally tractable approximation that enables 15-minute forecasts in seconds on embedded computing hardware [5].

This paper presents EgeoFireNet, an integrated fire monitoring system combining Swin Transformer-based detection and CA spread prediction on a 43-camera Jetson Orin NX edge computing network deployed across the Chios island Aegean pine forest zone. The principal contributions are: (i) a Swin-T fire/smoke detection model trained on a domain-specific Aegean pine fire dataset augmented with the FLAME2 benchmark, achieving mAP@0.5 = 0.889 at 22.4 FPS on edge hardware; (ii) a CA spread prediction module producing 15-minute geo-referenced forecasts at 4.2-second computation time on the Jetson Orin NX CPU; (iii) validated end-to-end system performance across 14 controlled test fire events confirming 87-second median alert latency; and (iv) an operational integration architecture linking the edge network to the Chios Regional Emergency Operations Centre via geo-tagged API alerts. Section 2 reviews the relevant literature; Sections 3 and 4 present the system design and experimental methodology; Section 5 reports results; Section 6 provides discussion; and Section 7 concludes.

2. Literature Review

The fire ecology of Pinus brutia forests in the Aegean region is characterised by a natural fire return interval of 25-60 years under pre-anthropogenic conditions, but fire suppression during the 20th century has substantially increased fuel loads and reduced the proportion of forests in fire-resilient structural stages. The accumulation of continuous ladder fuels and deep litter layers under suppression-fire regimes creates conditions for high-intensity crown fires that are qualitatively different from the surface fires of the historical fire regime, dramatically increasing the rate of fire spread under extreme weather conditions [2]. The Greek National Fire Risk Map, maintained by the Hellenic Fire Corps and updated annually, classifies approximately 1.8 million ha of Greek forest as “extreme” or “very high” fire risk, with the Aegean islands and eastern mainland coastal zone accounting for the highest density of critical risk sites.

The fire season in the Aegean pine zone extends from approximately mid-May through mid-October, with the peak risk period concentrated in July and August when the combination of low fuel moisture content (fine dead fuel moisture content dropping below 8-10% during heat waves), high temperatures (routinely 38-42°C), low relative humidity (< 25%), and strong north (Meltemi) or south (Sirocco) winds creates conditions for catastrophic fire development. The Meltemi wind, a persistent dry northerly that accelerates through the Aegean island channels is particularly dangerous due to its gusty, directionally variable character that causes rapid fire direction changes that outpace suppression force repositioning [2].

The evolution of computer vision architectures for fire detection has followed the broader trajectory of deep learning: early work used hand-crafted features (colour histograms, motion vectors) in classical classifiers; CNN architectures (ResNet, VGG, DenseNet) replaced these with learned spatial features; and the current frontier is occupied by vision transformer models that integrate spatial and contextual information through self-attention mechanisms. The systematic review by Ghali and Akhloufi [6] documented the transition across 47 fire detection studies published between 2019 and 2023, confirming that transformer-based models consistently outperform CNNs on the most challenging detection cases: early-stage smoke plumes with low pixel count, optically thin smoke under high-contrast backgrounds, and night-time NIR imagery with reduced contrast.

The Swin Transformer [4], introduced by Liu et al. at Microsoft Research in 2021, addresses the computational bottleneck of the original Vision Transformer (ViT) by restricting self-attention computation to local windows and introducing a shifted window mechanism for cross-window interaction. The hierarchical representation with feature map down sampling at each stage analogous to CNN feature pyramid networks makes Swin-T directly applicable to dense prediction tasks (detection, segmentation) without requiring the additional multi-scale processing needed for plain ViT architectures. Swin-T with 28M parameters achieves an ImageNet top-1 accuracy of 81.3%, comparable to ResNet-101 (77.3%) at

significantly lower computational cost, and provides pretrained weights that transfer effectively to domain-specific applications including fire detection [4].

The ensemble fire detection approach of Falcao et al., combining EfficientNetV2, DeiT, and Swin Transformer V2 on the D-Fire and Foggia datasets, achieved 96.46% accuracy and AUPRC of 95.14%, establishing transformer ensemble methods as the current state of the art. The distilled Vision Transformer (D-ViT) approach of Nawaz et al. [3] demonstrated that knowledge distillation from a large ViT teacher to a lightweight student model enables deployment at 28 FPS on NVIDIA Jetson Nano with accuracy retention within 2% of the full model establishing the practical pathway for edge-deployed ViT fire detection that EgeoFireNet extends to the more capable Jetson Orin NX hardware with the Swin-T architecture.

Fire spread models can be classified along a spectrum from empirical (tabulated rate-of-spread functions) through semi-empirical (Rothermel surface fire model, Cheney grass fire models) to physics-based (computational fluid dynamics fire spread codes). The Rothermel model [7], the basis of most operational fire spread tools including FARSITE, computes the rate of fire spread as a function of fuel moisture, wind speed, wind direction, terrain slope, and fuel bed characteristics. While the Rothermel model provides physically grounded predictions, its calibration for Mediterranean Pinus brutia fuel types requires site-specific fuel load and moisture content measurements that are not routinely available in operational settings.

Cellular automaton fire spread models represent an alternative approach in which the landscape is discretised on a regular grid and fire propagation between cells is governed by local state transition rules encoding the physical spread mechanisms. The CA approach trades physical rigour for computational tractability and parametric simplicity, enabling real-time operational spread forecasting at spatial resolutions (10-25 m) and time steps (1-15 minutes) appropriate for tactical fire management decision support [5]. The validation of CA models for Mediterranean vegetation types by the JRC-EFFIS team documented Dice similarity coefficients of 0.62-0.80 for short-horizon (15-30 minute) spread predictions across 48 validation events in Greece, Italy, and Portugal, providing the performance benchmark against which EgeoFireNet’s spread prediction module is evaluated.

The deployment of deep neural network inference on embedded GPU computing platforms generalised as Edge AI has been demonstrated across multiple environmental monitoring applications. Mu et al. [8] demonstrated UAV-based forest fire detection on embedded processors at 15 FPS, establishing real-time edge inference viability for aerial monitoring. The OpenWeedLocator [9] project confirmed that open-source Raspberry Pi-class embedded platforms can support accuracy-competitive agricultural monitoring applications in field environments. The NVIDIA Jetson Orin NX (100 TOPS at INT8), the highest-performance member of the Jetson Orin module family, provides a substantial compute upgrade over the Jetson Nano (21 TOPS) and Jetson Xavier NX (21 TOPS) used in earlier agricultural and environmental monitoring deployments, enabling higher-capacity vision transformer models to be deployed at real-time frame rates within the 15-25 W power envelope required for solar-powered tower infrastructure [3].

The specific challenges of edge AI deployment in the Aegean pine fire monitoring context are: (i) high ambient temperatures (up to 42°C on exposed tower sites during peak fire season) requiring active cooling of the edge compute enclosure; (ii) 4G/LTE connectivity with variable signal strength requiring local alert logic that does not depend on cloud acknowledgement; (iii) power supply interruption risk during fire events (grid outage) requiring battery backup with sufficient capacity for 48-hour autonomous operation; and (iv) dust, salt spray, and insect ingress in the maritime island environment requiring IP66-rated enclosures and periodic cleaning maintenance. These deployment constraints drove specific design choices in the EgeoFireNet infrastructure, detailed in Section 3.

The optimisation of camera placement for forest fire detection networks involves a trade-off between coverage area (which increases with tower height and lens focal length), detection sensitivity (which decreases with distance from fire), and infrastructure cost (which scales with tower count). Quantitative network design methods based on viewshed analysis from digital elevation models have been developed to maximise detection probability per unit infrastructure cost [10]. The specific challenge of the Chios island terrain characterised by steep ridgelines, deep valleys, and coastal escarpments that create complex viewshed geometries required iterative viewshed optimisation to achieve the 97% area coverage fraction achieved by the 43-tower EgeoFireNet network.

Multi-camera confirmation logic requiring concurrent detection by two or more cameras viewing the same geographic area before issuing a dispatch alert is a widely adopted strategy for reducing false alarm dispatch rates in fire camera networks. The trade-off is that confirmation logic increases detection latency by requiring a second camera's detection to be confirmed within a time window. In EgeoFireNet, the confirmation logic is implemented with a 90-second concurrent detection window and a fallback single-camera dispatch pathway for high-confidence detections (> 0.92 confidence) coinciding with FWI > 30, balancing false alarm reduction against early detection performance under extreme conditions.

Three specific gaps in the published literature motivate EgeoFireNet: (i) transformer-based fire detection systems have been evaluated primarily on benchmark datasets (D-Fire, Foggia, FiSmo) that do not represent the specific spectral and morphological characteristics of Pinus brutia smoke and flame; real-world validation on an operational multi-camera network with controlled test fires has not been reported; (ii) the integration of camera-derived fire perimeter segmentation as the initialisation source for operational CA spread prediction has been described in concept but not validated against real fire events in the Mediterranean context; and (iii) the end-to-end system latency from ignition to actionable alert encompassing detection, confirmation, spread forecast generation, and alert dispatch has not been empirically measured for any transformer-based camera fire detection system in an operational deployment.

3. Methodology

Chios island (858 km², population 52,000) is located in the northeastern Aegean Sea, 8 km off the Turkish coast, and hosts approximately 12,400 ha of Pinus brutia forest concentrated in the northern and central sectors of the island. The terrain is characterised by the Agio Galas ridge (maximum elevation 1,297 m a.s.l.) and multiple secondary ridges running broadly north-south, creating complex viewshed geometry that requires strategically placed high-elevation tower sites for wide-area coverage. The forest is classified by the Greek Forest Service as "extreme risk" throughout the July-August peak season, with wind-driven fire spread rates of 200-1,500 m/h documented in historical incidents. The 43-tower EgeoFireNet network was planned using LiDAR-derived 5 m DEM viewshed analysis and installed in partnership with the Chios Regional Fire Service over February-April 2024, with towers positioned to maximise joint coverage and minimise the number of single-camera blind spots.

Each tower hosts: a Hanwha Qnap dual-channel pan-tilt-zoom camera unit (1080p visible, 1080p NIR, 50 mm standard + 200 mm telephoto channels, motorised PTZ, −40° to +60° tilt range); an NVIDIA Jetson Orin NX 16 GB edge compute module in a Bud Industries NBD-16505 IP65 enclosure with active cooling fan (forced air, thermostat-controlled); a Teltonika RUTX11 industrial 4G/LTE router with dual SIM failover; a Kerlink iStation LoRaWAN gateway connected to the field sensor network; and a Victron Energy solar/battery power system (300 Wp monocrystalline panel, 200 Ah 24 V LiFePO4 battery bank providing 72-hour autonomy at design load). Total tower power budget: Jetson Orin NX 18 W + camera 8 W + router 6 W + fan 3 W + gateway 5 W = 40 W nominal.

The fire and smoke detection model architecture uses a Swin-T (Swin Transformer Tiny, 28M parameters) ImageNet-22K pretrained backbone [4] connected to a Mask R-CNN detection head for simultaneous object detection (bounding box) and instance segmentation (binary fire/smoke mask). The Swin-T feature pyramid produces feature maps at 1/4, 1/8, 1/16, and 1/32 of input resolution, which are processed by a Feature Pyramid Network (FPN) neck before the RPN and ROI Align stages of the Mask R-CNN head. This architecture enables simultaneous per-frame bounding box detection (for alert

triggering) and pixel-level segmentation (for fire perimeter extraction as CA spread prediction initialisation).

Training data for fine-tuning comprises the EgeoFireNet domain corpus: 18,600 annotated frames from the Aegean pine fire image archive (sourced from Hellenic Fire Corps documentation footage 2017-2023, UAV survey flights 2021-2023, and the Chios Regional Fire Service CCTV archive) plus 4,200 frames selected from the FLAME2 aerial multispectral benchmark [5]. The FLAME2 frames provide diversity in fire scale, viewing geometry, and fuel type beyond the Chios corpus. All frames were annotated by two independent trained annotators in CVAT v2.3 with bounding boxes (fire class, smoke class) and instance segmentation masks; annotation disagreements were resolved by a senior annotator. The 22,800-frame combined dataset was split 70:15:15 (train:validation:test) with stratification by source, fire class, and approximate fire size category.

Data augmentation addresses the distribution mismatch between training data (predominantly day-time visible-channel images) and the full operational deployment envelope (night-time NIR, variable haze, backlit conditions). Augmentation pipeline: per-channel colour jitter (±20% brightness, ±15% contrast, ±10% saturation); Gaussian noise injection (SNR 28-42 dB, simulating camera noise at night); random horizontal flip; mosaic compositing (4-image mosaic at 50% probability); copy-paste augmentation of fire instances onto non-fire background frames; simulated NIR channel conversion (luminance-only with spectral weighting); and random cloud/haze overlay (simulating early smoke diffusion into background). Fine-tuning was performed for 80 epochs on 4× NVIDIA A100 GPUs at Aristotle University HPC cluster (GRNET ARIS), with AdamW optimiser, initial learning rate 1×10^{-4}, cosine annealing, and layer-wise learning rate decay (backbone: 0.1× head rate). The final model was exported to ONNX and then TensorRT INT8 using the NVIDIA TensorRT 8.6 toolkit for Jetson Orin NX deployment.

The fire spread prediction module operates on a 10 m × 10 m grid covering the 12,400 ha Chios forest zone (1,240 × 1,000 cells). The CA state space is binary (burning / not-burning) with a memory buffer storing ignition time for each burned cell to enable fireline intensity estimation. State transition rules encode four coupled spread mechanisms:

(i) Wind-driven surface spread: the Rothermel rate-of-spread equation is evaluated for the wind speed and direction provided by the nearest LPWAN sensor node (one sensor per 4 km² average), parameterised using the Pinus brutia needle-litter fuel model (GR6 fuel category per the IFTDSS classification, adapted for Mediterranean conditions). Maximum rate of spread under 40 km/h wind at field-measured fuel moisture content (6-8% in dry conditions) is 450-650 m/h in the head fire direction.

(ii) Topographic modification: fire spread rate is multiplied by the slope factor φ_S = e^(3.533 × (tan β)^1.2) for upslope spread and φ_S = e^(−1.2 × tan β) for downslope spread, where β is the terrain slope angle computed from the 5 m DEM (Hellenic Cadastre, 2021).

(iii) Fuel model spatial variation: fuel load and fuel moisture are assigned per cell from the CORINE Land Cover 2018 and the Greek National Fuel Map (Velez-Munoz classification adapted for Greece), distinguishing dense Pinus brutia forest (GR6), degraded maquis (SH5), and rocky/sparse vegetation (NB1) cells. Firebreaks (roads, watercourses, rocky outcrops) are represented as cells with zero fuel load.

(iv) Spotting: long-range spot fires ignitions ahead of the fireline caused by wind-transported burning embers are implemented probabilistically: a spot fire is generated at random distance (uniformly distributed 0-1,500 m in the wind direction) with probability P_spot = 0.003 per burning boundary cell per time step under conditions FWI > 35 and wind speed > 25 km/h. This parameterisation follows the spotting model of Fernandes et al. calibrated for Pinus brutia.

The CA module is invoked at the moment of detection alert confirmation, with the fire perimeter initialised from the Swin Transformer instance segmentation mask geo-referenced using the camera calibration matrix and 5 m DEM. Real-time wind, humidity, and temperature from the LPWAN sensor network are ingested via MQTT at 60-second update intervals. A 15-minute forecast is generated by advancing the CA 90 time steps of 10-second duration. Parallelisation across the 1,240 × 1,000 cell grid on the Jetson Orin NX CPU (8-core ARM Cortex-A78AE, 2.2 GHz) using NumPy vectorised operations achieves a forecast generation time of 4.2 seconds.

The alert logic implements a three-tier escalation: (i) Level 1 (Monitor): single camera detection with confidence 0.65-0.85, logged and displayed on the Chios EOC dashboard without audio alert; (ii) Level 2 (Advisory): single camera detection confidence > 0.85 AND FWI > 15, or dual-camera concurrent confirmation within 90-second window at any confidence, triggers audible EOC alert and automated SMS to duty officer; (iii) Level 3 (Dispatch): Level 2 alert confirmed by duty officer OR single camera confidence > 0.92 with FWI > 30, triggers automatic resource dispatch request to Hellenic Fire Corps coordination centre with attached GeoJSON alert package. The GeoJSON alert package includes: fire location (GPS centroid from camera triangulation), detection confidence, camera imagery thumbnails from all detecting cameras, current CA spread forecast as a GeoJSON polygon (15-minute perimeter), and distance/bearing to nearest community and road access point. The package is formatted for direct import into the Hellenic Fire Corps' ARCGIS-based Common Operational Picture.

The Swin-T model is evaluated on the held-out 15% test set (3,420 frames: 2,280 fire frames, 1,140 smoke-only frames) using standard COCO detection metrics: mAP@0.5, mAP@0.5:0.95, per-class precision and recall at 0.5 confidence threshold, and F1-score. Instance segmentation performance is evaluated using the COCO mask mAP metric at IoU thresholds 0.5 and 0.75. Edge deployment performance (FPS, latency, power draw) is measured on a production Jetson Orin NX 16 GB module at nominal 15 W power mode using the TensorRT INT8 model at 640×640 input resolution, averaged over 500 inference cycles. A YOLOv8-nano baseline model trained on the same dataset under identical conditions provides the performance comparison benchmark.

14 controlled test fires were conducted across the Chios network from June to September 2024, in coordination with the Chios Regional Fire Service and the Hellenic Fire Corps Training Department. Each test fire used a standardised Pinus brutia needle-litter fuel pile of 1.5 kg dry mass (approximately 0.5 m^2 footprint) positioned at known GPS coordinates within the field of view of at least two network cameras at distances of 0.8-3.4 km. Ignition was by electronic remote igniter with GPS-timestamped trigger signal, providing a precise ignition time reference for alert latency measurement. Fires were extinguished within 3 minutes of ignition by pre-positioned fire service personnel.

Alert latency is defined as the elapsed time from ignition timestamp to the first confirmed Level 2 or Level 3 alert dispatch to the EOC. Spread prediction accuracy is evaluated for the 6 test events where the fire achieved sufficient areal extent (> 0.2 ha final burned area) for meaningful perimeter comparison; for these events, the 15-minute CA spread forecast (with the CA initialised from the Swin Transformer perimeter at detection time) is compared to the actual burned perimeter mapped by post-fire UAV photogrammetry at 2 cm GSD resolution. Comparison metrics: Sorensen-Dice similarity coefficient, false alarm ratio (predicted burning area outside actual perimeter), and miss ratio (actual burned area outside predicted perimeter).

4. Results

On the 3,420-frame Aegean pine test set, the Swin-T model achieves mAP@0.5 = 0.889 and mAP@0.5:0.95 = 0.627, with per-class performance: fire mAP@0.5 = 0.921, smoke mAP@0.5 = 0.857. Per-class F1-scores at 0.5 confidence threshold are: fire 0.942, smoke 0.882, combined 0.912. The YOLOv8-nano baseline achieves mAP@0.5 = 0.834 and F1 = 0.871, a 5.5 percentage-point mAP and 4.1 percentage-point F1 gap confirming the Swin Transformer's advantage for the Aegean pine fire detection task. The performance differential is most pronounced for smoke detection at small spatial scales: for smoke bounding boxes with area < 1,024 $pixels^2$ (subtending < 0.25% of the 640×640 frame area), Swin-T achieves AP = 0.813 vs. YOLOv8-nano AP = 0.722 a 12.6% improvement attributable to the Swin

Transformer's multi-scale attention mechanism capturing early-stage smoke signatures embedded in complex forested backgrounds.

Instance segmentation performance (mask mAP) is 0.761 at IoU 0.5 and 0.512 at IoU 0.75. The lower mAP@0.75 reflects the inherent challenge of precise fire boundary delineation in optically semi-transparent smoke environments, where the "fire/not-fire" boundary is gradual rather than sharp. For the CA spread prediction application, the mAP@0.5 segmentation accuracy is sufficient, as the CA model is relatively insensitive to precise perimeter shape and more sensitive to fire centroid location and approximate area both of which are well determined at mAP@0.5 level. Inference latency at TensorRT INT8 is 44.6 ms per frame (22.4 FPS) at 15 W power mode, with memory bandwidth utilisation of 62% confirming headroom for simultaneous CA computation on the Jetson Orin NX CPU cores.

Night-time performance (NIR channel input, grayscale converted to 3-channel pseudo-colour for Swin-T input) achieves mAP@0.5 = 0.821 - a 7.6 percentage-point reduction relative to daytime visible performance. Analysis of false negatives in the night-time subset reveals that 68% involve smoke with NIR transmittance > 0.85, where smoke plumes are nearly invisible against the night sky background. Future work will train a dedicated NIR-specific model branch and investigate thermal imaging integration to address this systematic night-time performance gap.

Across the 14 test fire events, the median end-to-end alert latency is 87 seconds (interquartile range 71-118 s). The latency decomposition is: smoke generation to first camera frame with detectable smoke 31 s (median); Swin-T detection to first Level 1 trigger 2.3 s; single-camera Level 1 to dual-camera Level 2 confirmation 41 s (median, dependent on second camera pan-to-target time); alert package assembly and EOC API transmission 6.2 s; SMS dispatch 3.8 s. The dominant latency component (47% of total) is the dual-camera confirmation window, which is a deliberate trade-off against false alarm rate. In the 3 test events where FWI exceeded 30 at ignition time, the single-camera dispatch pathway (confidence > 0.92 override) was triggered, reducing alert latency to 46 s (median) for those events.

Comparison with conventional aerial patrol detection in the Chios context: the Hellenic Fire Corps operates helicopter patrol over Chios on a 45-minute recurrence during Extreme danger days, and 90-minute recurrence on Very High danger days. Analysis of the 22 detected fires on Chios between 2018 and 2023 shows a mean patrol-based detection time of 23 minutes (range 8-67 minutes). The EgeoFireNet 87-second median alert latency represents a 15.9× improvement in mean detection time, and a factor of 5.5 improvement at the best-case patrol detection (8 minutes). This quantified detection improvement translates, using the fire outcome analysis cited in the introduction, from a 71% success rate for < 5-minute detection to a projected rate substantially above the 30% success rate achievable at 23-minute detection under High danger conditions.

For the 6 test events with sufficient burned area for spread prediction evaluation, the mean Sorensen-Dice similarity coefficient is 0.76 (range 0.68-0.84). The false alarm ratio (predicted burning area outside actual perimeter) averages 0.18, indicating an 18% overprediction of eventual burned area at 15 minutes. The miss ratio (actual burned area outside predicted perimeter) averages 0.09, indicating that the CA model misses 9% of the actual 15-minute spread extent. The overprediction bias is directionally preferred by fire managers: a conservative overestimate of spread extent supports pre-positioning resources ahead of the fire head, the primary tactical application of short-horizon spread forecasts. The highest-accuracy prediction (Dice = 0.84) occurred in Test Event 9 under steady 22 km/h northerly wind with uniform Pinus brutia fuel, where the simple wind-driven Rothermel model is most physically representative. The lowest-accuracy (Dice = 0.68, Event 6) occurred under rapidly shifting wind during a Meltemi gusting event, where wind direction changed 38° within the 15-minute forecast horizon a systematic limitation of any wind-driven CA model without real-time wind field updates.

CA forecast generation time of 4.2 seconds on the Jetson Orin NX 8-core CPU confirms real-time operational viability: the 4.2-second computation leaves 10.8 seconds for alert package assembly within the 15-second target from detection to first EOC alert API call. Memory allocation for the 1,240 × 1,000 ×

5-variable CA state array is 24.8 MB, well within the Jetson Orin NX 16 GB RAM capacity. The implementation uses NumPy 1.24 vectorised operations; profiling confirms that the wind-driven spread computation (97% of total CPU time) is dominated by the elementwise Rothermel model evaluation, which could be further accelerated by GPU offloading (estimated 0.8 s on the Jetson Orin NX GPU) in future implementations.

Over the 6-month operational period (June-November 2024, 43 cameras × 6 months = 258 camera-months), the system generated 1,847 Level 1 alerts, 63 Level 2 alerts, and 7 Level 3 dispatch alerts. Of the 63 Level 2 alerts, 12 were confirmed true fire detections (8 actual fires plus the 4 test events conducted during this period), 34 were confirmed false positives (primarily cloud shadow transitions in hazy conditions), and 17 remain unresolved (camera not reachable for confirmation). The Level 2 false alarm rate of 0.023 per camera per hour (34 false positives / [43 cameras × 24 h/day × 6 months × 30 d/month] = 0.023) is within the < 0.05 operational tolerance specified by the Hellenic Fire Corps for fixed camera networks.

System availability over the 6-month period was 97.8% (4 cameras out of service for > 24 hours due to: 2 solar charge controller failures, 1 4G router failure, 1 camera PTZ motor failure). No unrecovered data loss occurred due to local SD card buffering of all detection events. The primary maintenance activity was quarterly cleaning of camera domes and heat exchanger fins in the dusty and salt-marine environment, performed by the Chios Regional Fire Service maintenance team. Camera PTZ calibration drift (maximum 0.8° heading, 0.3° elevation) was automatically corrected by the weekly landmark-based calibration routine using GPS-surveyed mountaintop reflectors visible from each tower.

5. Discussion

The 5.5 percentage-point mAP@0.5 advantage of Swin-T over YOLOv8-nano on the Aegean pine test set is operationally significant: it corresponds to detecting a proportionally larger fraction of early-stage fire events before they grow beyond the suppression-feasible threshold. The advantage is most pronounced for small-scale smoke detections, precisely the category that matters most for early detection performance. The trade-off is a 2.2× higher inference latency (44.6 ms vs. 20.3 ms for YOLOv8-nano) and 3.2× higher parameter count (28M vs. 3.2M). At 22.4 FPS, the Swin-T model processes 1,344 frames per minute per camera substantially more than required for a 30-second temporal resolution detection system, leaving substantial compute headroom. The parameter count is not a binding constraint for the Jetson Orin NX with 16 GB RAM.

Future work will explore the training of larger Swin variants (Swin-S, 50M parameters; Swin-B, 88M parameters) on the expanded EgeoFireNet corpus to assess whether the detection performance ceiling has been reached with Swin-T, or whether the accuracy gap relative to the theoretical optimum can be further reduced. A further avenue is the integration of temporal information through video transformer architectures (VideoSwin, TimeSformer) that could exploit the temporal consistency of fire and smoke development to reduce false positives from transient cloud shadow and vehicle exhaust events.

The Rothermel-CA spread model achieves Dice = 0.76 under favourable steady-wind conditions but degrades to 0.68 under rapidly shifting winds, representing the fundamental limitation of any wind-driven CA model using point weather measurements. Three improvement pathways are identified: (i) assimilation of high-frequency (10-second) LPWAN wind sensor data across the network to update the CA wind field in near-real-time, reducing the lag between actual and modelled wind conditions; (ii) data-driven correction of the Rothermel model using historical fire-spread observations from Chios and comparable Aegean islands to recalibrate the fuel model parameters for local Pinus brutia stands; and (iii) machine learning emulation of the full CA spread process using a graph neural network trained on CA simulation outputs, enabling probabilistic ensemble spread forecasting that captures wind uncertainty within the same computation time [5].

The spotting model parameterisation (P_spot = 0.003 per boundary cell per time step at FWI > 35) was not validated against independent Chios fire event data due to the limited number of observed spotting events in the test fire corpus. Dedicated spotting model validation using UAV-observed historical fires from the 2021 and 2023 Chios seasons is planned, and the spotting probability will be recalibrated if available data warrant.

The capital cost of the EgeoFireNet 43-tower Chios installation is estimated at EUR 1.2 million (EUR 27,900 per tower including civil works, tower, camera, edge compute, power, and communications). The annual operating cost is approximately EUR 85,000 (camera cleaning, maintenance visits, communication contracts, software updates). The primary value proposition is the reduction in suppression cost attributable to earlier detection: the Greek Forest Protection Study [1] estimates that fires detected within 5 minutes incur mean suppression costs of EUR 12,000 per ha, versus EUR 340,000 per ha for fires not detected until 30+ minutes. Using the EgeoFireNet median detection time of 87 seconds and assuming detection of 3-5 fires per season on the monitored 12,400 ha, the avoided suppression cost relative to the 23-minute patrol baseline is estimated at EUR 1.8-3.2 million per season a payback period of well under one year on the installation cost.

Scaling to the 3.5 million ha of extreme/very-high risk Greek forest at Chios network density (43 towers / 12,400 ha = 3.5 towers/1,000 ha) would require approximately 12,200 towers at an estimated capital investment of EUR 340 million and annual operating cost of EUR 24 million. While substantial, this investment is comparable to one to two years of Greek aerial firefighting operational expenditure during severe seasons, and provides continuous coverage that aerial patrol cannot. A phased deployment prioritising the 1.0 million ha of highest-risk zone (Greek islands plus Attica/Corinthia/Laconia coastal zone) would require approximately 3,500 towers at EUR 98 million capital cost.

6. Conclusions

EgeoFireNet, a 43-camera edge-AI forest fire monitoring system combining Swin Transformer detection with cellular automaton spread prediction, has been presented and validated in an operational Aegean pine forest deployment on Chios island, Greece. The system represents the first documented operational deployment and comprehensive test-fire validation of a transformer-based fire camera network in the Eastern Mediterranean. Key quantitative findings are: (i) Swin-T fire/smoke detection achieves mAP@0.5 = 0.889 and F1 = 0.912 at 22.4 FPS on the Jetson Orin NX (TensorRT INT8), outperforming the YOLOv8-nano baseline by 5.5 percentage-points in mAP@0.5; (ii) CA fire spread prediction achieves Dice = 0.76 for 15-minute forecasts relative to validation test fires, with 4.2-second generation time on the edge node CPU; (iii) end-to-end alert latency from ignition to EOC dispatch is 87 seconds (median), a 15.9× improvement over conventional aerial patrol detection; (iv) operational false alarm rate of 0.023 Level 2 alerts per camera per hour is within the Hellenic Fire Corps tolerance; and (v) system availability of 97.8% over the 6-month operational period confirms deployment-environment robustness.

The EgeoFireNet architecture provides a validated, replicable technical baseline for national-scale deployment of transformer-based fire detection networks across the Greek high-risk forest zone, with a quantified economic case supporting cost-recovery within a single fire season. Future development priorities are: night-time performance improvement through thermal imaging integration and dedicated NIR model branches; spread prediction accuracy improvement through real-time wind field assimilation and ML-based Rothermel calibration; SAR satellite data integration for overnight gap coverage; extension of the training dataset to additional Eastern Mediterranean fuel types; and longitudinal evaluation of detection-outcome relationships using multi-season suppression records.

REFERENCES

[1] European Forest Fire Information System (EFFIS), Annual Fire Danger Season Analysis 2023. Ispra: European Commission Joint Research Centre, 2024.

[2] A. Mohapatra and T. Trinh, "Early wildfire detection technologies in practice A review," Sustainability, vol. 14, no. 19, pp. 12270, Sep. 2022.

[3] T. Nawaz, O. Maguire, K. G. Clarke, R. Holsgrove, A. Al-Mamoon, and R. Bhakta, "Forest fire detection using distilled vision transformers for deployment on edge devices," preprint, Staffordshire University, 2024.

[4] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in Proc. ICCV, pp. 10012-10022, 2021.

[5] B. H. Hopkins, L. O. O'Neill, F. A. Afghah, A. R. Razi, A. W. Watts, P. F. Fule, and J. C. Coen, "FLAME 2: Fire detection and modeLing: Aerial multi-spectral imagE dataset," IEEE DataPort, 2022.

[6] R. Ghali and M. A. Akhloufi, "Deep learning approaches for wildland fires remote sensing: Classification, detection, and segmentation," Remote Sens., vol. 15, no. 7, pp. 1821, Mar. 2023.

[7] R. C. Rothermel, A Mathematical Model for Predicting Fire Spread in Wildland Fuels. Ogden: USDA Forest Service Research Paper INT-115, 1972.

[8] L. Mu, Y. Yang, B. Wang, Y. Zhang, N. Feng, and X. Xie, "Edge computing-based real-time forest fire detection using UAV thermal and color images," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 18, pp. 6760-6771, 2025.

[9] G. Coleman, W. Salter, and M. Walsh, "OpenWeedLocator (OWL): An open-source, low-cost device for fallow weed detection," Sci. Rep., vol. 12, no. 1, pp. 170, Jan. 2022.

[10] A. Stergiopoulos, G. Botzoris, and C. Basbas, "Optimal placement of forest fire detection cameras using geographic information systems and analytic hierarchy process," Appl. Sci., vol. 13, no. 6, pp. 3563, Mar. 2023.

[11] J. Li, R. Xu, and Y. Liu, "An improved forest fire and smoke detection model based on YOLOv5," Forests, vol. 14, no. 5, pp. 833, May 2023.

[12] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient deep CNN-based fire detection and localization in video surveillance applications," IEEE Trans. Syst. Man Cybern. Syst., vol. 49, no. 7, pp. 1419-1434, Jul. 2019.

[13] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," in Proc. ICLR, 2021.

[14] M. Wang, X. Li, and J. Zhao, "Real time fire and smoke detection using vision transformers and spatiotemporal learning," Sci. Rep., 2026.

[15] K. Das, J. Poovvancheri, S. Flesc, R. C. Calidonn, and D. Chen, "Emerging trends in wildfire detection through the lens of computer vision," IEEE Access, vol. 14, pp. 20201-20228, 2026.

[16] L. Papa, P. Russo, I. Amerini, and L. Zhou, "A survey on efficient vision transformers: Algorithms, techniques, and performance benchmarking," IEEE Trans. Pattern Anal. Mach. Intell., pp. 1-20, 2024.

[17] A. Khan, B. Hassan, S. Khan, R. Ahmed, and A. Abuassba, "DeepFire: A novel dataset and deep transfer learning benchmark for forest fire detection," Mob. Inf. Syst., vol. 2022, pp. 5358359, 2022.