APP4FARM, a project for the sustainable use of resources in agriculture

C. Carnevale[1,2], A. Adessi[3], L. O'Faolain[4,5], M. Saberioon[6], V. Sberveglieri[7,8], A. Bellabarba[3,9], D. Belmonte[8,10], G. Biagi[4,5], F. Decorosi[3,9], S. De Nardi[1,2], G. C. R. Devarapu[4,5], E. Nunez-Carmona[7], G. Piccoli[1], E. Poeta[7], S. Raccagni[1], L. Sangiorgi[1,2], V. V. B. Yallapragada[4,5]

***Abstract*— Efficient nitrogen management remains a major challenge in modern agriculture due to the low nitrogen use efficiency (NUE) of cropping systems and the associated environmental impacts of reactive nitrogen emissions. APP4FARM proposes an integrated cyber-physical framework combining distributed gas sensing, soil microbial functional analysis, satellite data, and artificial intelligence to enable continuous monitoring and short-term forecasting of nitrogen-related dynamics in agricultural systems. The system integrates low-cost metal oxide (MOX) gas sensors with reference photonic sensors, in-field meteorological measurements, Copernicus satellite products, and soil molecular characterization based on quantitative PCR targeting nitrogen cycling functional genes. Biological validation in High Yield (HY) and Low Yield (LY) zones demonstrated significant differences in nitrification and denitrification gene abundance, supporting mechanistic interpretation of nitrogen transformation processes. Artificial intelligence models, including ARX, ANN, CNN, and Hybrid CNN (HCNN), were implemented for pointwise and spatial forecasting. In the German pilot, pointwise models achieved correlation coefficients up to 0.9967 for soil-related variables. For satellite-based prediction, HCNN models reached correlations above 0.95 for Soil Water Index at a 3-day forecast horizon, while ANN models provided robust predictions of NDVI and NDMI. The results demonstrate the feasibility of integrating multi-source sensing and AI-based forecasting within a Decision Support System capable of supporting environmentally sustainable fertilizer management. APP4FARM represents a scalable approach for linking real-time emission monitoring, biological soil processes, and predictive analytics to improve nitrogen efficiency and reduce greenhouse gas emissions in agroecosystems.**

I. INTRODUCTION

APP4FARM is a research and innovation project funded under the ICT-AGRI-FOOD transnational call, aimed at advancing digital solutions for sustainable agriculture through the integration of sensing technologies and artificial intelligence. The project addresses one of the most critical challenges in modern agroecosystems: the efficient and environmentally responsible management of nitrogen (N). Nitrogen is the primary nutrient underpinning agricultural productivity and a fundamental constituent of biomolecules such as proteins, nucleic acids, and chlorophyll. Although dinitrogen (N_2) constitutes approximately 78% of the Earth's atmosphere, its chemical inertness renders it largely unavailable to most living organisms. The industrial fixation of atmospheric N_2 into ammonia (NH_3) via the Haber–Bosch process has dramatically increased the global availability of reactive nitrogen, enabling unprecedented growth in agricultural production and supporting rapid population expansion. It has been estimated that nearly 80% of the nitrogen present in human tissues originates from synthetic fixation processes [1]. However, the anthropogenic amplification of nitrogen inputs has profoundly altered the global nitrogen cycle [2]. In agricultural systems, nitrogen use efficiency (NUE)—defined according to the mass balance principle as the ratio between nitrogen output and nitrogen input (NUE = N_{output}/N_{input}), as proposed by the EU Nitrogen Expert Panel—remains suboptimal [3]. Under non-optimal management conditions, up to 70% of applied nitrogen may be lost from the agroecosystem, generating both economic losses and significant environmental externalities [4]. Nitrogen losses occur through multiple pathways, including ammonia (NH_3) volatilization, nitrate (NO_3^-) leaching and runoff, and gaseous emissions of dinitrogen (N_2), nitrous oxide (N_2O), and nitrogen oxides (NO, NO_2) driven by nitrification–denitrification processes. These reactive nitrogen species are biologically, photochemically, and radiatively active. NH_3 and NO_x contribute to secondary particulate matter formation and tropospheric ozone production, while N_2O is a potent greenhouse gas and a contributor to stratospheric ozone depletion [2], [5]. The magnitude and dynamics of these emissions are influenced by multiple interacting factors, including nitrogen application rate, crop type, fertilizer formulation, soil organic carbon content, pH, texture, and climatic variables [6]. Climatic and hydrological conditions critically modulate nitrogen loss pathways: ammonia volatilization is enhanced under warm and humid conditions in the absence of rainfall, whereas nitrate leaching and denitrification are favored when water inputs exceed evapotranspiration [3]. These nonlinear interactions highlight the limitations of static indicators such as NUE alone, which quantify nitrogen surplus but do not resolve the temporal and

[1]Department of Mechanical and Industrial Engineering, Università di Brescia, Via Branze 38, I-25123, Brescia, Italy claudio.carnevale@unibs.it
[2]DATINNOVA srl, Via Cipro 1, I-25124, Brescia, Italy
[3]Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Piazzale delle Cascine, 18, 50144, Florence, Italy
[4]Center for Advanced Photonics and Process Analysis, Munster Technological University, T12P928 Cork, Ireland
[5]Tyndall National Institute, T12R5CP Cork, Ireland
[6]GFZ Helmholtz Centre for Geosciences Remote Sensing and Geoinformatics Telegrafenberg, 14473 Potsdam, Germany
[7]Institute of Bioscience and Bioresources (CNR-IBBR), National Research Council, URT-Re, Via J.F. Kennedy, 17/i, 42124 Reggio Emilia, Italy
[8]Nano Sensor Systems srl (NASYS), Via Alfonso Catalani, 9, 42121 Reggio Emilia, Italy
[9]Genexpress Laboratory, Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, I-50019 Sesto Fiorentino, Italy
[10]Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, Viale delle Scienze, Ed. 16, 90128 Palermo, Italy

spatial dynamics of specific emission fluxes. A mechanistic understanding of soil reactive nitrogen emissions, linked to underlying biogeochemical processes, is therefore essential for the development of advanced nitrogen management strategies capable of reconciling agronomic productivity with environmental protection. Despite this need, current measurement systems present substantial constraints. Chamber-based techniques provide high analytical sensitivity but operate discontinuously and are labor-intensive, while reference-grade analyzers are costly and unsuitable for distributed deployment. Within this context, APP4FARM proposes a systematic approach that merges heterogeneous sensing technologies with artificial intelligence to enable continuous, spatially distributed, and biologically meaningful monitoring of nitrogen dynamics in agricultural environments. The project integrates conductometric gas sensors based on semiconducting metal oxides (MOX) with photonic sensors based on Quartz-Enhanced Photoacoustic Spectroscopy (QEPAS), coupled with data-driven models including autoregressive approaches and recurrent neural networks for real-time anomaly detection and short-term forecasting. The overall objective of APP4FARM is the implementation of a smart-farming ICT infrastructure enabling extensive monitoring and forecasting of nitrogen losses to support efficient fertilizer management. The system combines (i) a monitoring platform collecting meteorological variables and nitrogen emission data with (ii) a Decision Support System (DSS) providing real-time information and forecasts up to three days in advance. This architecture aims to optimize agronomic performance while minimizing greenhouse gas emissions (notably N_2O) and atmospheric nitrogen oxides, thereby contributing to improved air quality and climate mitigation. Beyond the farm scale, APP4FARM enhances transparency along the agri-food supply chain by enabling data accessibility for regulatory authorities and stakeholders, supporting green labeling initiatives, and empowering consumers with traceable information on environmental impact and production conditions. Unlike existing agricultural monitoring solutions focused on isolated sensing or standalone forecasting approaches, APP4FARM integrates distributed gas sensing, microbial functional validation, satellite-derived indicators, and heterogeneous AI forecasting models within a unified cyber-physical framework. The proposed architecture enables both continuous environmental monitoring and biologically interpretable forecasting of nitrogen-related dynamics. The paper is organized as follows. Section II describes the APP4FARM system architecture, sensing infrastructure, and AI models. Section III presents the German pilot study and the validation results. Finally, Section IV summarizes the main conclusions and future developments.

II. Materials and Methods

A. Overall System Architecture

APP4FARM is designed as a distributed cyber-physical system integrating in-field sensing, laboratory soil analysis, external environmental data streams, and artificial intelligence within a unified ICT infrastructure. The architecture is structured into three functional layers: (i) data acquisition, (ii) core AI modelling and data fusion, and (iii) Decision Support System (DSS). The DSS integrates real-time measurements and short-term forecasts to support operational decisions related to fertilizer application timing and environmental risk mitigation. The system operates at pilot scale in High Yield (HY) and Low Yield (LY) fields, enabling spatially explicit monitoring of nitrogen dynamics and supporting fertilizer management decisions through short-term forecasting (up to 72 hours).

B. Data Acquisition

Three main information sources are integrated: (1) in-field gas sensor networks, (2) soil sampling and molecular analysis, and (3) external environmental data, including satellite products and nearby monitoring stations.

1) In-field Gas Sensor Network: Reactive nitrogen species are continuously monitored using distributed sensing nodes centered around Metal Oxide (MOX) conductometric gas sensors. MOX devices operate via resistance modulation induced by surface redox reactions between adsorbed oxygen species and target gases. Their low cost, compactness, and sensitivity make them suitable for spatially distributed deployment in agricultural pilots. Sensor arrays are adopted to enhance selectivity through multivariate pattern recognition, following embedded AI strategies for low-cost MOX systems [7]. Each node includes temperature and relative humidity compensation modules to mitigate environmental cross-sensitivity, critical under field conditions. Selected nodes integrate Quartz-Enhanced Photoacoustic Spectroscopy (QEPAS) modules to provide high-selectivity reference measurements. These reference-grade data are used for calibration, drift correction, and periodic validation of the MOX network. Meteorological variables (air temperature, humidity, solar radiation, wind speed and direction, precipitation) are co-registered with gas measurements, as these factors strongly influence nitrogen transformation and emission pathways [6].

2) Soil Sampling and qPCR Analysis: Soil sampling was conducted in both HY and LY pilots immediately after MOX sensor installation. Sampling was restricted to a circular plot with 2 m radius centered on the sensor (Figure 1). After removal of surface litter, 12 random subsamples were collected at 10 cm depth within predefined sectors surrounding the sensor. Samples were stored at $-80°C$ until analysis.

Total DNA was extracted from 0.5 g of soil following [8]. DNA concentration and purity were assessed using a NanoDrop Lite Plus spectrophotometer (Thermo Fisher Scientific). Quantitative PCR (qPCR) assays targeted total bacteria (16S rRNA) and functional genes involved in nitrogen cycling: *nifH* (nitrogen fixation), *amoA* (nitrification), *nirK*, *nirS*, and *nosZ* (denitrification). Amplifications were performed using an ABI StepOnePlus Real-Time PCR System with SYBR Green chemistry. Thermal cycling consisted of an initial activation step at 95°C for 10 min, followed by 40 cycles with gene-specific annealing temperatures, and melt curve analysis to confirm specificity. Gene abundances

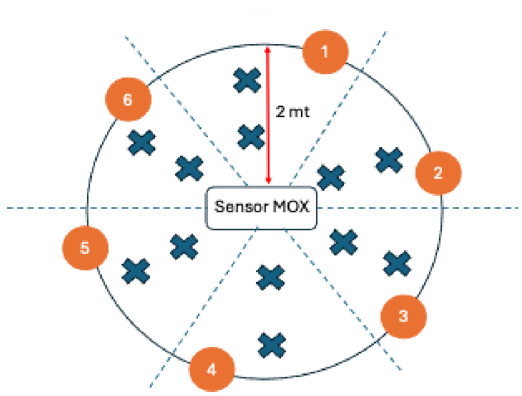

Fig. 1. Soil sampling scheme within the 2 m radius circular plot surrounding the MOX sensor in HY and LY pilots.

were expressed as copy numbers per gram of soil. The N_2O reduction potential was calculated as:

$$\frac{nosZ}{amoA + nirK}$$

while the production potential was computed as its reciprocal. Statistical analyses were performed in R (v4.4.3), after normality testing (Shapiro–Wilk), using Student's t-test or Wilcoxon rank tests depending on data distribution. This molecular characterization enables biological validation of field gas measurements by linking observed emissions to microbial functional potential in the surrounding soil matrix. Table I reports the primer pairs used for gene quantification.

3) External Environmental Data: External datasets are integrated to enhance spatial coverage and contextual interpretation. Satellite data are primarily sourced from the Copernicus program (Sentinel missions), including:

- Sentinel-2 multispectral imagery for vegetation indices (e.g., NDVI) and crop status,
- Sentinel-5P atmospheric products for NO_2 column density,
- Land surface temperature and soil moisture proxies.

These datasets provide spatially continuous information on crop vigor, atmospheric composition, and environmental stress conditions influencing nitrogen emissions. Additionally, data from nearby regulatory air-quality monitoring stations are incorporated to validate regional NO_x trends and improve model calibration. External measurements are harmonized temporally with in-field data streams and assimilated into the modelling framework using data fusion approaches [15], [16].

C. Core AI: Forecasting Models

The APP4FARM modelling layer implements a set of data-driven forecasting models aimed at predicting agro-meteorological and soil-related variables from multiscale inputs. The adopted models are grouped into (i) *pointwise* approaches operating on single-site time series and (ii) *spatial* approaches operating on gridded (raster) inputs. The following subsections summarize the model classes and the common training/validation protocol.

1) Pointwise models:

a) ARX and Piece-Wise Linear extensions.: AutoRegressive with eXogenous inputs (ARX) models are used to predict a target variable by combining its past values with lagged exogenous drivers. The generic ARX form is:

$$y(t) = \sum_{i=1}^{n_a} a_i \, y(t-i) + \sum_{j=1}^{n_b} b_j \, u(t - n_k - j + 1) + e(t) \quad (1)$$

where $y(t)$ is the output at time t, $u(t)$ is the exogenous input, n_a and n_b are autoregressive and exogenous orders, n_k is the input delay, $\{a_i\}$ and $\{b_j\}$ are model parameters, and $e(t)$ is the residual. To address nonlinearities while preserving computational efficiency, a piece-wise linear (PWL) strategy can be adopted by approximating nonlinear mappings with multiple linear regimes (segments) and optimized breakpoints.

b) ANN.: Artificial Neural Networks (ANNs) are used to capture nonlinear dependencies between inputs and outputs. A neuron computes:

$$o = \phi\left(\sum_{i=1}^{m} w_i x_i + b\right) \quad (2)$$

where $\{x_i\}$ are the input features, $\{w_i\}$ weights, b bias, $\phi(\cdot)$ a nonlinear activation function, and o the neuron output. Model training minimizes a loss function, usually the mean squared error (MSE) via backpropagation and gradient-based optimization.

2) Spatial models:

a) CNN.: Convolutional Neural Networks (CNNs) exploit local spatial correlations in gridded inputs (e.g., remote sensing maps). Given an input tensor X and a set of learnable kernels $\{K_j\}$, a convolutional layer produces feature maps:

$$Y_j = \sigma\left(\sum_{i} K_{j,i} * X_i + b_j\right) \quad (3)$$

where $*$ denotes convolution/cross-correlation, b_j is a bias term, and $\sigma(\cdot)$ is an activation function. Pooling layers can reduce spatial dimensionality and mitigate overfitting.

b) HCNN (Hybrid CNN).: To jointly process heterogeneous inputs, a Hybrid CNN (HCNN) merges (i) gridded inputs processed by convolutional/pooling stages and (ii) pointwise variables concatenated at a merge stage. The merged representation is then processed by dense layers to generate multi-step forecasts.

3) Training, hyperparameter tuning, and evaluation: ANN-based models employed ReLU activation functions in hidden layers and linear output layers, while CNN/HCNN architectures combined convolutional and dense layers optimized through random-search hyperparameter tuning. Hyperparameters are optimized via Random Search. Models are trained on 70–80% of the available data, with 20% of the training set reserved for validation. Early stopping is applied with a patience of 15 iterations, halting training if validation loss does not improve over 15 consecutive steps. Final evaluation is performed on a held-out test set (20–30%). Performance is quantified using normalized mean absolute error (NMAE) and Pearson correlation coefficient.

TABLE I

PRIMER PAIRS USED FOR QPCR QUANTIFICATION OF MICROBIAL GROUPS IN HY AND LY PILOTS.

Target group	Target gene	Primer	Sequence (5'–3')	Annealing T (°C)	Reference
Bacteria	16S rRNA	Bac341F	CCTACGGGAGGCAGCAG	60	[9]
		Bac805R	GGACTACHVGGGTWTCTAA		[10]
Nitrogen-fixing bacteria	nifH	PolF	TGCGAYCCSAARGCBGACTC	55	[11]
		PolR	ATSGCCATCATYTCRCCGGA		
Nitrifying bacteria	amoA	AmoA-1F	GGGGTTTCTACTGGTGGT	57	[12]
		AmoA-2R	CCCCTCKGSAAAGCCTTCTTC		
Denitrifying bacteria	nirK	F1aCu	ATCATGGTSCTGCCGCG	57	[13]
		R3Cu	GCCTCGATCAGRTTGTGGTT		
	nosZ	nosZF	CGYTGTTCHTCGACAGCCAG	55	[14]
		nosZ1733R	ATRTCGATCARCTGBTCGTT		
	nirS	Cd3aF1	GTSAACGTSAAGGARACSGG	52	[13]
		R3cd	GASTTCGGRTGSGTCTTGA		

III. CASE STUDY

A. Pilot description: German site

The German pilot combines long-term meteorological records, high-frequency in-field measurements, and satellite-derived gridded products. The integrated dataset supports both pointwise forecasting (single-site variables) and spatial forecasting (raster indices).

B. Monitoring Validation

Monitoring validation was performed to assess (i) hardware robustness, (ii) structural integrity of the sensing layer, and (iii) the capability of the MOX system to discriminate different emission regimes under controlled conditions prior to field deployment.

1) Hardware configuration and substrate optimization: MOX sensing elements were fabricated using SnO_2-based nanostructured materials (SnO_2, SnO_2-Au, SnO_2-Pd; 1% doping). During optimization, the substrate material was changed from quartz to alumina in order to improve thermal control, conditioning stability, and baseline reproducibility. This modification resulted in improved signal stability and reduced power consumption (from 3 W to 1.5 W). A plug-and-play configuration was adopted, enabling modular integration of three sensing elements within a compact architecture.

2) Structural characterization of the sensing layer: Following screen-print deposition and sintering (750°C, 3 h), the sensing layer morphology was characterized by Scanning Electron Microscopy (SEM). The microstructure (Fig. 2) exhibits high porosity and particle sizes ranging from tens to hundreds of nanometers. Such morphology increases the effective surface area and enhances gas–surface interaction, supporting sensitivity to reactive nitrogen compounds.

3) Discrimination capability under controlled conditions: Controlled experiments were conducted to evaluate the ability of the MOX system to discriminate between different emission regimes. Sensor responses were processed using multivariate pattern recognition techniques, and Linear Discriminant Analysis (LDA) was applied to evaluate clustering and separability. The following analyses were performed to biologically validate the sensing and forecasting framework.

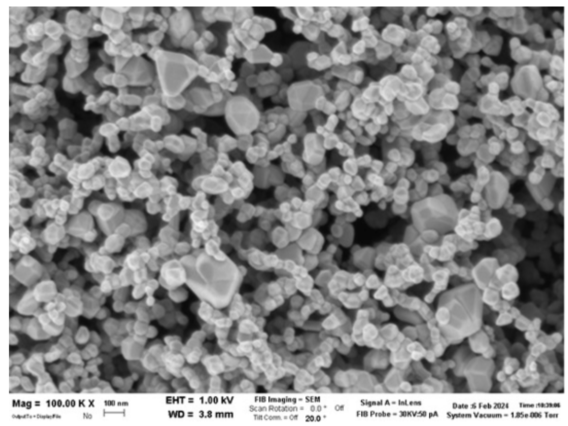

Fig. 2. SEM micrograph of the MOX sensing layer showing porous morphology and broad particle size distribution.

C. Microbial Functional Validation

To biologically validate nitrogen cycling dynamics in the two yield zones, quantitative PCR (qPCR) analyses were performed to quantify total bacterial abundance and functional genes involved in nitrogen transformations. Total bacterial abundance was estimated targeting the 16S rDNA gene. Functional genes included: *nifH* (biological nitrogen fixation), *amoA-AOB* (nitrification), *nirK* and *nirS* (nitrite reduction during denitrification), and *nosZ* (nitrous oxide reduction).

1) Functional gene abundance: Based on 16S rDNA copy numbers, total bacterial abundance was significantly higher in the High Yield (HY) pilot compared to the Low Yield (LY) pilot (Fig. 3A). The abundance of nitrogen cycling functional genes showed distinct patterns between yield zones (Fig. 3B–F). Specifically:

- *nifH* abundance was significantly higher in HY soils, suggesting enhanced biological nitrogen fixation potential.
- *amoA* abundance was significantly higher in HY soils, indicating increased nitrification potential.
- *nirK* and *nirS* were significantly enriched in HY areas, supporting enhanced denitrification capacity.
- In contrast, *nosZ* abundance did not differ significantly between HY and LY zones.

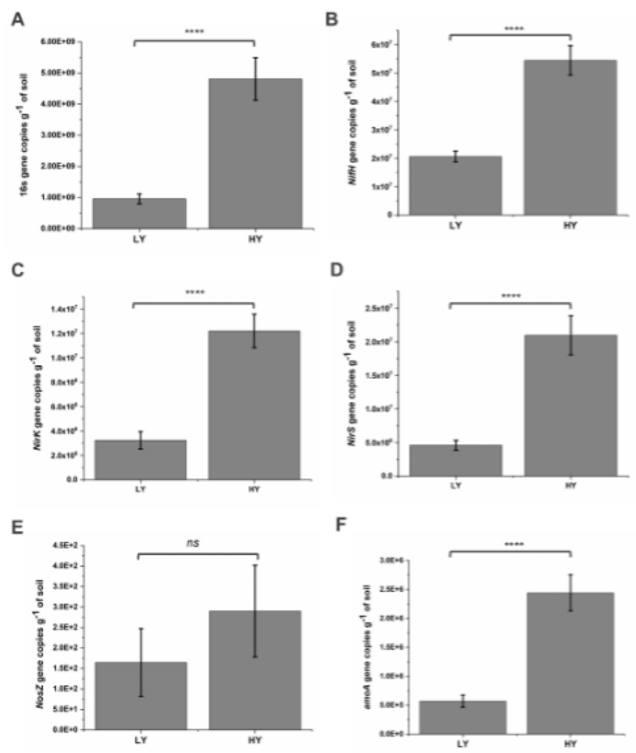

Fig. 3. Quantification of (A) total bacteria, (B) nitrogen-fixing bacteria, (C–E) denitrifying bacteria and (F) nitrifying bacteria in low and high-field pilots. Error bars indicate standard error. Statistical differences: ns (not significant); * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

2) N_2O production and reduction potential: To assess long-term N_2O emission potential, production and reduction potentials were calculated from gene copy ratios. Although differences were not statistically significant, HY soils exhibited higher predicted N_2O production potential and lower N_2O reduction potential compared to LY soils (Fig. 4). The lack of statistical significance may be attributed to variability observed within LY samples.

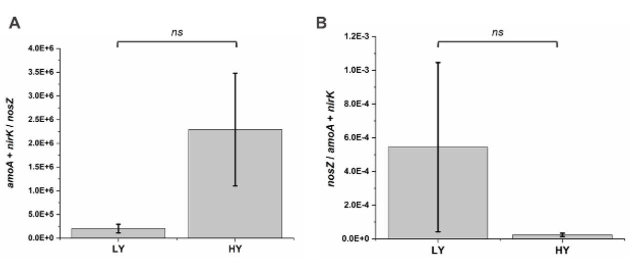

Fig. 4. N_2O production potential (A) and reduction potential (B) in low and high-field pilots. Error bars represent standard error. Statistical differences: ns (not significant); * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

3) Pointwise models – German pilot: Several model structures were tested (ARX, PARX, ARX_noInt, PARX_noInt, RT, RF, MLP) under different autoregressive and exogenous configurations to forecast agro-meteorological variables. The best-performing model for each output variable is reported in Table II, together with its correlation (CORR) and normalized mean absolute error (NMAE).

TABLE II

SELECTED POINTWISE MODEL ARCHITECTURE AND PERFORMANCE FOR THE GERMAN PILOT.

Output Variable	Best Model	CORR	NMAE
Dew Point	ARX_noInt(1,2,1)	0.8851	0.3755
Soil Permittivity	ARX_noInt(1,1,1)	0.9628	0.0912
Watermark	ARX(3,2,1)	0.9967	0.0400
Soil Moisture	PARX_noInt(1,1,1)	0.9660	0.1140
Soil Temperature	ARX(1,2,1)	0.9841	0.1042

The results show high predictive performance for soil-related variables, with correlation coefficients exceeding 0.96 for Soil Permittivity, Watermark, Soil Moisture, and Soil Temperature. Dew Point forecasting exhibits lower correlation and higher NMAE compared to the other variables, reflecting higher short-term variability and atmospheric sensitivity.

4) Satellite-based Forecast Validation: Satellite-derived variables were modelled to forecast Soil Water Index (SWI), Normalized Difference Vegetation Index (NDVI), and Normalized Difference Moisture Index (NDMI). Different model families were evaluated during development, including Convolutional Neural Networks (CNN), Hybrid CNN (HCNN), and Artificial Neural Networks (ANN). In the following, only the performance of the best-performing model for each target variable is reported, as identified during the model selection phase.

a) Soil Water Index (SWI): For SWI prediction, both CNN and HCNN architectures were tested. Due to the heterogeneous structure of the input data (gridded satellite information combined with pointwise meteorological drivers), the Hybrid CNN (HCNN) demonstrated superior predictive capability and was selected as the final model. Performance was evaluated for multi-step forecasting horizons (t, $t+1$, $t+2$, $t+3$ days) using Normalized Mean Absolute Error (NMAE) and Pearson correlation coefficient (CORR).

The HCNN achieved very high predictive accuracy, particularly for the deeper soil moisture layer ($T = 10$), where correlation values remained above 0.95 even at the $t + 3$ horizon. As expected, performance progressively decreases with increasing forecast horizon.

b) Vegetation Indices (NDVI and NDMI): Vegetation indices were modelled using ANN architectures. Although convolutional models were explored, the limited number of valid satellite acquisitions (due to cloud masking constraints) favored the use of ANN models, which showed better generalization under sparse temporal sampling. Performance metrics for NDVI and NDMI forecasting are reported below for the selected best-performing ANN configurations. Prediction accuracy decreases with increasing forecast horizon, particularly for NDVI at $t+3$, reflecting the higher variability of vegetation dynamics and the reduced temporal density of valid satellite observations. Nevertheless, correlation values remain within acceptable operational thresholds for integration into the Decision Support System (DSS).

TABLE III

SWI FORECASTING PERFORMANCE (GERMAN PILOT). BEST-PERFORMING MODEL: HCNN.

Output	Model	Step	NMAE	CORR
SWI $T = 2$	HCNN	t	0.0767	0.9435
		$t+1$	0.1009	0.8971
		$t+2$	0.1193	0.8604
		$t+3$	0.1344	0.8195
SWI $T = 10$	HCNN	t	0.0277	0.9898
		$t+1$	0.0372	0.9821
		$t+2$	0.0464	0.9725
		$t+3$	0.0554	0.9599

TABLE IV

VEGETATION INDEX FORECASTING PERFORMANCE (GERMAN PILOT). BEST-PERFORMING MODEL: ANN.

Output	Model	Step	NMAE	CORR
NDVI	ANN	t	0.0911	0.8714
		$t+1$	0.0899	0.8721
		$t+2$	0.0994	0.8584
		$t+3$	0.1045	0.7812
NDMI	ANN	t	0.1930	0.9028
		$t+1$	0.1770	0.9196
		$t+2$	0.2333	0.8434
		$t+3$	0.2246	0.8194

IV. CONCLUSIONS

This study presents the implementation and validation of the APP4FARM framework, integrating distributed gas sensing, microbial functional analysis, satellite data, and artificial intelligence to support nitrogen management in agricultural systems. Monitoring validation confirmed the robustness of the MOX-based sensing platform, including structural characterization of the sensing layer and the ability to discriminate emission regimes through multivariate analysis. Microbial functional validation revealed significantly higher abundance of nitrogen fixation, nitrification, and denitrification genes in High Yield zones, providing biological context for observed nitrogen transformation dynamics. Forecasting validation demonstrated strong predictive performance for both pointwise and spatial models. ARX-based models achieved high correlation for soil and meteorological variables, while Hybrid CNN models provided highly accurate multi-step forecasts of Soil Water Index. ANN models successfully predicted vegetation indices despite irregular satellite acquisition frequency. The combined results highlight the effectiveness of integrating heterogeneous sensing technologies with data-driven modelling to achieve continuous, biologically meaningful monitoring of nitrogen-related processes. By linking emission detection, soil functional potential, and short-term forecasting within a unified ICT infrastructure, APP4FARM provides a practical pathway toward improving nitrogen use efficiency, reducing reactive nitrogen losses, and supporting climate-smart agricultural practices. Although the current validation was performed on the German pilot site, future developments will include cross-site validation under different climatic, soil, and agronomic conditions.

V. ACKNOWLEDGMENTS

APP4FARM is part of the ERA-NET Cofund ICT-AGRI-FOOD, with funding provided by national sources [BMEL, Germany; DAFM, Ireland; MASAF, Italy] and co-funding by the European Union's Horizon 2020 research and innovation program, Grant Agreement number 862665.

REFERENCES

[1] R. W. Howarth, "Coastal nitrogen pollution: A review of sources and trends globally and regionally," *Harmful Algae*, 2008.

[2] J. N. Galloway, A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, and M. A. Sutton, "Transformation of the nitrogen cycle: recent trends, questions, and potential solutions," *Science*, vol. 320, no. 5878, pp. 889–892, 2008.

[3] EU Nitrogen Expert Panel, "Nitrogen use efficiency (nue) – guidance document for assessing nue at farm level," Wageningen University, Alterra, Wageningen, Netherlands, Tech. Rep., 2016.

[4] M. Anas, F. Liao, K. Verma, and et al., "Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency," *Biology Research*, vol. 53, p. 47, 2020.

[5] M. A. Sutton, C. M. Howard, J. W. Erisman, G. Billen, A. Bleeker, P. Grennfelt, and B. Grizzetti, Eds., *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives*. Cambridge University Press, 2011.

[6] E. Stehfest and L. Bouwman, "N2o and no emission from agricultural fields and soils under natural vegetation: summarizing available measurement data and modeling of global annual emissions," *Nutrient Cycling in Agroecosystems*, vol. 74, pp. 207–228, 2006.

[7] C. Bruno, A. Licciardello, G. Nastasi, F. Passaniti, C. Brigante, F. Sudano, A. Faulisi, and E. Alessi, "Embedded artificial intelligence approach for gas recognition in smart agriculture applications using low cost mox gas sensors," in *Smart Systems Integration (SSI)*, 2021.

[8] A. Bellabarba, G. Ujvári, M. Daghio, F. Rocchi, C. Becagli, R. Pastorelli, G. Buscioni, and C. Viti, "The impact of alfalfa intercropping and conventional tillage on n-cycling microbes: A tuscan vineyard case study," *Applied Soil Ecology*, vol. 213, p. 106240, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0929139325003786

[9] G. Muyzer, E. C. de Waal, and A. G. Uitterlinden, "Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16s rrna," *Applied and Environmental Microbiology*, vol. 59, no. 3, pp. 695–700, 1993.

[10] J. G. Caporaso, C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer, and R. Knight, "Global patterns of 16s rrna diversity at a depth of millions of sequences per sample," *Proceedings of the National Academy of Sciences*, vol. 108, no. Supplement 1, pp. 4516–4522, 2011.

[11] F. Poly, L. J. Monrozier, and R. Bally, "Improvement in the rflp procedure for studying the diversity of nifh genes in communities of nitrogen fixers in soil," *Research in Microbiology*, vol. 152, no. 1, pp. 95–103, 2001.

[12] J. H. Rotthauwe, K. P. Witzel, and W. Liesack, "The ammonia monooxygenase structural gene amoa as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations," *Applied and Environmental Microbiology*, vol. 63, no. 12, pp. 4704–4712, 1997.

[13] I. N. Throbäck, K. Enwall, Å. Jarvis, and S. Hallin, "Reassessing pcr primers targeting nirs, nirk and nosz genes for community surveys of denitrifying bacteria with dgge," *FEMS Microbiology Ecology*, vol. 49, no. 3, pp. 401–417, 2004.

[14] D. J. Scala and L. J. Kerkhof, "Nitrous oxide reductase (nosz) gene-specific pcr primers for detection of denitrifiers and three nosz genes from marine sediments," *FEMS Microbiology Letters*, vol. 162, no. 1, pp. 61–68, 1998.

[15] C. Carnevale, E. Turrini, R. Zeziola, E. De Angelis, and M. Volta, "A wavenet-based virtual sensor for pm10 monitoring," *Electronics*, vol. 10, no. 17, 2021.

[16] C. Carnevale, E. De Angelis, G. Finzi, E. Turrini, and M. Volta, "Optimal interpolation based data fusion techniques to improve deterministic air quality forecast," in *Springer Proceedings in Complexity*, 2021, pp. 145–150.