Machine Learning–Driven Insights into Nitrogen Dynamics and Greenhouse Gas Emissions in Commercial Potato Production Systems in Eastern Canada

Javed, B.[1,2], Cambouris, A. N.[1], Smith, E. L.[3], Dandrifosse, S.[4], Ziadi, N.[1], Karam, A.[2]

[1] Quebec Research and Development Centre, Agriculture and Agri-Food Canada, 2560 Hochelaga Blvd., Quebec City, QC, G1V 2J3, Canada.

[2] Soil and Agri-Food Engineering Department, Université Laval, 2425 rue de l'Agriculture, Québec, QC, G1V 0A6, Canada.

[3] Kentville Research and Development Centre, Agriculture and Agri-Food Canada, 361 Pictou Road, Bible Hill, NS, B2N 2T6, Canada.

[4] Walloon Agricultural Research Centre, 9 rue de Liroux, 5030 Gembloux, Belgium.

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract.

Nitrogen (N) management in potato systems is challenging because soil nitrate (NO_3-N) and nitrous oxide (N_2O) emissions vary substantially across space and time, driven by various factors including weather, soil, topographic, crop, and agricultural management systems. Treatment comparisons alone offer limited explanatory power under such complexity. A study was conducted in four commercial potato fields in Prince Edward Island, Canada, during 2020 and 2021. The objectives of this study were to identify the main drivers of soil N_2O and NO_3-N variability using Spearman correlation and random forest (RF) modeling, and to determine the most influential predictors. Fields were managed under uniform and variable-rate N strategies. Predictors spanned weather, probe-based soil measurements, Veris and topographic data, and Sentinel-2 spectral variables. Recent rainfall was represented using 1-d rainfall, defined as rainfall during the previous day, and weighted 7-d rainfall, defined as rainfall weighted across the previous 7 days. For N_2O, correlation analysis revealed significant positive associations with weighted 7d rainfall, soil moisture, soil NO_3-N, and 1-d rainfall, while shallow soil apparent electrical conductivity (ECa), relative humidity, and elevation were negatively associated. The RF model yielded RMSE = 0.027 kg ha^{-1} day^{-1} and R^2 = 0.56, with soil temperature, elevation, and soil moisture ranking as the three most influential predictors. For soil NO_3-N, the strongest correlations were observed with petiole NO_3-N (p = 0.58), chlorophyll vegetation index (CVI; p = 0.56), and Sentinel-2 SWIR1 (p = 0.48), followed by weighted 7d rainfall and growing degree days. The RF model achieved RMSE = 28.23 mg kg^{-1} and R^2 = 0.71, with SWIR1, weighted 7d rainfall, and petiole NO_3-N as the top three drivers. Collectively, these findings indicated that integrated environmental, spectral, and crop-status information explains N dynamics in commercial potato production far better than management grouping alone.

Keywords. *Precision agriculture, remote sensing, digital agronomy, features engineering, sustainable production, nitrogen management.*

Introduction

Nitrogen (N) management remains one of the main challenges in potato production because both nitrous oxide (N_2O) emissions and soil nitrate (NO_3-N) vary considerably across time and space. In commercial potato fields, this variability is influenced by weather conditions, soil properties, topography, crop status, and fertilizer management. As a result, treatment effects alone do not fully explain N dynamics under field conditions. Understanding how these factors contribute to soil n NO_3-N availability and greenhouse gas emissions is important for improving N use efficiency, while reducing environmental losses (Clément et al. 2021). Recent advances in precision agriculture have increased the availability of multi-source datasets that can be used to study this variability in greater detail. Weather records, probe-based soil measurements, Veris and topographic sensor-based data, and remote sensing indicators provide complementary information on the processes affecting N behavior in cropping systems. Approaches that combine these sources can improve the interpretation of soil NO_3-N and N_2O variability beyond what is possible from single-factor comparisons (Preza Fontes et al. 2019). Machine learning methods are useful in this context because they can evaluate multiple predictors simultaneously and help identify the variables most strongly associated with nitrogen dynamics under field conditions. This is particularly important in commercial potato systems, where environmental, soil, and crop-related factors interact across space and time (Gnisi et al. 2025). The objectives of this study were to identify the main drivers of soil N_2O and NO_3-N variability and determine the most influential predictors under commercial potato production conditions in eastern Canada.

Materials and Methods

The study was conducted in four commercial potato fields in Prince Edward Island, Canada, during the 2020 and 2021 growing seasons. Two sites were located at Oyster Cove (OC1, 8.6ha; OC2, 15.ha) and two at Black Pond (BP1, 8.8ha; BP2, 8.1ha). Soils were classified as Orthic Humo-Ferric Podzols of the Charlottetown series with sandy loam texture and pH ranging from 5.3 to 6.7. Experimental layouts included three management zones and site-specific N treatments under commercial production conditions. Soil NO_3-N and N_2O data were analyzed together with; environmental, soil, crop, and remote sensing predictors. Weather variables included rainfall and thermal indicators, while field measurements included soil moisture, soil temperature, topographic elevation, and soil apparent electrical conductivity. Crop-status information included petiole NO_3-N. Remote sensing predictors were derived from Sentinel-2 imagery and included spectral bands and vegetation indices relevant to potato N status, such as the chlorophyll vegetation index (CVI) and canopy chlorophyll content. Spearman correlation analysis was used to identify initial associations between each response variable and selected predictors. Random forest (RF) was tested to estimate N_2O and soil NO_3-N. The dataset was divided into training (75%) and testing (25%) subsets, and model tuning was conducted using five-fold cross-validation. Model performance was evaluated using root mean square error (RMSE) and coefficient of determination (R2). Variable importance from the best-performing RF models was used to identify the dominant predictors of N_2O and soil NO_3-N variability.

Results and Discussion

Spearman correlation analysis identified contrasting controls on N_2O and soil NO_3-N variability (Fig. 1a,b). For N_2O, fluxes were positively associated with recent rainfall metrics, including weighted 7-d rainfall and 1-d rainfall, as well as soil moisture and soil NO_3-N. In contrast, shallow soil apparent electrical conductivity, relative humidity, and elevation were negatively associated with N_2O fluxes, consistent with the influence of rainfall and soil water status on N_2O emissions (Wang et al., 2024). For soil NO_3-N, the strongest positive correlations were observed with petiole NO_3-N, CVI and Sentinel-2 SWIR1, followed by weighted 7d rainfall, growing degree days, and soil temperature. These results indicated that soil NO_3-N variability was more closely linked to crop status, recent weather conditions, and spectral information than to management grouping alone (Preza Fontes et al. 2019).

The RF model provided the best predictive performance for N_2O, with RMSE = 0.027 kg ha^{-1} day^{-1} and R^2 = 0.56, whereas the soil NO_3-N model achieved RMSE = 28.23 mg kg^{-1} and R^2 = 0.71 (Fig. 1c,d). Final RF importance showed that soil temperature, elevation, and soil moisture were the dominant predictors of N_2O, while soil NO_3-N estimation was driven primarily by Sentinel-2 SWIR1, weighted 7-d rainfall, and petiole NO_3-N (Fig. 1e, f). Together, these results showed that short-term environmental conditions were most important for N_2O, whereas spectral, rainfall, and crop-status information were more important for soil NO_3-N (Preza Fontes et al. 2019; Wang et al. 2024).

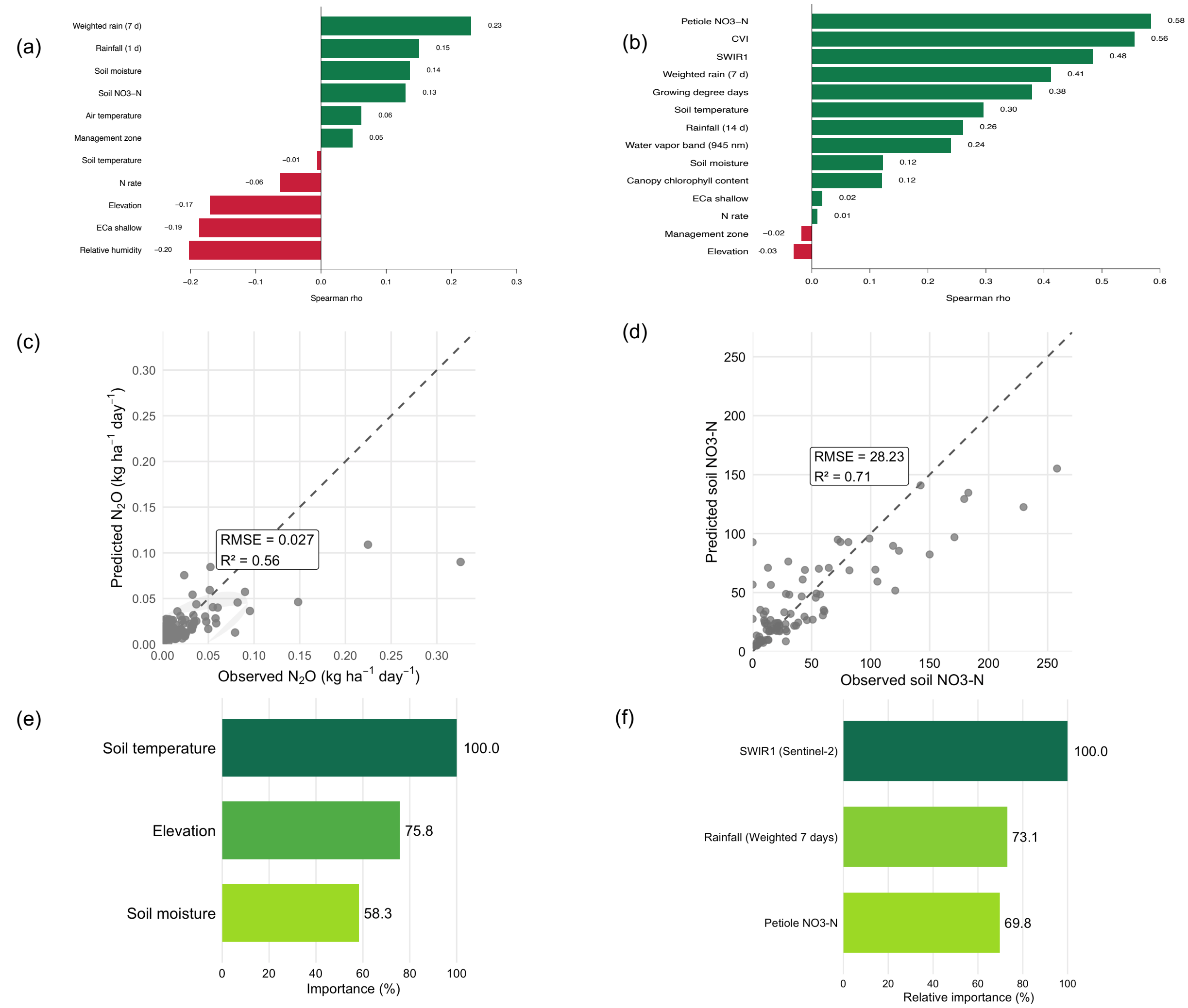

Figure 1: Summary of correlation and random forest results for N_2O and soil NO_3-N. Panels (a) and (b) show the Spearman correlation coefficients between selected predictors and N_2O and soil NO_3-N, respectively. Panels (c) and (d) show observed versus predicted values from the random forest models for N_2O and soil NO_3-N, respectively. Panels (e) and (f) show the three most influential predictors identified by the random forest models for N_2O and soil NO_3-N, respectively.

Conclusion

This study demonstrated that N_2O was driven mainly by short-term environmental factors, whereas soil NO_3-N was most strongly linked to spectral, rainfall, and crop-status information. Overall, these findings support the use of integrated environmental and remote sensing data for more site-specific N management in commercial potato production systems.

Acknowledgments

This research was supported by Agriculture and Agri-Food Canada (Project #J-002269).

References

Clément, C.-C., Cambouris, A. N., Ziadi, N., Zebarth, B. J., & Karam, A. (2021). Potato Yield Response and Seasonal Nitrate Leaching as Influenced by Nitrogen Management. Agronomy, 11(10), 2055. https://doi.org/10.3390/agronomy11102055.

Preza Fontes G, Bhattarai R, Christianson LE & Pittelkow CM (2019) Combining Environmental Monitoring and Remote Sensing Technologies to Evaluate Cropping System Nitrogen Dynamics at the Field-Scale. Front. Sustain. Food Syst. 3:8. doi: 10.3389/fsufs.2019.00008.

Gregor Gnisi, G., Weik, J., Ruser, R., Essich, L., Lewandowski, I., & Stein, A. (2025). *Machine learning-based prediction of nitrous oxide emissions from arable farming: Exploring management practices as predictor variables*. Ecological Indicators, 172, 113233. https://doi.org/10.1016/j.ecolind.2025.113233.

Wang, J., Liu, Q., Zhao, S., He, S., Yang, B., Wang, N., Hou, P., Xue, L., & Yang, L. (2024). Heavy rainfall stimulates more N2O emissions from wheat fields during basal and overwintering fertilization phases. *Agriculture, Ecosystems & Environment, 376*, 109227. https://doi.org/10.1016/j.agee.2024.109227.