DOI 10.5162/EUROSENSORS2025/T4.2.3

Correlation of gaseous emissions with phenological phases in tomato crops

Melissa Tamisari[1], Emanuela Tavaglione[2], Francesco Tralli[2], Matteo Valt[3], Barbara Fabbri[2]
[1] Department of Neuroscience and Rehabilitation, University of Ferrara, Via Luigi Borsari 46, 44121 Ferrara, Italy
[2] Department of Physics and Earth Science, University of Ferrara, Via Giuseppe Saragat 1/C, 44122 Ferrara, Italy
[3] Sensors and Devices Center, Bruno Kessler Foundation, Via Santa Croce 77, 38123 Trento, Italy

Corresponding Author's e-mail address: melissa.tamisari@unife.it

Summary:
Currently, plant phenological phases are determined through models based on historical data series and advanced measurement tools such as satellite imaging, which are not always accurate. This study explores innovatively the relationship between Volatile Organic Compounds emitted by tomato plants and their phenological phases using machine learning algorithms. Supervised models, particularly k-NN, achieved an accuracy of 96.8% in identifying phases. These results highlight the potential of sensors and AI to accurately monitor crop development.

Keywords: Machine Learning, Phenological Phases, Precision agriculture, Volatile organic compounds, MOX gas sensors.

Introduction

Precision agriculture represents a strategic approach aimed at improving agricultural productivity while simultaneously reducing resource use and maintaining high-quality standards. At the same time, it serves as a fundamental tool for climate adaptation, addressing the growing environmental challenges faced by modern agricultural systems. Indeed, rising temperatures and prolonged periods of drought are causing serious issues, including shifts in crop phenological cycles, leading to a decline in both the quantity and quality of yields. The aim of this study is to identify a relationship between Volatile Organic Compounds (VOCs) secreted by tomato crop, and empirically derived agronomic parameters — in particular, the phenological phase (PPhases)— via supervised machine learning algorithms [1]. PPhases are currently complemented by more sophisticated estimates obtained through models based on historical data series and measurement tools such which are not always accurate. The Growing Degree Days (GDD) index, which measures the accumulation of heat useful for plant growth during the crop cycle, is used to mark the transition from one PPhase to another. Since each phenological stage requires a specific amount of thermal energy to be completed, knowing a plant's thermal requirements allows for reasonably accurate estimates of the duration of its developmental stages. PPhases represent the main stages of the crop's growth cycle, marking key moments in plant development. The PPhases are: (1) pre-emergence, (2) sowing/transplanting (3) fruit set – first truss (4) veraison – first truss (5) veraison – second truss, (6) ripening: 50% of berries fully colored and (7) ripening: 100% of berries fully colored.

Method

From 2020 to 2022, we carried out industrial tomato monitoring campaigns lasting three to four months (typically from early June to early September). As part of the project, a multifunctional station was installed in tomato crop field, equipped with: a relative humidity (RH%) and temperature sensor (Sensirion SHT11); a small camera with a removable infrared filter, used to monitor the plants' radiation absorption; four metal oxide (MOX)-based gas sensors to monitor plant gas emissions during their phenological development [3]. In particular, the sensors used were SnO_2/Au, SnO_2/Pt, SnO_2/Pd, WO_3; monitoring of VOCs emitted by plants is primarily aimed at assessing their health status, which can be affected by insects or diseases [4]. Classification may be defined as a supervised learning method in which machine learning (ML) algorithms are trained to assign labels to input

DOI 10.5162/EUROSENSORS2025/T4.2.3

data based on patterns learned from a labeled training dataset [1]. The main goal of these algorithms is generalizing from the training examples in order to enable the model to correctly predict the label of unseen data. In light of this, applying ML techniques to a wide range of sensor signals, makes it possible to train a model able to identify the phenological phase. In this work, a model capable of doing this type of classification was trained due to the vast amount of data acquired during the years.

Results

A signal was recorded approximately every 14 seconds during the data acquisition period. The datasets we obtained generally cover phenophases from 2 to 6, which have been assigned to classification labels ranging from 1 to 5.

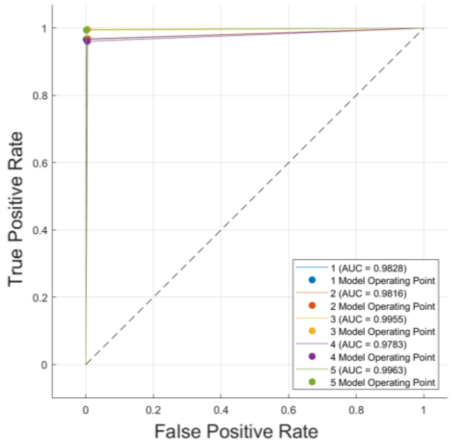

Fig. 1. ROC curves and AUC values for the k-NN algorithm (validation).

As an initial investigation, the 2020 dataset presented here was analyzed using several supervised machine learning methods, using a dataset split of 90% training and 10% testing, with 5-fold cross-validation. Even with relatively simple algorithms, the results proved to be promising. Among these, the k-nearest neighbors (k-NN) algorithm emerged as the best-performing method, achieving an accuracy of 98.8%, along with a high Receiver Operating Characteristic (ROC) curve score (Fig. 1), implying significant discriminatory power. The confusion matrix for this model is shown in Fig. 2. These promising results suggest that training the model on multi-year datasets at disposal could lead to a robust method for determining phenophases based on the signals acquired from an array of four chemoresistive gas sensors. A model of this kind could be integrated into decision support systems in agriculture, helping to improve the identification of phenophases and, consequently, to optimize agronomic strategies, making farming practices more efficient and sustainable.

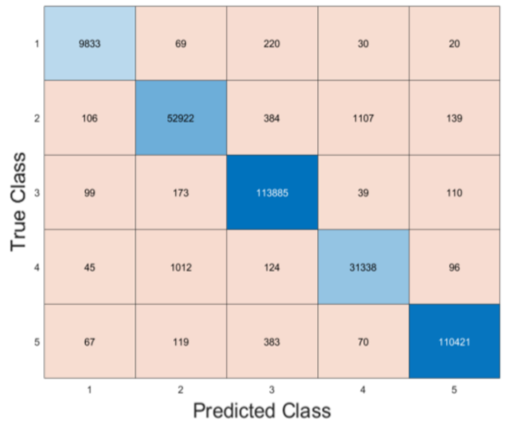

Fig. 2. Confusion matrix for the k-NN algorithm (validation).

References

[1] S. L. Stevan Jr., A. F. C. Garcia, B. A. Menegotto, J. C. Ferreira Da Rocha, H. V. Siqueira, R. A. Ayub, "Phenological stages analysis in peach trees using electronic nose" *Open Agriculture*,9, (2024), 20220337. doi: 10.1515/opag-2022-0337

[2] L. S. Pereira, R. G. Allen, M. Smith, D. Raes, Crop evapotranspiration estimation with FAO56: Past and future, Agricultural Water Management, 147, 4 (2015) doi: 10.1016/j.agwat.2014.07.031

[3] B. Fabbri, M. Valt, C. Parretta, S. Gherardi, A. Gaiardo, C. Malagù, F. Mantovani, V. Strati, V. Guidi, Correlation of gaseous emissions to water stress in tomato and maize crops: From field to laboratory and back, Sensors and Actuators B: Chemical, 303, 127227 (2020) doi: 10.1016/j.snb.2019.127227

[4] K. Takayama, R.M.C. Jansen, E.J. van Henten, F.W.A. Verstappen, H. J. Bouwmeester, H. Nishina, Emission index for evaluation of volatile organic compounds emitted from tomato plants in greenhouses, Biosystems Engineering, 113, 220 (2012) doi:10.1016/j.biosystemseng.2012.08.004

Acknowledgements

This work is done in collaboration with CER Consortium (Consorzio di Bonifica Canale Emiliano Romagnolo) and the Emilia-Romagna Region, as part of the POSITIVE project (Scalable Operational Protocols for Precision Agriculture).