Precision Agriculture
https://doi.org/10.1007/s11119-024-10153-w

# Fertilization and soil management machine learning based sustainable agronomic prescriptions for durum wheat in Italy

Marco Fiorentini[1] · Calogero Schillaci[2] · Michele Denora[3] · Stefano Zenobi[1] · Paola A. Deligios[1] · Rodolfo Santilocchi[1] · Michele Perniola[3] · Luigi Ledda[1] · Roberto Orsini[1]

Accepted: 16 May 2024


## Abstract

**Purpose** This research aims to develop a meta-machine learning model to optimize soil and nitrogen management for durum wheat in Italy. It addresses the challenges of increased food production on limited land amidst rising input costs, geopolitical changes, and climate change. The goal is to aid decision-makers in achieving maximum crop yield and income margins through effective agronomic strategies.

**Methods** The study developed a meta-machine learning model, integrating classification and regression models, and tested it at four sites in Marche and Basilicata, Italy, over several years. The model incorporated data from remote sensing, crop phenology, soil chemical properties, weather data, soil management, and nitrogen levels. A Random Forest model was used to classify crop phenology, while a Neural Network model predicted yield. Eleven nitrogen levels were compared across these sites.

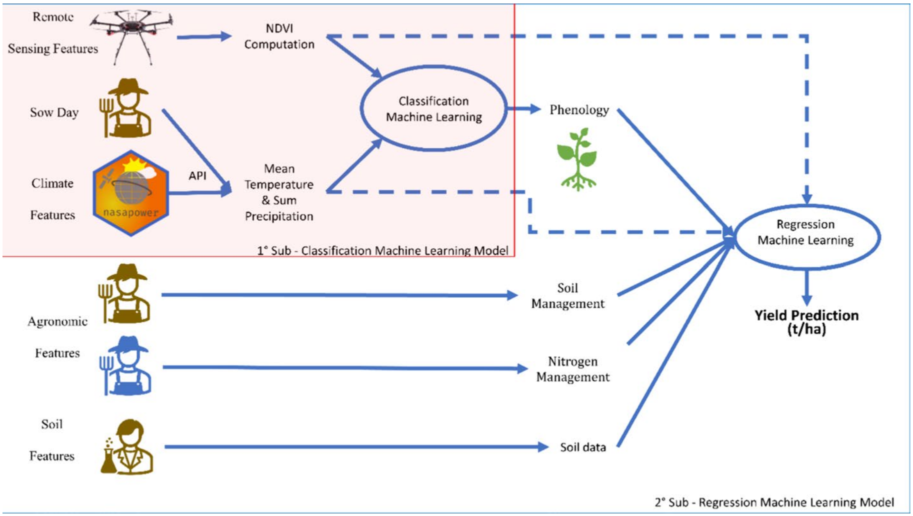
**Results** The Random Forest model achieved an accuracy of 0.98, kappa of 0.96, and recall of 0.98 for predicting crop phenology. The Neural Network model for yield prediction had an R squared of 0.90 and a Root Mean Square Error of 0.59 t ha-1. Key factors identified for model accuracy were temperature, precipitation, NDVI, and nitrogen input. Simulations of 30 soil management and fertilization combinations revealed that no-tillage management increased grain yield. The Marginal Fertilizer Yield Index determined optimal nitrogen application.

**Conclusions** The meta-machine learning model accurately predicted durum wheat yield and identified effective agronomic strategies, demonstrating the potential for broader application in field conditions. The model offers a promising approach to sustainable agriculture and climate change mitigation by utilising publicly available spatial datasets.

Extended author information available on the last page of the article

Published online: 05 July 2024

## Graphical abstract




**Keywords** Machine learning · Decision support system · Remote sensing · Nitrogen fertilization · Soil management · Durum wheat

## Introduction

Global population growth, soil sealing, rising raw material costs and increase of climate extremes are putting a strain on the world's agricultural sector. The world population is projected to reach 8.5 billion by 2030, with further increases to 9.7 billion by 2050 and 10.4 billion by 2100. However, a recent paper suggests that the world population may reach its peak soon (Le Page, 2023). The rapid population growth is expected to lead to heightened urbanization, bringing a surge in demand for food, raising concerns about food safety. The agricultural sector confronted with the crucial task of increasing food production while having access to limited agricultural land. Additionally, rising production input costs, influenced by geopolitical and climate changes, add to the list of challenges. These factors collectively put stress on the agricultural sector, promoting the agricultural scientific community to actively seek solutions to address these pressing issues. The scientific community is focusing on the development of innovative cropping systems (De Menna et al., 2016), on the spread of conservative agriculture (Capmourteres et al., 2018), on the application of the principles of organic farming (Deligios et al., 2021) and precision agriculture (Argento et al., 2021; Orsini et al., 2023). In the Mediterranean area, winter cereals hold a crucial position among the arable crops (Tsialtas et al., 2018). Achieving maximum yield in durum wheat is highly dependent on effective N fertilization, particularly when water availability is not a limiting factor (Orsini et al., 2019a).

Weather conditions, soil physical–chemical properties and soil management practices are also agronomic factors that play a pivotal role in enhancing durum wheat yields (Scott et al., 2024). Given the multitude of factors that must be considered when making agronomic decisions, the development of decision support systems (DSS) is grown in significance. These DSS have the capability to run simulations that assess various combinations of agronomic management practices and their relative impacts on yields, it also help (Terribile et al., 2024) in identifying and establishing the most effective agronomic management strategies (Hoogenboom et al., 2004; Kamilaris & Prenafeta-Boldú, 2018). Over the past 40 years, both process based and stochastic models have been developed to support agronomic decision-making. Some of the well-known models include DSSAT (Jones et al., 2003), APSIM (Holzworth et al., 2014), Century (Lugato et al., 2014), SALUS (Basso et al., 2006), ARMOSA (Valkama et al., 2020), and more recently, there have been successful attempts to combine process based models with machine learning (ML). This integration can be used for predicting yield and evapotranspiration (Attia et al., 2022; Zhang et al., 2023). Process based models are the most used models in agriculture. They involve simulating each component of the system, including plants, soil, climate, and the exchanges between these components. However, process based models rely on a predefined set of covariates that must be measured to use the model effectively. The selection of covariates is determined by the creators of the model, and some covariates can be challenging to measure at each field. Also, process based models require site-specific calibration, making them less suitable for use at larger scales (Pasquel et al., 2022). Stochastic methods, particularly ML approaches (Chlingaryan et al., 2018; Schillaci et al., 2022), have gained recognition in agriculture due to their ability to establish relationships between biotic and environmental predictors (Adamchuk et al., 2004; Kayad et al., 2019). These approaches, often built on top of remote sensing data, offer the advantage of scalability and can be applied to large agricultural areas. ML models, when appropriately designed and trained with spatially explicit covariates from each field, can provide economic and environmental benefits (Fiorentini et al., 2023). Several studies reported the potential of ML models in agriculture, spanning various applications such as yield prediction (van Klompenburg et al., 2020), nitrogen provision recommendation (Shi et al., 2020), weed (Talaviya et al., 2020) and pest detection (Wang et al., 2021), and soil digital mapping (Schillaci et al., 2021). While many of these studies have emphasized the accuracy and validation of the models, it is crucial that these models are designed to support agronomic decision-making. The primary focus should be on utilizing the high accuracy of ML models to provide actionable insights that assist agronomists and farmers in making informed decisions about agricultural management practices. In essence, the goal is to translate the predictive power of these models into practical and valuable guidance for optimizing fertilization, resource management, and sustainability in agriculture.

The aim of this study was to develop a ML-based system that allow simulating different combinations of soil and nitrogen management for durum wheat in Italy. This system was intended to optimize decision-making in durum wheat farming by providing insights into the best practices. To achieve this, a multi-data source approach was employed, integrating data on nitrogen fertilization levels (0, 40, 90, 95, 118, 120, 133, 139, 150, 180 and 222 kg N ha$^{-1}$), soil management practices (conventional and conservation agriculture), climate conditions (average monthly temperature and precipitation), soil properties (soil organic carbon, available nitrogen, C:N ratio), and remote sensing data ( average monthly NDVI) to train a meta-ML system.

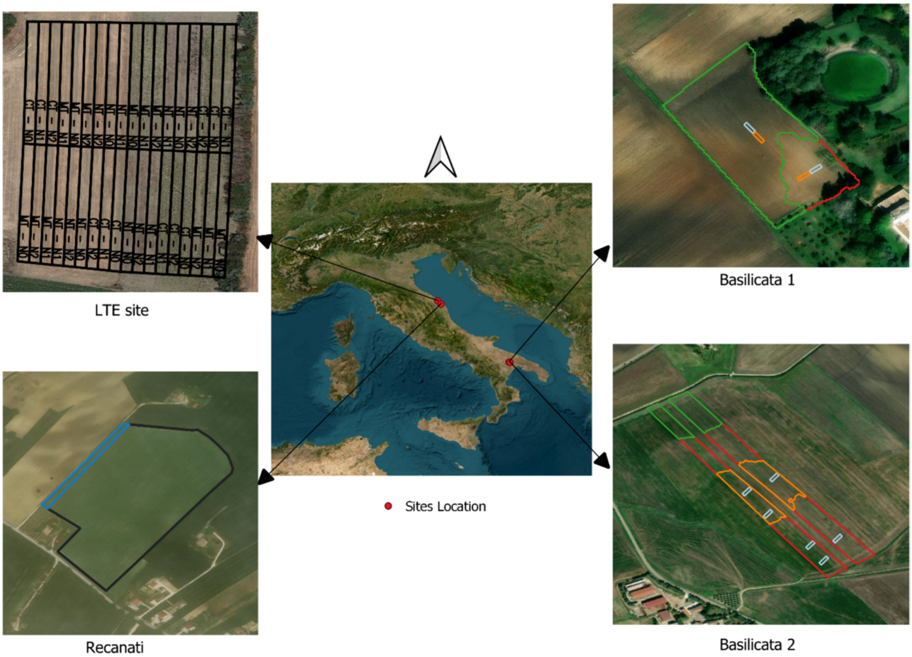



**Fig. 1** Location of experimental sites and experimental designs

**Table 1** Description of experimental sites, year, and nitrogen levels provided

| Site | Region | Name | Year of experiment | Soil management | Nitrogen levels (kg ha$^{-1}$) |
|---|---|---|---|---|---|
| 1 | Marche | LTE | 2018, 2019, 2020, 2021, 2022 | CT, NT, MT | 0, 90, 180 |
| 2 | Marche | Recanati | 2020 | CT | 90, 222 |
| 3 | Basilicata | Basilicata 1 | 2021 | CT | 0, 95, 120, 150 |
| 4 | Basilicata | Basilicata 2 | 2020 | NT | 40, 118, 133, 139, 150 |

*LTE* Long Term Experiment, *CT* Conventional tillage, *NT* No Tillage, *MT* Minimum tillage

## Materials and methods

### Experimental sites' overview

The framework was developed and tested at four sites distributed between the Marche and Basilicata Italian regions (Fig. 1). All sites had different experimental designs, and durum wheat (*Triticum turgidum* subsp. *durum* Desf.) was grown for several years (Table 1).

Table 2 displays all agronomic management activities performed at the four experimental sites throughout the growing seasons, while weather data are reported in Table 3.

**Table 2** Agronomic site management and dates of agronomic practices

| | Years | | | | |
|---|---|---|---|---|---|
| | 2018 | 2019 | 2020 | 2021 | 2022 |
| LTE Site | | | | | |
| Agro-technique | | | | | |
| Ploughing | 10/02/2017 | 09/26/2018 | 11/11/2019 | 11/08/2020 | 11/12/2021 |
| Harrowing and seedbed preparation | 11/20/2017 | 11/01/2018 | 11/25/2019 | 11/27/2020 | 11/21/2021 |
| Sowing | 11/21/2017 | 11/30/2018 | 12/27/2019 | 12/22/2020 | 12/24/2021 |
| Weed control: Pinoxaden | 03/28/2018 | 03/08/2019 | 02/10/2020 | 02/01/2021 | 02/15/2022 |
| N fertilization | 03/29/2018 | 03/18/2019 | 02/25/2020 | 02/23/2021 | 02/26/2022 |
| | 04/30/2018 | 05/02/2019 | 04/09/2020 | 04/16/2021 | 04/08/2022 |
| Pest control: Azoxystrobin, Cyproconazole | 04/24/2018 | 04/22/2019 | 04/20/2020 | 04/22/2021 | 04/17/2022 |
| Harvest | 07/06/2018 | 07/07/2019 | 07/06/2020 | 07/04/2021 | 07/09/2022 |
| Recanati | | | | | |
| Agro-technique | | | 2020 | | |
| Ploughing | | | 10/02/2019 | | |
| Harrowing and seedbed preparation | | | 11/11/2019 | | |
| Sowing | | | 11/15/2019 | | |
| Weed control: Pinoxaden | | | 04/22/2020 | | |
| N fertilization | | | 01/20/2020 | | |
| | | | 03/20/2020 | | |
| Pest control: Azoxystrobin, Cyproconazole | | | 04/22/2020 | | |
| Harvest | | | 06/28/2020 | | |
| Basilicata 1 | | | | | |
| Agro-technique | | | | 2021 | |
| Ploughing | | | | 28/09/2020 | |
| Harrowing and seedbed preparation | | | | 14/12/2020 | |
| Sowing | | | | 18/01/2021 | |

**Table 2** (continued)

| | Years | | | | |
|---|---|---|---|---|---|
| | 2018 | 2019 | 2020 | 2021 | 2022 |
| Weed control: Pinoxaden and Metsulfuron-methyl | | | | 23/03/2021 | |
| N fertilization | | | | 14/12/2020 | |
| | | | | 20/04/2021 | |
| Pest control: Pyraclostrobim | | | | 23/03/2021 | |
| Harvest | | | | 19/07/2021 | |
| Basilicata 2 | | | | | |
| Agro-technique | | | 2020 | | |
| Harrowing and seedbed preparation | | | 10/12/2020 | | |
| Sowing | | | 07/01/2020 | | |
| Weed control: Pinoxaden and Metsulfuron-methyl | | | 15/03/2020 | | |
| N fertilization | | | 16/12/2019 | | |
| | | | 14/04/2020 | | |
| Pest control: Pyraclostrobin | | | 15/03/2020 | | |
| Harvest | | | 20/07/2020 | | |

*LTE* Long Term Experiment

**Table 3** Mean temperature (°C) and cumulative rainfall (mm) during the agronomic year compared to the long-term data

| | Months | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | November | December | January | February | March | April | May | June | July |
| LTE site | | | | | | | | | |
| Cumulative rainfall (mm) | | | | | | | | | |
| 2017–2018 | 24 | 96 | 29 | 173 | 143 | 37 | 95 | 48 | 2 |
| 2018–2019 | 0 | 61 | 70 | 22 | 36 | 59 | 165 | 1 | 0 |
| 2019–2020 | 0 | 0 | 4 | 17 | 158 | 119 | 142 | 61 | 5 |
| 2020–2021 | 44 | 102 | 79 | 20 | 25 | 33 | 23 | 8 | 40 |
| 2021–2022 | 191 | 138 | 29 | 61 | 25 | 61 | 14 | 35 | 44 |
| Long-term data 2000–2020 | 78 | 74 | 55 | 60 | 69 | 54 | 55 | 53 | 37 |
| Mean temperature (°C) | | | | | | | | | |
| 2017–2018 | 10 | 7 | 9 | 5 | 9 | 16 | 19 | 23 | 26 |
| 2018–2019 | 5 | 7 | 5 | 8 | 12 | 13 | 15 | 25 | 28 |
| 2019–2020 | 13 | 6 | 7 | 11 | 10 | 14 | 19 | 22 | 25 |
| 2020–2021 | 11 | 8 | 6 | 9 | 10 | 12 | 18 | 25 | 27 |
| 2021–2022 | 11 | 7 | 6 | 9 | 9 | 13 | 20 | 26 | 27 |
| Long-term data 2000–2020 | 12 | 6 | 7 | 6 | 10 | 12 | 19 | 23 | 26 |
| Recanati | | | | | | | | | |
| Cumulative rainfall (mm) | | | | | | | | | |
| 2019–2020 | 13 | 64 | 30 | 26 | 128 | 44 | 44 | 50 | |
| Long-term data 2000–2020 | 96 | 84 | 50 | 48 | 51 | 65 | 51 | 47 | |
| Mean temperature (°C) | | | | | | | | | |
| 2019–2020 | 12.6 | 9 | 6.5 | 10.1 | 10.1 | 13.5 | 18.5 | 21.8 | |
| Long-term data 2000–2020 | 12.3 | 9.2 | 7.9 | 7.7 | 10 | 12.8 | 17.4 | 21.7 | |
| Basilicata 1 | | | | | | | | | |
| Cumulative rainfall (mm) | | | | | | | | | |
| 2020–2021 | | | 44 | 24 | 31 | 31 | 1 | 1 | 1 |
| Long-term data 2000–2020 | | | 64 | 54 | 61 | 50 | 38 | 33 | 33 |
| Mean temperature (°C) | | | | | | | | | |
| 2020–2021 | | | 5.8 | 7.7 | 7.3 | 9.9 | 16 | 23.1 | 26 |
| Long-term data 2000–2020 | | | 8.6 | 8.8 | 10.9 | 13.7 | 18 | 23.4 | 26 |
| Basilicata 2 | | | | | | | | | |
| Cumulative rainfall (mm) | | | | | | | | | |
| 2019–2020 | | | 79 | 69 | 44 | 31 | 140 | 46 | |
| Long-term data 2000–2020 | | | 64 | 54 | 61 | 50 | 38 | 33 | |
| Mean temperature (°C) | | | | | | | | | |
| 2019–2020 | | | 8 | 8.5 | 11.5 | 16.1 | 20 | 23.3 | |
| Long-term data 2000–2020 | | | 8.6 | 8.8 | 10.9 | 13.7 | 18 | 23.4 | |

*LTE* Long Term Experiment

### Long term experiment (LTE)

The long-term experimental (LTE) site was located at the “Pasquale Rosati” experimental farm of the Polytechnic University of Marche in Agugliano, Italy (43°32′N, 13°22′E, 100 a.s.l.). This site registered a mean annual precipitation of approximately 749 mm and a mean annual temperature of 15.5 °C. Monthly temperatures averages range from 6 °C in February to 26 °C in July. During the 2017–2018 crop growing season, the highest cumulative precipitation (647 mm) was recorded. However, in the subsequent growing seasons of 2019–2020 and 2020–2021, there was a decrease in precipitation by 36% (414 mm) and 42% (374 mm), respectively. Additionally, the 2018–2019 growing season had a mean annual temperature that was 1 °C lower than the other growing seasons (Table 3). The soil of the study area has a silt–clay texture and it is classified as Calcaric Gleyic Cambisols. It contains 9.1 g $kg^{-1}$ of organic carbon, 13.8 mg $kg^{-1}$ of phosphorus (Olsen-P), 323.2 mg $kg^{-1}$ of potassium, 26.7 cmol $kg^{-1}$ of cation exchange capacity, pH of 8.1, 29.6% of field water capacity, 17.6% of wilting point, and a slope of approximately 10%. The LTE, established in 1994, consists of a rainfed 2-year rotation of durum wheat cv. Tyrex and maize (*Zea mays* L.) cv. DK440 (hybrid, FAO Class 300, Dekalb). The crop rotation involved two adjacent fields to ensure the presence of all crops each year (Fig. 1). Within each field, three soil management (main plot, 1.500 $m^2$ each) systems (conventional, minimum tillage, and no-tillage) and three N fertilizer (sub-plot, 500 $m^2$ each) levels were arranged according to a split-plot experimental design with two randomized blocks and were repeated in the same plots every year. Conventional tillage (CT), which is representative of the business-as-usual tillage practice in the study area, was plowed along the maximum slope every year by a mouldboard (with two plows) to a depth of 0.40 m. In minimum tillage (MT) plots, a chisel was used at a depth of 25 cm in autumn for wheat and in spring for maize. The seedbed was prepared by double harrowing prior to sowing. Durum wheat was sown at the beginning of autumn at a seeding rate of 220 kg $ha^{-1}$, with a row distance of 0.17 m. Each year, N fertilization in the form of urea (46%) was split into two applications: 50% at the end of tillering and 50% before head emergence. Weeds, pests, and diseases were controlled chemically. The no-tillage (NT) soil was left undisturbed, except for specific practices such as sod seeding, managing crop residuals, weed chopping, and total herbicide spraying before seeding. The three N fertilizer treatments were N0, N90, and N180, corresponding to 0, 90, and 180 kg N $ha^{-1}$, respectively, distributed in two rates for wheat. The unfertilized N0 treatment was used as the control. The N90 treatment was compliant with the agri-environmental measures adopted within rural development plans at the local scale (https://www.regione.marche.it/Regione-Utile/Agricoltura-Sviluppo-Rurale-e-Pesca/Produzione-Integrata#Tecniche-Agronomiche). The N180 treatment represented the conventional management approach. The sequences of the agronomic practices applied at LTE are reported in Table 2.

### Recanati Site

The Recanati experiment was conducted during the 2019–2020 growing season at the “Guzzini” private farm in Recanati (43°23′N, 13°31′E, 115 a.s.l.), Central Italy. This site experiences a mean annual precipitation of 719 mm, and a mean annual temperature of 16.8 °C. Monthly means temperatures vary from 7.7 °C in February to 24.7 °C in August. During the 2019–2020 crop growing season, the farm site was characterized by a seasonal rainfall of approximately 400 mm, which was around 100 mm lower than the precipitation recorded at the Agugliano weather station during the same growing season. The mean temperatures at the farm site were 1.5 °C lower than those recorded at Agugliano (Table 3). The soil of the farm site has a silt–clay texture and is classified as Calcaric Gleyic Cambisols. It contains 8.4 g kg$^{-1}$ of organic carbon, 10.3 mg kg$^{-1}$ of phosphorus, 283.2 mg kg$^{-1}$ of potassium, 28.0 meq 100 g$^{-1}$ of cation exchange capacity, pH of 8.1, 28.5% of field water capacity, and 16.2% of wilting point and a slope of about 6%. The agricultural practices at the Recanati site were similar to those at the LTE site. This included a rainfed 2-year crop rotation with durum wheat cv. Tyrex and maize cv. DK440 (Hybrid, FAO class 300). The soil was managed using conventional tillage (CT), and two different N levels were applied in two adjacent areas (Fig. 1). Seedbed preparation was performed as reported at the LTE site, where plowing was performed along the maximum slope by a mouldboard plow (with two plows) at a depth of 0.40 m in autumn. Before seeding, harrowing was performed, and the wheat was sown at a rate of 240 kg ha$^{-1}$ with a row spacing of 0.13 m. The two N levels corresponded to 90 and 180 kg N ha$^{-1}$, and they were distributed at two rates. Each rate consisted of half of the total N and it was provided to the crop in the form of urea (46%). The first rate was provided at the end of tillering and the second rate before head emergence. Weeds, pests, and diseases were controlled chemically. The sequences of the agronomic practices applied at the farm site are shown in Table 2.

### Basilicata 1 Site

The Basilicata 1 experiment was conducted during the 2020–2021 growing season at the "Az. Agricola Pugliese" private farm located near Laterza (40°42′N, 16°42′E, 362 a.s.l.), Southern Italy. The mean annual precipitation is 669 mm, and the mean annual temperature is 18 °C. Monthly means temperatures range from 8.6 °C in January to 26.4 °C in July. During the 2020–2021 crop growing season, the farm site was characterized by a seasonal rainfall of 133 mm and a mean temperature of 13.7 °C (Table 3). The soil of the Basilicata site 1 has a silt–clay texture and is classified as Haplic Calcisol. It contains 9.1 g kg$^{-1}$ of organic carbon, 20.5 mg kg$^{-1}$ of phosphorus (Olsen-P), 75.2 mg kg$^{-1}$ of potassium, 17.9 cmol kg$^{-1}$ of cation exchange capacity, pH of 8.3, 33.4% of field water capacity, 20.6% of wilting point, and a slope of approximately 2%. The experiment was conducted using durum wheat (var. Kanakis, RAGT Semences, France). Seedbed preparation process was consistent with the LTE and Recanati sites. Plowing was performed along the maximum slope using a mouldboard plow (with two plows) at a depth of 0.40 m in autumn. Prior to seeding in January, a harrowing operation was performed in December. The seeding was done with a seed rate of 240 kg ha$^{-1}$ and a row spacing of 0.13 m. Weed control was performed in March using gliphosate and metsulfuron-methyl, while pyraclostrobin was used for pest control. The experiment was set up as a precision agriculture experiment. Variable-rate nitrogen fertilization was applied based on a prescription map, as depicted in Fig. 1. The map identified two homogenous areas based on the spatial variability of soil physical and chemical characteristics (Denora et al., 2022b). The amount of N applied (urea 46%) was calculated based on the estimated nitrogen uptake by the crop and the soil characteristics. The N levels provided to the crop were 120 kg N ha$^{-1}$ and 95 kg N ha$^{-1}$, split into two rates throughout the crop cycle. Additionally, within the two identified areas, specific plots measuring 25 m$^2$ received N levels of 150 kg N ha$^{-1}$and 0 kg N ha$^{-1}$ (Fig. 1). The sequences of the agronomic practices used at the farm site are shown in Table 2.

### Basilicata 2 Site

The Basilicata 2 experiment was carried out during the 2019–2020 growing season in the "Az. Agricola F. lli Lillo" private farm located in Matera (40°42′N, 16°39′E, 401 a.s.l.), Southern Italy. The mean annual precipitation is 670 mm, and the mean annual temperature is 18 °C, with monthly means ranging from 8.6 °C in January to 23.4 °C in June. During the 2019–2020 crop growing season, the farm site was characterized by a seasonal rainfall of 409 mm and a mean temperature of 14.5 °C (Table 3). The soil of the Basilicata site 2 has a silt–clay texture (Haplic Vertisol) with 11.6 g $kg^{-1}$ of organic carbon, 11.3 mg $kg^{-1}$ of phosphorus (Olsen-P), 185.2 mg $kg^{-1}$ of potassium, 21.2 cmol $kg^{-1}$ of cation exchange capacity, pH of 8.2, 32.2% of field water capacity, 18.9% of wilting point, and a slope of approximately 3%. The experiment was conducted using durum wheat (var. PR22D89, Pioneer Hi-Bred, Italia) and a no-tillage soil management approach similar to the one used at the LTE site. Before seeding, a harrowing operation was performed in December, while sowing was done in January with a seed rate of 240 kg $ha^{-1}$ and a row spacing of 0.13 m. Weed control was performed in March using gliphosate and metsulfuron-methyl. Pest control was managed using pyraclostrobin. The experiment was configured as a precision agriculture trial. Variable-rate nitrogen fertilization was performed with a prescription map (Fig. 1), where two distinct homogenous areas were identified based on the spatial variability of soil properties (Denora et al., 2022a). The amount of N applied (ammonium nitrate, 26%) was calculated based on the estimates of the crop's nitrogen uptake and the specific soil characteristics within each area. Different N levels were provided to the crop, including 150 kg N $ha^{-1}$, 139 kg N $ha^{-1}$, 133 kg N $ha^{-1}$, and 118 kg N $ha^{-1}$. Within the two designated areas, we identified plots of 4 $m^2$ that received a distinct N level of 40 kg N $ha^{-1}$ (Fig. 1). The sequences of the agronomic practices used at the farm site are shown in Table 2.

### Covariates and measurements

The selected covariates should comprehensively represent all components of the agricultural system, including crop-related factors, soil characteristics, climate parameters, and agronomic management practices (Table 4).

In this study, multispectral images were captured during different growth stages of durum wheat using two different UAV platforms equipped with multispectral cameras.

**Table 4** List of the covariates group, type, and measurement

| Covariates group | Covariate type | Covariate measurement |
|---|---|---|
| Crop | Remote sensing | NDVI |
| | Crop phenology | Zadoks scale |
| | Grain yield | Grain yield (t $ha^{-1}$) |
| Soil | Chemical components | Soil organic carbon, C/N ratio, nitrogen |
| Climate | Weather data | Mean temperature (°C), cumulative precipitation (mm) |
| Agronomic management | Soil management | No tillage, minimum tillage, conventional tillage |
| | | Continuous or intermittent |
| | Nitrogen level | kg N $ha^{-1}$ |
| | Sowing day | Date of sowing |

To determine the optimal timing for image acquisition, the Zadoks scale was employed. Images were obtained at three key growth stages: Z22 (tillering), Z35 (stem elongation), and Z60 (anthesis). These images were used to calculate the normalized difference vegetation index (NDVI) (Rouse et al., 1974). Additionally, at the maturity stage, grain yield (t $ha^{-1}$) was determined within randomly selected and georeferenced test areas, each consisting of a 1-m-long row, using a laboratory thresher (Fiorentini et al., 2019). For georeferencing purposes, the Leica Zeno 20 (Leica Geosystem, Heerbrugg, Canton St. Gallen, Switzerland) was employed at each experimental site. Multispectral images were acquired around 12:00 am using two different UAV platforms and multispectral cameras. In the Marche region sites, the MAIA S-2 Multispectral Camera (SAL engineering, Russi, Italy) was used to capture multispectral images for NDVI computation, using the following formula:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \tag{1}$$

where Red (665 nm) is the red band, and NIR is the near-infrared spectral band (865 nm).

For the Basilicata region sites, multispectral images were obtained using the Parrot Sequoia multispectral sensor (Parrot Drone SAS, Paris, France) for NDVI computation. Both UAVs were equipped with an incident light sensor to mitigate the effects of sunlight variations during flight, ensuring optimal imagery acquisition by recording changes in light levels, and subsequently correcting the spectral bands. To enhance the spatial accuracy of the acquired images, ground control points (GCPs) were strategically positioned within the study area and georeferenced before each UAV flight operation. Each image acquired during an UAV flight underwent a series of image processing steps, including geometric correction, coregistration, radiometric correction, file format conversion, and preparation of image datasets for the orthomosaic map generation process. The image processing workflow comprised three main steps: (1) orthomosaic reflectance map generation, (2) computation of the vegetation indices (VI) map, and (3) data extraction. After the UAV flights, raw images from the multispectral cameras were transferred to a computer. Agisoft Metashape (Agisoft LLC), a software based on the structure-from-motion (SfM) algorithm (Verhoeven, 2011), was used to generate the orthomosaic reflectance map. Subsequently, the orthomosaic reflectance map was imported into R statistical software (Core Team, 2014). The "stack" function of the raster R package (Hijmans et al., 2011) was used in the second and third steps of imaging processing. Based on previous studies (Orsini et al., 2019a, 2019b; Pro et al., 2021), NDVI was identified as a vegetation index highly correlated with key durum wheat crop parameters. To extract the NDVI values from specific sampling points, we imported a.xlsx file containing the georeferenced coordinates of the samples. This was achieved using the "read_excel" function of the *readxl* R package (Wickham, 2016). Subsequently, the "*extract*" function of the *raster* R package (Hijmans et al., 2011) was used to extract these values, organizing them into a data frame format. The resulting data-frame object was then exported in CSV file format using the "write.csv" function of the *utilis* R package (Wickham et al., 2019). Soil organic matter was determined according to the modified Walkley–Black method (Walkley & Black, 1934). The Van Bemmelen factor (= 1.724) has been used to derive soil organic carbon, while the soil texture was determined using the pipette method (Gee & Bauder, 1986). The C/N ratio, nitrogen (N) content and pH were measured at each experimental site. Specifically, thirty-six soil samples were collected at the LTE site; and six soil samples were obtained at the Recanati site; Basilicata 1, and Basilicata 2. All soil samples were collected prior to sowing at each respective site.

For climate data, the NASA POWER API (Sparks, 2018) was used to retrieve precipitation and temperature data. In this study, mean temperature (°C) and cumulative precipitation (mm) were calculated between the sowing date and the date of the multispectral image. In terms of agronomic management features, soil management and nitrogen management were used. Soil management was characterized by two categorical variables (tillage type and intensity). First, the type of soil management was categorized as no-tillage, minimum tillage, and conventional tillage. Second, the temporal aspect of soil management was considered, distinguishing between continuous soil management (repeated consistently over time) and intermittent soil management (applied every two years). It's worth noting that only the LTE site employed continuous soil management, while the farm sites used an intermittent soil management system. Nitrogen management was considered as continuous variable, denoting the amount of nitrogen (kg N ha$^{-1}$) applied to the crop.

## Modelling procedure

The entire modelling procedure was performed using the R statistical programming languages. Table 5 lists the R packages that were utilized, along with their respective purposes.

The ML-based system combines two ML models, each with a different aim and scope. The first ML model is designed to classify crop phenology according to the Zadoks scale (classification task), while the second ML model is focused on predicting crop yield (t ha$^{-1}$) (Fig. 2) regression task.

The first ML model was trained using climate and remote sensing features, while the regression ML model was trained with all the covariates listed in Table 4, along with the crop phenology predictions from the first ML model. Four ML algorithms were compared for both classification, and regression tasks: (1) Random Forest (Svetnik et al., 2003) was used as the benchmark algorithm; (2) eXtreme Gradient Boosting, aka XGBoost (Chen &

**Table 5** R packages used in the modeling phase and their objective

| Aim | R Package |
|---|---|
| Handling raster data | Raster |
| Timestamp data management | Lubridate |
| Import and export excel files | Readxl |
| Spatial data management | Sf |
| Handling tabular data | Tidyverse |
| Connect R, PostgreSQL and Postgis | Rpostgis |
| Framework of R packages for modeling and machine learning | Tidymodels |
| Variable Importance Plots | Vip |
| An end-to-end open-source deep learning platform | Tensorflow |
| Keras is a high-level API to build and train deep learning models | Keras |
| A suite of tools for tracking, visualizing, and managing TensorFlow training runs and experiments from R | Tfruns |
| R Packages to provide summary statistics about variables in data frames, tibbles, data tables and vectors | Skimr |
| Ensembling Learners builder for R | Stacks |
| Error metrics computation | Mlmetrics |

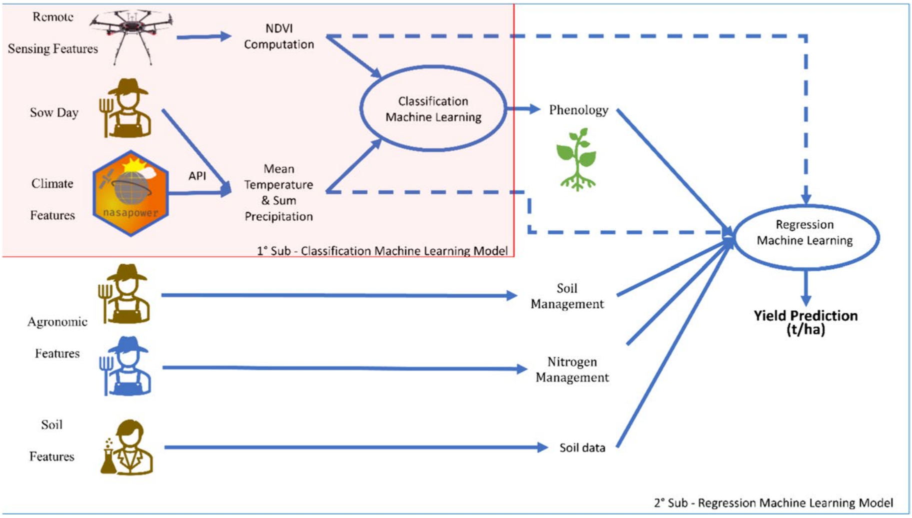



**Fig. 2** Machine learning based system

Guestrin, 2016); (3) Neural Network (McCulloch & Pitts, 1943); and (5) Ensemble Learner (Zhang & Ma, 2012). These ML algorithms were selected due to their ease of comparison in the scientific literature (Chlingaryan et al., 2018). In statistics and ML, ensemble methods use multiple learning algorithms to obtain enhanced predictive performance compared to what any individual learning algorithm could deliver in isolation (Zhou, 2009). All hyperparameters for the ML model were adjusted through a 50-grid search and fivefold cross-validation (Table 6). When assessing the algorithms for classification, the performance metrics included accuracy, kappa, precision, and recall. For the regression task, the evaluation was conducted based on multiple metrics, including the coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Moreover, the 'vip' function of the vip R package was used to visualize the variable importance, to describe which covariate had the most significant impact on improving model accuracy.

The ML algorithms were trained using data from the LTE and Basilicata 1 experimental sites, which consisted of 873 data points. Subsequently, the accuracy of the models was

**Table 6** Machine learning algorithms, abbreviation, R packages and hyperparameter tuned

| Model | Abbreviation | R package | Tuning hyperparameter |
|---|---|---|---|
| Random forest | RF | randomForest | mtry, N. trees, minimum number of data points in a node that are required for the node to be split further |
| XGBoost | XG | xgboost | Tree depth, N. trees, learning rate, mtry, minimum number of data points in a node that are required for the node to be split further, loss reduction |
| Neural network | NN | tensorflow | N. of neurons for each hidden layer, penalty, epochs |
| Ensemble learners | EL | stacks | Penalty |

assessed using the Recanati and Basilicata 2 experimental sites, which served as test datasets and contained 108 data points.

### Simulations and Marginal Fertilizer Yield Index (MFYI) calculations

Following the training and validation of the ML models, the best model for both the classification and regression tasks was selected based on the error metrics. These chosen models were then combined to create a meta-model, which allowed the regression model to use the predictions of the classification model. The meta-model was subsequently used to run simulations, considering all possible combinations of soil management practices (no-tillage, minimum tillage, and conventional tillage) and ten levels of nitrogen fertilization. The range of N levels varied from 40 kg N ha$^{-1}$ to 220 kg N ha$^{-1}$, with increments of 20 kg N ha$^{-1}$ between levels. In total, 30 simulations were conducted for each experimental site (Fig. 3).

Simulations served to quantify the yield (t ha$^{-1}$) for each combination of soil and nitrogen management. Subsequently, the Marginal Fertilizer Yield Index (MFYI) (Kyveryga et al., 2007) was calculated. The MFYI is a metric used to assess the economic efficiency of fertilizer use in agriculture. It helps identify the point at which further increases in fertilizer application do not result in proportionate increases in crop yield, indicating a diminishing return on fertilizer investment. The calculation of the MFYI involved the following steps: (1) Data collection, gather data on crop yield (e.g., in kg ha$^{-1}$) achieved at different levels of fertilizer application (e.g., nitrogen input in kg ha$^{-1}$); (2) Data sorting, arrange the data in ascending order based on the levels of fertilizer application; (3) Calculate the Marginal Yield: calculated as the difference in crop yield between two consecutive fertilizer application levels (Eq. 2); (4) Calculate the MFYI, divide the Marginal Yield by the corresponding increase in fertilizer input (Eq. 3).

$$Marginal\ Yield = (Crop\ yield\ at\ fertilization\ level\ x) - (Crop\ yield\ at\ fertilization\ level\ x - 1) \tag{2}$$

$$MFYI = \frac{Marginal\ Yield}{Increase\ in\ fertilizer\ input} \tag{3}$$

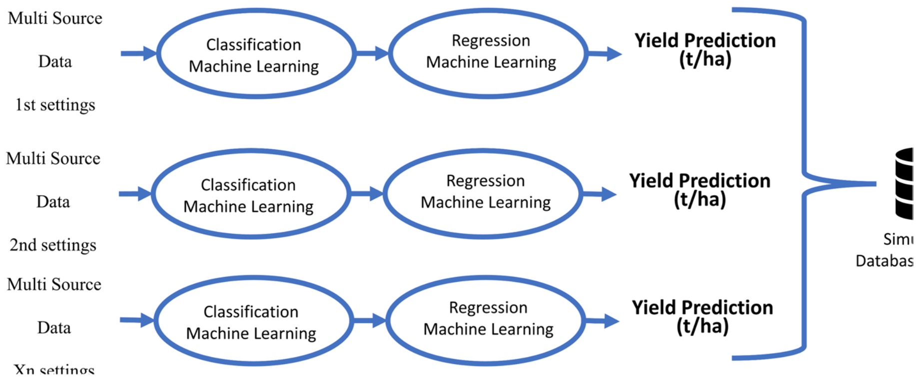



**Fig. 3** Outline of simulations and data flow for determination of best agronomic practice

## Results

### Descriptive statistics and data exploration

Descriptive statistics were performed to analyze grain yield (t $ha^{-1}$) in the entire dataset and separately for each experimental sites (Table 7). A total of 981 observations were analyzed and the dataset exhibited a mean grain yield of 3.48 t $ha^{-1}$, with a minimum value of 0.29 t $ha^{-1}$ and a maximum value of 8.35 t $ha^{-1}$ (Table 7). Notably, the Recanati site had the highest average grain yield (6.14 t $ha^{-1}$), while the LTE site showed the highest standard deviation (+ 1.87). Conversely, the LTE site had the lowest grain yield, with a value of 0.29 t $ha^{-1}$, while the Recanati site reported the highest grain yield of 8.35 t $ha^{-1}$.

### Classification and regression machine learning models

Table 8 provides the results of the five-fold cross-validation and hyperparameters tuning procedure, which involved a 50 grid-search. For the classification task, during the calibration phase, all models achieved high performance with an accuracy, kappa, precision, and recall of 1, except for the Neural Network. In the evaluation phase, the Random Forest and the Ensemble Learners models performed the best, with a mean error metric of 0.98. The other models, in descending order of accuracy, were XGBoost and Neural Network, with mean error metrics of 0.97 and 0.96, respectively. For the regression task, during the calibration phase, the Random Forest and Ensemble Learners were the best performing models reaching an $R^2$, RMSE, MAE and MAPE of 0.97, 0.35, 0.25 and 0.09, respectively (Fig. 4). In comparison, the XGBoost and Neural Network models reached an $R^2$ of 0.96 and 0.91, respectively. In the evaluation phase, using a dataset that was not part of the training procedure, each model showed a slight worsening of the error metrics. Among the models, the Neural Network performed the best, with a MAPE of 0.17; it was followed by the XGBoost, Ensemble Learners, and Random Forest with MAPE values around 0.18 for each (Table 8). The Neural Network demonstrated a smaller difference in error metrics between training and evaluation.

### Variable importance

Figure 5 displays the variable importance of the different machine learning models for the classification task.

Random Forest and XGBoost considered precipitation as the most important variable (~ 100), followed by temperature (28) and NDVI (18). The Neural Network, on the other

**Table 7** Descriptive statistics of yield (t $ha^{-1}$) measured at the four experimental sites

| Experimental sites | N. Observations | Years of experiment | Mean | St. Dev | Min | Max |
|---|---|---|---|---|---|---|
| All sites | 981 | 2018, 2019, 2020, 2021, 2022 | 3.48 | 1.91 | 0.29 | 8.35 |
| LTE | 810 | 2018, 2019, 2020, 2021, 2022 | 3.44 | 1.87 | 0.29 | 7.50 |
| Recanati | 108 | 2020, 2021 | 6.14 | 1.12 | 4.02 | 8.35 |
| Basilicata 1 | 63 | 2021 | 1.66 | 1.11 | 1.50 | 3.75 |
| Basilicata 2 | 54 | 2020 | 3.62 | 1.27 | 1.44 | 4.87 |

**Table 8** Hyperparameter tuning results, calibration, and evaluation of each individual model for both classification and regression task

| Objective | Model | Tuning hyperparameter | Calibration | | | | Evaluation | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Accuracy | Kappa | Precision | Recall | Accuracy | Kappa | Precision | Recall |
| Classification | Random forest | mtry = 2, N. trees = 275, minimum number of data points in a node that are required for the node to be split further = 19 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 0.96 | 0.98 | 0.98 |
| | XGBoost | Tree depth = 7, N. trees = 572, learning rate = 0.0101, mtry = 1, minimum number of data points in a node that are required for the node to be split further = 5, loss reduction = 0.0175 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 0.97 | 0.98 | 0.98 |
| | Neural network | N. of neurons = 9, penalty = 0.00000739, epochs = 233 | 0.99 | 0.98 | 0.98 | 0.98 | 0.97 | 0.95 | 0.97 | 0.96 |
| | Ensemble learners | penalty = 0.000001 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 0.98 | 0.98 | 0.98 |
| | | | $R^2$ | RMSE | MAE | MAPE | $R^2$ | RMSE | MAE | MAPE |
| Regression | Random forest | mtry = 9, N. trees = 1025, minimum number of data points in a node that are required for the node to be split further = 3 | 0.97 | 0.35 | 0.25 | 0.09 | 0.88 | 0.64 | 0.47 | 0.19 |
| | XGBoost | Tree depth = 5, N. trees = 680, learning rate = 0.0968, mtry = 9, minimum number of data points in a node that are required for the node to be split further = 36, loss reduction = 0.00000301 | 0.96 | 0.40 | 0.29 | 0.11 | 0.89 | 0.63 | 0.46 | 0.18 |
| | Neural network | N. of neurons = 8, penalty = 0.00296, epochs = 886 | 0.91 | 0.58 | 0.43 | 0.15 | 0.90 | 0.59 | 0.45 | 0.17 |
| | Ensemble learners | Penalty = 0.00001 | 0.97 | 0.35 | 0.25 | 0.09 | 0.89 | 0.61 | 0.46 | 0.18 |

The creation of the meta-model involved combining Random Forest for the classification task and the Neural Network for the regression task. Random Forest was chosen for its popularity and simplicity, with fewer parameters compared to the other models. Using this meta- ML model, 30 simulations were executed for each site

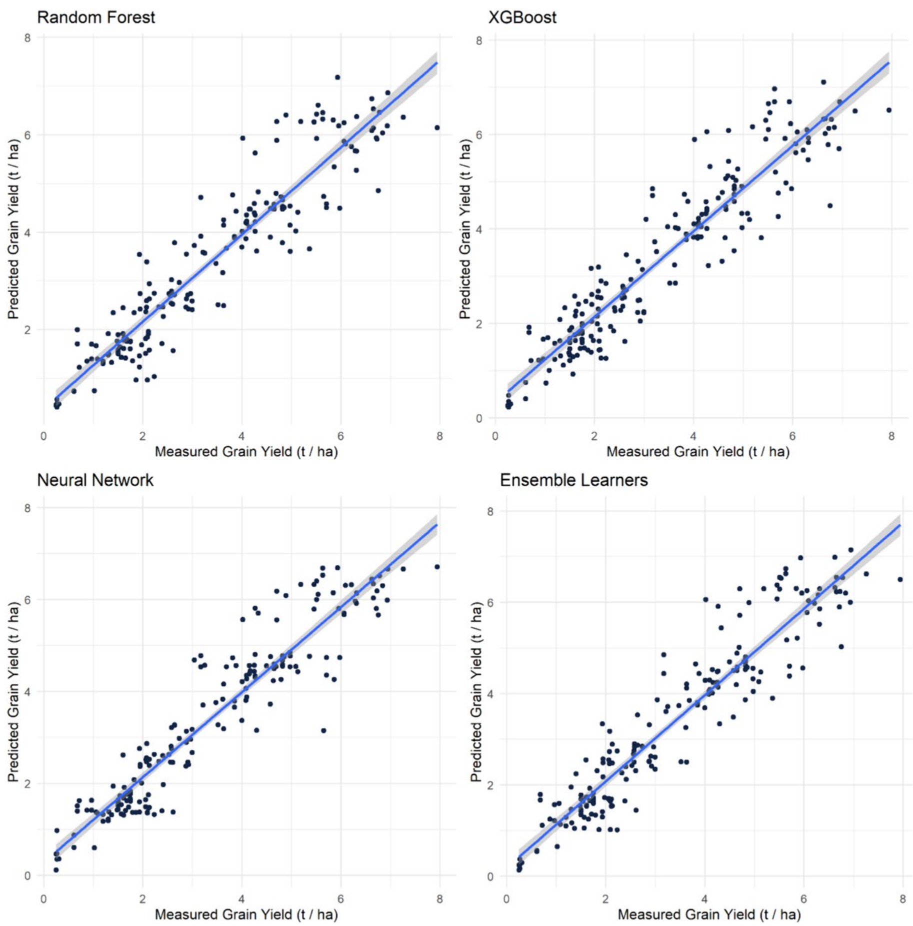



**Fig. 4** Scatterplot of the difference between predicted and measured gain yield (t ha$^{-1}$) in the regressive testing phase

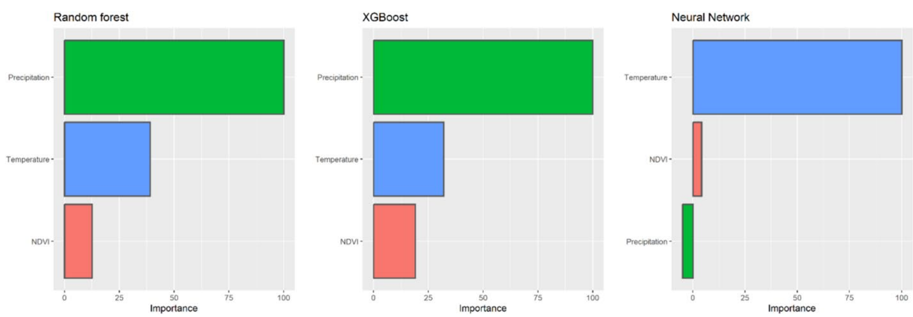



**Fig. 5** Variable importance of the machine learning models for the classification task

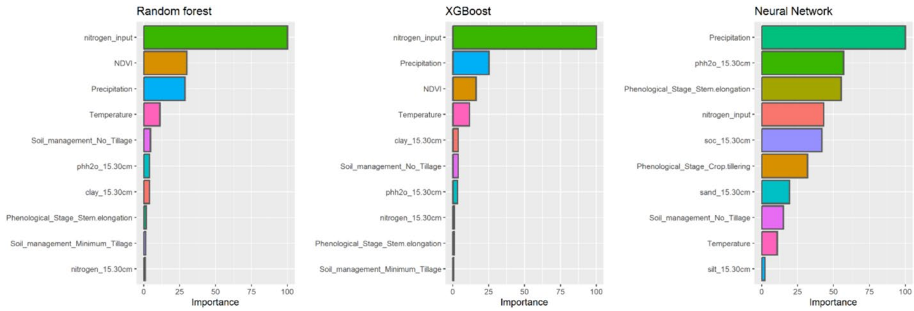



**Fig. 6** Variable importance of the machine learning models for the regression task

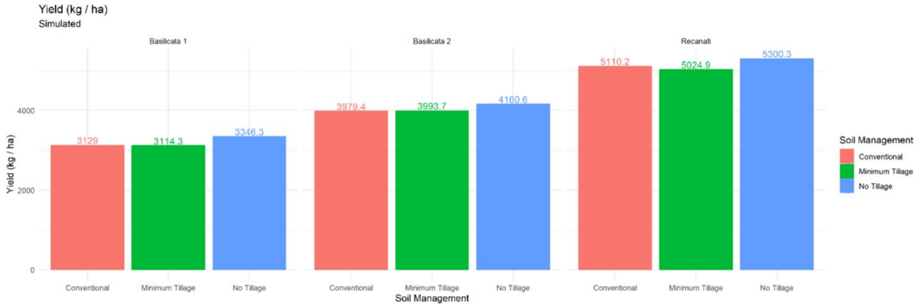



**Fig. 7** Results of the simulations at the three farm sites related to the soil management

hand, found temperature to be the most important variable (~ 100), with NDVI (7) also contributing, and with negative importance for precipitation (− 7). For the regression task, the variable importance of the models is shown in Fig. 6.

The variable importance results suggest that different ML models emphasize various features when predicting crop yield in the regression task. For Random Forest and XGBoost, nitrogen management was the most important feature (~ 100), followed by precipitation (27.5), NDVI (23.5), and temperature (15). Soil chemical components, soil management, and crop phenology had a relatively lower importance (less than 10). On the other hand, the Neural Network placed importance on multiple features, with precipitation, pH, phenological stage, nitrogen management, being some of the top-ranking features.

## Simulations and Marginal Fertilizer Yield Index calculations

Figure 7 presents the results of the simulations for different soil management practices at the three farm sites. The data showed that no-tillage consistently achieved the highest grain yield, with an average yield of 4,05 t $ha^{-1}$ at each of the farm sites. Minimum tillage was the next best performing practice, with an average yield of 3,97 t $ha^{-1}$, while conventional tillage had the lowest average yield of 3,88 t $ha^{-1}$ across the sites.

Figure 8 illustrates the results of the MFYI calculated for each simulation at the different farm sites. For Recanati site, the MFYI values exceeded 1 at nitrogen fertilization levels of 100 kg N $ha^{-1}$ and above. This indicates that for the Recanati site, applying 100 kg N $ha^{-1}$ of nitrogen or more resulted in a diminishing return on fertilizer investment. There


**Fig. 8** Results of the simulations at the three farm sites related to the MFYI and nitrogen input (kg N $ha^{-1}$)

was a noticeable increase in MFYI from 80 kg N $ha^{-1}$ to 100 kg N $ha^{-1}$ before decreasing. At the Basilicata 1 site, the MFYI values were above 1 for nitrogen fertilization levels of 120 kg N $ha^{-1}$ and above. Similar to Recanati, applying 120 kg N $ha^{-1}$ of nitrogen or more led to diminishing returns. There was an increase in MFYI from 80 kg N $ha^{-1}$ to 120 kg N $ha^{-1}$ before decreasing. At the Basilicata 2 site, the MFYI values did not reach 1, at any nitrogen fertilization level, with the highest value obtained being 0.90 at 100 kg N $ha^{-1}$. For this site, applying more nitrogen did not result in a significant increase in crop yield, as indicated by the MFYI values. There was a minor increase from 80 kg N $ha^{-1}$ to 100 kg N $ha^{-1}$ before decreasing.

## Discussions

### Classification and regression models, training and testing

In this study, a meta- ML model was developed, consisting of two linked ML models. The first model's role was to classify and determine the phenological stage of the crop, while the second model was responsible for predicting crop yield. The crop phenological stage classification model reached an accuracy of 0.98, kappa of 0.96 and recall of 0.98, even though it was classifying into only three stages: crop tillering, stem elongation and anthesis. This limited number of phenological stages is indeed a constraint of the study, as other research has considered a more extensive range of phenological stages (Tao et al., 2022). However, when comparing the error metrics with those from other studies, such as Arya et al., (2022), using ML in combination with climate and remote sensing data allows very high accuracy levels. This makes the approach suitable for large-scale applications with a high degree of confidence (Zheng et al., 2021). Several works have demonstrated the high accuracy of ML models and the multi-data source approach for precise yield predictions across various crops, including cotton (Filippi et al., 2019; Leo et al., 2021), durum wheat (Chergui, 2022; Dietrich et al., 2022; Fiorentini et al., 2022) and maize (Attia et al., 2022; Nyéki et al., 2021; Shahhosseini et al., 2020). Numerous studies have primarily focused on validating the framework, including the selection of covariates, data pre-processing, and model building. However, the scientific community nowadays is increasingly challenged to go beyond validation and to make these models accessible for determining the best agronomic management (McNunn et al., 2019). Having access to these models allows farmers to be able to benefit directly from scientific studies and use their outputs to apply environmentally and economically beneficial practices, as well as empowering the farmers with scientific evidences to make informed decisions. To ensure that these generated models are both usable and useful, one approach is to pose "if-else-questions" to the model. This approach helps in understanding how altering soil management and fertilization levels can affect crop yield and their economic and environmental impact. This information enables us to determine the most suitable agronomic practices for each specific site, taking into account the associated costs (Cammarano et al., 2023). It is important to note that while these are valuable for optimizing agronomic practices, they cannot predict or account for adverse climatic events that may occur during crop growing season. Therefore, it is recommended that in the event of adverse weather conditions, the "if-else-questions" should be rerun to redefine the most suitable agronomic management. In this study, the Neural Network model performed the best in the regression step, achieving an accuracy of 0.90, a result consistent with findings in another research (Filippi et al., 2019).

## Variable importance

Variable importance analysis is a procedure used to determine which of the variables used in training a ML model contributed to improving the model's performance (Carneiro et al., 2023). In the classification models, only climate data and remote sensing were used, and it was found that the most important variables, in descending order, were precipitation, temperature and NDVI. This is because all climate variables were calculated from the date of sowing to the date on which the multispectral image was acquired. This consideration is important because using this calculated climate data, there is not much variation, allowing the model to achieve high accuracy (Zheng et al., 2021). A common approach for modeling phenological development is to calculate degree days, which is based on sowing date, temperature data, and thresholds specific to different crops (Ram et al., 2016). In the regression models, all variables were used, including the phenological stage provided by the prediction of the classification model. The most important features, in descending order, were nitrogen input, precipitation, temperature and NDVI (Orsini et al., 2023). Soil management (Simoniello et al., 2022), crop phenology, pH, texture, nitrogen, and soil organic carbon have also contributed to the model, but to a lesser extent compared to the first set of covariates. The importance of nitrogen input as the most significant variable in predicting durum wheat grain yield aligns with several studies that have revealed the key role of nitrogen in crop productivity (Abad et al., 2004; Grahmann et al., 2014). Research conducted in Northwest India, where ML was applied to predict durum wheat yields, also identified nitrogen as a key covariate for the Random Forest model (Nayak et al., 2022). Nitrogen levels influence various physiological process in plants, such as of chlorophyll content (Fiorentini et al., 2019) and metabolic pathways involved in aminoacids production. These processes are fundamental for grain quality and quantity (Abad et al., 2004). Therefore, models that do not consider nitrogen fertilization in their covariates list may not be effective in forecasting grain yield accurately. Precipitation emerges as the second most critical variable because it plays a pivotal role in providing water for crop growth. This is especially significant in rainfed agriculture where irrigation is not applied. Precipitation not only supplies the necessary water for crop development but also serves to solubilize nutrients from fertilizers and aids in the mineralization of organic matter, making these nutrients available for plant uptake (Li et al., 2022). Temperature significant influences crop development, as it can be well-modeled using degree-days, which accurately represent various phenological development of the crop (Li et al., 2012). In China Han et al., (2020) showed that temperature, particularly minimum temperature, plays a crucial role in crop development. When assessing the importance of variables in a Random Forest model, the Enhanced Vegetation Index, minimum temperature, precipitation, NDVI, and soil moisture emerged as some of the most significant factor affecting crop growth yield. The NDVI is identified as the fourth most important variable. It is widely recognized vegetation index used in precision agriculture to assess the status of crops. It is worth noting that NDVI can exhibit saturation in cases where the leaf area index exceeds 2 (Aklilu Tesfaye & Gessesse Awoke, 2021). Additionally, NDVI based forecast has been proved highly accurate for inter-annual yield prediction (Toscano et al., 2019). However, ML algorithms can help mitigate this saturation by transforming nonlinear variables into linear ones (Chollet & Allaire, 2018). The choice of soil management practices is another critical factor for crop yield, and this variable is consistently found among the top ten most important variables in the ML models used in this study. Several authors showed the benefits of no-tillage practices in improving soil organic matter (Page et al., 2020; Valkama et al., 2020). No-tillage practices help maintain soil structure and reduce disturbance, promoting the activity of beneficial microorganisms (Wacker et al., 2022). These improvements in soil health have been associated with higher grain yield (Li et al., 2022). Crop phenology is also a crucial variable for predicting crop yield. It provides valuable information about the developmental stage of the crop and can indicate whether the crop is experiencing any stress. By incorporating crop phenology and relative NDVI values, machine learning models can better understand the crop's condition, stress levels, and growth stage (Richetti et al., 2018). In this study, the regression model used the predicted crop phenology from the classification model to enhance its accuracy and ability to estimate crop yield (Haghverdi et al., 2018).

## Simulations and marginal fertilizer yield index calculations

In this study, the trained meta-model was used to run yield simulations exploring the yield impact of the combination of various combinations of fertilization and soil management. Other widely recognized models like SALUS (Liu & Basso, 2017) and DSSAT (Jones et al., 2003; Wang et al., 2022) also run simulations. However, these crop models need site-specific calibration and rely on predefined covariates set by the creators. In this study, the covariates were selected to ensure model scalability (Hansen & Jones, 2000) and facilitate the use of it even within production contexts. This means that each covariate can be obtained for any field using publicly available datasets, such as soil grids (Poggio et al., 2021), the Land Use/Cover Area frame statistical Survey (LUCAS) database (Andrimont et al., 2020), Terraclimate (Abatzoglou et al., 2018), and Copernicus Sentinel 2 multispectral images, providing all the data required for running the meta-model with zero cost. We conducted thirty simulations using the meta-model across three farm sites, considering all possible combinations of three soil management practices (no-tillage, minimum tillage, and conventional tillage) and ten levels of nitrogen fertilization. This allowed us to determine the optimal agronomic management for each site, considering various combinations and their impact on yield, ultimately providing adaptable recommendations for yearly application. Other studies have used calibrated models to perform simulations, enabling projections based on climate change. These simulations aid in formulating on-season suitable agronomic practices to mitigate the impacts of rising temperatures (Basso et al., 2015; Cammarano et al., 2020). After running the simulations, we assessed the nitrogen levels that yield the maximum margin between nitrogen dosage and crop yield. This aspect is of utmost importance in the current context, given the influence of climate change, droughts, food security concerns, and the impact of geopolitical factors on agricultural input costs (Arndt et al., 2023). Our findings consistently revealed that no-tillage practices resulted in higher yields when compared to other soil management approaches across all the sites. Furthermore, we determined the most effective nitrogen dosage by utilizing simulated yield data to calculate the MFYI. This index helps in identifying the threshold beyond which further increases in fertilizer application fail to produce proportionate increases in crop yield, allowing the farmers to intuitively identify the level of nitrogen that effectively increases nitrogen use efficiency, as well as reduced environmental impact and increased economic returns. For the Recanati and Basilicata 2 sites, the optimal nitrogen level was found to be 100 kg N ha$^{-1}$, while for Basilicata 1, it was 120 kg N ha$^{-1}$.

## Conclusions

A meta-model was trained using a multi-data source approach, incorporating remote sensing, crop phenology, soil chemical components, weather data, soil management, and nitrogen levels to predict durum wheat yield. This meta-model comprises two interconnected ML models, each with a distinct role: the first model classifies phenological stages using Random Forest with a remarkable accuracy of 0.98, while the second model predicts yield using a Neural Network, achieving a high accuracy of 0.90. Within each individual model comprising the meta-model, we conducted variable importance analysis to identify the most influential covariates that contribute to accuracy improvement. In the classification model, the top three significant features were temperature, precipitation, and NDVI. In contrast, the regression model placed the greatest importance on nitrogen input, followed by precipitation, NDVI, and temperature. The meta-model was subsequently employed to run simulations involving different soil management and fertilization levels, aiding in the determination of the optimal agronomic practices. Notably, no tillage practices consistently yielded higher grain production. Additionally, the Marginal Fertilizer Yield Index was used to identify the optimal nitrogen application levels for the crop. It is worth emphasizing that the scalability of this model can be achieved through the utilization of publicly available datasets. In the realm of agriculture, the potential of ML models is promising, provided they are user-friendly, accessible to farmers, and perceived as valuable tools.

**Acknowledgements** Marche Polytechnic University through the University Strategic Project (PSA 2017) “PFRLab: Setting of a Precision Farming Robotic Laboratory for cropping system sustain-ability and food safety and security” and Marche Region through measure 16.1 Action 2 of the Rural Development Plan (PSR 2014/2020) “Precision agriculture: reduction of the environmental impact of production systems” (project ID 29000). HORIZON 2020 Project (grant number 952879), SolAcqua, “Accessible, reliable, and affordable solar irrigation for Europe and beyond”. CERESO Project, the Rural Development Program of Basilicata Region-mis. 16.2. The authors would like to thank the engineers and workers of the “Pasquale

Rosati” experimental farm of the Università Politecnica delle Marche for their technical contributions. NASA POWER API service for rapidly downloading and elaborating on the weather data.

**Data availability** Data will be made available for research purposes upond request to the first author.

## References

Abad, A., Lloveras, J., & Michelena, A. (2004). Nitrogen fertilization and foliar urea effects on durum wheat yield and quality and on residual soil nitrate in irrigated Mediterranean conditions. *Field Crops Research, 87*(2–3), 257–269. https://doi.org/10.1016/j.fcr.2003.11.007

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data, 5*(1), 170191. https://doi.org/10.1038/sdata.2017.191

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture, 44*(1), 71–91. https://doi.org/10.1016/j.compag.2004.03.002

Aklilu Tesfaye, A., & Gessesse Awoke, B. (2021). Evaluation of the saturation property of vegetation indices derived from sentinel-2 in mixed crop-forest ecosystem. *Spatial Information Research, 29*(1), 109–121. https://doi.org/10.1007/s41324-020-00339-5

Argento, F., Anken, T., Abt, F., Vogelsanger, E., Walter, A., & Liebisch, F. (2021). Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. *Precision Agriculture, 22*(2), 364–386. https://doi.org/10.1007/s11119-020-09733-3

Arndt, C., Diao, X., Dorosh, P., Pauw, K., & Thurlow, J. (2023). The Ukraine war and rising commodity prices: Implications for developing countries. *Global Food Security, 36*, 100680. https://doi.org/10.1016/j.gfs.2023.100680

Arya, S., Sandhu, K. S., Singh, J., & kumar, S. (2022). Deep learning: As the new frontier in high-throughput plant phenotyping. *Euphytica, 218*(4), 47. https://doi.org/10.1007/s10681-022-02992-3

Attia, A., Govind, A., Qureshi, A. S., Feike, T., Rizk, M. S., Shabana, M. M. A., & Kheir, A. M. S. (2022). Coupling process-based models and machine learning algorithms for predicting yield and evapotranspiration of maize in arid environments. *Water, 14*(22), 3647. https://doi.org/10.3390/w14223647

Basso, B., Hyndman, D. W., Kendall, A. D., Grace, P. R., & Robertson, G. P. (2015). Can impacts of climate change and agricultural adaptation strategies be accurately quantified if crop models are annually re-initialized? *PLoS ONE, 10*(6), e0127333. https://doi.org/10.1371/journal.pone.0127333

Basso, B., Ritchie, J. T., Grace, P. R., & Sartori, L. (2006). Simulation of tillage systems impact on soil biophysical properties using the SALUS model. *Italian Journal of Agronomy, 1*(4), 677. https://doi.org/10.4081/ija.2006.677

Cammarano, D., Miguez, F. E., & Puntel, L. (2023). *Process-Based Models and Simulation of Nitrogen Dynamics* (pp. 105–127). Springer. https://doi.org/10.1007/978-3-031-15258-0_5

Cammarano, D., Ronga, D., Di Mola, I., Mori, M., & Parisi, M. (2020). Impact of climate change on water and nitrogen use efficiencies of processing tomato cultivated in Italy. *Agricultural Water Management, 241*, 106336. https://doi.org/10.1016/j.agwat.2020.106336

Capmourteres, V., Adams, J., Berg, A., Fraser, E., Swanton, C., & Anand, M. (2018). Precision conservation meets precision agriculture: A case study from southern Ontario. *Agricultural Systems, 167*, 176–185. https://doi.org/10.1016/j.agsy.2018.09.011

Carneiro, F. M., Filho, A. L. de B., Ferreira, F. M., Junior, G. de F. S., Brandão, Z. N., da Silva, R. P., & Shiratsuchi, L. S. (2023). Soil and satellite remote sensing variables importance using machine learning to predict cotton yield. *Smart Agricultural Technology, 5*, 100292. https://doi.org/10.1016/j.atech.2023.100292

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. https://doi.org/10.1145/2939672.2939785

Chergui, N. (2022). Durum wheat yield forecasting using machine learning. *Artificial Intelligence in Agriculture, 6*, 156–166. https://doi.org/10.1016/J.AIIA.2022.09.003

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture* (Vol. 151, pp. 61–69). Elsevier.

Chollet, F., & Allaire, J. J. (2018). *Deep learning with R*. Manning Publications.

d'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P., Gallego, J., Reuter, H. I., Joebges, C., Lemoine, G., & van der Velde, M. (2020). Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union. *Scientific Data, 7*(1), 1–15. https://doi.org/10.1038/s41597-020-00675-z

De Menna, F., Malagnino, R., Vittuari, M., Molari, G., Seddaiu, G., Deligios, P., Solinas, S., & Ledda, L. (2016). Potential biogas production from artichoke byproducts in Sardinia, Italy. *Energies, 9*(2), 92. https://doi.org/10.3390/en9020092

Deligios, P. A., Farina, R., Tiloca, M. T., Francaviglia, R., & Ledda, L. (2021). C-sequestration and resilience to climate change of globe artichoke cropping systems depend on crop residues management. *Agronomy for Sustainable Development, 41*(2), 20. https://doi.org/10.1007/s13593-021-00680-5

Denora, M., Amato, M., Brunetti, G., De Mastro, F., & Perniola, M. (2022). Geophysical field zoning for nitrogen fertilization in durum wheat (*Triticum durum* Des.). *PLoS ONE*. https://doi.org/10.1371/journal.pone.0267219

Denora, M., Fiorentini, M., Zenobi, S., Deligios, P. A., Orsini, R., Ledda, L., & Perniola, M. (2022). Validation of rapid and low-cost approach for the delineation of zone management based on machine learning algorithms. *Agronomy*. https://doi.org/10.3390/agronomy12010183

Dietrich, S., Bebie, M., Cavalaris, C., & Kyparissis, A. (2022). Assessing durum wheat yield through sentinel-2 imagery: A machine learning approach. *Remote Sensing*. https://doi.org/10.3390/rs14163880

Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D. S. N., Pozza, L. E., Ugbaje, S. U., Jephcott, T. G., Paterson, S. E., Whelan, B. M., & Bishop, T. F. A. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture, 20*(5), 1015–1029. https://doi.org/10.1007/s11119-018-09628-4

Fiorentini, M., Schillaci, C., Denora, M., Zenobi, S., Deligios, P., Orsini, R., Santilocchi, R., Perniola, M., Montanarella, L., & Ledda, L. (2022). A machine learning modelling framework for *Triticum turgidum* subsp. durum Desf yield forecasting in Italy. *Agronomy Journal*. https://doi.org/10.1002/agj2.21279

Fiorentini, M., Schillaci, C., Denora, M., Zenobi, S., Deligios, P., Orsini, R., Santilocchi, R., Perniola, M., Montanarella, L., & Ledda, L. (2023). 52. Early prediction of durum wheat yield in Italy using a machine learning modelling framework. *Precision Agriculture, 23*, 423–428. https://doi.org/10.3920/978-90-8686-947-3_52

Fiorentini, M., Zenobi, S., Giorgini, E., Basili, D., Conti, C., Pro, C., Monaci, E., & Orsini, R. (2019). Nitrogen and chlorophyll status determination in durum wheat as influenced by fertilization and soil management: Preliminary results. *PLoS ONE, 14*(11), e0225126. https://doi.org/10.1371/journal.pone.0225126

Gee, G. W., & Bauder, J. W. (1986). Particle-size Analysis. *Methods of Soil Analysis* (pp. 383–411). Soil Science Society of America, American Society of Agronomy.

Grahmann, K., Verhulst, N., Peña, R. J., Buerkert, A., Vargas-Rojas, L., & Govaerts, B. (2014). Durum wheat (*Triticum durum* L.) quality and yield as affected by tillage–straw management and nitrogen fertilization practice under furrow-irrigated conditions. *Field Crops Research, 164*(1), 166–177. https://doi.org/10.1016/J.FCR.2014.05.002

Haghverdi, A., Washington-Allen, R. A., & Leib, B. G. (2018). Prediction of cotton lint yield from phenology of crop indices using artificial neural networks. *Computers and Electronics in Agriculture, 152*, 186–197. https://doi.org/10.1016/J.COMPAG.2018.07.021

Han, J., Zhang, Z., Cao, J., Luo, Y., Zhang, L., Li, Z., & Zhang, J. (2020). Prediction of winter wheat yield based on multi-source data and machine learning in China. *Remote Sensing*. https://doi.org/10.3390/rs12020236

Hansen, J. W., & Jones, J. W. (2000). Scaling-up crop models for climate variability applications. *Agricultural Systems, 65*(1), 43–72. https://doi.org/10.1016/S0308-521X(00)00025-1

Hijmans, R. J., van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., & Shortridge, A. (2011). Raster: raster: Geographic data analysis and modeling. In *R package, version* (pp. 2–0).

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P. M., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., & Keating, B. A. (2014). APSIM—Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software, 62*, 327–350. https://doi.org/10.1016/J.ENVSOFT.2014.07.009

Hoogenboom, G., Jones, J. W., Wilkens, P. W., Porter, C. H., Batchelor, W. D., Hunt, L. A., Boote, K. J., Singh, U., Uryasev, O., Bowen, W. T., & others. (2004). Decision support system for agrotechnology transfer version 4.0. *University of Hawaii, Honolulu, HI (CD-ROM)*.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J., & Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy, 18*(3–4), 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture* (Vol. 147, pp. 70–90). Elsevier. https://doi.org/10.1016/j.compag.2018.02.016

Kayad, A., Sozzi, M., Gatto, S., Marinello, F., & Pirotti, F. (2019). Monitoring within-field variability of corn yield using sentinel-2 and machine learning techniques. *Remote Sensing*. https://doi.org/10.3390/rs11232873

Kyveryga, P. M., Blackmer, A. M., & Morris, T. F. (2007). Alternative benchmarks for economically optimal rates of nitrogen fertilization for corn. *Agronomy Journal, 99*(4), 1057–1065. https://doi.org/10.2134/agronj2006.0340

Le Page, M. (2023). World population could soon peak. *New Scientist, 257*(3432), 11. https://doi.org/10.1016/S0262-4079(23)00552-3

Li, Q. Y., Yin, J., Liu, W. D., Zhou, S. M., Li, L., Niu, J. S., Niu, H. Bin., & Ma, Y. (2012). Determination of optimum growing degree-days (GDD) range before winter for wheat cultivars with different growth characteristics in North China Plain. *Journal of Integrative Agriculture, 11*(3), 405–415. https://doi.org/10.1016/S2095-3119(12)60025-2

Li, Z., Cui, S., Zhang, Q., Xu, G., Feng, Q., Chen, C., & Li, Y. (2022). Optimizing wheat yield, water, and nitrogen use efficiency with water and nitrogen inputs in China: A synthesis and life cycle assessment. *Frontiers in Plant Science, 13*, 1982. https://doi.org/10.3389/FPLS.2022.930484/BIBTEX

Leo, S., De Antoni Migliorati, M., & Grace, P. R. (2021). Predicting within-field cotton yields using publicly available datasets and machine learning. *Agronomy Journal, 113*(2), 1150–1163. https://doi.org/10.1002/agj2.20543

Liu, L., & Basso, B. (2017). Spatial evaluation of switchgrass productivity under historical and future climate scenarios in Michigan. *GCB Bioenergy, 9*(8), 1320–1332. https://doi.org/10.1111/gcbb.12417

Lugato, E., Panagos, P., Bampa, F., Jones, A., & Montanarella, L. (2014). A new baseline of organic carbon stock in European agricultural soils using a modelling approach. *Global Change Biology, 20*(1), 313–326. https://doi.org/10.1111/gcb.12292

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics, 5*(4), 115–133. https://doi.org/10.1007/BF02478259

McNunn, G., Heaton, E., Archontoulis, S., Licht, M., & VanLoocke, A. (2019). Using a crop modeling framework for precision cost-benefit analysis of variable seeding and nitrogen application rates. *Frontiers in Sustainable Food Systems*. https://doi.org/10.3389/fsufs.2019.00108

Nayak, H. S., Silva, J. V., Parihar, C. M., Krupnik, T. J., Sena, D. R., Kakraliya, S. K., Jat, H. S., Sidhu, H. S., Sharma, P. C., Jat, M. L., & Sapkota, T. B. (2022). Interpretable machine learning methods to explain on-farm yield variability of high productivity wheat in Northwest India. *Field Crops Research, 287*, 108640. https://doi.org/10.1016/j.fcr.2022.108640

Nyéki, A., Kerepesi, C., Daróczy, B., Benczúr, A., Milics, G., Nagy, J., Harsányi, E., Kovács, A. J., & Neményi, M. (2021). Application of spatio-temporal data in site-specific maize yield prediction with machine learning methods. *Precision Agriculture, 22*(5), 1397–1415. https://doi.org/10.1007/s11119-021-09833-8

Orsini, R., Basili, D., Belletti, M., Bentivoglio, D., Bozzi, C. A., Chiappini, S., Conti, C., Galli, A., Giorgini, E., Fiorentini, M., Malinverni, E. S., Mancini, A., Mazzanti, L., Monaci, E., Passerini, G., Pro, C., Santilocchi, R., Vignini, A., Zenobi, S., & Zingaretti, P. (2019a). Setting of a precision farming robotic laboratory for cropping system sustainability and food safety and security: Preliminary results. *IOP Conference Series: Earth and Environmental Science*. https://doi.org/10.1088/1755-1315/275/1/012021

Orsini, R., Fiorentini, M., & Zenobi, S. (2019b). Testing vegetation index categories as influenced by soil management and nitrogen fertilization in cereal based cropping systems. *IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), 2019*, 13–18. https://doi.org/10.1109/MetroAgriFor.2019.8909216

Orsini, R., Fiorentini, M., Zenobi, S., Santilocchi, R., Deligios, P. A., & Ledda, L. (2023). Automatic zone management definition in Mediterranean environment for *Triticum durum*. *Agrosystems, Geosciences & Environment,*. https://doi.org/10.1002/agg2.20422

Page, K. L., Dang, Y. P., & Dalal, R. C. (2020). The ability of conservation agriculture to conserve soil organic carbon and the subsequent impact on soil physical, chemical, and biological properties and yield. *Frontiers in Sustainable Food Systems*. https://doi.org/10.3389/fsufs.2020.00031

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture, 23*(4), 1489–1513. https://doi.org/10.1007/s11119-022-09885-4

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL, 7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021

Pro, C., Basili, D., Notarstefano, V., Belloni, A., Fiorentini, M., Zenobi, S., Alia, S., Vignini, A., Orsini, R., & Giorgini, E. (2021). A spectroscopic approach to evaluate the effects of different soil tillage methods and nitrogen fertilization levels on the biochemical composition of durum wheat (*Triticum turgidum* subsp. durum) leaves and caryopses. *Agriculture*. https://doi.org/10.3390/agriculture11040321

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*.

Ram, H., Gupta, N., & Saini, J. S. (2016). Growing Degree day requirements and yield ability of irrigated durum wheat as influenced by sowing time. *Agricultural Research Journal, 53*(3), 303. https://doi.org/10.5958/2395-146X.2016.00059.4

Richetti, J., Judge, J., Boote, K. J., Johann, J. A., Uribe-Opazo, M. A., Becker, W. R., Paludo, A., & Silva, L. C. de A. (2018). Using phenology-based enhanced vegetation index and machine learning for soybean yield estimation in Paraná State, Brazil. *Journal of Applied Remote Sensing, 12*(2), 1–15. https://doi.org/10.1117/1.JRS.12.026029

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the great plains with ERTS. *NASA Special Publication, 351*(1974), 309.

Schillaci, C., Inverardi, F., Battaglia, M. L., Perego, A., Thomason, W., & Acutis, M. (2022). Assessment of hail damages in maize using remote sensing and comparison with an insurance assessment: A case study in Lombardy. *Italian Journal of Agronomy*. https://doi.org/10.4081/ija.2022.2126

Schillaci, C., Perego, A., Valkama, E., Märker, M., Saia, S., Veronesi, F., Lipani, A., Lombardo, L., Tadiello, T., Gamper, H. A., Tedone, L., Moss, C., Pareja-Serrano, E., Amato, G., Kühl, K., Damatirca, C., Cogato, A., Mzid, N., Eeswaran, R., & Acutis, M. (2021). New pedotransfer approaches to predict soil bulk density using WoSIS soil data and environmental covariates in Mediterranean agro-ecosystems. *Science of The Total Environment, 780*, 146609. https://doi.org/10.1016/j.scitotenv.2021.146609

Scott, D. A., Liebig, M. A., Saliendra, N. Z., Toledo, D., DeGreef, M., Kobilansky, C., & Feld, J. (2024). Crop- and weather-dependent yield and wind erosion benefits from a conservation practices system. *Soil Science Society of America Journal*. https://doi.org/10.1002/saj2.20646

Shahhosseini, M., Hu, G., & Archontoulis, S. V. (2020). Forecasting corn yield with machine learning ensembles. *Frontiers in Plant Science*. https://doi.org/10.3389/FPLS.2020.01120

Shi, P., Wang, Y., Xu, J., Zhao, Y., Yang, B., Yuan, Z., & Sun, Q. (2020). Rice nitrogen nutrition estimation with RGB images and machine learning methods. *Computers and Electronics in Agriculture, 180*, 105860. https://doi.org/10.1016/j.compag.2020.105860

Simoniello, T., Coluzzi, R., D'Emilio, M., Imbrenda, V., Salvati, L., Sinisi, R., & Summa, V. (2022). Going conservative or conventional? Investigating farm management strategies in between economic and environmental sustainability in Southern Italy. *Agronomy, 12*(3), 597. https://doi.org/10.3390/agronomy12030597

Sparks, A. H. (2018). nasapower: A NASA POWER global meteorology, surface solar energy and climatology data client for R. *The Journal of Open Source Software, 3*(30), 1035. https://doi.org/10.21105/joss.01035

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences, 43*(6), 1947–1958. https://doi.org/10.1021/ci034160g

Talaviya, T., Shah, D., Patel, N., Yagnik, H., & Shah, M. (2020). Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture, 4*, 58–73. https://doi.org/10.1016/J.AIIA.2020.04.002

Tao, H., Xu, S., Tian, Y., Li, Z., Ge, Y., Zhang, J., Wang, Y., Zhou, G., Deng, X., Zhang, Z., Ding, Y., Jiang, D., Guo, Q., & Jin, S. (2022). Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives. *Plant Communications, 3*(6), 100344. https://doi.org/10.1016/j.xplc.2022.100344

Terribile, F., Acutis, M., Agrillo, A., Anzalone, E., Azam-Ali, S., Bancheri, M., Baumann, P., Birli, B., Bonfante, A., Botta, M., Cavaliere, F., Colandrea, M., D'Antonio, A., De Mascellis, R., De Michele, C., De Paoli, G., Monica, C. D., Di Leginio, M., Ferlan, M., & Basile, A. (2024). The LANDSUPPORT geospatial decision support system (S-DSS ) vision: Operational tools to implement sustainability policies in land planning and management. *Land Degradation & Development, 35*(2), 813–834. https://doi.org/10.1002/ldr.4954

Toscano, P., Castrignanò, A., Di Gennaro, S. F., Vonella, A. V., Ventrella, D., & Matese, A. (2019). A precision agriculture approach for durum wheat yield assessment using remote sensing data and yield mapping. *Agronomy, 9*(8), 437. https://doi.org/10.3390/agronomy9080437

Tsialtas, J. T., Theologidou, G. S., & Karaoglanidis, G. S. (2018). Effects of pyraclostrobin on leaf diseases, leaf physiology, yield and quality of durum wheat under Mediterranean conditions. *Crop Protection, 113*, 48–55.

Valkama, E., Kunypiyaeva, G., Zhapayev, R., Karabayev, M., Zhusupbekov, E., Perego, A., Schillaci, C., Sacco, D., Moretti, B., Grignani, C., & Acutis, M. (2020). Can conservation agriculture increase soil carbon sequestration? A modelling approach. *Geoderma, 369*, 114298. https://doi.org/10.1016/j.geoderma.2020.114298

van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture, 177*, 105709. https://doi.org/10.1016/J.COMPAG.2020.105709

Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeological Prospection, 18*(1), 67–73. https://doi.org/10.1002/arp.399

Wacker, T. S., Jensen, L. S., & Thorup-Kristensen, K. (2022). Conservation agriculture affects soil organic matter distribution, microbial metabolic capacity and nitrogen turnover under Danish field conditions. *Soil and Tillage Research, 224*, 105508. https://doi.org/10.1016/J.STILL.2022.105508

Walkley, A., & Black, I. A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science, 37*, 29–38. https://doi.org/10.1097/00010694-193401000-00003

Wang, E., He, D., Wang, J., Lilley, J. M., Christy, B., Hoffmann, M. P., O'Leary, G., Hatfield, J. H., Ledda, L., Deligios, P. A., Grant, B., Jing, Q., Nendel, C., Kage, H., Qian, B., Rezaei, E. E., Smith, W., Weymann, W., & Ewert, F. (2022). How reliable are current crop models for simulating growth and seed yield of canola across simulating growth and seed yield of canola across global sites and under future climate change? *Climatic Change*. https://doi.org/10.1007/s10584-022-03375-2

Wang, R., Liu, L., Xie, C., Yang, P., Li, R., & Zhou, M. (2021). AgriPest: A large-scale domain-specific benchmark dataset for practical agricultural pest detection in the wild. *Sensors*. https://doi.org/10.3390/s21051601

Wickham, H. (2016). *readxl: Read Excel Files*. https://cran.r-project.org/package=readxl

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning: Methods and Applications*. Springer.

Zhang, N., Zhou, X., Kang, M., Hu, B.-G., Heuvelink, E., & Marcelis, L. F. M. (2023). Machine learning versus crop growth models: an ally, not a rival. *AoB PLANTS*. https://doi.org/10.1093/aobpla/plac061

Zheng, C., Abd-Elrahman, A., & Whitaker, V. (2021). Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. *Remote Sensing, 13*(3), 531. https://doi.org/10.3390/rs13030531

Zhou, Z.-H. (2009). Ensemble Learning. *Encyclopedia of Biometrics* (pp. 270–273). Springer. https://doi.org/10.1007/978-0-387-73003-5_293

## Authors and Affiliations

**Marco Fiorentini[1] · Calogero Schillaci[2] · Michele Denora[3] · Stefano Zenobi[1] · Paola A. Deligios[1] · Rodolfo Santilocchi[1] · Michele Perniola[3] · Luigi Ledda[1] · Roberto Orsini[1]**

✉ Calogero Schillaci
calogero.schillaci@ec.europa.eu

1 Department of Agricultural, Food and Environmental Sciences (D3A), Agronomy and Crop Science, Marche Polytechnic University, Ancona, Italy

2 European Commission, Joint Research Centre (JRC), Ispra, Italy

3 Department of European and Mediterranean Cultures, Environment, and Cultural Heritage, University of Basilicata IT, Matera, Italy