PhD School on Agriculture, Environment and Bioenergy

(http://sites.unimi.it/dottorato_aab/)

(XXXVII cycle, 2021-24)

Project draft

1. Field of interest

Agronomy (AGR/02 - Agronomia e coltivazioni erbacee)

2. Project title

Integration of crop models and machine learning: methods and applications for agro-environmental simulations

3. Tutor

Prof. Luca Bechini

Co-tutor: Prof. Giorgio Ragaglini

4. Relevance of the topic and state of the art

Agriculture in the recent decades has increased the interest in collecting complex information about field status in order to better manage nutrients, water, chemicals (Wójtowicz et al., 2016) and tillage operations as well. There is a growing demand for tools able to identify cropping system needs and losses. Process based models (PBMs) are considered key instruments to face these challenges; these are computer-based tools that manage data from different sources and represent mathematically the dynamics of soil-crop-atmosphere systems (Wallach et al., 2018). However, PBMs have some limits when used at farm scale and in non-experimental conditions, because it is not always possible to retrieve input data of sufficient quality (Donohue et al., 2018). In recent years, machine learning (ML) has also been used to assess agro-environmental variables useful to support decision making (Hoffman et al., 2020; Saha et al., 2021). Machine learning relies on the application of artificial intelligence tools, represented by complex algorithms and statistical models able to analyse big data to identify underlying patterns and do inference. Their application in agriculture is very recent and promising, but still empirical. Recently, simulation models and machine learning techniques have been used together to overcome the limitation of the two approaches used separately. These attempts are still limited (e.g. Feng et al., 2019; Shahhosseini et al., 2021), and mostly related to large-scale assessments. However, coupling PBMs with ML has the potential to increase the explanatory ability of spatiotemporal variability of field conditions that, otherwise, cannot be completely accounted by simulation models. Indeed, ML approaches might be used also to improve in-season predictions by exploiting PBM outputs, long-term climatic and technological variables, and remotely sensed multispectral data.

Therefore, the objective of this project is to evaluate the potential of integrating PBMs and ML to predict crop growth and development, and water/carbon/nitrogen budgets, in relation to pedoclimatic conditions and soil/crop management.

5. Layout of the project (draft)

5.1. Materials & Methods

The project will be carried out through four different activities:

1) Building a conceptual framework

This task includes the formal design of the integrated simulation model, the identification of possible benchmarks and of the experimental trials to be used.

2) Data collection and pre-processing

This action includes:

- Collection of experimental data related to input and state variables from short and long-term experiments, considering both the national and the international level.
- Data quality assessment, scaling and building of harmonised datasets.
- Building simulation scenarios.

3) Simulations with PBMs

In this task, the ARMOSA (Perego et al., 2013) simulation model will be calibrated and used to simulate crop growth and development, water / C / N budgets, and environmental losses (e.g. nitrogen leaching and GHG fluxes).

4) Assessing machine learning algorithms (alone or integrated with PBMs)

Different machine learning algorithms (MLAs) will be considered in order to assess their prediction potential and stability across different training and test datasets and their ability in explaining the relevance of predictors on the variability of responses. MLAs will be fed with measured variables alone and with a combination of measured and PBM-simulated variables.

5) Analysis of the results

Results of the different simulation approaches (PBMs alone, MLAs alone and integrated PBM-MLA) will be evaluated against the measured agro-environmental variables of the case studies. The comparison of the three simulation approaches will be based on a set of different metrics and the most suited MLAs will be identified.

5.2. Schedule and major steps (3 years)

The research will be articulated in the following activities and tasks.

Activity	Task	Year 1	Year 2	Year 3
Building on a conceptual framework towards an experimental plan	Literature review and framework design	X		
Data collection and pre-processing	Dataset development	X		
PBM simulations	Parameterization of the simulation model	X	X	
Assessing MLAs	Implementation of machine learning		X	
	Coupling simulation model and machine learning		X	X
Analysis of the results	PBM evaluation on errors of prediction		X	
	MLA evaluation on			X

Activity	Task	Year 1	Year 2	Year 3
	errors of prediction			
	Integrated PBM-MLA evaluation on errors of prediction			X
Thesis and article writing			X	X

6. Available funds

Progetto CONSENSI, "Ottimizzazione della concimazione mediante la sensoristica e metodi dell'agricoltura di precisione", 66.661 €

Progetto SOS-AP, "SOluzioni Sostenibili per l'Agricoltura di Precisione in Lombardia: irrigazione e fertilizzazione rateo-variabile in maidicoltura e viticoltura".

7. Literature

Donohue, R.J., Lawes, R.A., Mata, G., Gobbett, D., Ouzman, J., 2018. Towards a national, remote-sensing-based model for predicting field-scale crop yield. Field Crops Research 227, 79–90. https://doi.org/10.1016/j.fcr.2018.08.005

Feng, P., Wang, B., Liu, D.L., Waters, C., Yu, Q., 2019. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. Agricultural and Forest Meteorology 275, 100–113. doi:10.1016/j.agrformet.2019.05.018

Hoffman, A.L., Kemanian, A.R., Forest, C.E., 2020. The response of maize, sorghum, and soybean yield to growing-phase climate revealed with machine learning. Environmental Research Letters 15, 094013. doi:10.1088/1748-9326/ab7b22

Perego, A., Giussani, A., Sanna, M., Fumagalli, M., Carozzi, M., Alfieri, L., Brenna, S., Acutis, M., 2013. The ARMOSA simulation crop model: overall features, calibration and validation results. Italian Journal of Agrometeorology 3, 23–38.

Saha, D., Basso, B., Robertson, G.P., 2021. Machine learning improves predictions of agricultural nitrous oxide (N_2O) emissions from intensively managed cropping systems. Environmental Research Letters 16, 024004. https://doi.org/10.1088/1748-9326/abd2f3

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific Reports 11, 1606. doi:10.1038/s41598-020-80820-1

Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2018. Working with dynamic crop models: methods, tools and examples for agriculture and environment. Academic Press.

Wójtowicz, M., Wójtowicz, A., Piekarczyk, J., 2016. Application of remote sensing methods in agriculture. Communications in Biometry and Crop Science 11, 31–50.