European Journal of Agronomy 139 (2022) 126569

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Mixing process-based and data-driven approaches in yield prediction

Bernardo Maestrini [a,*], Gordan Mimić [b], Pepijn A.J. van Oort [a], Keiji Jindo [a], Sanja Brdar [b], Ioannis N. Athanasiadis [a], Frits K. van Evert [a]

[a] Wageningen University & Research, Wageningen, the Netherlands
[b] BioSense Institute–Research and Development Institute for Information Technologies in Biosystems, University of Novi Sad, Novi Sad, Serbia

ARTICLE INFO

Keywords:
Crop models
Yield prediction
Data-driven
Process-based
Crop growth models
Dynamic crop growth models
Artificial intelligence
Neural networks
Metamodels
Surrogate models

ABSTRACT

Yield prediction models can be divided between data-driven and process-based models (crop growth models). The first category contains many different types of models with parameters learned from the data themselves and where domain knowledge is only used to select the predictors and engineer features. In the second category, models are based upon biophysical principles, whose structure and parameters are derived primarily from domain knowledge. Here we investigate if the integration of the two approaches can be beneficial as it allows to overcome the limitations of the two approaches taken individually - lack of sufficiently large, reliable and orthogonal datasets for data-driven approaches and the need of many inputs for process-based models. The applications of the two categories of models have been reviewed, paying special attention to the cases where the two approaches have been mixed. By analysing the literature we identified three major cases of integration between the two approaches: (1) using crop growth models to engineer features and expand the predictors space, (2) use data-driven approaches to estimate missing inputs for process-based models (3) using data-driven approaches to produce meta-models to reduce computation burden. Finally we propose a methodology based on metamodels and transfer learning to integrate data-driven and process-based approaches.

1. Introduction

Data-driven and process-based approaches are two extremes of a spectrum of modeling approaches and both are used in crop yield prediction. Yield prediction is important for many actors, for example in-season real-time yield prediction can be used to timely signal shortages of water or nutrients and thus for management optimization. Yield prediction can be relevant for processors planning the logistics of yield collection, storage and processing. At a regional to global level yield predictions can be important for timely signaling of upcoming food shortages. Over the last decades, yields have been predicted mainly with theory-driven approaches or simple regression models. However, the increased availability of data from regional to field scale has paved the way for new data-driven approaches. Here we aim at reviewing the two methodologies, paying particular attention to how the two approaches have been integrated in the past and how this can be done in the future. Of particular interest is the question whether integrating the two approaches may lead to more accurate predictions of yield than using either approach separately.

1.1. Process-based approaches

Process-based approaches for yield prediction rely on a large body of theories from the fields of crop ecology, crop physiology, meteorology, and soil science. A set of theories from these scientific fields is incorporated in so-called “crop growth models”. Crop growth models are based on physical processes, for example, they may include equations that describe the movement of water in the soil or the penetration of light in the canopy. Crop growth modelers make choices regarding the level of detail at which processes are described in their model. These simplifications may lead to biases in model outcomes. Such biases are often “fixed” through model calibration, ideally based on experiments and knowledge of realistic parameter bounds. Being based on physical, biological and chemical laws, process-based models and their scope of application is theoretically universal. In practice, one often finds a fair amount of time has to be spent on model calibration before attaining accurate yield predictions. Certain processes are better understood than others. Almost all crop growth models can simulate the effects of CO_2, weather, soil moisture and soil nitrogen on growth. Very few can simulate the effects of reducing factors (*sensu* Van Ittersum and

* Correspondence to: Postbus 16, 6700 AA Wageningen, the Netherlands.
E-mail address: bernardo.maestrini@wur.nl (B. Maestrini).

https://doi.org/10.1016/j.eja.2022.126569
Received 4 August 2021; Received in revised form 26 April 2022; Accepted 15 June 2022
Available online 8 July 2022

Rabbinge, 1997) such as weeds, pests, diseases (Donatelli et al., 2017; Silva and Giller, 2021), and nutrients other than nitrogen. For a recent overview on crop models and their use we refer to Boote (2019), and Silva and Giller (2021) whereas for inter-comparisons of crop growth models see Fleisher et al. (2017), Fronzek et al. (2018), Kimball et al. (2019).

1.2. Data-driven approaches

At the other end of the spectrum, data-driven approaches for yield prediction rely solely on data for their parameterization. Data-driven approaches apply functions to the inputs to get the desired output and the parameters of the function are trained (i.e., calibrated or parameterized) to minimize the difference between the observed and simulated values. The functions applied to transform the input vary wildly as the family of data-driven models is large and include linear and non-linear regressions, random forests, decision trees, support vector machines, neural networks, and many more. The main advantage of data-driven approaches is that no prior information on the relationship between the variables is required for their parameterization (training) and there is great flexibility in the type of input that can be used. Moreover data-driven approaches can easily accommodate new explanatory variables for which there is as of yet limited theory available. However, data-driven approaches have three main drawbacks: they do not use the domain knowledge available besides the data, their scope is limited to the scope where they have been trained (i.e., their results can not be extrapolated to new contexts), and they need large "training datasets" to discover patterns in data, especially if the patterns are non-linear and involve interactions between predictors.

1.3. Objectives and outline

Scientists tend to have training in at most one of these two methods. Crop growth modelers tend to have a strong background in the scientific disciplines of agronomy, meteorology, soil science and farming systems analysis. Data scientists tend to have a background in computer science, statistics, machine learning and big data. The first objective of this paper is to take stock of the current state of both approaches. To focus our work, we compare the two approaches in applications of yield prediction.

The second objective is to discuss the advantages that can derive from the combination of the two approaches. The benefit of integrating knowledge of physical processes in data-driven models has been discussed previously by Karpatne et al. (2017) and here we focus on the domain of agriculture and yield prediction. We investigate the thesis that a mixed approach can help overcome the limits of both. For example processes that are not comprehended in process-based models (e.g., pests and disease) may be included through machine learning and the prior-knowledge encoded in process-based models may facilitate algorithms learning relevant predictive features.

This review paper is organized as follows. In the next section, we describe the methods and the results of a systematic literature review on yield prediction studies. The systematic review is followed by a qualitative review on how the methods have been used in the literature and a section where we investigate the advantages in combining these approaches and propose metamodels and transfer learning as tools to combine them.

2. A systematic survey of yield prediction methods used in literature

To have an overview of the published methods to predict yield, we conducted a literature survey, read and extracted information on the methodology used in 109 selected papers. The purpose of the systematic review was to understand the scale and the input used for yield prediction by data-driven and process-based approaches. The systematic review was also aimed at identifying case studies where the two approaches have been integrated.

2.1. Method for the selection of papers categorized

We searched the Scopus abstract and citation database[1] using the following query:

DOCTYPE(ar or re) AND TITLE-ABS-KEY(crop yield) AND TITLE-ABS-KEY(predict OR forecast OR monitor* OR calculat*) AND TITLE-ABS-KEY ("decision support system" OR regional OR farm* OR tactical OR strategic OR scenario* OR policy* OR DSS OR (compar* AND method) OR review OR AI OR "machine learning" OR "deep learning" OR "random forest" OR "neural network" OR "support vector machines" OR "artificial intelligence" OR "big data" OR "statistic* model*") AND PUBYEAR > 2008.*

We can break down the query into the following blocks linked by an AND operator:

- Published between 2009 and 2019 (the query was run on the 12th of July 2019)
- Included a "predictive" word (*predict OR forecast* OR monitor* OR calculate**)
- Included a word from a purpose words list (e.g., *decision support system*), or a word from a method words list (e.g., *machine learning, crop growth model*).

We tested the quality of our query and ensured that a set of 12 papers - that we deemed relevant for this study - would be included in the query's output.

We read through the titles (~15603) returned by the query to select the ones that — based solely on the title — were within the scope of our review (~14 % of the titles). We selected the articles that were based on open field crops, because there are more process-based models for these crops, and therefore, our exercise of comparing the two approaches becomes more meaningful.

From the list of articles selected on the basis of their title, we read 109 of them. Eight articles out of 109 were review papers, therefore, they were read but not included in the systematic review.

For each of the 101 articles included (Table SI1), we recorded the following information:

1. Spatial scale. This category represents the spatial scale of the data used to validate the model or, if the model was not validated to calibrate the model. We set this information as a two-level categorical variable: field or farm scale, or above (typically county, province or state).
2. Model category: this information was a list of the methods used in the study to predict the yield, we reduced the heterogeneity of possible methodologies to two categories: data-driven (linear and non-linear regressions, A.I. methods), and process-based (crop growth models). In 12 cases both approaches were used (Table 1), in these cases, we recorded where the two approaches were used to compare them (e.

Table 1
Contingency table of the articles by scale and approach.

	Farm,field or plot	Higher than farm	Sum
Data-driven	23	24	47
Process-based	32	10	42
Mixed	3	6	9
Comparison modeling approaches	2	1	3
Sum	60	41	101

[1] https://www.scopus.com.

g., the performance of regression vs. process-based model) or to integrate them (mixed) effectively.

3. Input data used as a predictor: for this information, we aggregated the wide range of predictors used in literature in the following categories: weather, soil, management (including irrigation, fertilization, planting date), cultivar (including cases where cultivar duration was a predictor) and canopy reflectance, crop biological variables (e.g., flowering date, or leaf N content) and others (for the input that were used in less than two studies). The inputs category is reported in Table 2. The input was considered as the rawest input, and further variables derived from the original input were disregarded. For example Morel et al. (2014) calculated the fraction of interception of photosynthetic active radiation from normalized difference vegetation index (NDVI), a vegetation index derived from canopy reflectance, therefore in this case, the input was classified as reflectance. The category *other* included, for example, fluorescence, synthetic aperture radar, or cost of the labor. They were categorized as other as they appeared in less than three studies. The inputs used by process-based models are not reported because process-based models require a fixed amount of information, typically weather, soil, and management.

3. Results of the literature survey

3.1. The scale of the prediction

The analysis of the spatial scale of the predictions indicated that process-based models were employed more often at field scale or lower (32 vs. 10 cases, Table 1). In contrast, data-driven models were used equally at the different spatial scales (23 vs 24). The predominant use of process-based models at field scale or lower probably reflects the need of management information (e.g., fertilization, sowing date) to run them. Clearly, process-based models may be run at a regional scale varying only the weather, however, in these cases, the user needs to make assumptions about crop management.

3.2. Inputs and type of model

The analysis of the frequency of the predictors used in the data-driven studies revealed — not surprisingly — that reflectance was the most frequent predictor for data-driven approaches, followed by weather, management, and soil. van Klompenburg et al. (2020) conducted a similar study on the predictors of data-driven models in agriculture and found that weather (temperature, radiation and rainfall in their study) was the most prevalent predictor, however, they found that soil variables were more frequently used than remote sensing images. This discrepancy can be due to differences in the categorization of the predictor variables. When looking at the scale at which the different predictors have been used, we found that weather was more frequently used at regional scale (where differences may arise also between regions in the same year), whereas management was used more often at the field or farm-scale. Surprisingly we found 3 studies where management was used at regional scale, these were Jeong et al. (2016), who used a global dataset of wheat and maize predictions and they could include fertilizer use at national level as a predictor, Crane-Droesch (2018) who trained a neural network to predict yield at county scale in the Midwest of the U.S. and used information on sowing date derived from regional statistics, and Kern et al. (2018) who created a yield prediction model for Hungary and included fertilizer use at national level every year as a predictor.

We found that approximately half of the studies used process-based approaches (crop models) and half of the studies used data-driven approaches. Twelve studies used the two approaches (category mixed in Table 1), however only in nine cases the two approaches were integrated, whereas in the other three cases, the two approaches were used independently for model comparison (e.g., the performance of statistical vs crop model).

4. Qualitative review of yield prediction methods

4.1. Data-driven approaches

Approaches that are purely data-driven have the advantage that they do not require previous knowledge of the physical processes involved, although it could be argued that the decision of which data to use (often simply all the available data) could be regarded as a form of use of pre-existing knowledge. Here below, we report on the different purposes for which data-driven approaches have been used, dividing the subject between studies that used statistical regressions and studies that used machine learning.

4.2. Machine learning

Machine learning methods belong to a broad category that encompasses approaches of various complexity. A few studies focused on the comparison of different machine learning algorithms for yield prediction. For example, Gonzalez-Sanchez et al. (2014) compared various machine learning methods (e.g., Support Vector Machines, Multi-Layer Perceptron, neural networks) for yield prediction of irrigated fields in an irrigated basin in Mexico. Maya Gopal and Bhargavi (2019) tested several machine learning approaches (Artificial Neural Network (ANN), Support Vector Regression, k-Nearest Neighbors, and Random Forest) using regional data from Tamil Nadu state. Similarly, Jeong et al. (2016) and Kim et al. (2019) compared different methods (neural networks with different architectures, Support Vector Machines, Random Forests) to predict maize and soybean yield in the Midwest of the United States. Although several studies compared different methodologies, it is impossible to synthesize them and declare a winning methodology, as this is case-specific, it depends on the amount of data available and their correlation. We can generally observe that approaches characterized by many parameters require a large amount of training data, for example, many studies used neural networks to predict yield using the large USDA-NASS dataset (Bose et al., 2016; Crane-Droesch, 2018; Kim et al., 2019; NASS, 2017) whereas when fewer (and typically at a higher resolution) data are available simpler models may be preferable. For example, both Zheng et al. (2009) and Krupnik et al. (2015) used regression trees to predict yield at field level using data from 48 soybean fields in China and 48 wheat fields of Bangladesh farmers, respectively.

Data-driven models fed with the weather, soil characteristics, and remote sensing data are a good fit for predictions carried out at regional scale. For example, Alvarez (2009) used ANN with soil properties from surveys and monthly weather data to predict wheat yield at the regional level in Argentina, and Mann et al. (2019) used Random Forest to predict yield at the village scale in Ethiopia. Crane-Droesch (2018) used maize yield at the county level for the U.S. Midwest along with daily weather data, soil properties and the proportion of each county that is irrigated, to train a semiparametric neural network and predict yield in

Table 2
Model input used by the studies using data-driven approaches. The numbers indicate the percentage of studies that used that input. The total number of data-driven studies analysed was 47.

	Reflectance	Weather	Management	Soil	Other	Cultivar	Crop biological variable	Land surface temperature	Spatial	Number studies
Farm,field or plot	57	30	30	22	17	22	9	4	4	23
Higher than farm	62	62	12	17	29	0	0	8	8	24

future weather scenarios. Jeong et al. (2016) carried out regional and global crop yield prediction for three staple crops: wheat, maize, and potato from weather data and the amount of fertilizer used. More recent works include a machine learning baseline using both crop simulation outputs and weather, remote sensing and soil data to forecast yields at regional level (Paudel et al., 2021), which has been scaled up to 35 case studies in Europe, including nine countries that are major producers of six crops (soft wheat, spring barley, sunflower, grain maize, sugar beets and potatoes, Paudel et al., 2022).

However, machine learning models have been applied also at field scale, for example Saruta et al. (2013) used Support Vector Machines to build predictive models for yield and protein content of brown rice using explanatory variables representing the growth (plant height), nutrition conditions (nitrogen uptake) and the mean value of weather variables around the heading stage. Plant breeders may also benefit from machine learning approaches, in fact, their data are typically well-curated and span multiple years, sites and genotypes. Khaki and Wang (2019) used data from the maize breeders' trials to build a deep neural network for yield prediction. Because of the size of the dataset and wide range of environments (>100k records), a complex algorithm like a deep neural network outperformed other methods such as LASSO (Least Absolute Shrinkage and Selection Operator), shallow neural networks, and regression tree.

Machine learning applications to agriculture are becoming increasingly popular as they can consume remote sensing measurements that can be captured from an increasing number of platforms, such as satellites, drones, and tractors-mounted cameras. Most studies used reflectance measured from satellites (typically MODIS, Landsat and Sentinel), but also radar and surface temperature information have been used as predictors. For example, You et al. (2017) built a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network to predict soybean yield at the county level in the USA. They used publicly available remote sensing data of surface reflectance, land surface temperature and land cover type derived from the MODIS satellite and extracted features from the transformed normalized histograms of the raw images. Bose et al. (2016) used spiking neural networks for winter wheat yield estimation by NDVI image time series from MODIS, during the growing season in Shandong province (China). They used crop yield data aggregated at regional level and predicted the yield around six weeks before harvest with good accuracy. Johnson et al. (2016) developed crop yield forecast models using a Bayesian Neural Network for barley, canola and spring wheat grown on the Canadian Prairies using NDVI and enhanced vegetation index (EVI) from MODIS. Machine learning and remote sensing has also been used to estimate yield at field level, for example Fieuzal et al. (2017) used ANN with optical and radar satellite data acquired throughout the crop cycle for within-season maize yield prediction at field level in France. The application of machine learning algorithms to data collected using drones is challenged by the fact that a limited number of fields can be monitored within a single study. However Nevavuori et al. (2019) combined the data from yield monitors along with NDVI and the data from Red-Green-Blue (RGB) channels acquired with a drone, to build a model for wheat and barley within-field yield prediction using CNN.

4.3. Statistical models

We included statistical regressions in the data-driven category as their parameterization does not require prior domain knowledge. However, they differ from the kind of machine learning approaches described above as they are less data-hungry and generally allow for simpler interactions between the predictors. We discuss separately the case of regressions based on weather data and on remote sensing.

4.3.1. Regressions on weather and farm—level data

Weather is the primary information for predictions carried out at regional level. For example Conradt et al. (2016), Schlenker and Roberts (2009) and Lobell and Burke (2010) used weather to predict yield at county level respectively in Germany and in the United States. Similarly Chen et al. (2019) built a generalized additive model to predict wheat yield throughout the growing season in Western Australia. Explanatory variables included weather data and derivatives (such as growing season available water, 30-year average rainfall, early season rainfall and germination time), geolocation, soil type, land capability and wheat varieties. Peng et al. (2018) and Li et al. (2019) predicted maize yield at the county level in the US Corn Belt, with EVI from MODIS and using temperature, precipitation, and vapor pressure deficit. Kern et al. (2018) used multiple linear regression to predict the yield of winter wheat, rapeseed, maize and sunflower at county and country-level in Hungary, with NDVI from MODIS and monthly values of temperature, precipitation, vapor pressure deficit, radiation, soil water content, and country-mean nitrogen fertilizer data. The main limitation in the use of regional data is that no information on crop management can be used in aggregated form, and so the model is of limited use to farmers. Although management information can be available at aggregated level, their use in predictive regressions is challenged by the fact that they are correlated with weather variables, posing a multicollinearity problem. For example, average regional planting date (often available) is correlated with temperature and number of frost days (van Oort et al., 2012), and irrigation is typically a function of potential evapotranspiration and rainfall.

Regressions have also been used to make predictions using datasets collected through large famers surveys (Silva et al., 2017a, 2017b) either by collecting data from online decision support systems (Silva et al., 2020) or by collecting different datasets from different sources and different spatial resolutions (Sylvester-Bradley et al., 2019). This is a powerful approach but often showed limited predictive capacities. An important shortcoming of farmers' data is often the lack of variability in the input factors, for example, Nunes-Viera da Silva (Silva et al., 2017a, b) did not observe an effect N fertilization on yield in a dataset of Dutch farms (that typically receive high N dosages) whereas he observed such effect in a dataset from Philippines (where fertilization rate is more variable. Sylvester-Bradley et al. (2019) analyzed a large data set with more than 17k records of wheat yield in British farms and found that the "farm factor", accounted for the most variation, although numerous information on environment, management and genetics were recorded, suggesting that differences due to farmers' skills, attitude and behavior may be more relevant than differences due to machinery or the environment. Another important factor limiting the usability of farmers' data is the inaccuracy of the yield value based on farmer-reported data, for example, Burke and Lobell (2017) report that in developing countries, yield and field size are often over or underestimated, thus not allowing proper cross-validation with the predicted yield.

An advantage of regression models is that they can be trained with modestly-sized datasets and are more adapted to include predictors that are harder to collect compared to weather and reflectance, such as surveys or phenological data. For example, Chen et al. (2017) used a linear mixed model with temperature and precipitation data for the growing season along with the biological, chemical, mechanical data, irrigation and labor inputs measured as cost per hectare, to estimate the impact of climate change on maize yield at farm level in Hebei province (China). Mourtzinis et al. (2013) used multiple linear regression with cumulative precipitation, N fertilization rate and simple plant morphological measurements at a silking stage to predict maize yield at field level in Alabama (USA), while Mourtzinis et al. (2014) used ridge regression with cultivar, spectral reflectance from field sensors and weather data to predict soybean yield in North Central USA. Hernandez et al. (2015) also used ridge regression on reflectance data during anthesis and grain filling to predict wheat yield at field level grown under three water regimes in Chile. Zhu et al. (2019) proposed a crop yield forecasting statistical model that integrates weather variables and crop production information from different geographically correlated regions using a credibility estimator and closed form reinsurance pricing

formulas. The model was built using farm level yield data for Manitoba (Canada) to help insurance companies with risk management perspective for insurers and reinsurers.

4.3.2. Regressions on remote sensing data

There have been many attempts to estimate the crop yield with statistical models using remote sensing data. Son et al. (2013) built a quadratic model using EVI and LAI from MODIS to predict rice crop yield during the growing season in Vietnam. NDVI and the Fraction of Absorbed Photosynthetically Active Radiation from SPOT-VEGETATION instrument were used for wheat yield prediction with linear regression in Tunisia by Meroni et al. (2013), while their normalized values were used with a partial least squares regression for Europe (Kowalik et al., 2014). García-León et al. (2019) compared meteorological and satellite-based drought indices as yield predictors of Spanish cereals and Kogan et al. (2016) assessed cereal yield losses from drought in Saratov (Russia) using regression analysis. Lambert et al. (2018) used the peak LAI from Sentinel-2 satellite for a linear regression of cotton, maize, millet, and sorghum yield at village level in Mali and Pan et al. (2009) proposed a function that combines photosynthetic active radiation, the fraction of absorbed photosynthetic ative radiation and light-use efficiency from QuickBird satellite to estimate the crop yield of spring wheat, pea and alfalfa at the village level in the Loess Plateau (China).

In order to estimate the yield within the growing season using reflectance data, statistical approaches have been applied at the field level as well, mainly using linear regression. Kogan et al. (2018) predicted maize yield with vegetation health indices from different weeks during the season in experimental maize fields in Bulgaria. Al-Gaadi et al. (2016) used NDVI and Soil Adjusted Vegetation Index (SAVI) from Landsat-8 and Sentinel-2 satellites to predict potato yield in irrigated fields in Saudi Arabia. Geipel et al. (2014) utilized RGB reflectance data and a reflectance-derived crop height map to predict maize yield at three early- to mid-season growth stages. Gong et al. (2018) estimated rapeseed yield using drone-obtained canopy reflectance and abundance data at plot level in Central China, they used the product of vegetation indices (VIs) — such as NDVI and SAVI — and short-stalk-leaf abundance to create predictors. Zhou et al. (2017) used vegetation indices (VIs) from unmanned aerial vehicles to perform within-season rice yield prediction at plot level in China, showing that a high correlation with LAI performed well for yield prediction and the best results were obtained with NDVI at booting and heading stages. Magney et al. (2016) used daily values of NDVI from ground-based measurements with spectral reflectance sensors to quantify the rate and duration of phenological periods during the wheat growing season, although in their experiment, daily NDVI values showed insufficient predictive power of harvest metrics such as grain yield and protein concentration. Kumhálová and Matějková (2017) tried to explain yield variability of winter wheat and winter barley at plot level in the Czech Republic using correlation coefficients with NDVI from Landsat, QuickBird and WorldView-2 satellites and GreenSeeker handheld crop sensor.

4.4. Sources of errors in data-driven approaches

There are two major sources of errors in predictions produced using data-driven based models, uncertainty in the data used to calibrate (train) the models and errors derived from using the models out of their original scope. Errors in the data have different origins depending on the source of the dataset. Regional data may come from destructive measured data or from farmers' estimates. When the source of the data are self-reported yields, not surprisingly there may be systematic errors, for example Derriere and Jolliffe, (2018) found that small-holder farmers in Ethiopia tend to over-estimate production of small fields and under-estimate production of large fields. Similarly Irwin and Good, (2016) report that USDA Agricultural yield surveys (farm-reported yield data) may be biased downward compared to measured yield data, particularly in dry years.

However even survey measurements based on ground measurements of the crops may present large errors, for example Kosmowski et al. (2021) found that the Spearman correlation in the yield measured using two similar destructive methods on maize in regional surveys in Ethiopia was only 0.55. Furthermore regional survey data even when based on destructive sampling may be affected not only by non-systematic errors but even by bias, for example Nandram et al., (2014) reports that the USDA objective yield survey is overestimated compared to end-of-year surveys and suggests that this is caused by a systematic overestimation of plant density. Therefore despite data-driven models calibrated at regional level represent a unique opportunity to produce complex data-driven models and push forward the field of modeling, they have two major drawbacks, first that the data on which they are built they are often affected by large noise, or even bias, and second that they cannot incorporate management as an input and thus they cannot be used at lower scales, for example to take decision at farm level. At the farmscale the rapid widespread of agricultural machinery and data-platforms has prompted the collection of large sets of data from farmer fields. This represents a unique opportunity to collect data for data-driven approaches, that can be linked to publicly available sources (e.g. satellite images, soil surveys). There are still however some obstacles to the full exploitation of these resources, like the lack of data on management, whose recording is often lacking behind on decision support platforms, the reliability of harvest recorded by agricultural machinery, which is often highlighted when different machines are used to harvest the same field and the lack of a clear framework for the rights on using farmers' data to produce yield prediction models, the challenge of recording yield reduction factors data (pests and weeds).

The second major source of errors in yield predictions derived from using a data-driven approach is the usage of the model out of its original scope, for example when a model is applied to a different region or when a data-driven model calibrated using past data is used to predict yield under new climatic scenarios, thus while data-driven models calibrated on regional data have generally a wider scope, data-driven models calibrate using farmers' data or trials include management inputs that are more useful to the farmer but are more prone to be used in out-of-scope contexts.

4.5. Key progresses and knowledge gaps in data-driven models

In past two decades there has been a huge increase in the number of studies and publications testing various data-driven models for yield prediction of many different crops.The majority of progresses have been observed in the application of more complex algorithms (such as CNNs or LSTM networks), improvement of spatial, temporal and spectral resolution of remote sensing data and increased amount of publicly available data. However there are a growing number of data available from ground sensors (e.g. soil and leaf moisture sensors, water table depth sensor, towers to measure gas exchanges) that are not yet fully exploited in the context of machine learning because of the lack of large dataset containing both on-the-farm measurements and publicly available data. Such datasets would be particularly useful in the context of machine learning algorithms that are known for being more data hungry. Because of the big-data availability bottle neck in agriculture, agricultural data platform in the future will play a key role to have benchmark datasets with open-source soil, weather and satellite data, and large repository of yield data, especially at the field level.

4.6. Process-based approaches

Crop growth models date back to the 1950 s when De Wit and Van Bavel worked on computation of plant and soil processes (Jones et al., 2016; Wit, 1958). A lot has been written about the purposes of crop models and whether they aim at predicting (engineering purposes) or explaining (scientific purposes, Boote et al., 2013; Passioura, 1996). The

main advantage of crop growth models compared to statistical approaches is that they are based on the physical processes underlying crop growth (e.g., water movement in soil, light interception as a function of LAI), therefore their scope of application should be more universal than statistical approaches. The extent to which this hypothesis is true is highly debatable, for example, most of the studies that we encountered in our review had undergone some calibration. There is an extensive literature on model calibration, for an overview of current practices see the recent work from the Agmip calibration group (Seidel et al. 2018), and Wallach et al. (2019a), (2019b). Briefly the main issue with calibration is that model calibration may range from an adjustment of one or two crop parameters regulating crop maturity class, based on prior knowledge of cultivar maturity to the well-known *"most-cumbersome curve fitting exercise"* (De Wit), where up to 20 parameters are optimized Seidel et al. (2018). Beside the required calibration another important drawback of existing dynamic crop growth models is that they are driven by weather, soil, genetics, and management, so there is little room to incorporate information that can be retrieved from within-season measurements. However, it is now standard practice to include observation through data assimilation procedures, but this approach only allows to incorporate variable that exists in the model, for example leaf moisture time series or observations on pests and diseases could be measured but hardly assimilated in a crop growth model. We now analyze the most frequent uses for crop models that we encountered in our systematic review.

When applied at a regional scale, a crop growth model is run for different situations representing ideal fields, and the outputs are aggregated on the regional level using summary measures (mean, median ...). When using process-based models at a regional scale, there is always uncertainty on the management input that can only be estimated from knowledge on regional practices. However, Constantin et al. (2019) in the context of a model intercomparison for simulations at regional level found that varying the management over space and time using decision rules had a modest effect when simulations are carried out over long periods. Next to using them at regional scale crop growth models are used mostly to manage irrigation, fertilization, and sowing at farm level. That is because these models generally do include process knowledge on water and nutrient response. Irrigation is one of the most addressed targets between the management practices, for example, the AquaCrop model has been used to optimize water use efficiency of winter wheat in China (Xiangxiang et al., 2013) and Turkey (Kale Celik et al., 2017), barley in Iran (Tavakoli et al., 2015) or maize in Uganda (Mibulo and Kiggundu, 2018) and India (Abedinpour et al., 2012). Singels et al. (2019) used the Canesim model for sugarcane in South Africa to compare different irrigation strategies. Amarasingha et al. (2015) used the APSIM-Oryza model to compare the achievable rice yields with and without access to irrigation in Sri Lanka. Chen et al. (2010) used APSIM to evaluate the effects of irrigation supply on the productivity of wheat and maize in the double cropping system in China by estimating the amount of irrigation required to achieve the yield potential. In a similar vein, van Oort et al. (2016) simulated irrigation strategies for ground water timings and amounts that could be sustainably applied for irrigation without depleting groundwater. Paredes et al. (2018) used the SIMDualKc model to assess the impact of different sowing dates on water requirement and yield productivity of potato crops in Italy. Attia et al. (2016) used the CERES-Wheat model in Texas to simulate water use efficiency of winter wheat in response to deficit and full irrigation treatments.

Crop growth models can also be utilized to assess the effects of other management decisions on crop yield at the farm level. Rahman et al. (2019) used the CROPGRO-Cotton model for Pakistan to determine the potential impact of planting dates on cotton yield. The model was calibrated using a diverse range of field observations of phenology, growth, yield. Ngwira et al. (2014) used DSSAT for maize yield in Malawi to choose specific conservation agriculture practices. The results showed positive benefits of no-till systems if accompanied with crop rotation and crop residue retention, where maize–cowpea rotation was the most efficient management system. Bidogeza et al. (2012) also used DSSAT in Rwanda to determine the best fertility management options under a combined use of organic and inorganic fertilizers, showing the potential benefits in the production of maize and sorghum. Lopez et al. (2017) used the CERES model to simulate optimal rooting depth for sweet sorghum in the southeastern USA, indicating that values between 110 and 140 cm can maximize final biomass yield. Borus et al. (2018) evaluated the ability of the APSIM-Potato model to predict nitrogen uptake and potato production under Tasmanian conditions (Australia). The results showed that the model realistically reproduced the observed tuber yield with high precision under various management options. García et al. (2014) evaluated AquaCrop for within-season yield prediction of maize in Colombia where biomass, harvest index, and yield were aligned with the field experiments data. Gilardelli et al. (2018) performed the sensitivity analysis of the WOFOST model in the simulation of the yield of different crops (winter and durum wheat, winter barley, maize, sunflower) across several European sites under various conditions. Ammar and Davies (2019) hypothesized that the AquaCrop model can operate accurately under coarser-than-daily simulations by aggregated weather input data. They performed simulations with barley data in Canada and obtained low bias errors for crop production and water estimates at daily and semi-weekly time steps, whereas weekly simulations showed poorer performance. Thus, temporal aggregation of weather data resulted in a slight overestimation of both biomass and yield.

4.6.1. Sources of error in crop growth models

There is a general consensus on the classification of error sources in crop growth models in the following four categories: model structure, model parameters, uncertainty in model inputs, uncertainty in validation data (Beven and Kirkby, 2009). Model structure errors are errors that occur when the equations that describe a process (e.g. photosynthesis, soil water movement) are essentially wrong. Despite the great number of crop growth models presented over the years, the set of equations describing the processes is relatively limited, for example RUE or leaf level assimilation for photosynthesis or tipping bucket vs Richards equation for water movement in unsaturated soils. These approaches differ in complexity, but are all well accepted by the modeling community and are unlikely to be a major source of error because of mistakes in understanding the described processes. However there can be cases where a simpler approach neglects a process, that maybe relevant for a certain environment for example the tipping bucket approach neglects water capillary rise from the water table and thus may give systematic errors where the fluctuation of the water table plays an important role, see for example the work from Hack-ten Broeke et al., (2019, 2016) who coupled an hydrological model SWAP with a crop growth model WOFOST to account for fluctuating ground water. Error in the model parameters play a major role in model accuracy and a vast body of literature exists on model calibration (see in particular the work from the Agmip calibration group, Seidel et al., 2018; Wallach et al., 2021). The uncertainty in the model inputs is often an under looked type of error in crop growth models as well as in data-driven approaches. This is particularly true when large simulations sets are run over large area (e.g. gridded simulations), the main errors usually comes from uncertainties associated with rainfall (which has typically a larger variability than the other weather input, like air temperature) and soil characteristics. More and more studies rely on soil survey for soil input to the model, in this case soil physical properties are generally the more certain and these can then be used to simulate other soil properties used pedotransfer functions. However a larger uncertainty resides in estimates of soil organic matter, which is affected largely by past soil uses. Boundary conditions (i.e. initial water and soil nutrient content) deserve a special mention as they often represent an underlooked source of error. The problem of not knowing initial soil water content is often solved by starting the simulation with a sufficient anticipation so that soil water

status gets "reset" after a large rainfall event. The initial content of nutrients (e.g. nitrogen) is a major issue, as this is often an unknown parameter that can vary largely, depending on the weather in the months preceding the sowing.

4.6.2. Key progresses and knowledge gaps in process based models

Since the first developments of crop growth models, great progress has been made on translating fundamental process knowledge on soil processes, crop physiology and canopy meteorology into these models. Much progress has been made in incorporating these models into decision support systems and yield gap analyses at various scales. Where crop growth modeling started with a few crops, nowadays many crops can be simulated, although a bias remains towards annual crops (i.e. less models can simulate growth of perennials). Knowledge gaps do remain. The number of growth limiting factors that can be simulated with crop growth models is nowadays still not much different from that in the 1970 s, limited to water, nitrogen, CO2 and irradiance. Very few models can dynamically simulate dynamics of other nutrients in soil and plant. Very few models can dynamically simulate dynamics of weeds, pests and diseases and their interactions with crop growth.

Overall the review of the literature revealed at least three major frontiers were crop growth models research is advancing: assimilation of data in crop growth models for improving within-season forecasting, integration of crop growth models and genetics, inclusion of yield reduction factors.

Data assimilation (Jin et al., 2018) uses real time improvement of yield forecasts based on assimilation of in season measurements of LAI or other state variables. While data assimilation in crop growth models has been implemented already in the 1990 s (e.g. Bouman, 1995), its adoption has been very slow. There are still very few operational platforms offering this service. The integration of crop growth models and genetics is a promising research topic where a major objective is to link genetics and crop growth models parameters beyond the calibration of individual cultivars. As reviewed by Boote et al. (2021) a number of studies focused on linking quantitative trait loci to parameters of models describing growth or phenology (e.g.Wallach et al., 2018). Such approach will be greatly beneficial in assisting breeding for new crops as it will allow the prediction of the behavior of plant material in different environment based on genetic information. Efforts on incorporating plants and diseases in crop growth models proceed slowly because of the large variety of interaction mechanisms between plants and diseases. Nonetheless advances occur specially for canopy diseases, for example Caubel et al. (2017) coupled a crop growth model to a model predicting leaf rust spore dynamics in wheat, by reducing the leaf area index of the crop as a function of the spores dynamics over the season.

4.7. Mixed process-based and data-driven approaches

Following our review of the published studies on yield prediction, using data-driven or process-based approaches, here we intend to investigate the benefits that can derive from the integration of the two approaches. In our literature survey of approaches, we found that in most cases of mixed approaches process-based models have been used to build features or indicators (Bussay et al., 2015; Kloss et al., 2012; Mavromatis, 2016, 2014). Bussay et al. (2015) used crop models-based predictors to explain observed yield detrended for technology advancement. A similar approach was used by Mavromatis (2016, 2014) who calculated several crop indicators from simulation models for different regions in Greece and predicted yield differential (i.e., the difference between a year and the previous year) using a PCA regression.

Lobell et al. (2015) created a scalable satellite-based crop yield mapper of soybean and maize in the U.S. by "using crop model simulations to train statistical models for different combinations of possible image acquisition dates". Briefly, they created a set of APSIM (a dynamic crop model) metamodels (one for each combination of observation days) that take as input weather and remote sensing data. Huang et al. (2017) used Bayesian averaging to create a model ensemble for predicting maize yield for Liaoning province (China). Kloss et al. (2012) used synthetic data generated by crop models to develop stochastic crop water production functions (SCWPFs) for arid environments. Bannayan et al. (2004) used a pattern recognition algorithm to estimate crop model parameters for new cultivars.

In data-driven approaches the model parameters are learned from the data, however, prior-knowledge can be injected in these models through feature engineering. For example, evapotranspiration can be estimated from raw weather data using the Penman-Monteith equation and used as a predictor. In an ideal situation where a sufficiently large number of records are present, a data-driven model may still reach good predictions from the raw inputs, however if the training dataset is small the training process can be facilitated by extracting relevant features using apriori knowledge. We could further discern within this approach different levels of utilization of existing knowledge: from the selection of the inputs that are deemed as relevant based on expert opinion, to extraction of feature using static formulae, such as NDVI or potential evapo-transpiration, to the use of crop growth models to compute ancillary predictors, that are used to feed the model. For example Saha et al. (2021) used a crop growth model to increment the space of the predictors used in a regression-tree to predict nitrous oxide emissions from corn and soybean fields in Michigan. Their study showed that coupling crop growth models with machine learning approach can improve the amount of explained variation from 38 % to 51 % as well as reduce the bias thanks to the fact that the model is calibrated also on the observed data. Similarly Corrales et al. (2022) used a crop growth model to calculate several state variables for a set of observed soybean trials in the South of France and then produced a metamodel where the predictors where the state variables calculated by the model.

A data-driven approach can also be integrated into a process-based models by using data-driven models used to estimate unknown parameters. For example, soil characteristics (a required parameter for most crop models) can be estimated using empirical pedotransfer functions — i.e. models that predict soil characteristics from few soil properties like texture and organic matter — fed with soil characteristics and remote-sensing data (Jana and Mohanty, 2011). For example, Ewes et al., (2020) used pedotransfer functions trained on an European soil database (Wösten et al., 1999) to derive hydraulic soil properties used in a study on the assimilation of LAI in a winter wheat crop dynamic model. Fraga et al. (2020) and Yang et al. (2020) used soil parameters derived from the 3D soil hydraulic database of Europe at 250 m resolution by (Tóth et al., 2017). In a similar way Saha et al. (2021) envision that empirical models could be used to replace subroutines in crop growth models for processes where our capacity of describing the process in physical equations is often limited, as is the case for nitrous oxide emissions.

Metamodels are an effective way to integrate data-driven approaches and crop growth models, in the next section we will propose a new way to use them to improve prediction, whereas here we briefly report how metamodels have been used in agriculture. Metamodels are models of models (Wallach et al., 2019c) and are in essence a simple version of a complex model. They are also known as summary models, model emulators, surrogate models, proxy models, low-fidelity models (Razavi et al., 2012). The main purpose of metamodels is often to reduce the computation time (at the expense of fidelity to the original model, Razavi et al., 2012). Meta models found many applications in hydrology (for a review, see Castelletti et al., 2012) and engineering (reviewed by Simpson et al., 2001). In agriculture, they have been seldomly applied because most models are point-based models and therefore are much less computationally intensive than distributed models (i.e., models having state variables defined in different locations, e.g., the temperature in different points of a lake). Nonetheless, there are examples in the literature where metamodels of crop models have been used for various purposes, for instance running gridded simulations (Folberth et al., 2019) or large sets of simulations at regional level (Xu et al., 2021), couple models (Britz and Leip, 2009; Hack-Ten Broeke et al., 2016), run

factorial simulations (Shahhosseini et al., 2019), reduce the number of parameters (Brooks et al., 2001; Rakotovololona et al., 2019), reduce the computations burden of running an ensemble models on thousands of sites (Martinez-Feria et al., 2022). Tartarini et al. (2021) used a synthetic dataset to produce a metamodel based on weather-indices to be used by index-based insurances against extreme weather events (frosts, drought). Pylianidis et al. (2021) also investigated machine learning metamodels for predicting nitrogen response rate in pastures. Despite the vast majority of crop growth models are point-based and thus not very computationally intensive, there are exceptions. For example Qu and Drummond (2018) produced a spatially explicit model for wild blueberry pollination. Because the model is a spatially explicit agent-based simulation model that simulates individual bees pollinating wildberries, its runs are computationally very intensive, with one simulation run taking 2 h to run on 1 ha. Therefore Obsie et al. (2020) produced a metamodel to reduced the computation time at the expense of model precision (metamodel r^2 0.93).

Metamodels can be classified in two categories, the data-driven metamodels and the low-fidelity models. Data-driven metamodels are constructed by running the original model several times under different boundary conditions, and then the dataset is reproduced using various predictive techniques (Simpson et al., 2001), typically, linear regressions (Britz and Leip, 2009; Makowski et al., 2015), tree-based methods (Folberth et al., 2019; Pylianidis et al., 2021; Shahhosseini et al., 2019), neural networks (Florin et al., 2011). Low-fidelity models still maintain the dynamic structure of process-based models, but some of the contained processes are simplified (e.g. Tittonell et al., 2010).

Metamodels can also be classified as dynamic or non-dynamic. Dynamic metamodels are based on the simulation and update of the state variables over time. These are common in hydrology, but are rarer in agriculture as crop models are typically point-based and thus less computation intensive. For example Brooks et al. (2001) created a simpler version of Syrius wheat dynamic model (Jamieson et al., 1998) with the purpose of increasing the understandability of the model. Van Evert et al. (2006) created a model for optimal fertilizer strategy where a complex model for soil nitrogen dynamics was replaced by a linear model where the parameter regulating the fraction of gaseous and aqueous nitrogen losses was parameterized to match simulations from the more complex model. The purpose of the simplification was to speed up the search for the optimal fertilization strategy using a linear programming optimization algorithm.

Based on the reviewed literature we can identify three major cases for the integration of data-driven and process-based models for yield prediction: usage of crop growth models to engeneer new indices and expand the predictors space that will be used in a data-driven models, use data-driven models to calculate inputs for crop growth models that are not available (e.g. soil characteristics using pedotransfer functions), metamodels mainly with the purpose of reducing computation costs.

The integration of the approaches presents risks and pitfalls next to the benefits described above. First of all the integration of crop models into data-driven approaches, for example to augment the input sets, has the pre-requisite that the crop growth models correctly simulate the variables used in the data-driven approach. It is unclear however what would be the effect of introducing uncorrectly simulated values as a predictor in a data-driven model, it is auspicable that the model optimization algorithm is not affected by state variables that do not improve the predictions.

4.7.1. *Beyond simplification, metamodels for transfer learning*

Here we would like to propose a new methodology for the integration of process-based and data-driven based on the use of metamodels and transfer learning. This builds on the previous work of Pylianidis et al., (2022) who have investigated the use of machine learning metamodels using synthetic data for developing operational digital twins. They demonstrated it in a case study about predicting nitrogen response rate in pastures, and evaluated in-silico transfer learning scenarios when both historical data are available for a location or not.

Transfer learning consists in exploiting what has been learned from one domain (the source domain) into another, affine, domain (the target domain, Goodfellow et al., 2016). Transfer learning is typically employed for classification of texts or images. Reviewing transfer learning use cases goes beyond the scope of this study, and we refer the readers to publications from Farahani et al. (2021) and Weiss et al. (2016) for a comprehensive review. A typical example of transfer learning is the transfer of a model trained to classify the sentiment (good, bad) of food textual reviews to classify the sentiment of reviews on a different type of products, for example cars. Like in other fields, also in agriculture large data repositories are becoming increasingly available, however the large repositories rarely fit exactly the problem at hand, thus the need to transfer knowledge that can be learned from other contexts. Transfer learning has been adopted also in agriculture for example Suh et al. (2018) used transfer learning to adapt a generic neural network for image classification (Alexnet) to the binary classification of the health status of sugar beet leaves.

To understand how we propose to use transfer learning for yield prediction, let us consider the imaginary case where we have been assigned the task to produce a model for regional wheat yield prediction in the Netherlands using the weather data. We would have a limited number of records as the Netherlands has 12 provinces and probably recorded data for two or three decades. We could consider using a transfer learning approach where the domain are all the grain crops in Northern Europe, therefore using a larger dataset from a related domain. In this case the target domain is our wheat yield in the Netherlands, the source domain is grain yield in Northern Europe.

Here below we propose a method where transfer learning is used not only to transfer knowledge between related agricultural domains, but also to bring the knowledge encoded in process-based models into data-driven models, using metamodels. The process that we envision is the following: a metamodel of a crop growth model using neural network is trained on a synthetic dataset (source domain), the model is then transferred to the target situation using the smaller dataset that is available for the targeted problem. For example, a general metamodel of a potato crop growth model could be built, and then such a model could be tailor-made to make predictions for the fields of a specific farm or of a different target domain, for example, seed potatoes. To implement our method we foresee the following steps: first a metamodel based on a neural network is trained on the synthetic dataset, then the part of the weights of the metamodel are frozen (i.e. set as not trainable) and the metamodel is further trained on the few observed data. This approach would bring the following benefits: thanks to the part of the metamodel that was trained on the synthetic data such metamodel is able to extract features and patterns useful for yield predictions that could not be learned just from the few observed data (similar to the many works that used crop growth models to build features and indicators, as pointed out in the previous section), moreover the model will not be biased because the last part of the network has been calibrated for the problem at hand. The produced metamodel will also be flexible as it will allow for the integration of new predictors by simply adding a new input layer to the metamodel trained solely on the synthetic data and retraining the model on the few observed data so that the metamodel becomes tailor made for the case study at hand. Let's image for example that a farmer has installed a leaf moisture sensors in the field (a state variable normally not simulated in crop growth models), we could simply add a new input layer to the neural network metamodel (leaf wetness) and train only part of the neural network-based metamodel.

Ideally a model created using transfer learning and metamodels should outperform traditional process-based models and data-driven model. When the final model does not perform better than a model trained only on the (typically few) target data, the transfer process is called negative transfer (Weiss et al., 2016). Negative transfer often results from using a weakly related source domain. Such weak relation arises from differences in the distribution in the source and target

domain. The differences can be in the distribution of the input features — in the proposed case for example if the source metamodel was trained using Dutch weather and is used to predict yield in Italy — or in the output target — in the proposed approach if the metamodel has been trained on wheat (typical yield 2–8 t/ha) and used for potatoes (typical yield 30–70 t/ha) — or in the conditional distribution of the yield over the input, i.e. $P(y = x)$, if the processes in source and target domain are very different, for example if a metamodel trained on potatoes is used to predict yield of berries. Because of the risk of negative transfer learning, it is important to use a source data derived from a domain close to the target domain. When this advice is translated into our proposed approach it means that the synthetic dataset used to produce the metamodel should reflect the conditions of the target data, thus refraining from creating generic models that are adapted to new circumstances but creating when possible new synthetic data reflecting the target conditions.

The major challenge to the application of this methodology is the lack of large, standardized data sets of observations at field level to use for the transfer learning. In fact, few farms are large and diverse enough to produce such a large datasets, therefore aggregating data from different farms using, for example data platform such as Akkerweb (Evert et al., 2018) or dacom (https://www.dacom.nl/) may represent a valid opportunity create such large sets of data.

5. Conclusion

This study showed that in a sample of 101 articles, process-based and data-driven approaches were equally represented. From the analysis of the literature, it was clear that a major obstacle to the widespread adoption of complex data-driven approaches is the lack of suitable datasets, whereas crop growth models are facing the problem of not being able to simulate reduction factors while requiring often numerous input factors. We found partial confirmation that few studies used a mix of these two approaches to predict yield and we identified three major ways to mix the two approaches: use of crop-growth models to engineer features or indices, use data-driven models to estimate lacking crop growth models input, produce meta-models to reduce computation burden. Next to describe existing examples on how the two approaches can be mixed we propose a novel approach based on metamodels and transfer learning.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Bernardo Maestrini reports that financial support was provided by Dutch Ministry of Agriculture, Nature and Food. Frits K. van Evert reports financial support was provided by Dutch Ministry of Agriculture, Nature and Food. Bernardo Maestrini, Keiji Jindo and Frits van Evert, Gordan Mimić and Sanja Brdar reports that financial support was provided by Horizon 2020 EU program. Ioannis Athanasiadis contributed to develop the idea of using metamodel and transfer learning and to discussion on the integration methodologies.

Acknowledgements

We are grateful to two anonymous reviewers for their insightful comments. This work was funded by the Ministry of Agriculture, Nature and Food quality through the project *Data driven and High Tech (KB-38)*, and by the *Antares* project (grant agreement number 739570) that funded through the Horizon 2020 EU Program.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eja.2022.126569.

References

Abedinpour, M., Sarangi, A., Rajput, T.B.S., Singh, M., Pathak, H., Ahmad, T., 2012. Performance evaluation of AquaCrop model for maize crop in a semi-arid environment. Agric. Water Manag. 110, 55–66. https://doi.org/10.1016/j.agwat.2012.04.001.

Al-Gaadi, K.A., Hassaballa, A.A., Tola, E., Kayad, A.G., Madugundu, R., Alblewi, B., Assiri, F., 2016. Prediction of potato crop yield using precision agriculture techniques. PLoS One 11, 1–16. https://doi.org/10.1371/journal.pone.0162219.

Alvarez, R., 2009. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. Eur. J. Agron. 30, 70–77. https://doi.org/10.1016/j.eja.2008.07.005.

Amarasingha, R.P.R.K., Suriyagoda, L.D.B., Marambe, B., Gaydon, D.S., Galagedara, L. W., Punyawardena, R., Silva, G.L.L.P., Nidumolu, U., Howden, M., 2015. Simulation of crop and water productivity for rice (Oryza sativa L.) using APSIM under diverse agro-climatic conditions and water management techniques in Sri Lanka. Agric. Water Manag. 160, 132–143. https://doi.org/10.1016/j.agwat.2015.07.001.

Ammar, M.E., Davies, E.G.R., 2019. On the accuracy of crop production and water requirement calculations: process-based crop modeling at daily, semi-weekly, and weekly time steps for integrated assessments. J. Environ. Manag. 238, 460–472. https://doi.org/10.1016/j.jenvman.2019.03.030.

Attia, A., Rajan, N., Xue, Q., Nair, S., Ibrahim, A., Hays, D., 2016. Application of DSSAT-CERES-wheat model to simulate winter wheat response to irrigation management in the Texas High Plains. Agric. Water Manag. 165, 50–60. https://doi.org/10.1016/j.agwat.2015.11.002.

Bannayan, M., Hoogenboom, G., Crout, N.M.J., 2004. Photothermal impact on maize performance: a simulation approach. Ecol. Modell. 180, 277–290. https://doi.org/10.1016/j.ecolmodel.2004.04.022.

Beven, K.J., Kirkby, M.J., 2009. A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 6936. https://doi.org/10.1080/02626667909491834.

Bidogeza, J.C., Hoogenboom, G., Berensten, P.B.M., de Graaff, J., Oude Lansink, A.G.J. M., 2012. Application of DSSAT crop models to generate alternative production activities under combined use of organic-inorganic nutrients in Rwanda. J. Crop Improv. 26, 346–363. https://doi.org/10.1080/15427528.2011.641140.

Boote, K., 2019. Advances in Crop Modelling for a Sustainable Agriculture, first. ed. Taylor & Francis, London. https://doi.org/10.1201/9780429266591.

Boote, K.J., Jones, J.W., Hoogenboom, G., 2021. Incorporating realistic trait physiology into crop growth models to support genetic improvement. Silico Plants 3, 1–19. https://doi.org/10.1093/insilicoplants/diab002.

Boote, K.J., Jones, J.W., White, J.W., Asseng, S., Lizaso, J.I., 2013. Putting mechanisms into crop production models. Plant Cell Environ. 36, 1658–1672. https://doi.org/10.1111/pce.12119.

Borus, D., Parsons, D., Boersma, M., Brown, H., Mohammed, C., 2018. Improving the prediction of potato productivity: APSIM-Potato model parameterization and evaluation in Tasmania, Australia. Aust. J. Crop Sci. 12, 32–43. https://doi.org/10.21475/ajcs.18.12.01.pne570.

Bose, P., Kasabov, N.K., Bruzzone, L., Hartono, R.N., 2016. Spiking Neural networks for crop yield estimation based on spatiotemporal analysis of image time series. IEEE Trans. Geosci. Remote Sens. 54, 6563–6573. https://doi.org/10.1109/TGRS.2016.2586602.

Bouman, B.A., 1995. Crop modelling and remote sensing for yield prediction. Neth. J. Agric. Sci. 43, 143–161.

Britz, W., Leip, A., 2009. Development of marginal emission factors for N losses from agricultural soils with the DNDC-CAPRI meta-model. Agric. Ecosyst. Environ. 133, 267–279. https://doi.org/10.1016/j.agee.2009.04.026.

Brooks, R.J., Semenov, M.A., Jamieson, P.D., 2001. Simplifying sirius: Sensitivity analysis and development of a meta-model for wheat yield prediction. Eur. J. Agron. 14, 43–60. https://doi.org/10.1016/S1161-0301(00)00089-7.

Burke, M., Lobell, D.B., 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. Proc. Natl. Acad. Sci. USA 114, 2189–2194. https://doi.org/10.1073/pnas.1616919114.

Bussay, A., van der Velde, M., Fumagalli, D., Seguini, L., 2015. Improving operational maize yield forecasting in Hungary. Agric. Syst. 141, 94–106. https://doi.org/10.1016/j.agsy.2015.10.001.

Castelletti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P.C., 2012. A general framework for dynamic emulation modelling in environmental problems. Environ. Model. Softw. 34, 5–18. https://doi.org/10.1016/j.envsoft.2012.01.002.

Caubel, J., Launay, M., Ripoche, D., Gouache, D., Buis, S., Huard, F., Huber, L., Brun, F., Bancal, M.O., 2017. Climate change effects on leaf rust of wheat: implementing a coupled crop-disease model in a French regional application. Eur. J. Agron. 90, 53–66. https://doi.org/10.1016/j.eja.2017.07.004.

Chen, C., Wang, E., Yu, Q., 2010. Modeling wheat and maize productivity as affected by climate variation and irrigation supply in North China plain. Agron. J. 102, 1037–1049. https://doi.org/10.2134/agronj2009.0505.

Chen, K., O'Leary, R.A., Evans, F.H., 2019. A simple and parsimonious generalised additive model for predicting wheat yield in a decision support tool. Agric. Syst. 173, 140–150. https://doi.org/10.1016/j.agsy.2019.02.009.

Chen, Y., Han, X., Si, W., Wu, Z., Chien, H., Okamoto, K., 2017. An assessment of climate change impacts on maize yields in Hebei Province of China. Sci. Total Environ. 581–582, 507–517. https://doi.org/10.1016/j.scitotenv.2016.12.158.

Conradt, T., Gornott, C., Wechsung, F., 2016. Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: enhancing the predictive skill by panel definition through cluster analysis. Agric. Meteorol. 216, 68–81. https://doi.org/10.1016/j.agrformet.2015.10.003.

Constantin, J., Raynal, H., Casellas, E., Hoffmann, H., Bindi, M., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.C., Klatt, S., Kuhnert, M., Lewan, E., Maharjan, G.R., Moriondo, M., Nendel, C., Roggero, P.P., Specka, X., Trombi, G., Villa, A., Wang, E., Weihermüller, L., Yeluripati, J., Zhao, Z., Ewert, F., Bergez, J.E., 2019. Management and spatial resolution effects on yield and water balance at regional scale in crop models. Agric. For. Meteorol. 275, 184–195. https://doi.org/10.1016/j.agrformet.2019.05.013.

Corrales, D.C., Schoving, C., Raynal, H., Debaeke, P., Journet, E.P., Constantin, J., 2022. A surrogate model based on feature selection techniques and regression learners to improve soybean yield prediction in southern France. Comput. Electron. Agric. 192. https://doi.org/10.1016/j.compag.2021.106578.

Crane-Droesch, A., 2018. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environ. Res. Lett. 13, 114003 https://doi.org/10.1088/1748-9326/aae159.

Derriere, S, Jolliffe, D., et al., 2018. Land productivity and plot size: Is measurement error driving the inverse relationship? J. Dev. Econ. 84–98. https://doi.org/10.1016/j.jdeveco.2017.10.002.

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., Savary, S., 2017. Modelling the impacts of pests and diseases on agricultural systems. Agric. Syst. 155, 213–224. https://doi.org/10.1016/j.agsy.2017.01.019.

Evert, V., Been, T., Booij, A.J., Kempenaar, C., Kessel, J.G., 2018. Akkerweb: a platform for precision farming data, science, and practice, in: : Proceedings of the Fourteenth International Conference on Precision Agriculture. International Society of Precision Agriculture, Monticello, IL.

Farahani, A., Pourshojae, B., Rasheed, K., Arabnia, H.R., 2021. A Concise Review of Transfer Learning.

Fieuzal, R., Marais Sicre, C., Baup, F., 2017. Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. Int. J. Appl. Earth Obs. Geoinf. 57, 14–23. https://doi.org/10.1016/j.jag.2016.12.011.

Fleisher, D.H., Condori, B., Quiroz, R., Alva, A., Asseng, S., Barreda, C., Bindi, M., Boote, K.J., Ferrise, R., Franke, A.C., Govindakrishnan, P.M., Harahagazwe, D., Hoogenboom, G., Naresh Kumar, S., Merante, P., Nendel, C., Olesen, J.E., Parker, P. S., Raes, D., Raymundo, R., Ruane, A.C., Stockle, C., Supit, I., Vanuytrecht, E., Wolf, J., Woli, P., 2017. A potato model intercomparison across varying climates and productivity levels. Glob. Chang. Biol. 23, 1258–1281. https://doi.org/10.1111/gcb.13411.

Florin, M.J., McBratney, A.B., Whelan, B.M., Minasny, B., 2011. Inverse meta-modelling to estimate soil available water capacity at high spatial resolution across a farm. Precis. Agric. 12, 421–438. https://doi.org/10.1007/s11119-010-9184-3.

Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N., Obersteiner, M., 2019. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. Agric. Meteorol. 264, 1–15. https://doi.org/10.1016/j.agrformet.2018.09.021.

Fraga, H., Pinto, J.G., Santos, J.A., 2020. Olive tree irrigation as a climate change adaptation measure in Alentejo, Portugal. Agric. Water Manag. 237, 106193 https://doi.org/10.1016/j.agwat.2020.106193.

Fronzek, S., Pirttioja, N., Carter, T.R., Bindi, M., Hoffmann, H., Palosuo, T., Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P., Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M.F., Dumont, B., Ewert, F., Ferrise, R., François, L., Gaiser, T., Hlavinka, P., Jacquemin, I., Kersebaum, K.C., Kollas, C., Krzyszczak, J., Lorite, I.J., Minet, J., Minguez, M.I., Montesino, M., Moriondo, M., Müller, C., Nendel, C., Öztürk, I., Perego, A., Rodríguez, A., Ruane, A. C., Ruget, F., Sanna, M., Semenov, M.A., Slawinski, C., Stratonovitch, P., Supit, I., Waha, K., Wang, E., Wu, L., Zhao, Z., Rötter, R.P., 2018. Classifying multi-model wheat yield impact response surfaces showing sensitivity to temperature and precipitation change. Agric. Syst. 159, 209–224. https://doi.org/10.1016/j.agsy.2017.08.004.

García, Á., Riaño, J., N, H., Magnitskiy, S., 2014. Simulation of corn (Zea mays L.) production in different agricultural zones of Colombia using the AquaCrop model. Agron. Colomb. 32, 358–366. https://doi.org/10.15446/agron.colomb.v32n3.45939.

García-León, D., Contreras, S., Hunink, J., 2019. Comparison of meteorological and satellite-based drought indices as yield predictors of Spanish cereals. Agric. Water Manag. 213, 388–396. https://doi.org/10.1016/j.agwat.2018.10.030.

Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. 6, 10335–10355. https://doi.org/10.3390/rs61110335.

Gilardelli, C., Confalonieri, R., Cappelli, G.A., Bellocchi, G., 2018. Sensitivity of WOFOST-based modelling solutions to crop parameters under climate change. Ecol. Modell. 368, 1–14. https://doi.org/10.1016/j.ecolmodel.2017.11.003.

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., Peng, Y., 2018. Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods 14, 1–14. https://doi.org/10.1186/s13007-018-0338-z.

Gonzalez-Sanchez, A., Frausto-Solis, J., Ojeda-Bustamante, W., 2014. Attribute selection impact on linear and nonlinear regression models for crop yield prediction. Sci. World J. 2014. https://doi.org/10.1155/2014/509429.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Representation learning. In: In: Deep Learning. MIT Press, pp. 524–554.

Hack-Ten Broeke, M.J.D., Kroes, J.G., Bartholomeus, R.P., Van Dam, J.C., De Wit, A.J.W., Supit, I., Walvoort, D.J.J., Van Bakel, P.J.T., Ruijtenberg, R., 2016. Quantification of the impact of hydrology on agricultural production as a result of too dry, too wet or too saline conditions. Soil 2, 391–402. https://doi.org/10.5194/soil-2-391-2016.

Hack-ten Broeke, M.J.D., Mulder, H.M., Bartholomeus, R.P., van Dam, J.C., Holshof, G., Hoving, I.E., Walvoort, D.J.J., Heinen, M., Kroes, J.G., van Bakel, P.J.T., Supit, I., de Wit, A.J.W., Ruijtenberg, R., 2019. Quantitative land evaluation implemented in Dutch water management. Geoderma 338, 536–545. https://doi.org/10.1016/j.geoderma.2018.11.002.

Hernandez, J., Lobos, G.A., Matus, I., del Pozo, A., Silva, P., Galleguillos, M., 2015. Using ridge regression models to estimate grain yield from field spectral data in bread wheat (Triticum aestivum L.) grown under three water regimes. Remote Sens. 7, 2109–2126. https://doi.org/10.3390/rs70202109.

Huang, X., Huang, G., Yu, C., Ni, S., Yu, L., 2017. A multiple crop model ensemble for improving broad-scale yield prediction using Bayesian model averaging. Field Crop. Res. 211, 114–124. https://doi.org/10.1016/j.fcr.2017.06.011.

Irwin, S., Good, D., 2016. Opening Up the Black Box: More on the USDA Corn Yield Forecasting Methodology. Dep. Agric. Consum. Econ. Univ. Illinois 1–8.

Jamieson, P.D., Semenov, M.A., Brooking, I.R., Francis, G.S., 1998. Sirius: a mechanistic model of wheat response to environmental variation. Eur. J. Agron. 8, 161–179. https://doi.org/10.1016/S1161-0301(98)00020-3.

Jana, R.B., Mohanty, B.P., 2011. Enhancing PTFs with remotely sensed data for multi-scale soil water retention estimation. J. Hydrol. 399, 201–211. https://doi.org/10.1016/j.jhydrol.2010.12.043.

Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.M., Gerber, J.S., Reddy, V.R., Kim, S.H., 2016. Random forests for global and regional crop yield predictions. PLoS One 11, 1–15. https://doi.org/10.1371/journal.pone.0156571.

Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J., 2018. A review of data assimilation of remote sensing and crop models. Eur. J. Agron. 92, 141–152. https://doi.org/10.1016/j.eja.2017.11.002.

Johnson, M.D., Hsieh, W.W., Cannon, A.J., Davidson, A., Bédard, F., 2016. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agric. For. Meteorol. 218–219, 74–84. https://doi.org/10.1016/j.agrformet.2015.11.003.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., Wheeler, T.R., 2016. Brief history of agricultural systems modeling. Agric. Syst. https://doi.org/10.1016/j.agsy.2016.05.014.

Kale Celik, S., MADENOGLU, S., SONMEZ, B., 2017. Evaluating AquaCrop Model for winter wheat under various irrigation conditions in Turkey. J. Agric. Sci. 24, 205–217.

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-guided data science: a new paradigm for scientific discovery from data. IEEE Trans. Knowl. Data Eng. 29, 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168.

Kern, A., Barcza, Z., Marjanović, H., Árendás, T., Fodor, N., Bónis, P., Bognár, P., Lichtenberger, J., 2018. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices. Agric. Meteorol. 260–261, 300–320. https://doi.org/10.1016/j.agrformet.2018.06.009.

Khaki, S., Wang, L., 2019. Crop yield prediction using deep neural networks. Front. Plant Sci. 10, 1–10. https://doi.org/10.3389/fpls.2019.00621.

Kim, N., Ha, K.J., Park, N.W., Cho, J., Hong, S., Lee, Y.W., 2019. A comparison between major artificial intelligence models for crop yield prediction: case study of the midwestern United States, 2006–2015. ISPRS Int. J. Geo Inf. 8. https://doi.org/10.3390/ijgi8050240.

Kimball, B.A., Boote, K.J., Hatfield, J.L., Ahuja, L.R., Stockle, C., Archontoulis, S., Baron, C., Basso, B., Bertuzzi, P., Constantin, J., Deryng, D., Dumont, B., Durand, J. L., Ewert, F., Gaiser, T., Gayler, S., Hoffmann, M.P., Jiang, Q., Kim, S.H., Lizaso, J., Moulin, S., Nendel, C., Parker, P., Palosuo, T., Priesack, E., Qi, Z., Srivastava, A., Stella, T., Tao, F., Thorp, K.R., Timlin, D., Twine, T.E., Webber, H., Willaume, M., Williams, K., 2019. Simulation of maize evapotranspiration: An inter-comparison among 29 maize models. Agric. For. Meteorol. 271, 264–284. https://doi.org/10.1016/j.agrformet.2019.02.037.

Kloss, S., Pushpalatha, R., Kamoyo, K.J., Schütze, N., 2012. Evaluation of crop models for simulating and optimizing deficit irrigation systems in arid and semi-arid countries under climate variability. Water Resour. Manag. 26, 997–1014. https://doi.org/10.1007/s11269-011-9906-y.

Kogan, F., Guo, W., Strashnaia, A., Kleshenko, A., Chub, O., Virchenko, O., 2016. Modelling and prediction of crop losses from NOAA polar-orbiting operational satellites. Geomat. Nat. Hazards Risk 7, 886–900. https://doi.org/10.1080/19475705.2015.1009178.

Kogan, F., Popova, Z., Singh, R., Alexandrova, P., 2018. Early forecasting corn yield using ground truth data and vegetation health indices in bulgaria. Bulg. J. Agric. Sci. 24, 57–67.

Kosmowski, F., Chamberlin, Ayalew, H., Sida, T., Abay, K., Craufurd, P., 2021. How accurate are yield estimates from crop cuts? Evidence from smallholder maize farms in Ethiopia. Food Policy, 102122. https://doi.org/10.1016/j.foodpol.2021.102122.

Kowalik, W., Dabrowska-Zielinska, K., Meroni, M., Raczka, T.U., de Wit, A., 2014. Yield estimation using SPOT-VEGETATION products: a case study of wheat in European countries. Int. J. Appl. Earth Obs. Geoinf. 32, 228–239. https://doi.org/10.1016/j.jag.2014.03.011.

Krupnik, T.J., Ahmed, Z.U., Timsina, J., Yasmin, S., Hossain, F., Al Mamun, A., Mridha, A.I., McDonald, A.J., 2015. Untangling crop management and environmental influences on wheat yield variability in Bangladesh: an application of non-parametric approaches. Agric. Syst. 139, 166–179. https://doi.org/10.1016/j.agsy.2015.05.007.

Kumhálová, J., Matějková, Š., 2017. Yield variability prediction by remote sensing sensors with different spatial resolution. Int. Agrophys. 31, 195–202. https://doi.org/10.1515/intag-2016-0046.

Lambert, M.J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's

cotton belt. Remote Sens. Environ. 216, 647–657. https://doi.org/10.1016/j.rse.2018.06.036.
Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., Peng, J., 2019. Toward building a transparent statistical model for improving crop yield prediction: Modeling rainfed corn in the U.S. F. Crop. Res 234, 55–65. https://doi.org/10.1016/j.fcr.2019.02.005.
Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150, 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008.
Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333. https://doi.org/10.1016/j.rse.2015.04.021.
Lopez, J.R., Erickson, J.E., Asseng, S., Bobeda, E.L., 2017. Modification of the CERES grain sorghum model to simulate optimum sweet sorghum rooting depth for rainfed production on coarse textured soils in a sub-tropical environment. Agric. Water Manag. 181, 47–55. https://doi.org/10.1016/j.agwat.2016.11.023.
Magney, T.S., Eitel, J.U.H., Huggins, D.R., Vierling, L.A., 2016. Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. Agric. Meteorol. 217, 46–60. https://doi.org/10.1016/j.agrformet.2015.11.009.
Makowski, D., Asseng, S., Ewert, F., Bassu, S., Durand, J.L., Li, T., Martre, P., Adam, M., Aggarwal, P.K., Angulo, C., Baron, C., Basso, B., Bertuzzi, P., Biernath, C., Boogaard, H., Boote, K.J., Bouman, B., Bregaglio, S., Brisson, N., Buis, S., Cammarano, D., Challinor, A.J., Confalonieri, R., Conijn, J.G., Corbeels, M., Deryng, D., De Sanctis, G., Doltra, J., Fumoto, T., Gaydon, D., Gayler, S., Goldberg, R., Grant, R.F., Grassini, P., Hatfield, J.L., Hasegawa, T., Heng, L., Hoek, S., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Jongschaap, R.E.E., Jones, J.W., Kemanian, R.A., Kersebaum, K.C., Kim, S.H., Lizaso, J., Marcaida, M., Müller, C., Nakagawa, H., Naresh Kumar, S., Nendel, C., O'Leary, G.J., Olesen, J.E., Oriol, P., Osborne, T.M., Palosuo, T., Pravia, M.V., Priesack, E., Ripoche, D., Rosenzweig, C., Ruane, A.C., Ruget, F., Sau, F., Semenov, M.A., Shcherbak, I., Singh, B., Singh, U., Soo, H.K., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tang, L., Tao, F., Teixeira, E.I., Thorburn, P., Timlin, D., Travasso, M., Rötter, R.P., Waha, K., Wallach, D., White, J.W., Wilkens, P., Williams, J.R., Wolf, J., Yin, X., Yoshida, H., Zhang, Z., Zhu, Y., 2015. A statistical analysis of three ensembles of crop model responses to temperature and CO_2 concentration. Agric. For. Meteorol. 214–215, 483–493. https://doi.org/10.1016/j.agrformet.2015.09.013.
Mann, M.L., Warner, J.M., Malik, A.S., 2019. Predicting high-magnitude, low-frequency crop losses using machine learning: an application to cereal crops in Ethiopia. Clim. Change 154, 211–227. https://doi.org/10.1007/s10584-019-02432-7.
Martinez-Feria, R.A., Basso, B., Kim, S., 2022. Boosting climate change mitigation potential of perennial lignocellulosic crops grown on marginal lands. Environ. Res. Lett. 17, 044004 https://doi.org/10.1088/1748-9326/ac541b.
Mavromatis, T., 2014. Pre-season prediction of regional rainfed wheat yield in Northern Greece with CERES-Wheat. Theor. Appl. Climatol. 117, 653–665. https://doi.org/10.1007/s00704-013-1031-9.
Mavromatis, T., 2016. Spatial resolution effects on crop yield forecasts: an application to rainfed wheat yield in north Greece with CERES-Wheat. Agric. Syst. 143, 38–48. https://doi.org/10.1016/j.agsy.2015.12.002.
Maya Gopal, P.S., Bhargavi, R., 2019. Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms. Appl. Artif. Intell. 33, 621–642. https://doi.org/10.1080/08839514.2019.1592343.
Meroni, M., Marinho, E., Sghaier, N., Verstrate, M.M., Leo, O., 2013. Remote sensing based yield estimation in a stochastic framework - case study of durum wheat in Tunisia. Remote Sens. 5, 539–557. https://doi.org/10.3390/rs5020539.
Mibulo, T., Kiggundu, N., 2018. Evaluation of FAO aquacrop model for simulating rainfed maize growth and yields in Uganda. Agronomy 8. https://doi.org/10.3390/agronomy8110238.
Morel, J., Bégué, A., Todoroff, P., Martiné, J.F., Lebourgeois, V., Petit, M., 2014. Coupling a sugarcane crop model with the remotely sensed time series of fIPAR to optimise the yield estimation. Eur. J. Agron. 61, 60–68. https://doi.org/10.1016/j.eja.2014.08.004.
Mourtzinis, S., Arriaga, F.J., Balkcom, K.S., Ortiz, B.V., 2013. Corn grain and stover yield prediction at R1 growth stage. Agron. J. 105, 1045–1050. https://doi.org/10.2134/agronj2012.0393.
Mourtzinis, S., Rowntree, S.C., Suhre, J.J., Weidenbenner, N.H., Wilson, E.W., Davis, V. M., Naeve, S.L., Casteel, S.N., Diers, B.W., Esker, P.D., Specht, J.E., Conley, S.P., 2014. The use of reflectance data for in-season soybean yield prediction. Agron. J. 106, 1159–1168. https://doi.org/10.2134/agronj13.0577.
NASS, 2017. Crop Progress. Crop Prog. Rep.
Nandram, B, Berg, E., Barbosa, W., 2014. A hierarchical Bayesian model for forecasting state-level corn yield. Environ. Ecol. Stat. 21, 507–530. https://doi.org/10.1007/s10651-013-0266-z.
Nevavuori, P., Narra, N., Lipping, T., 2019. Crop yield prediction with deep convolutional neural networks. Comput. Electron. Agric. 163, 104859 https://doi.org/10.1016/j.compag.2019.104859.
Ngwira, A.R., Aune, J.B., Thierfelder, C., 2014. DSSAT modelling of conservation agriculture maize response to climate change in Malawi. Soil Tillage Res. 143, 85–94. https://doi.org/10.1016/j.still.2014.05.003.
Obsie, E.Y., Qu, H., Drummond, F., 2020. Wild blueberry yield prediction using a combination of computer simulation and machine learning algorithms. Comput. Electron. Agric. 178, 105778 https://doi.org/10.1016/j.compag.2020.105778.
Pan, G., Sun, G.J., Li, F.M., 2009. Using QuickBird imagery and a production efficiency model to improve crop yield estimation in the semi-arid hilly Loess Plateau, China. Environ. Model. Softw. 24, 510–516. https://doi.org/10.1016/j.envsoft.2008.09.014.
Paredes, P., D'Agostino, D., Assif, M., Todorovic, M., Pereira, L.S., 2018. Assessing potato transpiration, yield and water productivity under various water regimes and planting dates using the FAO dual Kc approach. Agric. Water Manag. 195, 11–24. https://doi.org/10.1016/j.agwat.2017.09.011.
Passioura, J.B., 1996. Simulation models: science, snake oil, education, or engineering? Agron. J. 88, 690–694.
Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., Athanasiadis, I.N., 2021. Machine learning for large-scale crop yield forecasting. Agricultural Systems 187, 103016. https://doi.org/10.1016/j.agsy.2020.103016.
Paudel, D., Boogaard, H., de Wit, A., van der Velde, M., Claverie, M., Nisini, L., Janssen, S., Osinga, S., Athanasiadis, I.N., 2022. Machine learning for regional crop yield forecasting in Europe. Field Crops Research 276, 108377. https://doi.org/10.1016/j.fcr.2021.108377.
Peng, B., Guan, K., Pan, M., Li, Y., 2018. Benefits of seasonal climate prediction and satellite data for forecasting U.S. maize yield. Geophys. Res. Lett. 45, 9662–9671. https://doi.org/10.1029/2018GL079291.
Pylianidis, C., Snow, V., Holzworth, D., Bryant, J., Athanasiadis, I.N., 2021. Location-specific vs location-agnostic machine learning metamodels for predicting pasture nitrogen response rate. In: Del Bimbo, A., Cucchiara, R., Sclaroff, S., Farinella, G.M., Mei, T., Bertini, M., Escalante, H.J., Vezzani, R. (Eds.), Lecture Notes in Computer Science (Pattern Recognition. ICPR International Workshops and Challenges), 12666. Springer, pp. 96–109.
Pylianidis, C., Snow, V., Overweg, H., Osinga, S., Kean, J., Athanasiadis, I.N., 2022. Simulation-assisted machine learning for operational digital twins. Environmental Modelling & Software 148, 105274. https://doi.org/10.1016/j.envsoft.2021.105274.
Qu, H., Drummond, F., 2018. Simulation-based modeling of wild blueberry pollination. Comput. Electron. Agric. 144, 94–101. https://doi.org/10.1016/j.compag.2017.11.003.
Rahman, M.H., ur, Ahmad, A., Wajid, A., Hussain, M., Rasul, F., Ishaque, W., Islam, M.A., Shelia, V., Awais, M., Ullah, A., Wahid, A., Sultana, S.R., Saud, S., Khan, S., Fahad, S., Hussain, S., Nasim, W., 2019. Application of CSM-CROPGRO-Cotton model for cultivars and optimum planting dates: Evaluation in changing semi-arid climate. Field Crop. Res. 238, 139–152. https://doi.org/10.1016/j.fcr.2017.07.007.
Rakotovololona, L., Beaudoin, N., Ronceux, A., Venet, E., Mary, B., 2019. Driving factors of nitrate leaching in arable organic cropping systems in Northern France. Agric. Ecosyst. Environ. 272, 38–51. https://doi.org/10.1016/j.agee.2018.10.005.
Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48. https://doi.org/10.1029/2011WR011527.
Saha, D., Basso, B., Robertson, G.P., 2021. Machine learning improves predictions of agricultural nitrous oxide (N2O) emissions from intensively managed cropping systems. Environ. Res. Lett. 16. https://doi.org/10.1088/1748-9326/abd2f3.
Saruta, K., Hirai, Y., Tanaka, K., Inoue, E., Okayasu, T., Mitsuoka, M., 2013. Predictive models for yield and protein content of brown rice using support vector machine. Comput. Electron. Agric. 99, 93–100. https://doi.org/10.1016/j.compag.2013.09.003.
Schlenker, W., Roberts, M.J., 2009. Do nonlinear temperature effects indicate severe damages to US crop yields under climate change. Proc. Natl. Acad. Sci. USA 106, 15594–15598. https://doi.org/10.1073/pnas.0910618106.
Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go. Eur. J. Agron. 94, 25–35. https://doi.org/10.1016/j.eja.2018.01.006.
Shahhosseini, M., Martinez-Feria, R.A., Hu, G., Archontoulis, S.V., 2019. Maize yield and nitrate loss prediction with machine learning algorithms. Environ. Res. Lett. 14, 124026 https://doi.org/10.1088/1748-9326/ab5268.
Silva, J.V., Giller, K.E., 2021. Grand challenges for the 21st century: What crop models can and can't (yet) do. J. Agric. Sci. https://doi.org/10.1017/S0021859621000150.
Silva, J.V., Reidsma, P., Laborte, A.G., van Ittersum, M.K., 2017a. Explaining rice yields and yield gaps in Central Luzon, Philippines: an application of stochastic frontier analysis and crop modelling. Eur. J. Agron. 82, 223–241. https://doi.org/10.1016/j.eja.2016.06.017.
Silva, J.V., Reidsma, P., van Ittersum, M.K., 2017b. Yield gaps in Dutch arable farming systems: analysis at crop and crop rotation level. Agric. Syst. 158, 78–92. https://doi.org/10.1016/j.agsy.2017.06.005.
Silva, J.V., Tenreiro, T.R., Spätjens, L., Anten, N.P.R., van Ittersum, M.K., Reidsma, P., 2020. Can big data explain yield variability and water productivity in intensive cropping systems. Field Crop. Res. 255. https://doi.org/10.1016/j.fcr.2020.107828.
Simpson, T.W., Peplinski, J.D., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-based engineering design: survey and recommendations. Eng. Comput. 17, 129–150. https://doi.org/10.1007/PL00007198.
Singels, A., Paraskevopoulos, A.L., Mashabela, M.L., 2019. Farm level decision support for sugarcane irrigation management during drought. Agric. Water Manag. 222, 274–285. https://doi.org/10.1016/j.agwat.2019.05.048.
Son, N.T., Chen, C.F., Chen, C.R., Chang, L.Y., Duc, H.N., Nguyen, L.D., 2013. Prediction of rice crop yield using MODIS EVI-LAI data in the Mekong Delta, Vietnam. Int. J. Remote Sens. 34, 7275–7292. https://doi.org/10.1080/01431161.2013.818258.
Suh, H.K., IJsselmuiden, J., Hofstee, J.W., van Henten, E.J., 2018. Transfer learning for the classification of sugar beet and volunteer potato under field conditions. Biosyst. Eng. 174, 50–65. https://doi.org/10.1016/j.biosystemseng.2018.06.017.
Sylvester-Bradley, R., Clarke, S., Kindred, D., Roques, S., Berry, P., Welham, S., 2019. Variation across scales indicates that best progress in crop yields should come from farmer-centric research. In: Proceedings of the Pap. Present. Twelfth Eur. Conf. Precis. Agric. ECPA 2019, 917-924. ⟨https://doi.org/10.3920/978-90-8686-888-9_113⟩.
Tavakoli, A.R., Mahdavi Moghadam, M., Sepaskhah, A.R., 2015. Evaluation of the AquaCrop model for barley production under deficit irrigation and rainfed condition

in Iran. Agric. Water Manag. 161, 136–146. https://doi.org/10.1016/j.agwat.2015.07.020.
Tittonell, P., Corbeels, M., van Wijk, M.T., Giller, K.E., 2010. FIELD-A summary simulation model of the soil-crop system to analyse long-term resource interactions and use efficiencies at farm scale. Eur. J. Agron. 32, 10–21. https://doi.org/10.1016/j.eja.2009.05.008.
Tóth, B., Weynants, M., Pásztor, L., Hengl, T., 2017. 3D soil hydraulic database of Europe at 250 m resolution. Hydrol. Process. 31, 2662–2666. https://doi.org/10.1002/hyp.11203.
Van Evert, F.K., De Visser, P.H.B., Heinen, M., 2006. Operational optimization of organic fertilizer application in greenhouse crops. Acta Hortic. 718, 165–171. https://doi.org/10.17660/actahortic.2006.718.18.
Van Ittersum, M.K., Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crop. Res 52, 197–208. https://doi.org/10.1016/S0378-4290(97)00037-3.
van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: a systematic literature review. Comput. Electron. Agric. 177, 105709 https://doi.org/10.1016/j.compag.2020.105709.
van Oort, P.A.J., Timmermans, B.G.H., van Swaaij, A.C.P.M., 2012. Why farmers' sowing dates hardly change when temperature rises. Eur. J. Agron. 40, 102–111. https://doi.org/10.1016/j.eja.2012.02.005.
van Oort, P.A.J., Wang, G., Vos, J., Meinke, H., Li, B.G., Huang, J.K., van der Werf, W., 2016. Towards groundwater neutral cropping systems in the Alluvial Fans of the North China Plain. Agric. Water Manag. 165, 131–140. https://doi.org/10.1016/j.agwat.2015.11.005.
Wallach, D., Hwang, C., Correll, M.J., Jones, J.W., Boote, K., Hoogenboom, G., Gezan, S., Bhakta, M., Vallejos, C.E., 2018. A dynamic model with QTL covariables for predicting flowering time of common bean (Phaseolus vulgaris) genotypes. Eur. J. Agron. 101, 200–209. https://doi.org/10.1016/j.eja.2018.10.003.
Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2019a. Chapter 7 - calibration of system models. In: Wallach, D., Makowski, D., Jones, J.W., Brun, F. (Eds.), Working with Dynamic Crop Models, third ed. Academic Press, pp. 251–274. https://doi.org/10.1016/B978-0-12-811756-9.00007-1.
Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2019b. Chapter 8 - parameter estimation with Bayesian methods. In: Wallach, D., Makowski, D., Jones, J.W., Brun, F. (Eds.), Working with Dynamic Crop Models, third ed. Academic Press, pp. 275–309. https://doi.org/10.1016/B978-0-12-811756-9.00008-3.
Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2019c. Chapter 11 - metamodeling. In: Wallach, D., Makowski, D., Jones, J.W., Brun, F. (Eds.), Working with Dynamic Crop Models (Third Edition), third ed. Academic Press, pp. 407–423. https://doi.org/10.1016/B978-0-12-811756-9.00011-3.
Wallach, D., Palosuo, T., Thorburn, P., Gourdain, E., Asseng, S., Basso, B., Buis, S., Crout, N., Dibari, C., Dumont, B., Ferrise, R., Gaiser, T., Garcia, C., Gayler, S., Ghahramani, A., Hochman, Z., Hoek, S., Hoogenboom, G., Horan, H., Huang, M., Jabloun, M., Jing, Q., Justes, E., Kersebaum, K.C., Klosterhalfen, A., Launay, M., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Nariman Zadeh, H., Olesen, J.E., Poyda, A., Priesack, E., Pullens, J.W.M., Qian, B., Schütze, N., Shelia, V., Souissi, A., Specka, X., Srivastava, A.K., Stella, T., Streck, T., Trombi, G., Wallor, E., Wang, J., Weber, T.K.D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao, C., Zhu, Y., Seidel, S.J., 2021. How well do crop modeling groups predict wheat phenology, given calibration data from the target population. Eur. J. Agron. 124. https://doi.org/10.1016/j.eja.2020.126195.
Weiss, K., Khoshgoftaar, T.M., Wang, D.D., 2016. A survey of transfer learning. J. Big Data. https://doi.org/10.1186/s40537-016-0043-6.
Wit, C.T. de, 1958. Transpiration and crop yields. Verslagen van Landbouwkundige Onderzoekingen, 64.6. 2, Wageningen University.
Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90, 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3.
Xiangxiang, W., Quanjiu, W., Jun, F., Qiuping, F., 2013. Evaluation of the AquaCrop model for simulating the impact of water deficits and different irrigation regimes on the biomass and yield of winter wheat grown on China's Loess Plateau. Agric. Water Manag. 129, 95–104. https://doi.org/10.1016/j.agwat.2013.07.010.
Xu, H., Zhang, X., Ye, Z., Jiang, L., Qiu, X., Tian, Y., Zhu, Y., Cao, W., 2021. Machine learning approaches can reduce environmental data requirements for regional yield potential simulation. Eur. J. Agron. 129. https://doi.org/10.1016/j.eja.2021.126335.
Yang, C., Fraga, H., van Ieperen, W., Santos, J.A., 2020. Assessing the impacts of recent-past climatic constraints on potential wheat yield and adaptation options under Mediterranean climate in southern Portugal. Agric. Syst. 182, 102844 https://doi.org/10.1016/j.agsy.2020.102844.
You, J., Li, X., Low, M., Lobell, D., Ermon, S., 2017. Deep Gaussian process for crop yield prediction based on remote sensing data. In: Proceedings of the Thirty First AAAI Conf. Artif. Intell. AAAI, 2017, 4559–4565.
Zheng, H., Chen, L., Han, X., Zhao, X., Ma, Y., 2009. Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: The importance of phosphorus application rates under drought conditions. Agric. Ecosyst. Environ. 132, 98–105. https://doi.org/10.1016/j.agee.2009.03.004.
Zhou, X., Zheng, H.B., Xu, X.Q., He, J.Y., Ge, X.K., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2017. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS J. Photogramm. Remote Sens. 130, 246–255. https://doi.org/10.1016/j.isprsjprs.2017.05.003.
Zhu, W., Porth, L., Tan, K.S., 2019. A credibility-based yield forecasting model for crop reinsurance pricing and weather risk management. Agric. Financ. Rev. 79, 2–26. https://doi.org/10.1108/AFR-08-2017-0064.