A review of methods to evaluate crop model performance at multiple and changing spatial scales

Daniel Pasquel, Sébastien Roux, Jonathan Richetti, Davide Cammarano, Bruno Tisseyre, James Taylor

REVIEW

Check for updates

A review of methods to evaluate crop model performance at multiple and changing spatial scales

Daniel Pasquel[1] · Sébastien Roux[2] · Jonathan Richetti[3] · Davide Cammarano[4,5] · Bruno Tisseyre[1] · James A. Taylor[1]

Abstract

Crop models are useful tools because they can help understand many complex processes by simulating them. They are mainly designed at a specific spatial scale, the field. But with the new spatial data being made available in modern agriculture, they are being more and more applied at multiple and changing scales. These applications range from typically at broader scales, to perform regional or national studies, or at finer scales to develop modern site-specific management approaches. These new approaches to the application of crop models raise new questions concerning the evaluation of their performance, particularly for downscaled applications. This article first reviews the reasons why practitioners decide to spatialize crop models and the main methods they have used to do this, which questions the best place of the spatialization process in the modelling framework. A strong focus is then given to the evaluation of these spatialized crop models. Evaluation metrics, including the consideration of dedicated sensitivity indices are reviewed from the published studies. Using a simple example of a spatialized crop model being used to define management zones in precision viticulture, it is shown that classical model evaluation involving aspatial indices (e.g. the RMSE) is not sufficient to characterize the model performance in this context. A focus is made at the end of the review on potentialities that a complementary evaluation could bring in a precision agriculture context.

Keywords Spatialization · Scaling methods · Crop model uncertainty · Sensitivity analysis · Spatial pattern

✉ Daniel Pasquel
    daniel.pasquel@inrae.fr

1   ITAP, Univ. Montpellier, Institut Agro, INRAE, Montpellier, France

2   MISTEA, Univ. Montpellier, Institut Agro, INRAE, Montpellier, France

3   CSIRO, Floreat, WA, Australia

4   Department of Agronomy, Purdue University, West Lafayette, IN, USA

5   Department of Agroecology, Aarhus University, Tjele, Denmark

🍂 Springer

Introduction

In many scientific domains, including agronomy, environmental sciences and hydrology, models are a way to simplify reality through a series of assumptions and by representing processes (Bouman et al., 1996; Sinclair & Seligman, 1996; van Ittersum & Donatelli, 2003). These models are often necessary to answer a specific question and are designed around this objective. Models come in various forms. Statistical models (also called empirical models) use a mathematical relation between different variables. The principal drawback of statistical models is that they are designed on observed data and are ill suited for use in sites or applications that were not involved in the model parameterization and development. They also cannot predict values in an uncertain context (e.g. impact of climate change on crop growth) (Jones et al., 2017). In contrast, purely mechanistic models (also called process-based models) rely on the modelling of biophysical processes. They are based on mathematical equations that describe physiological processes (and not on mathematical equations that simply link two variables, as in statistical modelling). They can be derived in the absence of any real data as long as the process has been described. Mechanistic models can be deterministic, i.e. without random variations within the model and equations, so that for a given set of inputs, the result will always be the same. Mechanistic models can also be stochastic, i.e. the models and equations include random effects, so that results will change between simulations even if the inputs remain constant. Most of the crop models are a combination between process-based and empirical models, resulting in mechanistic deterministic models.

Models are useful tools in agro-environmental fields because they can help understand many complex processes by simulating them. Indeed, models can be used as a surrogate to estimate data that are hard, expensive or cumbersome to measure. Models account for relationships between crop growth and environmental, management and genetic factors. Therefore, there is a huge interest in using crop modelling to see how crop growth is impacted by these factors or to quantify ecosystem services. In other words, crop models are system-based models that aim to simulate interactions between the "soil–plant-atmosphere-management" (Hoogenboom, 2000; Wallach et al., 2019). To achieve this, multidisciplinary approaches are needed and crop models can take into account biological, physiological, ecological, physical or economical components. Integrating these approaches in crop modelling has led to the development of large crop models, such as APSIM (Holzworth et al., 2014), DSSAT (Boote et al., 2019), STICS (Brisson et al., 2003), WOFOST (de Wit et al., 2019), CropSyst (Stöckle et al., 2003) or AquaCrop (Steduto et al., 2009). Crop models are explanatory tools that are typically used in scenarios testing. For example, Asseng et al. (2018) used an ensemble of crop models to understand the climate change impact and adaptation for wheat protein on a global level.

While crop modelling has become common within agricultural research domains for long-term strategic applications, it has traditionally been poorly used in shorter-term (single-season) production contexts (Asseng et al., 2013; Cammarano et al., 2020). This is changing. Modern agriculture has increasing access to data, including spatial data, which is providing increased possibilities to model agricultural systems, particularly using statistical modeling coupled to machine-learning approaches. It also provides an opportunity to integrate these data into conventional crop modelling platforms and to change the way these 'traditional' crop models can be used. One of the main ways that this is occurring is via the spatialization of crop models.

Most existing crop models are "point-based models" (Heuvelink et al., 2010). Spatialization is a way to apply these point-based models spatially across an area, by taking advantage of the new data available and applying these models to new scenarios without fundamentally changing the underlying model. Spatialization of crop models is of interest to the agricultural community as predictive crop modelling, particularly short to medium term predictions at field or subfield scales, is becoming an important part of modern site-specific management approaches. This is shifting the use of these 'traditional' crop models from long-term strategic applications, such as understanding long-term crop production potential under a changing climate, to short-term tactical applications and spatial applications. Examples of short-term tactical applications would be the determination of local fertilizer requirements given within-season production potential or generating in-season production estimates across local, regional or national scales to inform food security policy and actions.

The concept of model spatialization is not new. Faivre et al. (2004) defined spatializing a crop model as "using a crop model over areas larger than those over which it was developed". At this stage, the idea was to upscale crop models from field/farm level modelling to regional level modelling. With a push toward precision agriculture, the concept of spatialization has evolved and became less restrictive. Within this review, spatialization is more simply defined as "using a crop model on another scale than which it was initially designed". Thus, it could be applied at a larger scale or a smaller scale. This review will be focused on crop models but it is noted that the concepts developed could be applicable to any environmental models or other models in general.

Finally, the difference between spatializing a crop model and a spatial crop model is important. Point-based models do not take into account neighboring data or effects to compute a result at a point (or unit support) (Heuvelink et al., 2010). So with spatialized crop models, each point, regardless of its spatial footprint, is an independent simulation. An alternative would be to create crop models that do take into account spatial interactions between the unit supports to compute their results. These would be considered 'true' spatial crop models. However, this would require a fundamental change in the underlying crop model equations to achieve this and a considerable effort from the crop modelling community. Given the investment that has been made in current crop modelling platforms, short-term development seems better suited to spatializing crop models rather than redeveloping spatial models. Consequently, this review will focus on model evaluation with an emphasis on spatialized crop models, although it is recognized that some aspects in this review will be equally relevant to spatial crop model evaluation.

In the context of crop model use, shifting from strategic to tactical applications, model spatialization is expected to increase among agro-environmental models. Therefore, the purpose of this article is twofold: (i) to present an overview of different ways to spatialize a crop model and characterize more precisely spatialization methods and (ii) to review current ways that the outputs from these spatialized model are being evaluated and should be evaluated going forward. The article concludes with a comment on how these emerging spatialized crop models can be used in precision agriculture.

Issues and methods of model spatialization process

Model spatialization: goals and reasons

Crop model spatial footprint

Crop models were designed to understand and explain plant biophysical processes and developed with an assumption which considers a homogeneous unit support, i.e. same weather, soil and management in the simulated area; so they are point simulations or point-based models (Heuvelink et al., 2010). Crop models were also initially designed to operate at the field scale (van Ittersum & Donatelli, 2003). A common feature of crop models is that they have often been designed for a specific scale, and this scale refers to the scale of the processes that the models seek to predict. Nevertheless, Sinclair and Seligman (1996) highlight that processes described by models need to be described at a finer scale than the scale at which the models simulate outputs. Some models, such as DSSAT or STICS, were developed as a point represented by a small unit of support, for instance a homogeneous plot of one m^2, and then scaled to the field scale, and model the crop as a single entity grown within a field with homogeneous production conditions (Faivre et al., 2004). Others, e.g. the MAPP potato model (MacKerron et al., 2004), are based on small pot trials and simulate an individual plant, which is then grown in standard conditions at all points in the field. Note that this is not a spatialization of the model as all model inputs and parameters are kept spatially constant. Regardless of whether it is modelled at the individual plant or the individual field/plot level, the observation implies that the crop model was designed to the scale of the specific object or process of study (e.g. leaf, plant, plot, field, watershed, region, etc.). For this review, the term spatial footprint of the model is defined as the scale of model outputs and conditions model inputs (Fig. 1). Typically, the scale is the same for model inputs as for model outputs. Thus, to run a crop model, inputs need to correspond to the model spatial footprint and outputs will be obtained at the model spatial footprint scale. Therefore, if users need to have a different scale in output than the model spatial footprint, they will need to do some modifications, i.e. by spatialization (Ginaldi et al., 2019).

For this review, the spatialization process will not consider changes of scale to the molecular level, although the importance of the intersection of the 'omics' and the crop

Fig. 1 Notion of crop model spatial footprint. The spatial footprint is illustrated here with a crop model designed at the field scale. Model outputs in the native form are at the field scale and then spatialization (red arrows) can be used to change the scale of the outputs (Color figure online)
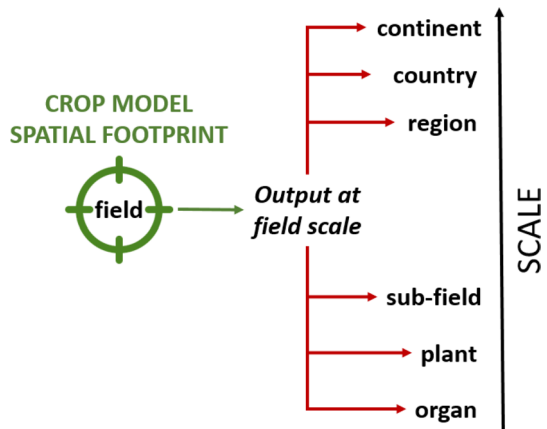
modelling community, especially using advances in high-throughput phenology platforms is noted. However, this intersection is more focussed on advancing genetic improvements, rather than for crop management applications.

Reasons for spatialization presented in the agro-environmental literature

Some users choose to spatialize crop models for specific purposes in order to obtain results that were not possible by using the crop model in its native spatial footprint. These reasons can be diverse but can be grouped together into several classes depending on the intended application (Fig. 2).

Site-specific crop management applications Users may aim to shift from strategic to tactical use of crop models with a desire to inform short-term management at finer spatial scales (site-specifically within fields). Thus, this refers to the use of crop models for precision agriculture purposes by aiming to have differential management across the field (Basso et al., 2001, 2011; Cammarano et al., 2021; Chen et al., 2017). For example, sensors are commonly used to provide variable rate applications of nitrogen (Colaço & Bramley, 2018), however, a spatialized crop model could be the main driver for the variable map or to understand spatial variability of soil–plant interactions.

Reveal and understand spatial heterogeneity Models are often constructed to improve understanding of crop development, however, model phenomena and processes, and thus
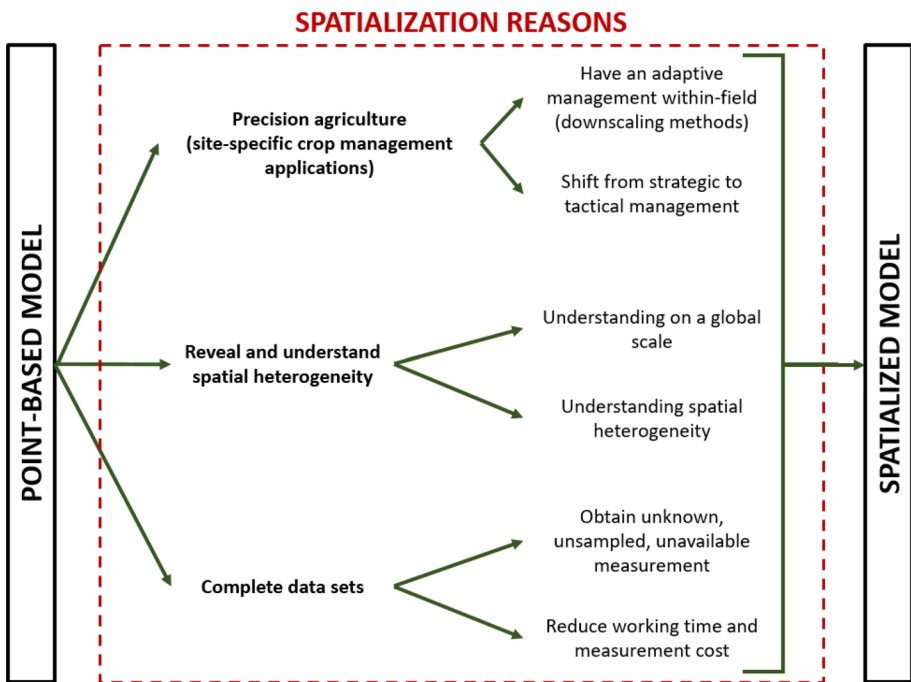

Fig. 2 Schematic illustrating the main reasons why users of point-based crop and environment models have decide to spatialize their models

the understanding, are usually restricted to a specific scale (Balkovič et al., 2013). Nevertheless, users could wish to have an understanding on a finer scale of a phenomenon that is simulated over a relatively large area, e.g. with climate change simulations (Huard et al., 2019) and/or to characterize the local spatial heterogeneity (Li et al., 2020) by downscaling a variable that was originally too coarse. Therefore, model spatialization is proving to be useful for developing an understanding of processes at different spatial scales using both upscaling and downscaling methods (Blanchoud et al., 2020; Domínguez-Álvarez et al., 2021).

Complete data sets As well as an improved understanding of processes, models can be used to predict unknown or unsampled points within a population or area. Spatialization can provide more accurate model simulations using spatially varying inputs within a known domain. When models have been calibrated and evaluated, their outputs can be used instead of real observations, thus model spatialization may be desirable to reduce the working time and cost of obtaining measurements in the field (Acevedo-Opazo et al., 2008, 2010; Baralon et al., 2012). Models can be used to obtain difficult, infeasible or unavailable measurements (Constantin et al., 2019).
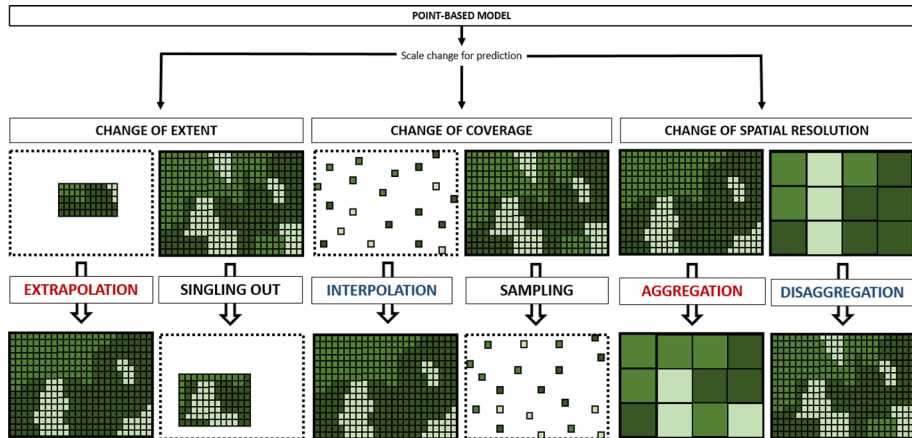
Methods used for crop model spatialization

In order to spatialize crop models, different methods have been applied according to the practitioner's objectives. These methods can relate to either or both the inputs and outputs of the models and will lead to a change of scale in the model input or output variables via the use of scaling methods at one particular point in the modelling process or, alternatively, scaling methods can be used in succession within a modified framework of crop model spatialization. An example of this would be the successive variable transformation of model inputs (e.g. to calculate an unknown model variable with a known/measured variable) then a change of scale of the model input/output variables.

Change of scale of model outputs

This is the simplest method of spatialization, whereby the model is run in its native form, without any changes to the inputs, model equations or the form of the outputs (i.e. there is no change of input/output). Once the output has been computed the scaling is achieved via spatial processing (e.g. geostatistical operators) only. The scaling methods (Fig. 3) for model spatialization can be classified into different categories depending on whether they increase or decrease the resolution: upscaling and downscaling methods (Blöschl, 2005; Ewert et al., 2011; Faivre et al., 2004). The aim of downscaling methods is to increase the variable resolution over a given area. Upscaling methods have the opposite goal, they generate a coarser resolution of the variables. Different approaches to up/downscaling have different consequences on the data and may lead to a change of extent, change of coverage or change of spatial resolution.

Change of extent Extrapolation is used for this purpose and aims to give a prediction on a wider area (e.g. farm, regional, national, etc.) than the inputs. Predictions are made into areas outside the spatial coverage of the original observations, i.e. the extent becomes

Fig. 3 Illustration of scaling methods used for model spatialization. Red processes refer to upscaling methods and blue processes refer to downscaling methods that use a spatial process. Black processes refer to a change of scale via direct extraction without using a spatial process (Color figure online). Adapted from Faivre et al. (2004)

larger (Acevedo-Opazo et al., 2010; Baralon et al., 2012; Roux et al., 2019), but the quality of prediction may be uncertain. The inverse process is termed 'singling out' to reduce the extent of the observations. This is a simple extraction process and the data quality is equivalent to the original observation(s).

Change of coverage Interpolation is used for this purpose and aims to provide estimates at locations where input variables are not available. Interpolation is performed over the entire area between known locations, for example by inverse distance weighting (IDW), kriging, spline functions or modern machine learning techniques. Reducing the coverage, or sub-setting the data, is performed using sampling approaches.

Change of spatial resolution Aggregation aims to give a coarse prediction scale of an event or a phenomenon, for example, by averaging the finer scale data to the desired coarser scale. Disaggregation is the opposite, obtaining a finer prediction scale of an event or a phenomenon that of the basic model pixel, it can be achieved by simply resampling the coarser data, such that a 10×10 m pixel could be disaggregated into 100 pixels of 1 m^2 with the same value, or by trying to differentially partition values spatially across the finer scale grid using some form of disaggregation model (Malone et al., 2013).

The lack of fine resolution data for some inputs or low computational capacity relative to the large quantity of fine resolution data available for other inputs are reasons that have led to the use of upscaling in many studies (Grosz et al., 2017). Data aggregation can be useful or even necessary in order to simplify the understanding of the processes represented and to be able to draw applicable conclusions (Jankowski et al., 2001). Some issues are related to the spatialization of models when moving from a local scale to a more global scale, in particular when using aggregation. This raises questions of whether or not to use averaged data, in order to try to quantify heterogeneity, or to keep and use very fine resolution data (Allain et al., 2018). For instance, the oversimplification of the considered process is cited as critical to the use of aggregation and

aggregated data (Scholes et al., 2013), but over-simplification may not be suitable for the intended model use.

Change of scale can be used to spatialize input or/and output data of a crop model. Thus, a crop model can be spatialized by running independent simulations in each unit (or pixel) of the desired area using spatial inputs or based on coarse independent simulations, after which the outputs undergo a change in scale.

Spatial alterations of the crop modelling framework

The methods outlined in Fig. 3 aim to manipulate the output data but could equally be applied to scale model inputs before running the model. This generates, as highlighted by Ewert et al. (2011), other potential methods to manipulate the models to achieve a spatialized crop models. These model manipulation methods correspond for instance to the modification of model parameters, the simplification of model structures or the use of nested models. Note that the processes in Fig. 3 are relevant to scaling inputs or outputs into a spatialized (crop) model. It is also potentially relevant even if the crop model framework has already been spatialized or a spatial crop model is being used.

In the case of input modification, by rescaling the model inputs to take advantage of modern sensing technologies, such as satellite imagery, spatially explicit model input data can be generated. This then generates the question of how and when these spatially explicit data can be incorporated into the model and at what moment the spatialization process takes place in the modelling framework. Given the diversity of crop model types and approaches and the diversity in the type and availability of spatially explicit model input data, it is not surprising that in a very short time there have been various methods of crop model spatialization proposed. The variety of methods include, for instance, studies about vine water status at different scales (Acevedo-Opazo et al., 2010; Baralon et al., 2012), adaption of wheat in a global warming context on a global scale (Asseng et al., 2018) or yield prediction at differing scales (Battude et al., 2016; Claverie et al., 2012). Figure 4 presents different crop model spatialization methodologies that have been synthesized from studies that have aimed to spatialize crop models.

In Fig. 4, the horizontal bars indicate the stages of the crop modelling process, from the collection of available data (top) to final model outputs (bottom), while the vertical arrows show potential pathways for modelling and the red arrows specifically indicate points where spatialization can occur. Figure 4 is constructed to indicate typical pathways for model spatialization. These include (a) a change of scale of model outputs and (b) a change of scale of the model inputs. Choosing whether to change the scale of model inputs or outputs or both is important because these methods will not have the same impact (Al-Shammari et al., 2021).

The available spatial variables can be either variables measured in the field, calculated data or output data from an upstream model (Fig. 3i). However, these available spatial variables (Fig. 4ii) may not necessarily be the same as the native variables (data) used as input by the original crop model (Fig. 4iv). In some cases, these spatial data/variables are the same and can be directly used in the model. However, in a majority of cases, sensing and modelling systems do not directly measure the correct model variable at the correct spatial resolution to be usable by the crop model. Therefore, variable transformation may be required to obtain the correct variable to run the model (Fig. 4iii). For example, many crop models use Leaf Area Index (LAI) but canopy sensors usually return a surrogate of LAI, such as a vegetation index (VI), at very high spatial resolutions. These available VI data
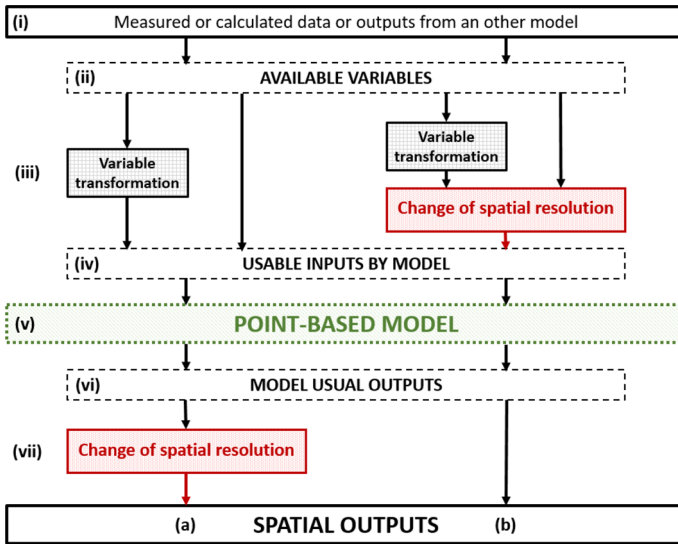
Fig. 4 Schematic illustration of pathways, from data collection to final output, to apply spatialization processes to classical point-based crop models to obtain spatial model outputs. Common pathways (a and b) in the literature are indicated and represent the main spatialization framework, but other methods can be used inside these pathways. Red boxes and red arrows correspond to the moment where spatialization really occurs in the pathway. Change of spatial resolution refers to methods that change the data resolution by processes described in Fig. 3 (extrapolation, interpolation, aggregation, disaggregation). Variable transformation refers to ancillary data being converted into model input variables. Black arrows correspond to simple transfers of data without changing data (Color figure online)

may be subject to a mathematical relationship, for instance a transfer function, to modify and transform the VI values into LAI values for the considered crop. In the example of pathway (b), available variables may be subject to a change of resolution before being used in the model. Any approaches outlined in Fig. 3 can be applied to the input data to change the support, coverage or extent of the model, leading to a change of spatial resolution of inputs (Teixeira et al., 2017) (Fig. 4iii). Often this change of scale is possible by using other ancillary data; the available data can be coupled with these ancillary data, such as high-resolution remotely-sensed imagery, to try to reduce the uncertainty in the available spatial input data (Kasampalis et al., 2018).

Once the scale of usable model inputs has been correctly adjusted (Fig. 4iv), the spatialized crop model can be run (Fig. 4v) and the usual model outputs are computed (Fig. 4vi). To reiterate, this is not a spatial modelling approach, but a punctual crop model applied at a different spatial resolution than its native design. The obtained output(s) (Fig. 4vi) may not necessarily be at the scale desired by the user, as indicated in pathway (a). Thus, the output data may also be subject to a change of spatial resolution. The approaches outlined in Fig. 3 can be applied to the output data to change the support, coverage or extent of the model output. Some studies have compared strategies to aggregating input or output data and have highlighted only a few differences between these strategies (Angulo et al., 2013b; van Bussel et al., 2011).

Pathways (a) and (b) described in Fig. 4 are the typical and shorter frameworks to spatialized predicted variables from crop models. They present simple versions where the

change in spatial resolution is done either after (a) or before (b) the modelling step. A third simple pathway (not shown), could also be considered where changes in spatial resolution occur both before and after modelling, i.e. a cross-over approach in Fig. 4 between pathways (a) and (b).

These are the simplest representations of model spatialization and more complex approaches are possible by adding other methods inside these pathways, because other data modifications may need to be done to obtain spatialized variables. For instance, in some studies, there is a resolution gap between different types of available data, especially weather-based inputs and other crop model inputs. Weather data is often designated at low resolutions (between 10 and 200 km), while crop and environmental inputs are designated at the field scale (Challinor et al., 2009). To bridge this gap, authors usually use scale transfer models, i.e. a corrective model on outputs from an upstream model, that involves upstream outputs that are generally too coarse in regards to the study area, to debiase them using local variables from the area (Choukri et al., 2020; Huard et al., 2019).

An important point is how to perform model calibration when using a model in a spatialized context. The necessity for calibration is well known to improve crop model predictions, especially for variables estimated over large areas (Jagtap & Jones, 2002). Crop model calibration involves several key steps to improve predictions, whereas, practitioners do not necessarily have the same approaches (Seidel et al., 2018). Calibration requires a large amount of data when crop models are used on a large scale, but these data are often difficult to obtain at this scale. To tackle this issue of spatialized calibration, Angulo et al. (2013a) tried three calibration strategies in an attempt to calibrate a crop model at the continental scale. Defining region-specific crop growth and phenology parameters, without considering output correction, improved the accuracy of crop model predictions on a large scale and seemed to be the best calibration strategy (Angulo et al., 2013a).

While this review is focused on methods to assess model spatialization, it is important to note the growing importance of data assimilation in the development of spatialized crop models (Jin et al., 2018). Data assimilation can be a method used inside the pathways (a) and (b). It is an approach used to recalibrate or to update a model to generate good short-term predictions. It is typically used for weather modelling, but is equally applicable to the shifting of strategic crop models to short-term tactical applications. To date, data assimilation has been mainly used for upscaling crop models to regional (Battude et al., 2016; Claverie et al., 2012) or national scales (for examples see Jin et al., 2018), but there is a growing interest in downscaling applications.

Uncertainty and error propagation when spatializing models

The uncertainty of a spatialized model will be a combined result of the model errors and the scaling errors, i.e. the uncertainty of the model itself plus the uncertainty of the scaled data plus the uncertainty of the spatialization method itself. Model uncertainty itself refers to parameter values and equations and will not be presented in this review.

Scaling errors

Scaling errors are linked to the methods used to scale model inputs and outputs. In some cases, a succession of scaling methods may be used and their combination will lead to an accumulation of uncertainty in the final result, which is often difficult to quantify (Ewert et al., 2011). The data aggregation effect (DAE) is a subject widely discussed in

upscaling studies (Zhao et al., 2015) and this effect is linked to uncertainties introduced with the methods used to achieve aggregation. Many studies have focused on weather DAE (Ewert et al., 2011; Hoffmann et al., 2015, 2016; van Bussel et al., 2011; Zhao et al., 2015) because weather is an important driver in crop modelling and often observed at a large scale, whereas crop models are designed at the field scale. Therefore, using these large scale data as an input in crop models could raise questions about the consequences of changing scale. Zhao et al. (2015) showed that DAE on weather inputs increased at coarser resolutions and was stronger with a higher spatial heterogeneity. However, some studies have shown that weather DAEs on crop yield and development are low (Ewert et al., 2011; Hoffmann et al., 2015). Apart from weather, many studies have also focused on soil data. For instance, Grosz et al. (2017) showed that for soil organic content (SOC), DAE smoothed extreme values. As model inputs were aggregated to higher scales (from 10 to 100 km), the amount of heterogeneity in the model output(s) decreased. However, Grosz et al. (2017) demonstrated that aggregating to 50 km resulted in a higher variability than the reference aggregation (the computational scale was 1 km). Hoffmann et al. (2016) showed that, with soil data aggregation only, the bias of yield prediction was below 15%. However when weather data were aggregated in addition to these soil data, the bias increased (Hoffmann et al., 2016). This shows that the scaling may have a significant effect and can be complex because some variables can be overestimated at certain scales. Thus, a comparison between different aggregation scales could be a good approach to crop model evaluation (Al-Shammari et al., 2021).

An important issue to consider is how scaling errors will vary if scaling methods are applied on model inputs or outputs (Ewert et al., 2011). For instance, DAE on input data can be reduced using a coarse output resolution or aggregating model outputs, whereas these methods can only lead to a low reduction of model structure error, i.e. regrouping of model parameters and model equations (Grosz et al., 2017). If model manipulation (e.g. modifying model parameters, simplifying model structures or using nested models) is used to spatialize a model, using downscaling methods to match between the scale of upstream model outputs and the scale of downstream model inputs can increase the quality of downstream model outputs (Cammarano et al., 2017). In reality, scaling error impacts are a trade-off between a resolution fine enough to represent the spatial variability and an acceptable computational time (Grosz et al., 2017; Zhao et al., 2015). In some cases, spatial aggregation can reduce errors from deficient input data or model structure (Heuvelink, 2002).

Reduce model uncertainty by multi-model ensembles

Using multi-model ensembles (MMEs) rather than just one model is a quite new approach in crop modelling and has been enabled due to international cooperative modelling programs (Wallach et al., 2018, 2019). The more models there are, the more the prediction error decreases (Wallach et al., 2018). Studies using crop MMEs have shown that using indicators, such as the ensemble mean (e-mean) and ensemble median (e-median) of simulated data, produces better estimates than the use of indicators from a single crop model, even if it is the best available model (Martre et al., 2015; Wallach et al., 2018). These MMEs allow an increased accuracy of crop growth simulations (Martre et al., 2015) and so are useful to reduce the uncertainty introduced by error propagation. Improving the models used in MMEs, for example by

re-calibration or incorporating or modifying simulated variables, can lead to a reduced number of models in these ensembles while simultaneously reducing uncertainty in the MME outputs (Maiorano et al., 2017). Examples of the successful use of small MMEs in agriculture include yield prediction and greenhouse gas emissions at the field scale (Ehrhardt et al., 2018).

Evaluation of model performances

Aim and importance of evaluating model performances

Model evaluation refers to the question of knowing how well model predictions are relevant to real observations, with the aim to ascertain the value computed by models. Moreover, this evaluation has to match with the proposed use of the model (Wallach et al., 2019). Model performances are case-dependent, so it is necessary to define at the beginning the model purpose (Bennett et al., 2013). The concepts of crop model 'evaluation' and 'validation' are slightly different. Validation refers to the process of determining if the model is adequate for its intended purpose or not and refers to the processes involved within the models (Tedeschi, 2006; Wallach et al., 2019). As argued by Wallach et al., (2019), crop models are never fully valid because they will always describe real world processes with assumptions and simplifications and thus are not identical to the real processes. Model evaluation is a black box concept and is not a question about the processes within the model but about the relevance of the model output (Wallach et al., 2019). Model inputs are subject to sources of uncertainty, such as measurement errors and inappropriate sampling resolutions (Crosetto et al., 2000). It is possible to consider parameter estimation when model evaluation is carried out; however, these are out of the scope of this review, and the focus in this review will be on output evaluation. Regarding output, evaluation can be performed qualitatively using graphs or quantitatively using indicators. Uncertainty and sensitivity analysis are part of the process of model evaluation (Wallach et al., 2019). These analyses aim to understand how variations in the output can be explained by variability in the model inputs.

To understand model evaluation, it is important to know what should be evaluated. To illustrate this, let's take an example of irrigation decision-making using a water stress model compared to a threshold defined outside the model. The model simulates plant water stress and can have a wide uncertainty. However, this uncertainty will not ultimately change the final decision, which is to irrigate or not, because the decision will depend on the model output relative to a threshold value identified by the decision-maker. So how should the performance of this model be evaluated? Only on the out-puts from the predictive model? Or on the whole process that culminates with the final decision-making, which is ultimately the real action that is of interest to the agronomic community? These questions highlight that the method of evaluation has to match with the use of the model, i.e. if the model is used to estimate a variable then it is the variable that needs to be evaluated, but if the model is used to make a decision then it is the decision that needs to be evaluated, and not just the variable that was taken into account for the decision-making. This question is an important one when discussing

model evaluation but it will not be detailed or discussed further in this review; indeed, this review is focused on evaluation of model output values.

Evaluation of methods used for evaluating spatialized crop model performance

Evaluation based on comparisons between observed and simulated data applied to crop models

The most common practice to evaluate a model is to compare observed data versus simulated data (outputs) using a metric or indicator that measures the distance between these observed and simulated data (Wallach et al., 2019). Various metrics exist for models in general, and many of these have been transferred for use in evaluating spatialialized crop models. Table 1 reviews common metrics that have been reported in the spatialized crop model literature to date. In all these cases, model evaluation has been performed aspatially. There have been no spatial characteristics taken into account for model evaluation even though the crop models were being used in a spatialized context. This simplified utilization of aspatial indicators may affect the evaluation of spatialized crop models. Although multiple indicators are shown in Table 1, the indicators are not equally used. The RMSE was the most frequently used indicator in published studies.

Uncertainty and sensitivity analysis methods applied to crop models

Uncertainty analysis is used to quantify the global uncertainty in the model outputs in comparison to the uncertainty in model inputs (Crosetto et al., 2000). Sensitivity analysis (SA) is used to study how the model output variations can be assigned to different sources of input variations and how the model depends on its inputs (Crosetto et al., 2000). There are different ways of varying inputs: inputs can vary around a reference value, termed a local sensitivity analysis (LSA); or inputs can vary through and across a whole feasible domain, which is called a global sensitivity analysis (GSA) (Pianosi et al., 2016). SA can be used for different purposes, for instance evaluating the consistency of the model behavior or evaluating the robustness of model outputs depending on input uncertainty and model hypothesis (Pianosi et al., 2016). Thus, SA can be used as a form of model evaluation in various ways. For instance, SA can estimate if an input's impact on the model output is acceptable. It can also identify the key inputs with the most influence on the output(s) and can prompt users to consider if there is enough knowledge about these inputs to make a considered decision (Wallach et al., 2019). SA has been used in crop modelling studies in order to have a better understanding of uncertainty propagation and to determine impacts on simulated outputs, (Acevedo-Opazo et al., 2010; Adam et al., 2011; Asseng et al., 2013; Baralon et al., 2012; Beaudoin et al., 2018; Duchemin et al., 2008; Teixeira et al., 2017). However, none of the reviewed literature on crop modelling has considered if there was a spatial component to the SA. Some questions arise from this observation such as: Is LSA suitable for assessing spatial effects? Should GSA be avoided in all situations?

Some methods accounting for spatial characteristics have been used with environmental models. However, GSA was created to explain uncertainty in scalar outputs by variations of scalar inputs and so cannot directly be used with spatial models (Saint-Geours et al., 2012). To generalize GSA methods on spatial models, Saint-Geours et al. (2012) defined two sensitivity indices (SI): one on scalar inputs (i.e. a constant over the extent) and one on

Table 1 Indicators used to measure the distance between observed and simulated data

	Indicator measuring distance between simulated and observed data		
Acronym	Description	Equation	References
Bias	–	$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$	(Battude et al., 2016; Therond et al., 2011)
MAE	Mean absolute error	$\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$	Adam et al. (2011)
MSE	Mean square error	$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	(Therond et al., 2011; Wallach et al., 2018; Zhao et al., 2016)
RMSE	Root mean square error	$\sqrt{MSE}$	(Angulo et al., 2013b; Balkovič et al., 2013; Basso et al., 2001, 2011; Battude et al., 2016; Beaudoin et al., 2018; Cammarano et al., 2021; Claverie et al., 2012; Duchemin et al., 2008; Wallach et al., 2018; Wang et al., 2017)
R^2	Determination coefficients	$1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{i,fitted})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$	(Acevedo-Opazo et al., 2010; Angulo et al., 2013b; Baralon et al., 2012; Basso et al., 2011; Battude et al., 2016; Duchemin et al., 2008)
SEC	Standard error of calibration	$\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij}-\hat{y}_{ij})^2}{(n\times m-1)}}$	(Acevedo-Opazo et al., 2010; Baralon et al., 2012)
SEP_j	Standard error of prediction (defined on spatial and temporal aspects)	$\sqrt{\frac{\sum_{i=1}^{n}(y_{ij}-\hat{y}_{ij})^2}{(n\times m)}}$	(Acevedo-Opazo et al., 2010; Baralon et al., 2012)
NSE	Nash–Sutcliffe efficiency	$1 - \frac{MSE}{MSE_{\bar{y}}}$	(Balkovič et al., 2013; Beaudoin et al., 2018; Duchemin et al., 2008; Therond et al., 2011; Wallach et al., 2018; Wang et al., 2017)
D-index	Willmott index of agreement	$1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(|\hat{y}_i-\bar{y}|+|y_i-\bar{y}|)^2}$	Cammarano et al. (2021)

y_i is the observed value, $\hat{y}_i$ is the corresponding simulated value, $\hat{y}_{i,fitted}$ is the simulated value of y_i using a regression model between y_i and $\hat{y}_i$, n is the number of observations and $\bar{y}$ is the average of observed values

spatial inputs. Both are applicable as site SI (depending on an output point) and block SI (depending on the size of the spatial support defined by the upscaling process) in the case of upscaling process (average or sum of a point-based model) in the model output. This approach was applied to a point-based hydrology model. It was shown that the SI depends on the size of the spatial support and also that the uncertainty and the influence of the inputs on the outputs was spatially heterogeneous (Saint-Geours et al., 2014). Saint-Geours et al. (2014) showed that the SI of spatial inputs decreased with an increase in unit support size and the SI of aspatial inputs increased if the unit support size increased. Their study showed that a ratio of SI can be determined with the unit support size. Uncertainty analysis and SA on an environmental model at different scales has also shown that SI varies with the modeling scale (Şalap-Ayça & Jankowski, 2018). These results have raised important questions such as: Which unit size (and scale) for the output should the user choose to limit uncertainty propagation of spatial or aspatial inputs? What level of uncertainty are model users ready to accept to make a decision, and how is this uncertainty spatially distributed?

Why is an evaluation of spatialized models different from current model evaluation important ?

As shown previously, the evaluation of spatialized crop models is currently done with aspatial indicators. Often, model evaluations are made without either accounting for a change of scale or the spatial character of the data. Tedeschi (2006) highlights that statistical analysis to evaluate predictive models is essential and needs to be appropriate for the model use in order to evaluate its precision and accuracy. For instance, there is an issue when input or output data are spatially autocorrelated, such that errors (i.e. difference between observed and simulated data) are not independent. The presence of this spatial autocorrelation can strongly reduce the reliability of many statistical metrics, including some popular ones shown in Table 1. Moreover, a lot of environmental variables present a continuity in their spatial structure so those variables are spatially dependent (Zhao et al., 2016).

Saint-Geours et al. (2014) showed that the output variance explained by spatial inputs decreases with an upscaling process, due to a data smoothing effect. This result highlights that evaluation should take into account changes of scale because model performance can depend on the scale at which it is run. The link between uncertainty propagation and scale change (upscaling and downscaling) is an area that requires more consideration (Saint-Geours et al., 2012).

To illustrate the issue and the need for new approaches to spatialized crop model evaluation, a simple case study is presented here. The aim is to demonstrate the limitation of aspatial statistics that have been widely used for evaluation of spatialized crop models in the recent literature (Table 1). In this case, the RMSE is used as the example statistic. In the case study, the intent is to define management zones (MZ) within a vineyard for precision viticulture. The predicted variable that is used to define these MZs is the predawn leaf water potential (PLWP). The purpose of this example is to show that with different theoretical spatialized models of PLWP, the outcomes of clustering based on PLWP predictions can be variable and independent of the RMSE.

This simulated example is built on observed data of PLWP on a 1.2 ha Shiraz vineyard in 2003. This vineyard is located in Pech Rouge (INRAE Gruissan, 43° 08′ 47″ N, 03° 07′ 19″ E) (See Acevedo-Opazo et al., 2010 for full details of the data set). To simulate the output from various theoretical spatialized crop models, three noise models were constructed, all built from the same values sampled from a normal distribution
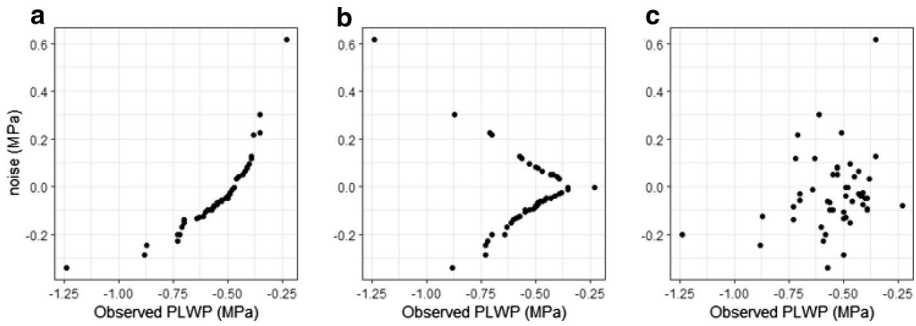
Fig. 5 Attribution of noise depending on observed PLWP values: **a** Model 1—Noise attributed is positively related to the PLWP values, **b** Model 2—Noise absolute value are positively applied to the observed PLWP values, and **c** Model 3—Noise randomly applied to the observed PLWP

Table 2 Observed field data and three simulated Predawn Leaf Water Potential (PLWP) models classified into Management Zones (MZs) based on a 3-class classification of the observed data (and the same thresholds used for the simulated data)

	Observed PLWP	Simulated PLWP		
	Real Data	Model 1	Model 2	Model 3
MZ				
RMSE	–	0.16	0.16	0.16
Cohen's Kappa	–	0.64	0.05	0.31

Metrics of model fit are shown as the RMSE between observed and simulated data (n=49) and Cohen's Kappa value associated with the similarity of the simulated MZ models to the observed MZ model

with a fixed mean (0) and variance (0.2). Various levels of spatial structure in the simulated PLWP were then obtained by altering how these noise distributions were associated with the observed PLWP values. Two of the noise distributions were dependent on the observed PLWP (some spatial structure), while the third distribution was random (i.e. independent from observed PLWP) (Fig. 5). These three noise models were added to the original data to simulate the output from three theoretical models. The original PLWP data were made into MZs based on a tiertile analysis and the threshold values from this analysis was used to create the MZs in the simulated PLWP maps (Table 2). The agreement between the MZ maps was determined using Cohen's Kappa statistic (Eq. 1) (Cohen, 1960).

$$kappa = \frac{P_0 - P_e}{1 - P_e} \qquad (1)$$

where P_0 is the proportion of agreement observed (i.e. the proportion of agreement between MZs of observed and simulated data) and P_e is the proportion of a random agreement (i.e. the proportion of agreement in the case MZ are derived from observed and spatially reorganized data at random).

The RMSE was calculated from the simulated PLWP (i.e. sum of observed PLWP and attributed noise) and observed PLWP. Thus, the RMSE should identify which simulation is the best. However, because the simulated noise models are built from the same distribution (but different spatial structure) the RMSE in these cases was identical (Table 2). Therefore, the conclusion is that all three simulated models were equally good, and the defined MZs should be equally good. However, the resulting MZ maps for the three simulated models do not support this, nor do the Cohen's Kappa values (Table 2). Even though the RMSE was constant, the Model 1 spatial pattern was much closer to the original data (higher Cohen's Kappa value) than Model 2 or 3. Model 2 had the least similar spatial pattern to the observed data (lowest Cohen's Kappa value). Thus, even though the RMSE was the same on these three simulations, the derived MZs were significantly different between simulations. Selecting the best MZ (i.e. from the best spatialized model) cannot be decided with only the RMSE.

Perspectives and needs for crop model spatialization and evaluation in a precision agriculture context

Crop model spatialization is currently often used with upscaling methods. Published studies have aimed to apply traditional point crop models over larger areas, for instance, at the field scale (Acevedo-Opazo et al., 2010), at multiple field scales (Baralon et al., 2012), at the regional scale (Balkovič et al., 2013; Battude et al., 2016; Beaudoin et al., 2018; Therond et al., 2011) and at the continental scale (Adam et al., 2011; Teixeira et al., 2017). Upscaling crop models to a larger area is more common because crop models can be used by land managers and policymakers to make decisions on these large areas (Jones et al., 2017). In contrast, there are considerably fewer studies that have aimed to use crop models at finer scales, i.e. attempting downscaling rather than upscaling. Precision agriculture is much more concerned with finer scale predictions and so downscaling approaches applied to crop models are of particular interest to the precision agriculture community.

Using crop models in a tactical management way represents a goal of precision agriculture. Nevertheless, to achieve this objective, crop models need to manage a large amount of ancillary spatial data (Chen et al., 2017). Chen et al. (2017) identify different kind spatial data: relatively stable data (e.g. soil type and depth), constantly changing data (e.g. LAI, soil moisture and temperature, solar radiation) and aspatial data (e.g. management activities, cultivar information). In addition to the nature of these data, the resolution of these data has to be taken into account and needs to match with the spatial footprint of the modelling (Adam et al., 2011). Spatialization, in the context of downscaling crop modelling approaches, leans heavily on using these high resolution data to define (relatively homogeneous) sub-units on which to apply a crop model (Basso et al., 2011; Cammarano et al., 2019, 2021; Guo et al., 2018).

If crop modelling using spatialized crop models is to become a common aspect of precision agriculture then new methods or statistical metrics that take into account the spatial

characteristics of the data and models will be needed. Evaluation of spatialized crop models needs to be improved and evolve beyond aspatial metrics. These new statistical metrics could take into account some spatial characteristics of the data. For instance, systematically using variography to estimate the spatial structure of inputs and outputs could be a method to identify if there is an issue in input data, in the model structure or with an interaction between both of them. Using geostatistic metrics on residual crop models could be a solution to improve the spatialized crop model evaluation. At a minimum, an evaluation of spatial autocorrelation should be performed, such as using Moran's autocorrelation coefficient (Moran, 1948) on the inputs, outputs or residuals to provide quantitative evidence of spatial autocorrelation in the data.

Furthermore, due consideration needs to be given as to how these spatial data can be best used as both inputs into the model and as data for the calibration and evaluation of the models. Spatial ancillary data are often derived data layers themselves with some level of error and uncertainty associated with them. The wide variety of gridded high-resolution digital soil property maps (https://esdac.jrc.ec.europa.eu/resource-type/soil-data-maps, accessed 24/05/2021) now available are a good example of this. Soil property information is essential for many crop models, and better soil information is critical to expanding the uses of crop models. However, these soil property maps are estimates, derived from modelling approaches themselves. They are not directly measured soil properties that can be entered with confidence into the models, but the temptation is to treat these spatial ancillary data as 'true' data for modelling purposes. This temptation should be avoided and robust modelling approaches that explicitly take into account input uncertainty, such as Monte Carlo methods, should be routinely used in spatialized modelling applications.

Finally, the evaluation on any spatialized (or spatial) crop model will be affected by the number and spatial location of any real observations used for validation. This is true for any scale of application. However, for finer scale spatial modelling that is to be used for short-term predictive modelling to aid in-season management, the selection of correct validation sites is critical as there is limited time to resample before the crop model output needs to be used to make (spatial) management decisions. As for any modelling approach, if the validation sites do not cover the distribution of both model inputs and outputs then model evaluation will be restricted and diminished. In the case of using spatialized (or spatial) models, the spatial distribution and relevance of these validation sites must also be considered when they are selected. Related to this is the need for any validation data to respect the spatial footprint of the model outputs, either in its native form or after scaling. This in turn creates potential issues for crop model validation if the model outputs are multi-scalar in nature.

All of this comes back to the type of metric that is best suited to evaluate spatialized (and spatial) crop models. None of the metrics in Table 1 were developed for or are suitable to address these issues. How would a comparison between a well-performed model, with poorly selected spatial validation sites at the incorrect scale, and a poorly-performed model, with well selected sites correctly sampled be properly made so that the better model was identified? Note that this question is not an issue of the quality of the analysis, but the location and the spatial footprint of the sampling. The assumption is that the analysis of the validation data is done equally well in both instances.

It is clear from the review of the literature performed here that there has not been a lot of consideration so far of spatial issues when applying crop models to precision agriculture. Despite precision agriculture being built on spatial data sets, spatial autocorrelation and its implications for statistical analysis, particularly for the assumptions behind many statistical methods, are often overlooked (Taylor & Bates, 2013). In many cases this

is because precision agriculturists do not always fully comprehend the statistical implications behind spatial data (compared to 'conventional' agri-data sets). This is generally true across all aspects of agricultural science that are seeking to include spatial data in their domain, including the crop modelling domain. To ensure the correct use of these spatial data, agricultural scientists and modelers will continue to need the support of the statistical community, particularly the geo-statistical community, to develop new metrics to support this new area of crop modelling.

Conclusion

Existing crop models are point-based models and spatialization allows the use of these crop models to predict spatially across an area. Spatialization of a crop model is realized for different reasons, such as applying site-specific crop management, improving understanding of processes or to complete data sets. Most published studies have addressed spatialization from an upscaling objective to inform regional, national or global decision-making. However, in a precision agriculture context, downscaling methods need to be used for the spatialization of most crop models and only limited research has been performed in this domain so far. In addition to crop model uncertainty itself, scaling methods to spatialize models will add uncertainty to the model predictions. The present review raised questions about the current approaches to the evaluation of spatialized crop models. Current evaluation methods in published studies have used mainly aspatial indicators. When spatializing crop models, spatial autocorrelation should be considered and assessed, otherwise, crop model evaluation could be wrong. Additionally, spatialized predictive crop model evaluation will be influenced by the number, location and spatial footprint of validation data. To overcome those issues, indicators and coefficients that take spatial autocorrelation in account when evaluating the performance of the spatialized (or spatial) crop model are urgently needed and should be developed via a collaboration of the crop modelling and biometry (statistical) communities.

Glossary

Crop model   A system-based model that is used to simulate daily dynamic interactions between the "soil-plant-atmosphere" (Wallach et al., 2019). In a broader sense, they complement field experiments and can be used to extrapolate/integrate observed data.

Point-based model A model designed to predict variables on a point (i.e. an unit support) without taking into account neighbouring data or effects to compute the variable. This is what is considered a 'classical' crop model.

Spatialization
(or model
spatialization)   Means to apply point-based models spatially across an area different from the native model area (unit support) on which it was designed.

Spatialized model A point-based model that has had a spatialization process applied to it (see above).

Spatial model	A model designed to compute output variables taking into account neighbouring data or/and effects.
Spatial footprint	The native model spatial area.
Finer scale	Corresponds to a more accurate resolution (disaggregation method, smaller pixels) or the use of a model on a smaller scale than which the model was initially designed.
Larger scale	Corresponds to a more coarse resolution (aggregation method, larger pixels) or the use of a model on a larger scale than which the model was initially designed.
Processes	Corresponds to an activity ensemble which are correlated or interactive. These activities can be biological, environmental, physical or chemical.
Variable	Variable is only used to describe a model input or model output.
State variable	A 'state variable' is a variable that is internal (calculated) and not an input and is not necessarily expected to be given as a model output.
Parameter	Parameter is only used to describe a part of a model. Parameters can be calibrated using different methods.
Calibration	A process to find the best values of model parameters by using the observed data, a consequence of calibration is that simulated data are better fitted to observed data.
Resolution	Refers to the minimum scale of both spatial and/or temporal phenomena (i.e. is non-specific). If a statement only refers to a spatial or a temporal phenomena, then it will be described as such (i.e. spatial resolution or temporal resolution), otherwise it may mean either or both.
Data assimilation	A suite of methods to combine simulated data from a model and observed data. It aims to find an optimal combination between both to improve model predictions (e.g. by recalibrating or updating a model) (Huang et al., 2019).
Data fusion	Methods to combine data from different sources into an integrated and unified compound with higher quality of information (Bleiholder & Naumann, 2009; Oliveira et al., 2021).
Evaluation	Refers to the question of knowing how well model predictions are relevant with measures collected in real-world situations. The aim is to ascertain the value computed by models.

References

Acevedo-Opazo, C., Tisseyre, B., Ojeda, H., Ortega-Farias, S., & Guillaume, S. (2008). Is it possible to assess the spatial variability of vine water status? *OENO One, 42*(4), 203–219. https://doi.org/10.20870/oeno-one.2008.42.4.811

Acevedo-Opazo, C., Tisseyre, B., Taylor, J. A., Ojeda, H., & Guillaume, S. (2010). A model for the spatial prediction of water status in vines (Vitis vinifera L.) using high resolution ancillary information. *Precision Agriculture, 11*(4), 358–378. https://doi.org/10.1007/s11119-010-9164-7

Adam, M., Van Bussel, L. G. J., Leffelaar, P. A., Van Keulen, H., & Ewert, F. (2011). Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions. *Ecological Modelling, 222*(1), 131–143. https://doi.org/10.1016/j.ecolmodel.2010.09.001

Allain, S., Plumecocq, G., & Leenhardt, D. (2018). Spatial aggregation of indicators in sustainability assessments: Descriptive and normative claims. *Land Use Policy, 76*, 577–588. https://doi.org/10.1016/j.landusepol.2018.02.038

Al-Shammari, D., Whelan, B. M., Wang, C., Bramley, R. G. V., Fajardo, M., & Bishop, T. F. A. (2021). Impact of spatial resolution on the quality of crop yield predictions for site-specific crop management. *Agricultural and Forest Meteorology, 310*, 108622. https://doi.org/10.1016/j.agrformet.2021.108622

Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S., & Ewert, F. (2013a). Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. *Agricultural and Forest Meteorology, 170*, 32–46. https://doi.org/10.1016/j.agrformet.2012.11.017

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P., & Ewert, F. (2013b). Characteristic 'fingerprints' of crop model responses to weather input data at different spatial resolutions. *European Journal of Agronomy, 49*, 104–114. https://doi.org/10.1016/j.eja.2013.04.003

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., et al. (2013). Uncertainty in simulating wheat yields under climate change. *Nature Climate Change, 3*(9), 827–832. https://doi.org/10.1038/nclimate1916

Asseng, S., Martre, P., Maiorano, A., Rötter, R. P., O'Leary, G. J., Fitzgerald, G. J., et al. (2018). Climate change impact and adaptation for wheat protein. *Global Change Biology, 25*(1), 155–173. https://doi.org/10.1111/gcb.14481

Balkovič, J., van der Velde, M., Schmid, E., Skalský, R., Khabarov, N., Obersteiner, M., et al. (2013). Pan-European crop modelling with EPIC: Implementation, up-scaling and regional crop yield validation. *Agricultural Systems, 120*, 61–75. https://doi.org/10.1016/j.agsy.2013.05.008

Baralon, K., Payan, J.-C., Salançon, E., & Tisseyre, B. (2012). Spider: Spatial extrapolation of the vine water status at the whole denomination scale from a reference site. *OENO One, 46*(3), 167–175. https://doi.org/10.20870/oeno-one.2012.46.3.1517

Basso, B., Ritchie, J. T., Cammarano, D., & Sartori, L. (2011). A strategic and tactical management approach to select optimal N fertilizer rates for wheat in a spatially variable field. *European Journal of Agronomy, 35*(4), 215–222. https://doi.org/10.1016/j.eja.2011.06.004

Basso, B., Ritchie, J. T., Pierce, F. J., Braga, R. P., & Jones, J. W. (2001). Spatial validation of crop models for precision agriculture. *Agricultural Systems, 68*(2), 97–112. https://doi.org/10.1016/S0308-521X(00)00063-9

Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., et al. (2016). Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sensing of Environment, 184*, 668–681. https://doi.org/10.1016/j.rse.2016.07.030

Beaudoin, N., Gallois, N., Viennot, P., Le Bas, C., Puech, T., Schott, C., et al. (2018). Evaluation of a spatialized agronomic model in predicting yield and N leaching at the scale of the Seine-Normandie Basin. *Environmental Science and Pollution Research, 25*, 23529–23558. https://doi.org/10.1007/s11356-016-7478-3

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., et al. (2013). Characterising performance of environmental models. *Environmental Modelling & Software, 40*, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Blanchoud, H., Schott, C., Tallec, G., Queyrel, W., Gallois, N., Habets, F., et al. (2020). How should agricultural practices be integrated to understand and simulate long-term pesticide contamination in the Seine River basin? In N. Flipo, P. Labadie, & L. Lestel (Eds.), *The Seine River Basin* (Vol. 90, pp. 141–162). Springer. https://doi.org/10.1007/698_2019_385

Bleiholder, J., & Naumann, F. (2009). Data fusion. *ACM Computing Surveys, 41*(1), 1–41. https://doi.org/10.1145/1456650.1456651

Blöschl, G. (2005). Statistical upscaling and downscaling in hydrology. In M. G. Anderson & J. J. McDonnell (Eds.), *Encyclopedia of hydrological sciences* (pp. 135–154). Wiley. https://doi.org/10.1002/0470848944.hsa008

Boote, K., Hoogenboom, G., Porter, C. H., Boote, K. J., Shelia, V., Wilkens, P. W., et al. (2019). The DSSAT crop modeling ecosystem. In K. Boote (Ed.), *Burleigh Dodds series in agricultural science* (pp. 173–216). Burleigh Dodds Science Publishing. https://doi.org/10.19103/AS.2019.0061.10

Bouman, B. A. M., van Keulen, H., van Laar, H. H., & Rabbinge, R. (1996). The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. *Agricultural Systems, 52*(2–3), 171–198. https://doi.org/10.1016/0308-521X(96)00011-X

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., et al. (2003). An overview of the crop model stics. *European Journal of Agronomy, 18*(3–4), 309–332. https://doi.org/10.1016/S1161-0301(02)00110-7

Cammarano, D., Basso, B., Holland, J., Gianinetti, A., Baronchelli, M., & Ronga, D. (2021). Modeling spatial and temporal optimal N fertilizer rates to reduce nitrate leaching while improving grain yield and quality in malting barley. *Computers and Electronics in Agriculture, 182*, 105997. https://doi.org/10.1016/j.compag.2021.105997

Cammarano, D., Holland, J., Basso, B., Fontana, F., Murgia, T., Lange, C., et al. (2019). Integrating geospatial tools and a crop simulation model to understand spatial and temporal variability of cereals in Scotland. In Precision agriculture '19 Presented at the 12th European Conference on Precision Agriculture, Wageningen Academic Publishers, Montpellier, France (pp. 29–35). https://doi.org/10.3920/978-90-8686-888-9_2

Cammarano, D., Rivington, M., Matthews, K. B., Miller, D. G., & Bellocchi, G. (2017). Implications of climate model biases and downscaling on crop model simulated climate change impacts. *European Journal of Agronomy, 88*, 63–75. https://doi.org/10.1016/j.eja.2016.05.012

Cammarano, D., Valdivia, R. O., Beletse, Y. G., Durand, W., Crespo, O., Tesfuhuney, W. A., et al. (2020). Integrated assessment of climate change impacts on crop productivity and income of commercial maize farms in northeast South Africa. *Food Security, 12*(3), 659–678. https://doi.org/10.1007/s12571-020-01023-0

Challinor, A. J., Ewert, F., Arnold, S., Simelton, E., & Fraser, E. (2009). Crops and climate change: Progress, trends, and challenges in simulating impacts and informing adaptation. *Journal of Experimental Botany, 60*(10), 2775–2789. https://doi.org/10.1093/jxb/erp062

Chen, H., Leinonen, I., Marshall, B., & Taylor, J. A. (2017). Conceptual spatial crop models for potato production. *Advances in Animal Biosciences, 8*(2), 678–683. https://doi.org/10.1017/S2040470017000851

Choukri, F., Raclot, D., Naimi, M., Chikhaoui, M., Nunes, J. P., Huard, F., et al. (2020). Distinct and combined impacts of climate and land use scenarios on water availability and sediment loads for a water supply reservoir in northern Morocco. *International Soil and Water Conservation Research, 8*(2), 141–153. https://doi.org/10.1016/j.iswcr.2020.03.003

Claverie, M., Demarez, V., Duchemin, B., Hagolle, O., Ducrot, D., Marais-Sicre, C., et al. (2012). Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. *Remote Sensing of Environment, 124*, 844–857. https://doi.org/10.1016/j.rse.2012.04.005

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Colaço, A. F., & Bramley, R. G. V. (2018). Do crop sensors promote improved nitrogen management in grain crops? *Field Crops Research, 218*, 126–140. https://doi.org/10.1016/j.fcr.2018.01.007

Constantin, J., Raynal, H., Casellas, E., Hoffmann, H., Bindi, M., Doro, L., et al. (2019). Management and spatial resolution effects on yield and water balance at regional scale in crop models. *Agricultural and Forest Meteorology, 275*, 184–195. https://doi.org/10.1016/j.agrformet.2019.05.013

Crosetto, M., Tarantola, S., & Saltelli, A. (2000). Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture, Ecosystems & Environment, 81*(1), 71–79. https://doi.org/10.1016/S0167-8809(00)00169-9

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., et al. (2019). 25 years of the WOFOST cropping systems model. *Agricultural Systems, 168*, 154–167. https://doi.org/10.1016/j.agsy.2018.06.018

Domínguez-Álvarez, A., De-Tena-Rey, M.-T., & García-Moruno, L. (2021). Modelling global solar radiation to optimise agricultural production. *Spanish Journal of Agricultural Research, 19*(1), e0201. https://doi.org/10.5424/sjar/2021191-16813

Duchemin, B., Maisongrande, P., Boulet, G., & Benhadj, I. (2008). A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index. *Environmental Modelling & Software, 23*(7), 876–892. https://doi.org/10.1016/j.envsoft.2007.10.003

Ehrhardt, F., Soussana, J., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., et al. (2018). Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N₂O emissions. *Global Change Biology, 24*(2), e603–e616. https://doi.org/10.1111/gcb.13965

Ewert, F., van Ittersum, M. K., Heckelei, T., Therond, O., Bezlepkina, I., & Andersen, E. (2011). Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agriculture, Ecosystems & Environment, 142*(1–2), 6–17. https://doi.org/10.1016/j.agee.2011.05.016

Faivre, R., Leenhardt, D., Voltz, M., Benoît, M., Papy, F., Dedieu, G., & Wallach, D. (2004). Spatialising crop models. *Agronomie, 24*(4), 205–217. https://doi.org/10.1051/agro:2004016

Ginaldi, F., Bajocco, S., Bregaglio, S., & Cappelli, G. (2019). Spatializing crop models for sustainable agriculture. In M. Farooq & M. Pisante (Eds.), *Innovations in sustainable agriculture* (pp. 599–619). Springer. https://doi.org/10.1007/978-3-030-23169-9

Grosz, B., Dechow, R., Gebbert, S., Hoffmann, H., Zhao, G., Constantin, J., et al. (2017). The implication of input data aggregation on up-scaling soil organic carbon changes. *Environmental Modelling & Software, 96*, 361–377. https://doi.org/10.1016/j.envsoft.2017.06.046

Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., et al. (2018). Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precision Agriculture, 19*(1), 55–78. https://doi.org/10.1007/s11119-017-9498-5

Heuvelink, G. B. M. (2002). Analysing uncertainty propagation in GIS: Why is it not that simple? In G. M. Foody & P. M. Atkinson (Eds.), *Uncertainty in remote sensing and GIS* (pp. 155–165). Wiley. https://doi.org/10.1002/0470035269.ch10

Heuvelink, G. B. M., Brus, D. J., & Reinds, G. (2010). Accounting for spatial sampling effects in regional uncertainty propagation analysis. In Presented at the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester (pp. 85–88).

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., et al. (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLoS ONE, 11*(4), e0151782. https://doi.org/10.1371/journal.pone.0151782

Hoffmann, H., Zhao, G. R., van Bussel, L. G. J., Enders, A., Specka, X., Sosa, C., et al. (2015). Variability of effects of spatial climate data aggregation on regional yield simulation by crop models. *Climate Research, 65*, 53–69. https://doi.org/10.3354/cr01326

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., et al. (2014). APSIM—Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software, 62*, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009

Hoogenboom, G. (2000). Contribution of agrometeorology to the simulation of crop production and its applications. *Agricultural and Forest Meteorology, 103*(1–2), 137–157. https://doi.org/10.1016/S0168-1923(00)00108-8

Huang, J., Gómez-Dans, J. L., Huang, H., Ma, H., Wu, Q., Lewis, P. E., et al. (2019). Assimilation of remote sensing into crop growth models: Current status and perspectives. *Agricultural and Forest Meteorology, 276–277*, 107609. https://doi.org/10.1016/j.agrformet.2019.06.008

Huard, F., Choukri, F., Raclot, D., Pépin, Y., Chikhaoui, M., Naimi, M., et al. (2019). Spatialisation fine des projections climatiques dans le bassin versant du Tleta au Nord Maroc (Fine spatialization of climate projections in the Tleta watershed in northern Morocco). *Revue Marocaine Des Sciences Agronomiques Et Vétérinaires, 7*(2), 294–303.

Jagtap, S. S., & Jones, J. W. (2002). Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agriculture, Ecosystems & Environment, 93*, 73–85.

Jankowski, P., Andrienko, N., & Andrienko, G. (2001). Map-centred exploratory approach to multiple criteria spatial decision making. *International Journal of Geographical Information Science, 15*(2), 101–127. https://doi.org/10.1080/13658810010005525

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. *European Journal of Agronomy, 92*, 141–152. https://doi.org/10.1016/j.eja.2017.11.002

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., et al. (2017). Brief history of agricultural systems modeling. *Agricultural Systems, 155*, 240–254. https://doi.org/10.1016/j.agsy.2016.05.014

Kasampalis, D. A., Alexandridis, T. K., Deva, C., Challinor, A., Moshou, D., & Zalidis, G. (2018). Contribution of remote sensing on crop models: A review. *Journal of Imaging, 4*(4), 52. https://doi.org/10.3390/jimaging4040052

Li, J., Zhang, H., & Xu, E. (2020). Spatialization of actual grain crop yield coupled with cultivation systems and multiple factors: From survey data to grid. *Agronomy, 10*(5), 675. https://doi.org/10.3390/agronomy10050675

MacKerron, D. K. L., Marshall, B., & McNicol, J. W. (2004). MAPP and the underlying functions that it contains. In D. K. L. MacKerron & A. J. Haverkort (Eds.), *Decision support systems in potato production—Bringing models to practice*. Wageningen Academic Publishers.

Maiorano, A., Martre, P., Asseng, S., Ewert, F., Müller, C., Rötter, R. P., et al. (2017). Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. *Field Crops Research, 202*, 5–20. https://doi.org/10.1016/j.fcr.2016.05.001

Malone, B. P., McBratney, A. B., & Minasny, B. (2013). Spatial scaling for digital soil mapping. *Soil Science Society of America Journal, 77*(3), 890–902. https://doi.org/10.2136/sssaj2012.0419

Martre, P., Wallach, D., Asseng, S., & Ewert, F. (2015). Multimodel ensembles of wheat growth: Many models are better than one. *Global Change Biology, 21*, 911–925. https://doi.org/10.1111/gcb.12768

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (methodological), 10*(2), 243–251. https://doi.org/10.1111/j.2517-6161.1948.tb00012.x

Oliveira, D., Martins, L., Mora, A., Damásio, C., Caetano, M., Fonseca, J., & Ribeiro, R. A. (2021). Data fusion approach for eucalyptus trees identification. *International Journal of Remote Sensing, 42*(11), 4087–4109. https://doi.org/10.1080/01431161.2021.1883198

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software, 79*, 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008

Roux, S., Gaudin, R., & Tisseyre, B. (2019). Why does spatial extrapolation of the vine water status make sense? Insights from a modelling approach. *Agricultural Water Management, 217*, 255–264. https://doi.org/10.1016/j.agwat.2019.03.013

Saint-Geours, N., Bailly, J.-S., Grelot, F., & Lavergne, C. (2014). Multi-scale spatial sensitivity analysis of a model for economic appraisal of flood risk management policies. *Environmental Modelling & Software, 60*, 153–166. https://doi.org/10.1016/j.envsoft.2014.06.012

Saint-Geours, N., Lavergne, C., Bailly, J.-S., & Grelot, F. (2012). Change of support in spatial variance-based sensitivity analysis. *Mathematical Geosciences, 44*(8), 945–958. https://doi.org/10.1007/s11004-012-9406-5

Şalap-Ayça, S., & Jankowski, P. (2018). Analysis of the influence of parameter and scale uncertainties on a local multi-criteria land use evaluation model. *Stochastic Environmental Research and Risk Assessment, 32*(9), 2699–2719. https://doi.org/10.1007/s00477-018-1535-z

Scholes, R., Reyers, B., Biggs, R., Spierenburg, M., & Duriappah, A. (2013). Multi-scale and cross-scale assessments of social–ecological systems and their ecosystem services. *Current Opinion in Environmental Sustainability, 5*(1), 16–25. https://doi.org/10.1016/j.cosust.2013.01.004

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? *European Journal of Agronomy, 94*, 25–35. https://doi.org/10.1016/j.eja.2018.01.006

Sinclair, T. R., & Seligman, N. G. (1996). Crop modeling: From infancy to maturity. *Agronomy Journal, 88*(5), 698–704. https://doi.org/10.2134/agronj1996.00021962008800050004x

Steduto, P., Hsiao, T. C., Raes, D., & Fereres, E. (2009). AquaCrop-The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. *Agronomy Journal, 101*(3), 426–437. https://doi.org/10.2134/agronj2008.0139s

Stöckle, C. O., Donatelli, M., & Nelson, R. (2003). CropSyst, a cropping systems simulation model. *European Journal of Agronomy, 18*(3–4), 289–307. https://doi.org/10.1016/S1161-0301(02)00109-0

Taylor, J. A., & Bates, T. R. (2013). A discussion on the significance associated with Pearson's correlation in precision agriculture studies. *Precision Agriculture, 14*(5), 558–564. https://doi.org/10.1007/s11119-013-9314-9

Tedeschi, L. O. (2006). Assessment of the adequacy of mathematical models. *Agricultural Systems, 89*(2–3), 225–247. https://doi.org/10.1016/j.agsy.2005.11.004

Teixeira, E. I., Zhao, G., de Ruiter, J., Brown, H., Ausseil, A.-G., Meenken, E., & Ewert, F. (2017). The interactions between genotype, management and environment in regional crop modelling. *European Journal of Agronomy, 88*, 106–115. https://doi.org/10.1016/j.eja.2016.05.005

Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., et al. (2011). Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. *Agriculture, Ecosystems & Environment, 142*(1–2), 85–94. https://doi.org/10.1016/j.agee.2010.05.007

van Bussel, L. G. J., Ewert, F., & Leffelaar, P. A. (2011). Effects of data aggregation on simulations of crop phenology. *Agriculture, Ecosystems & Environment, 142*(1–2), 75–84. https://doi.org/10.1016/j.agee.2010.03.019

van Ittersum, M. K., & Donatelli, M. (2003). Modelling cropping systems—Highlights of the symposium and preface to the special issues. *European Journal of Agronomy, 18*, 187–197. https://doi.org/10.1016/S1161-0301(02)00095-3

Wallach, D., Makowski, D., Jones, J. W., & Brun, F. (2019). *Working with dynamic crop models—Methods, tools and examples for agriculture and environment* (3rd ed.). Elsevier.

Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P. J., et al. (2018). Multimodel ensembles improve predictions of crop–environment–management interactions. *Global Change Biology*. https://doi.org/10.1111/gcb.14411

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R., et al. (2017). The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nature Plants, 3*, 17102. https://doi.org/10.1038/nplants.2017.102

Zhao, G., Hoffmann, H., van Bussel, L., Enders, A., Specka, X., Sosa, C., et al. (2015). Effect of weather data aggregation on regional crop simulation for different crops, production conditions, and response variables. *Climate Research, 65*, 141–157. https://doi.org/10.3354/cr01301

Zhao, G., Hoffmann, H., Yeluripati, J., Xenia, S., Nendel, C., Coucheney, E., et al. (2016). Evaluating the precision of eight spatial sampling schemes in estimating regional means of simulated yield for two crops. *Environmental Modelling & Software, 80*, 100–112. https://doi.org/10.1016/j.envsoft.2016.02.022