SN Computer Science (2023) 4:554
https://doi.org/10.1007/s42979-023-02103-6

REVIEW ARTICLE

Computational Models in Precision Fruit Growing: Reviewing the Impact of Temporal Variability on Perennial Crop Yield Assessment

Renata Bulling Magro[1] · Silvio André Meirelles Alves[1] · Luciano Gebler[1]

Received: 28 March 2023 / Accepted: 28 June 2023

Abstract

Early yield information of perennial crops is crucial for growers and the industry as it enables cost reduction and facilitates crop planning. However, assessing the yield of perennial crops using computational models poses challenges due to the diverse aspects of interannual variability that affect these crops. This review aimed to investigate and analyze the literature on yield estimation and forecasting modeling in perennial cropping systems. We reviewed 49 articles and categorized them according to their yield assessment strategy, modeling class, and input variable characteristics. The strategies of yield assessment were discussed in the context of their principal improvement challenges. Our investigation revealed that image processing and deep learning models are emerging techniques for yield estimation. On the other hand, machine learning algorithms, such as Artificial Neural Networks and Decision Trees, were applied to yield forecasting with reasonable time in advance of harvest. Emphasis is placed on the lack of representative long-term datasets for developing computational models, which can lead to accurate early yield forecasting of perennial crops.

Keywords Yield modeling · Spatio-temporal analysis · Computational intelligence · Machine learning · Decision support

Introduction

Precision Fruit Growing (PFG) is a branch of traditional Precision Agriculture that focuses on understanding and improving the production of perennial fruit species. This is achieved through the assessment of a cycle repeated annually—from tree implantation to eradication [31]. To effectively implement PFG, it is relevant to consider the orchard's lifetime, which, for apple trees, can range from 20 to 25 years [48].

The influence of previous cycles on the present yield of perennial crops adds complexity to these systems. This can be exemplified by irregular production between crop years [24]. Consequently, one of the first steps in applying PFG is to comprehend the factors that determine the spatial and temporal patterns of yield [11]. Given that the tree production in a particular year corresponds to the cumulative impact of various variables, it is not strategic to solely rely on annual variables when evaluating the orchard's yield. Therefore, the use of computational methods to assess fruit growing outcomes must align with the management concepts derived from Precision Agriculture. This approach involves acknowledging the inherent spatial and temporal variability within each productive area on the rural property, rather than evaluating system productivity based on arbitrary divisions (such as plots or blocks) generated for personal convenience, time, or location.

Additionally, considering temporal and spatial aspects for developing more efficient and relevant methods of evaluating the yield of perennial crops can contribute to their better operational use [65, 77]. These advanced evaluation methods have the potential to enhance decision-making processes and increase productivity [47]. Therefore, there is a growing interest from both industry and the academic community in exploring yield assessment models for various crops, including apple, coffee, citrus, and even jujube [1, 9, 55, 75]. In this context, this review investigated the existing literature on the use of computational models for evaluating the yield of perennial fruit crops, offering insights into the

✉ Luciano Gebler
luciano.gebler@embrapa.br

Renata Bulling Magro
renatamagro4@gmail.com

Silvio André Meirelles Alves
silvio.alves@embrapa.br

[1] Embrapa Uva e Vinho, Vacaria, Rio Grande do Sul, Brazil

Published online: 29 July 2023

main strategies and technologies employed for estimating and forecasting yield. Furthermore, it provides directions for future studies in this field.

The content of this review is organized as follows: the section "Contextual aspects related to the yield of perennial crops" addresses general aspects related to the final yield of perennial fruit crops; the section "Review protocol" presents the review methodology of the literature; the section "Yield assessment of perennial crops" discusses the methods of yield assessment based on differences between estimation and forecasting models; the section "Temporal data scale challenges in perennial fruit crop modeling: insights and future directions" summarizes challenges and perspectives for future studies of yield assessment modeling strategies; finally, conclusions on the subject are presented in the section "Conclusions".

Contextual Aspects Related to the Yield of Perennial Crops

Yield is a variable at the end of the productive period that integrates the cumulative effects of climate and management practices throughout the season. The yield of a crop can be influenced by exogenous or endogenous factors [30]. Endogenous factors are associated with morphological and physiological aspects that influence tree behavior from the induction of sprouting to the period of fruit growth and fruit maturation. The extent to which endogenous factors affect yield can be attributed to the biological response of the tree (genetic factor) or determined by the cultivar and specific environmental conditions. On the other hand, environmental aspects are classified as exogenous factors [35]. These correspond to external interferences, such as climatic variables and anthropic actions related to crop management.

In the case of perennial crops, the final yield is not only influenced by the annual interferences of endogenous and exogenous factors but also by the cumulative effects experienced over the lifetime of the trees [3, 48]. Some management actions can impact yield for longer periods than one season, such as pruning and thinning practices. In such cases, the effects produced on the tree can extend at least until the next season, indicating a strong relationship between yield and temporal response parameters.

The temporal variability, as well as the spatial variability, related to yield can be associated with a stochastic process, in which at least one variable of the system behaves randomly over time [65]. Furthermore, several authors reported the occurrence of alternate bearing between years of high and low yield in perennial crops [13, 29, 33, 52]. Thus, modeling yield response with reasonable time before harvest can be a complex task due to variability factors and alternate bearing.

However, the relevance of alternate bearing in the modeling strategy depends on the spatial scale at which the model will be applied [49]. On a local scale, such as a management zone, productive area, or individual plant, the variability of yield among years is more pronounced. In such cases, considering alternate bearing in the modeling strategy becomes necessary to ensure accurate results. On the other hand, when dealing with regional or national scales, the final balance of production tends to be equivalent. This is because areas of low production can be offset by others that have increased productivity in a given year and vice versa [49]

Review Protocol

For searching and selecting the articles, we used specific keywords, such as Machine Learning, Yield Modeling, Yield Forecasting, Yield Prediction, Perennial Crop, Precision Fruit Growing, Artificial Neural Networks, Long-term Forecasting, Artificial Intelligence, Spatio-temporal Forecasting, and Deep Learning. To ensure comprehensive coverage, Google Scholar, Science Direct, and Scopus databases were used. Initially, we screened the articles based on the relevance of their titles to our research topic. Next, we evaluated the objectives of the papers by reviewing their abstracts. Articles that focused on perennial fruit crops, yield assessment, and yield modeling were identified as suitable for inclusion in this literature review. As a result, we selected and integrated 49 articles into the present review.

The articles selected for this review span from 1978 to April 2022 (Fig. 1), as no specific publication time interval was defined. The selected research papers cover a period of 44 years, with a significant concentration of publications in the last 6 years (Fig. 1). Despite the field of yield assessment in PFG has a long history, the increased interest in this topic may be attributed to advancements in

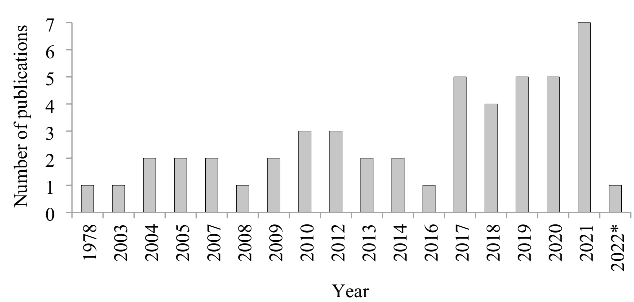

Fig. 1 Publications (research papers) related to yield assessment modeling in perennial fruit crops among the years 1978–2022 (*until April 2022)

technology and the adoption of more sophisticated computational model techniques.

Yield Assessment of Perennial Crops

The yield assessment of a productive area is the closing stage of a cycle, as it allows assessing whether there was a success in carrying out the corrective actions during the management phase. Since the yield assessment corresponds to a summary of the complete production cycle, this literature review focused on estimation and forecasting models of the final yield of perennial crops. However, this review does not cover the entire literature on computational models for assessing the other stages of crop growth. The articles compiled in this review were related to the harvest stage (final yield). This aspect guided the discussion of the strategies and challenges for modeling the final yield in perennial crops, considering spatio-temporal variability.

Strategies for Yield Assessment of Perennial Crops

After the selection of the articles, all obtained documents were analyzed and divided into categories according to their yield assessment modeling strategy:

1. Mathematical models developed based on crop data from the current season (n) for which the yield assessment is desired. These models are applied to estimate the yield and are referred to as Estimating models or Estimation models.
2. Mathematical models developed based on crop data from previous seasons ($n-1$) or before fruit ripening. This modeling strategy considers, at least partially, the temporal factors in the yield assessment. In this case, the crop conditions between data collection, data modeling, and the final yield assessment report may vary in function of the forecast horizon. These models are referred to as Predictive models or Forecasting models.

An overview of the main characteristics of the yield assessment modeling strategies of perennial crops is presented in Fig. 2.

The explanatory variable class used for each evaluation modeling strategy may differ. Yield estimation models usually use direct variables related to direct components of the tree, and sporadically use indirect variables (Fig. 2). The estimation modeling strategy seeks to objectively quantify yield components through field sampling throughout the current production cycle [42]. In this case, data corresponding to the number of flowers and fruits,

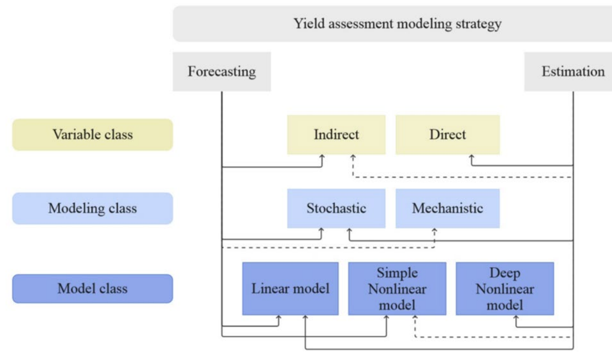

Fig. 2 An overview of the main characteristics of yield assessment modeling strategies (continuous line = higher frequency of use; dashed line = lower frequency of use)

as well as parameters such as average fruit weight and trunk cross-section diameter are commonly used to yield estimation [36]. This way, direct variables combined with statistical methods or computational intelligence may be able to quantify the yield of the current season.

Models for yield forecasting primarily rely mainly on indirect variables (Fig. 2) to establish quantitative cause–effect relationships among physiological variables, climatological variables, and orchard yields. In addition, when making yield forecasting, it is interesting that the variables come from pre-development periods of the crop or past production cycles [42]. That is, quantitative yield data from previous seasons are used [14, 61], or data are collected in periods before the full ripening of the fruits, such as during the budding period.

The relationship between the explanatory variables—whether direct or indirect—and the yield response can be mathematically represented in the modeling step. In terms of the estimation or forecasting objectives, we have identified two modeling classes. The main approach used in operational contexts is stochastic modeling, which involves establishing mathematical relationships between data through statistical or empirical models [14, 19, 20, 43, 55, 58, 75]. Although, when the modeler's knowledge of the mechanisms and processes of yield development is used to establish mathematical relationships among the explanatory variables, the modeling is considered deterministic or mechanistic [22, 49, 69] (Fig. 2).

Regarding the model class, three main subdivisions were listed:

1. Linear Models (LMs) compose the classical statistical models used to describe the behavior of a response variable (dependent variable) as a function of one or more predictor variables (independent variables).
2. Simple Nonlinear Models (SNLMs) correspond to more robust models. SNLMs can perform associations

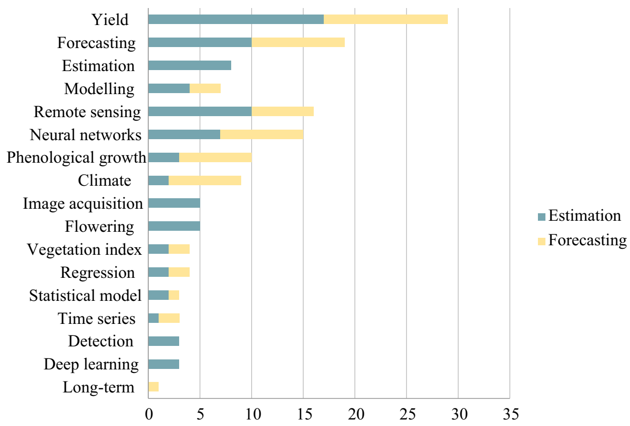

Fig. 3 Quantitative occurrence of keywords in the reviewed literature

between nonlinear data through relatively complex machine learning techniques.

3. Deep Nonlinear Models (DNLMs) correspond to a subarea of machine learning, which includes models based on deep learning techniques that work from massive data and numerous processing layers. DNLMs can manage and process unstructured data such as text and images and automate feature extraction, removing some of the dependency on human experts.

Next, we conducted an analysis of the most relevant keywords, considering the indexing words used in the analyzed documents (Fig. 3). The keywords "yield", "machine learning", and "remote sensing" emerged as the most significant topics. Additionally, the keywords "image processing", "flowering", "detection", and "deep learning" were prevalent in works related to yield estimation. The keyword "long term" appeared in one yield forecasting study. One notable aspect of this analysis was the keyword "forecasting" which was associated with both modeling strategies. We observed that despite using the term "forecasting", some studies utilized direct variables as input data to model the final crop yield in the current season, which may be better characterized as an estimation modeling strategy.

The result of yield modeling may vary depending on factors such as the type of input data and the statistical management of uncertainty allowed by the modeling strategy. These two factors are, then, related to the model type and the assessment time horizon. Therefore, in this review, the yield modeling strategies are discussed based on two main characteristics: (i) the association of the yield assessment with either estimation or forecasting and (ii) the class of models used for yield estimation or yield forecasting.

Yield Estimation Modeling

Yield estimation is the simplest approach to yield assessment, being performed in the same unit, time, and space as the measurement and sampling of input variables [42]. In this case, a common solution to obtain the relationships between the input variables and the final yield is the use of Linear Models (LM). LM appeared in the second place in the model class usage ranking (32%) (Fig. 4), considering that a linear and a known relationship can be established, at least partially, between input and output variables [5]. In addition to the simple linear regression model, multiple linear regression approaches that include more than one input parameter as an explanatory variable were frequently applied to yield estimation modeling [39, 62].

Some studies used regression models (linear or multiple) to estimate orchard yields according to flower density [4, 46, 62]. For Salvo et al. [62], the simplest approach to estimating the amount of fruit is based on the number of flowers, as presented by Jiménez and Diáz [38, 39] for evaluating the production of pears and apples (density of flowers per trunk cross-sectional area). The number of flowers can be obtained manually [38, 39, 62], but it is possible to employ automated methods of computer vision to detect and count flowers and fruits in the orchard. These methods combine image acquisition devices, including cameras acting in the visible band (RGB), and image processing and segmentation algorithms to estimate crop yield [27, 56].

Aggelopoulou et al. [4] used images of trees in full bloom to train a linear regression algorithm to estimate the orchard's yield from the number of flowers on each tree. The results indicated that the method can be used for yield estimation after the full flowering period, providing a yield response at the beginning of harvest in the actual season. However, the orchard area is a relevant factor to be considered in this type of experiment, mainly in terms of operating the system, due to the effort required to obtain the images. The authors manually collected images (static photographs) of 10% of the flowering trees due to the impossibility of capturing images of the entire orchard in the flowering temporal window, considering the technology available at the time.

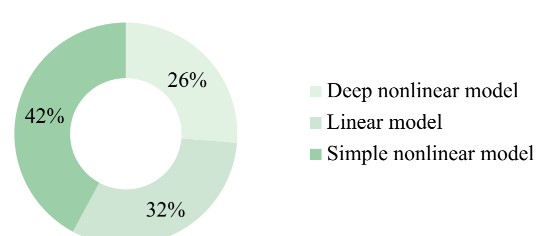

Fig. 4 Percentage of model class for yield estimation in the reviewed literature

In contrast to the collection of images by photography, vegetation indices (VI) are variables easier to obtain as they do not depend on manual acquisition. When combined with LM, VI may allow the improvement of the orchard yield estimation, as verified by Bai et al. [9] during their research on the yield of jujube fruits (*Ziziphus jujuba*). Bai et al. [9] determined that LM estimation results can be optimized by adding VI corresponding to phenological development to the input data. The incorporation of the fruit growth period development was able to improve the yield estimation of jujube fruit.

Anastasiou et al. [6] conducted a study to evaluate the effectiveness of satellite imagery and proximal remote sensing at various developmental stages of grape berry (veraison, mid-veraison, and maturity) in combination with a linear regression model for vineyard yield estimation. Their findings confirmed that the linear regression model can achieve satisfactory results for estimating the yield. Interestingly, the correlations between vegetation indices (VI) and yield varied depending on the method of image acquisition. The study revealed that using a proximal sensor resulted in higher correlations and better yield estimation results, even during the early stages of fruit development. This can be attributed to the superior spatial resolution provided by proximal sensors compared to satellite images. Consequently, proximal sensors have the potential to enhance yield estimation, even in the initial stages of fruit development [6].

Nevertheless, the use of LM can limit the accuracy of the estimation results, since the relationships between orchard yield and input variables are not always linear [12]. In a case study conducted by Logan et al. [47] to estimate the average fruit weight in an apple orchard of the 'Gala' cultivar, both linear (generalized linear model) and nonlinear (Random Forest) models were employed. The results showed that the generalized linear model performed well when using data collected 5 days before harvest. However, when data collected further in advance (12 days before harvest) were utilized, the Random Forest model exhibited better performance. Hence, for early yield estimation models, increased complexity is required, and nonlinear models may be useful for this task.

The SNLM category, which exhibited the highest frequency of use (42%) for yield estimation of perennial crops (Fig. 4), involves the acquisition of training and testing data. These data can be obtained, similar to the case of LMs, through photographic cameras or remote and proximal sensing methods, allowing for the direct acquisition of parameters that are closely associated with orchard yield. Parameters, such as the number of fruits per tree, fruit size, and canopy area of trees, are commonly employed as direct variables in the development of SNLM [20].

In a related study, Rahman et al. [58] explored the potential of high-resolution satellite images, captured throughout the fruit growth period, for estimating the yield of a mango orchard. They utilized an Artificial Neural Network (ANN) model and integrated vegetation indices (VI) with geographic information. The results of their study demonstrated yield estimation results with accuracy superior to 93%, considering the optimal simulation parameters.

Similarly, Črtomir et al. [23] employed an ANN to estimate the yield of an apple orchard using a single image per tree canopy. The findings of their study demonstrated the effectiveness of using parameters derived from orchard images in conjunction with SNLM for improved yield estimation. Specifically, these parameters proved to be efficient in providing more accurate yield estimates during a specific stage of development, preferably toward the end of the maturation period. However, it should be noted that the performance of SNLMs may reduce when it comes to allowing management interventions in the orchard during the current harvest season. If SNLMs could efficiently perform yield estimation tasks during earlier stages of fruit development, it could present an opportunity to increase orchard productivity.

Due to the need to explore more adaptable and efficient tools for potential operationalization of computational models in PFG [67], a portion corresponding to 26% of the works reviewed here used computer vision methods and different architectures of deep nonlinear models (DNLMs) for yield estimation (Fig. 4). Convolutional Neural Networks (CNN) showed promising results for image classification, particularly in fruit detection and counting. Santos and Gebler [63] developed a methodology that incorporates CNN algorithms based on multi-view geometry to enable fruit tracking. This methodology not only avoids double counting but also locates fruits in 3-D space, thereby facilitating subsequent yield estimation. Consequently, DNLMs offer even greater advantages compared to SNLMs, as they possess the capability to automatically learn features from raw image data [53].

Chen et al. [19] developed a Faster R-CNN model for strawberry yield estimation. This model demonstrated the capability to detect flowers, immature fruits, and mature fruits using RGB images captured by unmanned aerial vehicles (UAVs). The accuracy of the results varied between 0.76 and 0.91 for images with a resolution of 2 m. However, the model's performance in detecting immature fruits fell short of expectations, as it occasionally misidentified certain immature strawberries as dead leaves. Additionally, the authors noted challenges arising from fruit occlusion caused by overlapping leaves.

Compared to the Faster R-CNN model, the YOLO-V3 model integrates target detection and classification into a regression problem [68]. Tian et al. [68] obtained favorable results in detecting apples at various development stages (young apples, growing apples, and ripe apples) using a dense YOLO-V3 network. The authors demonstrated the YOLO-V3 capability to identify fruits under challenging conditions, such as overlapping fruits or fruits hidden by leaves and branches. Additionally, they were able to track fruit production throughout the growth stages. It is important to note that not all immature fruits will mature and contribute to the final yield due to potential losses during development. These findings exhibit promising advancements in fruit detection and classification throughout time, enabling more accurate estimation of the orchard's ultimate yield.

Yield Forecasting Modeling

The works related to yield forecasting frequently utilized the SNLM category (41%), with particular focus on Artificial Neural Networks (ANN) and Random Forests (RF). The LM category (35%) followed closely, as depicted in Fig. 5. Both categories fall under the umbrella of machine learning techniques. Based on articles published between 2000 and 2016, the prevailing research trend in Precision Agriculture is connected to the term "networks" [54]. This trend indicates a rising interest in the advancement of stochastic models based on machine learning techniques, which aligns with the findings of this study.

The main interest when using stochastic models is to describe the system's response, with little emphasis on understanding the underlying mechanisms. Invariably, due to their advantages in uncovering rules and patterns in large datasets [76], stochastic models have been widely implemented in PFG. Linear and nonlinear stochastic models have been applied to forecast yields in various perennial crops, including apple, grape, mango, and citrus [28, 36, 47].

In contrast, mechanistic models, which aim to explain why and how phenomena occur, have been less commonly used (24%) (Fig. 5). Developing mechanistic models involves representing and reasoning about the underlying nature of the process [15]. They are based on scientific principles rather than statistical convenience [17] and can

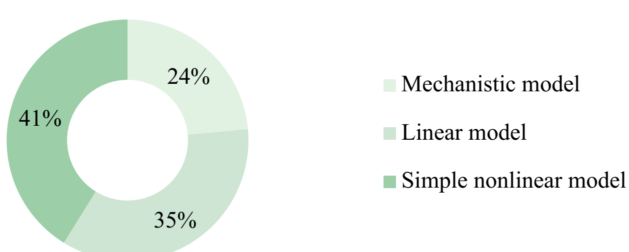

Fig. 5 Percentage of model class for yield forecasting in the reviewed literature

complement studies that rely on empirical approaches [16, 50].

Although mechanistic models are less conventional, they have gained popularity in PFG when the goal is to understand and establish the fundamental mechanisms underlying crop yield. These models provide a comprehensive understanding of fruit formation and how the environment interacts with the fruit growth development process [2, 22]. However, the process of mechanistic modeling is limited to the existing knowledge derived from previous research, which may rely on general interactions rather than accounting for cultivar-specific and location-specific interactions [42].

Moreover, mechanistic models pose challenges due to their high complexity. They require the specification and calibration of numerous parameters, including soil characterization, climate factors, and nutrient cycling [70]. Efficient data collection and minimizing parameter specifications are essential for operationalizing mechanistic models. Additionally, constructing an adequate model solution requires addressing aspects of the system that are still poorly understood and subject to questioning, particularly regarding the reasonableness of the formulated hypotheses.

Mechanistic Modeling Strategy: Focus on Phenology

The WOFOST model was optimized by Bai et al. [8] to forecast the yield of jujube (*Zizyphus jujuba*) orchards of different ages by incorporating the total dry weight (TDW) of new organs, including initial buds and roots. The results indicated that the age of the orchard is a crucial parameter for accurately predicting the yield of these fruit trees. Additionally, WOFOST effectively modeled the phenological development stages and predicted ripe fruits 2–3 days earlier than observed in the field. However, the researchers solely considered the influence of temperature on the phenological stages of jujube development. When evaluating other perennial species, it is important to consider factors such as day length (sunshine hours) and other variables to enhance the accuracy of yield forecasting [37].

There are some models available in the form of "R" packages that simulate the yield-related mechanisms, enabling the calculation and comparison of indices that model the biological processes of trees. For instance, the **Fruclimadapt** package was developed for temperate fruit species and can be applied to calculate bioclimatic indices for assessing potential risks to fruit quality and final yield [51]. This package facilitates the evaluation of long climate series, since there is no limit (aside from computational constraints) on the number of years that can be assessed in a single function run. On the other hand, the **ChillModels** package enables the calculation of multiple models to quantify the accumulation of chill and heat requirements for the development of temperate fruit species. This package acknowledges the importance of studying the seasonal effects on tree yield through climatic variables in terms of heat or chill units [57]. Similar to the study conducted by Bai et al. [8], both R packages allow the prediction of phenological stages and consider temperature as the most influential factor in regulating bud development.

Phenological stages, strongly influenced by climate, hold significant relevance in yield forecasting. They can be regarded as a valuable set of input variables for forecasting models, as each growth phase directly and indirectly impacts crop yield, particularly during maturation stages [34]. However, determining the precise timing of each phenological stage for perennial fruit crops can introduce uncertainties due to the inherent variability associated with these crops. As a result, there is no universally applicable pattern that instills confidence among producers.

Mechanistic Modeling Strategy: Focus on Productivity

As well as the phenological stages, the annual yield varies among different perennial crops [51]. Consequently, some studies have decomposed the yield forecasting modeling problem, focusing on the effects of specific input parameters on the other components involved. For instance, in the case of using the STICS model to forecast the yield of apple orchards, it becomes possible to define the main involved parameters, which include agricultural practices, soil type, climate conditions, and ecosystem services [25]. The STICS model is considered a mechanistic modeling approach, able to simulate the water, C, and N balances of various types of crops, both annual and perennial, herbaceous, and woody. Thus, the model may be able to consider the complexity of an orchard by assessing the interactions of these parameters with the final yield.

In brief, the input variables most used for mechanistic modeling in yield forecasting are climatic indicators focused on summarized data for a specific area, such as daily or monthly averages [8, 10, 49]. Additionally, specific data from orchards, such as physiological characteristics and vegetation indexes synthesized into indicators of phenological stages, are frequently utilized [16, 69]. Moreover, authors often incorporate topographical information and model-specific calibration parameters [22]. Furthermore, due to the interannual yield variability in perennial crops, including plant respiration as an input variable in mechanistic yield modeling can be advantageous [49].

Mechanistic models are developed based on the probable mechanisms that govern tree behavior and deductions of their consequences through the model. In the case of perennial crops, plant physiology has a significantly greater impact on forecasting results compared to plant nutrition or soil fertility, which contrasts with the common rules in Precision Agriculture applied to annual crops. Therefore, mechanistic models are well suited for yield forecasting as they can prioritize the important components of fruit development and quantify their effects on the final yield of the crop.

Stochastic Modeling Strategy

Yield forecasting takes into account that the explanatory variables exhibit characteristics that change over time. This implies that the variables' characteristics may differ between the time of data collection and the forecasting time horizon. Consequently, the relationships between the input variables and the predicted outcome are nonlinear and demonstrate stochastic characteristics, presenting challenges for the application of traditional statistical models [64]. For example, Sakai et al. [61] conducted a study using yield data from 48 citrus trees collected over 7 years to forecast the yield for the following season. Despite working with a relatively limited dataset, their findings provided clear evidence of the nonlinear nature of the data, particularly in relation to the biennial bearing phenomenon observed in fruit crops.

The complex nature of yield interference mechanisms in perennial fruit trees adds another layer of complexity to the system. This complexity, in turn, favors the use of stochastic SNLMs, as they can yield accurate results even without complete knowledge of the variable relationships and possess the capability to adapt to nonlinear data [10].

However, there are a large number of works that address the issue of yield forecasting from the perspective of LMs [1, 14, 18, 75]. Some authors justified the preference for LMs due to their greater ease of interpretation compared to black-box models [60]. However, the indirect variables commonly used as input to yield forecasting modeling have multiple characteristics and tend to be nonlinear, such as VI and climatic data [58]. Due to this, it is possible that LMs are very restricted models and do not correctly represent the most complex mathematical relationships of the system [74].

Arab et al. [7] demonstrated that ANN (with one hidden layer) trained with Normalized Difference Vegetation Index (NDVI) provided promising results for vineyard yield forecasting. The Mean Absolute Error (MAE) obtained with the ANN was 1.4 ton/ha, which was less than those obtained with the conventional regression models (MAE = 2 ton/ha). Moreover, Random Forest models demonstrated good accuracy for apple orchard yield forecasting on a regional scale, when the models were developed based on climatic data and NDVI from two consecutive years [10]. Based on these results, it was possible to admit that more complex models, such as ANN and Random Forest, are more suitable for yield modeling, because they allow the adjustment of data in a nonlinear way, even if the modeling objective is explanatory [40, 73].

The use of VI as input variables for yield forecasting has some limitations, because the correlations between these two parameters vary greatly depending on the period of image acquisition, which is related to the stage of crop development [71]. Ye et al. [75] observed that the yield forecasting models for a citrus orchard performed better when the training image acquisition occurred in May. During this month, new leaves were growing at a faster rate, and it was possible to differentiate trees with a greater number of new leaves in the spectral images, which is directly related to the yield. Therefore, identifying the optimal period for data acquisition is essential to achieve good accuracy in yield forecasting.

Furthermore, the optimal period for image data acquisition can be determined by considering the date when a yield parameter is no longer evolving [42]. However, it may not be easy to identify this key moment in the field, especially when the vegetative stages of trees occur asynchronously in time and space [66, 72]. This complexity is further compounded by the lack of consensus in the literature regarding the definition of these stages, especially for perennial crops cultivated in regions with limiting climatic conditions. In such regions, phenological stages and the corresponding periods of fruit development exhibit site-specific temporal variability. Consequently, it is not possible to establish a consistent phenological date pattern across years or cultivars that could aid in determining the optimal period for image acquisition.

Even though the literature indicates positive correlations between VIs and the yield of perennial crops, using VIs as sole input variables in forecasting models may lead to errors. Orchards are subject to various anthropic interferences that are not adequately captured when relying solely on parameters such as VIs in the development of yield forecasting models [59]. In perennial crops, yield is closely linked to different parameters, including both more stable variables such as soil and chemical properties, and less-stable variables like physiological parameters (e.g., number of leaves and tree vigor), management practices, and climatic parameters. These parameters are interconnected and contribute to yield variability. Li et al. [44] reported that common climatic variables, such as precipitation and sunshine hours, could have a greater impact on final apple yield compared to extreme climatic parameters (e.g., frost days, heat damage). Thus, to achieve efficient yield forecasting, it is essential to consider the temporal dynamics of environmental influences and orchard management across multiple growing years [26].

For instance, collecting data over consecutive years may enable the identification of patterns of variability that often coincide with fluctuations in climate and orchard management practices [42]. Incorporating input variables in the form of accumulated values of vegetation indices (VIs) throughout the various phenological stages across multiple growing seasons of perennial orchards may be representative of the correlation between these variables and final crop yield, resulting in improved accuracy of yield forecasting models [10].

The use of variables associated with physical properties and soil fertility in yield forecasting through SNLMs has been minimally explored, with only two publications reporting on this topic [55, 66]. This limited exploration can be attributed, in part, to the stability of these parameters within the considered input data time window of the reviewed documents. Furthermore, unlike annual crops, managing soil-related variables in perennial crops is challenging as extensive soil adjustment activities are not feasible. Despite these parameters may influence final yield, their management and correction options are restricted, which explains their reduced use as input data for perennial yield forecasting. However, although these variables may not be particularly useful for guiding actions to increase productivity, they can still be employed as parameters for initial forecasting model calibration, providing an overview of the area's conditions.

In the study conducted by Papageorgiou et al. [55], eight machine learning methods were proposed for yield forecasting of an apple orchard across three categories: low, medium, and high yield. These methods, including fuzzy cognitive maps, MLP ANN, Naive Bayes, K-means clustering, Decision trees, Recurrent Neural Networks (RNN), and association rules, were based on parameters of soil properties. Among the various machine learning models, the fuzzy cognitive maps model outperformed the others, correctly categorizing the yield in 42 out of the 56 analyzed cases, resulting in an accuracy of 75%. However, it is important to note that the accuracy of the results is limited to the training dataset, as the authors did not utilize a verification set to test the model.

The variables most used for yield forecasting through SNLMs include weather variables (temperature, precipitation, photoperiod, and solar radiation intensity) [41], physiological variables (such as orchard age, tree density, and tree spacing) [40], and VIs [18, 74]. Some studies also considered geographic variables (such as latitude and longitude) [18, 40], as well as soil properties (such as electrical conductivity and organic matter content) [55]. In general, climate-physiological data tend to be more suitable for yield forecasting, because it better represents the variation in perennial fruit yield compared to soil data.

However, despite the significant progress in the development of SNLMs for yield forecasting, these models have inherent limitations. They heavily rely on the quality and quantity of the data used for training the models. Datasets with high levels of noise, incompleteness, or outliers can significantly reduce the performance of SNLMs [28]. To overcome these limitations, strategies such as incorporating expert knowledge to address collinearity, identifying outliers, and using cross-validation techniques can be implemented [21, 47].

Temporal Data Scale Challenges in Perennial Fruit Crop Modeling: Insights and Future Directions

Yield Estimation Models

The accuracy of yield estimation models relies on site-specific data, because they are typically based on direct variables, such as the number of flowers, number of fruits, fruit color, and canopy area. However, collecting local data for each site can be resource-intensive and time-consuming, requiring sufficient allocation of resources and comprehensive data collection and pre-processing efforts [45]. To overcome these challenges, recent studies have explored image-based analysis techniques, such as computer vision, to facilitate data collection and improve the quality of yield estimation methods. However, the reliability of image-based estimation can be affected by factors, such as lighting conditions, occlusion, and variations in fruit appearance. Ongoing research aims to enhance image capturing and develop more robust DNLMs, such as multi-view imaging and 3-D reconstruction [63, 68].

One of the complexities in developing yield estimation models arises when extrapolating results from one year to another. While the relationship between the number of flowers and fruits may be established, it may not be applicable for extrapolations beyond the period of data collection, especially in regions with year-to-year crop yield variability. Factors, such as alternate bearing and canopy changes, make it difficult, if not impossible, to create a general model that provides accurate results for different years using data from just one or two harvest seasons of a fruit crop [12]. Therefore, it is crucial to examine the transferability and generalizability of models to ensure accurate estimation results in diverse agricultural settings.

Yield Forecasting Models

Yield development is a dynamic process that is influenced by both endogenous and exogenous factors, which act on the crop throughout its trajectory and can include the integration of previous influences [42]. The complex nature of trajectory effects and accumulated influences has been suggested by studies analyzing crop yield behavior over multiple seasons [48, 61]. These findings highlight the importance of considering time-series data, as it can provide valuable insights into yield development by capturing the effects of its trajectory.

However, only a few studies have explicitly addressed the significance of temporal variability in their assessments, which can be achieved by incorporating intra-seasonal variables into the modeling strategy. Climatic variables and vegetation indices (VI) may have the potential to offer chronological data on crop development. Conversely, most of the existing literature has primarily focused on point indicators based on specific phenological stages or seasonal time stages, such as the average temperature during the flowering period [41]. These indicators are often treated as independent variables when analyzed using classical methods like linear regression [32].

In addition, an analysis of the temporal database among the reviewed documents revealed that only 11% of the studies applied historical records spanning more than 10 years for yield forecasting modeling (Fig. 6). In contrast, most studies (47%) relied on data from a maximum of three previous seasons, which may provide an insufficient time-series to adequately assess the temporal variability of perennial crops' yields. Consequently, this limited timeframe may not fully capture the effects of past influences on the current season's yield.

Since the yield of perennial crops can exhibit significant year-to-year variation, the development of accurate machine learning models for forecasting alternate bearing of fruit crop yield remains unknown. This task requires high-quality yield data spanning multiple years, enabling the establishment of causal relationships between yield and environmental variables or management practices to develop robust yield forecasting models [18, 47].

Also, the analysis of time-series data used in the reviewed literature revealed that there is a lack of investigation on long-term analysis for early yield forecasting. Long-term analysis means that the forecasting of crop yield for season '*n*' has been conducted without incorporating any input variables from the current year '*n*'. Otherwise, relying solely on input data from the current crop season (*n*) can limit the

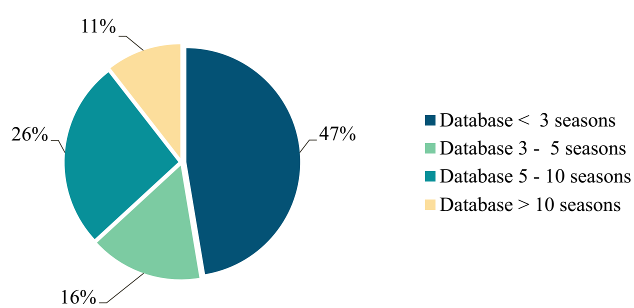

Fig. 6 An overview of the temporal scale relate to the input database used for yield forecasting modeling in the reviewed documents

forecasting horizon and reduce the time window for implementing corrective actions in crop management practices and decision-making processes, as the orchard production may already be established. Therefore, an ideal set of input data for long-term yield forecasting that provide actual early yield forecasting should include aggregated variables from previous seasons ($n-1$, $n-2$, $n-3$, and so on).

In a recent report, Brinkhoff and Robson [18] proposed a model for long-term yield forecasting in perennial crops, conducted with exclusively historical data from previous seasons (2014–2019). The authors successfully forecasted macadamia orchard yields for different regions within a specific time window, providing producers and the industry with valuable information for their decision-making processes. Thus, employing newer or more advanced machine learning models that incorporate time-series analysis can enhance the information derived from the data and contribute to the development of more robust long-term yield forecasting models with greater operational advantages.

Conclusions

In the context of PFG**,** the field of yield modeling for perennial crops through computational models is still in its early stages. This literature survey provides an overview of yield assessment modeling for perennial fruit crops and the commonly applied estimation and forecasting modeling strategies.

For yield estimation, LMs have been widely used and have yielded interesting results in the assessment of final crop yield based on input data from the current crop harvest season, such as the number of flowers and fruits. One advantage of LMs is their simplicity and interpretability; however, there are limitations when dealing with complex nonlinear relationships of perennial crop yield and spatio-temporal characteristics. In contrast, DNLMs can capture complex nonlinear patterns and interactions among explanatory variables, allowing for more accurate yield estimation. These models can automatically process large amounts of data and extract detailed crop characteristics from high-resolution images, such as leaf area, vegetation indices, and canopy structure.

On the other hand, the forecasting yield assessment typically refers to the prediction of the final yield within the same location but at a future time. Mechanistic modeling strategies enable the prioritization of crucial fruit development components and the quantification of their impact on the crop's final yield. However, the mechanistic approach requires specific calibration parameters that increase the modeling complexity. To overcome this issue, forecasting modeling has been frequently conducted by SNLMs. The SNLMs incorporate stochastic processes and statistical techniques to model the relationships between yield data and relevant influencing factors over time, such as weather and agronomic conditions. By accounting for the temporal dynamics and randomness in perennial crop yield, SNLMs may capture complex patterns of perennial crop development.

Choosing the yield assessment modeling strategy depends on the specific requirements of the application, the availability of data, and the expected assessment time horizon. Combined data from different locations can provide valuable information for estimation or forecasting modeling from a more generic perspective. However, it is equally important to retain the ability to interpret the results on a local scale, ensuring that the models can be effectively utilized from an agronomic perspective and applied in on-farm management practices. Moreover, while perennial crop yield is influenced by spatial and temporal variability, further research is needed to identify the main explanatory variables and establish consistent datasets over time and space.

Author Contributions LG has the idea for the article, contributed to the study conception and design, and critically revised the work. SAMA has critically revised the work and assisted with financial support. RBM performed part of the literature search, data collection and analysis, and drafted the work. All authors read and approved the final manuscript.

Declarations

Conflict of Interest The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abdulai I, Hoffmann MP, Jassogne L, Asare R, Graefe S, Tao H-H, Muilerman S, Vaast P, Van Asten P, Läderach P, Rötter RP. Variations in yield gaps of smallholder cocoa systems and the main determining factors along a climate gradient in Ghana. Agric Syst. 2020;181:102812. https://doi.org/10.1016/j.agsy.2020.102812.
2. Acock B, Pachepsky YA. Holes in precision farming: mechanistic crop models. In: Precision agriculture. Stanfford Journal; 1997. pp. 397–404.
3. Anderson NT, Walsh KB, Wulfsohn D. Technologies for forecasting tree fruit load and harvest timing—from ground, sky and time. Agronomy. 2021;11(7):1409. https://doi.org/10.3390/agronomy11071409.
4. Aggelopoulou AD, Bochtis D, Fountas S, Swain KC, Gemtos TA, Nanos GD. Yield prediction in apple orchards based on image processing. Precis Agric. 2010;12(3):448–56. https://doi.org/10.1007/s11119-010-9187-0.
5. Aggelopoulou KD, Wulfsohn D, Fountas S, Gemtos TA, Nanos GD, Blackmore S. Spatial variation in yield and quality in a

small apple orchard. Precis Agric. 2009;11(5):538–56. https://doi.org/10.1007/s11119-009-9146-9.
6. Anastasiou E, Balafoutis A, Darra N, Psiroukis V, Biniari A, Xanthopoulos G, Fountas S. Satellite and proximal sensing to estimate the yield and quality of table grapes. Agriculture. 2018;8(7):94. https://doi.org/10.3390/agriculture8070094.
7. Arab ST, Noguchi R, Matsushita S, Ahamed T. Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach. Remote Sens Appl Soc Environ. 2021;22:100485. https://doi.org/10.1016/j.rsase.2021.100485.
8. Bai T, Wang T, Zhang N, Chen Y, Mercatoris B. Growth simulation and yield prediction for perennial jujube fruit tree by integrating age into the WOFOST model. J Integr Agric. 2020;19(3):721–34. https://doi.org/10.1016/s2095-3119(19)62753-x.
9. Bai T, Zhang N, Mercatoris B, Chen Y. Jujube yield prediction method combining Landsat 8 Vegetation Index and the phenological length. Comput Electron Agric. 2019;162:1011–27. https://doi.org/10.1016/j.compag.2019.05.035.
10. Bai X, Li Z, Li W, Zhao Y, Li M, Chen H, Wei S, Jiang Y, Yang G, Zhu X. Comparison of machine-learning and CASA models for predicting apple fruit yields from time-series planet imageries. Remote Sens. 2021;13(16):3073. https://doi.org/10.3390/rs13163073.
11. Baluja J, Diago MP, Goovaerts P, Tardaguila J. Spatio-temporal dynamics of grape anthocyanin accumulation in a Tempranillo vineyard monitored by proximal sensing. Aust J Grape Wine Res. 2012;18(2):173–82. https://doi.org/10.1111/j.1755-0238.2012.00186.x.
12. Ballesteros R, Intrigliolo DS, Ortega JF, Ramírez-Cuesta JM, Buesa I, Moreno MA. Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. Precis Agric. 2020;21(6):1242–62. https://doi.org/10.1007/s11119-020-09717-3.
13. Beattie BB, Folley RRW. Production variability in apple crops. Sci Hortic. 1977;6(4):271–9. https://doi.org/10.1016/0304-4238(77)90084-x.
14. Beattie BB, Folley RRW. Production variability in apple crops. II. The long-term behaviour of the English crop. Sci Hortic. 1978;8(4):325–32. https://doi.org/10.1016/0304-4238(78)90054-7.
15. Bechtel W, Abrahamsen A. Explanation: a mechanist alternative. Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci. 2005;36(2):421–41. https://doi.org/10.1016/j.shpsc.2005.03.010.
16. Bregaglio S, Fischer K, Ginaldi F, Valeriano T, Giustarini L. The HADES yield prediction system—a case study on the Turkish hazelnut sector. Front Plant Sci. 2021. https://doi.org/10.3389/fpls.2021.665471.
17. Bretó C, He D, Ionides EL, King AA. Time series analysis via mechanistic models. Ann Appl Stat. 2009. https://doi.org/10.1214/08-aoas201.
18. Brinkhoff J, Robson AJ. Block-level macadamia yield forecasting using spatio-temporal datasets. Agric For Meteorol. 2021;303:108369. https://doi.org/10.1016/j.agrformet.2021.108369.
19. Chen Y, Lee WS, Gan H, Peres N, Fraisse C, Zhang Y, He Y. Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. Remote Sens. 2019;11(13):1584. https://doi.org/10.3390/rs11131584.
20. Cheng H, Damerow L, Sun Y, Blanke M. Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. J Imaging. 2017;3(1):6. https://doi.org/10.3390/jimaging3010006.
21. Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput Electron Agric. 2018;151:61–9. https://doi.org/10.1016/j.compag.2018.05.012.
22. Cola G, Mariani L, Salinari F, Civardi S, Bernizzoni F, Gatti M, Poni S. Description and testing of a weather-based model for predicting phenology, canopy development and source–sink balance in *Vitis vinifera* L. cv. Barbera. Agric For Meteorol. 2014;184:117–36. https://doi.org/10.1016/j.agrformet.2013.09.008.
23. Črtomir R, Urška C, Stanislav T, Denis S, Karmen P, Pavlovič M, Marjan V. Application of neural networks and image visualization for early forecast of apple yield. Erwerbs-obstbau. 2012;54(2):69–76. https://doi.org/10.1007/s10341-012-0162-y.
24. Dambreville A, Lauri P-É, Trottier C, Guédon Y, Normand F. Deciphering structural and temporal interplays during the architectural development of mango trees. J Exp Bot. 2013;64(8):2467–80. https://doi.org/10.1093/jxb/ert105.
25. Demestihas C, Plénet D, Génard M, Garcia de Cortazar-Atauri I, Launay M, Ripoche D, Beaudoin N, Simon S, Charreyron M, Raynal C, Lescourret F. Analyzing ecosystem services in apple orchards using the STICS model. Eur J Agron. 2018;94:108–19. https://doi.org/10.1016/j.eja.2018.01.009.
26. DeVetter L, Colquhoun J, Zalapa J, Harbut R. Yield estimation in commercial cranberry systems using physiological, environmental, and genetic variables. Sci Hortic. 2015;190:83–93. https://doi.org/10.1016/j.scienta.2015.04.016.
27. Dorj U-O, Lee M, Yun S. An yield estimation in citrus orchards via fruit detection and counting using image processing. Comput Electron Agric. 2017;140:103–12. https://doi.org/10.1016/j.compag.2017.05.019.
28. Elavarasan D, Vincent DR, Sharma V, Zomaya AY, Srinivasan K. Forecasting yield by integrating agrarian factors and machine learning models: a survey. Comput Electron Agric. 2018;155:257–82. https://doi.org/10.1016/j.compag.2018.10.024.
29. Fioravanço JC, Czermainski ABC. Biennial bearing in apple cultivars. Rev Ceres. 2018;65(2):144–9. https://doi.org/10.1590/0034-737x201865020005.
30. Fisher D. Understanding technology adoption through system dynamics modeling: implications for agribusiness management. Int Food Agribus Manag Rev. 2000;3(3):281–96. https://doi.org/10.1016/s1096-7508(01)00048-9.
31. Gebler L, Greco CR, Vieira AL, Kuse LR. Spatial influence of physical and chemical parameters on management zone definition in apple orchards. Engenharia Agrícola, Jaboticabal. 2015;35:1160–71. https://doi.org/10.1590/1809-4430-Eng.Agric.v35n6p1160-1171/2015.
32. Guilpart N, Metay A, Gary C. Grapevine bud fertility and number of berries per bunch are determined by water and nitrogen stress around flowering in the previous year. Eur J Agron. 2014;54:9–20. https://doi.org/10.1016/j.eja.2013.11.002.
33. Guitton B, Kelner J-J, Velasco R, Gardiner SE, Chagné D, Costes E. Genetic control of biennial bearing in apple. J Exp Bot. 2011;63(1):131–49. https://doi.org/10.1093/jxb/err261.
34. Guo Y, Fu Y, Hao F, Zhang X, Wu W, Jin X, Robin Bryant C, Senthilnath J. Integrated phenology and climate in rice yields prediction using machine learning methods. Ecol Indic. 2021;120:106935. https://doi.org/10.1016/j.ecolind.2020.106935.
35. Hastings A. Transients: the key to long-term ecological understanding? Trends Ecol Evol. 2004;19:39–45. https://doi.org/10.1016/j.tree.2003.09.007.
36. He L, Fang W, Zhao G, Wu Z, Fu L, Li R, Majeed Y, Dhupia J. Fruit yield prediction and estimation in orchards: a state-of-the-art comprehensive review for both direct and indirect methods. Comput Electron Agric. 2022;195:106812. https://doi.org/10.1016/j.compag.2022.106812.

37. Holzapfel CB, Lafond GP, Brandt SA, Bullock PR, Irvine RB, Morrison MJ, May WE, James DC. Estimating canola (*Brassica napus* L.) yield potential using an active optical sensor. Can J Plant Sci. 2009;89(6):1149–60. https://doi.org/10.4141/cjps09056.
38. Jiménez CM, Díaz JBR. Statistical model estimates potential yields in pear cultivars 'blanquilla' and 'conference' before bloom. J Am Soc Hortic Sci. 2003;128(4):452–7. https://doi.org/10.21273/jashs.128.4.0452.
39. Jiménez CM, Díaz JBR. Statistical model estimates potential yields in 'golden delicious' and 'royal gala' apples before bloom. J Am Soc Hortic Sci. 2004;129(1):20–5. https://doi.org/10.21273/jashs.129.1.0020.
40. Jin Y, Chen B, Lampinen BD, Brown PH. Advancing agricultural production with machine learning analytics: yield determinants for California's almond orchards. Front Plant Sci. 2020. https://doi.org/10.3389/fpls.2020.00290.
41. Keller M, Molitor D. Yield of Müller—Thurgau and Riesling grapevines is altered by meteorological conditions in the current and previous growing seasons. OENO One. 2017. https://doi.org/10.20870/oeno-one.2016.50.4.1071.
42. Laurent C, Oger B, Taylor JA, Scholasch T, Metay A, Tisseyre B. A review of the issues, methods and perspectives for yield estimation, prediction and forecasting in viticulture. Eur J Agron. 2021;130:126339. https://doi.org/10.1016/j.eja.2021.126339.
43. Lee MA, Monteiro A, Barclay A, Marcar J, Miteva-Neagu M, Parker J. A framework for predicting soft-fruit yields and phenology using embedded, networked microsensors, coupled weather models and machine-learning techniques. Comput Electron Agric. 2020;168:105103. https://doi.org/10.1016/j.compag.2019.105103.
44. Li M, Guo J, He J, Xu C, Li J, Mi C, Tao S. Possible impact of climate change on apple yield in Northwest China. Theor Appl Climatol. 2019;139(1–2):191–203. https://doi.org/10.1007/s00704-019-02965-y.
45. Li D, Song Z, Quan C, Xu X, Liu C. Recent advances in image fusion technology in agriculture. Comput Electron Agric. 2021;191:106491. https://doi.org/10.1016/j.compag.2021.106491.
46. Liakos V, Tagarakis A, Aggelopoulou K, Fountas S, Nanos GD, Gemtos T. In-season prediction of yield variability in an apple orchard. Eur J Hortic Sci. 2017;82(5):251–59. https://doi.org/10.17660/eJHS.2017/82.5.5
47. Logan TM, McLeod S, Guikema S. Predictive models in horticulture: a case study with Royal Gala apples. Sci Hortic. 2016;209:201–13. https://doi.org/10.1016/j.scienta.2016.06.033.
48. Lordan J, Francescatto P, Dominguez LI, Robinson TL. Long-term effects of tree density and tree shape on apple orchard performance, a 20 year study—part 1, agronomic analysis. Sci Hortic. 2018;238:303–17. https://doi.org/10.1016/j.scienta.2018.04.033.
49. Maselli F, Chiesi M, Brilli L, Moriondo M. Simulation of olive fruit yield in Tuscany through the integration of remote sensing and ground data. Ecol Model. 2012;244:1–12. https://doi.org/10.1016/j.ecolmodel.2012.06.028.
50. Minasny B, Finke P, Stockmann U, Vanwalleghem T, McBratney AB. Resolving the integral connection between pedogenesis and landscape evolution. Earth Sci Rev. 2015;150:102–20. https://doi.org/10.1016/j.earscirev.2015.07.004.
51. Miranda C, Urrestarazu J, Santesteban LG. Fruclimadapt: An R package for climate adaptation assessment of temperate fruit species. Comput Electron Agric. 2021;180:105879.
52. Monselise SP, Goldschmidt EE. Alternate bearing in fruit trees. In: Horticultural reviews. Hoboken: Wiley; 2011. p. 128–73. https://doi.org/10.1002/9781118060773.ch5.
53. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. Deep learning applications and challenges in big data analytics. J Big Data. 2015. https://doi.org/10.1186/s40537-014-0007-7.
54. Pallottino F, Biocca M, Nardi P, Figorilli S, Menesatti P, Costa C. Science mapping approach to analyze the research evolution on precision agriculture: world, EU and Italian situation. Precis Agric. 2018;19(6):1011–26. https://doi.org/10.1007/s11119-018-9569-2.
55. Papageorgiou EI, Aggelopoulou KD, Gemtos TA, Nanos GD. Yield prediction in apples using Fuzzy cognitive map learning approach. Comput Electron Agric. 2013;91:19–29. https://doi.org/10.1016/j.compag.2012.11.008.
56. Payne AB, Walsh KB, Subedi PP, Jarvis D. Estimation of mango crop yield using image analysis—segmentation method. Comput Electron Agric. 2013;91:57–64. https://doi.org/10.1016/j.compag.2012.11.009.
57. Pertille RH, Sachet MR, Guerrezi MT, Citadin I. An R package to quantify different chilling and heat models for temperate fruit trees. Comput Electron Agric. 2019;167:105067.
58. Rahman M, Robson A, Bristow M. Exploring the potential of high resolution worldview-3 imagery for estimating yield of mango. Remote Sens. 2018;10(12):1866. https://doi.org/10.3390/rs10121866.
59. Robson A, Rahman M, Muir J. Using worldview satellite imagery to map yield in avocado (*Persea americana*): a case study in Bundaberg, Australia. Remote Sens. 2017;9(12):1223. https://doi.org/10.3390/rs9121223.
60. Roznik M, Boyd M, Porth L. Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks. Remote Sens Appl Soc Environ. 2022;25:100693. https://doi.org/10.1016/j.rsase.2022.100693.
61. Sakai K, Noguchi Y, Asada S. Detecting chaos in a citrus orchard: reconstruction of nonlinear dynamics from very short ecological time series. Chaos Solitons Fract. 2008;38(5):1274–82. https://doi.org/10.1016/j.chaos.2007.01.144.
62. Salvo S, Muñoz C, Ávila J, Bustos J, Ramírez-Valdivia M, Silva C, Vivallo G. An estimate of potential blueberry yield using regression models that relate the number of fruits to the number of flower buds and to climatic variables. Sci Hortic. 2012;133:56–63. https://doi.org/10.1016/j.scienta.2011.10.020.
63. Santos T, Gebler L. A methodology for detection and localization of fruits in apples orchards from aerial images. In: Anais do XIII Congresso Brasileiro de Agroinformática. Porto Alegre: SBC; 2021. pp. 1–9. https://doi.org/10.5753/sbiagro.2021.18369
64. Sarron J, Malézieux É, Sané C, Faye É. Mango yield mapping at the orchard scale based on tree structure and land cover assessed by UAV. Remote Sens. 2018;10(12):1900. https://doi.org/10.3390/rs10121900.
65. Sartore L, Rosales AN, Johnson DM, Spiegelman CH. Assessing machine leaning algorithms on crop yield forecasts using functional covariates derived from remotely sensed data. Comput Electron Agric. 2022;194:106704. https://doi.org/10.1016/j.compag.2022.106704.
66. Sirsat MS, Mendes-Moreira J, Ferreira C, Cunha M. Machine Learning predictive model of grapevine yield based on agroclimatic patterns. Eng Agric Environ Food. 2019;12(4):443–50. https://doi.org/10.1016/j.eaef.2019.07.003.
67. Stajnko D, Cmelik Z. Modelling of apple fruit growth by application of image analysis. Agric Conspec Sci. 2005;70(2):59–64.
68. Tian Y, Yang G, Wang Z, Wang H, Li E, Liang Z. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. Comput Electron Agric. 2019;157:417–26. https://doi.org/10.1016/j.compag.2019.01.012.
69. Valdés-Gómez H, Celette F, García de Cortázar-Atauri I, Jara-Rojas F, Ortega-Farías S, Gary C. Modelling soil water content and grapevine growth and development with the stics crop-soil model under two different water management strategies. OENO One. 2009;43(1):13. https://doi.org/10.20870/oeno-one.2009.43.1.806.

70. Van Alphen BJ, Stoorvogel JJ. A methodology to define management units in support of an integrated, model-based approach to precision agriculture. In: Robert PC, Rust RH, Larson WE, editors, Proceedings of the fourth international conference on precision agriculture. American Society of Agronomy; 1999. pp. 1267–1278. https://doi.org/10.2134/1999.precisionagproc4.c30b.
71. Van Beek J, Tits L, Somers B, Deckers T, Verjans W, Bylemans D, Janssens P, Coppin P. Temporal dependency of yield and quality estimation through spectral vegetation indices in pear orchards. Remote Sens. 2015;7(8):9886–903. https://doi.org/10.3390/rs70809886.
72. Verdugo-Vásquez N, Acevedo-Opazo C, Valdés-Gómez H, et al. Towards an empirical model to estimate the spatial variability of grapevine phenology at the within field scale. Precision Agric. 2020;21:107–30. https://doi.org/10.1007/s11119-019-09657-7.
73. Yazdanpanah H, Ohadi D, Tabar M. Forecasting different phenological phases of apple using artificial neural network. J Res Agric Sci. 2010;6(2):97–106.
74. Ye X, Sakai K, Asada S-I, Akita T, Sasao A. Airborne hyperspectral imaging for investigating the dynamics of alternate bearing in citrus. Agric Inf Res. 2005;14(4):261–72. https://doi.org/10.3173/air.14.261.
75. Ye X, Sakai K, Manago M, Asada S, Sasao A. Prediction of citrus yield from airborne hyperspectral imagery. Precision Agric. 2007;8(3):111–25. https://doi.org/10.1007/s11119-007-9032-2.
76. Zhang D, Tsai JJ-P. Advances in machine learning applications in software engineering. Hershey: IGI Global; 2007.
77. Zhu Y, Yang G, Yang H, Guo L, Xu B, Li Z, Han S, Zhu X, Li Z, Jones G. Forecasting regional apple first flowering using the sequential model and gridded meteorological data with spatially optimized calibration. Comput Electron Agric. 2022;196:106869. https://doi.org/10.1016/j.compag.2022.106869.