Computers and Electronics in Agriculture 151 (2018) 61–69

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Review

# Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review

Anna Chlingaryan[a,*], Salah Sukkarieh[a], Brett Whelan[b]

[a] Australian Centre for Field Robotics, Dept. of Aerospace, Mechanical & Mechatronic Engineering, The University of Sydney, NSW 2006, Australia
[b] Centre for Carbon, Water and Food, School of Life and Environmental Sciences, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

*Keywords:*
Vegetation indices
Features extraction
Predictive modelling
Decision making
Information fusion

ABSTRACT

Accurate yield estimation and optimised nitrogen management is essential in agriculture. Remote sensing (RS) systems are being more widely used in building decision support tools for contemporary farming systems to improve yield production and nitrogen management while reducing operating costs and environmental impact. However, RS based approaches require processing of enormous amounts of remotely sensed data from different platforms and, therefore, greater attention is currently being devoted to machine learning (ML) methods. This is due to the capability of machine learning based systems to process a large number of inputs and handle non-linear tasks. This paper discusses research developments conducted within the last 15 years on machine learning based techniques for accurate crop yield prediction and nitrogen status estimation. The paper concludes that the rapid advances in sensing technologies and ML techniques will provide cost-effective and comprehensive solutions for better crop and environment state estimation and decision making. More targeted application of the sensor platforms and ML techniques, the fusion of different sensor modalities and expert knowledge, and the development of hybrid systems combining different ML and signal processing techniques are all likely to be part of precision agriculture (PA) in the near future.

## 1. Introduction

Improving crop yield production and quality while reducing operating costs and environmental pollution is a key goal in precision agriculture (PA). The potential growth and yield depends on many different production attributes such as the weather, soil properties, topography, irrigation and fertilizer management. The need for timely and accurate sensing of these inputs for large agricultural fields has led to increased adoption of remote and proximal sensing technologies (Campbell and Wynne, 2011) in PA (Curran, 1987). These sensing techniques provide acquisition of spectral, spatial and temporal information about the objects via ground-based vehicles, aircraft, satellites and handheld radiometers.

Remote sensing, such as satellite and airborne multi-spectral scanning, photography and video, enables precision weed management through the generation of timely and accurate weed maps (Lamb and Brown, 2001). Thermal remote sensing via airborne thermal imagery has the potential to identify spatial variations in crop water status (Tilling et al., 2006), which can enable improvements in the water management in irrigated cropping systems. Nowadays, this remote sensing technique is widely used in different crop species such as wheat (Gontia and Tiwari, 2008), maize (Taghvaeian et al., 2012), trees (Bellvert et al., 2016) and vineyards (Gutiérrez et al., 2018). At the same time many ground-based platforms have been developed and aimed at different PA tasks such as mapping of soil properties (Barnes et al., 2003), estimating evapotranspiration and drought stress (Maes and Steppe, 2012), weed mapping (Sui et al., 2008) and assessing crop water and nitrogen status (El-Shikha et al., 2007; Govender et al., 2009).

Remote sensing at visible and near-infrared wavelengths (vis-NIR) has been used to devise many spectral indices for estimating different vegetation properties. This includes the amount of chlorophylls and other photosynthetic/photoprotective pigments and the leaf area index (LAI) (Barati et al., 2011; Gao, 1996; Haboudane et al., 2004; Huete, 1988; Qi et al., 1994; Sims and Gamon, 2002; Viña et al., 2011; Zarco-Tejada et al., 2012). More than 100 vegetation indices along with their applicability, representativeness, environment and implementation precision have recently been reviewed by Xue and Su (2017). They concluded that for real-world applications the use of any existing vegetation indices requires careful consideration of the strengths and shortcomings of those indices and the specific environment where they will be applied. Making crop yield predictions using remotely sensed

* Corresponding author.
*E-mail address:* a.chlingaryan@acfr.usyd.edu.au (A. Chlingaryan).

https://doi.org/10.1016/j.compag.2018.05.012
Received 24 November 2017; Received in revised form 7 May 2018; Accepted 8 May 2018
0168-1699/ Published by Elsevier B.V.

vegetation indices has been attempted by Panda et al. (2010), Jaafar and Ahmad (2015) and many others.

Utilization of wireless sensors and actuators in PA, as well as algorithms for wireless sensor network data integration is now advancing (Zhou et al., 2012). Aqeel-ur-Rehman et al. (2014) presented a review on wireless sensor network technology and their applications in different aspects of agriculture and reported on existing system frameworks in PA.

Nitrogen (N) is considered by growers as a major mineral nutrient for plant growth and development because it is directly related to the photosynthetic process (Andrews et al., 2013). At the same time N has a high environmental and economic impact. Hence, the optimization of N fertilization for different crops has become a subject of many spectrometric studies (Cao et al., 2017; Chen et al., 2008; Goron et al., 2017; Lukina et al., 2001; MacKerron et al., 1993; Maresma et al., 2016; Quemada et al., 2014; Raun et al., 2005; Schepers and Raun, 2008).

The estimation of the plant N status can be divided into two main types: destructive and non-destructive. The most common method of destructive measurement is a chemical analysis which is associated with the Kjeldahl technique and is laborious, lengthy and costly (Jones and Moseley, 1993; Vigneau et al., 2011). Optical remote sensing of the plant N status is a non-destructive method based on canopy reflectance in the visible–NIR wavelengths (400–900 nm). This measurement is completed in-situ, lowering the number of field samples required and thus reducing the time and financial cost of field sample collection, preparation and laboratory analysis. Many studies have been dedicated to non-destructive measurement of the N status inferring in plants via remote sensing technology (Apostol et al., 2007; Lamb et al., 2002; Reyniers and Vrindts, 2006; Scharf et al., 2002; Tremblay et al., 2012) and spectral indices indicative of the plant N status have been derived from hyperspectral data (Chen et al., 2010; Tian et al., 2014; Yao et al., 2010).

Adoption of remote sensing in geology (Gupta, 2003), forestry (Holmgren and Thuresson, 1998; Hultquist et al., 2014), hydrology (Engman and Gurney, 1991), agriculture (Seelan et al., 2003) and other domains has led to the collection of significant volumes of data. The volume is continuously growing and it is beyond human ability to personally integrate, analyse and make the best informed decisions from the information. This is particularly the case when the data is not homogeneous, i.e. is sensed by sensors with different spatial, temporal and spectral modalities. Machine Learning (ML) is an emerging technology that can aid in the discovery of rules and patterns in large sets of data (Du and Jeffrey, 2007).

Crop yield prediction and N status estimation are considered together here because of the direct linkage in fertiliser management decisions. Crop yield goals are routinely utilised for calculating N requirements, both pre- and in-season. For devising potential site-specific management plans for N fertiliser, especially in-season, an estimation of both would be ideal. The aim of this review is to show the capability of different ML techniques to effectively handle these different but closely related tasks. A review is presented of recent studies in the area of crop yield prediction and N status estimation, which incorporate different ML techniques. It also covers comparative studies of different ML techniques as they are applied to the same task in PA. Some technical aspects of the ML techniques used in the reviewed studies are discussed.

## 2. Machine learning techniques

One of the main advantages of ML techniques is that they are capable of autonomously solving large non-linear problems using datasets from multiple (potentially interconnected) sources. Some ML techniques, such as Gaussian Processes (GPs) (Bishop, 2006; Rasmussen and Williams, 2005), Dirichlet Processes (DP) (Ferguson, 1973) and Indian Buffet Process (IBP) (Griffiths and Ghahramani, 2011) are probabilistic and enable consideration of sensor noise while conducting probabilistic fusion of information from different sensors (Castaldi et al., 2016; Dalponte et al., 2012; Pohl and Van Genderen, 1998) and providing confidence intervals for the predictions. ML enables better decision making and informed actions in real-world scenarios without (or with minimal) human intervention. ML provides a powerful and flexible framework for not only data-driven decision making but also for incorporation of expert knowledge into the system. These are some of the key characteristics of the ML techniques that make them widely used in many domains, and highly applicable to PA.

The major aim of PA in cropping systems is to provide information that will enable better decisions to be made on management across space and time (Whelan and Taylor, 2013). Specifically, information on variation in plant health and physiology, nutrient status or pest/disease burden may allow different treatments or treatment intensity to be applied to specific areas of crop. The practical identification and segregation of areas is commonly achieved by dividing a large field area into smaller management zones with identified requirements for different treatments. Conventionally, such delineation is based on maps of the crop field variability derived from soil and yield measurements. Nawar et al. (2017) provided a comprehensive review on management zone delineation approaches for PA applications. They illustrated how recent developments in sensing technologies, geostatistical analysis, data fusion and interpolation techniques have improved precision and reliability of management zone delineation, making it a viable strategy in commercial agriculture. They also compared traditional with advanced sensing technologies for delineating management zones.

Recently, Pantazi et al. (2015) demonstrated that the combination of data fusion modelling and clustering methods was able to improve the quality of management zone delineation. Specifically, they compared k-means clustering with the Self-Organizing Map (SOM) for delineating management zones maps for variable-rate N application. Furthermore, a hybrid SOM algorithm in combination with k-means was compared with k-means in terms of cluster separation and management zone formation based on data fusion of NDVI and soil parameters.

As reviewed by Behmann et al. (2015), ML techniques have been widely used for the early and accurate detection of biotic stress in crop, specifically, for detection of weeds, plant diseases and insect pests. Mehra et al. (2016) used ML techniques such as Artificial Neural Networks (ANNs), categorical and regression trees and Random Forests (RFs) to approach the problem of predicting the pre-planting risk of Stagonospora nodorum blotch (SNB) in winter wheat. They developed risk assessment models that could be useful in making disease management decisions prior to planting of the wheat crop. Also Tellaeche et al. (2008) showed that cost savings and reduced pollution could be achieved by a Bayesian framework based automatic decision making process for detecting weeds in corn crops.

Machine learning techniques applied to hyperspectral imaging data can be used to reveal physiological and structural characteristics in plants and enable tracking physiological dynamics due to environmental effects (Wahabzada et al., 2016). Goldstein et al. (2017) demonstrated that field data, such as soil moisture, weather, irrigation characteristics, and resulting yield could be fused via ML techniques to provide automated recommendations for irrigation. Gutiérrez et al. (2018) have used thermal imaging and a combination of two ML techniques (Rotation Forests and Decision Trees) to develop a new methodology for the on-the-go assessment of vineyard water status with potential for irrigation decision making.

Machine learning techniques can be used in field spectroscopy for offline and online prediction of the soil parameters studied in the field (Morellos et al., 2016). They can work not only with variables such as derived spectral indices, but also with the entire spectral response trace (Wittenberghe et al., 2014). Spectral indices depend on a small number of available spectral bands and therefore don't use the entire information conveyed by the spectral trace. Thus, there is always the question: which vegetation index or suite of vegetation indices is better for the given task? (Panda et al., 2010). ML techniques, such as Neural

Networks, can be employed for automated extraction of relevant features from the data (You et al., 2017). Continuum-removal (CR) can be used to effectively isolate individual absorption features of interest and, for example, estimate the chemical concentration in dry leaves (Clark and Roush, 1984; Huang et al., 2004; Kokaly, 2001; Vigneau et al., 2011). After the automated extraction of features is performed, the ML techniques (e.g. BPNNs) can be used for automated crop yield prediction and assessing which feature is more informative (Panda et al., 2010).

Due to the large number of bands in the hyperspectral data, some methods (e.g. Stepwise Multiple Linear Regression (SMLR)) are likely to suffer from multicollinearity when applied to canopy hyperspectral data (Bolster et al., 1996; Grossman et al., 1996). Multicollinearity means that one predicted variable can be linearly predicted from the others with a substantial degree of accuracy. A Partial Least Squares Regression (PLSR) approach can be adopted to reduce the large number of measured collinear spectral variables to a few non-correlated latent variables. It avoids the potential overfitting problems that are typically associated with SMLR (Grossman et al., 1996; Hansen and Schjoerring, 2003).

Although Artificial Neural Networks (ANNs) are widely used in remote sensing to predict vegetation parameters and crop yield (Farifteh et al., 2007; Kaul et al., 2005; Kuwata and Shibasaki, 2015) and are generally used to deal with non-linear models, their practical application has some difficulties, such as selection of the number and size of hidden layers, learning rate, the need for a large training dataset and the problem of overfitting. Support Vector Machines (SVMs) are also being used in PA and have the potential to resolve the problem of overfitting when analysing high-dimensional data such as hyperspectral imaging data. Applications of SVMs include crop type classification (Peña et al., 2014; Ustuner et al., 2014; Yang et al., 2016) and crop-weeds classification (Tellaeche et al., 2007). There are other ML techniques which may be used for different applications in PA. Advantages and disadvantages of some of them have been listed by Ali et al. (2015).

Despite the significant recent developments in ML and the successful application in many areas, ML techniques have some fundamental limitations when used naively in a purely data-driven fashion. The accuracy of the predictions and their uncertainties produced by the ML algorithms strongly depend on the data quality, model representativeness and the dependencies between the input and target variables in the collected datasets. Data with a high level of noise, erroneous data, presence of outliers and biases in the data, and incomplete datasets may significantly reduce the predictive power of the models. The appropriate definition of the ML model, such as the GP covariance function, ANN architecture and SVR parametrisations is also critical for optimal performance. Many strategies, such as incorporation of expert knowledge into the covariance function, outlier detection, transfer learning, and model selection through automated cross-validation can be employed to overcome these limitations.

## 3. Yield estimation

Achieving maximum crop yield at minimum cost with a healthy ecosystem is one of the main goals of agricultural production. Early detection and management of problems associated with crop yield restrictions can help increase yield and subsequent profit, and estimating yield is important to numerous crop management and business decisions.

In recent years different ML techniques have been implemented to achieve accurate yield prediction for different crops (Subhadra et al., 2016). The most successful ML techniques have been Artificial Neural Networks (Drummond et al., 2003; Fortin et al., 2011; Liu et al., 2001; Safa et al., 2004), Support Vector Regression (Ruß, 2009), M5-Prime Regression Trees (Frausto-Solis et al., 2009; Marinković et al., 2009; Ruß and Kruse, 2010; Wang and Witten, 1997) and k-nearest neighbour (Zhang et al., 2010).

Recently, Gonzalez-Sanchez et al. (2014) presented a comparative study of ANN, SVR, M5-Prime, kNN ML techniques and Multiple Linear Regression for crop yield prediction in ten crop datasets. To validate the models they used four accuracy metrics: Root Mean Square Error (RMS), Root Relative Square Error (RRSE), Normalized Mean Absolute Error (MAE) and Correlation Factor (R). Results showed that M5-Prime achieved the lowest errors across the produced crop yield models. The results of that study ranked the techniques from the best to the worst, according to RMSE, RRSE, R, and MAE results, in the following order: M5-Prime, kNN, SVR, ANN and MLR.

In another study Nari and Yang-Won (2016) applied four ML techniques, SVM, Random Forest (RF), Extremely Randomized Trees (ERT) and Deep Learning (DL) to estimate corn yield in Iowa State. Comparisons of the validation statistics showed that DL provided more stable results by overcoming the overfitting problem.

As soil and climatic conditions play a significant role in crop growth and yield, online proximal soil sensing for estimation of relevant soil properties remains a missing component in the management system. Pantazi et al. (2016) predicted the within field variation in wheat yield using online multi-layer soil data and satellite imagery crop growth characteristics. They used supervised Self-Organizing Maps (SOM) in this work. The data for a single cropping season was used and the performance of counter-propagation artificial neural networks (CP-ANN), XY-fused Networks (XY-F) and Supervised Kohonen Networks (SKN) for predicting wheat yield was compared. The average overall accuracy for SKN was 81.65%, for CP-ANN was 78.3% and for XY-F was 80.92%, showing that the SKN model had the best overall performance.

Spectral vegetation indices (VIs) are mathematical combinations (often ratios) of mainly red, green and infrared spectral bands. They are designed to find functional relationships between crop characteristics and remote sensing observations (Wiegand et al., 1979). Since the development of the Simple Ratio Index (SR) (Birth and McVey, 1968; Jordan, 1969; Knipling, 1970) and the Normalized Difference Vegetation Index (NDVI) (Johnson, 2014; Rouse et al., 1973; Tucker, 1979) a large number of vegetation indices have been developed, such as the two-band Enhanced Vegetation Index (EVI2) (Bolton and Friedl, 2013) and Normalized Difference Water Index (NDWI) (Satir and Berberoglu, 2016) to name a few. The availability of a large number of indices leads to the need to optimally choose and combine indices for maximally accurate crop yield estimation.

Panda et al. (2010) implemented Back-propagation Neural Network (BPNN) modelling to test the efficiency of the following four spectral vegetation indices: NDVI, green vegetation index (GVI), soil adjusted vegetation index (SAVI) and perpendicular vegetation index (PVI) in corn crop yield prediction. The results showed that the corn yield was best predicted using BPNN models that used the means and standard deviations of PVI grid images.

Although spectral vegetation indices are widely used, they depend only on a small number (usually two) of the available image bands and the full spectrum information in hyperspectral data is not exploited. In their recent publication You et al. (2017) used Convolutional Neural Network (CNN) and Long-Short Term Memory (LSTM) to automatically discover relevant features from raw data. Deep Gaussian Process was employed to integrate the spatio-temporal information from the data. They evaluated the proposed approach on the task of predicting county-level soybean production in the United States. The results of this study showed that the proposed approach outperformed competing techniques by a large margin.

Yuan et al. (2015) addressed the problem of selection of the most representative bands to reduce the dimensionality of the data while maximally preserving its original information. They proposed a group-wise band selection framework, with a joint sparsity constraint, which is fully unsupervised and computationally feasible. They conducted experiments on hyperspectral classification and colour visualisation. The results of those experiments demonstrated that, in these two applications, the proposed framework is more robust and reliable compared to the other state of the art alternatives (e.g. traditional pointwise-selection-based methods).

To make a timely prediction of crop yield, the Spiking Neural Networks (SNN) model has been presented by Bose et al. (2016). It introduces for the first time SNN as a promising technique for spatio-temporal data modelling, analysis, and land use/crop prediction. As deep learning has the capability to extract key features from the data for estimation, it can be expected to have less dependency on the input data. Because of this, even in areas where data acquisition is limited, deep learning can be expected to provide good quality estimation of crop yield (Kuwata and Shibasaki, 2015).

Many other studies have been conducted on the application of ML techniques to crop yield estimation from remotely sensed and *in-situ* data. Table 1. presents a review of the studies and provides a summary, methodology and discussion for each publication. This discussion is concentrated on some key technical aspects of the used ML techniques.

## 4. Precision nitrogen management

As nitrogen (N) plays a significant role in the process of photosynthesis, it is important for crop health and development. At the same time, environmental factors and cost require a prudent application of N. Because of these factors the problem of optimal N management has attracted the attention of numerous researchers (Cao et al., 2015, 2017; Dai et al., 2013; Magney et al., 2017; Zhao et al., 2013).

One of the approaches to optimal N management in PA is to use management zones, that is, identify subfield regions with homogeneous characteristics that require similar treatment. The most widely used methods for delineation of site-specific management zones are the fuzzy C-means and k-means algorithms (Schuster et al., 2011; Vrindts et al., 2005). These are popular clustering methods used extensively for unsupervised learning and identification of structure in datasets. However, determining subfield areas is a difficult task because of the complex correlations and spatial variability of soil properties and nutrient concentrations, which are responsible for variations in crop yield within the field.

The non-destructive methods for making N fertilizer recommendations for crops are commonly based on plant N status testing, using remote sensing and *in-situ* data (Cilia et al., 2014; Maresma et al., 2016; Tremblay et al., 2011). Cao et al. (2015) evaluated two sensor systems, based on a three band user-configurable Crop Circle ACS-470 sensor and a two fixed-band GreenSeeker sensor, in order to estimate winter wheat N status. The results of the comparison demonstrated that the winter wheat N status can be better predicted via the three band Crop Circle ACS-470 sensor.

A comprehensive review of the advantages and disadvantages of different methods for sensing the N status in plants has been conducted by Muñoz-Huerta et al. (2013). Diacono et al. (2013) produced a review of precision N management of a wheat crop. They investigated approaches/results of site-specific N management of wheat to analyse both performance and sustainability of this agricultural practice. However, to the best of the authors' knowledge, there are no published review papers focused on ML techniques for precision N management.

A number of spectrometric studies have been undertaken devoted to the estimation of the N status in plants using CR, vegetation indices (VIs) (Li et al., 2010; Maresma et al., 2016; Ruß and Kruse, 2011), SMLR and PLSR (Grossman et al., 1996; Haboudane et al., 2008; Hansen and Schjoerring, 2003; Huang et al., 2004; Tian et al., 2011; Yu et al., 2013; Zhu et al., 2006).

Yao et al. (2015) applied different linear (CR, VI, SMLR and PLSR) and nonlinear (ANN and SVM) regression methods in order to determine which method, input variable and model could estimate the Leaf Nitrogen Concentration (LNC) in winter wheat with higher accuracy, more robustness, less time and lower complexity. A comparative assessment of those six methods was conducted using the following six metrics: coefficients of determination for the calibration ($R_C^2$) and validation ($R_V^2$) sets, the root mean square errors of prediction (RMSEP) for the calibration and validation sets, the ratio of prediction to deviation (RPD), the computational efficiency (CE) and the complexity level (CL). The results of the comparison showed that the SVM method was more robust in coping with potential confounding factors for most varieties, ecological site and growth stage. However, the VI method utilising the Soil-Adjusted Vegetation Index (1200 and 705 bands) was most accurate for the estimation of the LNC in wheat.

Gaussian Processes (GPs) machine learning regression algorithms have been applied to estimate chlorophyll content, N content, leaf water content and specific leaf area from a field-based multi-species dataset (for trees) (Wittenberghe et al., 2014). As an input the GP used the entire spectral data and then several distinct wavebands were automatically chosen for estimation of different leaf parameters. Results illustrate that the information to predict a leaf parameter is not restricted to one or few spectral bands, but that six or more separate bands can be (equally) involved. Taking more bands into account helps better discriminate the source of change when some bands are also influenced by other variables.

Three methods (PLS, ANN, LS-SVM) have been used to estimate the N status non-destructively in rice using canopy spectral reflectance with visible and near-infrared reflectance spectroscopy (Shao et al., 2012). The comparative analysis showed that the LS-SVM outperformed the other methods and it was concluded that LS-SVM is a promising alternative for the regression analysis to quantify N status in rice.

The Random Forest (RF) algorithm is a data mining method developed by Breiman (2001). It can be employed to reduce the redundancy in complex high-dimensional hyperspectral datasets. The results by Elfatih et al. (2013) show that RF regression applied to hyperspectral data has the potential to accurately predict sugarcane leaf N concentration thus assisting in making informed decisions regarding site-specific application of N fertilizers.

There are many other studies dedicated to precision N management using not only ML but also other techniques such as kriging, multivariate methods and inverse distance weighting. Table 2. presents those studies and provides a summary, methodology and discussion for each publication. This discussion is concentrated on some key technical aspects of the used ML and other techniques.

## 5. Conclusions

Sensing technologies and ML techniques have rapidly advanced during the last decade. These developments are likely to continue providing cost-effective and more comprehensive datasets combined with more sophisticated algorithmic solutions enabling better crop and environment state estimation and decision making. We are at the beginning of a promising path that has the potential to significantly alter crop yield management.

A number of ML techniques have already been successfully applied to different PA tasks. This review particularly demonstrates that

(a) Back-propagation Neural Networks allow identification of the importance of different Vegetation Indices (VI) for more accurate crop yield estimation.
(b) The combination of Convolutional Neural Networks or Long-short Term Memory with Gaussian Processes enables feature extraction from the data and efficient reduction of the error maps.
(c) Gaussian Processes are useful for automatic selection of wavebands from the entire spectrum in order to predict different characteristics of plant leaves.
(d) M5-Prime Regression Trees are a suitable tool for multi-class crop prediction.
(e) Least Squares Support Vector Machine is a promising tool for regression analysis to quantify Nitrogen status.
(f) Fuzzy cognitive Map (FCM) can be used to model and represent expert knowledge for yield prediction and crop management.

**Table 1**
Publications that use machine learning techniques for crop yield estimation with a focus on their technical aspects.

| Paper | Summary | Methodology | Discussion |
|---|---|---|---|
| Pantazi et al. (2016) | This paper developed and evaluated a yield prediction model for wheat. For the yield prediction the fusion vectors have been used as input for the three ANNs. The fusion vectors consist of the values of the eight soil parameters collected with the on-line soil sensor, the satellite imagery calculated NDVI values and historic yield data from the previous two years. | Self-Organizing Map Models (SOMs):<br>● Counter-Propagation Artificial Neural Networks (CPANN)<br>● XY-Fused Networks (XY-Fs)<br>● Supervised Kohonen Networks (SKNs) | The presented approach incorporates the yield limiting factors in a multi-layer fusion model. |
| Stas et al. (2016) | The paper presented a comparison of two machine learning techniques (BRT and SVM) for prediction of winter wheat yield in Henan province of China. Three types of NDVI-related predictors have been used: Single NDVI, Incremental NDVI and Targeted NDVI.<br>The results of comparison, which are based on a cross-validation error (RMSE), showed that BRT model consistently outperforms SVM. | ● Boosted Regression Trees (BRT)<br>● Support Vector Machines (SVM) | When a limited number of training samples is available, ML techniques used in this paper are better able to cope with large set of predictors (compared to MLR) |
| Heremans et al. (2015) | In this paper two regression tree methods (BRT and RF) were used in order to evaluate the accuracy of winter wheat yield, using NDVI data from the SPOT-VEGETATION sensor together with meteorological variables and fertilization levels, in the North China. The aim was not only to compare the performance of the methods but also to assess the potential for early-season predictions of winter wheat yield at the prefecture level (five prefectures were involved).<br>The comparison of methods was based on cross-validation $R^2$ and RMSE. The results showed that BRT outperforms RF for four out of the five prefectures. | ● Boosted Regression Trees (BRT)<br>● Random Forest (RF) | BRT is sensitive to noise, prone to overfitting and much slower than bagging. At the same time, boosting has been shown to be more accurate than bagging.<br>RF can be used to improve the performance of bagging. In terms of accuracy, RFs are comparable to boosting but don't have the mentioned limitations. RF has much lower computational cost than boosting |
| Liang et al. (2015) | The paper presented a non-destructive method - the hybrid inversion method, for estimation of leaf area index (LAI) values of crops. The method used different regression algorithms and allowed determining the relationships between optimal simulated VIs and simulated LAI values. To establish hybrid inversion model ANN and RFR have been used.The comparison of the used algorithms showed that RFR was a better method for modelling with the higher $R^2$ and lower RMSE values for different datasets and various VIs. | ● Curve fitting<br>● Artificial Neural Network (ANN)<br>● Random Forest Regression (RFR) | In contrast to full-spectrum approaches, using VIs to estimate LAI requires a reduction in the number of model input parameters and therefore may result in lower inversion accuracy. However, RFR can enable good performance with several or even a single parameter if that input parameter is highly correlated and representative. |
| Wu et al. (2015) | This paper developed and compared two inversion models, using Statistical Regression model and BPNN model, to estimate the LAI of a temperate meadow steppe in China.<br>The results of comparison showed that BPNN method (accuracy: 82.2%) outperforms Statistical Regression model (accuracy: 78.8%). | ● Statistical Regression Model<br>● BPNN | BPNN refers to a broad family of ANNs where the error is calculated at the output layer (using the observations) and is propagated back through the layers of the ANN. The optimisation process adjusts the weights in each layer by minimising the pre-defined loss function. |
| Jin et al. (2016) | In the paper the particle swarm optimization (PSO) algorithm was used to assimilate field spectroscopic data into the AquaCrop model to improve the estimation accuracy of winter wheat yield under different planting dates and irrigation management strategies.<br>The results showed that the PSO algorithm is an effective method for improving the estimates of biomass and yield of winter wheat. | Particle Swarm Optimization (PSO) algorithm | Particle Swarm Optimization (PSO) algorithm can be used to minimise the difference between the regression based and AquaCrop model based estimates. This helps to improve the predictions. |
| Li et al. (2016) | The paper aimed to produce accurate and timely predictions of grassland LAI for the meadow steppes of northern China, using different regression approaches and hybrid geostatistical methods.<br>The comparison of predictions via hybrid geostatistical methods, followed by different regression models was presented. The results showed that the RF model provides the most accurate predictions among the regression models. | ● Partial Least Squares Regression (PLSR)<br>● ANNs<br>● RFs<br>● Regression Kriging (RK)<br>● Random Forests Residuals Kriging (RFRK) | RFs can provide better resistance to the over-fitting problem and to noise in the data compared with other regression methods. However, RF method ignores spatial autocorrelation information.<br>RFRK is an extension of RF and is very similar to RK. It helps to include the spatial autocorrelation into the RF |
| Papageorgiou et al. (2013) | The paper mainly aimed to present a method, based on FCM learning technique, to develop a computational intelligent tool for categorizing apple fruit yield.<br>The results showed that the knowledge-based FCM learning approach predicts properly the phenomenon, gives a front-end decision about the class of apple fruit yield, and provides similar results to those obtained from horticulturist experts. | ● Fuzzy Cognitive Mapping (FCM) | The FCM method has many advantages such as simplicity, adaptability and capability of approximating abstractive structures. FCMs are knowledge-based and, therefore, provide meaningful results. As FCMs are mainly constructed manually, data-driven learning algorithms are required when dealing with a large number of variables. |
| Papageorgiou et al. (2011) | The main aim of the paper was to connect yield defining parameters with yield in cotton crop production in Central Greece. The simulation approach based on the soft computing technique of Fuzzy Cognitive Maps was investigated (FCM tool).<br>The data from six subsequent years were used to estimate the average classification accuracy of the yield production, using the FCM tool. The results of estimation were compared with results of some ML techniques obtained from the same data.<br>The results of comparison based on the overall accuracy of each method showed that the FCM technique performed better in most of the cases. | ● Fuzzy Cognitive Mapping (FCM)<br>● ANNs<br>● Decision Trees (DTs)<br>● Bayesian Networks (BNs) | Fuzzy cognitive Map (FCM) represents a combination of neural networks and fuzzy logic, and can be used for information representation and decision making in complex processing environments. In particular, FCMs can be used to model and represent expert knowledge for cotton yield prediction and crop management |
| Kaul et al. (2005) | The paper described the development of ANN models as an accurate technique for corn and soybean yield prediction in Maryland nutrient management planning.<br>The results showed that ANN yield prediction is more accurate than the MLR-based yield model. | ● Artificial Neural Network (ANN)<br>● Multiple Linear Regression (MLR) | ANN and MLR are among the techniques that can be used for agricultural modelling and prediction. The MLR is a simple methodology which is also easy to apply. ANN is a much more sophisticated technique the difficulties of the practical application of which are described in Section 2. |

**Table 2**
Publications that use machine learning and other techniques for precision nitrogen management.

| Paper | Summary | Methodology | Discussion |
|---|---|---|---|
| Song et al. (2017) | The paper showed that based on the data collected from some georeferenced locations the Ordinary Kriging Analysis allows the interpolation of maps for<br>● wheat grain protein content (GPC)<br>● GPC yield<br>● wheat canopy fluorescence parameters, including Simple Florescence Ratio and Nitrogen Balance Indices (NBI)<br>● soil Nitrate-Nitrogen (NO3-N) content and soil Time Domain Reflectometry (TDR)<br>The comparison of the fluorescence parameter maps, soil $NO_3$-N and soil TDR maps with the wheat GPC and the GPC yield maps demonstrated that the NBI spatial variability map in the late stage of wheat growth can be used to distinguish areas that produce higher GPC | Ordinary Kriging Analysis (OKA) | Ordinary Kriging Analysis is a widely used approach in PA. However, ML via Gaussian Processes based modelling will provide<br>● automation<br>● quick update of the models as new data becomes available<br>● optimised data collection (based on maximising the information gain)<br>● probabilistic outcomes with representative uncertainties<br>● possibility for further fusion of data from different sensors |
| Magney et al. (2017) | The paper aimed<br>● to evaluate the usefulness of RapidEye spectral VIs to predict cumulative N uptake in wheat<br>● to examine the usefulness of remotely sensed N uptake maps for PA applications<br>It was concluded that<br>● the top performing VI was the Normalized Difference Red-Edge index (NDRE)<br>● N uptake maps from RapidEye imagery could have important implications for PA | Seventeen commonly used spectral VIs | Although VIs have been used for a number of different applications in satellite remote sensing for PA, this paper reports that VIs from RapidEye imagery can be used for estimating wheat N uptake. |
| Morellos et al. (2016) | To predict total nitrogen (TN), organic carbon (OC) and moisture content (MC) in fresh (wet and unprocessed) soil samples two multivariate and two machine learning methods have been compared.<br>The results indicated that machine learning methods outperformed the multivariate methods for the prediction of all three soil properties. | Multivariate methods:<br>● Principal Component Regression (PCR)<br>● Partial Least Squares Regression (PLSR),<br>Machine learning methods:<br>● Least Squares Support Vector Machines (LS-SVM)<br>● Cubist | The advantage of ML methods is that they are capable of tackling non-linear problems in the dataset.<br>The ML techniques can be used in field spectroscopy for off-line and online prediction of the soil parameters studied in the fields (if the soil type and variability is similar to the one studied in this paper) |
| Castaldi et al. (2016) | The paper proposed data fusion process in order to improve the choice of satellite bands for grain N uptake prediction.<br>The results showed that the best spectral regions vary over the growing season of the wheat crop. | Combination of<br>● Stepwise Regression with Backward Selection<br>● Stepwise Variance Inflation Factors (VIFs) analysis<br>● Linear Mixed Effect Model (LMEM) | LMEM can be a very efficient technique to estimate the spatial variability of the soil and crop variables accurately across the field with limited data, thus saving time and reducing the costs. |
| Wang et al. (2017) | The paper investigated the modelling performances of four different chemometric techniques and three vegetation indices. Results showed that the best modelling and prediction accuracy were found in the model established by PLSR and spectra measured with a black background. A higher coefficient of determination between the leaf N concentration and fruit yield was found at 50 days after full bloom. | Four techniques:<br>● Principal Components Regression (PCR)<br>● Partial Least Squares Regression (PLSR)<br>● Stepwise Multiple Linear Regression (SMLR)<br>● BPNN<br>Three indices:<br>● Difference Spectral Index<br>● Normalized Difference Spectral Index<br>● Ratio Spectral Index | PCR, PLSR, and BPNN use all available wavelengths simultaneously, while SMLR selects useful wavelengths from the available spectrum and ignores the remaining wavebands.<br>To improve the performance of the methods normalization can be used on the raw spectra collected by the probe, and wavelengths with very large atmospheric influence can be removed. |
| Guo et al. (2015) | The paper compared two different approaches (SLR and RFRK) to predict and map the spatial distribution of soil organic matter for the rubber plantation.<br>Results showed that RFRK outperforms SLR, by providing lower prediction errors (ME, MAE, and RMSE) and higher R2 | ● Stepwise Linear Regression (SLR)<br>● RFRK<br>● Generalized Additive Mixed Model (GAMM)<br>● Classification And Regression Tree (CART) | RFRK model required no assumptions about the relationships between the target variable and the predictor variables. Those relationships could be nonlinear and hierarchical. This can be revealed by using GAMM and CART. |
| Dai et al. (2014) | The paper presented ANN-kriging methodology in order to predict accurate Soil Organic Matter (SOM) content maps.<br>A comparison of proposed method with the other interpolation methods was performed to assess the prediction accuracy.<br>The results indicated that ANN-kriging provides the lower RMSE. | ● ANN-kriging<br>● ANN<br>● Inverse Distance Weighting (IDW) | It is suggested that the proposed ANN-kriging methodology can be used to improve the accuracy of SOM content mapping at large scale. |

Based on the current dynamics in algorithmic developments and sensor technologies the following future trends can be expected:

(a) More optimised, targeted application of the currently available sensors and established ML techniques to specific PA tasks
(b) Interconnected treatment of spatial, spectral and temporal domains and incorporation of expert knowledge into the ML techniques aimed at modelling and decision making in different aspects of PA.
(c) The combination of multiple ML as well as signal processing techniques into hybrid systems to benefit from the strengths of those techniques and compensate for their individual shortcomings.
(d) Spatial and spectral fusion of information from sensors with different spatial resolution and spectral characteristics.
(e) The dynamic combination of stationary (e.g. in-ground probes, weather station) and mobile (e.g. ground and aerial vehicles, satellites) equipment to enable active optimal data collection, information fusion and model update for high value areas.

## Acknowledgement

This work has been supported by the Australian Centre for Field Robotics.

## References

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., Notarnicola, C., 2015. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. Remote Sens. 7, 15841.
Andrews, M., Raven, J.A., Lea, P.J., 2013. Do plants need nitrate? The mechanisms by which nitrogen form affects plants. Ann. Appl. Biol. 163, 174–199.
Apostol, S., Viau, A., Tremblay, N., 2007. A comparison of multiwavelength laser-induced fluorescence parameters for the remote sensing of nitrogen stress in field-cultivated corn. Can. J. Remote Sens. 33, 150–161.
Aqeel-ur-Rehman, Abbasi, A.Z., Islam, N., Shaikh, Z.A., 2014. A review of wireless sensors and networks' applications in agriculture. Comp. Stand. Interf. 36, 263–270.
Barati, S., Rayegani, B., Saati, M., Sharifi, A., Nasri, M., 2011. Comparison the accuracies of different spectral indices for estimation of vegetation cover fraction in sparse vegetated areas. Egypt. J. Remote Sens. Space Sci. 14, 49–56.
Barnes, E.M., Sudduth, K.A., Hummel, J.W., Lesch, S.M., Corwin, D.L., Yang, C., Daughtry, C.S.T., Bausch, W.C., 2003. Remote- and ground-based sensor techniques to map soil properties. Photogramm. Eng. Remote Sens. 6, 619–630.
Behmann, J., Mahlein, A., Rumpf, T., Römer, C., Plümer, L., 2015. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Prec. Agric. 16, 239–260.
Bellvert, J., Marsal, J., Girona, J., Gonzalez-Dugo, V., Fereres, E., Ustin, S., Zarco-Tejada, P., 2016. Airborne thermal imagery to detect the seasonal evolution of crop water status in peach, nectarine and saturn peach orchards. Remote Sens. 8, 39.
Birth, G.S., McVey, G.R., 1968. Measuring the color of growing turf with a reflectance spectrophotometer. Agron. J. 60, 640–643.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, New York Inc.
Bolster, K.L., Martin, M.E., Aber, J.D., 1996. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectances: a comparison of statistical methods. Can. J. For. Res. 26, 590–600.
Bolton, D.K., Friedl, M.A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84.
Bose, P., Kasabov, N.K., Bruzzone, L., Hartono, R.N., 2016. Spiking neural networks for crop yield estimation based on spatiotemporal analysis of image time series. IEEE Trans. Geosci. Remote Sens. 54, 6563–6573.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Campbell, J.B., Wynne, R.H., 2011. Introduction to Remote Sensing. fifth ed. Guildford Press.
Cao, Q., Miao, Y., Feng, G., Gao, X., Li, F., Liu, B., Yue, S., Cheng, S., Ustin, S.L., Khosla, R., 2015. Active canopy sensing of winter wheat nitrogen status: An evaluation of two sensor systems. Comput. Electr. Agric. 112, 54–67.
Cao, Q., Miao, Y., Li, F., Gao, X., Liu, B., Lu, D., Chen, X., 2017. Developing a new crop circle active canopy sensor-based precision nitrogen management strategy for winter wheat in North China Plain. Prec. Agric. 18, 2–18.
Castaldi, F., Castrignanò, A., Casa, R., 2016. A data fusion and spatial data analysis approach for the estimation of wheat grain nitrogen uptake from satellite data. Int. J. Remote Sens. 37, 4317–4336.
Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., Li, B., 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. Remote Sens. Environ. 114, 1987–1997.
Chen, D., Suter, H., Islam, A., Edis, R., Freney, J.R., Walker, C.N., 2008. Prospects of improving efficiency of fertiliser nitrogen in Australian agriculture: a review of enhanced efficiency fertilisers. Soil Res. 46, 289–301.
Cilia, C., Panigada, C., Rossini, M., Meroni, M., Busetto, L., Amaducci, S., Boschetti, M., Picchi, V., Colombo, R., 2014. Nitrogen status assessment for variable rate fertilization in maize through hyperspectral imagery. Remote Sens. 6, 6549.
Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. J. Geophys. Res.: Solid Earth 89, 6329–6340.
Curran, P.J., 1987. Remote sensing in agriculture: an introductory review. J. Geogr. 86, 147–156.
Dai, X., Zhou, X., Jia, D., Xiao, L., Kong, H., He, M., 2013. Managing the seeding rate to improve nitrogen-use efficiency of winter wheat. Field Crops Res. 154, 100–109.
Dai, F., Zhou, Q., Lv, Z., Wang, X., Liu, G., 2014. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. Ecol. Indicat. 45, 184–194.
Dalponte, M., Bruzzone, L., Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. Remote Sens. Environ. 123, 258–270.
Diacono, M., Rubino, P., Montemurro, F., 2013. Precision nitrogen management of wheat A review. Agron. Sustain. Dev. 33, 219–241.
Drummond, S.T., Sudduth, K.A., Joshi, A., Birrel, S.J., Kitchen, N.R., 2003. Statistical and neural methods for site-specific yield prediction. Trans. Am. Soc. Agric. Eng. 46, 5–14.
Du, Z., Jeffrey, J.P.T., 2007. Advances in Machine Learning Applications in Software Engineering. IGI Global, Hershey, PA, USA.
Elfatih, M.A., Fethi, B.A., Riyad, I., 2013. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. Int. J. Remote Sens. 34, 712–728.
El-Shikha, D.M., Waller, P., Hunsaker, D., Clarke, T., Barnes, E., 2007. Ground-based remote sensing for assessing water and nitrogen status of broccoli. Agric. Water Manage. 92, 183–193.
Engman, E.T., Gurney, R.J., 1991. Remote sensing in hydrology. Chapman and Hall, London.
Farifteh, J., Van der Meer, F., Atzberger, C., Carranza, E.J.M., 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). Remote Sens. Environ. 110, 59–78.
Ferguson, T.S., 1973. A Bayesian analysis of some nonparametric problems. Ann. Statist. 1, 209–230.
Fortin, J.G., Anctil, F., Parent, L., Bolinder, M.A., 2011. Site-specific early season potato yield forecast by neural network in Eastern Canada. Prec. Agric. 12, 905–923.
Frausto-Solis, J., Gonzalez-Sanchez, A., Larre, M., 2009. A New Method for Optimal Cropping Pattern, In: Aguirre, A.H., Borja, R.M., Garciá, C.A.R. (Eds.), MICAI 2009: Advances in Artificial Intelligence: 8th Mexican International Conference on Artificial Intelligence, Guanajuato, México, November 9-13, 2009. Proceedings. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 566–577.
Gao, B., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.
Goldstein, A., Fink, L., Meitin, A., Bohadana, S., Lutenberg, O., Ravid, G., 2017. Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge. Prec. Agric.
Gontia, N.K., Tiwari, K.N., 2008. Development of crop water stress index of wheat crop for scheduling irrigation using infrared thermometry. Agric. Water Manage. 95, 1144–1152.
Gonzalez-Sanchez, A., Frausto-Solis, J., Ojeda-Bustamante, W., 2014. Predictive ability of machine learning methods for massive crop yield prediction. Spanish J. Agric. Res. 12, 313–328.
Goron, T., Nederend, J., Stewart, G., Deen, B., Raizada, M., 2017. Mid-season leaf glutamine predicts end-season maize grain yield and nitrogen content in response to nitrogen fertilization under field conditions. Agronomy 7, 41.
Govender, M., Dye, P.J., Weiersbye, I.M., Witkowski, E.T.F., Ahmed, F., 2009. Review of commonly used remote sensing and ground-based technologies to measure plant water stress. Water SA 35.
Griffiths, T.L., Ghahramani, Z., 2011. The Indian buffet process: an introduction and review. J. Mach. Learn. Res. 12, 1185–1224.
Grossman, Y.L., Ustin, S.L., Jacquemoud, S., Sanderson, E.W., Schmuck, G., Verdebout, J., 1996. Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. Remote Sens. Environ. 56, 182–193.
Guo, P., Li, M., Luo, W., Tang, Q., Liu, Z., Lin, Z., 2015. Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach. Geoderma 237, 49–59.
Gupta, R.P., 2003. Remote Sensing Geology. Springer-Verlag, Berlin Heidelberg, Germany.
Gutiérrez, S., Diago, M.P., Fernández-Novales, J., Tardaguila, J., 2018. Vineyard water status assessment using on-the-go thermal imaging and machine learning. PLOS One 13, e0192037.
Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens. Environ. 90, 337–352.
Haboudane, D., Tremblay, N., Miller, J.R., Vigneault, P., 2008. Remote estimation of crop chlorophyll content using spectral indices derived from hyperspectral data. IEEE Trans. Geosci. Remote Sens. 46, 423–437.
Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 86, 542–553.
Heremans, S., Dong, Q., Zhang, B., Bydekerke, L., Van Orshoven, J., 2015. Potential of ensemble tree methods for early-season prediction of winter wheat yield from short time series of remotely sensed normalized difference vegetation index and in situ meteorological data. APPRES 9, 097095.
Holmgren, P., Thuresson, T., 1998. Satellite remote sensing for forestry planning—A review. Scandinavian J. For. Res. 13, 90–110.
Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R., Foley, W.J., 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. Remote Sens. Environ. 93, 18–29.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Hultquist, C., Chen, G., Zhao, K., 2014. A comparison of Gaussian process regression, random forests and support vector regression for burn severity assessment in diseased forests. Remote Sens. Lett. 5, 723–732.
Jaafar, H.H., Ahmad, F.A., 2015. Crop yield prediction from remotely sensed vegetation indices and primary productivity in arid and semi-arid lands. Int. J. Remote Sens. 36, 4570–4589.
Jin, X., Kumar, L., Li, Z., Xu, X., Yang, G., Wang, J., 2016. Estimation of winter wheat biomass and yield by combining the AquaCrop model and field hyperspectral data. Remote Sens. 8, 972.
Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128.
Jones, D.I.H., Moseley, G., 1993. Laboratory methods for estimating nutritive quality. In: Davies, A., Baker, R.D., Grant, S.A., Laidlaw, A.S. (Eds.), Sward Measurement Handbook, second ed., pp. 265–283.
Jordan, C.F., 1969. Derivation of leaf-area index from quality of light on the forest floor. Ecology 50, 663–666.
Kaul, M., Hill, R.L., Walthall, C., 2005. Artificial neural networks for corn and soybean yield prediction. Agric. Syst. 85, 1–18.
Knipling, E.B., 1970. Physical and physiological bases for the reference of visible and near infrared radiation from vegetation. Remote Sens. Environ. 1, 155–159.
Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sens. Environ. 75, 153–161.
Kuwata, K., Shibasaki, R., 2015. Estimating crop yields with deep learning and remotely sensed data, 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 858–861.
Lamb, D.W., Brown, R.B., 2001. PA—Precision agriculture. J. Agric. Eng. Res. 78, 117–125.
Lamb, D.W., Steyn-Ross, M., Schaare, P., Hanna, M.M., Silvester, W., Steyn-Ross, A., 2002. Estimating leaf nitrogen concentration in ryegrass (Lolium spp.) pasture using the chlorophyll red-edge: Theoretical modelling and experimental observations. Int. J. Remote Sens. 23, 3619–3648.
Li, F., Miao, Y., Hennig, S.D., Gnyp, M.L., Chen, X., Jia, L., Bareth, G., 2010. Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. Prec. Agric. 11, 335–357.
Li, Z., Wang, J., Tang, H., Huang, C., Yang, F., Chen, B., Wang, X., Xin, X., Ge, Y., 2016. Predicting grassland leaf area index in the meadow steppes of northern china: a comparative study of regression approaches and hybrid geostatistical methods. Remote Sens. 8, 632.
Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S., Lin, H., 2015. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. Remote Sens. Environ. 165, 123–134.
Liu, J., Goering, C.E., Tian, L., 2001. A neural network for setting target corn yields. 44, pp. 705–713.
Lukina, E.V., Freeman, K.W., Wynn, K.J., Thomason, W.E., Mullen, R.W., Stone, M.L., Solie, J.B., Klatt, A.R., Johnson, G.V., Elliott, R.L., Raun, W.R., 2001. Nitrogen fertilization optimization algorithm based on in-season estimates of yield and plant nitrogen uptake. J. Plant Nutr. 24, 885–898.
MacKerron, D.K.L., Young, M.W., Davies, H.V., 1993. A method to optimize N-application in relation to soil supply of N, and yield of potato. In: Fragoso, M.A.C., Van Beusichem, M.L., Houwers, A. (Eds.), Optimization of Plant Nutrition: Refereed papers from the Eighth International Colloquium for the Optimization of Plant Nutrition, 31 August – 8 September 1992, Lisbon, Portugal. Springer, Netherlands, Dordrecht, pp. 635–640.
Maes, W.H., Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. J. Exp. Bot. 63, 4671–4712.
Magney, T.S., Eitel, J.U.H., Vierling, L.A., 2017. Mapping wheat nitrogen uptake from RapidEye vegetation indices. Prec. Agric. 18, 429–451.
Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., Martínez-Casasnovas, J., 2016. Analysis of Vegetation Indices to Determine Nitrogen Application and Yield Prediction in Maize (Zea mays L.) from a Standard UAV Service. Remote Sensing 8, 973.
Marinković, B., Crnobarac, J., Brdar, S., Antić, B., Jaćimović, G., Crnojević, V., 2009. Data mining approach for predictive modeling of agricultural yield data. Proc. First International Workshop on Sensing Technologies in Agriculture, Novi Sad, Serbia, pp. 1–5.
Mehra, L.K., Cowger, C., Gross, K., Ojiambo, P.S., 2016. Predicting pre-planting risk of stagonospora nodorum blotch in winter wheat using machine learning models. Front. Plant Sci. 7, 390.
Morellos, A., Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., Wiebensohn, J., Bill, R., Mouazen, A.M., 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst. Eng. 152, 104–116.
Muñoz-Huerta, R., Guevara-Gonzalez, R., Contreras-Medina, L., Torres-Pacheco, I., Prado-Olivarez, J., Ocampo-Velazquez, R., 2013. A review of methods for sensing the nitrogen status in plants: advantages disadvantages and recent advances. Sensors 13, 10823.
Nari, K., Yang-Won, L., 2016. Machine learning approaches to corn yield estimation using satellite images and climate data: a case of iowa state. J. Korean Soc. Surv., Geodesy, Photogramm. Cartogr. 34, 383–390.
Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Chapter Four - Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 175–245.
Panda, S.S., Ames, D.P., Panigrahi, S., 2010. Application of vegetation indices for agricultural crop yield prediction using neural network techniques. Remote Sens. 2, 673–696.
Pantazi, X.E., Moshou, D., Mouazen, A., Alexandridis, T., Kuang, B., 2015. Data Fusion of Proximal Soil Sensing and Remote Crop Sensing for the Delineation of Management Zones in Arable Crop Precision Farming. In: 7th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA 2015), Kavala - Greece.
Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L., Mouazen, A.M., 2016. Wheat yield prediction using machine learning and advanced sensing techniques. Comput. Electr. Agric. 121, 57–65.
Papageorgiou, E.I., Markinos, A.T., Gemtos, T.A., 2011. Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. Appl. Soft Comput. 11, 3643–3657.
Papageorgiou, E.I., Aggelopoulou, K.D., Gemtos, T.A., Nanos, G.D., 2013. Yield prediction in apples using Fuzzy Cognitive Map learning approach. Comput. Electr. Agric. 91, 19–29.
Peña, J., Gutiérrez, P., Hervás-Martínez, C., Six, J., Plant, R., López-Granados, F., 2014. Object-based image classification of summer crops with machine learning methods. Remote Sens. 6, 5019.
Pohl, C., Van Genderen, J.L., 1998. Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. Int. J. Remote Sens. 19, 823–854.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126.
Quemada, M., Gabriel, J., Zarco-Tejada, P., 2014. Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization. Remote Sens. 6, 2940.
Rasmussen, C.E., Williams, C.K.I., 2005. Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press.
Raun, W.R., Solie, J.B., Stone, M.L., Martin, K.L., Freeman, K.W., Mullen, R.W., Zhang, H., Schepers, J.S., Johnson, G.V., 2005. Optical sensor-based algorithm for crop nitrogen fertilization. Commun. Soil Sci. Plant Anal. 36, 2759–2781.
Reyniers, M., Vrindts, E., 2006. Measuring wheat nitrogen status from space and ground-based platform. Int. J. Remote Sens. 27, 549–567.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring Vegetation Systems in the Great plains with ERTS, NASA. Goddard Space Flight Center 3d ERTS-1 Symposium. United States, Washington, DC, pp. 309–317.
Ruß, G., Kruse, R., 2010. Feature Selection for Wheat Yield Prediction. In: Bramer, M., Ellis, R., Petridis, M. (Eds.), Research and Development in Intelligent Systems XXVI: Incorporating Applications and Innovations in Intelligent Systems XVII. Springer, London, London, pp. 465–478.
Ruß, G., Kruse, R., 2011. Machine learning methods for spatial clustering on precision agriculture data. In: Kofod-Petersen, A., Heintz, F., Langseth, H. (Eds.), Eleventh Scandinavian Conference on Artificial Intelligence. IOS Press, pp. 40–49.
Ruß, G., 2009. Data mining of agricultural yield data: a comparison of regression models. In: Perner, P. (Ed.), Advances in Data Mining. Applications and Theoretical Aspects: 9th Industrial Conference, ICDM 2009, Leipzig, Germany, July 20 - 22, 2009. Proceedings. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 24–37.
Safa, B., Khalili, A., Teshnehlab, M., Liaghat, A., 2004. Artificial neural networks application to predict wheat yield using climatic data. In: Proc. 20th Int. Conf. on IIPS, pp. 1–39.
Satir, O., Berberoglu, S., 2016. Crop yield prediction under soil salinity using satellite derived vegetation indices. Field Crops Res. 192, 134–143.
Scharf, P.C., Schmidt, J.P., Kitchen, N.R., Sudduth, K.A., Hong, S.Y., Lory, J.A., Davis, J.G., 2002. Remote sensing for nitrogen management. J. Soil Water Conserv. 57, 518–524.
Schepers, J.S., Raun, W.R., 2008. Nitrogen in Agricultural Systems. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI.
Schuster, E.W., Kumar, S., Sarma, S.E., Willers, J.L., Milliken, G.A., 2011. Infrastructure for data-driven agriculture: identifying management zones for cotton using statistical modeling and machine learning techniques. In: 2011 8th International Conference & Expo on Emerging Technologies for a Smarter World, p. 1.
Seelan, S.K., Laguette, S., Casady, G.M., Seielstad, G.A., 2003. Remote sensing applications for precision agriculture: A learning community approach. Remote Sens. Environ. 88, 157–169.
Shao, Y., Zhao, C., Bao, Y., He, Y., 2012. Quantification of nitrogen status in rice by least squares support vector machines and reflectance spectroscopy. Food Bioprocess Technol. 5, 100–107.
Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81, 337–354.
Song, X., Yang, G., Yang, C., Wang, J., Cui, B., 2017. Spatial variability analysis of within-field winter wheat nitrogen and grain quality using canopy fluorescence sensor measurements. Remote Sens. 9, 237.
Stas, M., Orshoven, J.V., Dong, Q., Heremans, S., Zhang, B., 2016. A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT-VGT. In: 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1–5.
Subhadra, M., Debahuti, M., Gour Hari, S., 2016. Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper. Indian J. Sci. Technol. 9.
Sui, R., Thomasson, J.A., Hanks, J., Wooten, J., 2008. Ground-based sensing system for weed mapping in cotton. Comput. Electr. Agric. 60, 31–38.
Taghvaeian, S., Chávez, J., Hansen, N., 2012. Infrared thermometry to estimate crop water stress index and water use of irrigated maize in Northeastern Colorado. Remote Sens. 4, 3619.
Tellaeche, A., Burgos-Artizzu, X.P., Pajares, G., Ribeiro, A., 2008. A vision-based method for weeds identification through the Bayesian decision theory. Pattern Recogn. 41, 521–530.
Tellaeche, A., BurgosArtizzu, X.P., Pajares, G., Ribeiro, A., 2007. A vision-based classifier in precision agriculture combining bayes and support vector machines. In: IEEE International Symposium on Intelligent Signal Processing, pp. 1–6.
Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., Zhu, Y., 2011. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. Field Crops Res. 120, 299–310.

Tian, Y.C., Gu, K., Chu, X., Yao, X., Cao, W.X., Zhu, Y., 2014. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. Plant Soil 376, 193–209.

Tilling, A.K., O'Leary, G.J., Ferwerda, J.G., Jones, S.D., Fitzgerald, G.J., Rodriguez, D., Belford, R., 2006. Remote sensing of nitrogen and water stress in wheat. Field Crops Res. 104, 77–85.

Tremblay, N., Fallon, E., Ziadi, N., 2011. Sensing of crop nitrogen status: opportunities, tools, limitations, and supporting information requirements. HortTechnology 21, 274–281.

Tremblay, N., Wang, Z., Cerovic, Z.G., 2012. Sensing crop nitrogen status with fluorescence indicators. A review. Agron. Sustain. Dev. 32, 451–464.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.

Ustuner, M., Sanli, F.B., Abdikan, S., Esetlili, M.T., Kurucu, Y., 2014. Crop Type Classification Using Vegetation Indices of RapidEye Imagery. ISPRS Technical Commission VII Symposium, Istanbul, Turkey.

Vigneau, N., Ecarnot, M., Rabatel, G., Roumet, P., 2011. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. Field Crops Res. 122, 25–31.

Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. Remote Sens. Environ. 115, 3468–3478.

Vrindts, E., Mouazen, A.M., Reyniers, M., Maertens, K., Maleki, M.R., Ramon, H., De Baerdemaeker, J., 2005. Management zones based on correlation between soil compaction, yield and crop data. Biosyst. Eng. 92, 419–428.

Wahabzada, M., Mahlein, A., Bauckhage, C., Steiner, U., Oerke, E., Kersting, K., 2016. Plant Phenotyping using Probabilistic Topic Models: Uncovering the Hyperspectral Language of Plants. 6, 22482.

Wang, J., Shen, C., Liu, N., Jin, X., Fan, X., Dong, C., Xu, Y., 2017. Non-destructive evaluation of the leaf nitrogen concentration by in-field visible/near-infrared spectroscopy in pear orchards. Sensors 17, 538.

Wang, Y., Witten, I., 1997. Inducing model trees for continuous classes, Proc. 9th Eur. Conf. Mach. Learn. 128–137.

Whelan, B.M., Taylor, J.A., 2013. Precision agriculture for grain production systems. CSIRO Publishing.

Wiegand, C.L., Richardson, A.J., Kanemasu, E.T., 1979. Leaf area index estimates for wheat from LANDSAT and their implications for evapotranspiration and crop modeling. Agron. J. 71, 336–342.

Wittenberghe, V.S., Verrelst, J., Rivera, J.P., Alonso, L., Moreno, J., Samson, R., 2014. Gaussian processes retrieval of leaf parameters from a multi-species reflectance, absorbance and fluorescence dataset. J. Photochem. Photobiol. B: Biol. 134, 37–48.

Wu, Q., Jin, Y., Bao, Y., Hai, Q., Yan, R., Chen, B., Zhang, H., Zhang, B., Li, Z., Li, X., Xin, X., 2015. Comparison of two inversion methods for leaf area index using HJ-1 satellite data in a temperate meadow steppe. Int. J. Remote Sens. 36, 5192–5207.

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: a review of developments and applications. J. Sens. 2017, 17.

Yang, J., Gong, W., Shi, S., Du, L., Sun, J., Song, S., 2016. Laser-induced fluorescence characteristics of vegetation by a new excitation wavelength. Spectrosc. Lett. 49, 263–267.

Yao, X., Zhu, Y., Tian, Y.C., Feng, W., Cao, W., 2010. Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. Int. J. Appl. Earth Observ. Geoinform. 12, 89–100.

Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y.C., Cao, W., Zhu, Y., 2015. Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. Remote Sens. 7, 14939.

You, J., Li, X., Low, M., Lobell, D., Ermon, S., 2017. Deep gaussian process for crop yield prediction based on remote sensing data. Association for the Advancement of Artificial Intelligence.

Yu, K., Li, F., Gnyp, M.L., Miao, Y., Bareth, G., Chen, X., 2013. Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain. ISPRS J. Photogramm. Remote Sens. 78, 102–115.

Yuan, Y., Zhu, G., Wang, Q., 2015. Hyperspectral band selection by multitask sparsity pursuit. IEEE Trans. Geosci. Remote Sens. 53, 631–644.

Zarco-Tejada, P.J., González-Dugo, V., Berni, J.A.J., 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. Remote Sens. Environ. 117, 322–337.

Zhang, L., Zhang, J., Kyei-Boahen, S., Zhang, M., 2010. Simulation and prediction of soybean growth and development under field conditions. Am.-Eurasian J. Agric. Environ. Sci. 7, 374–385.

Zhao, G., Miao, Y., Wang, H., Su, M., Fan, M., Zhang, F., Jiang, R., Zhang, Z., Liu, C., Liu, P., Ma, D., 2013. A preliminary precision rice management system for increasing both grain yield and nitrogen use efficiency. Field Crops Res. 154, 23–30.

Zhou, Y., Song, G., Wang, M., 2012. Wireless sensor network data fusion algorithm based on neural network in the area of agriculture. Sens. Transducers J. 16, 128–136.

Zhu, Y., Li, Y., Zhou, D., Tian, Y.C., Yao, X., Cao, W., 2006. Quantitative relationship between leaf nitrogen concentration and canopy reflectance spectra in rice and wheat. Acta Ecol. Sin. 26, 3463–3469.