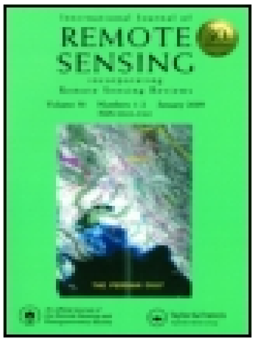

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Barley yield and fertilization analysis from UAV imagery: a deep learning approach

H. J. Escalante, S. Rodríguez-Sánchez, M. Jiménez-Lizárraga, A. Morales-Reyes, J. De La Calleja & R. Vazquez

INTERNATIONAL JOURNAL OF REMOTE SENSING
https://doi.org/10.1080/01431161.2019.1577571

Barley yield and fertilization analysis from UAV imagery: a deep learning approach

H. J. Escalante[a], S. Rodríguez-Sánchez[b], M. Jiménez-Lizárraga[c], A. Morales-Reyes[a], J. De La Calleja[d] and R. Vazquez[e]

[a]Computer Science Department, Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México; [b]Faculty of Mechanical and Electrical Engineering, UANL, San Nicolás de los Garza, NL, México; [c]Faculty of Physical and Mathematical Sciences, UANL, San Nicolás de los Garza, NL, México; [d]Universidad Politécnica de Puebla, México; [e]Faculty of Agronomy, UANL, San Nicolás de los Garza, NL, México

ABSTRACT

In a world with great challenges in food security, optimising cereal production is critical. Cereals are the most important food source for human consumption. The fourth important cereal worldwide after wheat, rice and maize, is Barley, and its production strongly depends on fertilization treatments. The adoption of suitable fertilizer management strategies often results in large economic benefits to producers. However, determining optimal fertilizer doses for a specific barley variety is complex. The collection of data and their analyses can be cost prohibitive for small farmers regarding time and money. This paper introduces an approach to support producers with automatic tools for the analysis of fertilization management of barley. The proposed methodology aims to simultaneously estimate nitrogen fertilization and barley yield, from information derived from aerial RGB images captured by a UAV. Our long term goal is to provide a low-cost and wide-are-coverage solution for the estimation of barley variables that can be leveraged to increase barley yield without increasing costs. A low-cost UAV is used to capture RGB crop field images. Then, a deep convolutional neural network is used for the automated extraction of features from the images. Extracted features are feed into predictive models that estimate the variables of interest. Experimental results reveal that the proposed methodology is able to reach an accuracy above 83% when estimating nitrogen fertilization and a high correlation and low RMSE in the estimation of yield in grams. Experimental results are promising and will pave the way for the development of deep learning methods for barley analysis from aerial imagery that can be accessed by the average farmer.

ARTICLE HISTORY
Received 8 August 2018
Accepted 29 November 2018

1. Introduction

Agriculture plays a key role in the challenge of feeding the world's population, and gradually is evolving to a more high-tech adoption in its productive system. The

CONTACT M. Jiménez-Lizárraga manalejimenez@yahoo.com Faculty of Physical and Mathematical Sciences, UANL, San Nicolás de los Garza, NL 66451, México

Agrismart industry is characterized by cyber systems, wireless, GPS, sensors, algorithms, actuators and controller devices jointly integrated to improve productivity. However, although this new technology has provided huge benefits to farming, it is also bringing new challenges. Producers are now facing problems to manage large amounts of data (Robert, 2002; Zhang et al. 2002), and turning into smart information that has the capacity to support decision-making and keep up-to-date on crop status. Hence, there is an urgent need for tools specifically designed for storing, processing and analyzing agriculture data with acceptable accuracy and fast response.

Specifically, the efficient use and targeting of fertilizer inputs on smallholder farms have been highlighted as a key to reach sustainability. However, analyzing fertilization variables is a complex task as usually, each barley variety requires an specific study to come up with an acceptable doses and management recommendations.

In this paper, we leverage machine learning to develop predictive models for the analysis of aerial imagery of crop fields, in order to provide the producer with information about fertilization treatment. The study focuses on a barley crop and relies on data from a Nitrogen (N) fertilization experiment that took place in Mexico and involved six barley varieties. Our long term goal is to provide a low-cost and wide-are-coverage solution for the estimation of barley variables that can be leveraged to increase barley production without increasing costs, through a fertilization management strategy selection. Barley was selected for their economic importance as it is the fourth largest cereal worldwide after maize, rice and wheat, with a world production of 144 millions tonnes (statistics from the Food and Agriculture Organization of the United Nations [FAO], 2014). What is more, Barley is the most important cereal in economical terms for the region in which the study took place in Mexico. We focus on the analysis of N fertilization and Barley yield as it is known that proper nutrition for satisfactory crop growth and production is correlated with fertilization (Ryan et al. 2008). One should note that although our study focuses on Barley, it can be extended to work with other crops.

The use of satellites to precise monitoring of crop growth in smallholder fields is very extensive, but limited due to its high cost, time consumption, resolution, and weather conditions like the presence of strong clouds cover, among others, opening new opportunities to UAV's.

In our study, a low-cost UAV is used to capture RGB crop field images from a fertilization study. In such study, different barley varieties were selected, and three types of fertilization levels were controlled and monitored. In an specific moment, a set of images was captured using a UAV. Then, a deep convolutional neural network is used for automatically extracting features from the images and standard classification techniques for recognition. Extracted features are feed into predictive models that estimate barley variables of interest, namely: fertilization treatment (three rates) and barley yield in grams (real variable). Our study is motivated by the outstanding performance that convolutional neural networks have obtained in a wide diversity of domains, including agricultural imagery. In fact, we show that it is feasible to estimate barley indicators with reasonable effectiveness from RGB images, representing an attractive solution for growers to be informed about the status of their crop fields.

On the one hand, we aim to infer the amount of N in crop fields. This is a highly relevant problem that can be helpful to determine the forthcoming fertilization needs/ treatments and that can have a direct impact into the yield. On the other hand, we aim to estimate the amount of barley a field can yield. This is a critical estimation as it will allow the producer to anticipate the storage requirements, expected profits and production capabilities. Nowadays there is not an abundance of information relating yield prediction based on reliable and accessible techniques for farmers producers. In this work imagery captured by an inexpensive sensor mounted in a UAV is seen as a promising alternative source of knowledge. The yield analyzing technique using models to recognize different fertilization treatments allows us to study the cause and effect relationship between fertilization and crop yield and is a feasible, fast and cheap method; with which the barley production can be improved. To the best of our knowledge, this is the first work that aims at simultaneously estimating these two variables by using RGB aerial imagery and a deep learning methodology.

The main contributions of this paper are as follows:

- We propose a novel methodology for the estimation of two decisive barley variables, fertilization levels and yield from visual information.
- We show that RGB images captured by UAVs are potentially useful for the analysis of agricultural variables.
- We show evidence on the usefulness of deep learning-based methods for the analysis of aerial agricultural imagery. We foresee this work will motivate further research in this direction.
- We provide a low-cost solution for the estimation of barley variables. Despite the proposed solution works in RGB images, it obtains satisfactory performance in the estimation of the considered barley variables.

The remainder of this paper is organised as follows. Next section reviews related work on the different aspects/methodologies involved. Section 3 details data acquisition. Then, Section 4 describes the methodology for data analysis and predictive model construction. Section 5 presents experimental results that aim at evaluating the performance of the developed techniques. Finally, Section 6 summarizes our findings and outlines future work directions.

2. Related work

This section reviews related work on estimation of barley variables by using computer vision and machine learning methodologies. Although there are no works on performing this task from aerial images, we also provide a brief review of related work on the analysis of UAV-based imagery of crop trials. Finally, we also review relevant literature on the use of deep learning for analysis of agricultural imagery.

2.1. *Data analysis from UAV-based imagery of crop trials*

The use of aerial and satellite imagery for the analysis of agricultural variables have been studied for a while, generating extensive studies, see, e.g. (Goel et al. 2003; Jones and Vaughan 2010). A variety of sensors and technologies are used from that scale to collect data, these include hyperspectral and RGB sensors, lidar systems, etc. (see Thenkabail, Lyon, and Huete (2012) and the references therein). In this context, machine learning techniques also have been applied to analyze such data. For instance, in (Uno et al. 2006) an airborne spectrographic imager was used to estimate the yield for corn. The authors relied on standard modelling methods including dimensionality reduction methods and a multilayer perceptron that took as input vegetation indexes extracted from the images. Although highly relevant, images were captured with an expensive device, relied on indexes and focused on corn. Other works have also reported the estimation of related variables using the same type of sensor (Cilia et al. 2014; Goel et al. 2003).

More recently, researchers have relied on (cheaper) imagery captured with multi-spectral airbone sensors mounted over unmanned aerial vehicles (UAVs), see (Zarco-Tejada et al. 2008). In (Bendig et al. 2014), authors estimate barley biomass from imagery captured by a low-cost UAV. Whereas in (Arroyo et al. 2017) the authors processed aerial images obtained with a UAV to determine the nitrogen concentration in corn crops from vegetation indexes. Related works in the same direction are been increasingly published. However, to the best of our knowledge, no work has focused on the estimation of the considered barley variables by using images obtained from low-cost UAVs (for a survey on UAVs application to agriculture see (Zhang and Kovacs 2012)).

2.2. *Barley analysis with computer vision and machine learning*

The analysis of barley data with computer vision and machine learning methods is not new. Regarding fertilization treatments Pagola et al. (2009) presented a methodology based on the analysis of colour in leaf samples (Pagola et al. 2009). Xu et al. (2014) estimated nitrogen concentration based on hyperspectral imagery and the combination of a number of indexes (Xu et al. 2014), still the analysis was restricted to leafs. Finally, in terms of yield, we did not find any reference devoted to its analysis regarding barley. There are, however, other attempts to analyze barley yield, but from a completely different perspective (Broner and Comstock 1997).

From this review, it is evident that there is a gap in the analysis of barley imagery. Few works have been proposed so far, and those existing perform the analysis at a very fine grain. In addition, no work has approached the problem of estimating two variables simultaneously, and more importantly, from aerial imagery. In that sense, our proposed methodology is novel and highly relevant.

2.3. *Deep learning for analysis of agricultural data*

Deep learning is a modelling framework that comprises models with multiple layers of parameters. These solutions have succeeded in many applications from computer vision, to speech analysis to natural language processing. It is only recently that deep learning and representation learning methods, in general, are being used for the analysis of agricultural data. Table 1 provides a summary of the most relevant works.

Grinblat et al. (2016) used a deep convolutional neural network (DCNN) as feature extractor for plant identification from plant vein images. In the same line, Dyrmann, Karstoft, and Midtiby (2016) implemented a DCNN for weed species recognition from closed up and segmented images. (Lu et al. 2017) used pretrained DCNNs for recognizing wheat diseases from close up images. Tang et al. describe a DCNN for weed

Table 1. Summary of related works on deep learning for analysis of agricultural imagery.

Reference	Task	Categories	Data	Network
(Grinblat et al. 2016)	Plant identification	White bean, red bean, soybean	Vein Images	2-Convolution layers
(Lu et al. 2017)	Wheat analysis	Healty/infected wheat (7)	In-field images	VGG/VGG-16
(Tang et al. 2017)	Soybean analysis	Weed (3) and soybean	In-field images	5-Convolution layers
(Cheng et al. 2017)	Plant with pests	Crop pests(10)	In-field images	AlexNet
(Dyrmann, Jorgensen, and Midtiby 2017)	Weed recognition	Weed vs no weed	In-field images	GoogleNet
(Dyrmann, Karstoft, and Midtiby 2016)	Plant identification	22 weed categories	In-field images	7-Conv. layers
This work	Barley analysis	Fert. (3), variety (6), yield (real)	Aerial images	AlexNet

identification (Tang et al. 2017), similarly, (Dyrmann, Jorgensen, and Midtiby 2017) describes a DCNN-based method for weed detection. (Cheng et al. 2017) used a residual DCNN for pest recognition in plants. All of these works are evidence that DCNN has a great potential to be used in the analysis of agricultural imagery, even when there are inherent difficulties when working with these types of images. Namely, homogeneous categories, complex backgrounds lack of annotated images, dealing with low-resolution images, among several others. We foresee deep learning will be established as the *de facto* image analysis methodology in agriculture in the forth-coming years.

Several interesting facts can be highlighted from Table 1. Firstly, it is noticeable that these works have been published during the last year. Being DCNN a cutting edge methodology that is being increasingly used in agricultural imagery analysis. Secondly, the variety of tasks and categories considered in the studies are quite diverse, mostly focusing on plants and weed, but also with applications in cereals. The variety of DCNN architectures is quite diverse as well. Finally, it is very interesting that among the surveyed methods, no article was found using deep learning methodologies for the analysis of aerial imagery. In fact, most works have focused on closed up images. Aerial images represent a more complicated problem, but also, one with greater impact (the areas that can be covered with our methodology considerably larger than those that can be covered within field techniques).

Moreover, a key feature of deep learning models is that they can automatically learn/extract representations from low-resolution raw data and such representations have proved to be particularly helpful for predictive models in a number of tasks, including precision agriculture. Herein, we show encouraging results using this methodology for the analysis of UAV-based imagery of crop trials. It is important to emphasize that we obtain acceptable predictive performance by using images that are easy to collect and that do not require of specialized sensors/lenses. Making it attractive for the average grower.

2.4. *Discussion*

We have reviewed related work in three core aspects around the proposed method. From the review, it is clear that there is a gap that this work aims to fill. In particular, we can state that the main novelty of our work relies on the following directions: (1) *simultaneous* analysis of two barley variables, (2) *from aerial imagery* and, (3) *using deep learning* feature extractors applied to RGB images. To the best of our knowledge, there is no current effort in this direction. Therefore, this paper can have a great impact into precision agriculture. What is more, as shown in Section 5, the obtained results using RGB imagery only are quite competitive. Thus, we foresee this article will pave the way for the analysis of similar agricultural variables with deep learning methods.

3. Data acquisition

In the last decade, there has been an increase of remote sensing applications with UAVs for precision agriculture rather than conventional satellite imagery (Hunt et al. 2008; Odido and Madara 2003; Tripicchio et al. 2015; Valente et al. 2011; Zhang and Kovacs 2012; Huang et al. 2013). The use of UAV cameras, added sensors, and its integration with Machine learning techniques has interesting potential for within-season crop management. For this analysis, we rely on RGB imagery captured with a low-cost aerial vehicles capable of covering large portions of fields in an affordable way.

This section then provides details on the UAV-based imagery captured. First, the barley crop study is described, and later the characteristics of the equipment used in the UAV are presented. Finally, this section ends with information about specific traits on the UAV-based imagery acquisition.

3.1. *Barley crop study*

An experiment on fertilization dose for different barley crop was conducted at the Agronomy Research Field on the Agronomy Campus of the Universidad Autonoma de Nuevo Leon at Marin, in the state of Nuevo Leon, Mexico. The location of the field is 25° 53′ 36″ N, 100° 02′ 38″ W, with an altitude of 370 m above the sea level. Figure 1 shows an aerial image of a crop field that was subject of study. To create various crop growth scenarios, 6 barley (Hordeum vulgare) crops were grown (V1 = Cuahutemoc, V2 = Menonita, V3 = Mezcalera V4 = Marn, V5 = Chichimeca, and V6 = UANL-138), under three nitrogen fertilization treatments (0, 100, 200 kg N ha^{-1}), and four replications each. Sample images for the different barley variables are provided in Figure 3. A total of 72 rectangular plots were established for harvesting, each one belongs to a specific type of barley and a level of nitrogen fertilization. The field consisted of soil preparation with tracking and furrowing operations, using a randomized block design, six furrows of 5 m long, and 0.7 m among furrows. The UAV-based imagery of the barley crop was captured at the time when the barley crop showed physiological maturity suitable for harvesting.

3.2. *Experimental platform*

The UAV used to capture the image is an hexacopter that we built, integrated with six brushless motors of 700 KVAs and four electronic speed controllers of 40A. The guidance and control of the UAV are developed by the autopilot (the Pixhawk of 3D

Figure 1. Aerial image of the barley crop field. Latitude: 25.85, Longitude: −100.02.

Robotics) that has been connected to a DX7s 7-Ch DSMX Radio System of Spektrum and to a telemetry to know all the movements and trajectories of the hexacopter in real time. The Parrot Sequoia is the multispectral sensor selected for the image capture, which possess five channels, four channels collect the following bands: Red, Green, Red-Edge and Near-Infra-Red with 660, 550, 735 and 790 nanometers of (central) wavelengths, all of them with a resolution of 1.2 Mpx. Additionally, a RBG channel of 16 Mpx, see Figure 1. For this study, only RGB imagery was used, this is because we wanted to explore the extend to which this information carries out predictive information for barley variables estimation.

3.3. *Field conditions*

The image was captured at 24.4 m of altitude with a Ground Sampling Distance (GSD) of 0.67 cm per pixel, that is one pixel of the image represents linearly 0.67 cm on the ground, this high resolution allows us a wealth of information required for the recognition purposes. The formula to calculate the altitude with the GSD is given by GSD = pixel size × AGL/Focal length; where pixel size = Sensor size (width or height in mm)/sensor resolution (number of pixels in either width or height). AGL is the elevation above the ground in meters and the focal length is in millimetres. The RGB channel of the Parrot Sequoia has a sensor resolution of 4608 × 3456 pixels, a focal length of 4.88 mm for pixel size of 1.34 μm; for a GSD of 0.67 cm per pixel, the necessary AGL is 24.4 m. The sensor was placed at the frame centre below of the hexacopter with a horizontal angle. For the mosaic given in Figure 1 a total of 52 individual images where taken on 4 May 2017; for the building of the mosaic, the software Pix4d® was used. The photos were taken at the stage of physiological maturity that is suitable for harvesting, (at this stage, the differences between the varieties of barley are more evident, and we support ourselves with a grain moisture analyzer to initiate the harvest moment). The characteristics and the orientation of the terrain where left-hand side of the picture is the north gives the tilt of the map.

4. Methodology

This section describes the proposed methodology for the analysis of barley imagery acquired as described in the previous section. We approach three related tasks, in each of them, RGB information exclusively is used to extract an image representation that is in turn used with predictive models. The overall approach is depicted in Figure 2.

Images captured with the UAV are preprocessed and feed to a deep convolutional neural network, the response of an intermediate layer of the network with the image produces a vectorial representation for the image. This vector is then used with standard classification and or regression models to estimate the different variables of interest. The remainder of this section provides details on the studied variables, the feature extraction process and the predictive model estimation.

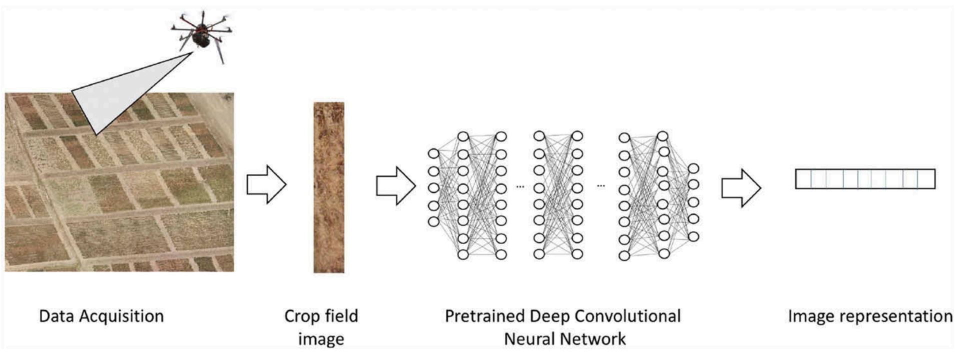

Figure 2. General diagram of the feature extraction process.

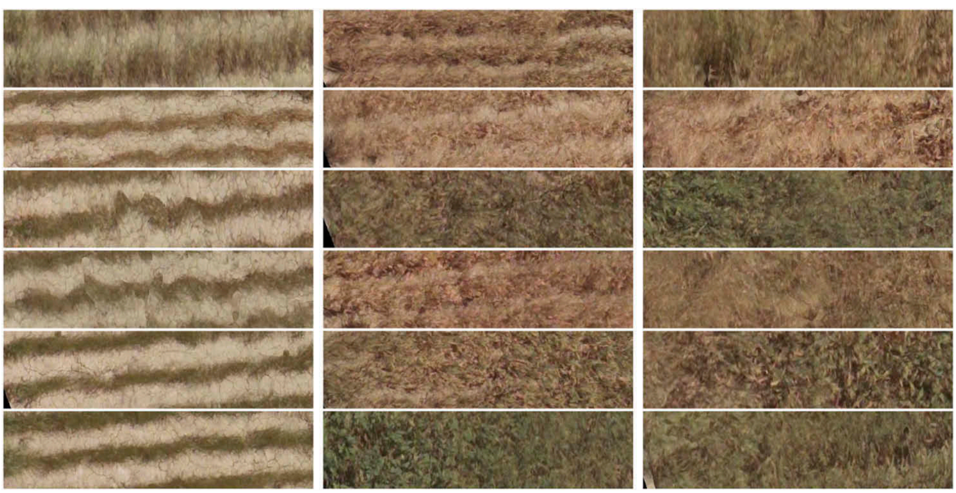

Figure 3. Sample images for the different barley varieties (rows) and fertilization treatments (columns).

4.1. *Barley variables*

As previously mentioned, the barley variables of interest are as follows:

- *Fertilization treatment estimation*. To determine the amount of nitrogen according to three predefined categories: 0, 100, 200 kg N ha^{-1}
- *Barley yield estimation*. To estimate the yield (in grams) that each field can produce.

The first variable corresponds to a classification problem, whereas the second one is a regression one. Regardless of the variable, the same feature extraction procedure was adopted. Our aim in doing so, was to assess the actual benefit of pretrained models and the generalization of features learned with them.

Sample images for the different barley variables are provided in Figure 3. Fertilization treatments are shown in columns, whereas the six barley varieties are shown as rows in the plot. It can be seen from this plot that there are some combinations of fertilization-variety that are extremely difficult to visually identify (e.g. all crop field images in the leftmost column). This illustrates the difficulty of the approached problem.

The relevance of predicting these variables is clear: having a model that autonomously indicates the nitrogen level, and, most importantly, the expected yield of a crop field would have an important economic impact. This is because it can be known in advance what crop fields require of more nitrogen fertilization and which ones are over fertilized. Providing the farmer of relevant information to make decisions on the treatment she/he applies to each crop field. Also, and most interestingly, one could know in advance the storage requirements for barley as well as an estimate on the profits and production capacity. Reducing costs and anticipating earnings. All of this, starting with a RGB image obtained with a simple machine.

4.2. *Feature extraction*

As previously mentioned, one of the main novelties of this work is the usage of features automatically derived from an image to predict barley variables. Representation learning is the subfield of machine learning that studies models that can extract automatically a descriptive representation from raw data (Bengio, Courville, and Vincent 2013). These types of methods are advantageous whenever it is not evident how to represent data in a way that descriptive or discriminative information is captured. Notable achievements of representation learning have been reported elsewhere (Hinton, Osindero, and Teh 2006; Krizhevsky, Sutskever, and Hinton 2012), of particular interest are those based on deep convolutional neural networks (DCNNs) (Krizhevsky, Sutskever, and Hinton 2012; Vinyals et al. 2015; Simonyan and Zisserman 2015).

DCNNs are layered models mostly used for image analysis (Lecun, Bengio, and Hinton 2015). They are called deep, because they are composed of multiple layers of learnable parameters. Where parameters can capture spatial, appearance and texture patterns. The distinctive feature of DCNNs is that they contain layers of convolutional parameters (i.e. they learn convolutional filters at different layers) that together with non-linear activations and pooling operations can learn quite useful image representations. Coupled with efficient learning via stochastic gradient descend, regularization mechanisms (e.g. dropout) and the availability of large amounts of images have provoked these models to be the state of the art solution for applications involving scene understanding. Accordingly, in this work, we use a DCNN for feature extraction from barley aerial imagery. To the best of our knowledge, this is the first work that uses deep learning for feature extraction in this sort of data.

Learning the parameters of a DCNN (which allow one to extract features from images automatically) is not an easy task. In fact, large collections of images (of the order of millions) have been used in the past for training successful DCNN models (Krizhevsky, Sutskever, and Hinton 2012; Vinyals et al. 2015; Chatfield et al. 2014; Simonyan and Zisserman 2015; Lecun, Bengio, and Hinton 2015). In our case, imagery is scarce hence this is not an option. However, there is an alternative for domains with limited data: *transfer learning from pretrained models* (Yosinski et al. 2014). The idea of this formulation is to get a model learned for a related task (e.g. object recognition) and use it to extract features for the domain of interest (e.g. agriculture or medical images). The reference model is called pretrained, because its parameters were already learned for a related task. Hence, an image of the domain of interest is feed into the pretrained model. Usually, the response of an intermediate (instead of the last) layer of the DCCN is taken. This is because the final layers are closely related to the task at hand and provide limited generalization capabilities. Figure 4 illustrates the pretraining strategy to feature extraction.

In this work, we relied on generic models pretrained for image classification. Specifically, we considered models trained for recognizing the 1000 generic classes

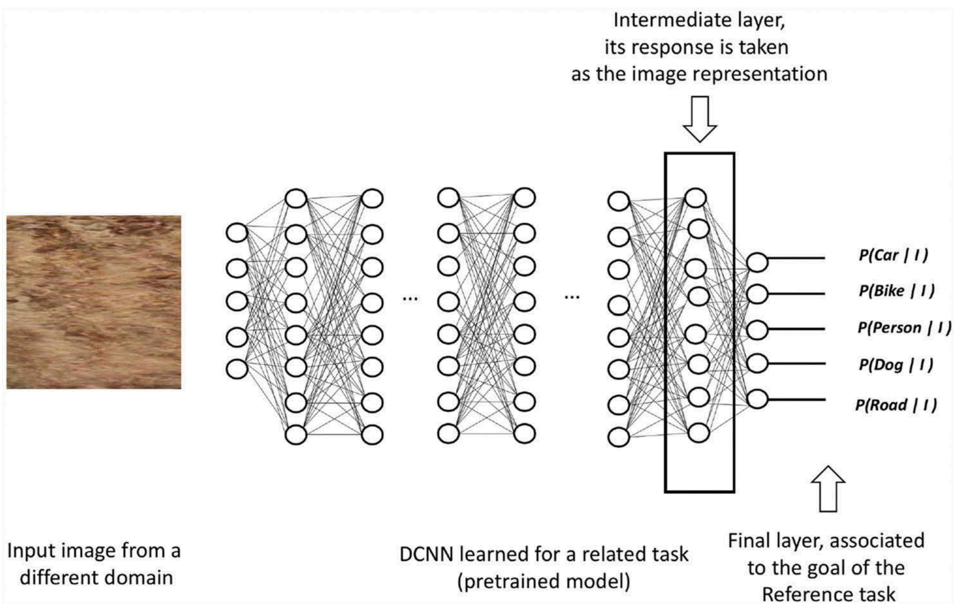

Figure 4. Illustration of the pretraining-based approach to feature extraction with DCNNs. A DCNN model pretrained for another task is obtained, and the image of the domain of interest is then feed into the model. The response of the DCNN at an intermediate layer is used as the descriptor of the image. Please note that the response of the final layers is discarded as it usually corresponds to the initial task.

Table 2. Description of the pretrained models considered in this study.

DCNN	Layers	1st layer filters	Layer	Reference
AleXNet	8: 5 conv./3 FC.	96	Relu	(Krizhevsky, Sutskever, and Hinton 2012)
VGG	8: 5 conv./3 FC.	64	Relu	(Chatfield et al. 2014)
VGG-19	19: 16 conv./3 FC.	64	Relu	(Simonyan and Zisserman 2015)

from ImageNET (Russakovsky et al. 2015). The parameters of the reference models where trained with millions of images, and they have proved to be extremely helpful for image classification, see (Krizhevsky, Sutskever, and Hinton 2012; Chatfield et al. 2014; Simonyan and Zisserman 2015). In addition, the considered models have been also widely used for transfer learning in a number of tasks (Mollahosseini, Chan, and Mahoor 2016; Tian et al. 2015; Pellegrin et al. 2017; Karpathy and Fei-Fei 2015). However, to the best of our knowledge, this is the first time they are used for agriculture image classification.

A description of the considered pretrained models is provided in Table 2. The three considered models comprise a set of convolutional layers (including non-linear activations and pooling with stride) and fully connected layers. In fact, all three models have three fully connected layers: the last layer associated to the 1000 classes, and the two pre-last layers of dimensionality 4096. In this paper, we used the output of the pre-last layer as image descriptor (see Figure 4). Each specific DCNN model has their own configuration of convolutional layers, strides, and pooling, we refer the reader to the corresponding references for more information. What it is important to mention is that VGG-19 differs from the other models in that it is deeper: 19 layers vs 8 layers of AlexNet and VGG. The latter two models differ in that VGG uses a larger stride and was further fine tunned with additional data.

We used the pretrained implementations from the MatConvNet framework (Vedaldi and Lenc 2015).The input for the three models is an image of size 240×240 pixels, since the crop field images have a different size (but all of them have the same size to each other), we used bicubic interpolation to resize the original images. Hence, each crop field image is processed as follows for feature extraction: (1) the image is resized to 240×240 pixels; (2) the resultant image is feed into the network, and the response of the weights in each layer is calculated; (3) the output of the pre-last layer is then used as image descriptor (a 4096 dimensional vector). The image descriptors based on pre-trained DCNNs are then used with standard models for estimating the barley variables.

4.3. *Learning predictive models and validation*

As previously mentioned, once that features have been extracted for each image, these are feed into classification and regression models that aim to learn a mapping from the image descriptor into the barley variables of interest. This section provides a description of such models.

As we mentioned earlier, the fertilization estimation is a classification problem, it maps the 4096-dimensional feature vector into the three categories of interest. In this work, we considered the classifiers described in Table 3. These are the most representative ones and cover most of the machine learning paradigms (i.e. probabilistic, instance

Table 3. Description of the classifiers considered for classification experiments.

Classifier	Description	Parameters
Naïve Bayes	Probablistic classifier, assumes independence between attributes	-
Bayesian Net	Probabilistic model, conditional independence is exploited	-
SVM	Support vector machine classifier, SMO implementation	Linear kernel
ANN	Multilayer Perceptron	50 Hidden units, 100 epochs, $\eta = 0.3$
Adaboost	Boosting ensemble	Decision stumps as weak learners
Logitboost	Additive logistic ensemble	Decision stumps as weak learners
Bagging	Bagging ensemble	REPTrees
1NN	1-Nearest neighbor classifiers	$k = 1$
3NN	3-Nearest neighbor classifiers	$k = 3$
LWR	Locally weighted regression	Decision stump as learner

Table 4. Description of the regression models considered for experiments. All models, but one (last row) was taken from Weka (Hall et al. 2009); whereas model ANN-C was taken from the CLOP clop toolbox (Saffari and Guyon, 2006).

Regressor	Description	Parameters
SVR	Regression with support vector machines	Linear kernel
GP	Gaussian process	-
MLP-W	Multilayer Perceptron	50 Hidden units, 100 epochs, $\eta = 0.3$
Simple LR	Linear regression using a single variable	-
1NN	1-Nearest neighbor classifiers	$k = 1$
Random forest	Ensemble of decision trees	Decision stumps as weak learners
LR	Linear regression	-
Bagging	Bagging ensemble	REPTrees
MLP-C	Multilayer Perceptron	50 Hidden units, 100 epochs, $\eta = 0.3$

based, linear-non-linear functions, ensembles, etc.). Default parameters were used in each case and the Weka implementations were adopted (Hall et al. 2009).

Regarding yield estimation, this is a regression problem, accordingly the considered models are described in Table 4. As before, we considered the most representative methods from the state of the art.

For validation, and given the scarcity of data (recall we have only 72 samples available), we relied on leave one out cross validation for estimating the performance. Under this procedure, all but one sample is used for training the model, then the trained model is used to predict the variable of interest in the leaved-out (test) instance. This process is repeated 72-times, each time changing the test sample. These values are averaged and reported. The next section reports experimental results obtained with the models described in this section using the leave one out validation scheme.

5. Experiments and results

This section describes experimental results that aim at evaluating the performance of the predictive methods described in Section 4. We have divided results into two sections covering different aspects. After presenting experimental settings, we analyze the performance of models for recognizing fertilization treatment. Next, we evaluate the performance of regression models that estimate the barley yield. Then, we analyze the pretrained models trying to get insights into the features of each model. We close this section with a discussion on the analysis of the main findings of this work.

5.1. *Experimental settings*

For the evaluation of the predictive models of barley related variables, we followed a standard machine learning evaluation protocol in which separate partitions were used for learning the model and evaluating their predictive performance. As previously mentioned, we adopted a leave one validation protocol. Regardless of the variable to be predicted and the considered classifier, the same protocol was adopted.

Each sample was processed as described in Section 4: images were scaled and feed into a pretrained model, then the activation up to certain layer is taken as the representation of the image; representations are used to train and evaluate classification (for fertilization) or regression models (for yield estimation).

5.2. *Fertilization*

First, we analyze the recognition performance of a set of classification models for recognizing the fertilization treatment.

Figure 5 shows the results of this experiment for fertilization. We report the average leave one out performance for each of the considered classifiers (see Table 3) and for each of the pretrained models. Where performance is measured by the accuracy, i.e. the percentage of correctly classified images.

From Figure 5, it can be seen that the best performance was obtained by Adaboost, ANN and SVM, with accuracies of 83.3%, 81.9%, and 80.5%, respectively. These are very positive results, if we consider only visual information (extracted from the RGB channels) is being used and that three classes are available (random guessing will result in 33% of accuracy). Regarding the different pretrained models, it can be seen that, overall, better performance was obtained with the AlexNet model, followed by VGG and VGG-19. This result, seems to indicate that the performance of methods is tied to the complexity of the model. More layers in VGG and VGG-19 seem to cause a decrement in the classification performance. This can be due to the fact that more complex models are more *tailored* to the task they were designed for, so, a simpler model, can be more generic and capture useful information for representing images.

Table 5 shows the confusion matrix for the best classifier (Adaboost). It can be seen that the control class A (i.e. no treatment) is the one that is better recognized, whereas the most difficult one is C (200 kg N ha^{-1}), which is commonly confused with class B (100 kg N ha^{-1}). This is a somewhat expected result, given the visual similarity of fields with both treatments, see Figure 3.

5.3. *Barley yield estimation*

In the next experiment, we aimed to evaluate the performance of different regression methods when predicting the barley yield directly from field images.

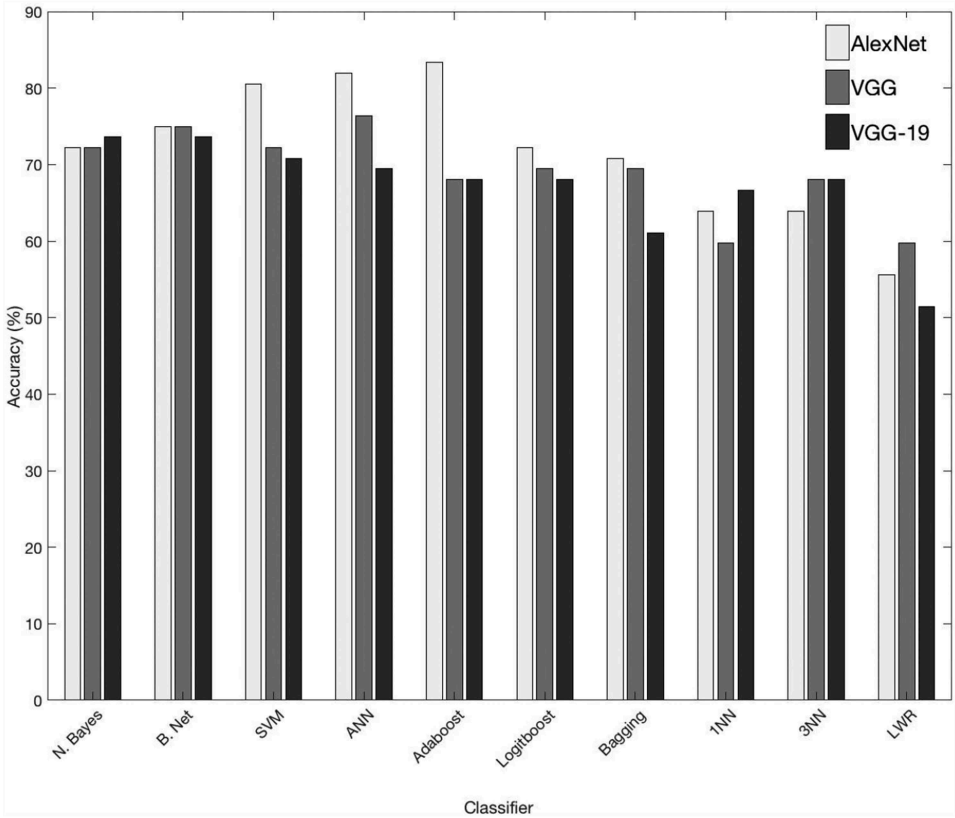

Figure 5. Leave one out performance for the considered classifiers when using different pretrained networks for recognizing the fertilization treatment.

Table 5. Confusion matrix for the SVM classifier when predicting fertilization treatment.

		Predicted		
True		A	B	C
	A	95.8%	4.2%	0.0%
	B	0.0%	83.3%	16.7%
	C	0.0%	29.1%	70.9%

The leave one out performance of the considered models (see Table 4) for the prediction of barley yield is reported in Figure 6. In these plots, we report the correlation coefficient and the root mean squared error between the real values and the predictions from the models.

It can be seen that in this case there is not a clear trend in the results. Regarding the correlation coefficient, we can see that several models exceed a correlation of $\rho = 0.5$, evidencing that these models capture the general behaviour of the barley yield variable. The combination of regressor and pretrained model MLP-W and VGG achieves a $\rho = 0.63$, which is a quite positive result. Other models with similar performance are SVR and GP. Regarding pretrained models, the only trend that can be seen from the

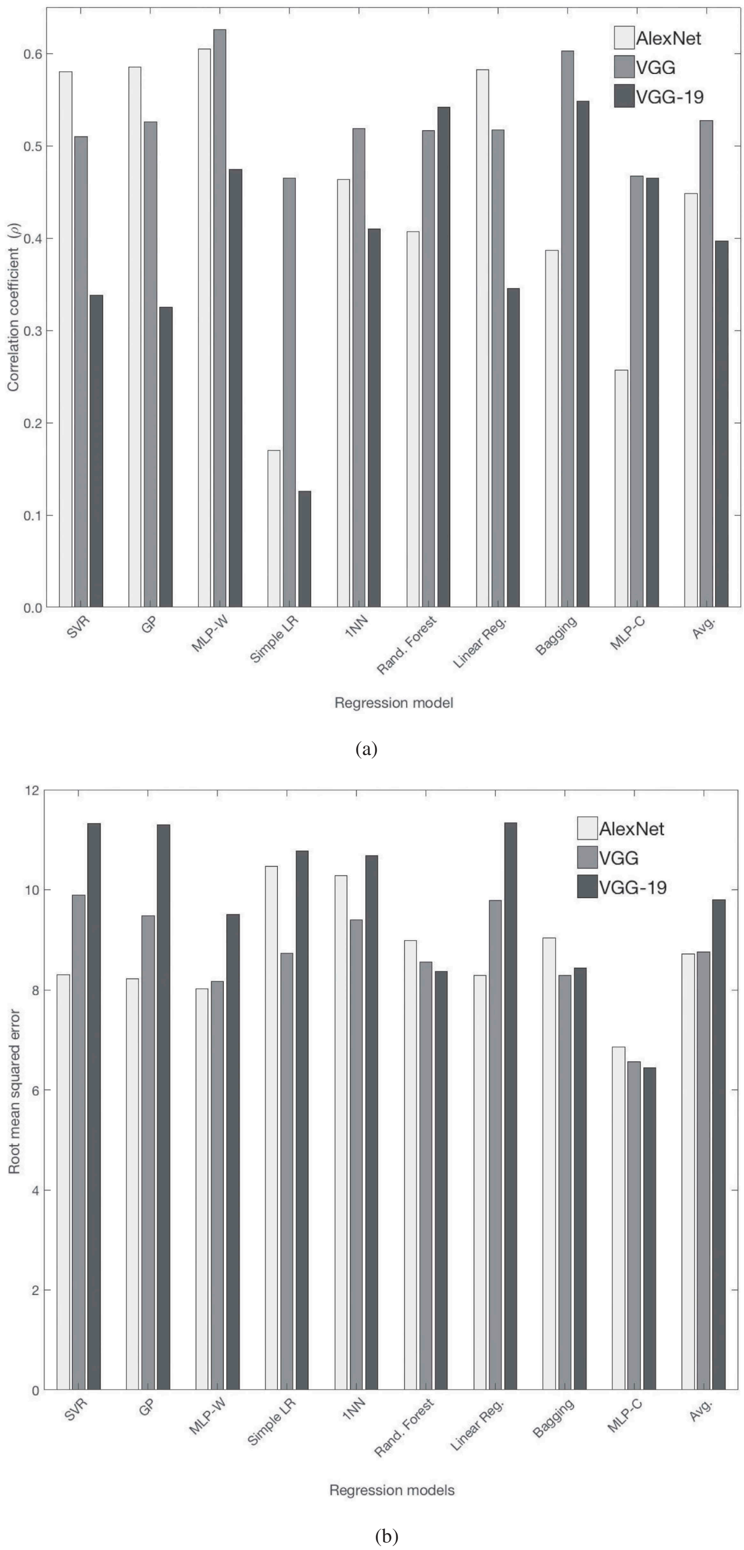

Figure 6. Leave one out performance for the considered regression models when using different pretrained networks. (a) Correlation coefficient; (b) RMSE.

results is that the worst correlation was obtained by the most complex model VGG-19. This is in agreement with results in Section 5.2. However, the performance between the other models follows a different trend, in fact, VGG seems to obtain a better correlation coefficient.

Regarding the RMSE, results are somewhat similar (inversely related, obviously) and complement correlation analysis. Interestingly, this time the best performance was obtained by MLP-C. This is interesting, because the correlation of this model was relatively low when compared to other models. To further analyze the performance on barley yield estimation, Figure 7 shows the per-sample raw error obtained by the model with the lowest RMSE (MLP-C). It can be seen that the magnitude of errors varies across samples. In fact, there are only a few samples for which the error is large, whereas for many samples the error is acceptable.

For the sake of analyzing the performance per fertilization treatment and variety, Figure 8 shows the correlation coefficient and RMSE per fertilization treatment and barley variety. From this figure, it can be seen that variety 2 is the one for which error is the highest, while the lowest correlation corresponded to a variety 5. Interestingly, fields without fertilization treatment (A), obtained the highest error.

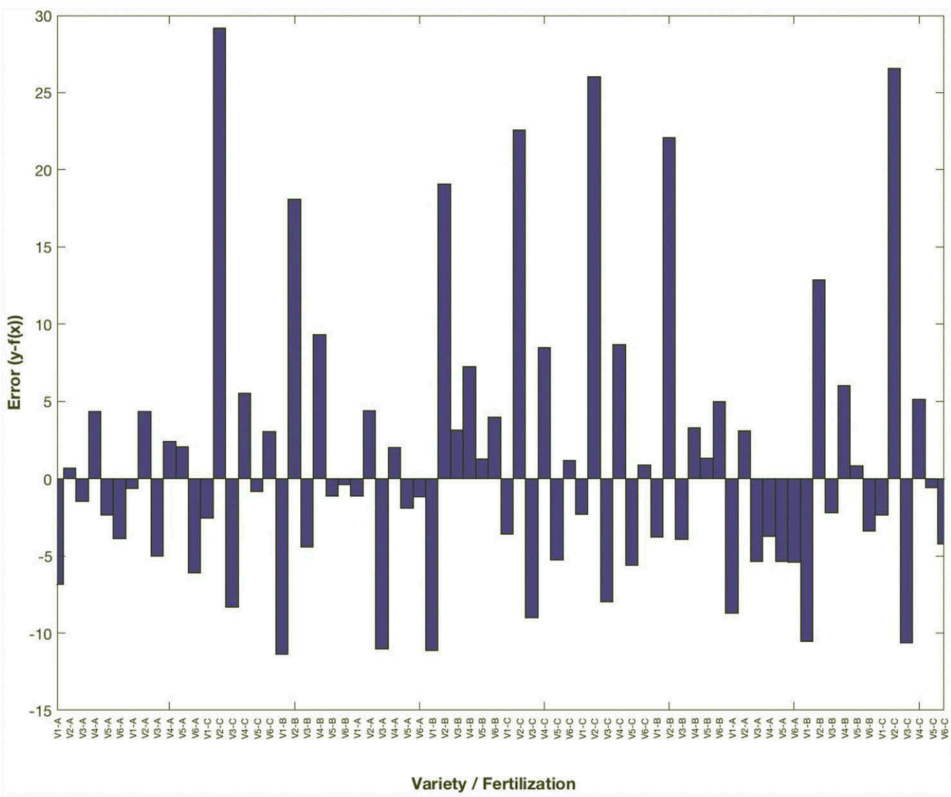

Figure 7. Leave one out raw error per sample for the method that obtained the lowest RMSE (MLP-C).

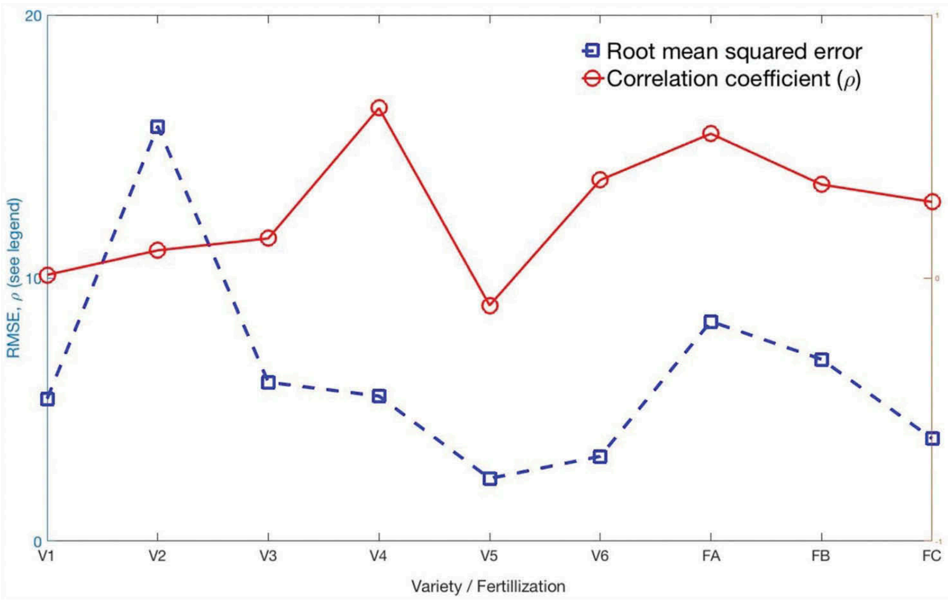

Figure 8. Leave one out correlation coefficient and RMSE per type fo sample for the method that obtained the lowest RMSE (MLP-C).

5.4. *Analysis of pretrained models*

In this section, we analyze the response of the first layer of filters of the different pretrained models. The goal of this section is to gain insights into the suitability of the different models for estimating the barley variables considered in the study. Figures 10, 11 and 12, respectively, show the response of AleXNet, VGG and VGG-19 pretrained models when feed the image shown in Figure 9.

Although it is difficult to interpret and analyze the usefulness or descriptiveness of each response, it is clear that the different filters highlight and capture complimentary information that together can effectively describe the content of crop field imagery. In fact, we observe that filter responses are quite similar across models. Perhaps having 96 filters from AlexNet is advantageous over the 64 filters from the VGG and VGG-19 models, as more filters can capture more low-level information from the image. This may be another reason on why AlexNet obtained better performance on the estimation of fertilization treatment and variety.

Figure 9. Crop field image used for the analysis of this section.

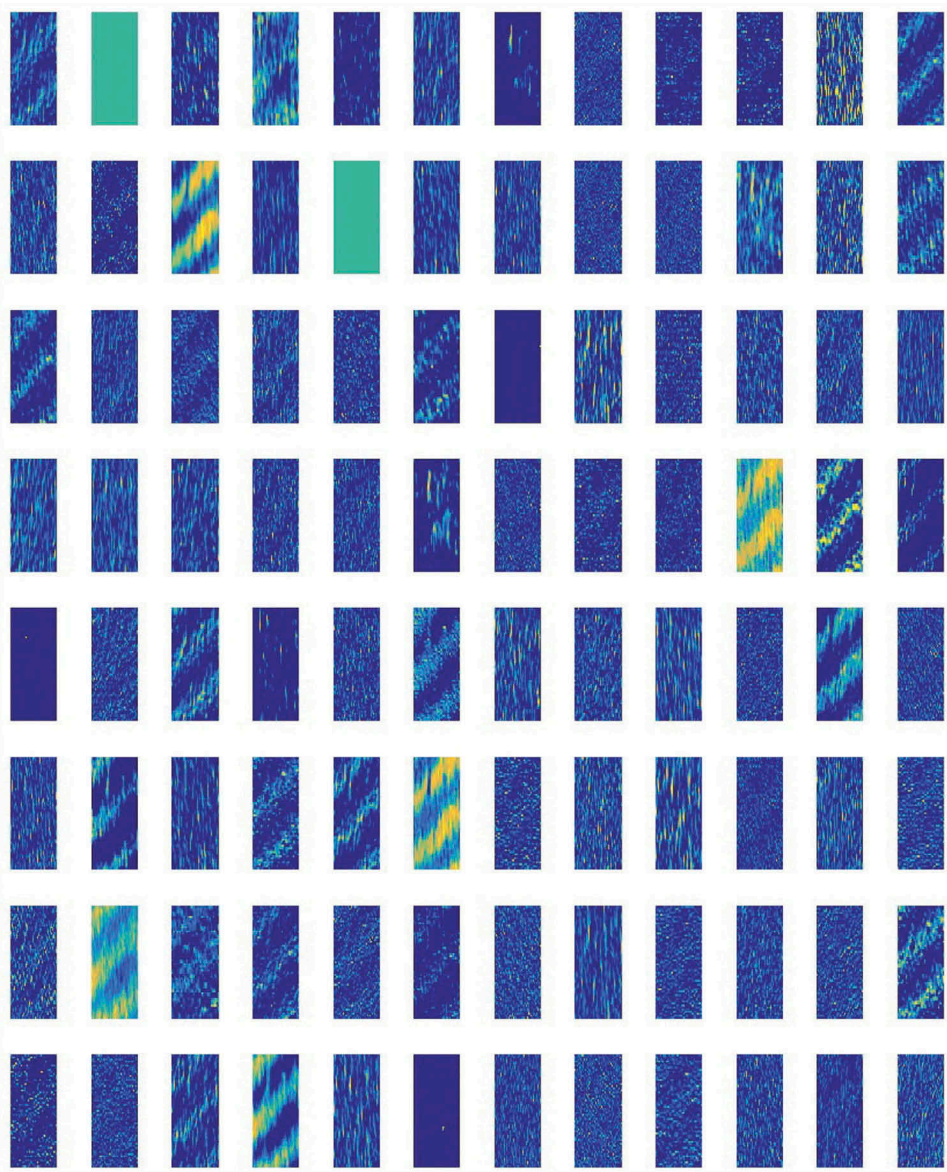

Figure 10. Visualization of the response of the first layer of filters of the AlexNet model for an input image. Ninety-six filters are available.

5.5. *Discussion*

Results obtained in this section comprise new knowledge will have a positive impact into barley variables analysis with computer vision techniques.

It was demonstrated the feasibility of the identification of fertilization treatment from aerial RGB images. The pretrained architectures were useful for obtaining image descriptors with highly discriminative information. Performance above 80% in this barley variable was obtained, where it was easier to recognize crops without nitrogen, and

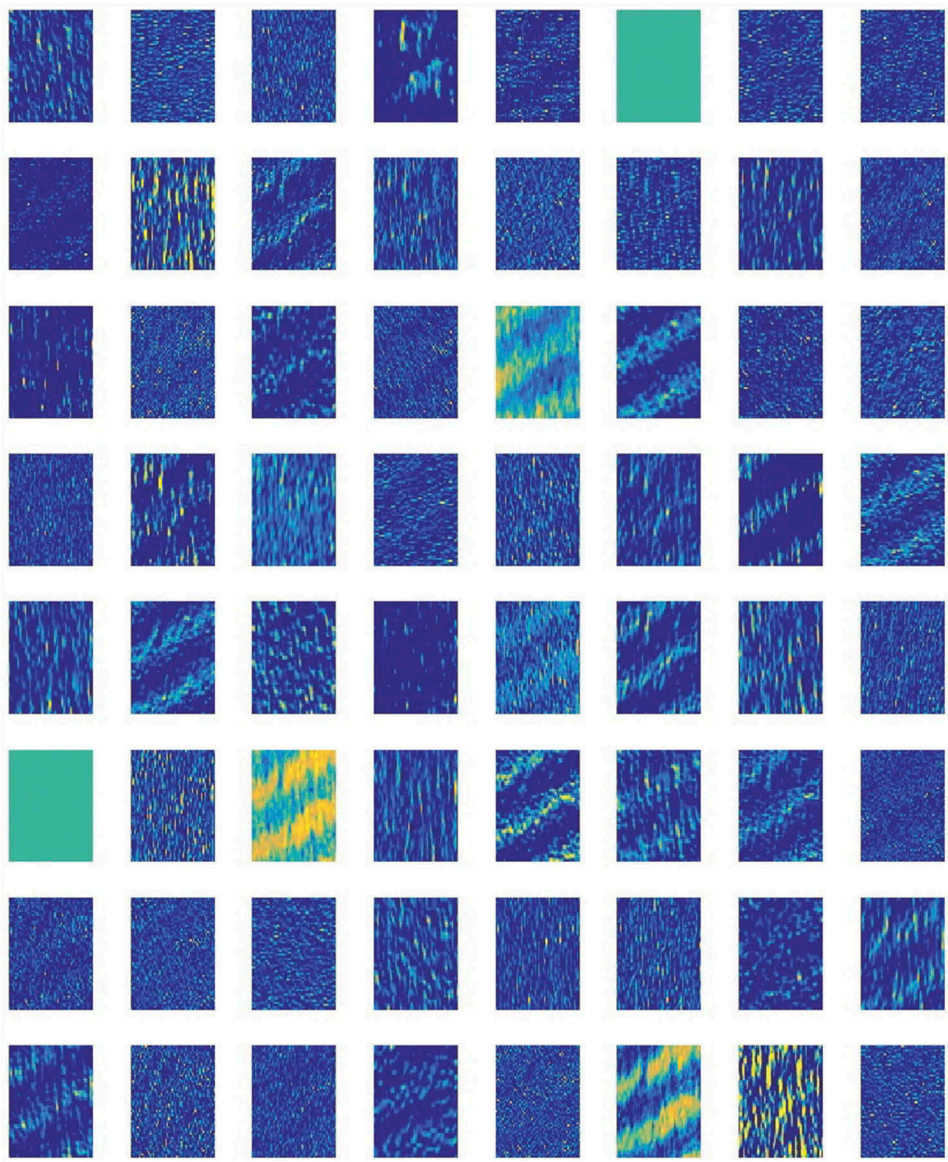

Figure 11. Visualization of the response of the first layer of filters of the VGG model for an input image. Sixty-four filters are available.

where other treatments were slightly more difficult to recognize. It was found that a less tailored/complex model was more helpful for this problem. Suggesting that a model trained with agricultural images (even from different grains) can have a greater impact in the task.

Results on yield estimation, on the other hand, were promising. High correlation and low RMSE was obtained with certain combinations of models and pretrained model. This result is encouraging and can have an important impact in the economy of barley

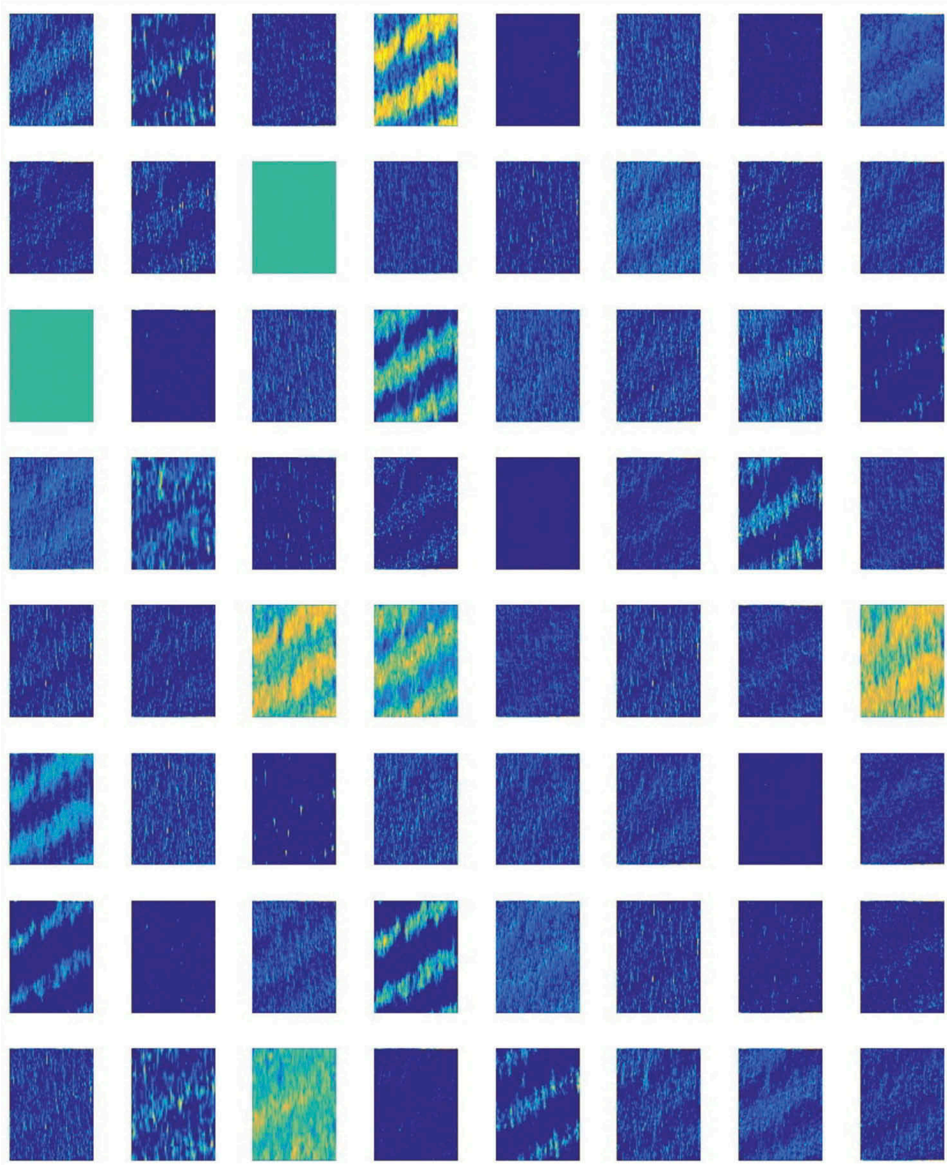

Figure 12. Visualization of the response of the first layer of filters of the VGG-19 model for an input image. Sixty-four filters are available.

agriculturists: by having a low-cost method that can estimate the yield with this degree of precision.

In general terms, we have reported evidence on the usefulness of deep learning-based features for the analysis of aerial imagery of barley. Our results will motivate further research on the analysis of different grains, on the usage of other deep learning formulations and on the use of aerial imagery in problems related to precision agriculture.

6. Conclusions

We have described a novel methodology for the simultaneous estimation of two barely related variables from aerial imagery captured by a UAV device. Low-cost imagery is captured by a UAV and pretrained deep learning convolutional neural networks are used as feature extractors. Then, standard classification and regression models are used for prediction. The conclusions derived from this work are as follows:

- Recognizing the fertilization treatment of different varieties of barley from RGB aerial imagery and with acceptable performance is possible.
- Estimation of yield with the proposed methodology results in acceptable estimates that seem to vary according to the barley variety.
- Pretrained deep learning models resulted extremely helpful to characterize aerial RGB barley imagery. A similar methodology was adopted to estimate two barley variables without relying on hyperspectral imagery and/or vegetation indexes. Better results are expected when including the latter information.
- UAV data have shown to provide information that, can jointly work with machine learning tools to efficiently automatize fertilization studios.
- The use of remote sensing with satellites have historically been restricted by the cost and availability of fine resolution data, but with the use of low-cost UAV, these impediments are solved.
- In this paper, we show that even with a relative small amount of data and images of the experimental field; the features of the pretrained DCNN models are not limited, in fact, the representation is able to capture descriptive information that yields to good performance.

In addition to these interesting conclusions, we emphasize that one of the most important outcomes of our study is providing evidence of **acceptable performance when predicting variables of practical relevance for agriculturist, all of this, starting from RGB imagery that is captured by low-cost UAVs**. Therefore, results from this paper can have great impact into society and economy.

Several future work directions are envisioned by the authors. On the one hand, we are working on training and end to end DCNN using UAV-based imagery of crop trials. We think a model trained on this type of images could have a huge impact in different fields, including biology, ecology, and forestry. Also, we are studying methods that can take advantage of pretrained DCNNs enhanced with information from vegetation indexes and hyperspectral imagery.

Acknowledgments

The authors wish to acknowledge the financial support of the Secretary of Environment and Natural Resources and National Council of Science and Technology of Mexico (SEMARNAT-CONACYT) under project 263080, the CB-CONACyT under grant CB-2014-241306, and the financial support of the thematic network BigDSS-Agro (P515RT0123) part of the CYTED program.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Consejo Nacional de Ciencia y Tecnología [263080, CB-2014-241306, SEMARNAT-CONACYT 263080]; BigDSS-Agro [P515RT0123].

References

Arroyo, J. A., C. Gomez-Castaneda, E. Ruiz, E. Munoz de Cote, F. Gavi, and L. E. Sucar. 2017. "Assessing Nitrogen Nutrition in Corn Crops with Airborne Multispectral Sensors." In *IEA/AIE 2017*, Part II, Vol. 10351 of *LNAI*, 259–267. Arras, France.

Bendig, J., A. Bolten, S. Bennertz, J. Broscheit, S. Eichfuss, and G. Bareth. 2014. "Estimating Biomass of Barley Using Crop Surface Models (Csms) Derived from UAV-Based RGB Imaging." *Remote Sensing* 6 (11): 10395–10412. doi:10.3390/rs61110395.

Bengio, Y., A. Courville, and P. Vincent. 2013. "Representation Learning: A Review and New Perspectives." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (8): 1798–1828. doi:10.1109/TPAMI.2013.50.

Broner, I., and C. R. Comstock. 1997. "Combining Expert Systems and Neural Networks for Learning Site-Specific Conditions." *Computers and Electronics in Agriculture* 19: 37–53. doi:10.1016/S0168-1699(97)00031-8.

Chatfield, K., K. Simonyan, A. Vedaldi, and A. Zisserman. 2014. "Return of the Devil in the Details: Delving Deep into Convolutional Nets." *Proceedings of BMVC* 1–12, Vol 1.

Cheng, X., Y. Zhang, Y. Chen, W. Yunzhi, and Y. Yue. 2017. "Pest Identification via Deep Residual Learning in Complex Background." *Computers and Electronics in Agriculture* 141: 351–356. doi:10.1016/j.compag.2017.08.005.

Cilia, C., C. Panigada, M. Rossini, M. Meroni, L. Busetto, S. Amaducci, M. Boschetti, V. Picchi, and R. Colombo. 2014. "Nitrogen Status Assessment for Variable Rate Fertilization in Maize through Hyperspectral Imagery." *Remote Sensing* 6 (7): 6549–6565. doi:10.3390/rs6076549.

Dyrmann, M., H. Karstoft, and H. S. Midtiby. 2016. "Plant Species Classification Using Deep Convolutional Neural Network." *Biosystems Engineering* 151: 72–80. doi:10.1016/j.biosystemseng.2016.08.024.

Dyrmann, M., R. N. Jorgensen, and H. S. Midtiby. 2017. "RoboWeedSupport - Detection of Weed Locations in Leaf Occluded Cereal Crops Using a Fully Convolutional Neural Network." *Advances in Animal Biosciences: Precision Agriculture* 8 (2): 842–847. doi:10.1017/S2040470017000206.

Goel, P. K., S. O. Prasher, J. A. Landry, R. M. Patel, R. B. Bonnell, A. A. Viau, and J. R. Miller. 2003. "Potential of Airborne Hyperspectral Remote Sensing to Detect Nitrogen Deficiency and Weed Infestation in Corn." *Computers and Electronics in Agriculture* 38 (2): 99–124. doi:10.1016/S0168-1699(02)00138-2.

Grinblat, G. L., L. C. Uzal, M. G. Larese, and P. M. Granitto. 2016. "Deep Learning for Plant Identification Using Vein Morphological Patterns." *Computers and Electronics in Agriculture* 127: 418424. doi:10.1016/j.compag.2016.07.003.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. 2009. "The WEKA Data Mining Software: An Update." *SIGKDD Explorations* 11: 1. doi:10.1145/1656274.1656278.

Hinton, G. E., S. Osindero, and Y. Teh. 2006. "A Fast Learning Algorithm for Deep Belief Nets." *Neural Computation* 18: 1527–1554. doi:10.1162/neco.2006.18.7.1527.

Huang, Y., S. J. Thomson, W. Clint Hoffmann, Y. Lan, and B. K. Fritz. 2013. "Development and Prospect of Unmanned Aerial Vehicle Technologies for Agricultural Production Management." *International Journal of Agricultural and Biological Engineering* 3: 1–10.

Hunt, E., W. D. Hively, C. S. Daughtry, G. W. McCarty, S. J. Fujikawa, T. Ng, M. T. Chitella, D. S. Linden, and D. W. Yoel. 2008. "Remote Sensing of Crop Leaf Area Index Using Unmanned Airborne Vehicles." In *Proceedings of the Pecora 17 Symposium*, Denver, CO, USA.

Jones, H. G., and R. A. Vaughan. 2010. *Remote Sensing of Vegetation*. Oxford: University Press.
Karpathy, A., and L. Fei-Fei. 2015. "Deep Visual-Semantic Alignments for Generating Image Descriptions." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA.
Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2012. "ImageNet Classification with Deep Convolutional Neural Networks." *Proceedings of NIPS*, Lake Tahoe, UT, USA.
Lecun, Y., Y. Bengio, and G. E. Hinton. 2015. "Deep Learning." *Nature* 521 (7553): 436–444. doi:10.1038/nature14539.
Lu, J., H. Jie, G. Zhao, F. Mei, and C. Zhang. 2017. "An In-Field Automatic Wheat Disease Diagnosis System." *Computers and Electronics in Agriculture* 142: 369–379. doi:10.1016/j.compag.2017.09.012.
Mollahosseini, A., D. Chan, and M. H. Mahoor. 2016. "Going Deeper in Facial Expression Recognition Using Deep Neural Networks." In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, Lake Placid, NY, USA.
Odido, D., and D. Madara. 2003. "Emerging Technologies: Use of Unmanned Aerial Systems in the Realisation of Vision 2030 Goals in the Counties." *International Journal of Applied Science and Technology* 3 (8): 107–127.
Pagola, M., R. Ortiz, I. Irigoyen, H. Bustince, E. Barrenechea, P. Aparicio-Tejo, C. Lamsfus, and B. Lasa. 2009. "New Method to Assess Barley Nitrogen Nutrition Status Based on Image Colour Analysis Comparison with SPAD-502." *Computers and Electronics in Agriculture* 65: 213218. doi:10.1016/j.compag.2008.10.003.
Pellegrin, L., H. J. Escalante, M. Montes-y-Gómez, and F. A. González. 2017. "Local and Global Approaches for Unsupervised Image Annotation." *Multimedia Tools and Applications* 76 (15): 16389–16414. doi:10.1007/s11042-016-3918-9.
Robert, P. 2002. "Precision Agriculture: a Challenge for Crop Nutrition Management." *Plant and Soil* 247 (1): 143–149. doi:10.1023/A:102117151.
Russakovsky, O., J. Deng, S. Hao, J. Krause, S. Satheesh, M. Sean, Z. Huang et al. 2015. "ImageNet Large Scale Visual Recognition Challenge." *International Journal of Computer Vision (IJCV)* 115 (3): 211–252. DOI:10.1007/s11263-015-0816-y.
Ryan, J., S. Masri, S. Ceccarelli, S. Grando, and H. Ibrikci. 2008. "Differential Responses of Barley Landraces and Improved Barley Cultivars to Nitrogen-Phosphorus Fertilizer." *Journal of Plant Nutrition* 31: 381–393. doi:10.1080/01904160801894939.
Saffari, A., and I. Guyon. 2006. "Quickstart Guide for Clop." *Technical report*, Graz University of Technology and Clopinet.
Simonyan, K., and A. Zisserman. 2015. "Very Deep Convolutional Networks for Large Scale Image Recognition." In *Proceedings of ICLR*, 1–14, San Diego, CA, USA.
Statistical Yearbook of the Food and Agricultural Organization for the United Nations, FAO 2014, http://www.fao.org/3/a-i3621e.pdf.
Tang, J., D. Wang, Z. Zhang, H. LiJun, J. Xin, and X. Yang. 2017. "Weed Identification Based on K-Means Feature Learning Combined with Convolutional Neural Network." *Computers and Electronics in Agriculture* 135: 63–70. doi:10.1016/j.compag.2017.01.001.
Thenkabail, P. S., J. G. Lyon, and A. Huete. 2012. *Hyperspectral Remote Sensing of Vegentation*. CRC Press-Taylor and Francis group, Boca Raton, London, New York.
Tian, Y., P. Luo, X. Wang, and X. Tang. 2015. "Deep Learning Strong Parts for Pedestrian Detection." In *Proceedings of IEEE International Conference on Computer Vision*, Santiago, Chile.
Tripicchio, P., M. Satler, G. Dabisias, E. Ruffaldi, and C. A. Avizzano. 2015. "Towards Smart Farming and Sustainable Agriculture with Drones: Intelligent Environments." *Intelligent Environments Conference* 2015: 140–143.
Uno, Y., S. O. Prasher, R. Lacroix, P. K. Goel, Y. Karimi, A. Viau, and R. M. Patel. 2006. "Artificial Neural Networks to Predict Corn Yield from Compact Airbone Spectrographic Imager Data." *Computers and Electronics in Agriculture* 47 (2): 49–161.
Valente, J., D. Sanz, A. Barrientos, J. D. Cerro, A. Ribeiro, and C. Rossi. 2011. "An Air-Ground Wireless Sensor Network for Crop Monitoring." *Sensors* 11: 6088–6108. doi:10.3390/s110606088.

Vedaldi, A., and K. Lenc. 2015. “MatConvNet – Convolutional Neural Networks for MATLAB.” In *Proceedings of the 23rd ACM international conference on Multimedia Pages 689–692,* Brisbane, Australia, ACM New York, USA.

Vinyals, O., A. Toshev, S. Bengio, and D. Erhan. 2015. “Show and Tell: A Neural Image Caption Generator.” In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA.

Xu, X., C. Zhao, X. Song, X. Yang, and G. Yang. 2014. “Estimating Leaf Nitrogen Concentration In Barley By Coupling Hyperspectral Measurements With Optimal Combination Principle.” *Intelligent Automation & Soft Computing* 20 (4): 611–623. doi:10.1080/10798587.2014.934596.

Yosinski, J., J. Clune, Y. Bengio, and H. Lipson. 2014. “How Transferable are Features in Deep Neural Networks?”

Zarco-Tejada, P. J., J. A. J. Berni, L. Suáez, and E. Fereres. 2008. “A New Era in Remote Sensing of Crops with Unmanned Robots.” *Proceedings of SPIE* 7480: 2–4.

Zhang, C., and J. M. Kovacs. 2012. “The Application of Small Unmanned Aerial Systems for Precision Agriculture: A Review.” *Precision Agriculture* 12: 693–712. doi:10.1007/s11119-012-9274-5.

Zhang, N., M. Wang, and N. Wang. 2002. “Precision Agriculture—a Worldwide Overview.” *Computers and Electronics in Agriculture* 36 (2–3): 113–132. doi: 10.1016/S0168-1699(02)00096-0.