

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/334306432>

# Wheat yield forecast using contextual spatial information

Conference Paper · July 2019

DOI: 10.3920/978-90-8686-888-9\_88

CITATIONS

0

READS

143

4 authors, including:



**Mario Fajardo**

The University of Sydney

13 PUBLICATIONS 60 CITATIONS

[SEE PROFILE](#)



**B. M. Whelan**

The University of Sydney

69 PUBLICATIONS 1,757 CITATIONS

[SEE PROFILE](#)



**Patrick Filippi**

The University of Sydney

15 PUBLICATIONS 47 CITATIONS

[SEE PROFILE](#)

## Wheat yield forecast using contextual spatial information

M. Fajardo, B. Whelan, P. Filippi, T. Bishop

*Precision Agriculture Laboratory | Sydney Institute of Agriculture | The University of Sydney, 1 Central Av., Eveleigh, 2016 Australia;*  
*mario.fajardo@sydney.edu.au*

### Abstract

In recent years, the sources of data on farming practices have increased exponentially. On-the-go fuel consumption, grain moisture and yield per hectare are now easily stored by farmers. Fine resolution (5m) wheat yield forecast is presented here using two machine learning approaches: a) bootstrapped regression trees (BRT) where predictions are pixel-wise and b) convolutional neural networks (CNN) where predictions use neighbouring pixels. This study made use of publicly available data (e.g., Sentinel 2 imagery) and farm-owned data (e.g., yield data). Results showed better performance of CNN over BRT, and when higher resolution data was included.

**Keywords:** convolutional neural networks; bootstrapped regression trees; yield forecasting; remote sensing.

### Introduction

As mechanization of agricultural practices increases, different types of information become easily available such as vehicle fuel consumption, speed, harvested grain protein and moisture. Despite this increasing amount of data, its practical use in agricultural decision-making tools remains minimal. There is also an ever-increasing amount of publicly available sources of information created daily e.g. worldwide satellite imagery.

One of the more common uses of remotely sensed data, is the calculation of normalized difference vegetation index (NDVI) using multispectral imagery (Irons et al., 2012) with the subsequent application of a clustering process (K-means for example) for the purpose of establishing management zones to optimize the use of fertilizers. However, the development of more complex analyses is still in the early stages. (e.g., yield forecasting).

Within-paddock and seasonal variability of crop yield is the reason why forecasting yield is still a challenge (Hasan et al., 2019). The main studies that have explored the possibility of yield forecasting are those from different governments or international organizations related to food-security or policy-making and their spatial coverage is state-wide or even country-wide (Baruth, et al., 2008; Bauer, 1975; Doraiswamy et al., 1979; Erickson, 1984).

Within-paddock yield varies from location to location and is largely seasonally influenced (due to sudden rainfall events for example). Traditionally, farmers will estimate their yield goals based on previous experience and personal knowledge about their farms. Recent advances in machine learning algorithms and machine vision have catalysed the improved use of the ever increasing amount of digitized data (Dempewolf et al., 2014; Krizhevsky, et al., 2012; You, et al., 2017).

In this study, forecasting within-paddock yield variability was attempted at a fine resolution (5m) using two of the increasingly popular machine learning approaches. On the one hand a BRT approach is undertaken, where the predictions are made pixel-wise, and on the other hand a CNN approach is used where 2D convolutions are included as

sliding windows, which adds spatial 'patterns' to the model. The goal of the study was to test and report the results of a basic modelling example, using only a few layers of information in order to test the potential for forecasting wheat yield just using publicly available sources of information.

## Material and Methods

### Study area

The study was conducted on a large aggregation of several farms (with 47 fields or paddocks in total) owned by a single corporation and located in the southern agricultural region of Western Australia (120.726 ° / -33.698 ° Lat/Lon, WGS 84). The soils of the area are typically sandy with notable amounts of gravel, and one of the most widespread soil types are sodosols (Isbell, 2002). Only paddocks with wheat yield information were used for this investigation.

The study area is characterised by a Mediterranean climate, with cool, wet winters, and hot, dry summers and average annual rainfall of the aggregations ranges from 420 to 542 mm. This study used data from 2013 to 2017 growing seasons, with a variation in the amount of total rainfall received between each year ranging from 389.0 mm to 689.3 mm for the different fields (Bureau of Meteorology, 2019).

### Data layers

Multiple data layers were available (up to 44 layers including satellite imagery, rainfall and terrain attributes) from different public data-sources. All the layers were compiled in a space-time array of ~80M pixels, covering a total area of ~6,800ha across a time-span of 4 years i.e., 2013-2017. A total of 35,948 pixels with wheat yield data were selected from this array as the initial training dataset. For this initial study, only 8 of those layers were used in order to test the efficacy of a very basic example of a deep network analysis, and to keep within a reasonable computational time frame. The specific layers were chosen to represent well known properties used in yield predictions, namely biomass, soil physical and chemical properties and terrain attributes.

The models used two calculated NDVI layers, either from Landsat 8 (30m resolution) (Irons et al., 2012) or Sentinel 2 (10m resolution) (Drusch et al., 2012) satellites, from the clearest (cloud-free) images of both autumn and winter seasons. In addition, surface clay content (%) and total nitrogen (%) (both 90m resolution) were included as soil attributes modelled by the terrestrial ecosystem research network (TERN) project (Rossel et al., 2015) and a digital elevation model (DEM) (30 m resolution) layer used as a terrain attribute (Farr et al., 2007). Finally, surface electromagnetic induction data (DualEM 21S, DualEM, Ontario, Canada) (5 m resolution) was included in the modelling as a soil attribute. All the layers were first set to a uniform resolution by a nearest neighbour algorithm and later processed using the Google earth engine platform (Gorelick et al., 2017).

### Model specifications

Two modelling types were employed, the first followed a BRT approach and the second a CNN approach, and both modelling approaches were validated against an independent dataset, following a “leave one field out” validation. Samples from all the fields but one, were used in creating a model that then was used for predicting the one field excluded. Using this procedure, a distribution of predictions was obtained rather than a single performance parameter.

### Bootstrapped Regression Trees (BRT)

A “cubist” algorithm was selected as the regression trees model (Kuhn et al., 2014a; Quinlan, 1986). This model builds successive discrimination points that segregate the data into different “branches”, and at the end of each branch a different multivariate linear regression is created to model the relationship between the variables in each branch. These regression models are combined to form an algorithm that predicts the final outcome.

To make the process more robust, a bootstrapping technique was applied to the BRT process. Using this technique, 50 different regression trees were created using a 90% subset of the original dataset, but for each tree a different subset was randomly chosen. Subsets are used to create different samples of the data population so that a distribution of predicted outcomes can be achieved. At the end, the mean of the predictions from the 50 trees was used as the final result.

### Convolutional neural networks (CNN)

CNNs take most of the concepts of an artificial neural network (ANN) (LeCun et al., 1989) which are an attempt to use concepts from neurobiology, where many *neurons* (or nodes) are linked by *synapses* (or weights) between them. A typical neural network design will involve an input layer with  $n$  nodes, and one or more *hidden* layers that will carry and modify information from the input layer to a final output layer, which results in a function of the multiple weights of each of the hidden layers.

If, in the input process, a *tensor* (n-dimensional array) with at least 2 dimensions is used instead of a one dimensional vector (as is the case of regression trees), then information from more than one point source (e.g. groups of pixels rather than single pixels) can be used. The approach is graphically depicted in Figure 1.

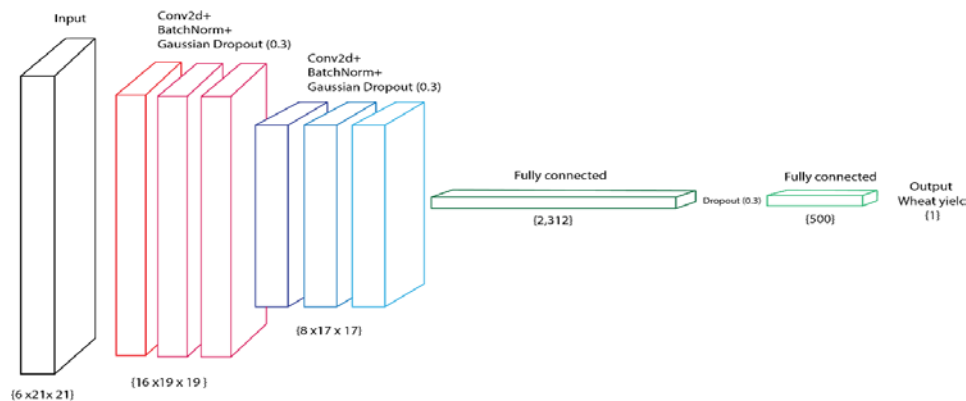


Figure 1. CNN architecture used for wheat yield prediction, numbers in brackets represent dimensions of each array.

As shown in Figure 1, the input layer corresponds to the extracted values from each of the raster layers (Sentinel2, clay, etc.) from a  $21 \times 21$  pixel window, (Figure 2). The next layer receives this array and applies 16 different filters (i.e., a  $3 \times 3$  pixel kernel moving throughout the array), and this process is the one that gives the “convolution” name to this kind of network. As this convolution uses a moving window, then the resulting output is a smaller  $19 \times 19$  stack.

The next step involves a batch normalization followed by a dropout layer (30 % of the information is left aside), which has been shown to improve the generalization of the model (Srivastava et al., 2014). The next steps involve a second convolution + normalization + dropout layers which increases the “depth” of the network (i.e., more

layers). Deep networks have been shown to produce better results when using convolutional processes (Krizhevsky et al., 2012).



Figure 2. Example of eight sampling locations (top). Detail of the same locations with central pixel highlighted in white (middle) and detail of two sample locations with a 21 x 21 pixels contextual sample (bottom).

Finally, after a third dropout, a “flatten” layer process is used to deliver a one-dimensional output of a single value of wheat yield. The architecture of a deep network represents a whole area of research by itself, and the one used in this study follows previously used approaches that achieved some success when predicting crop yield (You et al., 2017). More information about neural network layers and filters can be found in Chollet and Allaire (2018).

#### Sampling design and modelling approaches

In order to optimize the samples and the observed yield variability, a conditioned latin hyper-cube design (Minasny & McBratney, 2006; Roudier, 2011) was employed to select the sampling locations. Within each field, a fixed percentage (2%) of the total number of pixels (5m resolution) was sampled. Hence, using this design, each field will have a balanced contribution to the model in terms of within-paddock variability (Table 1).

Table 1. Different model designs and respective number of samples

Model Type	Covariates	Years	Number of fields	Number of samples
CNN	5	2013-2017	47	35,948
CNN-Sent	5	2016-2017	10	16,132
CNN-DualEM	6	2016-2017	7	7,961
RegTrees	5	2013-2017	47	35,948
RegTrees-Sent	5	2016-2017	10	16,132
RegTrees-DualEM	6	2016-2017	7	7,961

Three different sets of information for both BRT and CNNs were used, resulting in six different modelling approaches. The main difference between approaches was the use of either Landsat 8 or Sentinel as the “NDVI” input and the use of DualEM. The first approach used all the observed fields with a total of 35,948 samples with Landsat 8 imagery as the “NDVI” input, clay, total nitrogen and DEM (5 covariates) and was simply named “RegTrees” or “CNN” depending on the type of modelling framework (Table 1). The second approach (with a total of 16,132 samples) employed Sentinel 2 imagery instead of Landsat 8 (5 covariates as well) and was named either “RegTrees-Sent” or “CNN-Sent” and finally, the third (with a total of 7,961 samples) used Sentinel 2 imagery plus DualEM information, this final approach was named as “RegTrees-DualEM” or “CNN-DualEM” (Table 1). As mentioned, CNN modelling requires contextual information, hence for each of the selected locations, a window of 21 x 21 pixels was selected as input for the CNN modelling and only the central pixel of this matrix was used as input for the regression trees. In this way, both training datasets have the same number of samples (Figure 2).

All the analyses were undertaken using Python (for the raster extraction using Google earth engine) and R for the sampling design and modelling. BRT were implemented using R package “Cubist” (Kuhn et al., 2014b) and CNN modelling was implemented using the R package “Keras” with a “TensorFlow” back-end (Chollet & Allaire, 2018)

### Model performance

In order to compare the different modelling procedures, four metrics were employed. These were the well-known coefficient of determination ( $R^2$ ), root mean squared error (RMSE), Lin’s concordance correlation coefficient and a bias statistic (mean predicted values – mean observed values). All yield information was scaled between 0 to 1 for privacy reasons.

### Results

It was observed that as the resolution (Landsat 8 < Sentinel 2) and complexity of the modelling (RegTrees < CNN) increased, the performance increased accordingly (Figure 3).

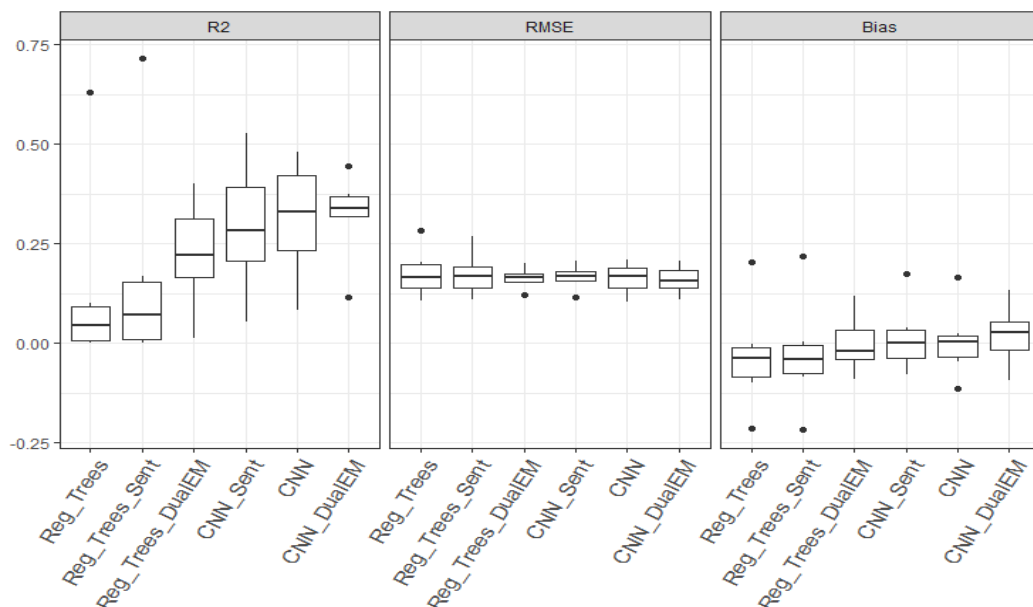




Figure 3. Performance of the different modelling approaches.

These results seem consistent with the amount of information included in each of the modelling approaches i.e., the BRT approach uses just the central pixel of a CNN observation; this is 1 pixel vs 441 pixels (21 x 21 window). In terms of RMSE, the results agreed with previous attempts at modelling (at bigger spatial scales). Even though, as far as the authors know, there is no comparable study modelling within-paddock forecasting, You et al. (2017) reported a substantial improvement of accuracy when using CNNs compared with regression trees at county scale in the U.S.A.

The results also showed the benefit of farm-owned data such as electromagnetic induction studies (DualEM), to final predictions. There was an obvious improvement for both BRT and CNNs compared with models without DualEM data in terms of higher  $R^2$ , less bias and less dispersion of the predictions (Figure 3). In terms of RMSE for the different modelling approaches, the differences were not as obvious, with all the models presenting RMSEs values around 10 to 15% of the total yield production (i.e., for a field with a yield of 4 t/ha an error of about 400~600 kg/ha). In terms of total yield, this result seems useful, but still a rough estimate.

In order to estimate the quality of the predictions for decision-making purposes, a spatial assessment is needed. Figure 4 shows an example of the spatial predictions derived from the CNN-Sent model for a randomly selected field (from all the fields of Figure 3). There are spatial patterns evident that are captured by the model which details the potential yield and its distribution. It is important to understand that the current modelling approach only uses information from the previous harvest (autumn-winter imagery) excluding harvest information from that particular field.

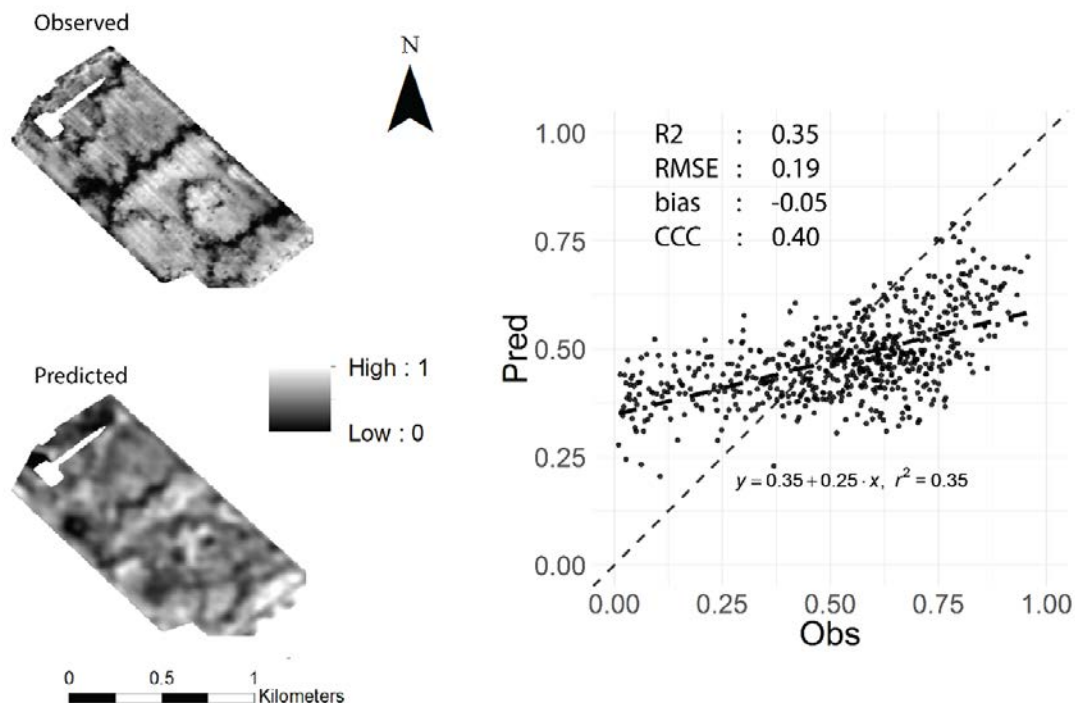


Figure 4 Spatial prediction using the CNN-Sent model for a randomly selected field in the data set (yield data standardised to 0-1 for privacy reasons).

Finally, considering the limited information included in the current models, within-paddock forecasting using publicly available data shows potential, especially if a more complex modelling approach (e.g., deep networks) such as CNNs is used. Future work will involve the use of data with greater yield variability (inclusion of more farms from different agro-climatic regions) and the addition of more layers in the modelling process, such as gamma-radiometric surveys and climate (temperature and rainfall) data.

## Conclusions

The study showed encouraging results for within-paddock wheat yield forecasting using only machine learning techniques. The model performance increased as more detailed information was used (higher spatial resolution) and when contextual information was considered in the modelling exercise. However, there is still room for improvement in the use of CNNs coupled with high-resolution yield information.

## References

- Baruth, B., Royer, A., Klisch, A., & Genovese, G. (2008). The use of remote sensing within the MARS crop yield monitoring system of the European Commission. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, **37**, 935-940.
- Bauer, M. E. (1975). The role of remote sensing in determining the distribution and yield of crops. *Advances in Agronomy*, **27**, 271-304.
- Bureau of Meteorology. (2019). Australian Government.  
<http://www.bom.gov.au/climate/data/> (last accessed February 2019)
- Chollet, F., & Allaire, J. J. (2018). *Deep Learning with R*. New York, USA: Manning Publications Company, 360p.
- Dempewolf, J., Adusei, B., Becker-Reshef, I., Hansen, M., Potapov, P., Khan, A., et al. (2014). Wheat yield forecasting for Punjab Province from vegetation index time series and historic crop statistics. *Remote Sensing*, **6**, 9653-9675.
- Doraiswamy, P.C., Hodges, T. and Phinney, D.E. (1979). Crop yield literature for AgRISTARS crops corn, soybeans, wheat, barley, sorghum, rice, cotton and sunflowers. AgRISTARS Technical Report SR-L9-00405, Houston, Texas: Lockheed Electronics Company, Inc., 105 p.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, **120**, 25-36.
- Erickson, J. (1984). The LACIE experiment in satellite aided monitoring of global crop production. In: G. M. Woodwell (Ed.) *The role of terrestrial vegetation in the global carbon cycle: measurement by remote sensing*, Chichester, UK and New York, USA: John Wiley & Sons Ltd. pp 191-217.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The Shuttle radar topography mission. *Reviews of Geophysics*, **45**, RG2004, pp33, doi:10.1029/2005rg000183
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, **202**, 18-27.
- Hasan, M.M., Croke, B.F.W. & Karim, F. (2019) Spatial and seasonal variations and inter-relationship in fitted model parameters for rainfall totals across Australia at various timescales. *Climate*, **7**, 4-15. doi.org/10.3390/cli7010004
- Irons, J. R., Dwyer, J. L., & Barsi, J. A. (2012). The next Landsat satellite: The Landsat data continuity mission. *Remote Sensing of Environment*, **122**, 11-21.
- Isbell, R. F. (2002). *The Australian soil classification*. CSIRO Publishing, 141p.



- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In: NIPS'12 Proceedings of the 25<sup>th</sup> International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, USA: Curran Associates Inc., pp 1097-1105.
- Kuhn, M., Weston, S., & Coulter, N. (2014a). C50: C5.0 Decision trees and rule-based models. R package version 0.1.0-19. <https://CRAN.R-project.org/package=C50>. (last accessed February 2019).
- Kuhn, M., Weston, S., Keefer, C., & Coulter, N. (2014b). Cubist: Rule- and instance-based regression modeling. <https://CRAN.R-project.org/package=Cubist>. (last accessed February 2019).
- LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E. et al. (1989). Back-propagation applied to handwritten zip code recognition. *Neural Computation*, **1**, 541-551.
- Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, **32**, 1378-1388.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, **1**, 81-106.
- Rossel, R. V., Chen, C., Grundy, M., Searle, R., Clifford, D., & Campbell, P. (2015). The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. *Soil Research*, **53**, 845-864.
- Roudier, P. (2011). clhs: a R package for conditioned Latin hypercube sampling. Retrieved from <http://code.scnzgrid.org/index.php/p/clhs/> (last accessed February 2019).
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, **15**, 1929-1958.
- You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep gaussian process for crop yield prediction based on remote sensing data. In: AAAI-17 Proceedings of the 31<sup>st</sup> AAAI Conference on Artificial Intelligence. San Francisco, USA: Association for the Advancement of Artificial Intelligence, pp 4559-4565.