Monitoring Grain Production across the Australian Continent

Roger A. Lawes[1], Joanne Chai[2], Yang Chen[1], Randall J. Donohue[3], Gonzalo Mata[1], Zvi Hochman[4], Franz Waldner[4], Chris Sharman[5], Roger Butler[6]

[1] CSIRO Agriculture and Food, Underwood Ave, Floreat WA 6014, Australia; roger.lawes@csiro.au
[2] CSIRO Data61, Underwood Ave, Floreat WA 6014, Australia
[3] CSIRO Land and Water, GPO Box 1700, Canberra, ACT 2061, Australia;
[4] CSIRO Agriculture and Food, 306 Carmody Rd, St Lucia QLD 4067, Australia;
[5] CSIRO Data61, College Rd, Sandy Bay TAS 7005, Australia.
[6] CSIRO Data61, 1 Technology Court, Pullenvale, Qld 4069, Australia.

Abstract
To successfully assess crop yield across Australia there is a need to monitor what has been sown and its progress as the season evolves. Crop type and species need to be identified at the paddock scale to calculate areas. Crop models then need to be applied to each individual paddock to generate a yield estimate. Finally, the information needs to be packaged at a resolution of interest. To generate state and national scale crop monitoring capability, a co-ordinated, multifaceted data gathering, data training, image capture, data acquisition and crop modelling operations were developed. Crop yield forecasting required new modelling techniques, as existing approaches were overwhelmed by the volume of data. We describe the detailed process of how we monitor and forecast crop production across the Australian landscape. Near real time crop monitoring products are now available across the Australian continent. This paper describes the overarching workflow of crop monitoring, forecasting and data dissemination to assist agribusiness to respond to the prevailing climate.

Key Words
Crop yield estimation, yield monitoring, crop type identification, soil water app.

Introduction
In agricultural systems, it is important to monitor crops and understand yields, as many agriculture businesses require productivity information to make critical business decisions. To support this need crop monitoring services have advanced considerably in recent years. Complex analytics, remote sensing, modelling and data dissemination are now being used by organisations such as Indigo Ag https://www.indigoag.com/ and Gro Intelligence https://gro-intelligence.com/ to generate information products about agricultural production in the US. Across the European Union, Americas, Africa and Australasia, global co-operatives have formed to produce information about the global food supply through the GEOGLAM Crop Monitor https://cropmonitor.org/ and the Agricultural Marking Information System (AMIS; www.amis-outlook.org). Collectively, these services provide insights into crop production at a global scale, and in some instances individual organisations are able to offer farmers information about individual crops.

In Australia, farmers can source satellite imagery for particular fields either through the United States Geological Society (https://www.usgs.gov) or the European Space Agency (https://www.esa.int). Information can also be sourced through third party organisations that provide land viewer services such as https://eos.com/landviewer/ where indices such as NDVI and SAVI are generated from both Sentinal-2 and Landsat 8 imagery. However, these systems do not provide reasonable measures of yield, or yield performance, and additional information is often required to monitor crop yields.

Historically, the Yield Prophet system has been used to fill this need (Hochman et al., 2009). More recently, researchers have demonstrated that farmers' yield data, crop models and satellite data can be analysed with crop models to generate site specific fertiliser recommendations. Others have constructed new crop models, built using satellite imagery to monitor grain yield and generate a prediction that farmers may find useful (Donohue et al 2018). However, in each case additional information needs to be sourced from somewhere to generate a reliable prediction that is accessible by agri-business.

To build a national crop monitoring system, training data about crop species are required in every season. Training data comprising of fields that are known to be sown to each of the relevant crop species are required across the entire continent. Information about fields not in cereal crop, such as ryegrass pastures, are also required. Minor crops pose a particular challenge, as training data is limited (Waldner et al 2019). Wheat, barley and oats dominate and as a crop they are largely interchangeable and suited to similar soils and climates. Differences and planting choices vary with pricing signals and disease pressure. These training data must be linked with satellite imagery, which is then used to classify and identify specific crop species (Figure 1). Information about all crops in a region is required and in the context of crop classification, a region is defined as a row and path of a particular satellite scene.

Here we provide a brief synopsis of the process of data collection, farmer engagement, model development and data dissemination required to generate a national crop monitoring service. The system was deployed across the 20 million ha of Western Australia in 2017 and 2018 and will be deployed at a national scale in 2019.

Methods

Data

A scalable national crop monitoring capability requires information on all aspects of crop production, for the entire continent. This capability requires two components, crop identification at the field scale and yield monitoring of that particular crop.

Crop Identification – training data

A training dataset was acquired for Western Australia in 2017 and 2018. For 2017 roadside surveys were conducted, where trained crop scouts identified crop species and recorded that crop species with a GPS. Photographs were used to verify the crop classification. Overall several thousand fields were captured in 2017 along the major roads of Western Australia. The rows and paths of Landsat 8 scenes that cover the Western Australian wheatbelt highlight where classification training data was plentiful and where it was arguably insufficient.

In 2018, training data was sourced through three mechanisms. Firstly, it was acquired with some targeted road surveys that occurred late in the growing season. Secondly, a relationship was formed with a third party organisation, who supplied information about crop type for the explicit purpose of classifying crop types. Finally, an app was constructed to provide farmers with information about their crop yields. As part of this process, farmers were required to provide information about the crop type.

Satellite information is required to discriminate between crop types. Satellites available include Landsat and Sentinal-2. In winter dominant cropping systems, cloud frequently blankets crops during the growing season. Therefore in some instances, there is a dearth of useable optical satellite imagery. Once imagery has been acquired, we used random forests (Brieman et al 2001) to classify the crop types. We present the results from 2018 to illustrate how crop classification is conducted. In general the accuracy improves through the season as both the amount of available imagery and quantity of useful training data increases (data not shown).

Crop Modelling

A satellite driven crop model, known as C-Crop (Donohue et al 2018) was deployed across Western Australia to monitor the yields of wheat, barley and canola. The model uses a time series of MODIS imagery to calculate FPAR and climate data sourced from the Bureau of Meteorology to predict above ground biomass. The model is parameterised from yield data collected from combine harvesters to generate a crop yield estimate. This model does not require soil type, rainfall or management information, and is therefore suitable for yield estimation with parsimonious data.

To deliver information to growers via an app based platform (https://research.csiro.au/graincast/), output from the APSIM model (Holzworth et al 2014) is generated, where all parameters are estimated using the national soil grid (McKenzie et al 2012) and the climate grid, courtesy of the Bureau of Meteorology (Raupach et al 2009). Starting conditions are estimated by way of simulation, where the run is initiated 30 years prior. The objective is to provide information to growers that approximates yield potential and soil-water status, where the only input required from the user is to define crop type.

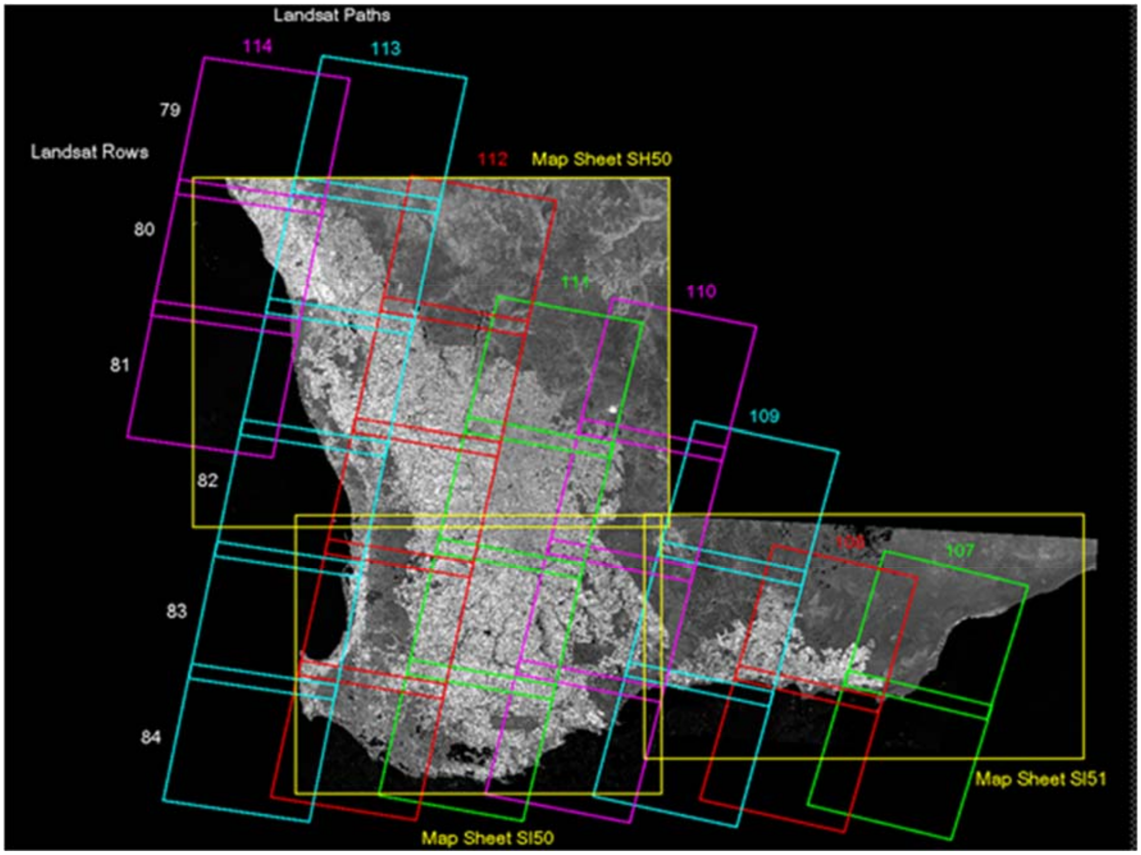

Figure 1. The Landsat Rows and Paths across the Western Australian wheatbelt, where the paths range from 114 on the western edge to 107 on the eastern edge. The rows range from 84 at the southern extent to 80 for the northern extremity of the Western Australian wheatbelt. Training data is required in each unique row and path combination.

Data Dissemination – Agribusiness

Extensive consultation occurred with an agribusiness organisation to provide data in a timely, manageable manner. Data about crop yield and crop type were provided in mid-September, mid-October and finally in mid-November as a raster GIS (shapefile) format. Key informatics of relevance to the agribusiness were tabulated, and error estimates were provided. Errors were calculated from the Random Forest analysis with regard to crop classification. Errors around yield were provided and based on the model error reported in Donohue et al 2018. Whilst data were delivered electronically, each time data was delivered, an engagement process took place with the client.

Data Dissemination – Producers

To disseminate information to growers about crop yield, an app was constructed. Growers were interviewed to determine what information they prefer and the interview process was conducted by a user experience (UX) team. The UX team worked with classically trained agriculture scientists with extensive agronomic knowledge. The UX team developed an app with a simple user interface, that focused on simplicity, rather than focusing on scientific capability.

Results

In 2018, the classification algorithm estimated that 9.7 million ha in WA were cropped, 6.5 million ha were in pasture and the remaining area was either bare, remnant vegetation or salt land. The cropped area equated to 46% of the total landscape in the Western Australian wheatbelt. The cropped area was portioned into wheat (53% of cropped area and 5.1 million ha), barley (15% cropped area and 1.4 million ha), canola (12% and 1.2 million ha) and legumes (15% and 1.5 million ha). When the C-Crop model was deployed to each paddock classified as either wheat, barley or canola, the final tonnage estimates were 12.2 million tonnes of wheat, 3.7 million tonnes of barley and 1.7 million tonnes of canola.
Classification errors, where a crop or pasture is misclassified as another species, and errors around yield

prediction, influence the value of a digital product. Canola was classified with 95% accuracy. Barley was classified with 86% accuracy, with a bias of 12% confusion with wheat. Wheat was classified with 79% accuracy, with a 16% confusion with barley. Legumes were classified with 64% accuracy with a confusion of 18% for canola and 11% for wheat. Similarly, there were errors associated with the yield estimates. For wheat, across 36 towns, 1 standard deviation equated to 0.64 t/ha. The implication of this standard error is that across the state, the tonnage estimate equated to 12.2 million tonnes with a standard deviation of 3.9 million tonnes. Future model developments (see Chen et al 2019), are expected to improve model performance.

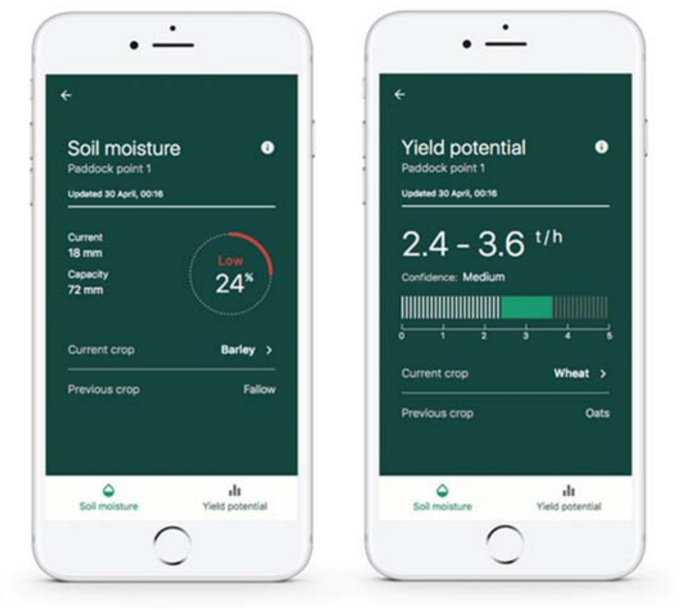

Figure 2. Output from the app delivered to growers in 2018, showing a soil moisture estimate and a yield potential estimate. All outputs were generated with the APSIM model.

Conclusion

National scale crop monitoring is complex. Here we have developed a workflow that generates training data, utilises satellite imagery to classify crop species, monitors crop yields with the next generation of crop models and deploys output to multiple end users by way of raster images or through an app based platform. Each of the fore mentioned components is in its infancy, but as the technology develops, we expect every aspect of the workflow to improve and result in reductions in the errors associated with the prediction and monitoring of grain production.

References

Breiman L. (2001) Random forests. Machine learning, 45(1), 5-32.

Chen Y, Donohue RJ, McVicar TR, Walnder F, Mata, G. Ota N, Houshmandfar A, Dayal K. and Lawes R. (2019) Crop yield estimation based on photosynthesis and meteorological-driven stress indices. 19th Australian Agronomy Conference. Wagga Wagga.

Donohue RJ, Lawes RA, Mata G, Gobbett D. and Ouzman, J. (2018) Towards a national, remote-sensing-based model for predicting field-scale crop yield. Field Crops Research, 227: 79-90.

Holzworth, DP et al., (2014) Apsim–evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software, 62: 327-350.

McKenzie NJ, Jacquier DW, Maschmedt DJ, Griffin EA. and Brough DM. (2012) The Australian Soil Resource Information System (ASRIS) Technical Specifications. Revised Version 1.6, June 2012. The Australian Collaborative Land Evaluation Program.

Waldner F, Chen, Y, Lawes, R and Hochman, Z. (2019) Mapping rare and infrequent crops from space. 19th Australian Agronomy Conference. Wagga Wagga.

Raupach MR, Briggs PR, Haverd V, King EA, Paget M. and Trudinger CM. (2009) Australian Water Availability Project (AWAP): CSIRO Marine and Atmospheric Research Component: Final Report for Phase 3, CAWCR Technical Report No. 013, July 2009