Space-based monitoring of Australian paddock-scale crop yields

Donohue, RJ[1,2], Lawes, RA[3]

[1] CSIRO Land and Water, Building 801 Clunies Ross St, Black Mountain ACT 2601, Randall.Donohue@csiro.au
[2] ARC Centre of Excellence for Climate Extremes, Australia
[3] CSIRO Agriculture and Food, Underwood Ave, Floreat, WA 6014

Abstract
Existing crop yield models typically estimate paddock-scale yields across small areas, or regional-scale yields nationally. This dichotomy restricts our ability to scale between paddock, region and country yields. We developed a nation-wide, paddock-scale crop yield model based on satellite imagery, called CCrop. For its simplicity, C-Crop is effective at estimating canola and wheat yields, with about 70% accuracy. Ongoing model refinements will improve accuracy, encompass more crop types and perform mid-season forecasting. Coupled with crop type maps, C-Crop now provides the ability to conduct nation-wide, fine-scale monitoring of Australian grain production.

Key Words
Yield, satellite, C-Crop, paddock, national

Introduction
An ideal crop yield model predicts yields at paddock scales across large areas, even up to a national extent. However, a general trade-off exists between local scale accuracy and geographic extent, meaning that such an ideal model is difficult to achieve (Lobell 2013; Strand 1981). Examples of locally accurate but geographically restricted models are APSIM (Holzworth et al. 2014) and CERES-Wheat (Ritchie and Otter 1985) whereas an example of the opposite is that of Potgieter et al., (2005).

In Australia, the development of a paddock-scale, national yield model has been inhibited by the absence of suitable data that are both fine-scale and national extent for model parameterisation and calibration (Gobbett et al. 2016). Recently, a major step in overcoming this obstacle has been the collection and collation, from across Australia, of high-resolution paddock yield data acquired from grain harvesters fitted with yield monitoring equipment.

With the aim of producing a paddock-scale, national yield model for Australia, we recently developed C-Crop, a highly simplified remote-sensing-driven grain yield model (Donohue et al. 2018). Here we provide an overview of C-Crop and present some preliminary results assessing the use of C-Crop in a forecasting mode.

Methods
The C-Crop model
C-Crop uses elevation, air temperature and satellite-derived greenness data as inputs (along with an indication of crop type). From these 4 inputs, the gross and net photosynthetic fluxes are estimated throughout the growing season, along with the total above-ground carbon mass of the crop. There is one crop-specific parameter (a crop-specific maximum photosynthetic rate) and two fitted parameters (the mean tissue respiration rate at 10°C and the mean tissue longevity). The latter have been calibrated such that the mass estimate at harvest is effectively the yield prediction. Sowing and harvest dates are set at mid-April and mid-October. The model runs at a 250 m grid resolution, and at a 16-day time-step. See Donohue et al. (2018) for more details.

The model was calibrated at the paddock scale, using cross-validation, drawing on 31 harvester-derived paddock yield maps for canola and 160 for wheat. Calibration minimised the difference between the predicted (mid-October) and observed paddock yields.

Mid-season forecasting
We explored how well C-Crop performs in a simplistic forecasting mode. To do this we calibrated the model every 16 days starting in June, minimising the difference between each date's modelled carbon mass

and the end of season, harvester-derived yields. This produced a separate parameter set for each date of forecast.

Results

Canola and wheat yields were predicted with errors of 0.54 and 0.73 t ha^{-1}, respectively (which equate to relative errors of 33 and 32%; Figure 1). While these are reasonably large errors, they are notable in that they are paddock-scale errors, derived using national-extent input datasets and a very simple crop model.

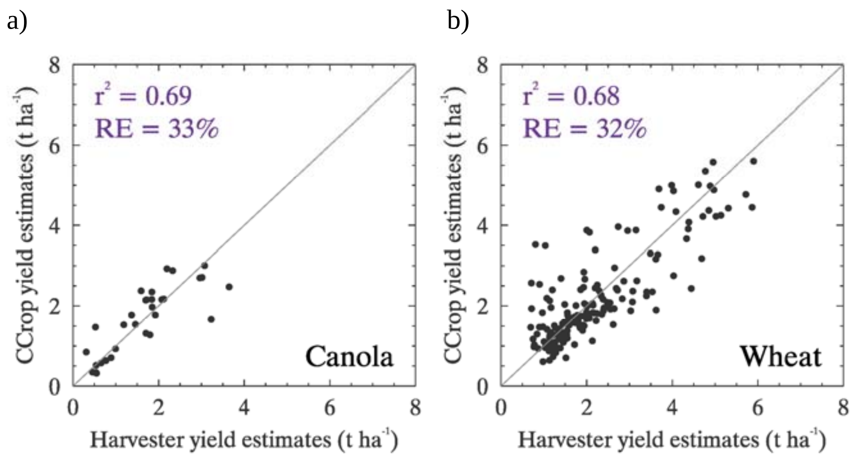

Figure 1. Assessment of C-Crop yield estimates against harvester-derived paddock yield values for canola (a) and wheat (b). RE is the relative error, calculated as the RMSE divided by the mean yield. The goodness of fit statistics are from the validation runs of the original cross-validation analysis. Reproduced from Donohue et al. (2018).

The performance of C-Crop in forecasting mode is shown in Figure 2. The correlation coefficient remained fairly constant through all prediction months at 0.6 – 0.7. In contrast, there is a marked temporal progression in the errors. Predictions made prior to August had very high errors (especially for canola), and indicate that early-mid season predictions are highly unreliable. However, from August onwards errors dropped to around 30-50%, culminating in the lowest errors (~33%) at the notional time of harvest.

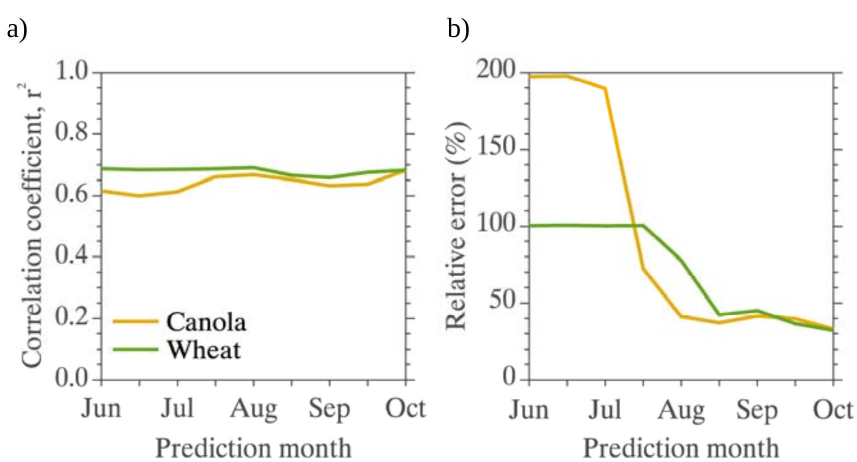

Figure 2. Assessment of C-Crop yield predictions made at sequential times throughout the growing season for canola and wheat. Plot a) shows the r^2 and plot b) shows the relative error, calculated as the RMSE divided by the mean yield.

One way to use these forecasts that minimises the effects of the high errors (assuming the errors are relatively constant across years) is seasonal profiling (or seasonal analogues; Figure 3). This compares current season forecasts to equivalent forecasts from previous years, allowing for an assessment of whether this season is tracking as a good, average or poor season. With C-Crop, such seasonal profiling can be done at a state level (Figure 3a) or for an individual paddock (3b) with equal ease (as long as underlying crop type information is available).

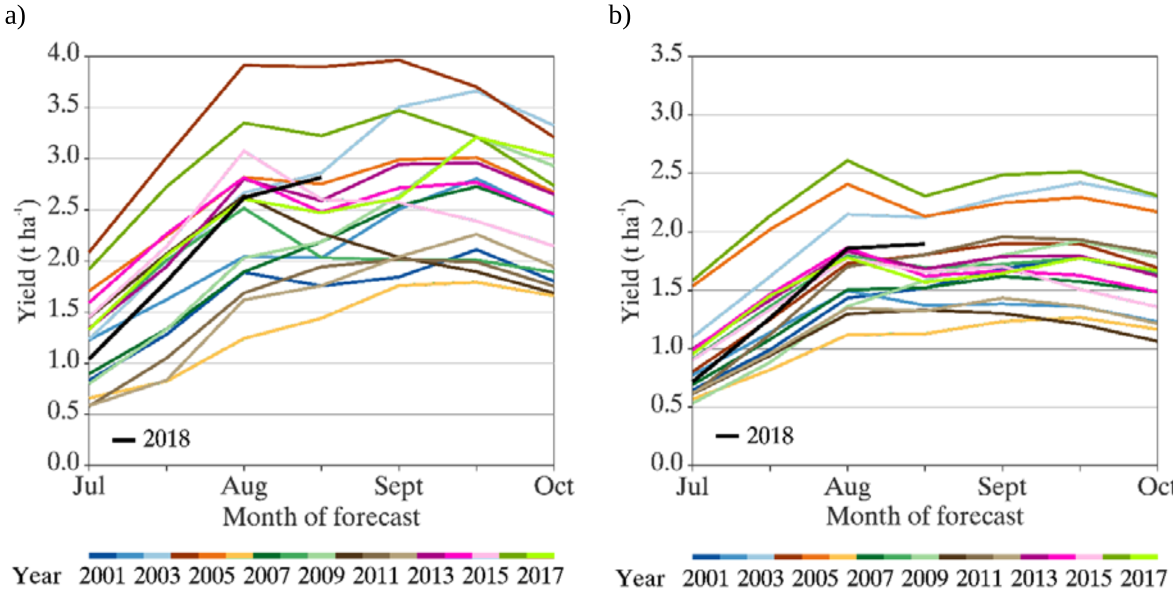

Figure 3. Progression of 2018 yield forecasts (as at mid-August) compared with previous year's seasonal forecast profiles. Plot a) shows the profiles for all of WA's croplands (which, for convenience of example, assumes all crops were wheat). Plot b) shows the profiles for a single wheat paddock in WA.

Conclusion

While C-Crop yield estimates are perhaps surprisingly accurate given the simplicity of the model, the model needs further improvement to increase its accuracy to be of maximum value to growers and the grain industry. The use of C-Crop depends on having underlying crop type information, and contemporary national crop type maps are required before national yield assessments can be performed. These two issues are the focus of ongoing research.

Used in a very simplistic forecasting mode, C-Crop provides moderate-accuracy yield predictions that are most usefully employed in a historically relative mode. We are now seeking to incorporate meteorological forecasts into C-Crop in order to increase the yield forecast accuracies.

C-Crop yield estimates can currently be made with an approximate 30-day latency. We are working to generate a rapid version of C-Crop with a latency of 7 days or less.

References

Donohue, R.J., Lawes, R.A., Mata, G., Gobbett, D.L., & Ouzman, J. (2018). Towards a national, remote-sensing-based model for predicting field-scale crop yield. *Field Crops Research*, 227, 79-90. DOI: 10.1016/j.fcr.2018.08.005.

Gobbett, D., Hochman, Z., Horan, H., Garcia, J.N., Grassini, P., & Cassman, K. (2016). Yield gap analysis of rainfed wheat demonstrates local to global relevance. *The Journal of Agricultural Science*, 155, 282-299. DOI: 10.1017/S0021859616000381.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., & Keating, B.A. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62, 327-350. DOI: https://doi.org/10.1016/j.envsoft.2014.07.009.

Lobell, D.B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, 143, 56-64. DOI: 10.1016/j.fcr.2012.08.008.

Potgieter, A.B., Hammer, G.L., Doherty, A., & de Voil, P. (2005). A simple regional-scale model for forecasting sorghum yield across North-Eastern Australia. *Agricultural and Forest Meteorology*, 132, 143-153. DOI: 10.1016/j.agrformet.2005.07.009.

Ritchie, J.T., & Otter, S. (1985). Description and performance of CERES-Wheat: a user-oriented wheat yield model. ARS wheat yield project (pp. 159-175): Agricultural Research Service, USDA

Strand, B.W. (1981). Spatial Scale of Crop-yield Models: A Review of the Relationship Between Scale of Models and Accuracy. Washington, D.C.: U.S. Department of Agriculture, Economics and Statistics Service, Statistical Research Division.