Using remote sensing and big data analytics to assess rotation effects on wheat yield across the entire WA wheatbelt.

Lawes R[*], Mata G, Herrmann C, Richetti J and Fletcher A

CSIRO Centre for Environment and Life Science, 147 Underwood Avenue, Wembley, WA 6014, Email: roger.lawes@csiro.au

Abstract
Crop rotations are an important component in agricultural systems. However, rotation experiments require considerable resources to implement. Here we use a suite of big data products to quickly evaluate crop rotations across the entire WA wheatbelt. In general farmers achieve yield benefits in the order of 10% using a break crop, but the magnitude of the break crop effect varies from season to season. These analyses can add value to other forms of data. Seasonal differences in the wheat yield response to a break crop were detected across the WA wheatbelt. Crop rotation also varied with location.

Keywords
Big data, crop rotation, crop models.

Introduction
In recent years, the ability to capture yield, environment and management information about a specific field has increased. Satellite information can be processed to monitor crop yield (e.g. Donohue et al. 2018), and crop species (e.g. Waldner et al. 2019). Additional information such as the climate, can also be extracted and these data sources can be combined to enable an interpretation of agronomic processes, such as the benefit or otherwise of crop rotation. Crop rotation effects on productivity in Western Australia are well understood at the experimental scale. Seymour et al. (2012) analysed historical experimental data and concluded that the yield benefit to cereal crops grown in rotation with grain legumes was in the order of 700 kg/ha compared to cereal crops grown in sequence. Notably lower benefits were identified for cereal crops following pastures (~ 400 kg/ha) and oilseeds (~ 400 kg/ha).

Yet, despite the findings that diverse crop rotations benefit wheat yields, cereal dominant rotations are still common. One possible reason for the reluctance to grow broad-leaf break crops is that the widely published benefits of break crops may not materialise for farmers (Seymour et al. 2012). However, aside from the recent analysis of wheat yield gaps, that demonstrated rotation can contribute to those yield gaps (Lawes et al., 2021), few exhaustive studies of wheat crop yield, given the previous crop, have been published. Here we utilise the recent advances in satellite driven crop monitoring technologies to understand the geographic spread of break crops across the entire WA wheatbelt. Satellite estimates of wheat yields are then combined with these data, and with statistical analysis, the industry wide benefits of crop rotation are estimated. The approach illustrates how big data analytics can help agronomic advancement and move beyond traditional experimentation.

Methods
Crop identification
Crop species were identified from 2017 to 2020 using a combination of training data, collected annually, and satellite imagery from the Landsat-8 and Sentinel-1 platforms. Machine learning (random forests) were applied to available optical scenes, collected across the WA landscape during the growing season. The Sentinel-1 SAR platform provided insights about crop species when cloud cover obscured the optical signal. Further methodological details about the machine learning, data processing, and training data collection are provided in Waldner et al. (2019). Over the course of the study, Crop ID was estimated with an accuracy of between 70 and 80%, across all crops (data not shown).

Crop yield estimation
Wheat crop yields were estimated on every field of the WA wheat-belt from 2017 – 2020 using the C-Crop model (Donohue et al. 2018). Briefly the C-Crop model uses the 250m MODIS platform, and climate information to predict yields, where the crop model allocates carbon, calculated via F-Par, to roots, stems and grains. The model predicts wheat yield with an r^2 of 0.71 and RMSE of ~ 700 kg/ha, when assessed against the mean crop yield of an individual field. Field estimates were derived from farmers yield maps. The

information from Crop ID is subsequently used to identify which paddocks are wheat, and through these two tools, the expected yield of every paddock is captured.

Creating the analytical data-frame

Across Western Australia the ePaddocks product, that maps the boundaries of every field in Australia using a single Sentinel-2 image, was used to combine information about crop yield and crop ID into an identifiable object (Waldner et al. 2020). This was repeated for each year of the survey (2017-2020), and using a spatial join, we were then able to identify the previous crop grown for every wheat crop across the entire state. These data were paired with climate information from the SILO grid (Jeffrey et al., 2001), where growing season rainfall (April to October) was also calculated for every field. For each wheat crop, the year, previous crop, pasture or fallow management option, and the growing season rainfall, were available.

Statistical analysis

A basic random effects regression model, with year, previous crop (canola, legume crop, pasture or fallow), growing season rainfall, and all interactions, were fit to the estimated wheat yield calculated from the C-Crop model. Across the wheat belt, 133,800 wheat paddocks were identified and included in the analysis. The predicted values for the rotation effects, on average, and for each year are explored in detail. All other effects were averaged to estimate the main effects of interest.

Regional mapping of crop frequency

The analysis focussed on understanding the rotational benefits at the individual field level. However, the data platform that was created was also converted to a 20 km by 20 km grid to graphically illustrate where particular break crops are grown, as the agricultural landscape across the WA wheatbelt is diverse. Other resolutions were not explored, as the objective was to visualise crop choice across the entire state, at a level of granularity that was visually appealing. To summarise, the area of cereal, canola, pasture, pulse and fallow was expressed as a percentage relative to the total managed area (crop + pasture + fallow) for the 20 km grid cell. The information was averaged over 4 years, temporal variations in crop choice were not considered here. These data are presented to illustrate where particular rotational choices dominate.

Results

Wheat yields averaged 1.88 t/ha, and ranged from 0.03 t/ha to 6.4 t/ha. The main effect of crop rotation marginally affected wheat yield across 4 seasons of the entire WA wheatbelt. Wheat after canola, and wheat after legumes yielded 2.00 and 1.96 t/ha. Wheat after cereal averaged 1.84 t/ha. Wheat after pasture and wheat after fallow averaged 1.87 and 1.82 t/ha. The growing season rainfall across all data averaged 213mm, and this was used to estimate the predicted yields from the regression model. The standard error on these data was 0.3 t/ha, so treatment differences between rotations across the whole dataset were not significant. However, the interaction terms in the statistical model were highly significant (Figure 1). Again, effects were presented assuming the average growing season rainfall of the whole data set (213 mm). Despite this averaging of rainfall, rotation effects on wheat yield did vary from season to season. In 2017 rotation effects accounted for a 250 kg/ha difference in yields. In 2018 the difference was 370 kg/ha, in 2019 it was 230 kg/ha and in 2020 it was 160 kg/ha. In 2017 the best break for wheat was a legume crop. In 2019 the best break was a pasture. In 2018 and 2020 the best break was canola (Figure 1). The strong interaction on the effectiveness of the break crop effect, and the dynamic nature of the break crop effect illustrates how big data analytics can help define how applicable a regional analysis might be to a local situation.

Crop rotation choices were not uniformly distributed across the WA wheatbelt. Patterns emerged, partly influenced by rainfall zone and soil type, but influenced by factors beyond what was captured in the data. Future analyses would attempt to isolate these effects. Pulses featured prominently in rotations in Northern Coastal Agricultural Region representing approximately 20-30% of crop area, but were included in many rotations at low levels across the state (Fig 2B). Oilseeds were prominent in the high to medium rainfall zones through-out the state (Figure 2A). On the south coast, these crops were often grown year in year out with cereal. Fallows featured in the drier regions of the state (Fig 2C), while pastures were present in the coast zone and in the far eastern reach of the wheatbelt (data not shown). Through the central wheatbelt, cereal dominant rotations were common. The generic analysis of crop rotations provides a reasonable assessment of the role crop rotations play on farms. However, at local levels, these findings may or may not be appropriate as the spatial distribution of residuals (data not shown), illustrated that in the southern region the statistical model often under predicted crop yields, and over predicted in the eastern wheat belt.

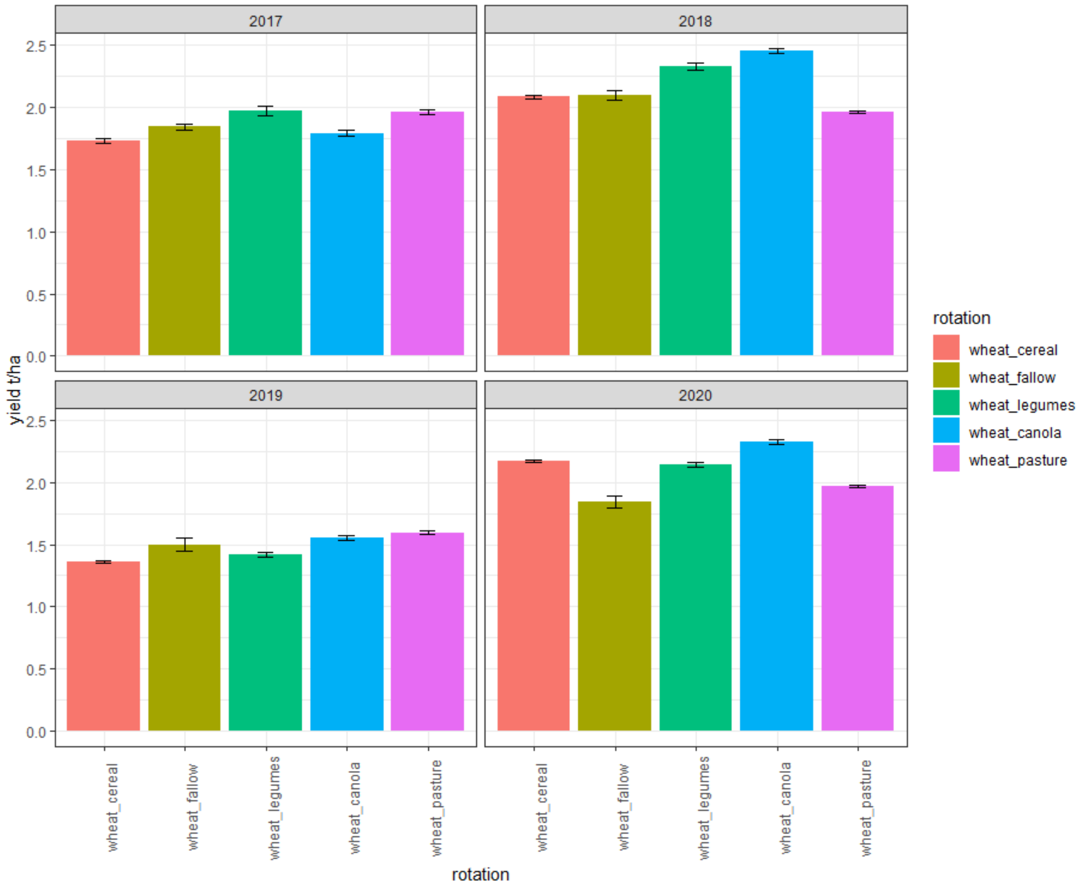

Figure 1. The predicted means for wheat crop yields, following a cereal, fallow, legume crops, canola or pasture from 2017 to 2020 in Western Australia.

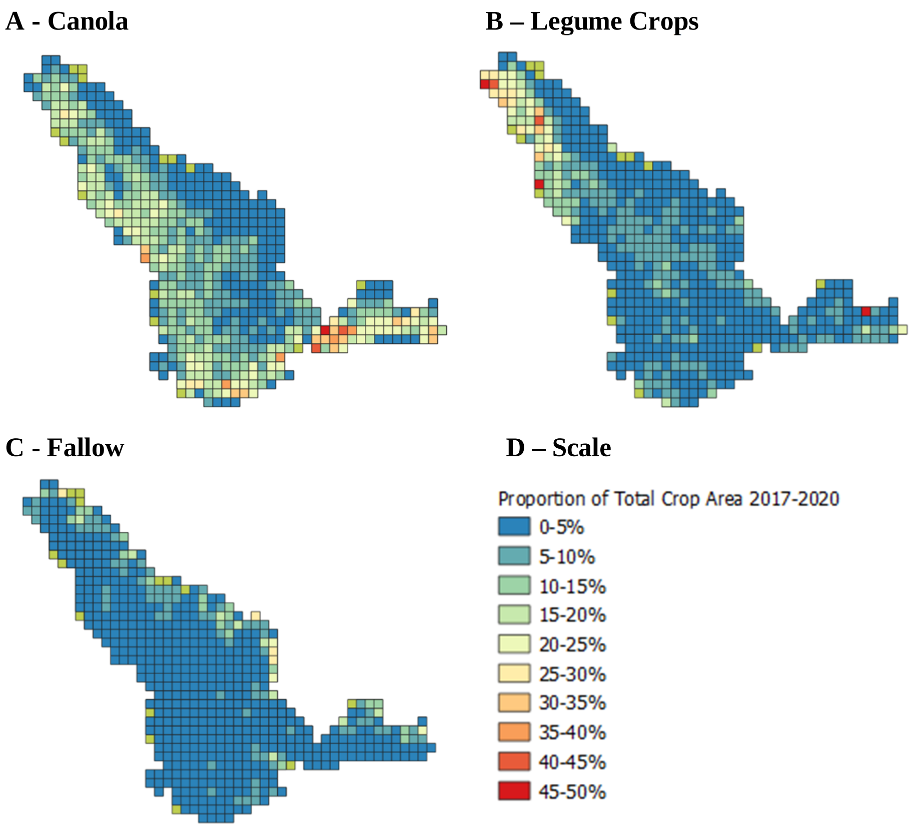

Figure 2. The spatial distribution of the location of various cropping practices across the wheatbelt for A: Canola, B: Legume crops, C: Fallow. Figure 2D is the scale of cropped area that each of these practices represents.

Conclusion

Big data analytics, that use data at the field scale, and combines information about climate, crop type, crop yield and field boundaries can assist to evaluate the performance and location of an agronomic practice like crop rotation at the state level. In time, we expect these analyses to improve, as new data such as soil type, and information about the farmers field are introduced into the analysis. Data science will need to evolve to resolve some of the issues that arise when data such as climate, soil type and shire boundary originate from different scales of resolution.

Acknowledgements

The analytical tools (C-Crop, Crop ID and ePaddocks) used in this analysis were developed independently by CSIRO. The specific analysis of crop rotations and break crop effects was jointly funded by GRDC and CSIRO through the project CSA00056.

References

Donohue R, et al. (2018). Towards a national, remote-sensing-based model for predicting field-scale crop yield. Field Crops Research 227, 79-90. (https://doi.org/10.1016/j.fcr.2018.08.005).

Jeffrey SJ, et al. (2001). Using spatial interpolation to construct a comprehensive archive of Australian climate data, Environmental Modelling and Software 16, 309-330. (https://doi.org/10.1016/S1364-8152(01)00008-1).

Lawes R, et al. (2021). Applying more nitrogen is not always sufficient to address dryland wheat yield gaps in Australia. Field Crops Research 262, 108033. (https://doi.org/10.1016/j.fcr.2020.108033).

Seymour M., et al. (2012). Break-crop benefits to wheat in Western Australia – insights from over three decades of research. Crop and Pasture Science 63, 1-16. (https://doi.org/10.1071/CP11320).

Waldner F, et al. (2019). Needle in a haystack: Mapping rare and infrequent crops using satellite imagery and data balancing methods. Remote Sensing of Environment 233, 111375. (https://doi.org/10.1016/j.rse.2019.111375).

Waldner F and Diakogiannis F (2020). Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network. Remote Sensing of Environment 245, 111741. (https://doi.org/10.1016/j.rse.2020.111741).