The potential of using LAI time series to predict plant available water capacity (PAWC) of soils

Di He [1], Enli Wang [1]

[1] CSIRO Agriculture and Food, GPO Box 1700, Canberra ACT 2601, ACT, Australia, di.he@csiro.au, enil.wang@csrio.au

Abstract
Plant available water holding capacity (PAWC) interacts with climate to determine crop yield and is thus a key factor for predicting spatial yield variations. However, PAWC data at the required spatial resolution are not available because direct soil measurement is expensive and time-consuming. Here, we explore a new approach to inversely estimate PAWC from crop LAI time series using process-based modelling with APSIM together with machine learning. We used the APSIM model to simulate daily LAI of wheat in response to a wide range of PAWC across Australia. Vegetation metrics are derived from simulated LAI time series and were used together with climatic variables to build a machine learning model for predicting PAWC. The model explained 29% to 83% variation of PAWC across ten sites with contrasting climate. This implies a potentially more effective way of PAWC estimation, an alternative to direct soil sampling method.

Key Words
soil-plant interaction, LAI, machine learning, random forest, APSIM

Introduction

Plant available water holding capacity (PAWC) refers to the maximum amount of water held between drained upper limit (DUL) and crop lower limit (CLL). PAWC has been identified as a key factor causing variability in yields of dryland crops (Lawes et al., 2009; Wong and Asseng, 2006). It interacts strongly with climate to determine crop growth dynamics and final yield across regions as well as within paddock (He et al, 2019). It is therefore essential to quantify the spatial variation in soil PAWC for reliable crop yield estimation and site-specific agriculture management in dryland areas where soil variability is high, such as Australia. However, accurate PAWC data at fine spatial resolution are not yet available due to the difficulties to directly measure soil properties across space with traditional soil sampling methods.

Previous studies showed that wheat yield increased with PAWC (Wong and Asseng, 2006) with the rate and extent of increase dependent on year or climate type (Lawes et al., 2009). Another study in South Australian further showed that phenological metrics extracted from a vegetation index were strongly correlated with PAWC (Araya et al., 2016). Theoretically, an inverse modelling approach can be developed to predict soil PAWC with measured vegetation dynamics, or crop biomass/yield combined with climate data.

In order to build such a modeling approach to predict PAWC, crop growth and yield on various soils with a wide range of PAWCs across climatic regions are needed. The limited field data from a few sites that are currently available do not allow extension of the results to large spatial scales. A modelling approach is better suited. He et al. (2019) developed a list of synthetic soils with a wide range of PAWC using soil texture classes and combined them with APSIM (Holzworth et al., 2014) modelling to develop relationships between wheat yield and soil PAWC. Their approach can be extended to inversely predict soil PAWC from crop yield or crop growth dynamics.

In this study, we attempt to extend the work of He et al. (2019) and (Araya et al., 2016) to examine the potential of using vegetation dynamics (e.g. crop LAI dynamics) to inversely predict soil PAWC, and to analyze the relative importance of climate and vegetation dynamics in inverse modelling.

Methods

APSIM model simulations

Ten sites within Australia wheat belt were selected along a north-south and an east-west rainfall transect (Fig 1). Across the sties, mean annual rainfall ranges from 382 mm to 647mm and rainfall patterns change from summer dominant (Emerald) to winter dominant (Ballart). We created 48 synthetic soil profiles using the 6 soil texture classes in the Australian Soil Resource Information System (ASRIS) (Carlile et al., 2001) together with pedeotranfer functions (He et al., 2019). The soils have a PAWC range of 15 mm to 286 mm.

We used APSIM to simulate wheat daily LAI on each of the 48 soil profiles from 1889 to 2017 across the 10 sites assuming sufficient nutrient supply. Due to the uncertain initial soil water conditions, the first 9 years simulation results were discarded. A total of 120 years of simulated LAI were used to construct vegetation metrics and train a machine learning model to predict PAWC from these vegetation metrics and climatic variables.

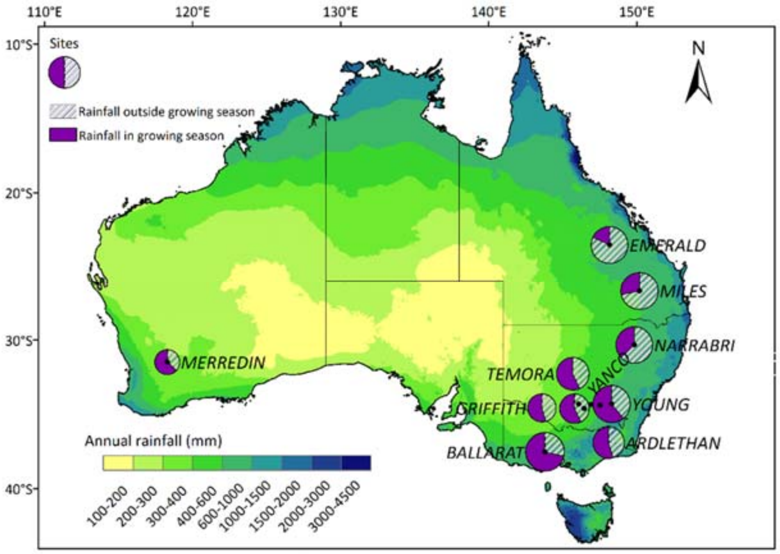

Figure 1 Distribution of annual rainfall across Australia and the sites selected in this study. Along the North-South (N-S) transect are Emerald, Miles, Narrabri, Young, and Ballarat. Along the East-West (E-W) transect are Young, Temora, Ardlethan, Yanco, Griffith, together with Merredin in West Australia (He et al., 2019).

Vegetation metrics and climatic variables

Nine vegetation metrics were calculated from the simulated daily LAI of wheat (Table 1) based on (Araya et al., 2016) who derived the metrics from NDVI time series. Crop growth in the previous growing season impact initial water conditions for the current growing season, so we further calculated 4 metrics from previous growing season. These vegetation metrics together with a list of climatic variables (Table 1) were used to train a machine learning model.

Machine learning (ML) model

A model based on machine learning with the Random forest algorithm was developed to predict PAWC for each of the 10 sites with two scenarios of inputs. Scenario 1 (S1) used only vegetation metrics, while Scenario 2 (S2) used both vegetation metrics and climatic variables. The model was trained and tuned using a 10-fold cross-validation repeated 5 times in R 3.5.1 with 100-year dataset stratified sampling from the total datasets (120-year) grouped by annual rainfall. Accuracy of the model was tested by using the remaining 20-years of independent datasets. We calculated the variable importance with the *varImp* function in the Caret packages in R software.

Results and discussion

Fig 2 shows the comparison of predicted PAWC using the metrics from two consecutive years and the actual PAWC. The skills of machine learning model built under S1 and S2 are similar, with R^2 ranges from 0.29 to 0.85. Climatic variables contributed little to improve model skills. The ML model built with only vegetation metrics explained over 75% of the variation of PAWC with RMSE less than 40 mm for summer rainfall dominated sites, with skills lower at the dry sites and winter rainfall dominated sites, like Merredin (RMSE = 61.4 mm. R^2 = 0.42) and Griffith (RMSE = 71.4 mm, R^2 = 0.29) (Fig 2). Even though there is large uncertainty in some winter rainfall dominated sites. It still implies that in some typical years LAI dynamics of crops grown under nutrient non-limiting conditions can be used to prediction PAWC under contrasting climate conditions, particularly for sites with summer dominant rainfall or high rainfall. Prediction skills can be improved in some representative years.

The relative importance of vegetation metrics is similar across sites. Leaf area duration (LAD), Leaf area duration after the maximum LAI (LADS2) in the current and previous seasons are the most important

www.agronomyaustralia.org/conference-proceedings

vegetation metrics to predict PAWC. For the dry sites, i.e. Merredin, Griffith and Yanco, LAI change rates also become important (Fig 3).

Table 1. Vegetation metrics, climatic and bioclimatic variables used in model training.

Variable type	Variable	Definition
Vegetation metrics in current year	LAIMax	Maximum LAI
	LAIMaxDay	Days from sowing to maximum LAI
	LAIMaxDayafter	Days from maximum LAI to maturity
	LAD	Total LAI of this growing season
	LADS1	Total LAI before reaching maximum LAI
	LADS2	Total LAI after reaching maximum LAI
	LADRate	TLaiS1 - TLaiS2
	LAIIncreaseRate	Maximum LAI/LAIMaxDay
	LAIDecreseRate	Maximum LAI/LAIMaxDayafter
vegetation metrics in previous season	LaiMaxpre	Maximum LAI in previous growing season
	LADpre	Total LAI of previous growing season
	LADS1pre	Total LAI before reaching maximum LAI in previous growing season
	LADS2pre	Total LAI after reaching maximum LAI in previous growing season
Annual climatic variables	TminTota	Annual total minimum temperature
	TminMean	Annual mean minimum temperature
	TmaxTota	Annual total maximum temperature
	TmaxMean	Annual mean maximum temperature
	RadiationTota	Annual total radiation
	RadiationMean	Annual mean radiation
	AnnualRain	Annual rainfall
Growing season climatic variables	STminTota	Growing season total minimum temperature
	STminMean	Growing season mean minimum temperature
	STmaxTota	Growing season total maximum temperature
	STmaxMean	Growing season mean maximum temperature
	SRadiationTota	Growing season total radiation
	SRadiationMean	Growing season mean radiation
	SeasonRain	Growing season rainfall
	Rainpattern	Growing season rainfall/Annual rainfall
	RainDistribution	RainBeforeLAImax – RainAfterLAImax
	RainBeforesow	Rainfall before sowing
	RainBeforeLaimax	Rainfall from sowing to maximum LAI
	RainAfterLaimax	Rainfall from maximum LAI to maturity

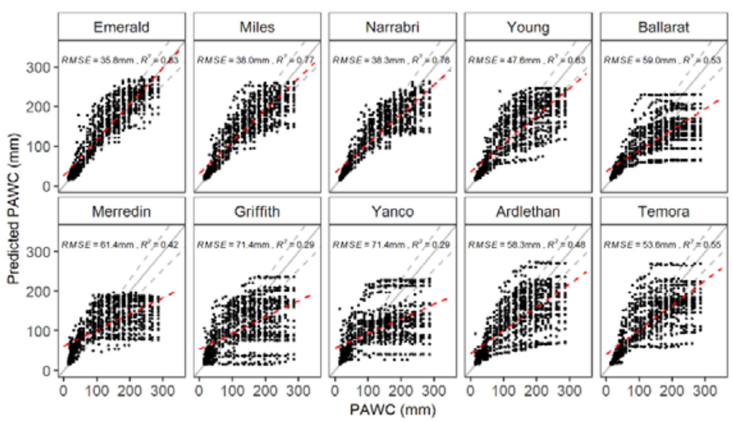

Figure 2. Comparison of predicted and actual PAWC at the 10 study sites. The red dash line is linear regression line. The grey dash lines show 85% confidence interval. RMSE is root mean square error.

Our results demonstrate the potential to predict soil PAWC using LAI dynamics under water limited condition with no other stresses. Our approach can be further extended to use LAI or NDVI dynamics derived from remote sensing data for prediction of soil PAWC at corresponding spatial skills. In reality, uncertainties in management practice, nutrient and other stresses may occur, which could complicate the model and lead to lower prediction skills. Nonetheless, our results imply that skilful PAWC predictions can be potentially made from more accurate measurement of LAI dynamics of crops, which is an easier and less expensive alternative to traditional soil sampling method.

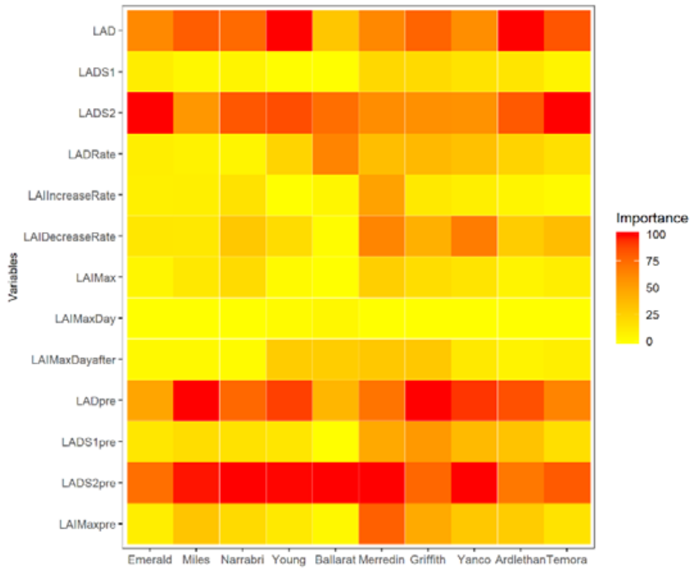

Figure 3. The relative importance of vegetation metrics in predicting PAWC.

Conclusion

This study reveals the potential to use machine learning model to predict PAWC using LAI time series of crops under nutrient non-stressed conditions. Under such conditions the model with vegetation metrics could explain over 76% variation of PAWC across summer rainfall dominated sites. Leaf area duration (LAD), Leaf area duration after the maximum LAI (LADS2) in the current and previous growing seasons are the most important vegetation metrics for PAWC estimation.

References

Araya, S., Lyle, G., Lewis, M., Ostendorf, B., 2016. Phenologic metrics derived from MODIS NDVI as indicators for Plant Available Water-holding Capacity. Ecol. Indic. 60, 1263–1272. https://doi.org/10.1016/j.ecolind.2015.09.012

Carlile, P., Bui, E., Moran, C., Minasny, B., McBratney, A.B., 2001. Estimating Soil Particle Size Distributions and Percent Sand , Silt and Clay for Six Texture Classes Using the Australian Soil Resource Information System Point Database . Canberra.

He, D., Wang, E. 2019. On the relation between soil water holding capacity and dryland crop productivity. Geoderma, under review.

Holzworth, D.P., Huth, N.I., DeVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G. et al., 2014. APSIM - Evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009

Lawes, R.A., Oliver, Y.M., Robertson, M.J., 2009. Integrating the effects of climate and plant available soil water holding capacity on wheat yield. F. Crop. Res. 113, 297–305. https://doi.org/10.1016/J.FCR.2009.06.008

Morgan, C.L.S., Norman, J.M., Lowery, B., 2003. Estimating plant-available water across a field with an inverse yield model. Soil Sci. Soc. Am. J. 67, 620–629. https://doi.org/10.2136/sssaj2003.6200

Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2011.08.010

Wang, E., Cresswell, H., Xu, J., Jiang, Q., 2009. Capacity of soils to buffer impact of climate variability and value of seasonal forecasts. Agric. For. Meteorol. 149, 38–50. https://doi.org/10.1016/j.agrformet.2008.07.001

Wong, M.T.F., Asseng, S., 2006. Determining the Causes of Spatial and Temporal Variability of Wheat Yields at Sub-field Scale Using a New Method of Upscaling a Crop Model. Plant Soil 283, 203–215. https://doi.org/10.1007/s11104-006-0012-5