Extraction of phenotypic traits using unmanned aerial vehicle for field experiments in the agronomy and crop breeding

Pengcheng Hu[1], Bangyou Zheng[1*], Scott Chapman[1, 2]

[1] CSIRO Agriculture and Food, Queensland Biosciences Precinct 306 Carmody Road, St Lucia, 4067, QLD, Australia
[2] School of Agriculture and Food Sciences, The University of Queensland, Gatton QLD 4343, Australia
* Corresponding Author: Tel: +61 (0)7 3214 2620 email: Bangyou.Zheng@csiro.au

Abstract
With advances in camera and robotic technologies, unmanned aerial vehicles (UAV) provide the means to capture images at sufficiently large scale to extract crop phenotypic traits. The current challenges are to manage efficiently the meta-information acquired, involving the processing of large stores of images, visualising intermediate and final outputs, and estimating phenotypic traits with robust algorithms. A cloud-based platform, PhenoCopter, was developed to handle the challenge of data processing. Here we demonstrate how this UAV based platform can enable scientists to extract phenotypic traits including ground coverage, canopy height, aboveground biomass and leaf area index (LAI). Across crops and situations, the UAV-based platform can be shown to be reliable for estimating phenotypic traits.

Key Words
UAV, machine learning, big data, breeding

Introduction

Advances in high-throughput phenotyping provide valuable tools to non-destructively and non-invasively extract phenotypic traits (e.g. coverage, height, biomass, leaf area index) in large-scale agronomic and crop breeding trials. Different types of image sensors mounted on a wide range of craft quickly capture large numbers of high quality images (Großkinsky et al., 2015), which are processed to extract phenotypic traits of different crops (Minervini et al., 2015). A growing challenge in high-throughput phenotyping is to store, remotely access, manage and visualize the datasets and images from large scale field experiments, and then automatically analyse huge datasets to solve different problems (Cobb et al., 2013; Rahaman et al., 2015). With the development of technology, consumable unmanned aerial vehicles (UAVs) are available for general usage and they have been used to collect complex aerial imagery through mounting diverse sensors (Chapman et al., 2014). UAV-based platforms are flexible in the spatial and temporal resolution of data acquisition depending on the experimental objectives (Chapman et al., 2014). Here we show how a cloud-based platform can be utilised to process images collected by UAV platform (PhenoCopter)and develop algorithms to extract phenotypic traits (e.g. ground coverage, plant height, aboveground biomass and leaf area index (LAI)).

Materials and Methods

Acquisition of images using unmanned aerial vehicle

The UAV system was flown for the image data acquisition on different species (e.g. wheat and sorghum) at several sites and at several times during the growing season of the crops. A high-resolution visual camera (e.g. Sony DSC-RX100M3) was mounted in landscape format to collect high resolution visual images. Controlled flight patterns and substantial image overlap are required to obtain satisfactory reconstruction results from UAV image sets. Autonomous flight plans ('lawnmower' designs) were planned to have substantial overlap (ca. 70% forward (image height) and 80% side (image width)) at elevations of 10 to 50 m, depending on the stage of development of the target crops.

The PhenoCopter cloud-based platform has a pipeline comprising several steps: collecting meta information, geo-tagging and checking raw images, reconstructing and tiling ortho-mosaic, segmenting the mosaic into individual plots, extracting plot phenotyping (See Fig. 2 in Chapman et al., 2014 for detail description of the whole pipeline) (Figure 1.) Demo flights can be found online at https://phenocopter.csiro.au.

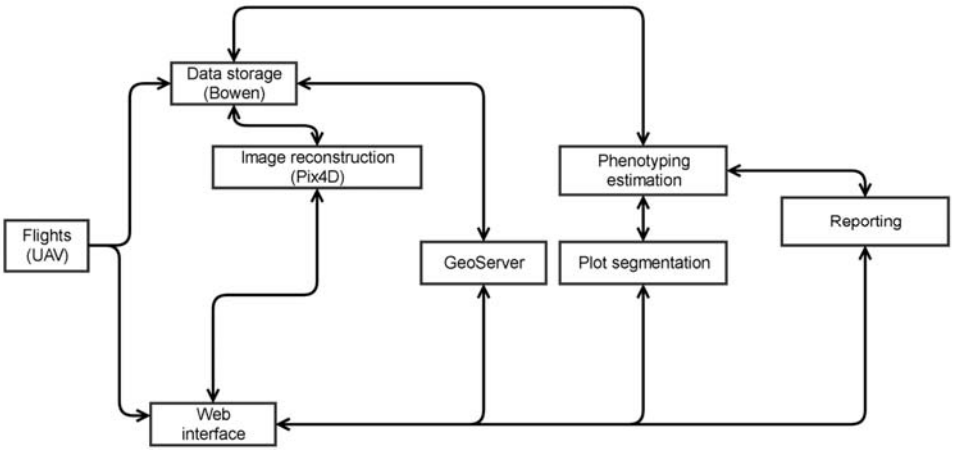


Figure 1. A pipeline of the PhenoCopter platform for processing data captured by unmanned aerial vehicle

Estimation of phenotypic traits from images captured by UAV

Ground coverage: Ground coverage refers to the fraction of soil surface covered by vegetation, as viewed from a nadir position. With a small amount of manual training (10 to 20 minutes), a machine-learning algorithm is applied to classify images into two classes: vegetation and background (Duan et al., 2017; Guo et al., 2013) and then to compute ground coverage. The approach generates the decision tree using the nine colour features (R, G, B, H, S, V, L, a* and b*) of the training dataset by using the CART (classification and regression trees) algorithm. The constructed tree model was applied to segment each pixel into either vegetation or background classes.

Canopy height: A self-calibration method was developed to estimate canopy height which is calculated as the difference between the estimated upper boundary and ground level from a digital surface model (DSM). Each plot in the field experiment was evenly split into 10 slices along the plots to mimic the manual measurement. Percentiles of all DSM pixels (i.e. 99^{th}) were used to represent the upper boundary in each slice. The upper boundary of the plot was calculated as the average of all slices. A smaller number of plots were selected as the self-calibration plots, which were evenly distributed across the whole block, and the other plots were considered as non-calibration plots. The ground levels of the self-calibration plots were calculated as the upper boundary of this plot minus its measured plant height. The estimated ground levels were assumed as the "true values". An inverse distance weighted (IDW) interpolation algorithm was applied to predict ground level in non-calibration plots using the estimated ground levels in the eight nearest calibration plots. Plant heights in the non-calibration plots were calculated as the difference between the calculated upper boundary from DSM and predicted ground level using IDW (Hu et al., 2018).

Aboveground biomass and leaf are index: Vegetation indexes from visual camera were used to build models for the estimation of aboveground biomass and LAI. Visible vegetation indexes were calculated from ortho-mosaics for each plot in a wheat trial, respectively (e.g. VARI, ExR and ExG). Aboveground biomass (AGB) and LAI were estimated using partial least squares regression models (PLSRMs) and robust partial least squares regression models (RPLSRMs), following with a k-fold cross-validation (k equals to 8 in this study).

Results

Images of the wheat canopy were well segmented into two classes (i.e. vegetation and background), including finer detail around the edges of tiny leaves (Figure 2 left). A wide range of ground coverages (0 = nil, 1.0 = full coverage) of the wheat trial were observed from sowing to flowering (Figure 2 below). The ground coverages rapidly increased from 0.06 ± 0.05 at 12 days after sowing to 0.79 ± 0.27 at 53 days after sowing, then remained constant until flowering time (0.80 ± 0.24). The highest standard deviation of ground coverages among genotypes and treatments (nitrogen and water) was observed at 44 days after sowing around the stem elongation stage.

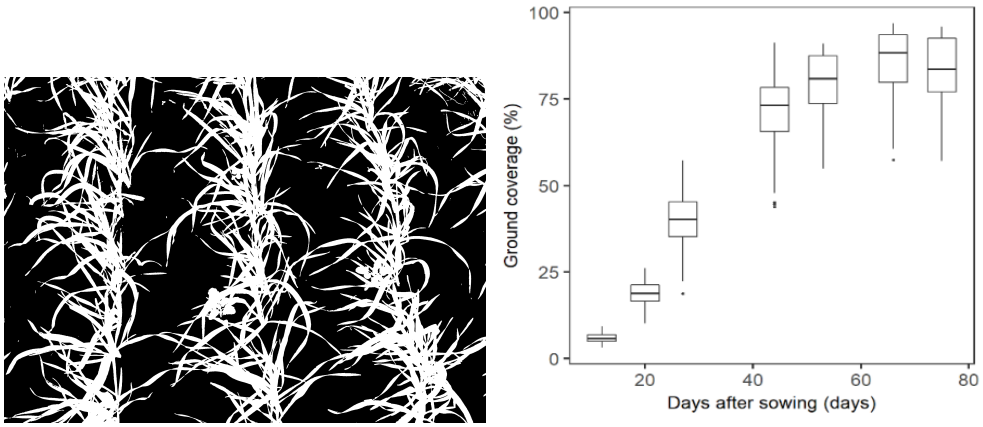


Figure 2. Demonstration of vegetation segmentation of a wheat canopy (left) and the dynamic of ground coverages in a wheat trial (right). Ground coverages were computed from high-resolution images at seven sampling time from sowing to flowering.

Canopy heights were manually measured on 1144 plots at the maturity stage of a sorghum breeding trial. The estimated canopy height from UAV had a good agreement with manual measurements (RMSE (Root Mean Square Error) = 0.07 m, MAE (mean absolute error) = 0.05 m, R^2 = 0.63; Figure 3). The repeatability was 0.78 from manual measurement and 0.74 from estimated heights of UAV. The self-calibration method (using ca. 10% of the trial plots in 4 rows across the trial) obtained accurate estimation of canopy height with low error.

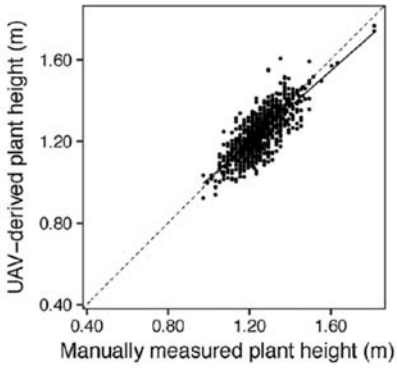
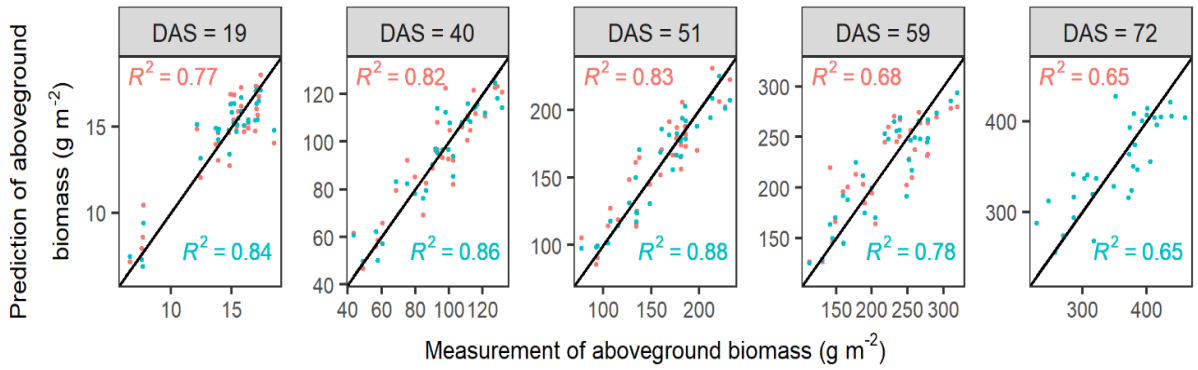


Figure 3. Comparison of manually measured and UAV-derived canopy height of the sorghum breeding trial. Canopy heights were manually measured for 1144 plots at the maturity stage. Figure adapted from Hu et al. 2018.

For each flight, aboveground biomass and LAI were predicted from vegetative indexes using PLSRMs and RPLSRMs. The good agreement between observed and predicted aboveground biomass and LAI were obtained for both PLSRMs and RPLSRMs (aboveground biomass: R^2 =0.65 ~ 0.83 and RRMSE = 0.10 ~ 0.14 for PLSRM and R^2 = 0.65 ~ 0.88 and RRMSE = 0.08 ~ 0.11 for RPLSRM; LAI: R^2 = 0.73 ~ 0.89 and RRMSE = 0.08 ~ 0.16 for PLSRM and R^2 = 0.69 ~ 0.94 and RRMSE = 0.06 ~ 0.19 for RPLSRM, Figure 4) from sowing to flowering in the wide range of aboveground biomass (from 5 to 500 g m^{-2}) and LAI (from 0.1 to 5 m^2 m^{-2}). In general, RPLSRM achieved better performance than PLSRM.



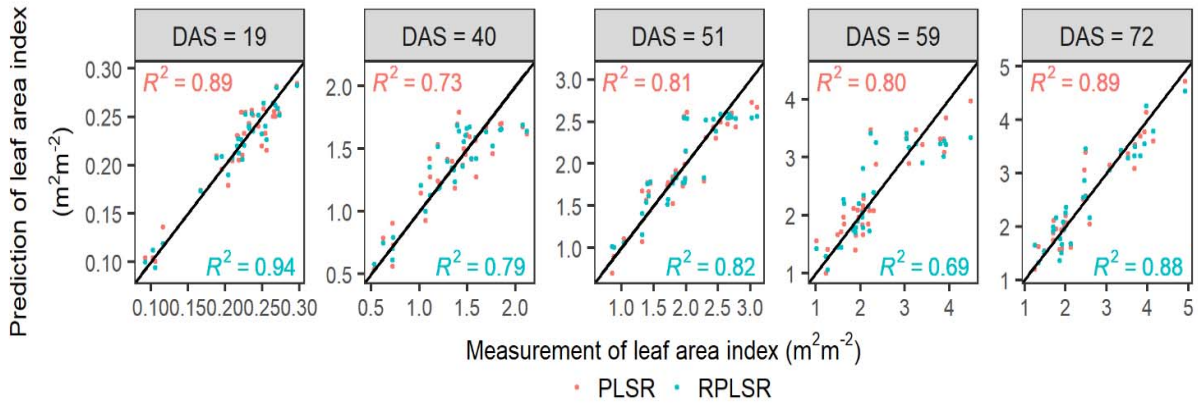


Figure 4. Comparison of measured and predicted aboveground biomass (top) and LAI (bottom) using partial least squares regression models (PLSRMs) and robust partial least squares regression models (RPLSRMs) based on the k-fold cross-validation procedure.

Conclusion

Combined with validated algorithms, the cloud-based platform manages massive datasets from unmanned aerial vehicles and after manual plot segmentation can automatically and efficiently extract phenotypic traits, such as ground coverage, plant height, aboveground coverage and leaf area index. Comparing with manual observations, the unmanned aerial vehicle platform provides acceptable accuracies to extract phenotypic traits in field experiments of agronomy and breeding programs.

References

Chapman, S.C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M.F., Holland, E., Zheng, B., Ling, T.J., Jimenez-Berni, J., 2014. Pheno-Copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. Agronomy 4, 279–301. https://doi.org/10.3390/agronomy4020279

Cobb, J.N., DeClerck, G., Greenberg, A., Clark, R., McCouch, S., 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126, 867–887. https://doi.org/10.1007/s00122-013-2066-0

Duan, T., Zheng, B., Guo, W., Ninomiya, S., Guo, Y., Chapman, S.C., 2017. Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by UAV. Funct. Plant Biol. 44, 169–183. https://doi.org/10.1071/FP16123

Großkinsky, D.K., Svensgaard, J., Christensen, S., Roitsch, T., 2015. Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap. J. Exp. Bot. 66, 5429–5440. https://doi.org/10.1093/jxb/erv345

Guo, W., Rage, U.K., Ninomiya, S., 2013. Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. Comput. Electron. Agric. 96, 58–66. https://doi.org/10.1016/j.compag.2013.04.010

Hu, P., Chapman, S.C., Wang, X., Potgieter, A., Duan, T., Jordan, D., Guo, Y., Zheng, B., 2018. Estimation of plant height using a high-throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: example for sorghum breeding. Eur. J. Agron. 95, 24–32. https://doi.org/10.1016/j.eja.2018.02.004

Minervini, M., Scharr, H., Tsaftaris, S.A., 2015. Image analysis: the new bottleneck in plant phenotyping. IEEE Signal Process. Mag. 32, 126–131. https://doi.org/10.1109/MSP.2015.2405111

Rahaman, M.M., Chen, D., Gillani, Z., Klukas, C., Chen, M., 2015. Advanced phenotyping and phenotype data analysis for the study of plant growth and development. Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00619