Paddock scale modelling and mapping of dry matter yield using UAV derived datasets: A case from dairy farming systems in Victoria

Senani Karunaratne, Elizabeth Morse-McNabb, Anna Thomson, Dani Stayches, Joe Jacobs

Agriculture Victoria Research, Ellinbank VIC 3821, website: www.agriculture.vic.gov.au; E: senani.karunaratne@ecodev.vic.gov.au

Abstract
Traditionally, quantification of dry matter (DM) yield at a paddock or farm scale is undertaken using a rising plate meter (RPM) which provides paddock scale estimates via calibrated equations to predict pasture DM yield. This approach ignores the inherent spatial variability within a paddock which may limit optimum utilisation. In this study, unmanned aerial vehicle (UAV) datasets were coupled with modern machine learning data analytical methods to model and map pasture DM yield variability across individual paddocks at 1 m spatial resolution. The results revealed that the near infrared spectral band had the highest influence in predicting the pasture DM yield. However, the use of additional UAV-derived data sources, such as digital surface and digital terrain models as proxies for pasture height, further improved the prediction. Height derived from the UAV datasets was identified as the second most important variable in prediction of the pasture DM yield. Derived models were cross-validated and also independently validated through data splitting which resulted in concordance values of 0.90 and 0.40 respectively. Model comparison with the calibration equation derived using a RPM revealed that both methods reported equal validation of the results based on the cross-validation. However, the RPM model surpassed the independent validation results of the UAV-machine learning modelling approach. There is potential to explore a wider spectral range and other ancillary datasets for model improvement, in order to improve these machine learning models for prediction of DM yield across the paddock scale.

Key Words
Machine learning, model comparison, pasture on demand, smart farming, smart feeding

Introduction
Due to a growing pressure on land resources, the increasing costs of production and fluctuating milk prices, the dairy industry needs to efficiently grow and manage pasture as a cost-effective source of raw material for primary production. Garcia et al. (2013) derived a direct relationship between pasture utilisation and farm profitability, using data collated from the South-Eastern Australian mainland and Tasmania. This investigation demonstrated a $ 251.24/ha profit increase per additional tonne/ha of pasture dry matter consumed. Despite this finding, the industry still reports a wide gap between actual and achievable pasture utilisation.

State-of-art sensor technology offers new hope for measuring and managing pasture DM yield across the farm landscape. Previous attempts to predict pasture DM yield using satellite datasets, such as MODIS in an Australian context under the "Pasture from Space" project (Henry et al., 2004), did not provide estimates at a spatial resolution required for paddock scale management. Additionally, data availability came with less frequency than desired due to a low temporal resolution of the satellite orbit times, thereby raising issues regarding merging data with management decisions. There is potential to utilise high spatial and temporal resolution satellite data (*e.g.* Sentinel 1 / 2) and unmanned aerial vehicle datasets for improving the prediction of pasture DM yield. Unmanned aerial vehicle datasets have the advantage of very high spatial resolution and flexible spectral resolution and they avoid atmospheric issues during collection, such as cloud cover. The aims of this work described here were: (a) utilise UAV derived covariates to train machine learning models to predict pasture DM yield; (b) validate these models and compare the results with the industry standard method currently used in the dairy industry and (c) discuss the potential of UAV data for paddock scale management of pasture.

Methods
Acquisition of UAV datasets
Four UAV surveys were undertaken on 3, 11, 21 and 25 September 2018 covering three paddocks (total area 2.89 ha) at the Ellinbank research farm (38.2408 S, 145.9414 E). Flight missions were carried out at an altitude of 40 m with a speed of 6 m/s. These paddocks were predominantly (> 90 %) perennial ryegrass (*Lolium perenne* L.) in composition. Images were acquired using a Parrot Sequoia multispectral camera which measured 4 separate spectral bands, namely, red (640-680 nm), near infrared (770-810 nm) (NIR), red-edge

(730-740 nm) (REG), and green (530-570 nm) with corresponding spatial resolution of ~3 cm. At the time of UAV data acquisition, panels of known reflectance were laid on the ground to enable validation of reflectance values for image calibration. Additionally, ground control points (GCP) (*n*=10) made of 'black and white' squares (35 x 35 cm) were placed around the survey area and their precise location and elevation were recorded using Real Time Kinematic (RTK) GPS to enable images to be geo-rectified.

Field sampling for calibration data

Within each paddock, 3 locations were identified using a purposive sampling scheme. At each location, 3 nested samples were collected from 'high', 'medium' and 'low' pasture heights based on visual estimation of the local pasture biomass. A rectangular quadrat (35 x 70 cm) was laid over those identified locations. Once the sampling sites were identified, those locations were also marked with GCPs, so the calibration cuts could be identified in the orthomosaic UAV image after image pre-processing. Once the UAV survey was completed, a rising plate meter (RPM) was used to measure pasture height in each quadrat, with pasture then cut to ground level using battery-operated shears. The cut pasture samples (*n*=88) were dried at 100^{o}C until constant weight was achieved, weighed and used to calculate DM yield (kg DM/ha). Finally, the recorded data (height and DM yield) was used to derive a paddock specific RPM calibration equation to predict DM yield. This method is considered the industry standard.

Image pre-processing

Pix4D software was used for image pre-processing which included radiometric calibration and geometric correction using 3D GCPs. Additionally, a digital surface model (DSM) and digital terrain model (DTM) were derived from each flight's dataset. Point cloud analysis and structure from motion (SfM) methods within Pix4D software was then used to determine an approximate height of the pasture (Height = DSM - DTM).

Deriving environmental covariates

Due to limited radiometric calibration range both the green and red bands were discarded from the analysis. As model drivers, reflectance values of NIR and the Red-Edge bands, height derived from the SfM and two vegetation indices were included, namely, Red Edge Index (REI) and Red Edge Normalised Vegetation Index (RENVI);

$$REI = \frac{NIR}{REG}$$

$$RENVI = \frac{NIR - REG}{(NIR + REG)}$$

The traditional Normalised Difference Vegetation Index (NDVI) could not be derived without the red band.

Modelling framework and model quality assessment

A random forest model was used to determine the covariates of greatest influence in predicting DM yield. Random Forest model is used as a regression model which has the ability to handle both linear and non-linear relationships and multicollinearity among the predictor variables. The number of splits and trees were optimised, the 'mtry' parameter set to 2 and with number of trees set to 1000. Model validation was carried out using the leave-one-out cross validation (LOOCV) technique and splitting the dataset as model calibration (80 %) and validation (20 %) for independent validation. To compare the model results with an industry standard, a local calibration model using measured DM yield (quadrat cuts) and pasture height (RPM) was derived as a simple linear regression model and validated using the same two approaches explained above. Models were validated after calculating several indices including mean error (ME) which explained the model bias, root mean squared error (RMSE) as a measure of model accuracy and Lin's concordance correlation (LCC) which measured the variation of predicted values from a 45-degree line (1:1 line) in a validation plot. The model prediction quality is superior when both ME and RMSE reported lower values and LCC was close to unity.

Results

Relationships between pasture biomass and environmental covariates

Correlation analysis results revealed that there was a high correlation between measured pasture DM yield and pasture height recorded using the RPM. Importantly, all the UAV derived covariates reported a positive correlation with measured DM yield (Figure 1a). The relationship between RPM derived height and the UAV derived proxy for height of pasture also reported a moderate positive correlation (0.34).

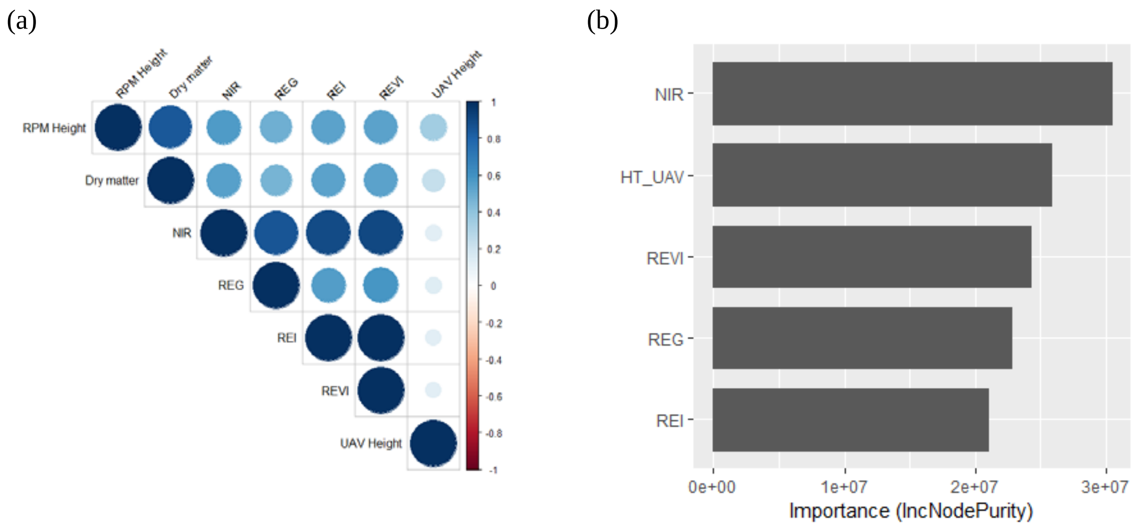


Figure 1: (a) A simple correlation analysis between measured pasture DM yield and UAV derived covariates; (b) A variable importance plot (VIP) from random forest model for the prediction of pasture DM yield

Drivers of the pasture DM yield at a paddock scale

The current modelling work only considered 5 covariates to model pasture DM yield. These covariates can be easily be acquired from a standard UAV and a compatible multispectral camera, with appropriate radiometric calibration. Results of the variable importance plot (VIP) derived from the Random Forest modelling indicated that of the 5 covariates considered, the NIR band had the highest influence in the prediction of pasture DM yield followed by the SfM-derived proxy for pasture height, followed by the REVI vegetation index (Figure 1b). Traditionally, NDVI is used for the prediction of pasture DM yield using either UAV or satellite datasets. Since it is well known that NDVI values can be quickly saturated at relatively low DM yields (Mutanga et al., 2004) exploring alternative predictors that can handle higher biomass scenarios is important as more than 70% of the Australia diary industry is concentrated in the high rainfall regions in the South-Eastern Australia. Our study did not calculate NDVI due to saturation of the red band. One of the key results of this analysis is the ability to use UAV point cloud analysis coupled with SfM, a proxy for pasture height that is a well-established indicator for the prediction of pasture DM yield.

Model quality assessment

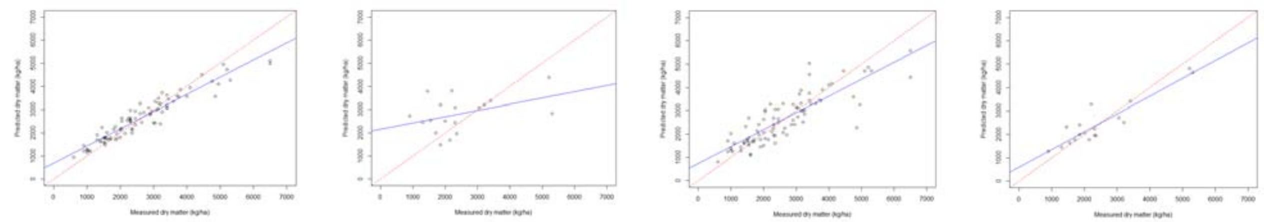
A comparison of the results of LOOCV and independent validation for both the UAV-machine learning model and the locally-calibrated RPM model (Model resulted in adjusted R^2 value of 0.71) are depicted in the Table 1. For LOOCV, the UAV machine learning model performed equally as well as the RPM derived model. However, for the independent validation, the RPM derived model outperformed the UAV-machine learning model (Table 1, Figure 2).

Even though the UAV machine learning model showed potential for modelling pasture DM yield, there is room to improve the modelling framework. For example, training and independent validation machine-learning models can be influenced by (i) the number of samples used for the model calibration, (ii) coverage of a wide distribution of pasture DM yield, and (iii) potential use of new covariates as model drivers. Further work is in progress to collate time series pasture calibration datasets gathered on a weekly basis, aimed at increasing the number of samples and to develop further robust calibration models using machine-learning algorithms and a data-cube approach as outlined by Filippi et al. (2018). These refinements may involve incorporating other ancillary datasets such as soil and climate information into the modelling framework, so that such models can be trained in a spatial context.

Table 1: Model quality assessment

	UAV-machine learning model		RPM model	
	LOOCV	*Independent*	*LOOCV*	*Independent*
ME (kg DM/ha)	3.2	336.5	8.1	9.1
RMSE (kg DM/ha)	431.9	1114.8	647.2	472.8

LCC (unit less)	0.92	0.40	0.83	0.90



(a) (b) (c) (d)

Figure 2: Validation plots for (a) UAV-machine learning LOOCV; (b) UAV-machine learning independent samples; (c) RPM model with LOOCV; (d) RPM model with independent samples

Mapping pasture DM yield across the paddock scale

Predicted pasture biomass maps for the four survey dates, derived from the application of the UAV machine learning model. are depicted in the Figure 3. Based on farm records, a grazing event took place prior to the 21 September UAV survey and Figure 3 depicts a clear decline of pasture DM yield across the paddock following that event.

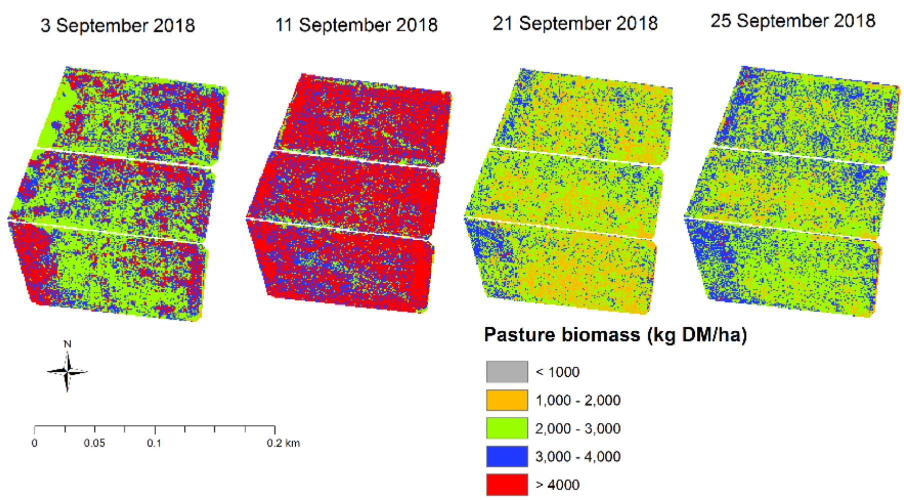


Figure 3: Spatio-temporal maps for predicted pasture biomass across the paddock B

Conclusions

This study demonstrates the potential usage of UAV derived datasets coupled with machine learning to predict pasture DM yield, offering the opportunity to depict the range (variability) in pasture DM yield across a given paddock. This is an important step forward in optimising pasture utilisation. While there is scope to improve the performance of models, we conclude that modern modelling and mapping approaches can be useful to the Australian dairy industry by creating ‘smart feeding’ systems that couple precision DM estimation with the precise allocation and management of pasture.

References

Filippi P, Jones E, Wimalathunge N, Somarathna P, Pozza L, Ugbaje S, Jephcott T, Paterson S, Whelan B, Bishop T (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture, In Press

Garcia SC, Clark C, Kerrisk K, Islam M, Farina S, Evans J (2013). Gaps and variability in pasture utilisation in Australian pasture-based dairy systems. 22nd International Grasslands Congress, Sydney, 15 – 19 September 2013.

Henry DA, Mata G, Gherardi SG, Donald GE, Gittins SP, Smith RCG, Edirisinghe A, Oldham CM (2004) Pastures from space - validating remotely sensed estimates of feed-on-offer. Science Access 1, 234-234.

Mutanga, O. and Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. International journal of remote sensing, 25(19), 3999-4014.