Estimating Pasture Biomass with Planet Labs CubeSats

M Asher[1], P Raedts[2], MT Harrison[2], J Hills[2] and RP Rawnsley[2]
[1]4521 72nd St, Urbandale IA 50322, USA

[2]Tasmanian Institute of Agriculture, University of Tasmania, Tasmania, Australia.

Abstract

Dairy farms require efficient monitoring of pasture biomass. However, physical proxy or destructive measurements are time-consuming, subject to sampling and collection error, and cannot characterise spatial biomass within entire paddocks. Until recently, spatial resolution and temporal frequency of satellite data have been insufficient to achieve effective management. We developed a model for biomass prediction based on data from new Planet Labs satellites to evaluate its utility in the context of commercial dairy farming. While the model had minimal residual error on the farm with the highest quality data (RMSE 260 kg/ha), validation of the model on new farms was poor. Accurately estimating biomass for the calibration farm is promising, but further work is necessary to develop a model which extrapolates robustly.

Introduction

The profitability of New Zealand (NZ) and Australian dairy farms is closely related to the amount of pasture consumed (t DM/ha) (Chapman *et al.* 2008). To optimise pasture consumption, dairy farm managers require timely estimates of pasture biomass. Ground-based measurements are time consuming and subject to sampling errors. Utilising satellite imagery to estimate biomass is an actively researched alternative, including work focusing on pastures (Hill *et al*. 2004 and Mata *et al.* 2011). However, widespread adoption has been limited since no existing approaches provide regular, global estimates at a spatial resolution that is able to capture the variation between or within typical-sized dairy paddocks (2-10 ha). Approaches have been limited by the number of cloud-free images from available satellite data sources namely MODIS (250 m near daily from 2002), Landsat-8 (30 m every 16 days from 2013) and Sentinel-2 (10 m every 5 days from 2017). Aside from work to combine remote sensing data with biophysical simulation models (Hill *et al.* 2004), methods have also been limited to simple regression algorithms, despite the availability of more effective approaches such as gradient boosted trees and convolution neural networks. In 2017, Planet Labs launched a constellation of over 160 CubeSats (3000 cm^3 each) which image the entire globe daily with 3 m resolution. This frequency and resolution appear promising with respect to pasture management. However, the CubeSat satellites are comparatively low cost with low radiometric quality and cross-sensor inconsistencies (Houborg and McCabe 2016), which may limit reliability for biomass estimation. The CubeSat bands include a wider range of frequencies than the corresponding bands on for example Sentinel 2, they overlap with one-another and vary between satellites. Here our aim was to assess the accuracy of the new Planet Labs dataset for the estimation of pasture biomass with sufficient reliability to be useful in operational decision-making of Australian dairy farms.

Materials and Methods

This study used high resolution satellite imagery from the PlanetScope constellation and moderate resolution imagery from MODIS to estimate paddock level biomass at four dairy farms in Northern Tasmania, Australia from June 2017 to May 2018. Every paddock had weekly biomass measurements. Three farms gathered data using rising-plate meters (RPM) and similarly trained operators (one operator per farm, formula 125x + 500), and one used a C-Dax sensor (farm D in Table 1, formula 17.5x + 850). For each paddock and date, we used MODIS NDVI (MOD13Q1.006), Planet (PSScene4Band Analytic Surface Reflectance) NDVI, red, blue, green, NIR, sun elevation, view angle, sun azimuth, and the ratios red:green, red:blue, NIR:green, NIR:blue and blue:green as features. We used a linear interpolation in time to compute the output of interest, biomass, at each of the dates for each paddock from the available weekly paddock level ground measurements. Outliers were removed using an elliptic envelope assuming a contamination of 0.1. A machine learning technique, XGBoost (Chen and Guestrin, 2016), was employed to predict biomass from the above features.

To assess the performance of the XGBoost model, we computed the cross-validation root mean square error (RMSE) (Table 1) on a holdout dataset of 90 days. A different model was trained for each farm. We also conducted a cross-validation using the same model for all farms and leaving one farm out at a time. Normalised RMSE was computed as the RMSE divided by the mean of the observed data. Mean bias was computed as the mean of the observed data less the mean of the predicted data.

Table 1. XGBoost validation RMSE. RMSE1 = validation for one farm leaving 90 days out, RMSE2 = validation using all farms leaving one farm out

Farm	RMSE1	RMSE2
A	610	524
B	472	426
C	511	519
D	374	438

Results & Discussion

To be useful to commercial dairy farming, pasture biomass predictions must be as precise as manual methods and be available daily. A farm specific model could be trained to provide paddock-level growth rates for a dairy farm. Grazing time was not explicitly accounted for in the model – pasture growth was modelled on satellite and biomass observations only. However, the calibrated model performed poorly on remaining farms resulting in overestimation at low biomass levels and significant underestimation of observed biomass at levels greater than ~2000 kg/ha (Maatanga and Skidmore, 2004). The poor model evaluation can be partially explained by:

- Planet Labs data has known radiometric quality and registration issues (Houborg and McCabe, 2016)
- The four farms taken together are a sparse dataset in the selected features. Tree based models have no mechanism for extrapolation.
- Pasture biomass exhibits significant intra-paddock variability. The selected model accounts for paddock averages only.

- NDVI does not fully capture biomass trends. This may be explained in part by pasture's tendency for darker colours with slower growth rates in winter along with lighter colors during times of faster growth.
- The model performed significantly better on the farm with C-Dax data and more uniform perennial ryegrass pastures. The cause of this is unknown.

A number of avenues exist to improve the model:

- Create a higher quality time series of input imagery by synthesizing MODIS, Landsat, Sentinel and Planet imagery for example (Houborg and McCabe 2016).
- Include a land use classification feature in the model to filter out non-pasture pixels.
- Collect higher quality training data for additional farms and date ranges.
- Explore other modeling approaches, for example deep learning methods which incorporate the spatial aspect of the imagery rather than simple paddock averages.

In conclusion, our work with the machine learning technique and imagery from Planet Labs and MODIS indicates that precise estimates of pasture biomass (RMSE 265 kg DM/ha) are achievable. However, further work is required to diagnose the issues causing the poor skill of the model in predicting biomass on validation paddocks.

References

Chapman DF, Kenny SN, Beca D, Johnson IR (2008) Pasture and forage crop systems for non-irrigated dairy farms in southern Australia. 1. Physical production and economic performance. Agricultural Systems 97(3), 108-125.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 785–94. KDD '16. New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785.

Hill, Michael J, Graham E Donald, Michael W Hyder, and Richard C G Smith. 2004. "Estimation of pasture growth rate in the south west of Western Australia from AVHRR NDVI and climate data." Remote Sensing of Environment 93 (4). Elsevier: 528–45.

Houborg, Rasmus, and Matthew F McCabe. 2016. "High-resolution NDVI from Planet's constellation of earth observing nano-satellites: a new data source for precision agriculture." Remote Sensing 8 (9). Multidisciplinary Digital Publishing Institute: 768.

Mata, G, N Purdie, R N Handcock, D Dalley, N Ota, and L Rossi. 2011. "Validating satellite monitoring of dairy pastures in Canterbury with Lincoln University Dairy Farm and commercial farm data." Proceedings of the New Zealand Grassland Association 73: 109–14.

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, DOI: 10.1080/01431160310001654923