Final report

Spatio-temporal prediction of pasture dieback using UAVs and remote sensing

Project code: **B.PAS.0510**

Prepared by: Prof Peter Grace
QUEENSLAND UNIVERSITY OF TECHNOLOGY

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

The project explores the feasibility of using satellite imagery and climatic data to predict and identify pasture dieback (PD) infestation. We acquired 187 PD sites and conducted a time series analysis on Landsat and Sentinel 2 imagery to estimate PD occurrence date. We then used the binary generalized extreme value additive model (BGEVA) to identify the probability (i.e., chance) of PD occurrence relative to the climatic conditions. The BGEVA model reveals that a combination of monthly average maximum temperature between 15-20 ^{o}C and monthly average rainfall between 8-10 mm results in the highest chance of PD occurrence. Seven UAV surveys were conducted, and the resulting high-resolution UAV imagery was used to train a machine learning model. The model classifies unhealthy/dead grass and can be applied to Sentinel 2 images for large scale analysis. Satellite imagery and SILO's gridded weather data were employed to create a Random Forest (RF) model predicting pasture's potential growth. The RF model identifies whether the unhealthy pasture is due to unfavourable weather conditions or other disturbances such as changes in land management or pest infestation. Finally, we created different proof-of-concept web apps demonstrating how our models can be deployed for PD warning and detection at scale.

Executive summary

Background

Pasture dieback (PD) is causing widespread damages to pastures and beef production in Queensland. However, PD causal agents are still poorly understood. This project integrates remotely sensed imagery and publicly available climatic data for cost-effective identification, mapping and monitoring of pasture dieback infestation over time at scale.

Objectives

The main objectives of the project are to:

- Characterise PD spectral signature using high resolution hyperspectral camera.
- Develop a predictive model for PD proliferation based on UAV, satellite imagery, and environmental variables.
- Engage and communicate across stakeholder groups to discuss digital delivery platform (s) suitable to meet industry needs.

Methodology

Time series analysis and statistical modelling were used to predict the chances of PD occurrence based on climatic conditions. Machine learning models were used for PD classification based on satellite and UAV imagery inputs.

Results/key findings

- Three difference models were built to (1) predict the chances of PD occurrence relative to climatic conditions, (2) classify unhealthy grass, and (3) identify if the unhealthy grass was due to unfavourable weather conditions or other disturbances such as changes in land management or pest infestation.
- The models performed well based on the data they were trained on. However, more training data are required for model applications across multiple climatic regions and pasture species.
- Engagement with stakeholders shows interests for model deployment on an on-demand and easy-to-setup platforms or web applications.

Benefits to industry

Our research showcases the usefulness of satellite and gridded climate data in monitoring and predicting PD occurrence. Our models could help graziers plan to mitigate the likely impacts of PD on livestock management and grazing.

Future research and recommendations

Future research needs to focus on acquiring more data to improve the model performance as well as building platforms for deployment of the resulting models at scale.

Table of contents

1. Background

Pasture dieback (PD) is a condition that kills sown and native summer growing pastures. It causes ill thrift and death in a range of introduced and native grasses across Queensland and into northern NSW, resulting in large losses in beef production areas. However, PD symptoms and causal agents are still poorly understood. For many producers and industry advisors the accurate identification of PD from other pasture conditions remains fraught. The Pasture Dieback Science Forum (June 2020) rated the development of accurate identification, mapping and prediction tools for PD to be a very high priority.

Geographically referenced soil, climate and management information on the infected areas is critical when developing predictive spatial modelling tools for accurate identification, monitoring and assessing the impact of mitigation strategies.

Mapping and predicting disturbances allow stakeholders to prioritise management actions at particular locations of concern over large areas. The field of remotely sensed observations to discriminate outbreaks in plant populations has rapidly evolved over the past 10 years with the continued deployment of new airborne and satellite-based platforms offering high resolution hyper spectral and multi spectral data products. These products have been used to detect a range of diseases in agricultural systems in Australia and abroad. Spectral signatures are specific to a species or the biochemical composition of leaves and can be used to identify PD outbreaks.

Remote sensing also provides a cost-effective non-bias, objective assessment of infections and outbreaks. Time series analysis of remotely sensed pasture data at sub-field resolution could potentially be performed on a regular (weekly to monthly) basis using publicly available and commercial data products once a protocol has been developed. Using multi-layer Geographic Information Systems (GIS) combined with spatially explicit remotely sensed data products provides the opportunity to better understand the distribution (when and where) and environmental factors associated with pasture dieback outbreaks in Australia.

One of the most challenging issues in remote sensing is the accurate identification of an infection at scale. Whilst UAV borne sensors with finer spatial resolution and increased spectral bandwidth offer greater accuracy for detection at the plant level (Sandino et al., 2018) they cannot rival satellite-based data in terms of spatial extent and cost effectiveness. However, both Unmanned Aerial Vehicles (UAVs) e.g., drone, and satellite-based imagery are entirely compatible and when used together provide a valuable technology for accurate disease detection at scale (Dash et al., 2018). Satellites offer the ability to cover large distances, at various timesteps. Historical datasets such as Landsat (30m2 by 30m2 and 10-day capture) and Sentinel ($10m^2$ by $10m^2$ and 5-day capture) can be combined with new services such as Planet ($3m^2$ by $3m^2$ and daily capture). Hyperspectral and multispectral imagery from UAVs can be used to calibrate satellite-based imagery leading to greater accuracy at scale.

Spectral phenotyping is a cost-effective method for the non-destructive characterisation of a plant's biochemical and physiological status. Its full potential of this data product is realised in combination with bioclimatic and soil chemical and physical properties and the development of robust predictive causal relationships of dieback in space and time. The performance of predictive modelling approaches not only depends on the abiotic conditions but also the distribution of sampled observations, which are critical for effective model calibration and validation of the statistical model. Whilst traditional generalised linear models have been used for species distribution modelling, the

choice of analytical approach in predictive model development also depends on the availability of relevant data and the biological understanding of the phenomena. In the case of PD the biological understanding is still not fully resolved, therefore multiple statistical approaches are required.

This proof-of-concept project integrated remotely sensed imagery from Unmanned Aerial Vehicles (UAVs) and satellite-based platforms to identify spectral signatures of pasture dieback for cost effective identification, mapping and monitoring over time and at scale using satellite-based systems. Once developed this capability will also allow further monitoring of the impact and mitigation practices to reduce infestations. Remotely sensed imagery was combined with soil and climate data to develop a predictive model of PD in space and time for forecasting risk – this may shed light on likely environmental triggers (biotic and abiotic) that cause episodic outbreaks of PD.

2. Objectives

The project specifically addresses the MLA pasture dieback program objectives of "technologies to enable early detection and subsequently monitoring processes that would enable detection at grazing property level and regional scale". This short proof of concept project aims to combine high spatial resolution UAV imagery (collected with narrow- band hyperspectral/multispectral camera) with broad-band satellite imagery to provide enhanced predictive capability for detecting, monitoring and forecasting changes in PD at scale. The specific project's objectives include:

- Determine the effectiveness of both UAV and satellite imagery for the detection of pasture dieback.
- Analysis of satellite-based time series imagery and gridded climatic conditions to identify environmental conditions when the plant is susceptible to infection or to identify infected plants prior to complete senescence.
- Develop a predictive spatial-statistical model for PD proliferation based on optimal design combining both in situ and remotely sensed data collected at regional sites. Calibration of the predictive model using characterised and sites that are bio-climatically diverse.
- Engage and communicate across stakeholder groups to discuss digital delivery platform (s) that provide multiple levels of interpretation and interrogation suitable to meet industry needs.

3. Methodology

An overview of our methodological approach is presented in Fig. 1.

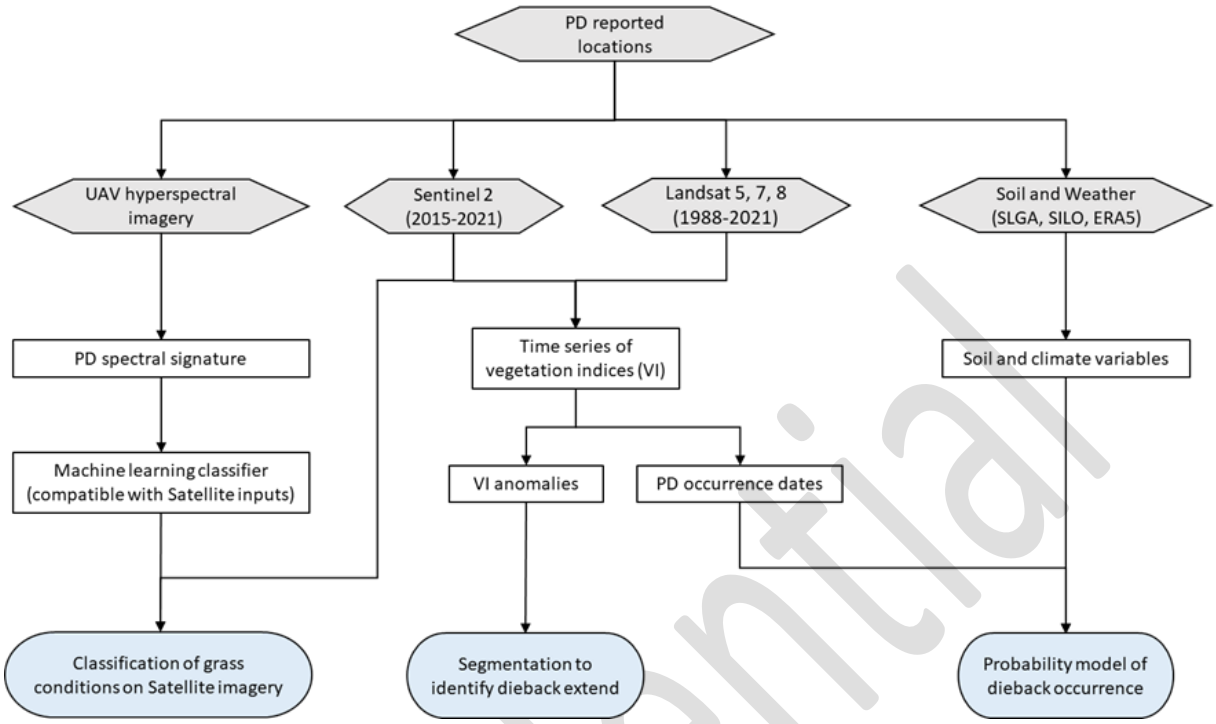


Fig. 1. Overview of the methodological approach.

3.1 PD ground truth data collection

The ground truth dataset consists of 187 locations that were reported with PD infestation at some point in time (Fig. 2, Appendix 8.1). These locations were acquired through ground surveys or provided to us from other representatives of the Pasture Dieback research program, such as Matrix professionals, AgForce Queensland, QUT, and UQ. Since the extent of PD infestation were not reported for these sites, they were identified as points with latitude and longitude coordinates in our ground truth dataset. There were 95 sites where only property addresses were reported instead of the dieback coordinates. These sites were, therefore, only used for our climate correlation analysis and were excluded from our time series analysis with satellite imagery. We further excluded 24 ground truth sites from our initial analysis due to either lack of cloud-free satellite data or unclear PD occurrence history. This resulted in a set of 56 PD sites that were used for our time series and segmentation analyses. Majority of these sites are located in Central Queensland around the Banana region (Fig. 2). The whole PD ground truth collection and the final set of PD sites used for analysis can be explored using the "PD_sites.html" file.

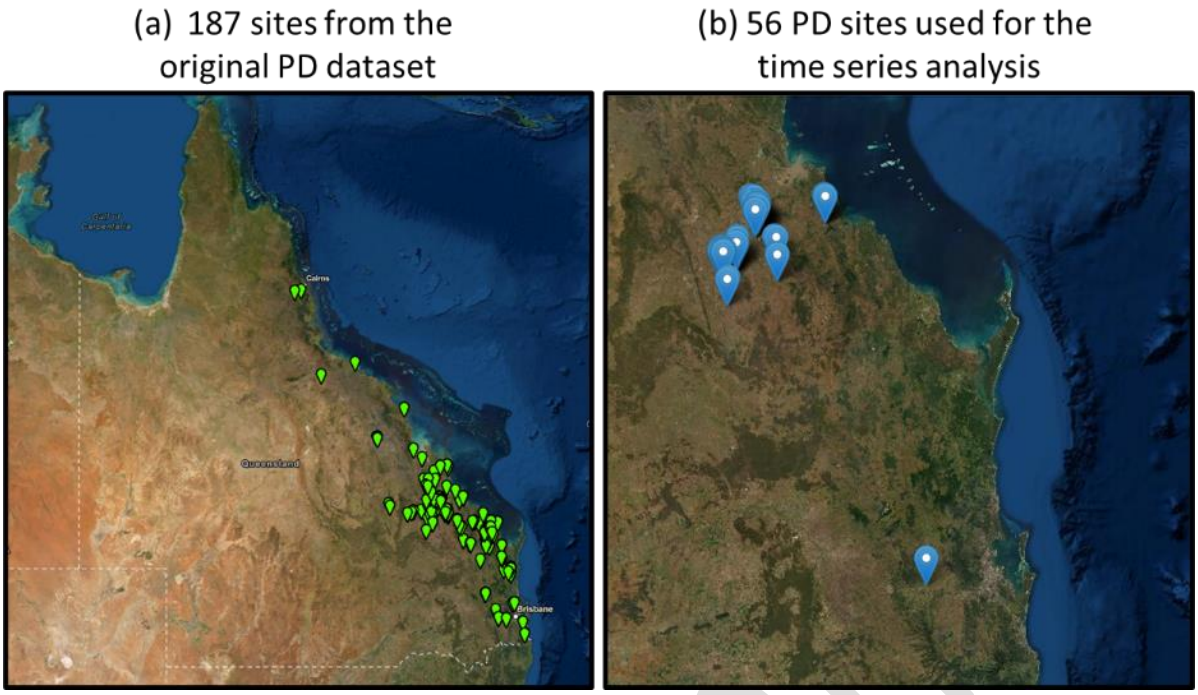


Fig. 2. Maps of Pasture dieback ground truth sites.

3.2 UAV surveys and imagery analysis

3.2.1 UAV imagery analysis

We conducted seven UAV surveys (or data collection campaigns) across five locations in southeast and central QLD, and northern NSW. The flight campaigns occurred from the 5th of February to the 21st of April 2021 because of suboptimal weather conditions to fly the aircraft and payloads. The sites were selected after consultation and feedback from landowners of Pasture Dieback (PD) at various grass species, and in collaboration with QUT researchers who have regularly analysed the grasses for biological tests. An illustration of the locations and specific flight test details are shown in **Fig. 3** and Table 1.

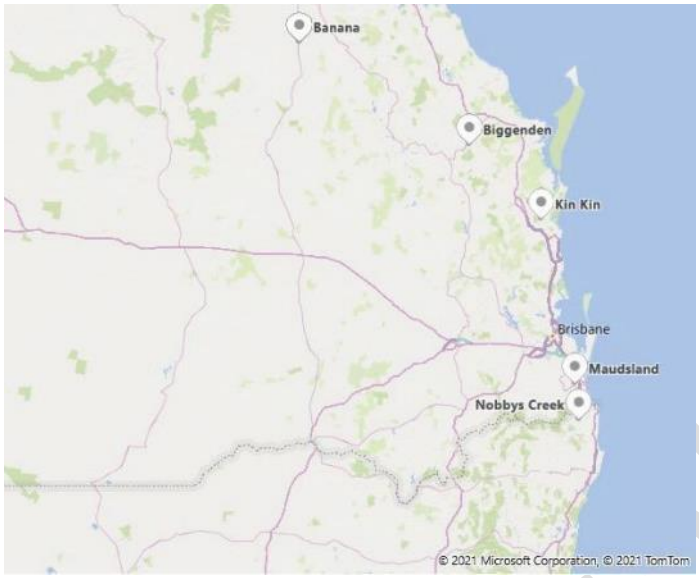


Fig. 3. Selected sites for Pasture Dieback (PD) analysis using UAV imagery.

Table 1. Properties of surveyed sites using UAVs.

Test	Date	Location	Processed Area	Dominant Grass Sp.
1	2021-02-05	Maudsland, QLD	6.62 Hectares	Rhodes, Setaria
2	2021-02-11	Nobbys Creek, NSW	2.82 Hectares	Narrow leaf paspalum
3	2021-02-23	Banana, QLD	6.05 Hectares	Buffel, Urochloa
4	2021-02-24	Biggenden, QLD	5.00 Hectares	Bisset Bluegrass
5	2021-02-25	Kin Kin, QLD	7.86 Hectares	Setaria, Kikuya
6	2021-03-09	Maudsland, QLD	6.62 Hectares	Rhodes
7	2021-04-21	Nobbys Creek, NSW	2.82 Hectares	Narrow leaf paspalum

The data was collected between 10:00 a.m. and 2:00 p.m. using a DJI M600-Pro hexa-rotor UAV. The UAV was tuned on site to operate under three different payload (or camera) configurations: 1) high-resolution RGB; 2) multispectral; and 3) hyperspectral. The order of the scans per flight campaign and relevant settings are detailed in Table 2.

Table 2. Properties of surveyed sites using UAV technology.

Flight	Sensor	Type	Spectral Bands	Altitude AGL (m)	Resolution (cm/px)	Overlapping (%)
1	Headwall Nano	Hyperspectral	274 (400nm to 1000nm)	40	5.9	10 to 20
2	Headwall Nano	Hyperspectral	274 (400nm to 1000nm)	60	8.9	10 to 20
3	Micasense RedEdge	Multispectral	5 (Blue, Green, Red, RedEdge, NIR)	60	2.7	70
4	Canon 5DSR	High-resolution RGB	3 (Red, Green, Blue)	60	0.7	70

UAV campaigns applied the following workflow to ensure compliance with CASA and QUT regulations, and achieve high-quality data:

1) Single communication point to seek authorisation from every landowner to collect imagery for research purposes (i.e., objective 1).
2) Design of flight missions highlighting surveyed areas and access routes of each property using projected satellite rasters and polygons (Fig. 4).
3) Compilation of the approval letter attaching the flight mission and other relevant data that ensures transparency in the research activities on each property.
4) Placement of physical Ground Control Points (GCPs) around the surveyed areas to generate high-fidelity mosaics (Fig. 5).
5) In-site sensor calibration (Fig. 5).

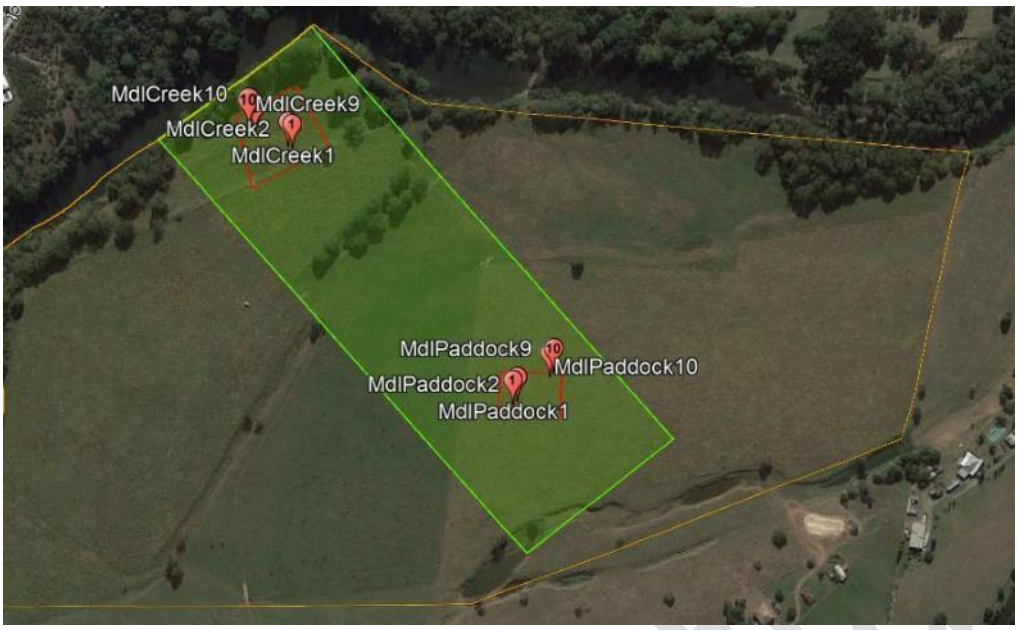


Fig. 4. Flight mission and polygon drawings for UAV data collection at Maudsland, QLD. The yellow, green and red polygons outline the property boundaries, covered area by UAV data and exclusion trials for ground surveys.

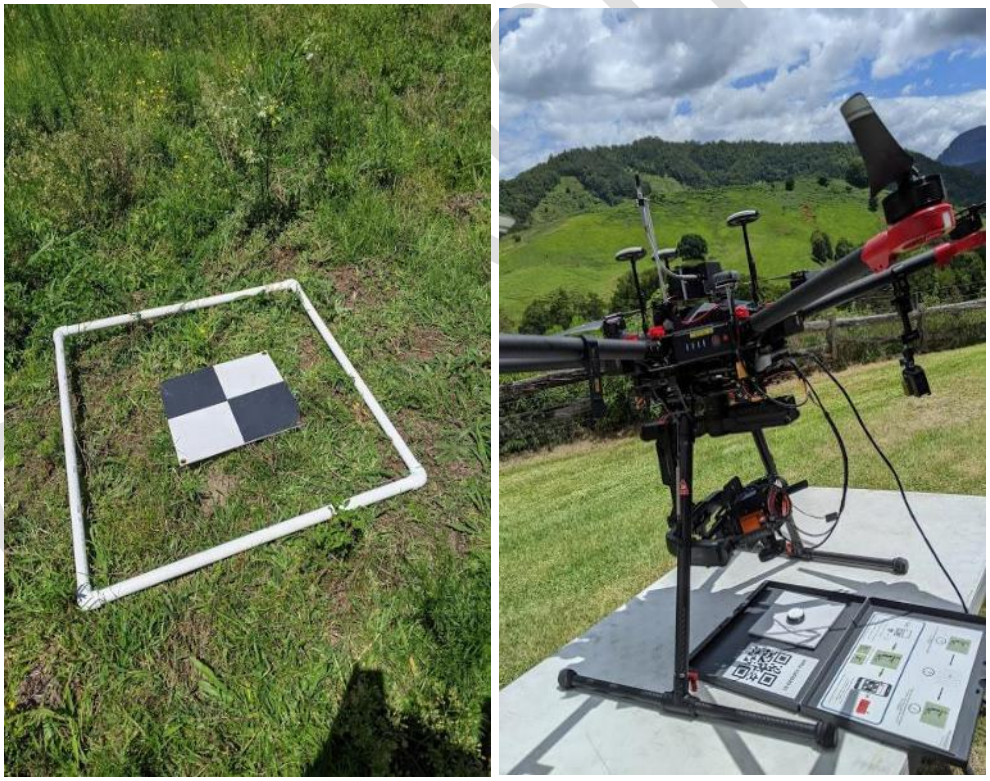


Fig. 5. (left) Placement of a Ground Control Point (GCP) at Nobbys Creek. A minimum of five GCPs must be placed across the extent of the surveyed area. (right) Collection of white reference at Nobbys Creek to correct atmospheric disturbances in spectral imagery.

Georeferenced mosaics of high-resolution RGB, multispectral and hyperspectral data have generated more than 4.6 TB of data. All research data is stored and protected at a dedicated QUT Research Data Storage Service (RDSS) that complies with the QUT MOPP D/2.8.

3.2.2 UAV imagery analysis

The processing and analysis of UAV data is individually applied to each site, owing to the various land conditions and grass species present there. The analysis workflow applied covers the following methods:

1. Spectral correlation analysis between resistant pasture and pasture with dieback symptoms. These correlations are generated by collecting statistics on raw spectral response and common vegetation indices found in the literature.
2. Spatial analysis of affected areas by tuning a Machine Learning (ML)-based model which provides a classification map of surveyed areas by the UAV. The model requires labelling key scene blobs represented in the imagery with a list of defined classes to identify PD. This labelling process is also known as the ground truth.
3. Comparison and usage of tuned models to upscaled satellite data.

A high-level illustration of the workflow is shown in Fig. 6.

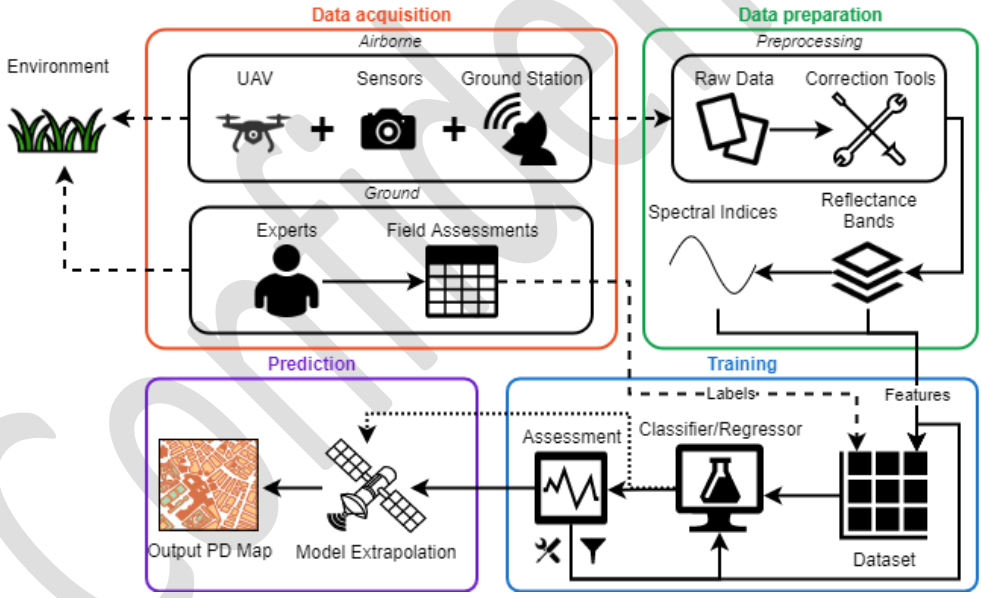


Fig. 6. Overview of the UAV data processing pipeline to detect Pasture Dieback (PD).

The ability to label the correct projected areas (or pixels) of pasture for spectral analysis is essential to run the methods mentioned above. Labelling specific regions for PD presented a particular labelling challenge because of; 1) The nature of how dieback is observed on the grass in early stages, where dispersed colourisation changes happen at the grass leaftops; and 2) Low image resolution captured from multispectral and especially hyperspectral sensors, as shown in Fig. 7.

Fig. 7. (left) Early stages of dieback symptoms visible on the plant leaftops in Nobbys Creek, NSW. (right) Visual projection of the surveyed area using high-resolution RGB imagery (left) and its georeferenced hyperspectral image (right). Due to a decrease in resolution in hyperspectral imagery, precise labelling of areas with dieback is essential for the spectral analysis.

The labelling challenges on UAV data were addressed by applying a data-fusion approach that consists of:

1.) Processing the high-resolution RGB dataset to generate high-fidelity georeferenced mosaics using and EMLID RTK GNSS device in Agisoft Metashape processing software.

2.) Correcting raw hyperspectral transects in radiance and reflectance and applying an orthorectification process using IMU data in Headwall SpectralView processing software.

3.) Georeferencing the orthorectified hyperspectral transects using GCPs and spline transformations in ArcGIS Pro.

4.) Labelling the regions of interest (areas with clear distinction of PD) with cross-validation from ground experts' feedback in ENVI processing software.

To simplify the spectral analysis, a total of five classes were compiled for PD assessments as shown in Table 3**Error! Reference source not found.**. An example of the georeferencing and labelling outputs are illustrated in Fig. 8**Error! Reference source not found.**. In ML theory, the labelling process and its outputs (and referred from here onwards) are known as the 'ground truth' for supervised classification.

Table 3. Number and description of classes for PD analysis.

Class ID	Name	Description
1	Healthy/Reflective Grass	Areas resistant or unaffected by dieback (including grass with mechanical damages).
2	Purple Leaftops	Areas displaying symptoms of dieback at early stages
3	Yellow/Dead Grass	Grass with "gold" colouration or unrecoverable grass at advanced/long-term stages of dieback.
4	Soil	Areas displaying any types of bare soil.
5	Non-grass	Areas containing other vegetation species in the grassland such as weeds and trees. Shadowed areas are also included here.

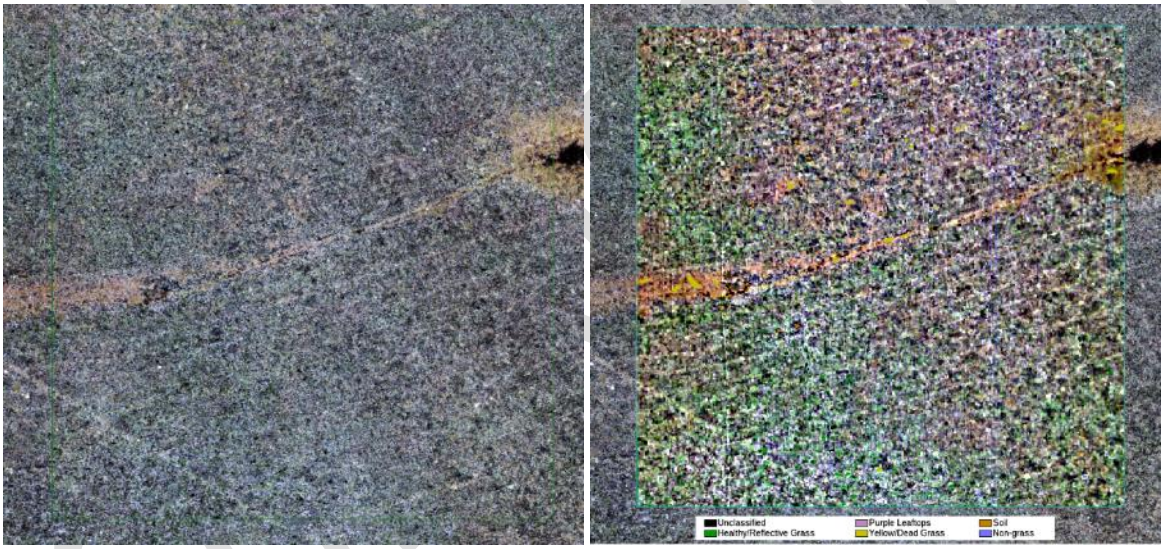


Fig. 8. (left) Georeferenced high-resolution RGB region from Biggenden, QLD. (right) PD labelling process on top of the overlayed and corrected hyperspectral transect. The labelled regions per class were evaluated in collaboration with biological experts by photo-interpreting the high-resolution UAV scans.

In addition to processing spectral response values in reflectance, a set of spectral indices commonly found in the literature were also calculated and included in the analysis. This assessment estimates the mean values of each spectral index from the 'ground truth' labelled regions to identify key indices with direct correlation of PD. The complete list of indices is shown in

Table 4.

Table 4. List of vegetation indices applied to the SHW tool.

Index	Description	Reference
NDVI	Normalized Difference Vegetation Index	(Harris Geospatial, 2021a)
SRI	Simple Ratio Index	(Harris Geospatial, 2021e)
MSAVI	Modified Soil Adjusted Vegetation Index 2	(Harris Geospatial, 2021a)
EVI	Enhanced Vegetation Index	(Harris Geospatial, 2021a)
ARVI	Atmospherically Resistant Vegetation Index	(Harris Geospatial, 2021e)
RENDVI	Red Edge Normalized Difference Vegetation Index	(Harris Geospatial, 2021e)
MRESRI	Modified Red Edge Simple Ratio Index	(Harris Geospatial, 2021e)
MRENDVI	Modified Red Edge Normalized Difference Vegetation Index	(Harris Geospatial, 2021e)
SGI	Sum Green Index	(Harris Geospatial, 2021a)
VRI1	Vogelmann Red Edge Index 1	(Harris Geospatial, 2021e)
REPI	Red Edge Position Index	(Harris Geospatial, 2021e)
PRI	Photochemical Reflectance Index	(Harris Geospatial, 2021e)
SIPI	Structure Insensitive Pigment Index	(Harris Geospatial, 2021)
RGRI	Red Green Ratio Index	(Harris Geospatial, 2021)
CRI1	Carotenoid Reflectance Index 1	(Harris Geospatial, 2021c)
CRI2	Carotenoid Reflectance Index 2	(Harris Geospatial, 2021c)
ARI1	Anthocyanin Reflectance Index 1	(Harris Geospatial, 2021c)
ARI2	Anthocyanin Reflectance Index 2	(Harris Geospatial, 2021c)
WBI	Water Band Index	(Harris Geospatial, 2021b)
GCI	Green Chlorophyll Index	(Harris Geospatial, 2021a)

The ‘ground truth’ data was fed into a supervised classifier model in order to map and visualise trends of PD distribution on the area. The model is based on extreme gradient boosting, or XGBoost

(Chen & Guestrin, 2016), a bleeding-edge decision tree-based model. An advantage of using XGBoost compared to other ML models is that XGBoost provides relevance metrics to the input data the model was fit. This information eases the understanding for the end-user on what were the features that the model relies the most to find correlations on the data.

3.3 Time series analysis of satellite imagery

3.3.1 Satellite imagery acquisition

For each ground truth site, we created a buffer area of 0.0055 decimal degree (~150 ha) (area buffer) around the reported coordinate. Sentinel 2 and Landsat 5, 7, 8 satellite data were acquired, clipped to the buffering area, and corrected for various distortion effects on the reflectance before being used in the time-series analysis. The correction series involves atmospheric correction to compute surface-leaving radiance, bi-directional reflectance modelling to remove the effects of topography and angular variation in reflectance, and adjustments for terrain illumination. Fmask[1] algorithm was used to detect and mask out clouds, cloud shadows and water from the acquired satellite images. The time series of noisy satellite images were merged to create monthly geo-median composites which are largely free of clouds and other noisy data. To further ensure the input image quality, we only retained the monthly composites with less than 5% of invalid pixels within the buffering area. Invalid pixels are pixels identified as cloud, cloud shadow, water, or snow. The process was performed on a virtual machine (VM) on the National Computational Infrastructure (NCI) through the Open Date Cube (ODC) (Dhu et al., 2019), an open-source geospatial data management and analysis software.

3.3.2 Time series analysis

The time series analysis of satellite imagery was carried out to track plant vigour and landscape dynamics at each ground truth location. This helps identify when dieback infestation might occur at each ground truth site as well as specific spectral patterns associated with the occurrence of dieback.

Several Vegetation Indices (VIs) were calculated based on the acquired satellite data, including Enhanced Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI), Modified Soil-adjusted Vegetation Index (MSAVI). These VIs are good indicators for plant vigour and were plotted for each ground truth site over time to inspect abnormal trends.

Since the PD infestation extents are not available, we created a smaller analysis buffer of 0.0005 decimal degree (~1.3 ha) around each ground truth site coordinate (latitude/longitude). This analysis buffer was deemed as an area of interest (AOI) for the analysis. We used time series data to identify when this AOI was first infested with dieback by tracking abnormal values or patterns in their corresponding VIs over time. To eliminate the influences of bare soil and low-growth periods, we retained only periods where the average EVI of the area buffer is above its 50th percentile of the whole time series. An example of time series plots of green band, EVI, MSAVI, and near infra-red band is

[1] http://www.pythonfmask.org/en/latest/

shown in **Fig. 9** while the full analysis for all ground truth sites can be viewed in the "*TimeSeries_Sentinel_raw.html*" file.

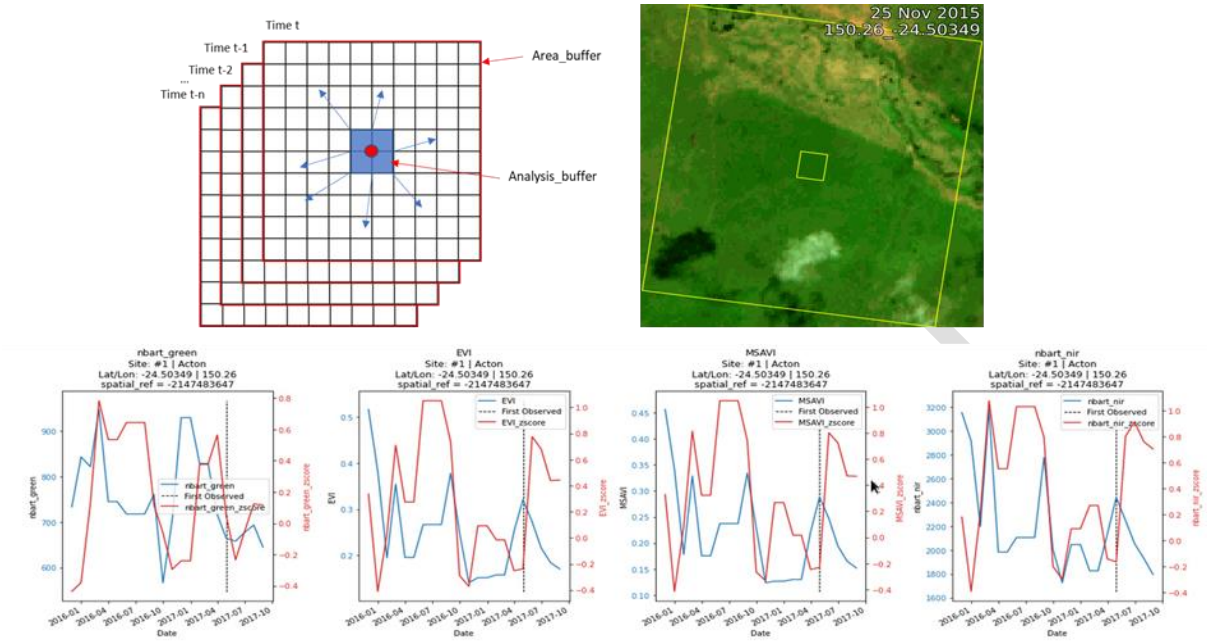


Fig. 9. An Illustration of the analysis buffer (smaller yellow square), area buffer (larger yellow square) and time series plots of different VIs.

On the plots, the dashed line is the date that PD first reported at the site. The blue lines are VI values of the analysis buffer area. The red lines are the z-scores of the VIs. The Z-score shows how the average VI of the analysis buffer area compared to the mean VI of the larger area buffer. Z-score below 0 indicates the analysis buffer is less productive than the landscape average. Z-score was calculated using the Eq. (1)

$$\text{Z-score} = \frac{VI - VI_a}{stdVI_a} \qquad \text{Eq. (1),}$$

where, VI and VI_a are the mean vegetation index value of the analysis buffer and area buffer, respectively, and the $stdVI_a$ is the standard deviation of the vegetation index of the area buffer.

3.3.3 Tracking spatial and temporal anomalies of VIs

VI raw values and VI z-scores (i.e., a pixel value relative to its neighbouring pixels) only indicate the short-term patterns of plant vigour that might be influenced by the current climatic conditions at the PD sites. We further calculated the anomalies of VIs to assess how the current VIs compared to their long-term averages. Landsat 5, 7, 8 imageries (30x30m resolution) for the period from 1986 - 2010 were acquired, resampled to a 10x10m resolution, and spatially aligned with Sentinel-2. The imageries were then filtered based on the 50th percentile of EVI for consistency with the data processing procedure used for Sentinel 2 before being used to compute the long-term averages of the VIs. We calculated VI temporal and spatial anomalies for the analysis area using the Eq. (2) and (3). The temporal anomaly (TA) indicates how the VI value of the analysis area differs from its long-term average (across all time t) (Marshall et al., 2014) while the spatial anomaly (SA) depicts how the VI

value of the analysis area differs from its surrounding areas on the landscape. The anomaly values below 0 indicate that the area is less productive at the current period than its long-term trend. An example of time series plots of temporal and spatial anomalies of green band, EVI, MSAVI, and near infra-red band is shown in **Fig. 10** while the full analysis for all ground truth sites can be viewed in the "*TimeSeries_Sentinel_anomalies.html*" file.

$$TA = \frac{VI - VI_{LT}}{stdVI_{LT}} \qquad \text{Eq. (2)}$$

$$SA = \frac{zscoreVI - zscoreVI_{LT}}{stdzscoreVI_{LT}} \qquad \text{Eq. (3),}$$

where, VI, zscoreVI, are the raw and Z-score values of a vegetation index of the analysis area; VI_{LT}, $zscoreVI_{LT}$, are the analysis area's long-term average VI and its Z-score values; $stdVI_{LT}$ and $stdVI_{LT}$ are the standard deviation of the VI_{LT}, $zscoreVI_{LT}$.

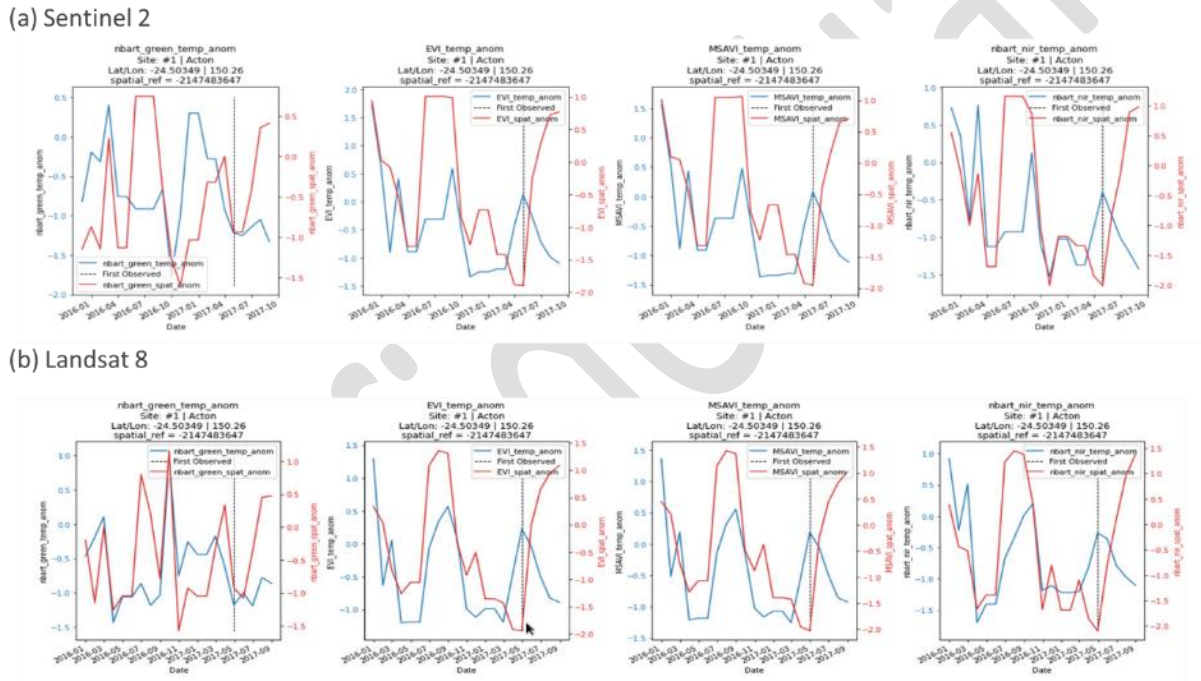


Fig. 10. An example of comparing Sentinel 2 and Landsat 8 as the satellite source for the anomaly analysis.

The analysis for all PD sites using Landsat satellite imagery can be viewed in the "*TimeSeries_Landsat_raw.html*" and "*TimeSeries_Landsat_anomalies.html*" files.

3.4 PD occurrence analysis

Based on the time series analysis, we estimated the PD occurrence date for the PD ground truth locations. The occurrence of PD was considered as the date near the reported date with large dips in both temporal and spatial anomalies on the time-series line charts. In many cases where these dips were not observed, expert opinion was sought to help identify the occurrence date. More information on PD occurrence dates can be found in the Appendix 8.1. PD occurrence dates plus

climatic data were used to train a model to predict the probability of pasture dieback occurrence (**model 2**).

Daily climate data were acquired from SILO[2] (5 km resolution) for the period from 1980 to 2021 and were averaged to the monthly time step. For each climate variable, we calculated its climatic anomaly using Eq. (4). The climatic anomaly compares the weather at a particular period to its long-period average. The measure unit of anomaly is the number of standard deviation from the long-term mean. The long-term period used in our anomaly calculation is 1980 – 2010. Monthly raw data and monthly anomalies up to 24 months prior to the PD occurrence date were used in our probability analysis.

$$\frac{m_i - M_i}{S_i} \qquad \text{(Eq. 4)}$$

where, m_i is mean of current period (i.e., week/month), M_i is mean of the corresponding long-term period (i.e., week/month), and Si is standard deviation of the corresponding long-term period (week/month) for climatic variable i (e.g., rainfall, min temperature).

The most commonly used approach to analyse the occurrence of a particular event (e.g., pasture dieback outbreak) is the logistic regression. However, on spatial and temporal scales, a pasture dieback outbreak represents a rare event. Given the highly skewed distribution of rare events, with the classical logistic regression, the probability of a rare event is underestimated as the response curve approaches zero at the same rate it approaches one. To overcome this problem, the binary generalized extreme value additive model was used to predict the probability of pasture dieback occurrence using only environmental variables. This approach was selected with the aim to identify environmental factors that contribute to the PD proliferation and also to develop a PD outbreak early warning system that provides an indication of PD occurrence risk at scale.

The environmental variables used as predictors were monthly averages of maximum temperature and daily rainfall obtained from 93 locations for 24 months prior to the observed pasture outbreak. Maximum average monthly temperature and average daily rainfall were selected to test the existence of a biological threshold of PD incidence at different levels of temperature and rainfall.

Additional analyses were performed to assess the influence of climatic anomalies on PD extend (affected area at farm scale). A functional generalised additive model was used to analyse the effect of max temperature and rain anomalies on the on-farm PD outbreaks extend (ha). The time series anomalies have been included in the model as functional terms.

3.5 Predicting potential vegetation growth

Detecting abnormal biomass growth using EVI temporal and spatial anomalies can be computationally challenging for larger areas since we need to acquire and process a large amount of time-series satellite data to calculate the long-term average as well as using moving window technique to calculate the spatial anomaly (i.e., how a pixel performs relative to its neighbours). As a result, we used machine learning to examine a less computationally intensive approach. For this, we trained a Random Forest model to predict the maximum EVI (maxEVI) of a summer growing season for each pixel based on the corresponding weather conditions (Model 3). The model's inputs

[2] https://www.longpaddock.qld.gov.au/silo

are the weather data of the current summer along with the maximum EVI of the two previous summers and their corresponding weather data (see Appendix 8.4 for the complete list of model's predictors). Our preliminary model training indicated that adding soil-related variables from SLGA did not improve the model performance while substantially increasing the complexity of the model's input preparation for large-scale analyses. We, therefore, excluded these variables from our model training. The predicted maxEVI defines the potential growth of a pixel given a specific weather condition of a summer growing season and how it performed over the previous two growing seasons. The predicted maxEVI will then be compared with the observed maximum EVI of the season to identify areas with abnormal biomass growth. Large differences between the climate-driven potential maximum EVI and the actual maximum EVI indicate disturbances (e.g., changes in land management or pest infestation) to the landscape rather than unfavourable weather conditions.

To create the training dataset for the Random Forest modelling, we sampled 200 random BOM's weather cells (5km x 5km) along the coastal area of Queensland (Fig. 11). For each sampled cell, we defined an Area of Interest (AOI) of 169 Landsat pixels (~ 15 ha). Landsat 8 data from 2013 to 2021 were acquired and processed for each pixel, and the EVI and other weather statistics were also calculated. For each pixel, the maxEVI was identified as the maximum value of EVI of that pixel across all summer months for a particular year. The sampling resulted in 202,800 (200 cells x 169 pixels x 6 years) data points for the machine learning model training.

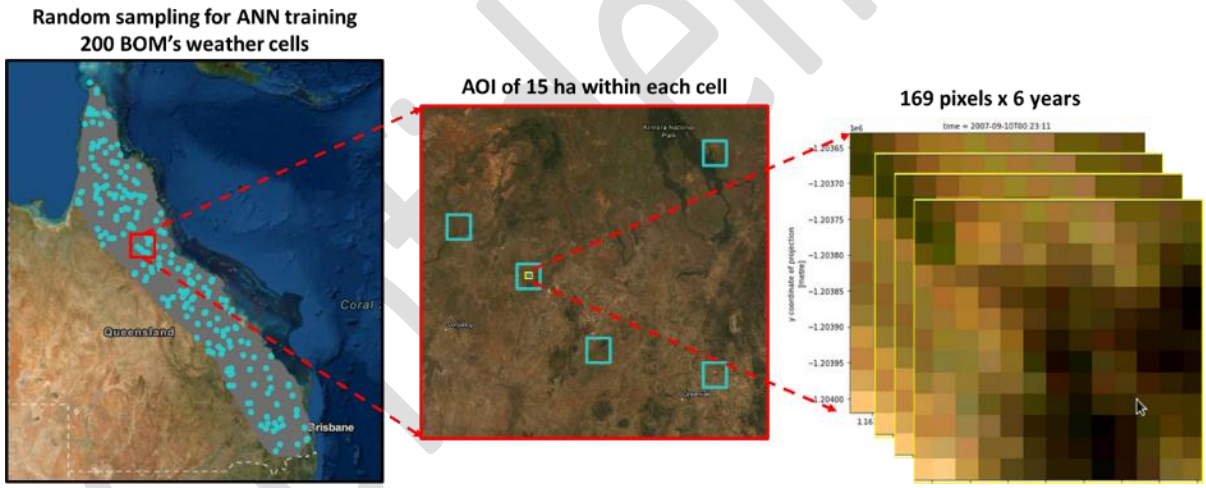


Fig. 11. Creating the training dataset for the Random Forest model to predict maximum EVI of a growing season.

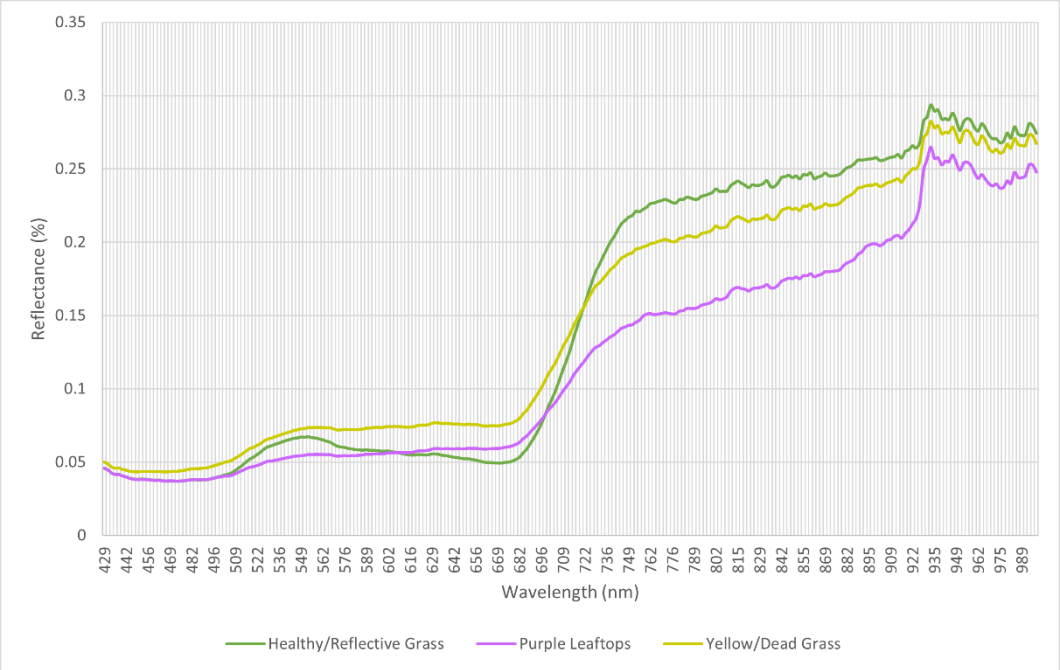
4. Results

4.1 Spectral signature of PD

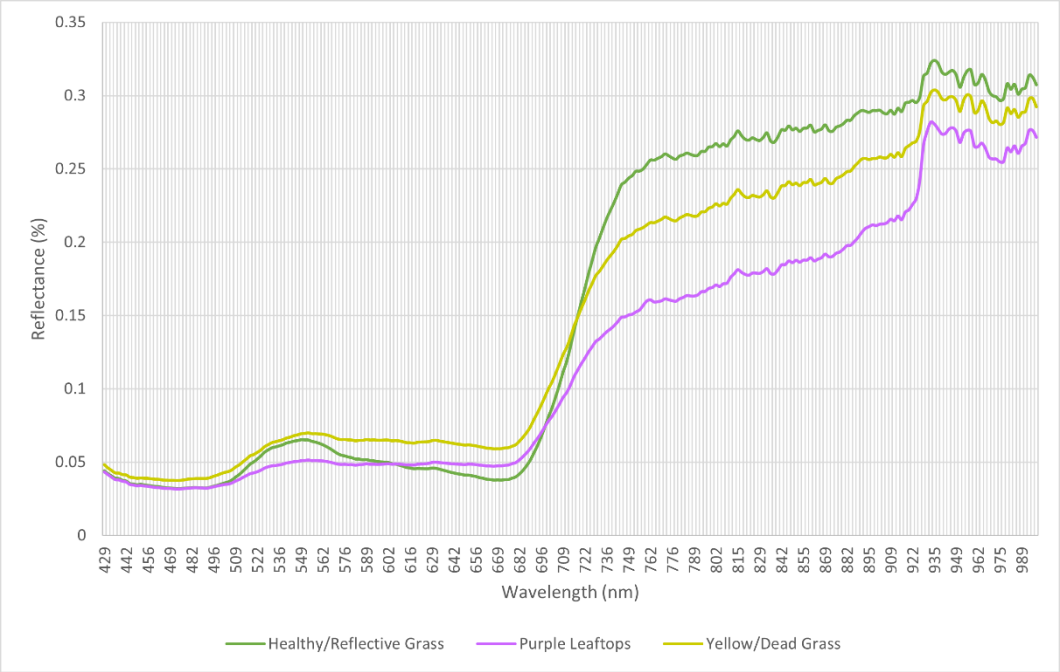
Owing to the high variety in geographical properties and weather conditions in the studied sites from UAV data, this section presents and discusses the combined mean spectral response from; 1) all the seven locations; 2) from 'coastal' locations (i.e., Maudsland, Nobbys Creek and Kin Kin); and 3) from 'Mainland Central QLD' locations (i.e., Banana Station and Biggenden). Spectral responses of individual locations can be found in Appendix 8.2.

4.1.1 Spectral Response in Reflectance

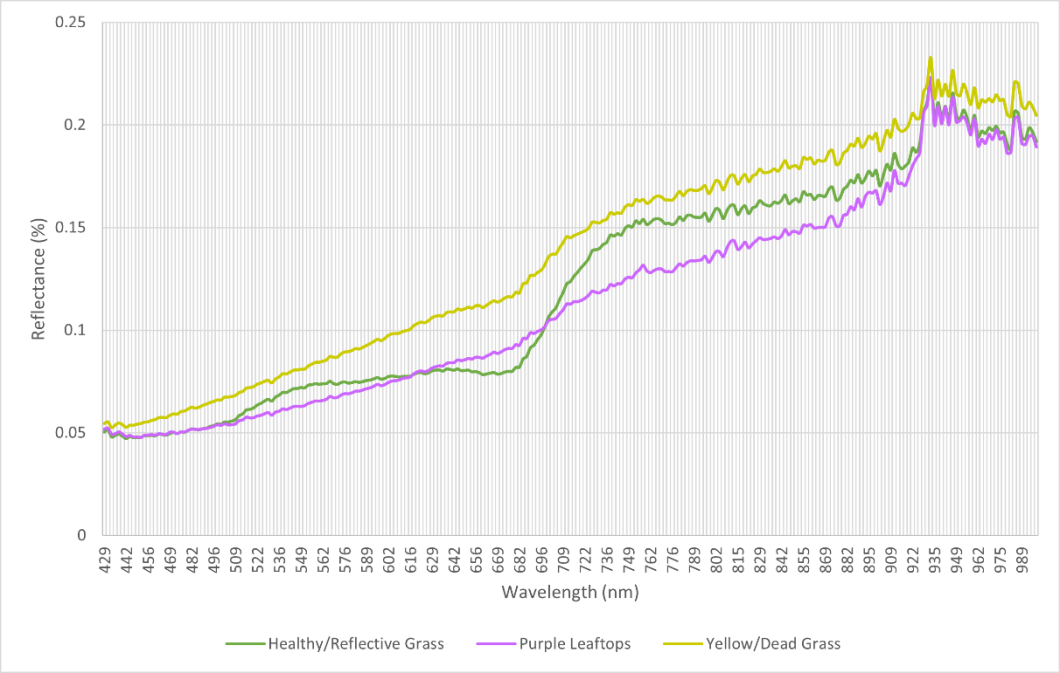
The combined mean spectral responses from the "ground truth" labelled areas are shown in Fig. 12.



(a)



(b)



(c)

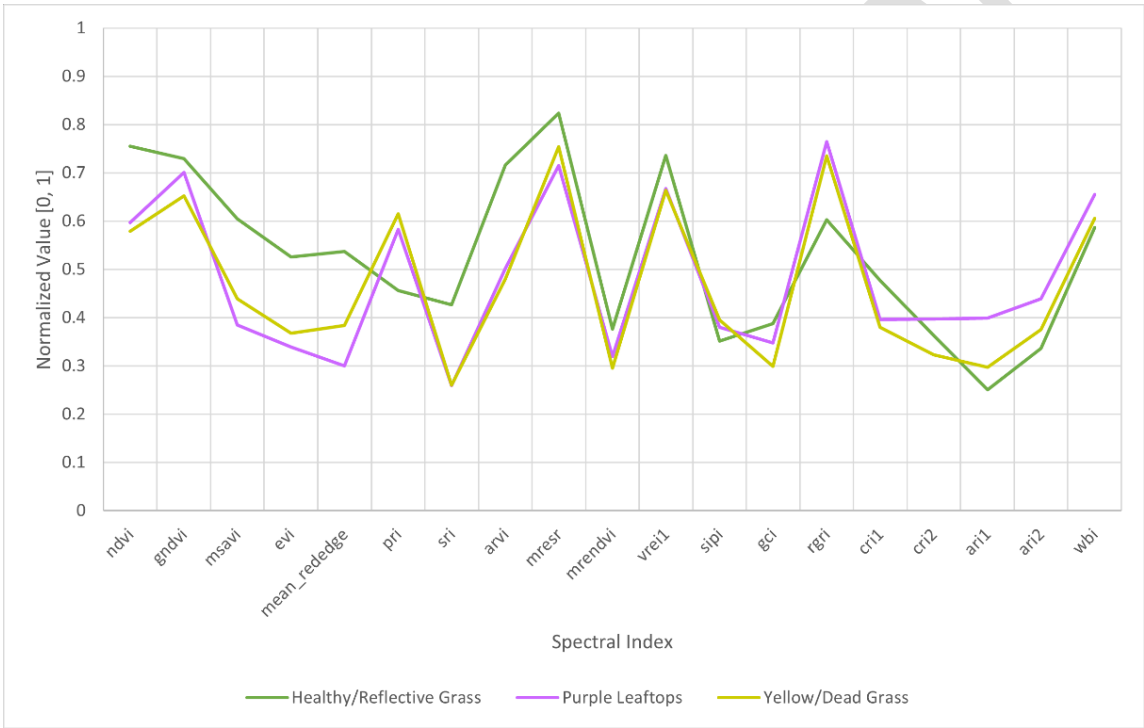
Fig. 12. Mean spectral response for (a) all the sites; (b) "coastal" locations; and (c) "mainland QLD" locations.

Several patterns were observed from the output signatures from healthy grass to dead grass. In coastal locations, a substantial decrease in the visible green (560 nm), red edge (717 nm) and near infrared (840, 930 nm) bands, and an increase in the visible red (668 nm) bands were found between

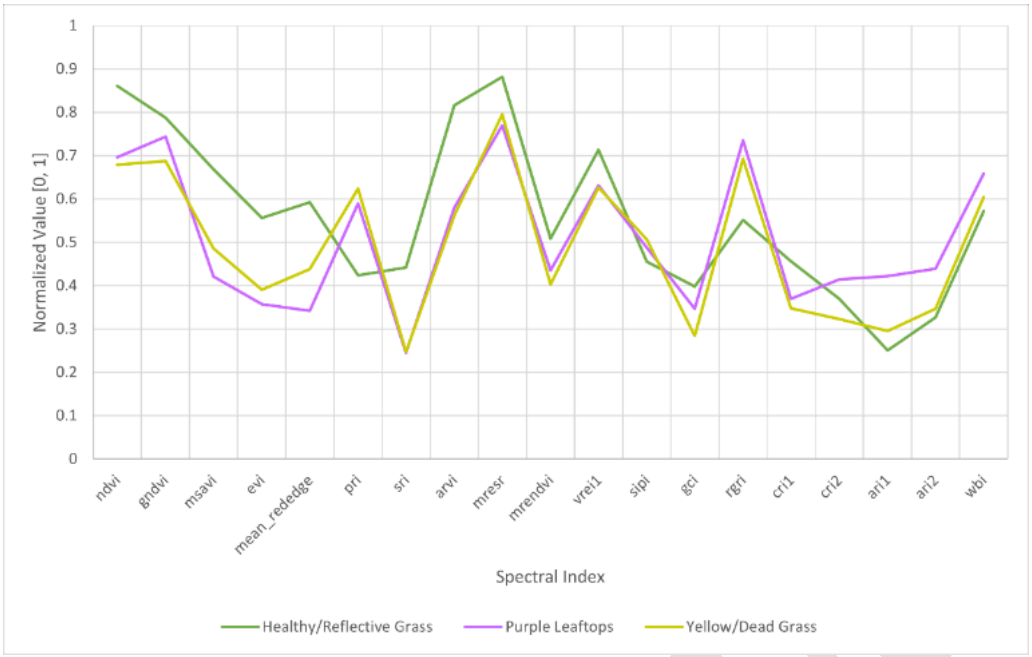
the “healthy” and “purple leaftop” classes. However, the trend for yellow/dead grass followed a different trend with reports of higher reflectance values in the near infrared range, despite such grass to be under advanced exposure to dieback or no longer recoverable. The spectral responses at mainland QLD locations also differed from coastal sites. Even though the “healthy” and “purpling” grass recorded similar variations in the above-mentioned wavelength ranges, the reflectance values of the “yellow/dead” class surpassed the recorded values of the “healthy” class in all the spectral domain. Intra-class variations between the 930 nm and the 1000 nm range were lower for mainland QLD locations compared to coastal sites.

4.1.2 Spectral Index Values

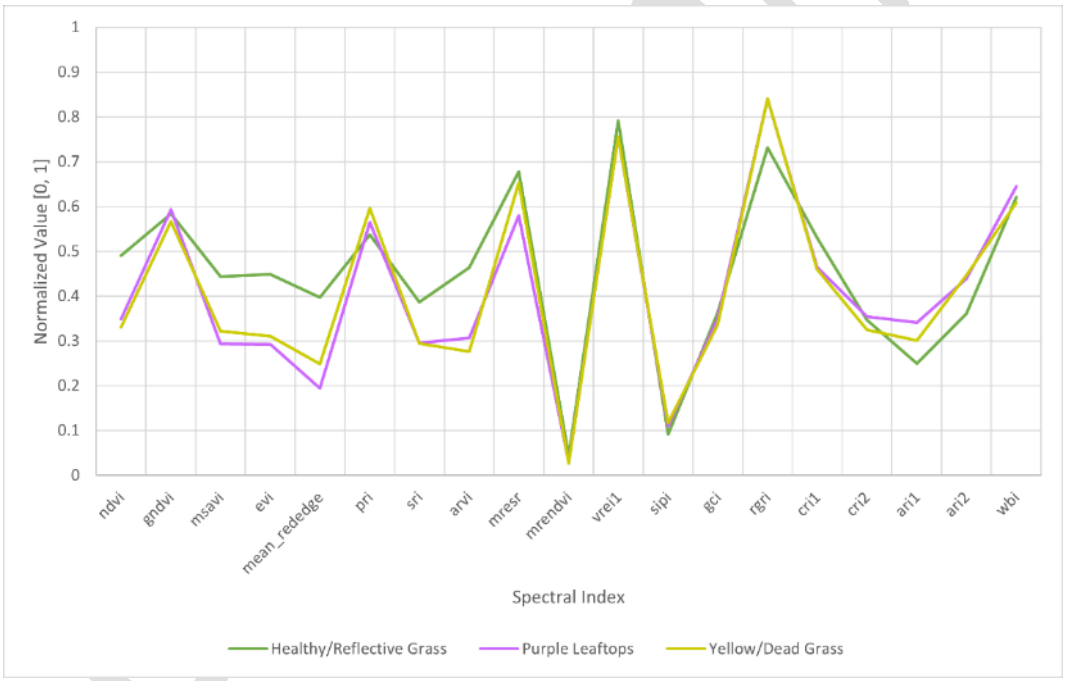
The combined mean normalised spectral index values from the “ground truth” labelled areas are illustrated in Fig. 20.



(a)



(b)



(c)

Fig. 13. Mean normalised spectral index values for (a) all the sites; (b) "coastal" locations; and (c) "mainland QLD" locations.

Overall, the spectral indices that aided tracking the sequence of PD symptomatology (1. Healthy/Reflective Grass, 2. Purple Leaftops, 3. Yellow/Dead Grass) were NDVI, GNDI, PRI, GCI, CRI1. The CRI2 and WBI indices particularly reported higher values for the "purple leaftop" class over the "healthy" class and followed by the "yellow/dead" class, whose order differ from most of the remaining indices. Variations for some spectral index values were high among locations (complete

plots shown in Appendix 8.2). The variance plot illustrated in Fig. 14. suggests that CRI2, ARI2, PRI and WBI spectral indices could be used for a generic analysis at various grass species.

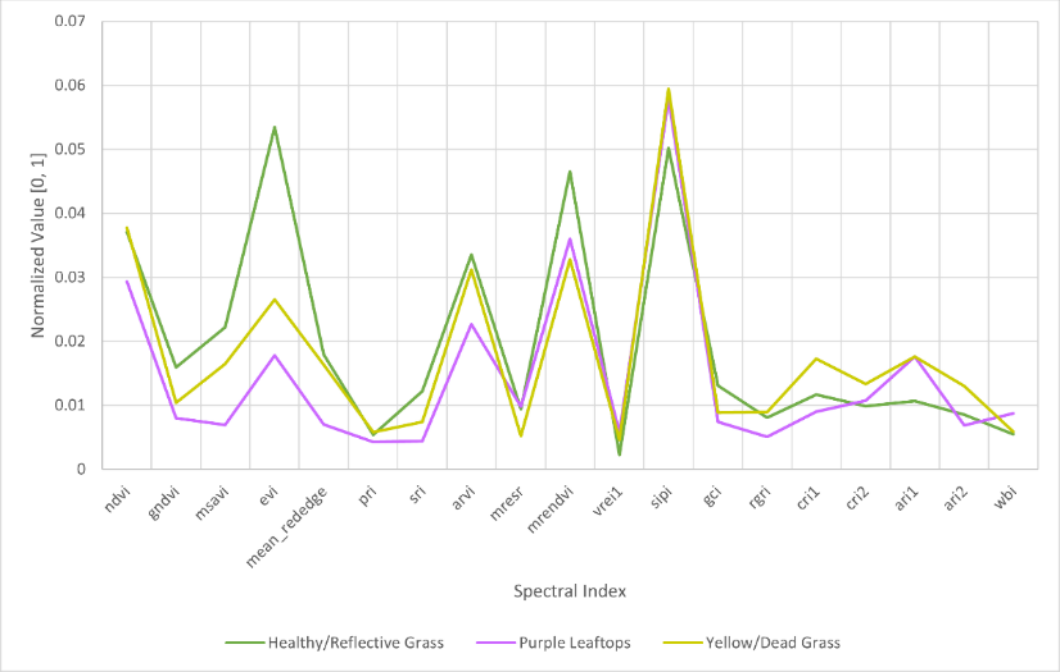


Fig. 14. Variance of normalised spectral index values for all the sites.

4.2 Time-series analysis of PD

The overall pasture vigour was successfully captured for each ground truth location using time-series satellite imagery (Fig. 15). We compared the three vegetation indices (VIs) (including, EVI, NDVI, and MSAVI) and the two spectral bands (including, green and near infra-red (NIR)) in capturing the vegetation dynamics of the analysis area. The three VIs showed very similar patterns over time. Of the two spectral bands, NIR showed more resembling signals to the VIs while being less sensitive to cloud and noisy data than the green band. Repeating the analysis using Landsat 8 imagery showed similar patterns for VIs and NIR (**Fig. 10**). This indicates that coarser resolution Landsat imagery can be used as the satellite source for the landscape dynamic analysis to increase the data availability and reduce the computational burden and data storage requirement.

Using the raw values of vegetation indices for the time series analysis showed no distinct abnormal patterns between the period when PD was first reported and those before that. However, comparing VIs with their long-term values (VI anomalies) did show dips on the time-series plots at or around the reported dates of first PD observation for many PD sites. This indicates that VIs anomalies are potential indicators for PD identification. Given the coarse resolutions of satellite imagery, it is only possible to identify unhealthy or dead pasture on homogeneous landscapes when their extend are large enough, often in hectares. It is not easy to spot PD occurrence and its early-stage dynamics using satellite imagery from Sentinel 2 and Landsat even with the support from pasture experts in addition to using higher resolution imagery from Planet Scope.

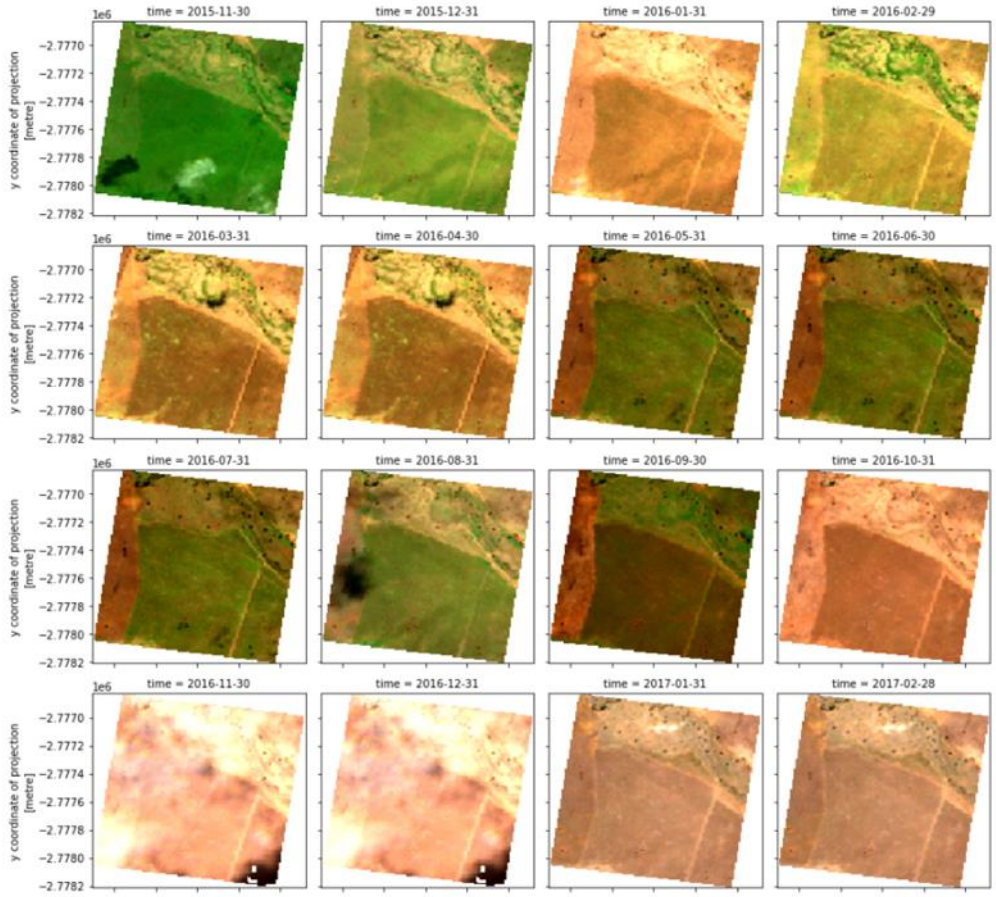


Fig. 15. An example of visual assessment of Sentinel 2 time series imagery at a PD ground truth site. Although satellite scenes have been processed with the state-of-the-art cloud detection algorithms, there were several scenes with undetected cirrus clouds.

Using image segmentation based on VI anomalies helped identify the extend of abnormal pasture growth at many PD sites (Fig. 16). The sensitivity analysis on the hyperparameters of the Quick Shift segmentation algorithm revealed an optimal kernel size of 5 and a cut-off threshold of -0.5 for Sentinel 2 while those for Landsat are 5 and -0.6, respectively. The kernel size defines the smoothing of the sample density. Higher kernel size means fewer clusters will be created. The cut-off threshold defines the anomaly value to separate between healthy and unhealthy pasture. The EVI spatial anomaly (i.e., EVI value of a pixel relative to its neighbouring pixels) is a better metric for the segmentation as compared to the raw EVI value and its temporal anomaly (i.e., EVI value of a pixel relative to its long-term average). Using Landsat as the imagery source for the segmentation did provide quite similar segmentation results.

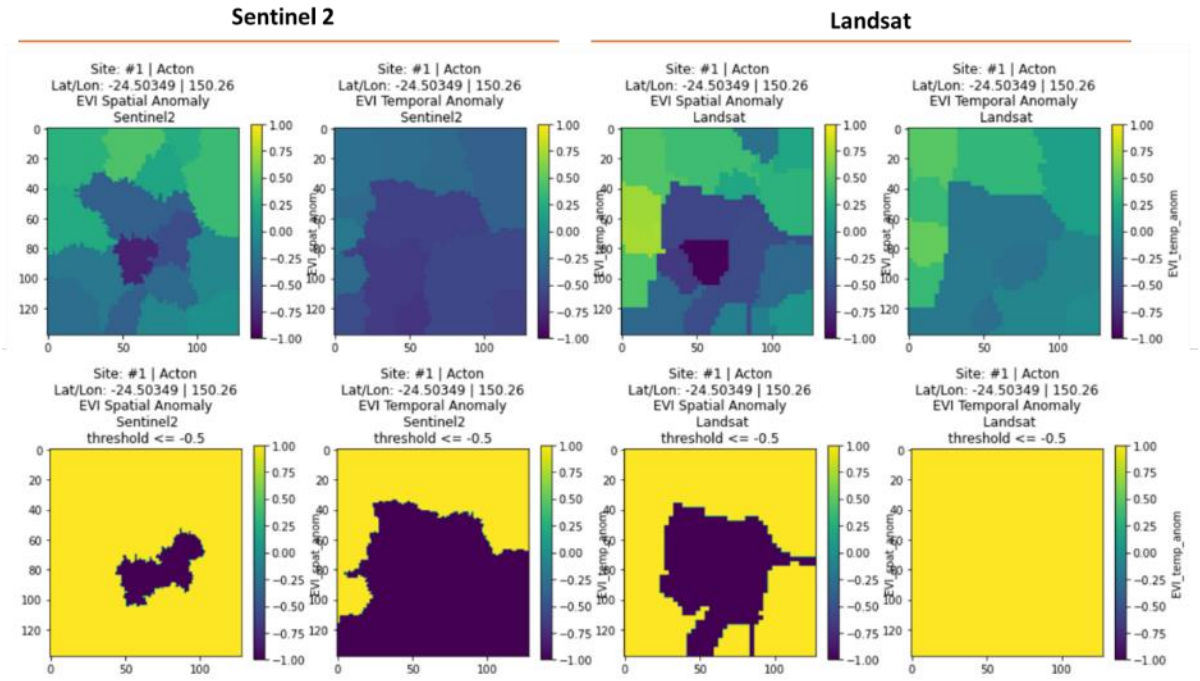


Fig. 16. An example of image segmentation (kernel size = 5) of EVI temporal and spatial anomalies (upper plots) and filtering of segments of low anomaly values at the cut-off threshold of -0.5 and below (lower plots).

4.3 Classification of PD

A dedicated XGBoost model per site was created and tuned using the “ground truth” labelled data. Table 5 shows a summary of the performance metrics. A comprehensive classification report can be found at Appendix 8.3.

Table 5. Accuracy summary of trained ML classifiers per surveyed site using UAVs and hyperspectral imagery.

Class	Maudsland (2021-02-05)	Maudsland (2021-03-09)	Nobbys Creek (2021-02-11)	Nobbys Creek (2021-04-21)	Banana Station	Biggenden	Kin Kin
Healthy/Reflective Grass	91%	98%	95%	97%	96%	96%	99%
Purple Leaftops	96%	97%	97%	100%	95%	76%	99%
Yellow/Dead Grass	91%	98%	91%	90%	97%	96%	93%
Soil	90%	83%	91%	100%	100%	96%	99%
Non-grass	98%	98%	89%	95%	99%	80%	100%
Mean accuracy	96%	98%	94%	97%	99%	94%	99%

Despite image registration challenges between high-resolution RGB and hyperspectral datasets, the tuned XGBoost classifers achieved an overall accuracy above 94% at the surveyed sites. Models of collected data at Maudsland, QLD (2021-02-05), Nobbys Creek, NSW (2021-02-11) and Biggenden QLD were less performant owing to unoptimal weather and illumination conditions which negatively impacted the quality of the collected data. From a pasture class perpective, Healthy/Reflective grass was the class with the highest accuracy outputs, followed by yellow/dead grass and purple leaftops. A reduction in accuracy in the Biggenden model was caused by the challenging labelling process of sampling purple leaftops, represented for most of the hyperspectral scans on a pixel-scale.

Once the models were tuned using the "ground truth", class predictions were applied to unlabelled data to validate general class distribution on the studied exclusion trials and ground assessments from biological experts. An illustration for Maudsland, QLD, Nobbys Creek, NSW, and Banana Station, QLD are shown below:

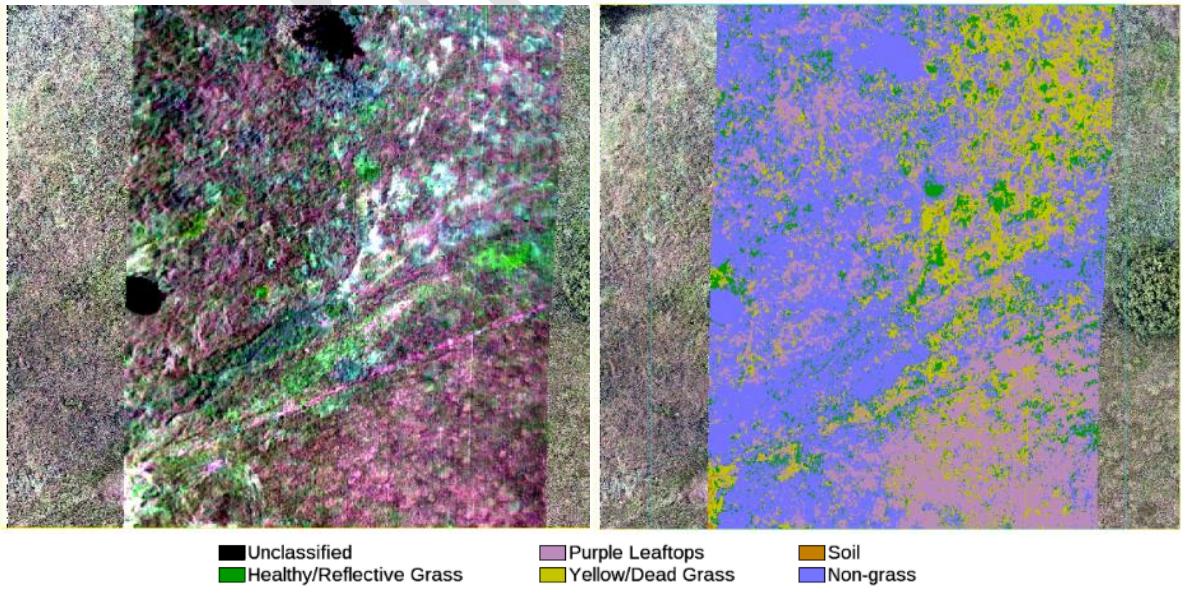


Fig. 157. Predicted classification map of Maudsland, QLD (2021-02-11).

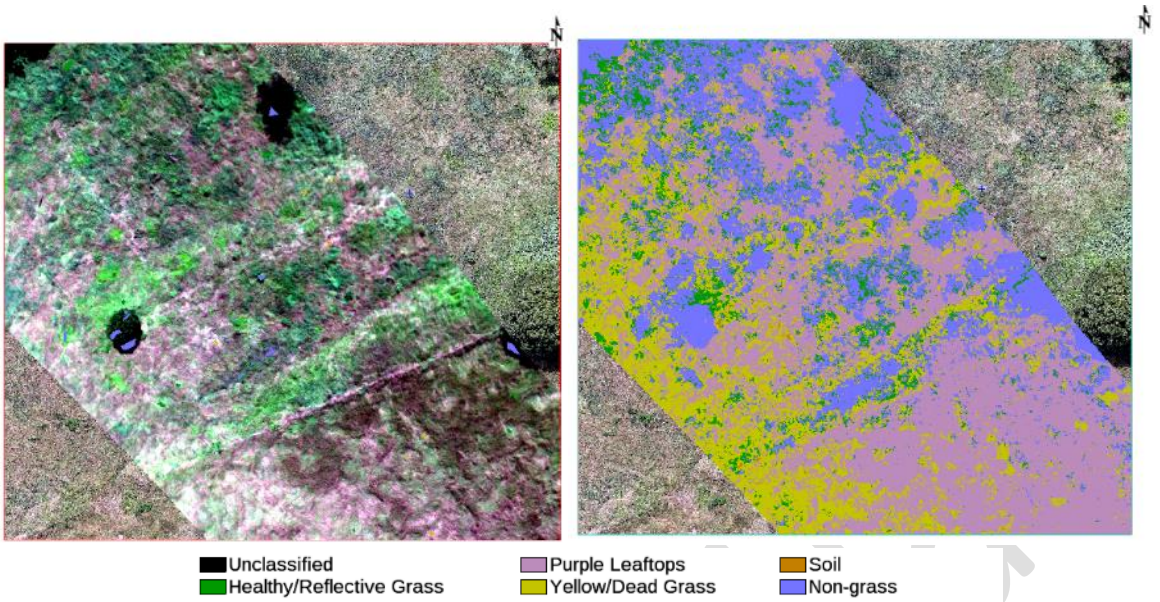


Fig. 168. Predicted classification map of Maudsland, QLD (2021-03-09)

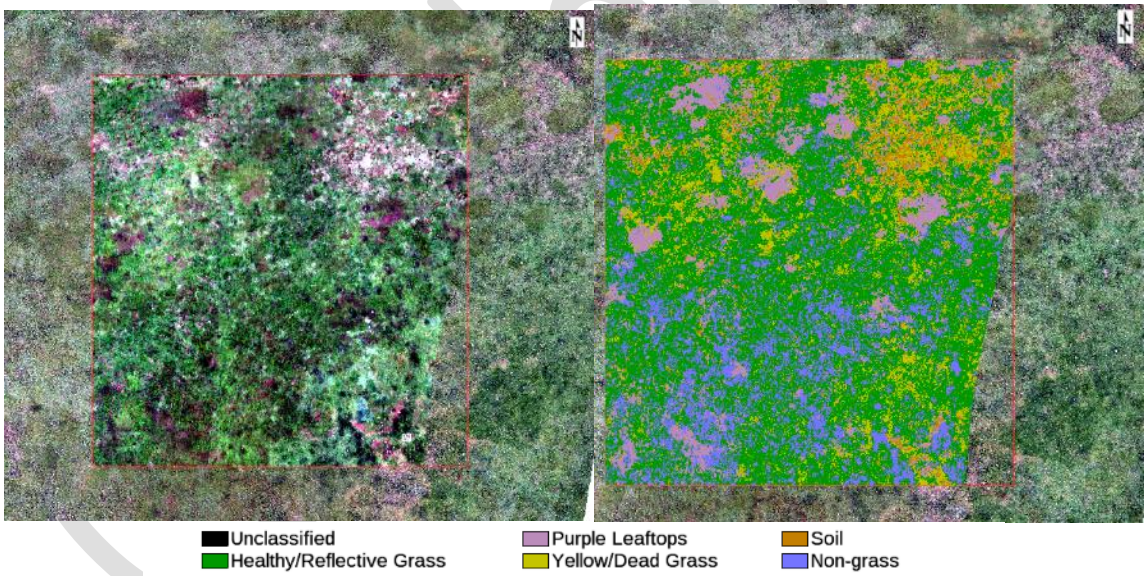


Fig. 19. Predicted classification map of Nobbys Creek, QLD (2021-02-11)

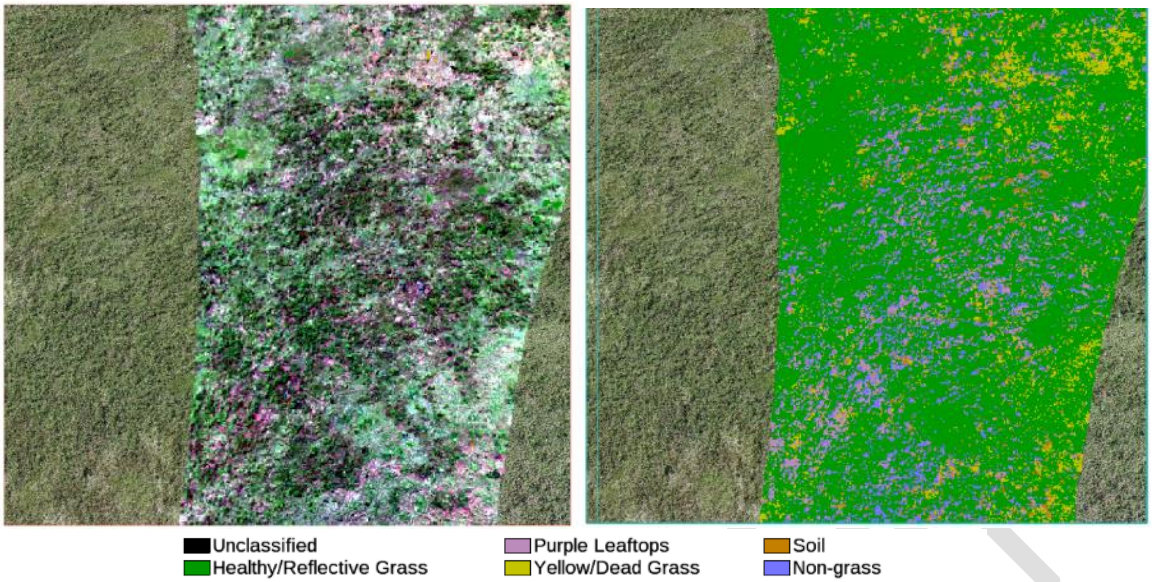
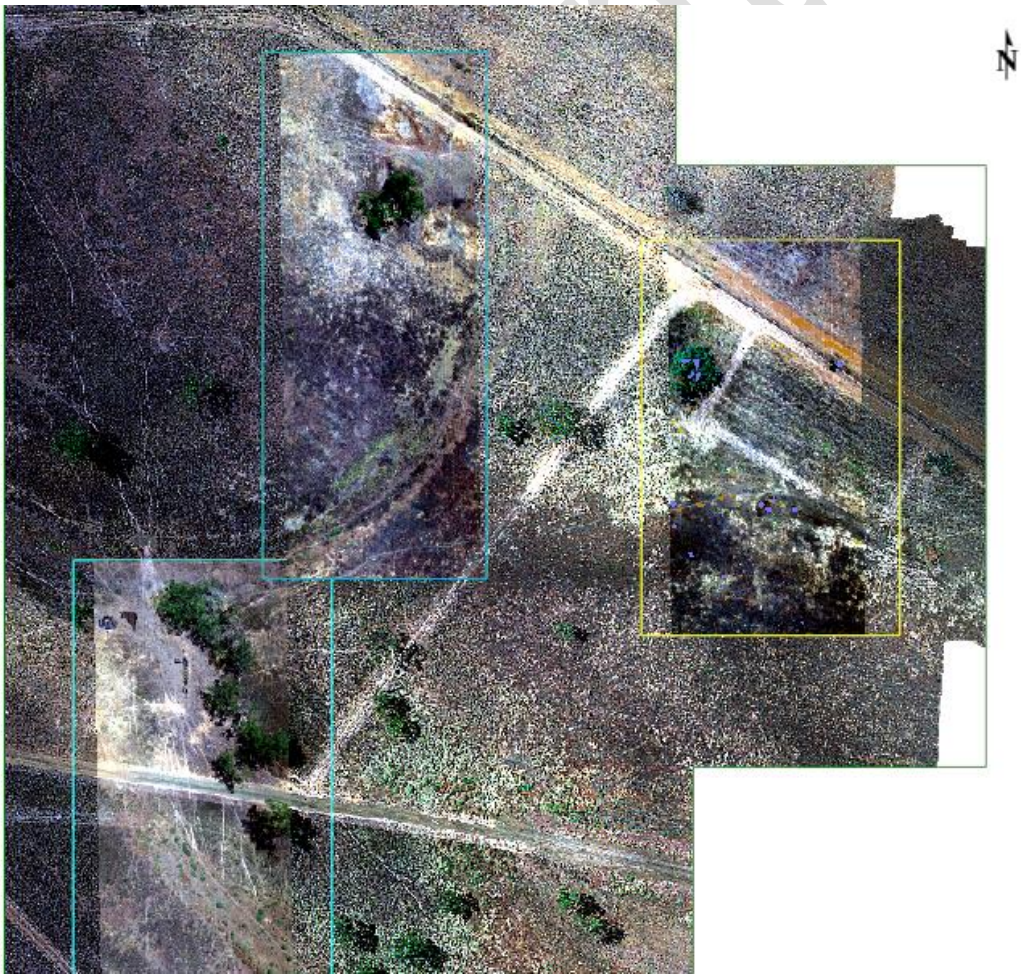


Fig. 20. Predicted classification map of Nobbys Creek, QLD (2021-04-21)



(a)

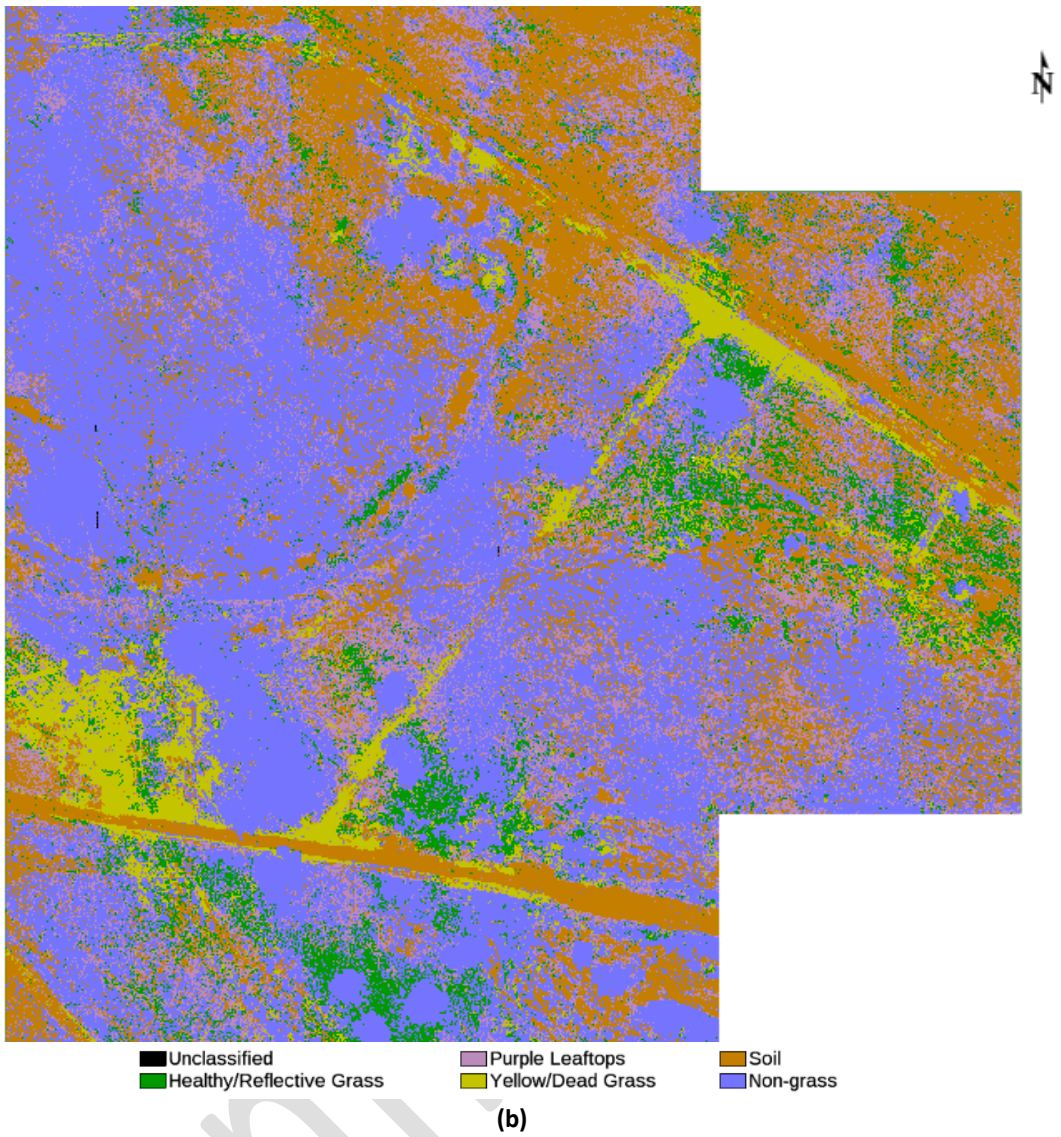


Fig. 17. Banana Station classification map. (a) Hyperspectral samples overlayed at high-resolution RGB mosaic (b) Predicted classes.

The generated classification maps illustrated not only affected regions by PD, but also land occupation by non-grass species in former grasslands. For example, recent in-site inspections by biological experts at Banana Station (Fig. 17) on early May 2021 have reported plant invasions from *Parthenium hysterophorus* sp. and *Spinifex* sp.

4.3.1 Image Registration to Satellite Data

A second model tuning has been proposed to compare PD detection between UAV and available Sentinel 2 Satellite data using multispectral imagery. The input data for this model are the corresponding vegetation indices that can be calculated from the Red, Green, Blue and NIR bands. The "ground truth" generated for hyperspectral data was reused for the second model as the datasets are georeferenced. The classification report of the fit model on multispectral data is shown in Table 6.

Table 6. Classification report on PD classes on multispectral data.

Class	Precision	Recall	F1-score	Support
Healthy/Reflective Grass	99%	99%	99%	3160
Purple Leaftops	93%	89%	91%	830
Yellow/Dead Grass	96%	96%	96%	1345
Non-grass	98%	99%	99%	7858
Soil	100%	98%	99%	382
Accuracy			98%	13575
Macro average	97%	96%	97%	13575
Weighted average	98%	98%	98%	13575

The overall accuracy for the multispectral classifier is 98% from a total of 13575 labelled samples. Despite achieving lower precision and recall scores for Purple Leaftops of 93% and 89% respectively, these initial results are encouraging considering the decrease in spectral data compared to hyperspectral imagery. Nevertheless, further is required to observe consistency on these findings.

The trained XGBoost classifier also outputs a ranking of relevant features for the classifier for PD and a visualisation of the predicted map, which are shown below. The GNDVI, SIPI MSAVI and EVI vegetation indices reported the highest relevance for the model to distinguish the defined PD classes.

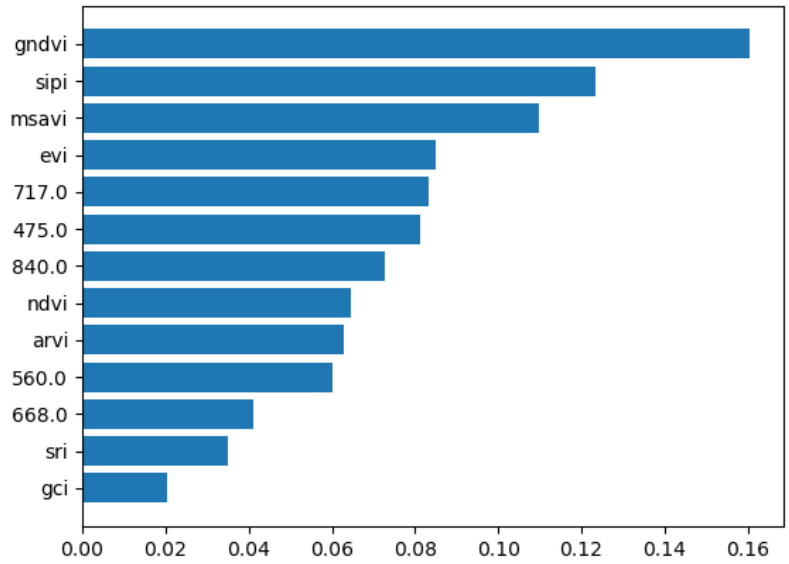


Fig. 18. Ranking of relevant data to find PD correlations on multispectral data.

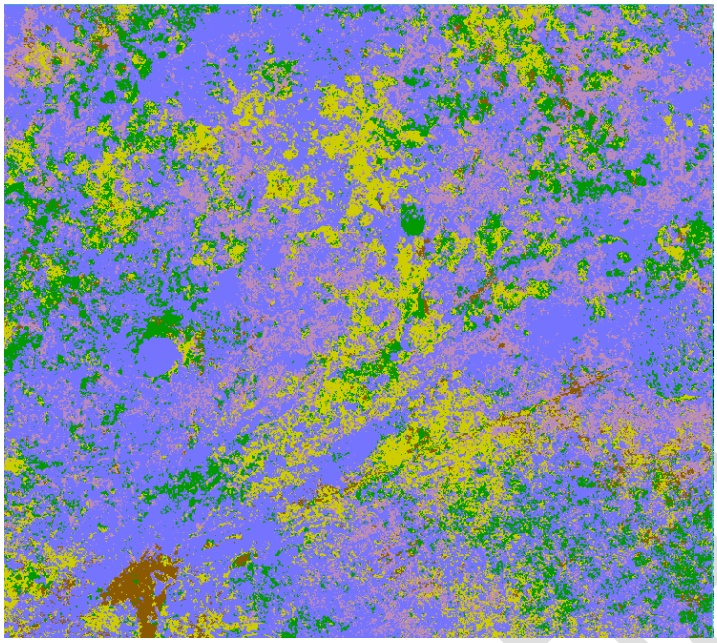


Fig. 193. Predicted map at the studied trial in Maudsland, QLD using multispectral data.

4.4 Association of climatic conditions with PD occurrence

4.4.1 Climate anomalies at the PD sites

Analysis of the climate anomalies of the PD sites reveals higher monthly average maximum and minimum temperature, but lower monthly average solar radiation for the period of 2015-2020 as compared to the corresponding 30-year long-term averages (Fig. 24). Particularly, higher temperature, lower radiation and rainfall are observed for the winter months at the PD sites. However, despite the clear climatic trends at the PD sites, only the radiation in June is statistically lower than its long-term mean.

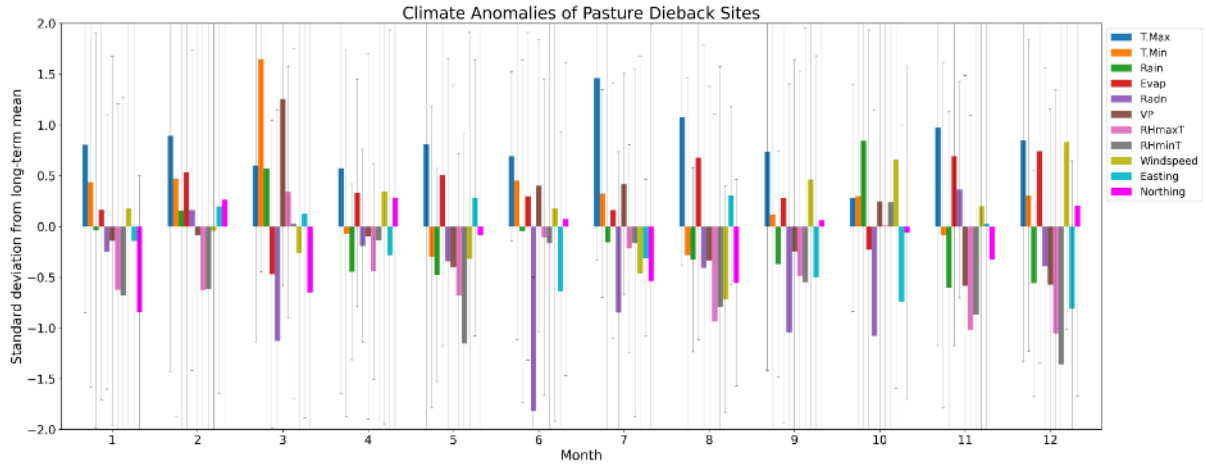


Fig. 24. Climate anomalies for the period of 2015 – 2020 (relative to 30-year-average data from 1980-2010)

4.4.2 PD Occurrence analysis

The probability model (model 2) fitted with observed climatic variables showed the highest probability (56%) of pasture dieback occurrence (Fig. 25) observed at an average monthly maximum temperature between 15 and 20 ^{o}C in combination with a monthly average rainfall between 8 to 10 mm. Increasing temperature drastically decreases the probability (i.e., lower the chance) of PD occurrence. The lower probability in PD occurrence was also observed with increases in rainfall, but to a lesser extent as compared to changes in temperature.

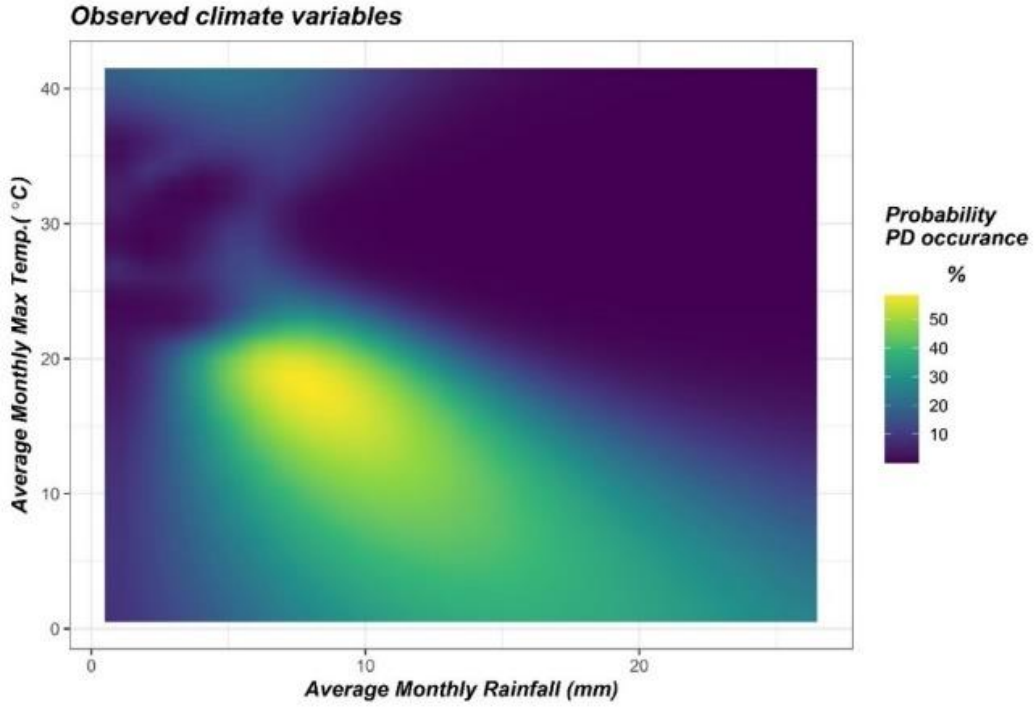


Fig. 25. Effect of average monthly rain and max temperature on PD occurrence.

The probability model was also fitted using the climate anomaly variables for the period of 2015-2020. While our analysis shows increases in temperature at the PD sites (positive anomalies) (Fig. 24), the highest probability of PD occurrence was only observed for the months with temperature anomaly below -0.5 and rainfall anomaly below -0.25 (Fig. 26).

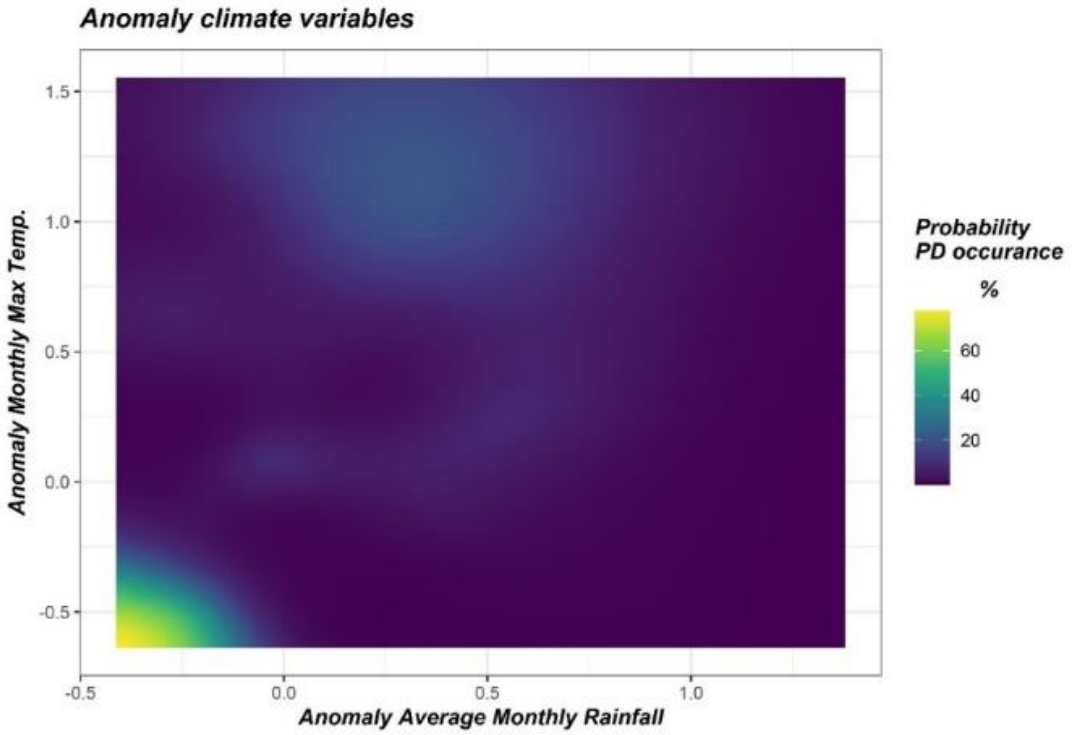


Fig. 26. Effect of long-term rain and max temperature anomalies on PD occurrence.

The analysis of the factors influencing the extend of PD showed that rainfall and maximum temperature below long-term averages tend to increase the PD infested extend (Fig. 27). However, at high max temperature and high rainfall anomalies, the land area affected by PD is limited. This is highly in agreement with our probability model.

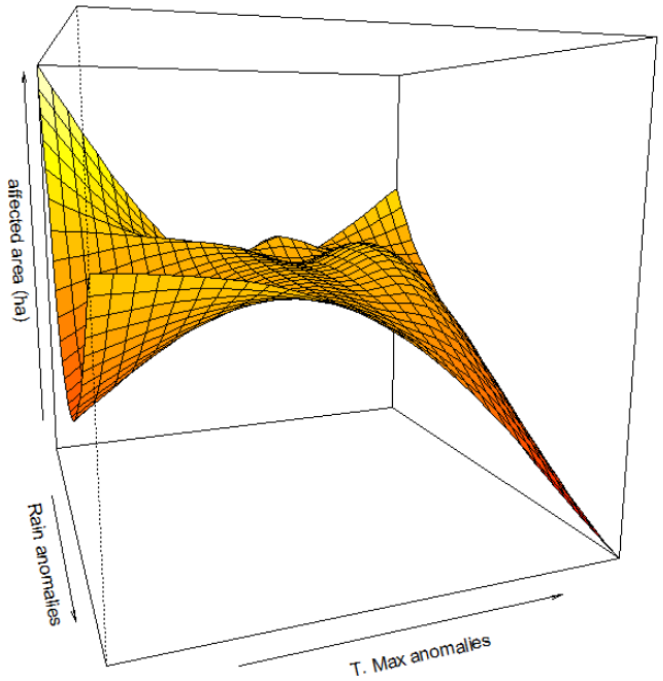


Fig. 27. Effect of rain and max temperature anomalies on the extension of PD affected area (ha).

The results suggest the existence of a particular climatic threshold that favours PD development. These results might help in identifying causal agents responsible for the severe outbreaks observed in the past years.

4.5 Prediction of potential pasture growth

A random forest machine learning model was created to predict potential vegetation growth (max EVI) of a summer growing season based on the season's weather condition and historical vegetation growth (model 3). The random forest model shows good performance with the accuracy of 93%, root mean squared errors of 0.04, and coefficient of determination (R^2) of 0.89 (Fig. 28). The top three most important predictors of the machine learning model are max EVI of the two previous summers (EVI_1 and EVI_2), and total rainfall in January (Fig. 29). Partial independent plots showing how maximum EVI varies with variations in these predictors are reported in Appendix 8.4 (Fig. 37). This model can be applied on a landscape to help identify areas with large differences between the potential and the actual vegetation growth. These differences signal disturbances (e.g., changes in land management or pest infestation) to the landscape rather than just unfavourable weather conditions.

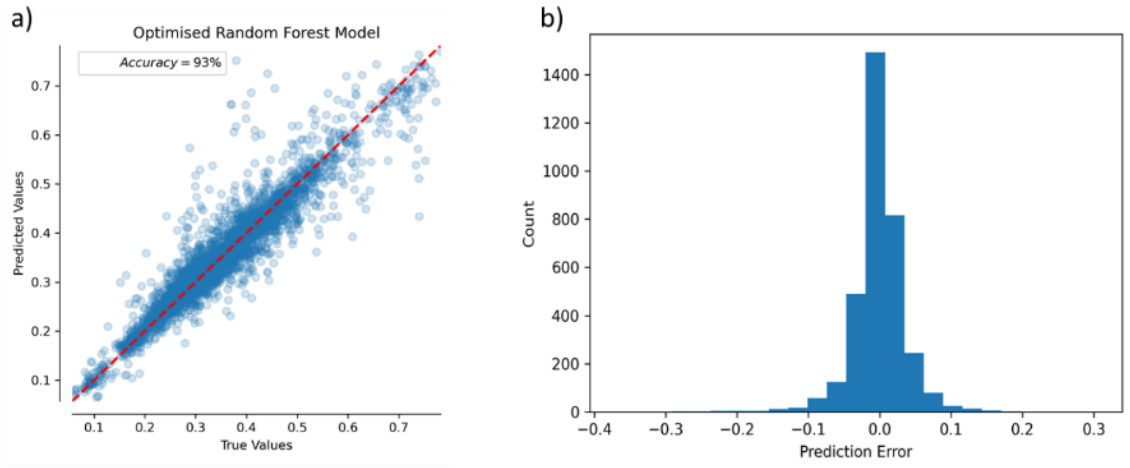


Fig. 28. Performance of random forest model predicting potential summer vegetation growth. a) Predicted vs. observed max EVI values; b) Prediction error of max EVI

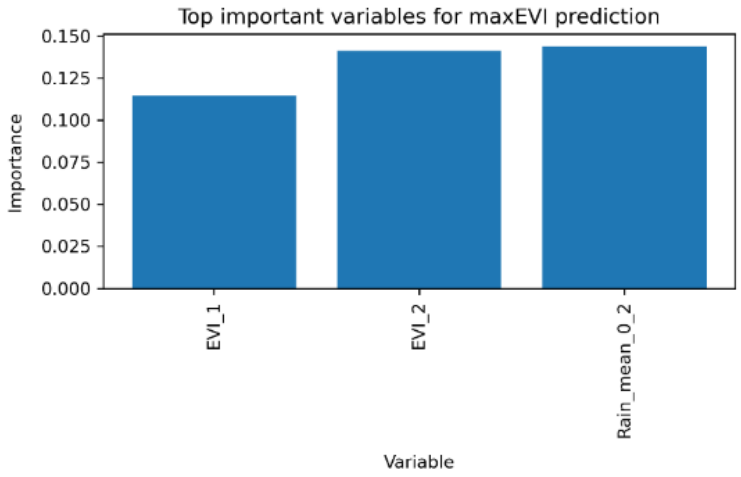


Fig. 29. Top three most important variable of the random forest model predicting potential summer vegetation growth.

4.6 Demonstration of large-scale applications of PD detection and prediction algorithms

The following section demonstrates how our findings and algorithms can be applied to larger landscapes to help identify and predict PD.

4.6.1 PD risk analysis for Queensland pasture

We applied our PD occurrence probability model (model 2) to the whole Queensland state (excluding desert region) to identify areas and period of the year that are more susceptible to PD outbreaks. The model trained with observed monthly average maximum temp and rainfall was applied to BOM's

gridded climate data for 2017 and 2020 (Fig. 30). The results show that during May 2017 and 2020 coastal areas from Rockhampton to Townsville had the highest probability of PD occurrence. These results align well with the recent observed PD outbreaks.

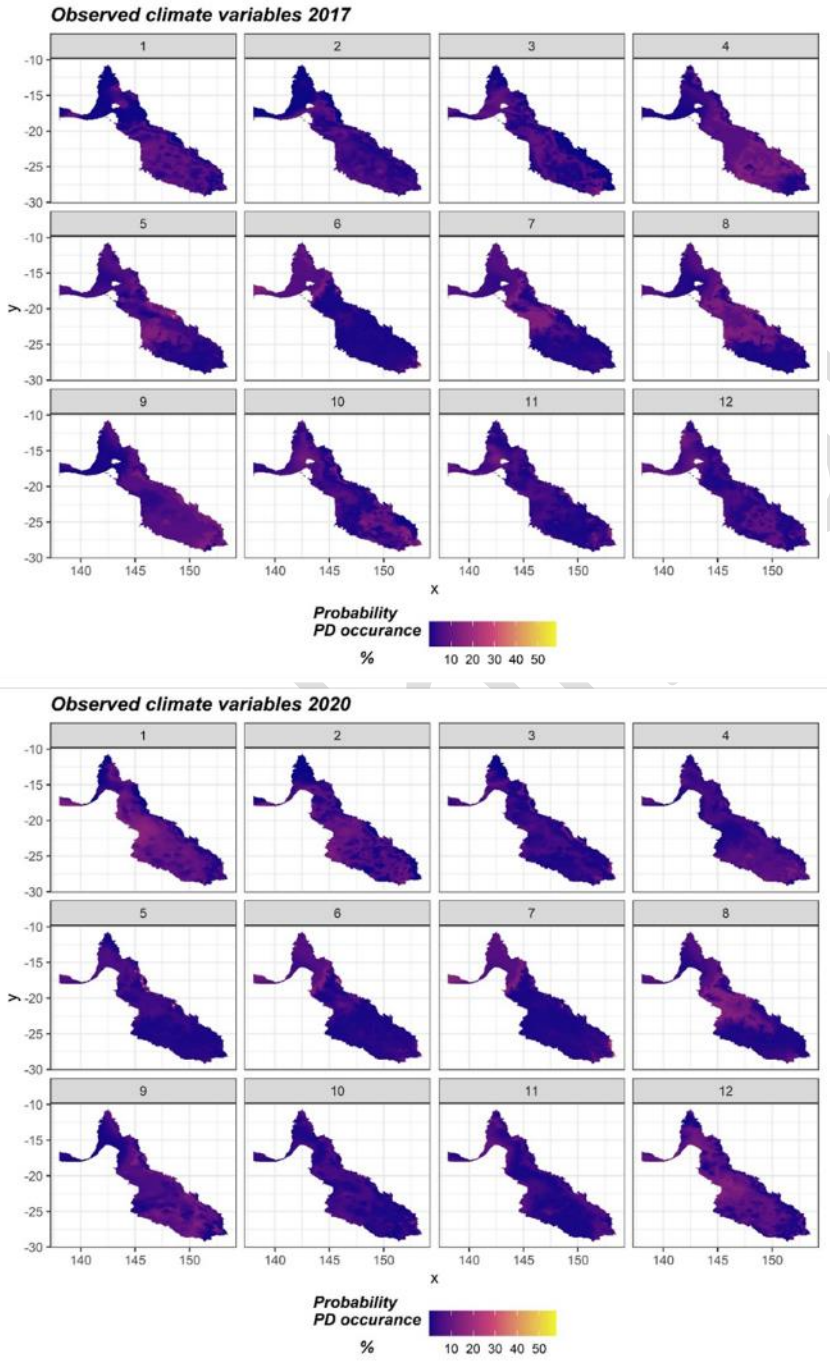


Fig. 30. Mapping the probability of PD for Queensland's pastures in (a) 2017 and (b) 2020

4.6.2 Identification of potential unhealthy pasture growth

Our potential pasture growth model (model 3) was applied to a study area of 2,500 ha in Banana, QLD around the Dawson Mine where we have PD ground truth data (Fig. 31a). The model computed the difference between the actual pasture growth and the potential growth based on the weather conditions in summer 2018 using on a cloud-free Landsat 7 imagery composite from January 2018 to April 2018 (Fig. 31b). The results show a strong agreement between the PD ground truth locations (green pins) and the areas of abnormally low pasture growth (dark blue).

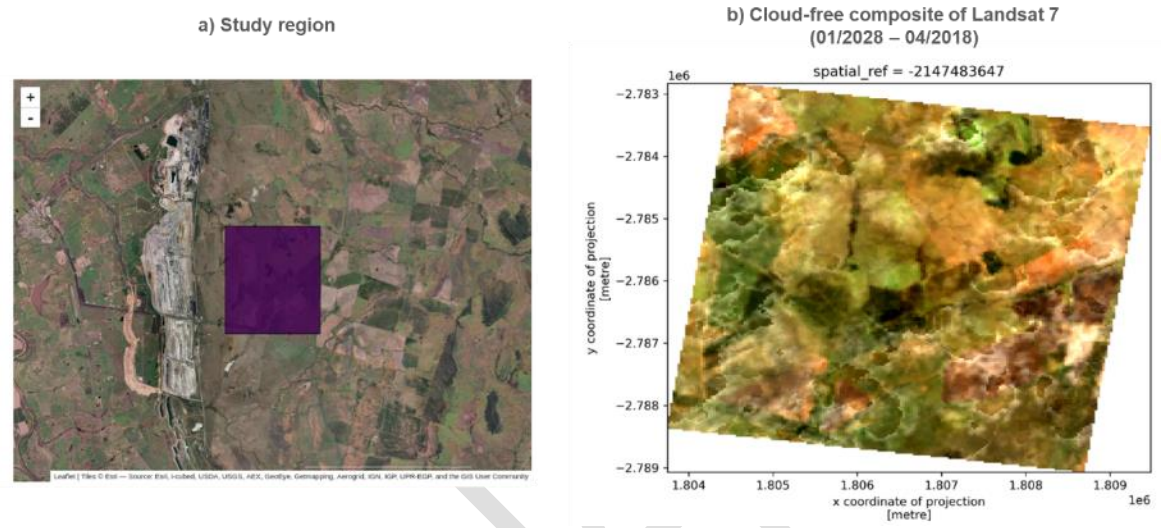


Fig. 31. a) Study site for Random Forest model application, b) Cloud-free composite of max EVI for the first 4 months in 2018.

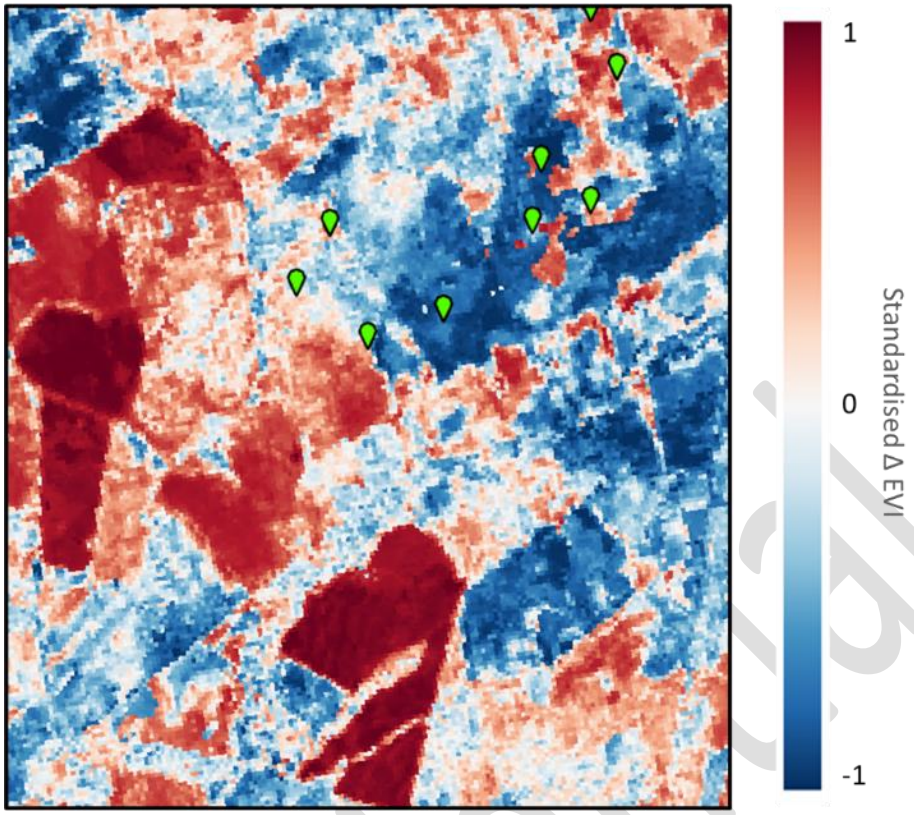


Fig. 32. Mapping of abnormal pasture growth unrelated to weather conditions. The green dots are reported PD sites in 2017. The dark red areas are cultivated crops that were barren over the last 2 summers. Dark blue areas are regions supposed to be highly productive based on the weather conditions and the past performance. However, actual pasture growth reveals abnormal low biomass productivity.

4.6.3 Classification of unhealthy pasture

We conducted a demonstration that extrapolates the trained classifiers from UAV data into Sentinel datasets. The following layers were used as inputs for the trained model: NDVI, MSAVI, EVI, GNDVI, SRI, ARVI, SIPI and GCI spectral indices. For this case study, raw reflectance bands were not included (Red, Green, Blue, NIR) as there were mismatches in the reported intensity on both datasets. The predicted classification map of the Banana region in 2017 is depicted below.

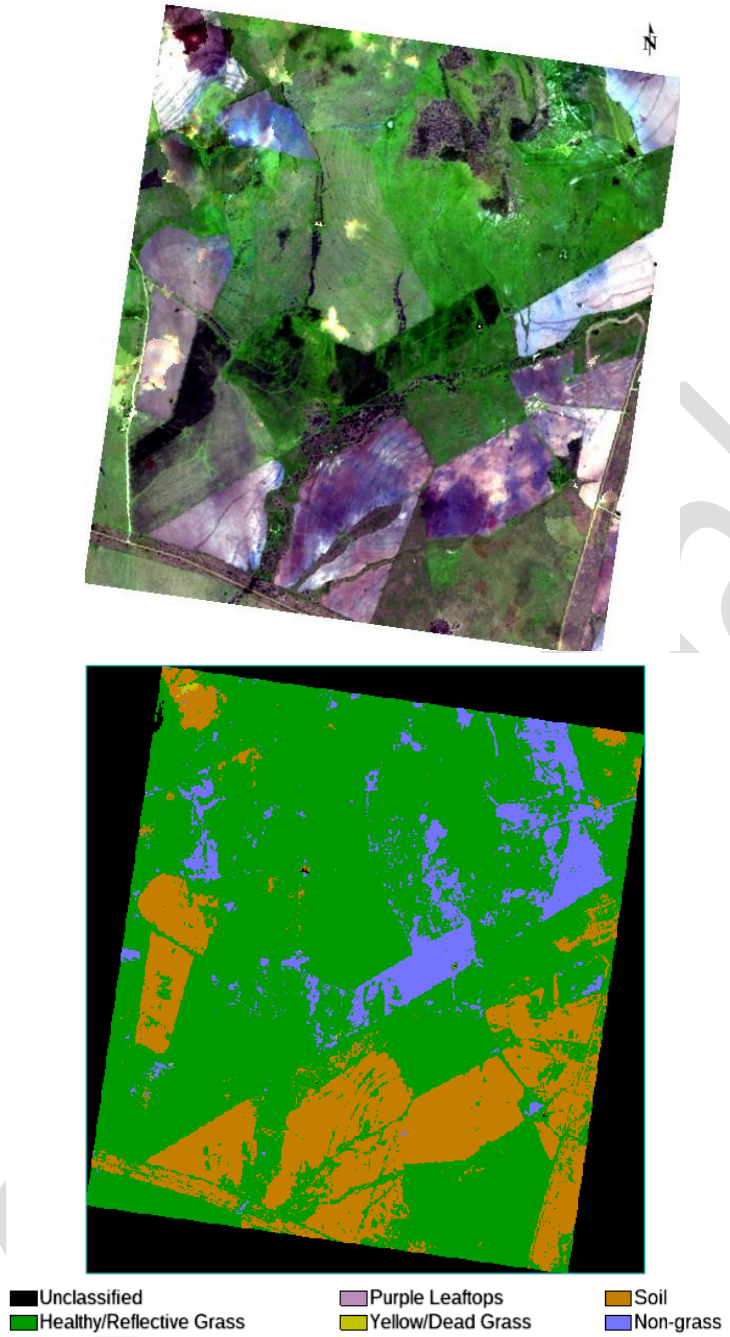


Fig. 33. Classification predictions of tuned hyperspectral data using the coarser spatial resolution Sentinel 2 data.

The current model was able to successfully distinguish between bare soil, grasslands, and other regions likely to grow other vegetation apart from grass. However, the current model was unable to discriminate between healthy grass and reported areas with dieback. Further research is necessary to expand on algorithms to correct reflectance differences between UAV and satellite data to upscale trained models for different sites and grass species.

4.7 Stakeholder engagement and benefits to industry

Several meetings have been organised with different stakeholders to discuss methodology improvement, data collection, and options for result presentation. These stakeholders include different research groups from QUT and UQ who are working on PD, staff from QDAF and AgForce, and local farmers. The majority of the stakeholders were keen to see broader applications of the resulting models for PD warning and detection. However, they emphasised the need for an on-demand and easy-to-setup platforms/applications.

Successful implementation data platforms/applications will provide landholders with information on the spatial extent of pasture dieback on their properties without having to rely on visual assessments over large tracts of land. It will also provide graziers with information on the climatic conditions that contribute to the development of PD so that they can plan ahead to mitigate the likely impacts of dieback on livestock management and grazing. Based on these discussions, we created different web applications to facilitate the employment of our models for analysis on different regions with no complicated setup. These pilot applications are by no mean in their mature form. They only serve as a proof-of-concept and thus require future development and maintenance. The list of the web applications is provided below:

a. Google Earth Engine application for continuous change detection and classification (time-series analysis – model 3) (Fig. 34): https://code.earthengine.google.com/b6c5288a7cbabde35cba1259f808ffe6

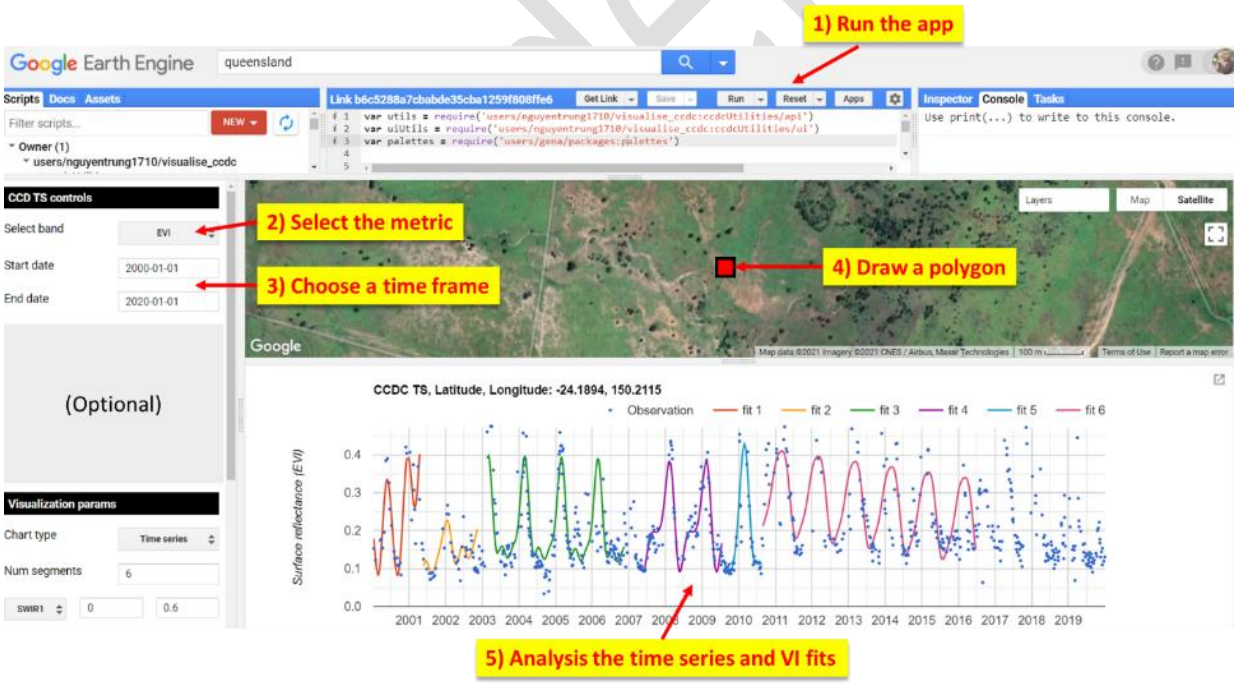


Fig. 34. Pilot web application for time series analysis of pasture growth.

b. Classification of unhealthy/dead grass (model 1): https://colab.research.google.com/drive/1T8PI4_nSF5EmKjhTHb2KXNQHz2d0mVal

c. Prediction of potential biomass growth and identify pasture disturbance (model 3): https://colab.research.google.com/drive/1PUQ-PXyfwTKCg7qMQ8fcNdfS-pIUtQ0I

d. Acquisition of soil, weather, elevation, wind speed data: https://colab.research.google.com/drive/1zgZ_PjCPxops-7xUSt1ww8HwlxgfhkxQ

5. Conclusion

The keys findings of the research on spatial-temporal prediction of pasture dieback using UAVs and remote sensing:

- The NDVI, GNDI, PRI, GCI and CRI1 spectral indices indicated the highest correlation to track PD symptomatology on the surveyed sites using UAV data. The CRI2 and WBI spectral indices particularly displayed higher values to track early stages of PD (purpling leaf tops). Important spectral wavelengths that could expand PD detection from Sentinel 2 imagery are bands at 510nm, 531 nm, 570 nm, 700nm, 900nm and 970 nm.
- There were differences in the PD spectral signature for different sampling locations. In coastal locations, a substantial decrease in the visible green (560 nm), red edge (717 nm) and near infrared (840, 930 nm) bands, and an increase in the visible red (668 nm) bands were found between healthy and early stages of PD. Yellow and dead grass followed a different trend with reports of higher reflectance values in the near infrared range. In surveyed locations at Central QLD (Biggenden and Banana Station), the reflectance values of yellow and dead grass surpassed the recorded values of healthy grass in all the spectral domain. Intra-class variations between the 930 nm and the 1000 nm range were lower at Central QLD locations compared to coastal sites in SE QLD and Northern NSW.
- Due to the highly variable spectral signature at each ground truth site, tuning a generalised PD classification model from UAV data that can be applied on sentinel 2 imagery was infeasible at each UAV surveyed site, which contained a different dominating grass species, displayed different colourisation changes (symptomatology) and spectral signatures for PD.
- Time series analysis of temporal and spatial anomalies of vegetation indices is useful to help identify unhealthy or dead pasture. However, given the coarse resolution of satellite imagery and the patchy nature of PD, it is not possible to detect PD occurrence and its early-stage dynamics using Sentinel 2 and Landsat.
- Analysis on the correlation of PD onset and climate reveals that a combination of average monthly maximum temperature between 15 and 20 ^{o}C and monthly average rainfall between 8 to 10 mm results in the highest chance of pasture dieback occurrence.
- Machine learning approach can be used to predict potential vegetation growth for a growing season and identify disturbances (e.g., changes in land management or pest infestation) to the landscape rather than just unfavourable weather conditions.

- More ground truth data are needed to improve the performance of the statistical and machine learning models created in this project.

- Application of the resulting models at the landscape and regional scales are of interest to stakeholders. Such applications are required to be on-demand and need not involve complicated setup on the user's end.

6. Future research and recommendations

Since PD can develop through many stages, the identification of PD stages for UAV surveys as well as timely UAV flights are important to improve the UAV classification models (model 1). Repeated flights are also required to cover various stages of PD development. Besides, awareness of the dominant grass species prior to UAV or satellite image analysis is recommended as it helps ground and aerial teams recognise the grass symptomatology, select a corresponding classification model for mapping PD and label new data. Due to the nature of how purpling leaftops appear on pasture species, obtaining precise coordinates (with GNSS RTK-based devices) of small grass regions where dieback is visualised is essential as such areas are likely to be projected as single or several pixels in multispectral and hyperspectral images. Any additional information from the ground that provides context of the location and nearby vegetation will ease the labelling process. For future research, better coordination between the airborne sensing team, biological scientists, and pasture specialists is recommended. This ensures the availability of all data in developing algorithms of PD presence, including more timely and optimised flights and the more accurate labelling and validation of collected UAV imagery.

The results of the PD occurrence probability/risk are highly affected by the uncertainty related to the estimation of PD occurrence, the accuracy of interpolated gridded weather data, as well as the intrinsic uncertainty associated with the model predictions. Furthermore, our time series analysis suggested that the variation of spatial or temporal patterns around the reported dates of PD in several PD ground truth sites indicates great uncertainties in PD identification (e.g., inaccurate coordinates or confusion of different conditions for PD). Compared with other studies that used a similar approach to investigate the occurrence or rare events, the entire dataset used in this study was relatively small. Therefore, more ground truth data with better identification of PD and estimation of PD occurrence date are required to improve the probability model performance (model 2).

Satellite data are often contaminated by clouds. While we have used the state-of-the-art cloud detection and masking algorithms to process the satellite scenes before analysis, many scenes significantly covered with clouds have been problematic. This affects our time series analysis as well as the accuracy of the model predicting the potential EVI (model 3). To create a reliable time series for our analysis, we used a combination of Landsat 5, 7, 8 and Sentinel 2 imagery to reduce the data gap. Data from these sources were only aligned geographically (i.e., based on geo-location) in this project. However, given the differences in the satellite sensors and the spectral ranges that these sensors cover, spectral correction and alignment might need to be carried out before conducting analysis to ensure fair comparisons. Alternatively, we can employ derived metrics from the raw satellite values as indicators for detecting abnormal pasture growth in replacement of using the raw data. Examples of these metrics include phenology, length of growing season, percent difference between maximum and minimum vegetation growth.

The Random Forest model for potential growth (model 3) was only trained using Landsat 8 data based on the historical growth of the previous growing seasons. As a result, the model performance is highly sensitive to the pixel quality. Missing data (e.g., due to clouds) in one season could result in inaccurate estimation of potential growth for the predicting season. Future research should focus on training a machine learning model that can use a combination of Landsat 7, 8 and Sentinel 2 imagery as inputs to reduce the uncertainty associated with missing data.

Our findings and models are only practical if they can be employed by the landowners and other stakeholders. The proof-of-concept applications presented in section 4.8 demonstrate examples of how our models can be deployed at different levels from no setup and no maintenance platforms for simple simulations to supported platforms that run on high-performance computer clusters for large-scale and more frequent analyses. It is highly recommended that such implementations are prioritised in future projects to bring our research outcomes to fruition.

7. References

Chen, T., & Guestrin, C. (2016). XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, 785–794. https://doi.org/10.1145/2939672.2939785

Dash, J.P., Pearse, G.D., Watt, M.S., 2018. UAV Multispectral Imagery Can Complement Satellite Data for Monitoring Forest Health. Remote Sens. 10, 1216. https://doi.org/10.3390/rs10081216

Dhu, T., Giuliani, G., Juárez, J., Kavvada, A., Killough, B., Merodio, P., Minchin, S., Ramage, S., 2019. National Open Data Cubes and Their Contribution to Country-Level Development Policies and Practices. Data 4, 144. https://doi.org/10.3390/data4040144

Marshall, V.M., Lewis, M.M., Ostendorf, B., 2014. Detecting new Buffel grass infestations in Australian arid lands: evaluation of methods using high-resolution multispectral imagery and aerial photography. Environ. Monit. Assess. 186, 1689–1703. https://doi.org/10.1007/s10661-013-3486-7

Sandino, J., Gonzalez, F., Mengersen, K., Gaston, K.J., 2018. UAVs and Machine Learning Revolutionising Invasive Grass and Vegetation Surveys in Remote Arid Lands. Sensors 18, 605. https://doi.org/10.3390/s18020605

8. Appendix

8.1 Ground truth dataset

Latitude	Longitude	Extend (ha)	Date Observed	Species	Site	Status	Estimated PD occurrence Date
-24.4664	150.1185		2015-11	No data	Acton	No data	2015-11
-24.5035	150.26		2017-04	No data	Acton	No data	2017-02
-24.5101	150.2599		2017-04	No data	Acton	No data	2016-01
-24.5101	150.2518		2017-04	No data	Acton	No data	2016-01
-24.4147	150.246		2017-03	No data	Allawah	No data	2017-03
-24.4559	150.2587		2017-03	No data	Allawah	No data	2016-06
-24.4614	150.2565		2017-03	No data	Allawah	No data	2017-03
-24.4663	150.2626		2017-03	No data	Allawah	No data	2017-03
-23.674	149.9528		2017-03	No data	Balcarres	No data	2016-06
-24.5826	150.0643		2017-11	No data	Tremere	No data	2017-01
-24.588	150.0662		2017-11	No data	Tremere	No data	2016-12
-24.5879	150.0691		2017-11	No data	Tremere	No data	2017-03
-24.5818	150.0681		2017-11	No data	Tremere	No data	2017-03
-24.5601	150.0782		2017-11	No data	Tremere	No data	2017-05
-24.5535	150.0893		2017-11	No data	Tremere	No data	2017-05
-24.5582	150.09		2017-11	No data	Tremere	No data	2017-05
-24.5629	150.0945		2017-11	No data	Tremere	No data	2017-05
-24.5672	150.0962		2017-11	No data	Tremere	No data	2017-05
-24.5705	150.0976		2017-11	No data	Tremere	No data	2017-05
-24.5729	150.0888		2017-11	No data	Tremere	No data	2017-05
-24.5716	150.0869		2017-11	No data	Tremere	No data	2017-05
-24.5893	150.0974		2017-11	No data	Tremere	No data	2017-05
-24.593	150.0951		2017-11	No data	Tremere	No data	2017-05
-24.5963	150.1		2017-11	No data	Tremere	No data	2017-05
-24.5946	150.1052		2017-11	No data	Tremere	No data	2017-05
-24.5853	150.1119		2016-11	No data	Namgoori	No data	2016-04
-24.5891	150.1113		2016-11	No data	Namgoori	No data	2016-04
-24.5879	150.1153		2016-11	No data	Namgoori	No data	2016-04
-24.5759	150.1153		2016-11	No data	Namgoori	No data	2016-04
-24.5795	150.1171		2016-11	No data	Namgoori	No data	2016-04
-24.1086	150.4342		2017-04	No data	Maynard	No data	2017-04
-24.1054	150.4374		2017-04	No data	Maynard	No data	2017-04
-24.052	150.4019		2017-06	No data	Maynard	No data	2017-04
-24.0622	150.4154		2017-06	No data	Maynard	No data	2017-04
-24.0664	150.4205		2017-06	No data	Maynard	No data	2017-04
-24.0617	150.4416		2017-06	No data	Maynard	No data	2017-04
-24.1112	150.4899		2017-11	No data	Maynard	No data	2017-04
-24.1238	150.4888		2017-11	No data	Maynard	No data	2017-04

-24.1282	150.4914		2017-11	No data	Maynard	No data	2017-04
-24.1309	150.4803		2017-11	No data	Maynard	No data	2017-04
-24.1296	150.4656		2017-11	No data	Maynard	No data	2017-04
-24.1401	150.4619		2017-11	No data	Maynard	No data	2017-04
-24.152	150.4857		2017-11	No data	Maynard	No data	2017-04
-24.1407	150.4771		2017-11	No data	Maynard	No data	2017-04
-24.0444	151.2142		2017-09	No data	LSBjorn	No data	2016-05
-24.043	151.2155		2016-05	No data	LSBjorn	No data	2016-05
-24.0439	151.2188		2017-09	No data	LSBjorn	No data	2016-05
-24.0574	151.2267		2017-04	No data	LSBjorn	No data	2016-05
-24.0588	151.2268		2017-04	No data	LSBjorn	No data	2016-05
-24.867	150.1452		2016-10	No data	Loma	No data	2016-03
-24.8617	150.1511		2016-03	No data	Loma	No data	2016-03
-24.8671	150.154		2016-03	No data	Loma	No data	2016-03
-24.8472	150.1613		2016-11	No data	Loma	No data	2016-03
-24.8528	150.1743		2017-01	No data	Loma	No data	2016-03
-24.8585	150.1759		2017-01	No data	Loma	No data	2016-03
-24.8636	150.174		2017-01	No data	Loma	No data	2016-03
-24.6174	150.7		2016-12	No data	Fieldview	No data	2016-08
-24.6287	150.7019		2016-08	No data	Fieldview	No data	2016-08
-24.1911	150.4498		2017-03	No data	Windswept	No data	2016-04
-24.1897	150.4573		2017-03	No data	Windswept	No data	2016-04
-24.1889	150.4584		2017-03	No data	Windswept	No data	2016-04
-24.1905	150.4559		2017-03	No data	Windswept	No data	2016-04
-24.4516	150.6692		2016-05	No data	Glenlivet	No data	2016-04
-24.4566	150.6939		2016-05	No data	Glenlivet	No data	2016-04
-25.3875	151.2221		2016-02	No data	Bullock	Dead	2016-02
-24.9851	151.1725		2016-02	No data	Panicum	Dead	2016-02
-24.4654	149.78		2016-02	No data	Denby - Chalks-	Affected	2016-02
-24.3223	149.9805		2016-02	No data	Baralaba-	Affected	2016-02
-24.8684	151.7596		2016-02	No data	Takilberan-	Dead	2016-02
-27.5603	152.3344		2018-08	No data	Gatton	Affected	2018-08
-27.549	152.3327		2020-09	No data	Gatton	Affected	2020-09
-27.5437	152.3379		2018-03	No data	Gatton	Affected	2018-03
-27.9169	153.2713		2020-07	No data	Maudsland Creek	Recovering	2020-07
-27.9189	153.2737		2020-07	No data	Maudsland Paddock	Recovering	2020-07
-28.2834	153.3347		2020-07	No data	Sutton Park NSW	Recovering	2020-07
-22.1635	148.2281			Buffel	Winchester Downs (Emily's)	Regrown	2015-11
-22.1499	148.2616			Buffel replaced by native	Winchester Downs (Emily's)	Dead	2015-11
-22.1721	148.26			Buffel replaced parthenium weed	Winchester Downs (Emily's)		2015-11

-22.1815	148.2578			Buffel	Winchester Downs (Emily's)	Regrown	2015-11
-22.2051	148.2624			Buffel	Winchester Downs (Emily's)		2015-11
-24.155	150.471			No data	Jambin		2017-06
-24.733	149.997			No data	Theodore		2016-05
-25.595	152.127			No data	Biggenden		2018-03
-24.815	151.6			No data	Gaeta		2016-03
-24.2642	148.6474		2015-12	No data	Lowesby		2015-12
-24.3434	148.6949		2015-12	No data	Lowesby		2015-12
-23.4934	150.0549		2015-12	No data	Weir Park		2015-12
-23.5037	150.0623		2015-12	No data	Weir Park		2015-12
-23.5138	150.0502		2021-01	No data	Weir Park - moving fast		2021-01
-23.5111	150.0347		2021-01	No data	Weir Park - moving fast		2021-01
-27.0369	151.9926		2021-03	No data	Mount Binga		2021-03
-27.0414	151.9707		2021-03	No data	Mount Binga		2021-03
-27.0504	151.9916		2021-03	No data	Mount Binga		2021-03
-20.1602	146.329	2	2012-01	Buffel	Farm Addresses	No data	2012-01
-23.2374	150.2102	150	2014-01	Bluegrass	Farm Addresses	No data	2014-01
-24.145	151.0873	40	2014-02	Bluegrass	Farm Addresses	No data	2014-02
-23.1126	150.7037	20	2014-06	Bluegrass	Farm Addresses	No data	2014-06
-21.2356	149.1869	5	2014-11	Pangola and callide rhodes grass	Farm Addresses	No data	2014-11
-23.0414	150.6629	3	2015-03	Bisset	Farm Addresses	No data	2015-03
-24.6161	150.0641	200	2015-05	Buffel	Farm Addresses	No data	2015-05
-24.527	150.1956	250	2015-12	Buffel	Farm Addresses	No data	2015-12
-24.4088	150.4823	8000	2015-12	All Introduced species of grasses	Farm Addresses	No data	2015-12
-25.4671	152.0778	6	2016-01	Bluegrass	Farm Addresses	No data	2016-01
-27.3527	152.9202	10	2016-01	Kikuyu	Farm Addresses	No data	2016-01
-19.7516	147.4966	100	2016-01	A little in buffel mainly in UROCHLOA grass	Farm Addresses	No data	2016-01
-23.4972	150.2534	100	2016-01	Buffel	Farm Addresses	No data	2016-01
-24.8137	151.6002	100	2016-01	Bissett,Rhodes, giant rat's tail possibly signal (dieback stopped too seen as a result of cool weather to be sure), native bluegrass and black spear,	Farm Addresses	No data	2016-01
-24.167	149.9763	100	2016-01	Buffel	Farm Addresses	No data	2016-01
-24.3634	150.3547	160	2016-01	All 3 varieties of buffel and blue grass	Farm Addresses	No data	2016-01
-23.8374	150.949	12	2016-03	Bluegrass	Farm Addresses	No data	2016-03
-24.9778	152.1045	30	2016-03	Bisset Bluegrass and Pangola	Farm Addresses	No data	2016-03

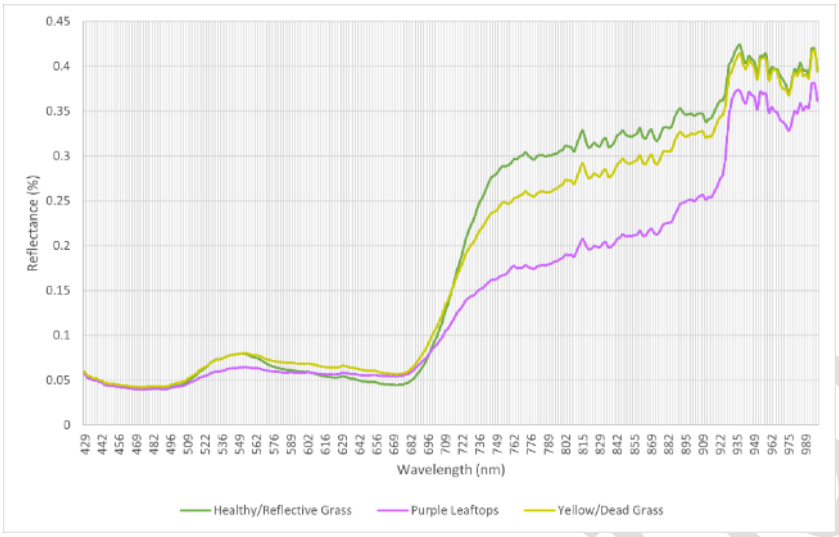
-24.9723	152.1082	100	2016-03	Bisset Bluegrass and Pangola	Farm Addresses	No data	2016-03
-25.4037	151.9705	100	2016-03	Bisset Bluegrass and Pangola	Farm Addresses	No data	2016-03
-24.5297	150.6324	120	2016-03	Nodata	Farm Addresses	No data	2016-03
-22.8136	149.8025	60	2016-04	Buffel	Farm Addresses	No data	2016-04
-25.1058	151.9567	80	2016-04	Bluegrass	Farm Addresses	No data	2016-04
-25.3847	152.0149	150	2016-04	Bluegrass	Farm Addresses	No data	2016-04
-24.5658	150.5844	250	2016-04	Bluegrass and Buffel	Farm Addresses	No data	2016-04
-24.4941	149.5082	400	2016-04	Buffel	Farm Addresses	No data	2016-04
-25.3536	151.9882	50	2016-05	Bluegrass	Farm Addresses	No data	2016-05
-24.0442	151.2156	1	2016-06	Bluegrass	Farm Addresses	No data	2016-06
-24.527	149.4379	100	2016-06	Buffel	Farm Addresses	No data	2016-06
-24.4785	150.1163	1200	2016-06	Buffel	Farm Addresses	No data	2016-06
-23.084	150.4288	10	2016-07	Indian couch, Medway couch	Farm Addresses	No data	2016-07
-24.52	150.1266	50	2016-07	Buffel	Farm Addresses	No data	2016-07
-23.7219	150.6486	100	2016-07	both buffel and bluegrass.	Farm Addresses	No data	2016-07
-24.6224	150.7103	5	2016-08	Buffel	Farm Addresses	No data	2016-08
-24.1102	150.2039	20	2016-08	Buffel	Farm Addresses	No data	2016-08
-23.4946	149.8552	100	2016-08	Plus Other Native species,Buffel	Farm Addresses	No data	2016-08
-24.8189	151.5377	15	2016-09	Roads grass	Farm Addresses	No data	2016-09
-25.2066	151.9396	50	2016-09	Nodata	Farm Addresses	No data	2016-09
-24.1566	149.9521	1040	2016-09	Buffel	Farm Addresses	No data	2016-09
-17.3937	145.6326	9	2016-10	Setaria and Brachiaria Pastures	Farm Addresses	No data	2016-10
-25.2299	152.2496	50	2016-10	Bluegrass , Couch , Kikuyu , Rhodesgrass	Farm Addresses	No data	2016-10
-25.5125	152.0316	20	2016-11	Bluegrass	Farm Addresses	No data	2016-11
-27.7859	152.4345	45	2016-11	Bluegrass	Farm Addresses	No data	2016-11
-23.4922	150.0154	250	2016-11	Buffel	Farm Addresses	No data	2016-11
-24.4083	150.4831	1000	2016-11	buffel and bluegrass	Farm Addresses	No data	2016-11
-24.9593	151.0934	10	2016-12	Rhodes Grass (Katombora)	Farm Addresses	No data	2016-12
-25.0218	152.0512	35	2016-12	Bisset Bluegrass	Farm Addresses	No data	2016-12
-23.6837	149.9305	60	2016-12	USA Buffel,Gayndah Buffel, Green panic, creeping blue grass, Qld blue grass, Urochloa,	Farm Addresses	No data	2016-12
-23.9058	150.1158	400	2016-12	USA Buffel,Gayndah Buffel, Green panic, creeping blue grass, Qld blue grass, Urochloa,	Farm Addresses	No data	2016-12

-24.5628	150.7089	10	2017-01	Panic Buffel	Farm Addresses	No data	2017-01
-26.415	152.8527	10	2017-01	Bluegrass	Farm Addresses	No data	2017-01
-25.2059	152.0129	20	2017-01	Bluegrass	Farm Addresses	No data	2017-01
-26.3142	152.6103	30	2017-01	Bluegrass	Farm Addresses	No data	2017-01
-24.5627	150.709	50	2017-01	Green panic	Farm Addresses	No data	2017-01
-25.5748	152.0018	100	2017-01	Bluegrass	Farm Addresses	No data	2017-01
-27.3342	152.9822	1.2	2017-02	rhodes and native	Farm Addresses	No data	2017-02
-24.5672	149.2977	1	2017-03	Bluegrass	Farm Addresses	No data	2017-03
-27.8191	152.7026	1	2017-03	forest Blue	Farm Addresses	No data	2017-03
-24.9506	152.2167	3	2017-03	Bisset Creeping Blue Grass,Bluegrass	Farm Addresses	No data	2017-03
-22.5386	149.504	5	2017-03	Bluegrass	Farm Addresses	No data	2017-03
-25.1763	151.9553	10	2017-03	Blue grass, Rhodes, Signal	Farm Addresses	No data	2017-03
-24.9703	152.0895	25	2017-03	callide rhodes, green panic, kikuyu,	Farm Addresses	No data	2017-03
-24.5714	151.9068	45	2017-03	Callide rhodes/stylo mixed some black spear and Baihai grass	Farm Addresses	No data	2017-03
-24.5392	150.6275	70	2017-03	Buffel	Farm Addresses	No data	2017-03
-25.1018	149.9422	200	2017-03	Buffel, green panic and Secca stylo legume.	Farm Addresses	No data	2017-03
-24.8458	152.4109	0.5	2017-04	Bluegrass	Farm Addresses	No data	2017-04
-24.1895	150.4462	25	2017-04	Buffel, Gatton panic, Indian Couch and strangely Reeds around a dam	Farm Addresses	No data	2017-04
-26.4423	152.8542	60	2017-04	Native Paspalum, Giant Paspalum, Green panic and Gatton Panic.	Farm Addresses	No data	2017-04
-26.447	152.8643	60	2017-04	Native Paspalum, Giant Paspalum, Green panic and Gatton Panic.	Farm Addresses	No data	2017-04
-24.4548	150.6508	100	2017-04	Buffel	Farm Addresses	No data	2017-04
-24.5346	150.6102	300	2017-04	Buffel	Farm Addresses	No data	2017-04
-23.7146	150.0012	20	2017-05	Buffel	Farm Addresses	No data	2017-05
-25.0109	152.1081	30	2017-05	Bluegrass	Farm Addresses	No data	2017-05
-26.0086	151.8082	100	2017-05	Buffel	Farm Addresses	No data	2017-05
-24.8128	152.1756	1.5	2017-07	Rhodes grass	Farm Addresses	No data	2017-07
-17.4381	145.4161	2	2017-07	Bluegrass	Farm Addresses	No data	2017-07
-24.8578	150.1681	100	2017-07	green panic	Farm Addresses	No data	2017-07
-24.9899	152.1571	12	2017-09	Rhodes	Farm Addresses	No data	2017-09
-26.2559	152.8748	4	2017-10	paspalm	Farm Addresses	No data	2017-10
-25.5116	151.4542	7	2017-10	Bluegrass	Farm Addresses	No data	2017-10

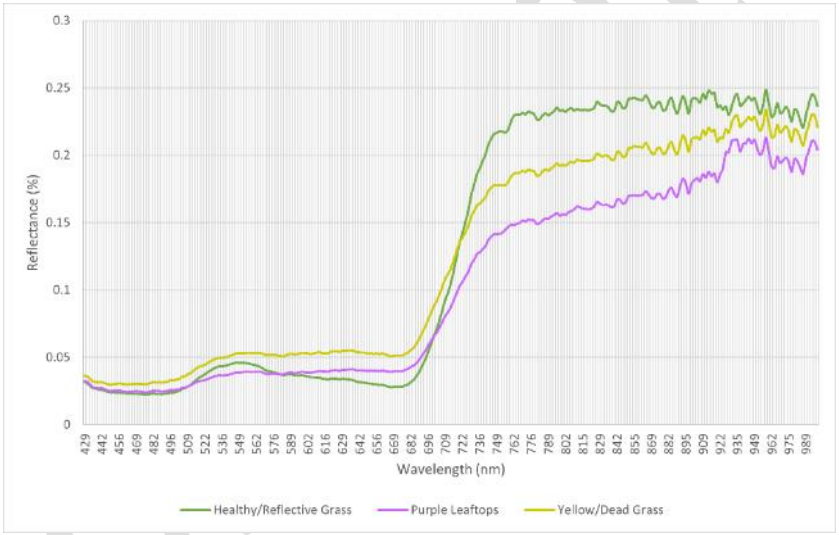
-26.2361	152.859	10	2017-11	Broad leaf paspalum, Signal grass	Farm Addresses	No data	2017-11
-25.2172	152.2685	14	2017-11	Bluegrass & Rhodes Grass	Farm Addresses	No data	2017-11
-26.4512	152.8641	15	2017-11	paspalum	Farm Addresses	No data	2017-11
-25.5064	151.4798	3	2017-12	Bluegrass	Farm Addresses	No data	2017-12
-25.5625	152.5561	5	2017-12	Bluegrass	Farm Addresses	No data	2017-12
-26.2846	152.8669	5	2017-12	kikua couch rhodes	Farm Addresses	No data	2017-12
-25.5697	152.5457	25	2017-12	Bluegrass	Farm Addresses	No data	2017-12
-24.9948	152.1606	30	2017-12	rhodes;Bluegrass	Farm Addresses	No data	2017-12
-25.9742	152.5564	30	2017-12	Rodes grass,	Farm Addresses	No data	2017-12
-25.9245	152.5395	12	2018-01	killed Bluegrass affected Rhodes and Bahia	Farm Addresses	No data	2018-01
-25.1658	152.2112	0.5	2018-02	Bluegrass	Farm Addresses	No data	2018-02
-26.3873	152.7749	1	2018-03	Paspalum	Farm Addresses	No data	2018-03
-24.7592	151.0157	180	2018-03	Gayndah and American Buffel,Green Panic, creeping Blue Grass	Farm Addresses	No data	2018-03
-24.7855	151.0086	180	2018-03	Gayndah and American Buffel,Green Panic, creeping Blue Grass	Farm Addresses	No data	2018-03

8.2 Spectral responses of PD per UAV surveyed site

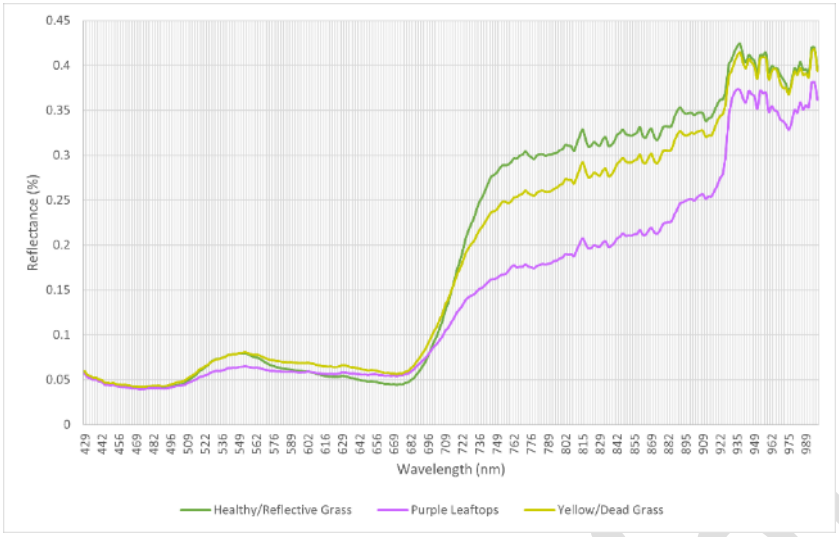
Spectral responses of relevant pasture classes per site are illustrated in Fig..



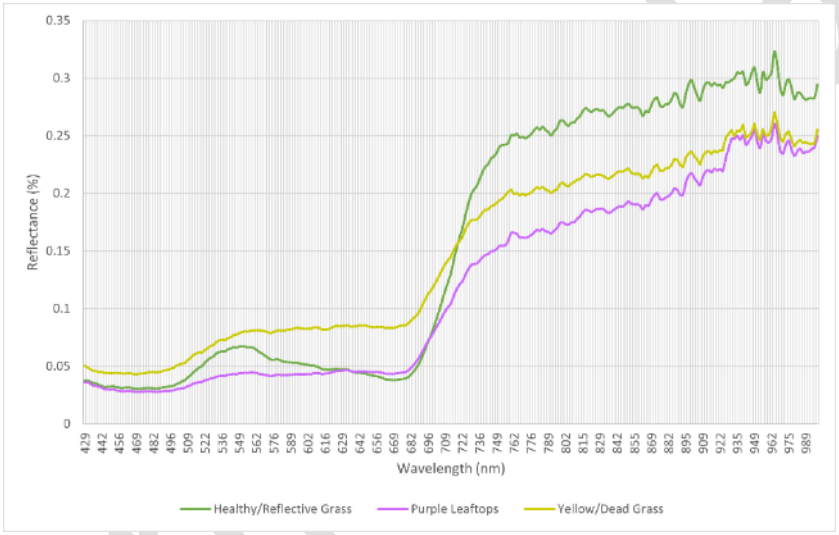
(a)



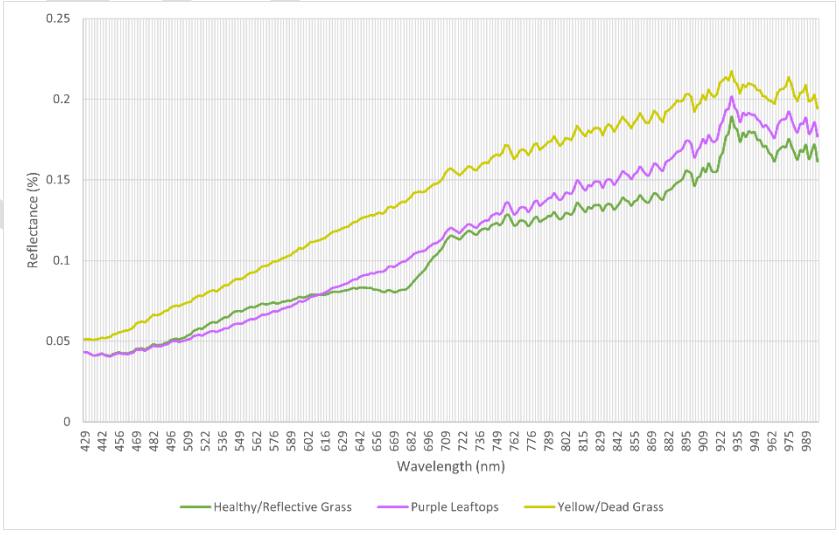
(b)


Reflectance (%)
Wavelength (nm)
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

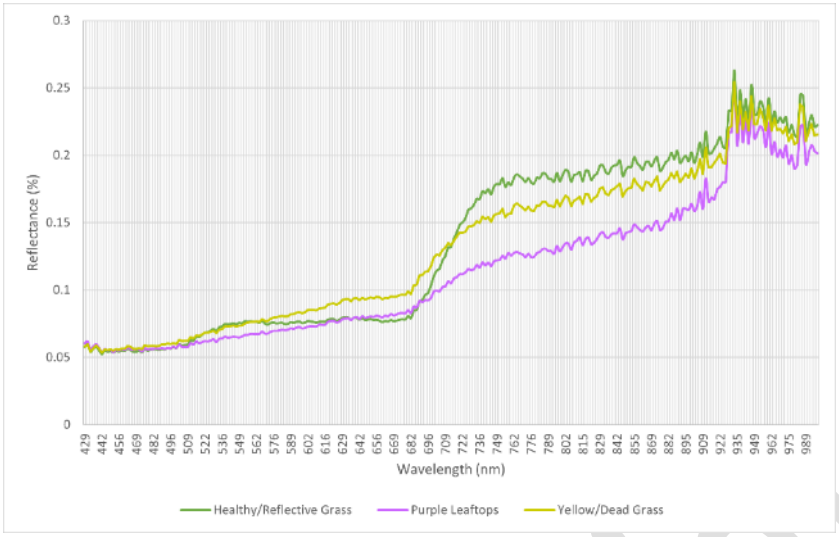
(c)


Reflectance (%)
Wavelength (nm)
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

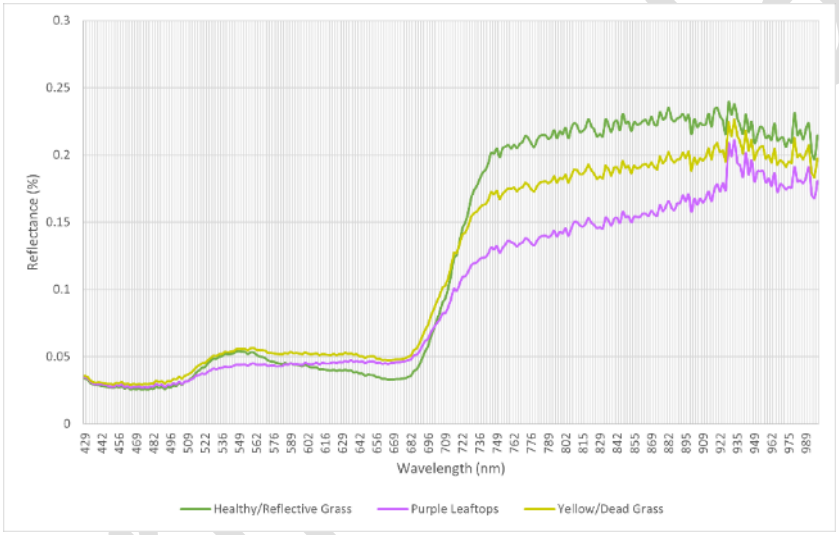
(d)


Reflectance (%)
Wavelength (nm)
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

(e)



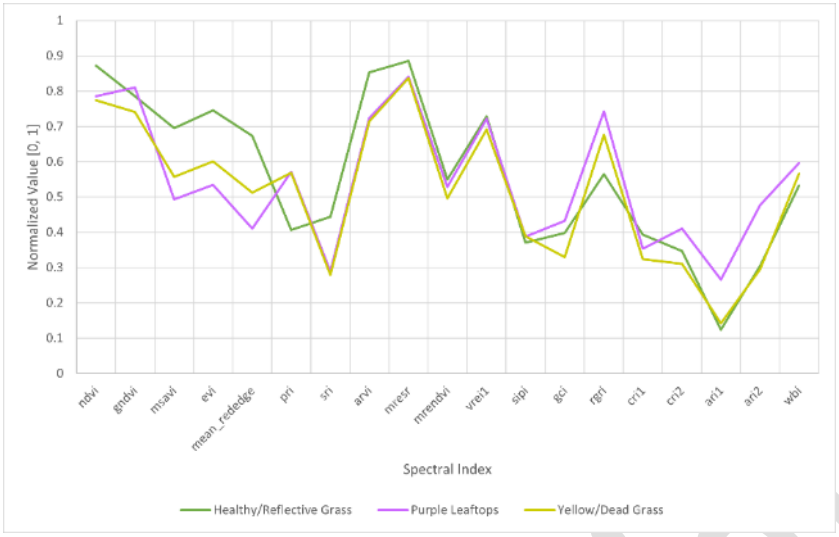
(f)



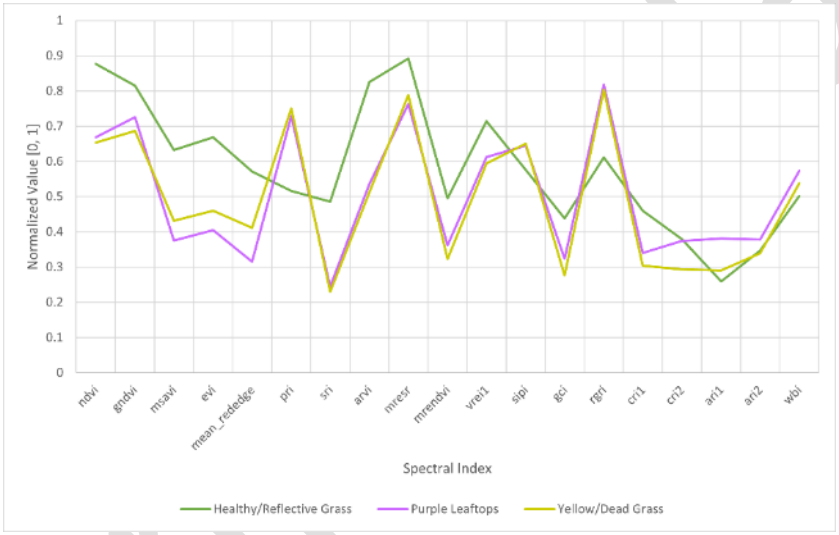
(g)

Fig.35. Individual spectral response of selected pasture classes for PD analysis. **(a)** Maudsland, QLD (2021-02-05). **(b)** Maudsland, QLD (2021-03-09). **(c)** Nobbys Creek, NSW (2021-02-11). **(d)** Nobbys Creek, NSW (2021-04-21). **(e)** Banana Station, QLD. **(f)** Biggenden, QLD. **(g)** Kin Kin, QLD.

Normalised spectral index values of relevant pasture classes per site are illustrated in Fig. 20.


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

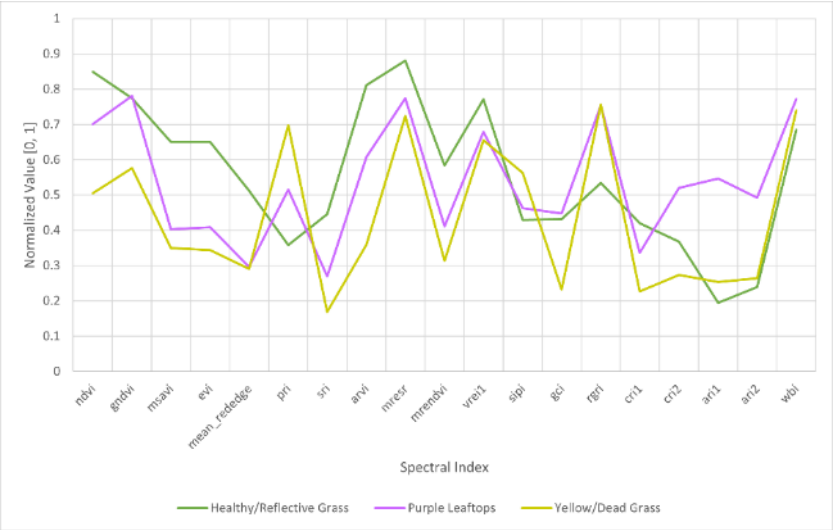
(a)


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

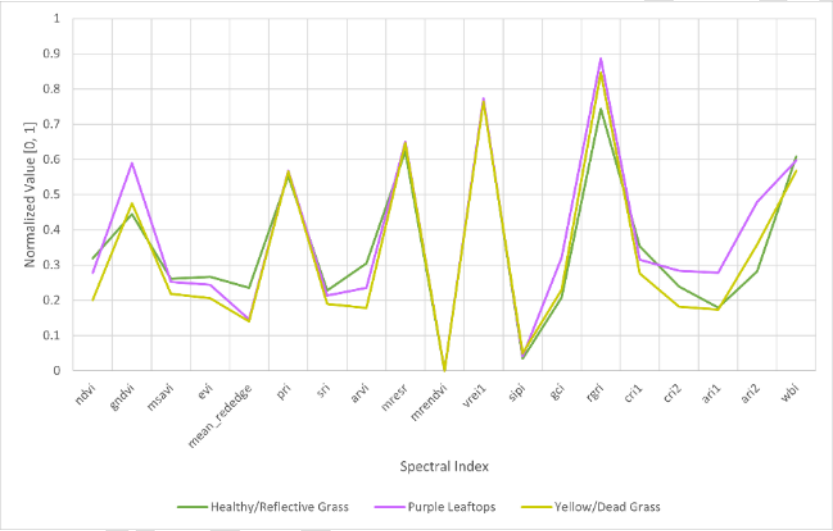
(b)


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

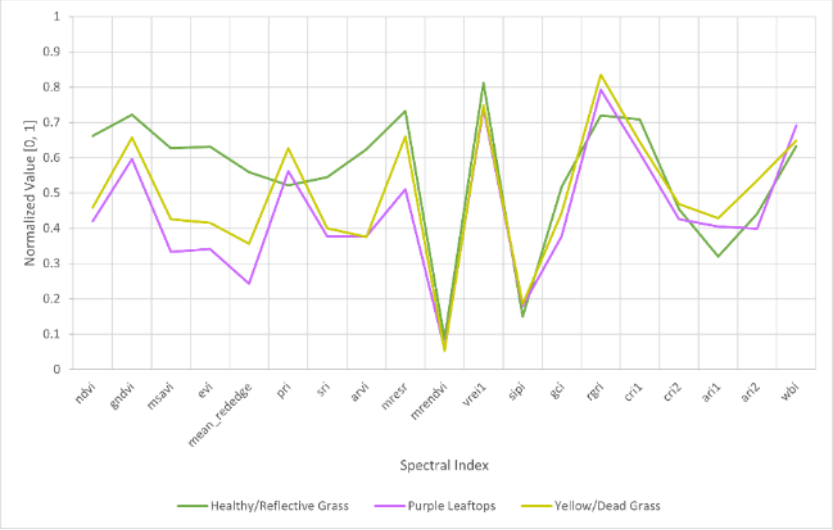
(c)


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

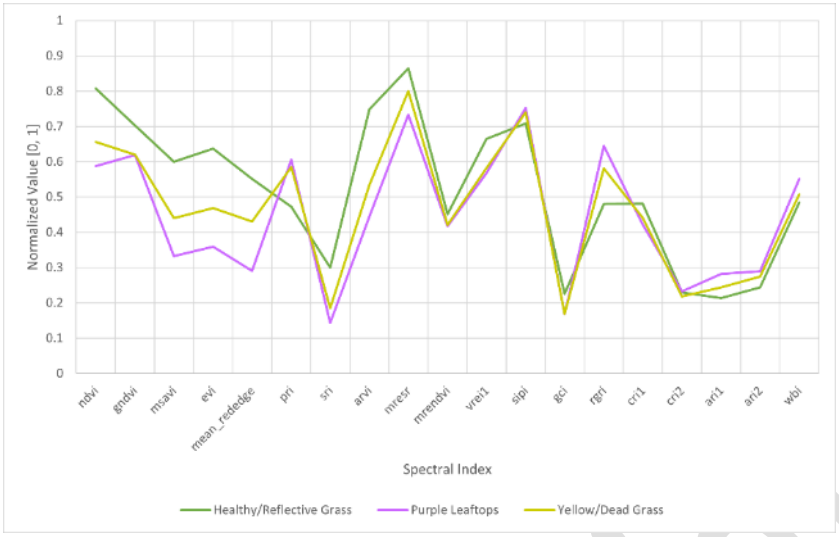
(d)


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

(e)


Normalized Value [0, 1]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ndvi
gndvi
msavi
evi
mean_rededge
pri
sri
arvi
mresr
mrendvi
vrei1
sipi
gci
rgri
cri1
cri2
ari1
ari2
wbi
Spectral Index
Healthy/Reflective Grass
Purple Leaftops
Yellow/Dead Grass

(f)



(g)

Fig. 20. Individual normalised spectral index values of selected pasture classes for PD analysis. **(a)** Maudsland, QLD (2021-02-05). **(b)** Maudsland, QLD (2021-03-09). **(c)** Nobbys Creek, NSW (2021-02-11). **(d)** Nobbys Creek, NSW (2021-04-21). **(e)** Banana Station, QLD. **(f)** Biggenden, QLD. **(g)** Kin Kin, QLD.

8.3 Performance metrics of PD classifiers

Detailed performance metrics of PD using hyperspectral data are shown below.

Table 7. Performance metrics of ML classifier for Maudsland, QLD (2021-02-05).

Class	Precision	Recall	F1-score	Support
Healthy/Reflective Grass	95%	88%	91%	286
Purple Leaftops	97%	96%	96%	97
Yellow/Dead Grass	92%	89%	91%	123
Soil	97%	84%	90%	37
Non-grass	97%	99%	98%	1254
Accuracy			96%	1797
Macro avg	95%	91%	93%	1797
Weighted avg	96%	96%	96%	1797

Table 8. Performance metrics of ML classifier for Maudsland, QLD (2021-03-09).

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	97%	99%	98%	89
Purple leaftops	97%	97%	97%	31
Yellow/dead grass	96%	100%	98%	52
Soil	100%	71%	83%	7
Non-grass	99%	97%	98%	115
Accuracy			98%	294
Macro avg	98%	93%	95%	294
Weighted avg	98%	98%	98%	294

Table 9. Performance metrics of ML classifier for Nobbys Creek, NSW (2021-02-11).

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	92%	97%	95%	365
Purple leaftops	95%	99%	97%	235
Yellow/dead grass	97%	86%	91%	110
Soil	94%	88%	91%	72
Non-grass	93%	86%	89%	132
Accuracy			94%	914
Macro avg	94%	91%	93%	914
Weighted avg	94%	94%	94%	914

Table 10. Performance metrics of ML classifier for Nobbys Creek, NSW (2021-04-21).

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	95%	100%	97%	70
Purple leaftops	100%	100%	100%	30
Yellow/dead grass	100%	82%	90%	11
Soil	100%	100%	100%	4
Non-grass	100%	90%	95%	20
Accuracy			97%	135
Macro avg	99%	94%	96%	135
Weighted avg	97%	97%	97%	135

Table 11. Performance metrics of ML classifier for Banana, QLD.

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	97%	95%	96%	110
Purple leaftops	97%	93%	95%	42
Yellow/dead grass	96%	97%	97%	113
Soil	100%	100%	100%	643
Non-grass	99%	100%	99%	682
Accuracy			99%	1590
Macro avg	98%	97%	98%	1590
Weighted avg	99%	99%	99%	1590

Table 12. Performance metrics of ML classifier for Biggenden, QLD.

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	97%	95%	96%	114
Purple leaftops	74%	79%	76%	33
Yellow/dead grass	96%	96%	96%	272
Soil	92%	100%	96%	66
Non-grass	100%	67%	80%	18
Accuracy			94%	503
Macro avg	92%	87%	89%	503
Weighted avg	94%	94%	94%	503

Table 13. Performance metrics of ML classifier for Kin Kin, QLD.

Class	Precision	Recall	F1-score	Support
Healthy/reflective grass	99%	99%	99%	348
Purple leaftops	99%	99%	99%	188
Yellow/dead grass	91%	96%	93%	93
Soil	100%	97%	99%	35
Non-grass	100%	99%	100%	691
Accuracy			99%	1355
Macro avg	98%	98%	98%	1355
Weighted avg	99%	99%	99%	1355

8.4 Input features of Random Forest model of maxEVI

The climatic variables are notated as follows: <variable>_<season>_<month>. For example,

- T.Max_0_Oct: Monthly average maximum temperature in Oct of the previous year
- T.Max_0_Apr: Monthly average maximum temperature in April of the prediction year
- T.Max_1_Oct: Monthly average maximum temperature in Oct of the season before the previous season T.Max_1_Apr: Monthly average maximum temperature in April of the previous year

A subset of the training data is available here.

Variable	Definition
EVI	Maximum EVI of the predicting season (target)
EVI_1	Maximum EVI of the previous season
EVI_2	Maximum EVI of the season before the previous season
T.Max_0_(Oct-Apr)	Monthly average maximum temperature of the predicting season.
T.Min_0_(Oct-Apr)	Monthly average minimum temperature of the predicting season.
Rain_0_(Oct-Apr)	Monthly average rainfall of the predicting season.
Evap_0_(Oct-Apr)	Monthly average evaporation of the predicting season.
Radn_0_(Oct-Apr)	Monthly average radiation of the predicting season.
VP_0_(Oct-Apr)	Monthly average vapor pressure of the predicting season.
T.Max_1_(Oct-Apr)	Monthly average maximum temperature of the previous season.
T.Min_1_(Oct-Apr)	Monthly average minimum temperature of the previous season.
Rain_1_(Oct-Apr)	Monthly average rainfall of the previous season.
Evap_1_(Oct-Apr)	Monthly average evaporation of the previous season.
Radn_1_(Oct-Apr)	Monthly average radiation of the previous season.
VP_1_(Oct-Apr)	Monthly average vapor pressure of the previous season.
T.Max_2_(Oct-Apr)	Monthly average maximum temperature of the season before the previous season.
T.Min_2_(Oct-Apr)	Monthly average minimum temperature of the season before the previous season.
Rain_2_(Oct-Apr)	Monthly average rainfall of the season before the previous season.
Evap_2_(Oct-Apr)	Monthly average evaporation of the season before the previous season.
Radn_2_(Oct-Apr)	Monthly average radiation of the season before the previous season.
VP_2_(Oct-Apr)	Monthly average vapor pressure of the season before the previous season.

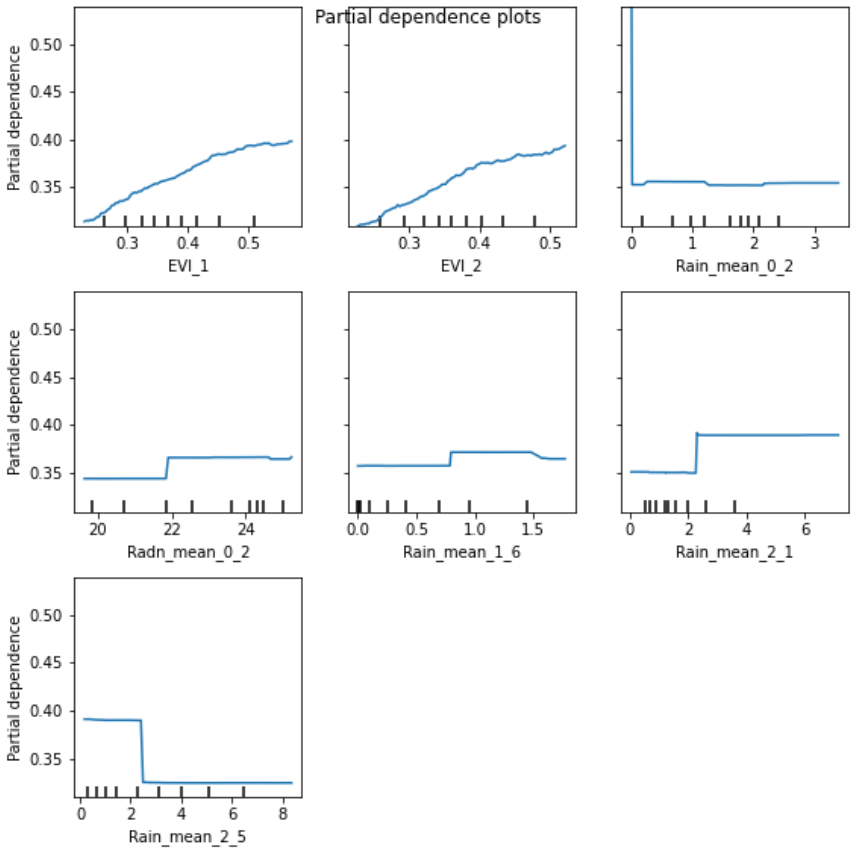


Fig. 37. Partial dependence plots of Random Forest model for MaxEVI