Reflecting on Hyperspectral Reflectance

20 Years of Classifying Vegetation Spectra

Background

- Mapping of plant species is possible using hyperspectral reflectance spectra.
- Reflectance spectra covers the visible, near-infrared, and shortwave-infrared regions of the electromagnetic spectrum (400—2500 nm).
- Small differences in leaf morphology and biochemicals between species produces subtle differences in reflectance spectra.

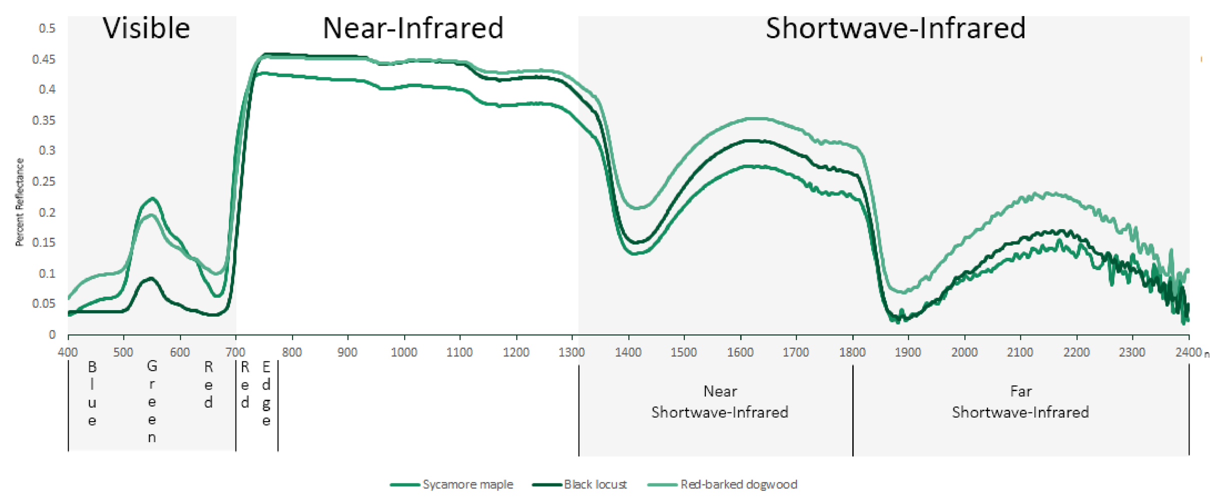

Figure 1: Delineation of main spectral regions in the 400—2400 nm domain. Reflectance spectra of three plant species (Sycamore maple, Black Locust, Red-barked dogwood),

- Hyperspectral data is susceptible to the curse of dimensionality—where the inclusion of a large number of explanatory features in a model can reduce its accuracy.
- Feature selection methods select a sub-sample of the most informative discriminatory features, avoiding the curse.
- A large number of feature selection methods have been used in the literature, with no single best performing method found.

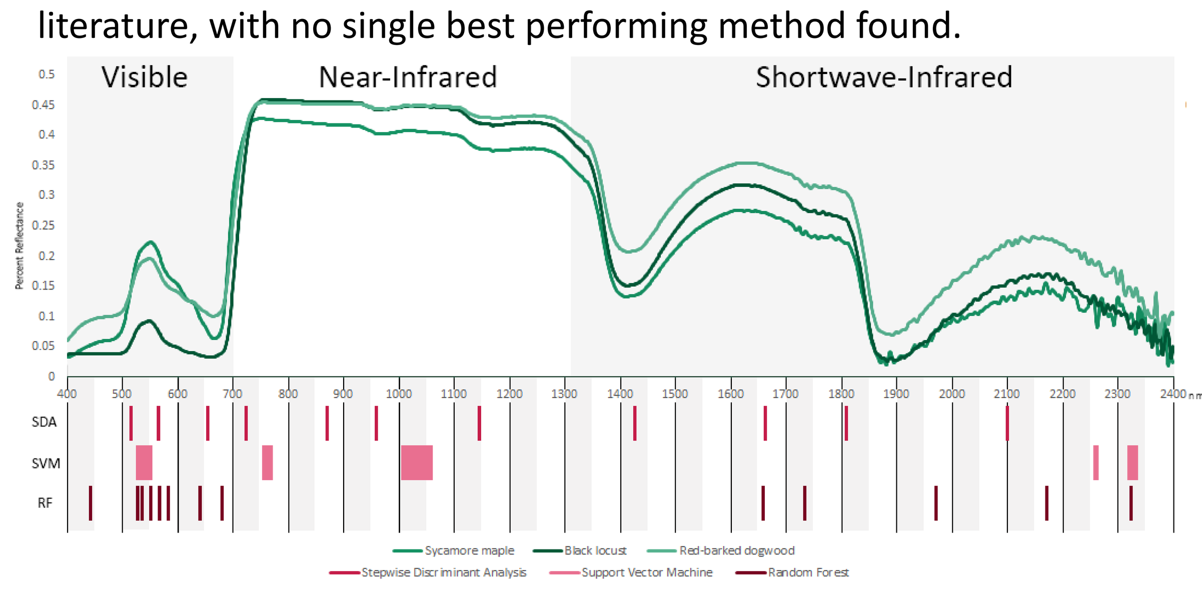

Figure 2: Wavebands selected by three feature selection methods, Stepwise Discriminant Analysis (SDA), Support Vector Machine (SVM), Random Forest (RF).

- Different feature selection methods select for different features when applied to the same dataset.
- The same feature selection method selects for different features when applied to different datasets.
- Pre-processing techniques can improve spectral separability of species, though can potentially decrease it.

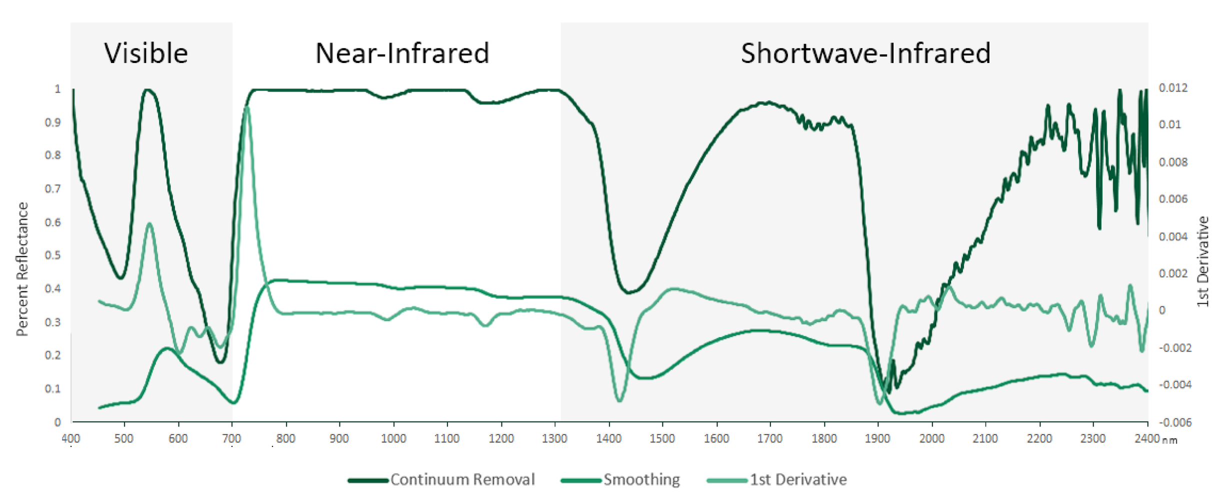

Figure 3: Sycamore maple spectra transformed by three pre-processing methods, continuum removal, Savitzky-Golay smoothing, 1st Derivative

AIM

- Do specific wavebands or spectral regions provide more discriminatory information for species classification by having a greater selection frequency?
- Do taxonomically similar plants select for similar features?

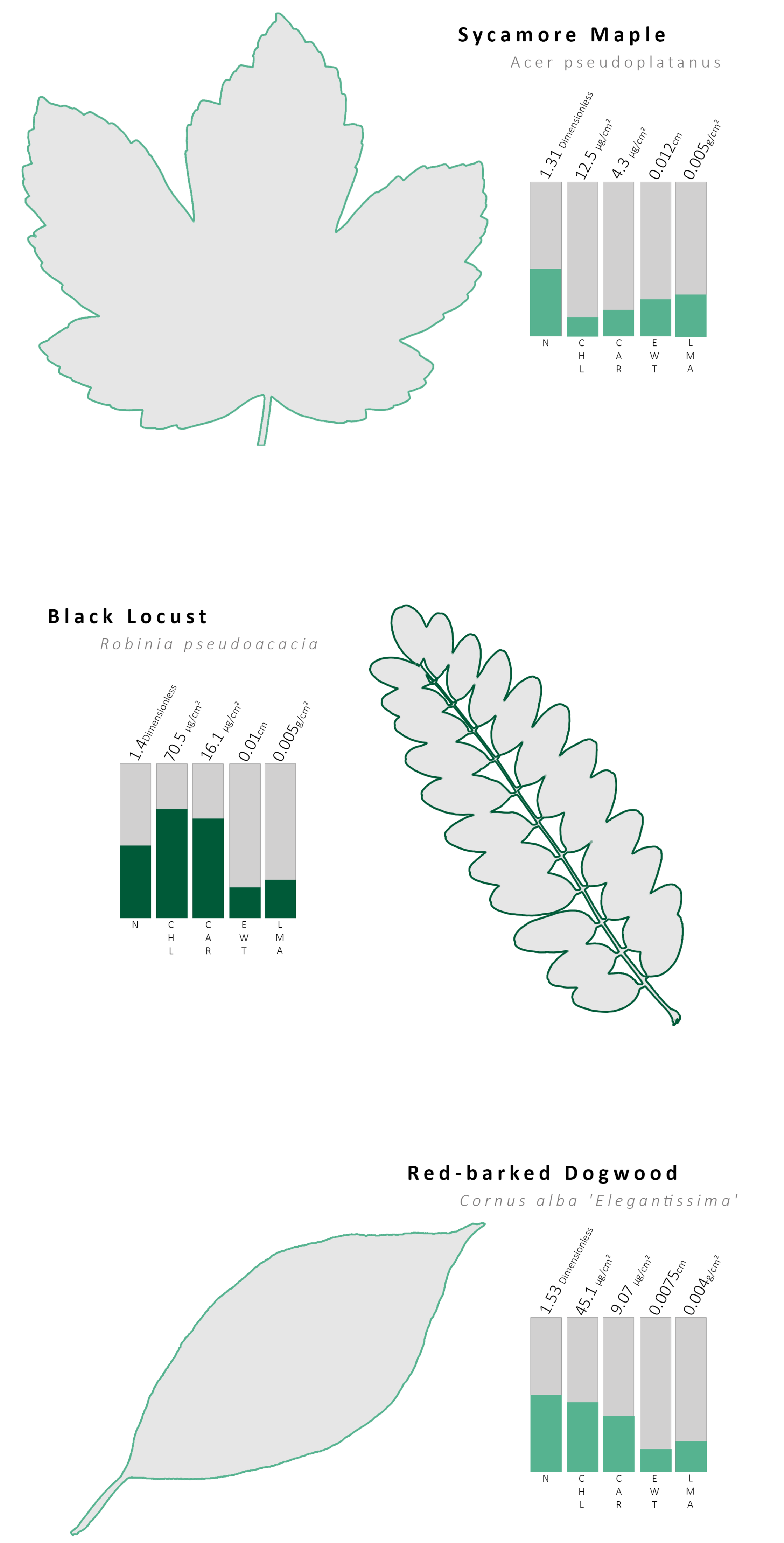

Figure 4: Leaf outlines of Sycamore maple, Black Locust, Red-barked dogwood, along with five leaf metrics measured at the time of spectral recording. **N** a dimensionless value representative of leaf structure. **CHL** total chlorophyll level. **CAR** total carotenoid level. **EWT** leaf equivalent water thickness. **LMA** leaf mass per area.

Method

- A literature review was performed of hyperspectral vegetation classification studies between 1996—2016.
- Waveband selections from the literature were collated into a table with 50 nm increments.
- Selection frequency calculated by the percentage of studies where a waveband was selected for each 50 nm increment.

Results

- 34 studies reported waveband selection.

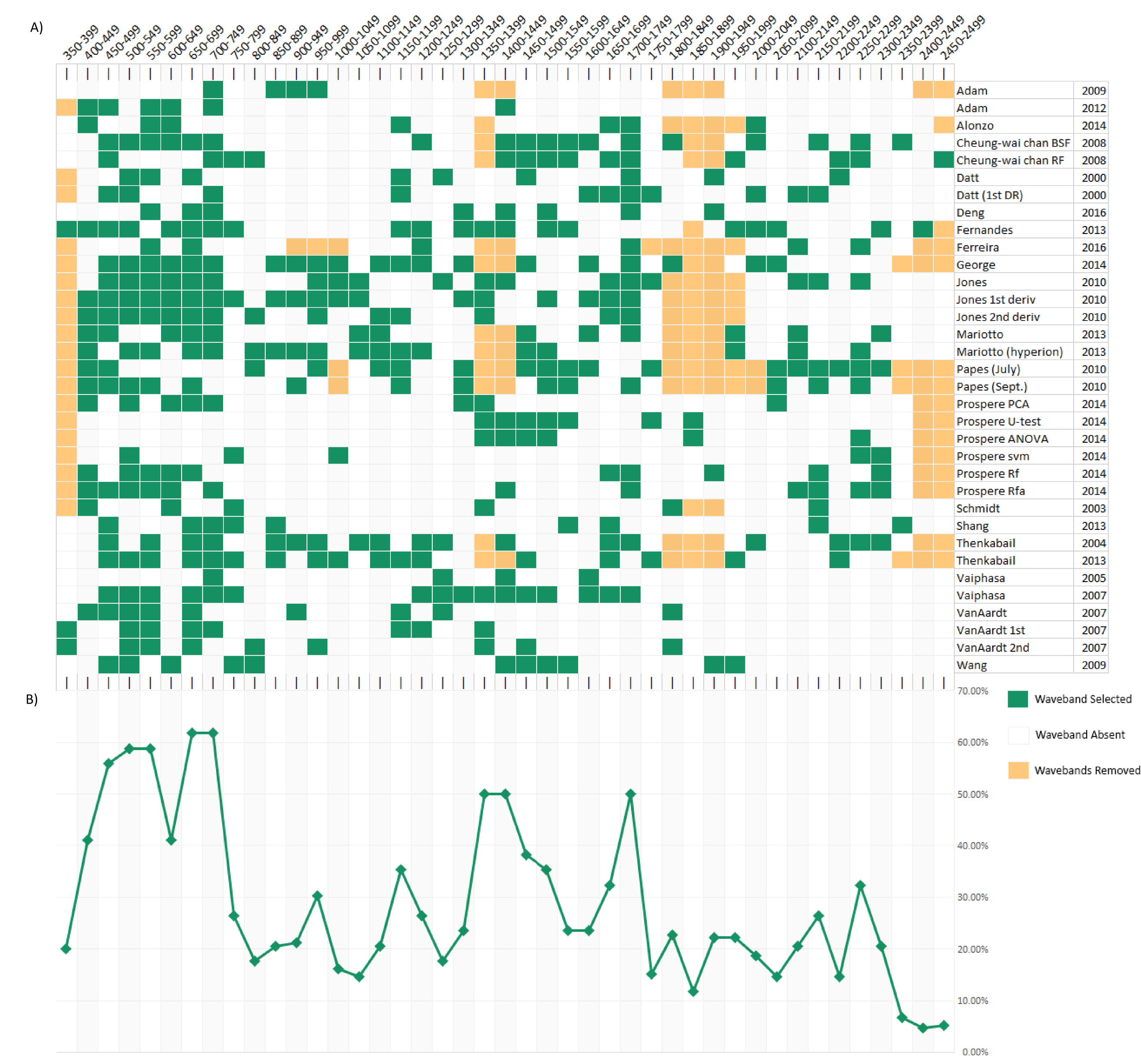

Figure 5: A) Presents and absence table of selected wavebands per individual study. B) Selection frequency of each 50 nm wavelength bin.

- Visible wavelengths have the greatest selection frequency.
- Selection frequency of the Near-infrared is low, except for the Red Edge (700-749 nm).
- The Near Shortwave-Infrared has higher selection frequency than the Far Shortwave-Infrared.

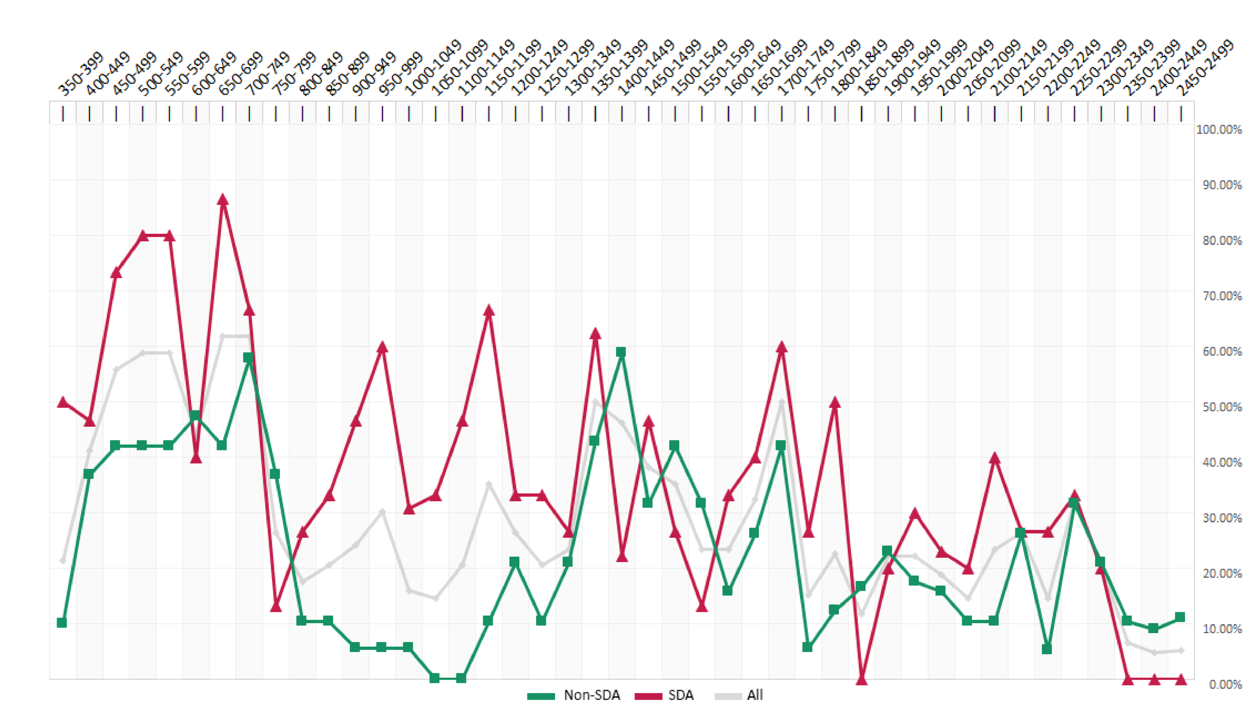

Figure 6: Selection frequency of each 50 nm bin split by studies that used stepwise discriminant analysis (SDA), and those that did not.

- The feature selection method Stepwise Discriminant Analysis (SDA) heavily favours selection of Near-Infrared Wavelengths.
- Blue, Green and Red regions are also selected more frequently by SDA.

Conclusion

- Visible region is selected with greater frequency than other regions.
- Large degree of variability in selected wavebands caused by the use of different pre-processing and feature selection methods.
- No clear evidence for or against similar waveband selection for taxonomically similar plants, due to aforementioned variability.
- Some feature selection methods demonstrate a bias towards selecting certain spectral regions.

PhD Student: Andrew Hennessy

Supervisors: Megan Lewis

Ken Clarke

THE UNIVERSITY of ADELAIDE