Proximal Near-Infrared Spectral Reflectance Characterisation of Weeds Species in New Zealand Pasture

Wayne S. Holmes, Melanie Po-Leen Ooi, Morgan Look,
Dept. of Engineering,
Unitec Institute of Technology,
New Zealand

Ye Chow Kuang
School of Engineering,
University of Waikato,
New Zealand

Ray Simpkin
Transducers and Sensing,
Callaghan Innovation,
New Zealand

Dan Blanchon
Dept. of Environmental and Animal Sciences,
Unitec Institute of Technology,
New Zealand

Serge Demidenko
School of Science & Technology
Sunway University,
Malaysia

Abstract—**Recent extensive study was performed on Jacobaea vulgaris that characterized the plant leaves using hyperspectral reflectance against the successional stage of vegetation. It reported a high similarity of the leaf spectral reflectance measurement over the plant's different stages of growth. This paper extends the earlier study by characterizing the proximal spectral reflectance measurements of three species of common New Zealand pasture weeds: (a) *Jacobaea Vulgaris*; (b) *Rubus*; and (c) *Ulex* growing in three different common soil pastures, specifically: (i) *Typic Orthic Gley*; (ii) *Typic Orthic Granular*; (iii) *Typic Orthic Brown*. The research goes on to determine the inter- and intra- species proximal spectral reflectance variation of the studied common weeds. Finally, it examines the suitability and extent of accuracy of different statistical analysis methods when applied on proximal spectral reflectance measurement to identify the three common species of weeds growing on New Zealand pastures.**

Keywords—spectral reflectance, field spectroscopy, proximal imaging, spectral imaging, weed detection and identification, pasture.

I. INTRODUCTION

In the context of precision agriculture, researchers demonstrated the potential use of proximal hyperspectral imaging to assess pasture quality [1]. Studies [2-4] demonstrated that it can be a fast and inexpensive tool to successfully evaluate pasture quality indicators such as levels of crude protein and neutral detergent fiber, organic matter digestibility, etc. However, the proximal hyperspectral imaging approach employs canopy spectral measurements to assess such bulk properties of the imaged area rather than evaluating individual plants. The obvious advantage of the canopy measurements is its speed since the assessment covers a large area at a single time. The downside is that it is unable to accurately locate individual unwanted plants or weeds that reduce the overall quality of the pasture. And thus it cannot assist to these weeds and plants elimination.

The management of invasive organisms in New Zealand is a priority for the country due to the risk to its unique flora and fauna as well as due to real and potential impacts on the agriculture and horticulture production [5]. Pastoral farming takes up nearly half of the total New Zealand land area [6]. At the same time, since 2007, 187 invasive plants have been reported as weeds in these pastures. Of these, *Jacobaea Vulgaris* (common name: *Ragwort*), *Ulex* (common name: *Gorse*), and *Rubus* (common name: *Blackberry*) are in the top ten reported significant pastoral weed species in the country [6].

The obvious benefit of precisely locating and destroying single weed plants without affecting the neighboring pasture plants is the elimination of broadcast herbicide spraying. This directly reduces the associated treatment costs, increases pasture productivity, and improves environmental sustainability. In the New Zealand context, *kaitiakitanga* [7] or stewardship/ guardianship is a very relevant concept, whereby any use of herbicides has to be clearly justified. The capability to perform an individual weed plant treatment promotes sustainable land management while recognizing the importance and integrating the *kaitiakitanga* compliant practices.

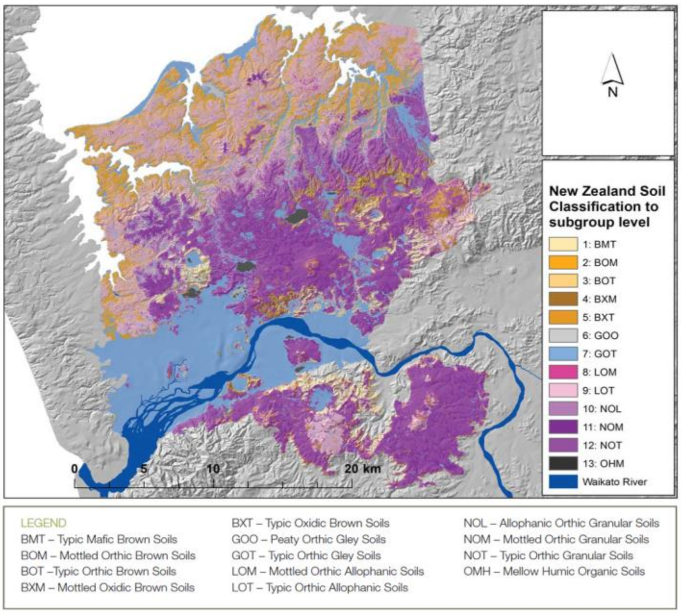

Fig. 1. Mapping of soil types in Franklin County, New Zealand

Plant discrimination using hyperspectral measurements was found to be viable through the studies of the spectral signatures of individual plants during their lifecycles, and also their component parts such as flowers and leaves [8-10]. The studies showed that the spectral signature of an individual plant was affected by a chemical make-up of the plant material, which, in turn, was influenced by season, nutrients, stages of growth, etc. Also there is a considerable variation between component parts of the plant such as leafs and flowers. Furthermore, soil type has a significant impact on the dry matter and plant chemistry [10], and hence it possibly has an impact on the measured spectral reflectance. This is due to the variance in the nutrient absorption that could affect the biochemistry of the leaves. These factors are important considerations for pasture management systems whereby different soil types may be encountered within a single pasture land. For example, it can be clearly seen that there are significant variations of the soil type over a relatively small geographical area of the Franklin county in South Auckland (Fig. 1).

Earlier performed extensive studies on the characterization of the common weed *Jacobaea Vulgaris* [7, 8] involved the proximal leaf spectral measurements based on season and succession. The study shows that leaves of the *Jacobaea Vulgaris* have higher similarity in their spectral measurement at different growth stages. It is important to further extend the study to characterize the spectral reflectance for leaves of the same species of the plant grown on different soil types. The soil composition can account for environmental factors that could change the color of the leaves and can be correlated to their physiochemical compounds. It is also important to determine whether such measurements can be employed to distinguish between common weed species.

This paper has two main objectives:

1. To characterize the near-infrared proximal spectral reflectance of three common weed species that grow on New Zealand pastures of three soil types:
    - The selected weed species are: (a) *Jacobaea Vulgaris*; (b) *Rubus*; (c) *Ulex*;
    - The selected soil types are (a) *Typic Orthic Gley*; (b) *Typic Orthic Granular*; (c) *Typic Orthic Brown.*
2. To determine the extent and suitability of different statistical analysis methods when they are applied to the proximal near-infrared spectral reflectance measurements to identify specific plant species.

II. MATERIALS AND METHODS

The study presented in this paper closely follows the spectral plant properties protocol laid out by Jiminez [11]. It specifies the measurement parameters for in-field measurement of plant spectra. Adopting this approach will lead to formation of a reference library of common pasture species. This paper introduces the inclusion of soil composition as an additional physiochemical parameter. The reason for such an addition is that soil is a primary factor in the nutrient uptake and hence the plant chemical make-up. Furthermore, as mentioned above, New Zealand pastures typically include more than one soil type.

a. Study locale

The reported study was carried out in Franklin County, South Auckland, New Zealand. It can be observed from Fig. 1 that the three major soil types encountered in this relatively small space domain are: (i) *Typic Orthic Gley*; (ii) *Typic Orthic Granular*; (iii) *Typic Orthic Brown*. Table 1 provides a short description for each of these soil types. This paper covers a study performed at 3 sites, whereby 16 permanent quadrats of 1m² area were chosen so as to incorporate as many of the target weeds species within the quadrat as possible. Each quadrat was aligned to the cardinal compass points. Its location on the site was fixed by two datum points: their distance and angle to the South-West corner of the quadrat were recorded. This will allow for later measurements to be made in subsequent seasons in the exact locations thus negating the need for repeated soil testing while also facilitating the extension of this research to study seasonal factors.

Table 1. Sire Soil Specification

Site: quadrat numbers	**Soil Type**	**Description**
A: 10	*Typic Orthic Gley*	Poorly drained. Often limited oxygen level. Range of nutrients (iron, manganese, nitrates and sometimes sulphites) can be insufficient. Grey color. Clay minerals dominant. High levels of organic matter. Could be low in nitrogen. Shallow plant rooting.
B: 4	*Typic Orthic Granular*	Tends to be clay soils, kaolin minerals dominant. Sticky and plastic-type. Limited plant rooting depth. High aluminum content. High phosphorus retention.
C: 2	*Typic Orthic Brown*	Brown color. Clay to minerals ratio ~ 2:1. Secondary iron oxides could give a yellowish-brown color. Soils not waterlogged in winter. Moderate to very high phosphorus retention. Good drainage and biologically active soils. Deep plant rooting.

Fig. 2. Example of 1m² quadrat used in the study. The soil here is *Typic Orthic Gley*; and the plant is *Jacobaea vulgaris*. A spherical white reference made from Polytetrafluoroethylene is used for taking hyperspectral images.

b. Spectral measurements

Near-infrared (NIR) measurements for each of the 16 quadrats were taken using the Pika NIR hyperspectral camera with the range of 900 nm to 1700nm [12]. For each quadrat, a total of three hyperspectral images were taken to allow for the assessment of repeatability. Natural full sunlight was used as the illumination source. The measurements were made as close as possible to the solar noon (as described in [13]) to reduce shadowing effects that could cause distortions. In order to quantify the effect of light variations due to cloud conditions, a *Polytetrafluoroethylene* (PTFE) spherical white reference was placed at the northern end of each quadrat. It served as both a visual marker in the image to denote orientation as well as a white *Lambertian* reflector [14].

c. Experimental setup and sample sizes

The list of the experiments along with their objectives are summarized in Table 2. The spectral reflectance measurements were done for the *Jacobaea Vulgaris*, *Rubus* and *Ulex* for all the three afore-mentioned soil types (Table 1).

Table 2. Experiments objectives and samples

No	**Mean spectral measurements: experiment samples and sites**	**Experiment objectives**
1	• 10 *Jacobaea Vulgaris* - Site A • 2 *Jacobaea Vulgaris* - Site B • 4 *Jacobaea Vulgaris* - Site C • 3 *Ulex* - Site A • 2 *Ulex* - Site B • 5 *Rubus* - Site A • 4 *Rubus* - Site B • 3 *Rubus* - Site C	To characterize the proximal NIR mean spectral reflectances for selected weed species found on three different pasture soils
2	• 10 *Jacobaea Vulgaris* - Site A • 3 *Ulex* - Site A • 5 *Rubus* - Site A	Statistically determine whether NIR spectral measurements can distinguish between individual plant species on the *Typic Orthic Gley* soil
3	• 2 *Jacobaea Vulgaris* - Site B • 2 *Ulex* - Site B • 4 *Rubus* - Site B	Statistically determine whether NIR spectral measurements can distinguish between individual plant species on the *Typic Orthic Granular* soil
4	• 4 *Jacobaea Vulgaris* - Site C • 5 *Rubus* - Site C	Statistically determine whether NIR spectral measurements can distinguish between individual plant species on the *Typic Orthic Brown* soil
5	• 16 *Jacobaea Vulgaris* -Sites A, B and C • 5 *Ulex* - Sites A and B • 12 *Rubus* - Sites A, B and C	Statistically determine whether NIR spectral measurements can distinguish between individual plant species regardless of the soil type

d. Data pre-processing and statistical analysis

Due to the uncontrolled nature of the natural lighting, considerable pixel-to-pixel variance within the same hyperspectral image was encountered. Some parts of the studied plants were in a shade while others were exposed to full sunlight. To account for this, the data cubes for each hyperspectral quadrat image were extracted and the mean spectra of the minimum of 5000 pixels was acquired for the target species (Table 2). The mean spectra of the PTFE white reference were used to convert the image data from mean radiance to mean reflectance, hence catering for the variations in the natural light spectral output. Finally, in order to further remove the intensity variations, the spectral data were again normalized to the highest frequency (900nm). It allowed for direct comparisons of the spectral measurements.

Statistical comparison was performed for each data group. It was presented in 20nm wavelength bands between 900nm to 1700nm using a single-factor *analysis of variance* (ANOVA) [15]. The statistical significance level was set at $p < 0.05$ for 95% confidence. In addition to the ANOVA, the discriminant analysis [16] was employed. The discriminant analysis assigns weights to evidences across different wavelengths to arrive at a statistically optimal decision surface that segregates different subgroups (species/sites). In the linear discriminant analysis, weights are assigned to each wavelength channel respectively, and the weighted responses are linearly combined to produce a decision score. Unlike the one-way ANOVA, the discriminant analysis does not require the user to specify an arbitrary decision threshold (e.g., 5% significance level). Instead, the algorithm determines a single overall threshold after assigning appropriate weights for each channel based on the combined *signal-to-noise ratio* (SNR). Due to the submission length constraint, this paper presents only the discriminant transformation corresponding to the one-against-all classification [16].

III. RESULTS AND DISCUSSION

a. Spectral reflectance measurements and signal-to-noise ratio

At the first glance, the mean reflectance characteristics appear to be able to discriminate between different species and soil types as shown in Fig. 3. However, the variance in the spectral reflectance is very high in the field measurements. Fig. 4 shows an example of the signal-to-noise ratio (SNR) for the pair-wise differentiation between any two reflectance curves from Fig. 3(d).

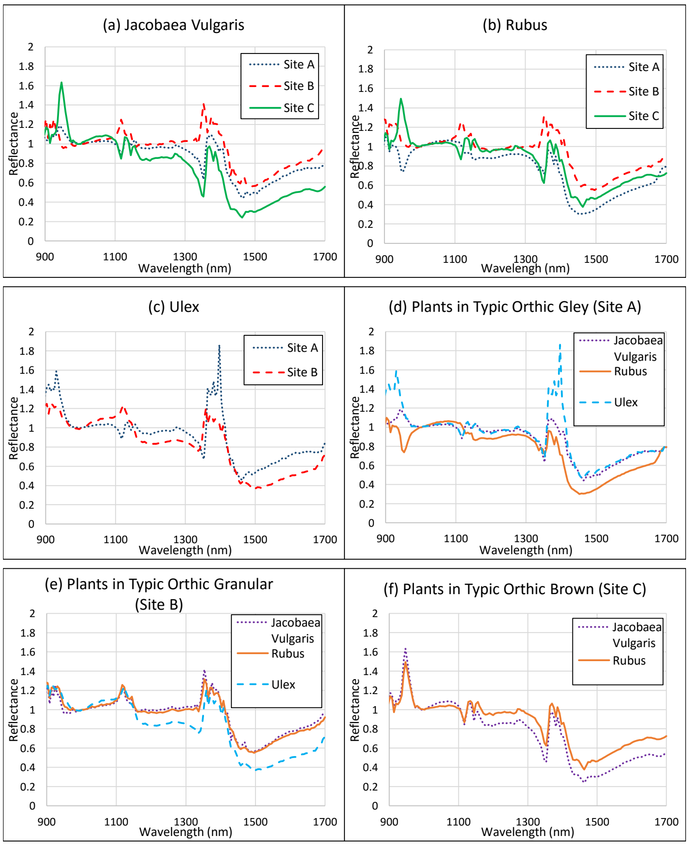

Fig. 3 Mean proximal NIR spectral reflectance on: 1) intra-species variation for (a) *Jacobaea Vulgaris*, (b) *Rubus*, and (c) *Ulex*; and 2) inter-species variation for (d) *Typic Orthic Gley*; e) *Typic Orthis Granular*; and f) *Typic Orthic Brown*

Fig. 3(d) seems to suggest that the region between 900 nm - 1000 nm can be used to differentiate the species due to the large differences in their mean reflectance values. However, the SNR shows that unlikely be possible since the difference in the mean reflectance is only 10% - 30% of the standard deviation in that range. Fig. 4 further shows that the SNR is very low across all the wavelength channels. Therefore any attempt to perform species identification through a small number of hand-pick wavelength channels is likely to be unreliable.

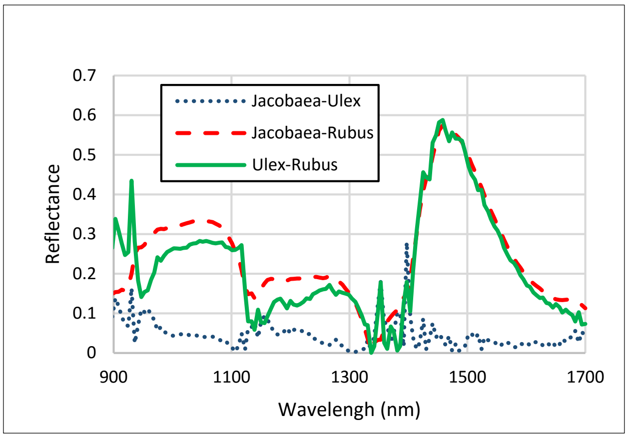

Fig. 4. Signal-to-noise ratio for the measured pairs: SNR is defined as the ratio between reflectance difference to the standard deviation (at each wavelength)

A method to improve SNR from noisy measurements is combining independent data methodologically such that the independent noise cancels out. Research presented in [8-9] described two approaches to this end: (1) the use of the one-way ANOVA to determine statistically significant differences, and (2) application of the discriminant analysis. These two methods are elaborated below in the context of the pasture weed species recognition.

b. One-way ANOVA

The results of applying the one-way ANOVA analysis are presented in Fig. 5 for each of the measurement site. The data is presented such that if the analysis results in a *p*-value that is less than 0.05, the relevant wave band is given a value of ‘1’. Any spectral band that has ‘0’ value, doesn’t yield any statistically significant differences. Ideally, for identification tasks, the comparison within the same group of species should has a value of ‘0’ showing a little variation in the reflectance. Conversely, spectral bands with differing species should yield the value of ‘1’. In the case of the Site A, it can be seen that such a spectral band exists between 1480 nm and 1500 nm (shown as the transparent bar). In the case of the Site B, and Site C, as well as when all the sites are analyzed, no such a band could be determined. This is an expected outcome due to the low SNR (as shown in Fig. 4).

Notably, it can be seen that unlike the other target species, the results for *Rubus* are significantly different between the three sites. This is especially apparent in the Site B where there is large variation across the majority of the wavelength bands between the four individual plants. The measurements for *Rubus* were taken in early spring just as plants awakened from the winter dormancy. It can be speculated that this has caused the variation in the spectral measurements. This phenomenon will be studied further as an extension of the reported research into seasonal factors.

In order to employ the analysis of this type in the future research extension, the challenges of the poor SNR must be addressed by utilizing some relevant pre- or post-processing.

c. Single discriminant transformation

The reported research applied two different discriminant analysis methods: 1) the linear discriminant analysis, i.e., *single discriminant transformation*, and 2) boosting the first discriminant vector with additional discriminant transformations. For the single discriminant analysis, the weighted evidence is mathematically equivalent to a linear transformation of a high-dimensional spectral measurement into a low-dimensional feature space.

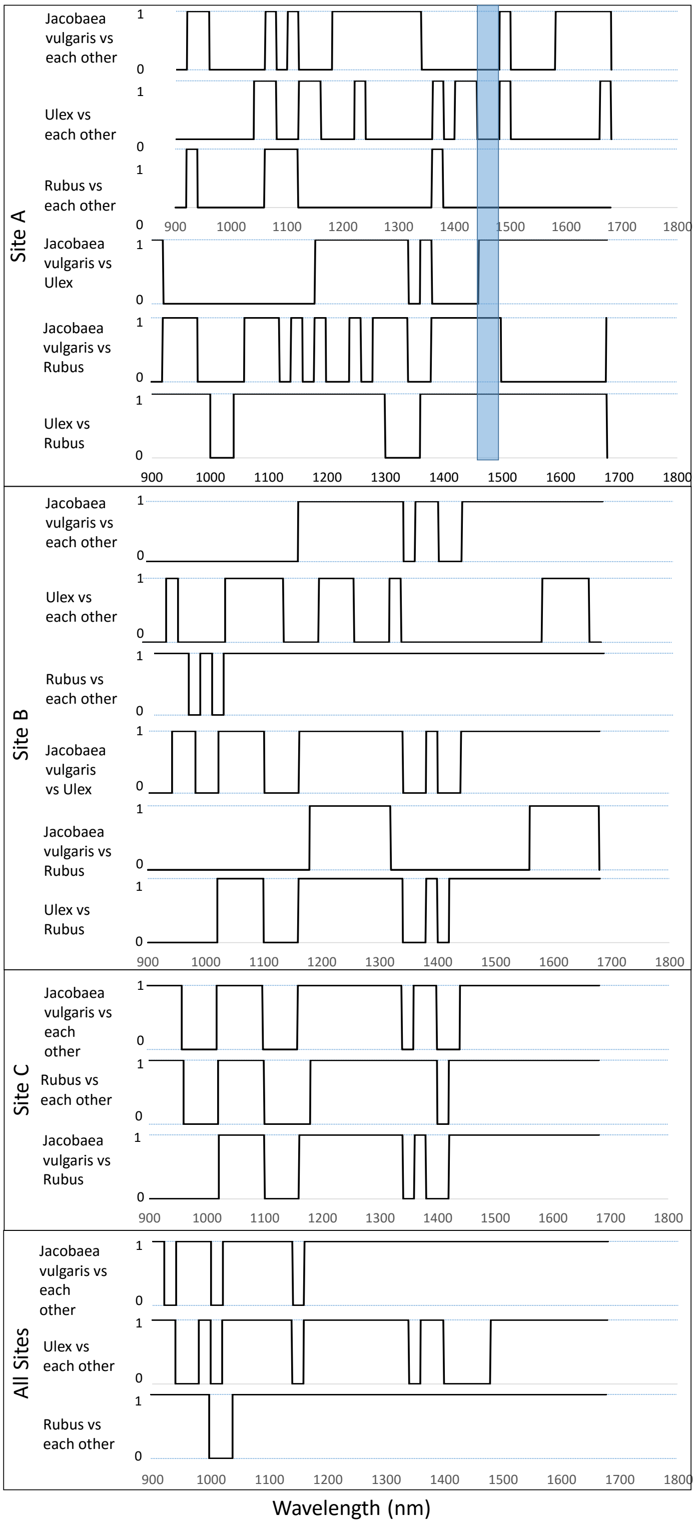

Fig. 5 One-way ANOVA of near-infrared spectral reflectance measurement by wavelength, whereby any $p<0.05$ is set as ‘1’ to indicate SSD

Fig 6 presents results of the single discriminant analysis application to separate the *Jacobaea Vulgaris* from the *Ulex* and

Rubus on three different pasture soil types. It clearly shows that the linear discriminant analysis is able to successfully separate most of the measurements taken for *Jacobaea Vulgaris* against other two species. Similar results were observed for all the studied plant species. However, despite the generally good separation, the transformation still yielded some small percentage of inseparable samples as circled in red in Fig. 6.

d. Boosted discriminant analysis with two transformations

Often there are more than one effective discriminant transformations required. The reported in this research discriminant analysis employs the boosted discriminant algorithm described in [17]. Boosting methodology can be used to combine multiple transformations to construct a more reliable model. The employed algorithm is particularly effective for the spectral data analysis as it finds the most effective discriminant transformations in stages, and then adds them incrementally to improve the overall accuracy. Fig. 7 shows the result of applying a two-discriminant transformation. It is clear that the boosted discriminant analysis improves the margin of classification. The boundary separation between the target plant species against other pasture plant types is very distinct. Therefore, it can be concluded that proximal near-infrared spectral reflectance can be employed to accurately identify pasture weed species when it is used in conjunction with the boosted discriminant analysis.

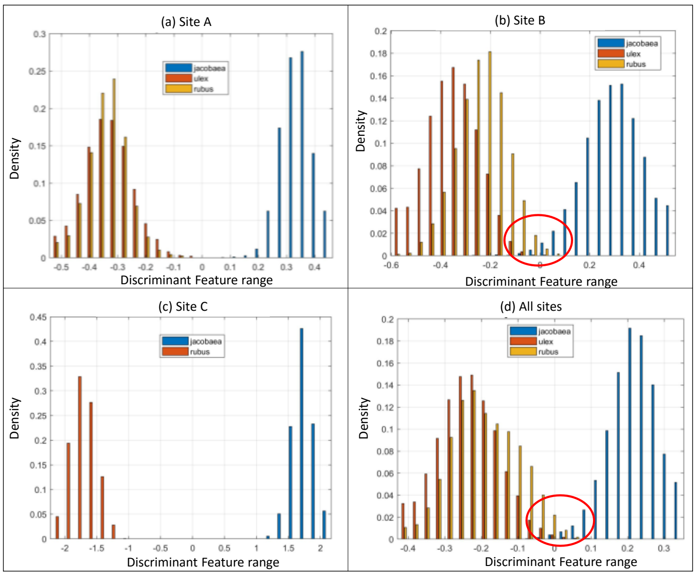

Fig. 6. Single discriminant transformation using linear discriminant analysis for *Jacobaea Vulgaris* against *Ulex* and *Rubus* in (a) Site A; (b) Site B; (c) Site C; and (d) All sites

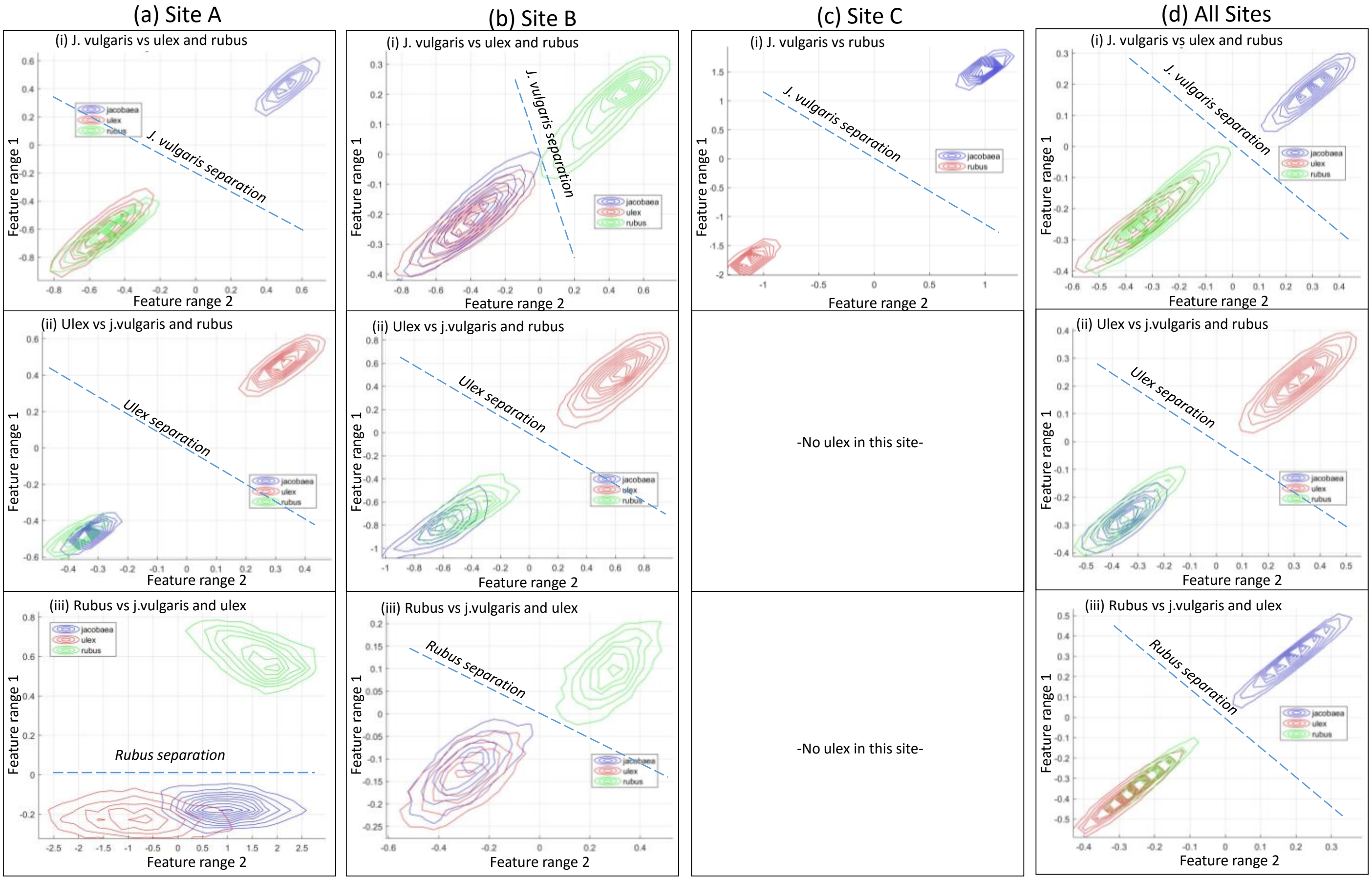

Fig. 7. Boosted discriminant analysis with two transformations by soil type. Contours represent the density of the species in the two-dimensional discriminant transform space. The individual plant species are clearly separable following the indicated separation boundaries.

In short, the results show that boosted discriminant analysis is a flexible methodology that enables reliable discrimination between different plant species under a very low signal-to-noise ratio conditions even when using a simple one-against-all scheme. The method appears to be reproducible across different soil conditions as well. Low SNR is likely to be unavoidable if field measurements are taken using natural lighting. This can be mitigated by either introducing a strong light source, or enhancing the measurements process by employing the boosted discriminant analysis.

An important possible extension of this work is the application of different techniques of the boosted discriminant analysis (e.g., sparse discriminant analysis) to identify key wavelength bands that best distinguish one plant species from others. This would significantly reduce the computational time taken to analyze the measurements. However, this is likely to result in some accuracy trade-offs. Therefore the extension of this research must cover an investigation into the optimal tradeoff for the real world field implementation.

IV. CONCLUSION AND FUTURE WORK

This study has shown that proximal near-infrared hyperspectral imaging with discriminant analysis has a strong potential to be a useful tool in the identification and classification of typical pasture weeds in the New Zealand agricultural industries. It also shows that spectral reflectance measurements must be analyzed with the appropriate statistical tools because of the often poor signal-to-noise ratio due to significant variations in pasture field measurements.

An interesting observation from the study is that seasonal factors appear to have rather significant effects on the mean spectral reflectance response of the target species. Therefore, further on-going work in this project will study the seasonal and plant life cycle effects in the same 16 quadrats used in the reported research in order to develop a deeper understanding on the relationships between NIR proximal spectral reflectance and seasonal/successional factors. Further extension of this research will also include an investigation into different boosted discriminant analysis techniques to compare their performances in terms of accuracy and computational speed, and to recommend the optimal trade-off for real-time weeds identification on a given pasture.

ACKNOWLEDGEMENT

This research was supported by the Strategic Research Fund, Unitec Institute of Technology RI18010, Auckland, New Zealand.

REFERENCES

[1] W. Strothmann, A. Ruckelshausen, J. Hertzberg, C. Scholz, and F. Langsenkamp, "Plant classification with In-Field-Labelling for crop/weed decrimination using spectral features from a multi-wavelength laser line profile system," *Computers and Electronics in Agriculture,* pp. 79-93, 2017

[2] D. Zhao, P. Starks, M. Brown, W. Phillips, and S. Coleman, "Assessment of forage biomass and quality parameters of bermudagrass using proximal sensing of pasture canopy reflectance," *Japanese Society of Grassland Science,* pp. 39-49, 2007

[3] K. Kawamura, N. Watabane, S. Sakanoue, H.-J. Lee, Y. Inoue, and S. Odagawa, "Testing genetic algorithm as atool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression," *Japanese Society of Grassland Science,* pp. 205-216, 2010

[4] R. Pullanagari, I. Yulle, M. Tuohy, M. Hedley, R. Dynes, and W. King, "In-Field hyperspectral proximal sensing for estimating quality parameters of mixed pasture," *Precision Agriculture,* pp. 351-369, 2011

[5] M. N. Clout and S. J. Lowe, "Invasive species and environmental changes in New Zealand," in *Invasive species in a changing world*, H. Mooney and R. Hobbs, Eds., Washington D.C., Island Press, pp. 369-383, 2000

[6] G. W. Bourdot, S. Fowler, G. Edwards, D. Kriticos, *et al.*, "Pastoral weeds in New Zealand: Status and potential solutions," *New Zealand Journal of Agricultural Research,* vol. 50, pp. 139-161, 2007

[7] "What is kaitiakitanga?", Environment Guide, http://www.environmentguide.org.nz/issues/marine/kaitiakitanga/what-is-kaitiakitanga/ (accessed on November 30, 2018)

[8] S. Carvalho, "Hyperspectral reflectance of leaves and flowers of an outbreak species discriminates season and successional stage of vegetation," *International Journal of Applied Earth Observation and Geoinformation,* vol. 24, pp. 32-41, 2013

[9] S. Carvalho, M. Macel, P. Mulder, A. Skidmore, and W. v. d. Putten, "Chemical variation in Jacobaea vugaris is influenced by the interaction of season and vegetation successional stage," *Elsesvier Phytochenistry,* pp. 86-94, 2014

[10] Y. Suzuki, H. Okamoto, M. Takahashi, T. Kataoka, and Y. Shibata, "Mapping the spatial distribution of botanical composition and herbage mass using hyperspectral imageaing," *Japanese Society of Grassland Science,* pp. 1-7, 2012

[10] L. Joosten, P. Mulder, P. Klinkhamer, and J. v. Veen, "Soil-borne microorganisms and soil-type affect pyrrolizidine alkaloids in Jacobea vulgaris," *Plant Soil,* pp. 133-143, 2009

[11] M. Jimenez and R. Diaz-Delgado, "Towards a Standard Plant Species Spectral Library Protocol for Vegetation Mapping: A case Study in the Shrubland of Donana National Park," *ISPR International Journal of Geo-Information,* vol. 4, pp. 2475-2495, 2015

[12] Pika NIR-320 - Near Infrared Hyperspectral Imaging Camera. Resonon. https://resonon.com/Products/imagers_pika_nir-320.html (accessed on November 30, 2018)

[13] J. W. Salisbury, "Spectral Measurements Field Guide," Defense Technology Information Center, Report No. ADA362372, 1998

[14] A Dictionary of Earth Sciences (3 ed.), M. Allaby - Ed., Oxford University Press, 2008

[15] ANOVA Test: Definition, Types, Examples. In: Statistics How To. https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/hypothesis-testing/anova/ (accessed on November 30, 2018)

[16] G. McLachlan, Discriminant Analysis and Statistical Pattern Recognition, Wiley, 2004

[17] H. K. Sok, M. P.-L. Ooi, Y. C. Kuang, and S. Demidenko, "Multivariate alternating decision tree," *Pattern Recognition,* vol. 50, pp. 195-209, 2016