Remote sensing bio-control damage on aquatic invasive alien plant species

Na‘een Agjee a, Onisimo Mutanga a and Riyad Ismail a

a School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, P/Bag X01 Scottsville, Pietermaritzburg, 3209, South Africa

DOI: http://dx.doi.org/10.4314/sajg.v4i4.8

Abstract
Aquatic Invasive Alien Plant (AIAP) species are a major threat to freshwater ecosystems, placing great strain on South Africa’s limited water resources. Bio-control programmes have been initiated in an effort to mitigate the negative environmental impacts associated with their presence in non-native areas. Remote sensing can be used as an effective tool to detect, map and monitor bio-control damage on AIAP species. This paper reconciles previous and current research concerning the application of remote sensing to detect and map bio-control damage on AIAP species. Initially, the spectral characteristics of bio-control damage are described. Thereafter, the potential of remote sensing chlorophyll content and chlorophyll fluorescence as pre-visual indicators of bio-control damage are reviewed and synthesised. The utility of multispectral and hyperspectral sensors for mapping different severities of bio-control damage are also discussed. Popular machine learning algorithms that offer operational potential to classify bio-control damage are proposed. This paper concludes with the challenges of remote sensing bio-control damage as well as proposes recommendations to guide future research to successfully detect and map bio-control damage on AIAP species.

1. Introduction
Aquatic Invasive Alien Plant (AIAP) species are a major threat to the structure and functioning of freshwater ecosystems throughout South Africa (van Wilgen et al., 2004). Once introduced, AIAP species have the capability to proliferate profusely and spread rapidly thus successfully colonising freshwater ecosystems (Verma et al., 2003; Kull and Rangan, 2008). The unrestrained expansion of AIAP species is attributed to the lack of natural enemies as well as the prevalence of eutrophic freshwater systems (Law, 2007). The colonization of freshwater bodies by AIAP species cause severe social, economic and environmental problems (Coetzee et al., 2007; Villamagna and Murphy, 2010). AIAP species have been found to reduce indigenous biodiversity, increase sediment deposition, reduce water quality and alter water regimes (Görgens and van Wilgen, 2004; Richardson and van Wilgen, 2004). Consequently, bio-control programmes have been established in
an effort to control the extent of AIAP species and mitigate the adverse effects associated with their presence in freshwater ecosystems (Coetzee et al., 2011).

Assessing bio-control damage on AIAP species is essential to evaluating the efficacy of bio-control agents (i.e. larval and adult insects) and the success of bio-control initiatives. Current monitoring techniques which involve visually surveying AIAP species for bio-control damage are subjective, spatially restrictive, time consuming and laborious to conduct on a large scale (Everitt et al., 2002). The advent of remote sensing technologies could play an important role in addressing the challenges faced in obtaining information on bio-control damage. Field based remote sensing forms an integral component of characterising the spectral reflectance of AIAP species under insect induced stresses (Everitt et al., 2007). Satellite based remote sensing provides a synoptic view of the Earth’s surface (Verma et al., 2003) which can capture complete and accurate information on bio-control damage repeatedly (Dennison et al., 2009; Nagler et al., 2014). Information that was once inaccessible over large water bodies can be acquired instantaneously in a cost effective and timeous manner (Joshi et al., 2004). Satellite imagery can be incorporated into a geographical information system for multi-temporal analysis (Albright et al., 2004) which could be implemented to quantify the change in bio-control damage over time. The implementation of semi-automated classification algorithms (Benz et al., 2004) could provide near real time information on the extent and severity of bio-control damage. Therefore, remote sensing as a tool could potentially provide a comprehensive mapping and monitoring solution to identify AIAP species, predict bio-control target areas, detect bio-control damage and effectively plan bio-control strategies (Hestir et al., 2008).

Over the past decade, few studies have applied the use of remote sensing to detect and map bio-control damage on AIAP species (Everitt et al., 2005; Everitt et al., 2007; Everitt et al., 2012). These studies have primarily focused on detecting and discriminating different severities of bio-control damage based on the colour of the foliage using aerial photography and conventional classification techniques (Everitt et al., 2005; Everitt et al., 2007; Everitt et al., 2012). It is evident that there is a significant gap in the body of knowledge on remote sensing bio-control damage on AIAP species. It is prudent to critically assess what new remotely sensed data sources and methods offer for bio-control damage detection. Measuring key biochemical and biophysical parameters could potentially provide greater insight into the extent of physiological stress experienced by AIAP species under bio-control. The greater availability of remotely sensed data at higher spatial and spectral resolutions coupled with the development of machine learning algorithms could potentially improve classification accuracies (Abdel-Rahman et al., 2014; Adelabu et al., 2014). However, as research related to remote sensing bio-control damage is limited, it is necessary to investigate remote sensing of insect induced stresses in other related fields, including agriculture and forestry,
to determine its potential application in detecting bio-control damage. Lessons learnt will assist in identifying key areas of research that should be explored further to develop operational methods that will meet the information requirements of bio-control initiatives.

In light of the above, this paper aims to review the application of remote sensing to detect and map bio-control damage on AIAP species. More specifically this paper focuses on the spectral characteristics of bio-control damage as well as remote sensing chlorophyll content and chlorophyll fluorescence as pre-visual indicators of bio-control damage. The use of multispectral and hyperspectral sensors for mapping different severities of bio-control damage are also discussed. Furthermore, this paper proposes popular machine learning algorithms that offer operational potential to classify bio-control damage. This paper concludes with the challenges of remote sensing bio-control damage as well as proposes recommendations to guide future research to successfully detect, map and monitor bio-control damage on AIAP species.

2. Detecting bio-control damage on AIAP species

2.1. A description of bio-control damage and its spectral characteristics

Bio-control agents inflict both morphological and physiological damage to AIAP species (Allen Dray, Jr. and Center, 2002; Moran, 2004). With reference to water hyacinth (*Eichhornia crassipes*), *Neochetina* spp. larvae tunnel through the plant causing internal damage to the plant tissue (Julien, 2001). Thereafter, herbivory by adult bio-control agents remove large portions of epidermal tissue from the leaf surface forming rectangular scars on the surface of the leaf (Julien, 2001). Feeding by bio-control agents negatively affect the leaf thickness, internal cell structure, pigment concentration and nutrient concentration thereby reducing the photosynthetic performance of the leaf. In addition, a large volume of water is lost through the feeding scars on the leaf surface thereby exposing the leaf to desiccation (Julien *et al.*, 1999). Feeding scars also serve as sites for plant pathogen infection making the plant susceptible to disease infection (Julien, 2001). It is a combination of the morphological and physiological stresses mentioned above that alters the spectral reflectance of infested plants at the leaf surface (Everitt *et al.*, 2005; Robles *et al.*, 2010).

The spectral reflectance of bio-control damage is primarily influenced by the pigments, intracellular airspaces, cellular structure and water content at the leaf surface. The spectral reflectance of bio-control damage in the visible region (VR: 400-700nm) is primarily governed by a reduction in chlorophyll concentration and leaf pigments (Fletcher, 2013). Infested plants exhibit a low reflectance in the green region together with low absorption pits in the blue (450nm) and red (670nm) regions. In the near infrared region (NIR) (700-1300nm), plants under stress, exhibit a low reflectance owing to extensive damage to the cellular structure of the leaf (Carter, 1993). The red-
edge region is characterised by an abrupt change in reflectance between the VR (670nm) and NIR (800nm), caused by the combined effects of strong chlorophyll absorption in the red wavelengths and high leaf structure-driven reflectance in the NIR (Main et al., 2011). In the short wave infrared region (SWIR) (1300-2500nm) incident radiation is both reflected and absorbed with reduced absorption, centered at 1400nm and 1900nm (Huang and Apan, 2006). It is the difference in the leaf optical response between healthy and infested plants that allow for the detection of bio-control damage on AIAP species (Figure 1). Different severities of bio-control damage can be discriminated based on subtle differences in the reflectance spectra between severities. Identifying diagnostic features that are related to the biochemical and biophysical status of the plant (Adam et al., 2010) offers greater potential in discriminating different severities of bio-control damage as well as ascertaining the physiological status of the plant.

![Figure 1: Spectral reflectance of water hyacinth (Eichhornia crassipes) plants a) healthy plant b) severely damaged plant.](image)

2.2. Remote sensing chlorophyll fluorescence

Herbivory by bio-control agents negatively affects the functioning of chloroplasts at the leaf surface thereby reducing the photosynthetic capacity and performance of the leaves (Fletcher, 2013). Light energy absorbed by chlorophyll that is not used for photochemistry is dissipated through non-photochemical quenching or chlorophyll fluorescence; thereby protecting the chloroplasts from light induced oxidative damage (Zarco-Tejada et al., 2000a). Quantifying chlorophyll fluorescence would provide an indication of the functional status of photosynthetic apparatus because it is a direct indicator for photosynthesis (Zarco-Tejada et al., 2013). Therefore, chlorophyll fluorescence can be used as a direct proxy for photosynthesis and a bio-indicator for vegetation stress. Changes in chlorophyll function often precede changes in chlorophyll content
hence changes in chlorophyll fluorescence would occur before leaves become chlorotic. Chlorophyll fluorescence could therefore be a good pre-visual bio-indicator of physiological stress experienced by AIAP species.

Remote sensing chlorophyll fluorescence could potentially determine the severity of physiological stress experienced by AIAP species before visual bio-control damage is seen. Chlorophyll fluorescence signal retrieval from reflectance spectra is a challenge as reflectance spectra is a combination of fluorescence emission and surface reflectance. A study conducted by Zarco-Tejada et al. (2000a-b) demonstrated that chlorophyll fluorescence can be detected at the leaf and canopy level in the red-edge spectral region. However, estimating chlorophyll fluorescence in vegetation canopies under natural illumination would provide a better understanding for detecting chlorophyll fluorescence using airborne or spaceborne remote sensing. Zarco-Tejada et al. (2001) estimated chlorophyll fluorescence under natural illumination from hyperspectral data. The results demonstrated that chlorophyll fluorescence changes can be tracked under natural illumination conditions using a fibre spectrometer over small canopies. These results are promising and illustrate the potential detection of chlorophyll fluorescence on AIAP species experiencing bio-control under natural illumination.

Many studies have adopted a reflectance based approach by computing optical indices that are related to chlorophyll fluorescence (Zarco-Tejada et al., 2000a; Zarco-Tejada et al., 2000b; Dobrowski et al., 2005). Chlorophyll fluorescence indices offer improved performance when compared to other popular vegetation indices, such as Normalized Difference Vegetation Index (NDVI), as they provide a good indication of canopy physiology. Majority of published vegetation indices characterize the amount of vegetation and pigment concentration but are not sensitive to the photosynthetic status of the plant (Dobrowski et al., 2005). Chlorophyll fluorescence indices track vegetation physiological effects that are correlated with some phenomena responsible for fluorescence. These indices exploit the effect of chlorophyll fluorescence on reflectance in the red-edge region (600 – 800nm). Dobrowski et al. (2005) explored the link between canopy reflectance derived chlorophyll fluorescence and plant physiological status. This was determined by investigating the influence of steady-state fluorescence on the red-edge spectral region. It was concluded that fluorescence ratio indices (R690/R600 and R740/R800) calculated in the red-edge spectral region tracked plant stress and photosynthetic status. In their study, Belasque et al. (2008) investigated the potential of fluorescence spectroscopy to detect stress caused by pitch canker (bacterial disease caused by Xanthomonas axonopodis pv. citri) and mechanical injury. This study (Belasque et al. 2008) employed ratios between fluorescence at different wavelengths (452 and 685nm, 452nm and 735nm, 685nm and 735nm) to monitor the stress caused by the bacterial infection. It was reported that the ratio of two chlorophyll fluorescence bands can be used to detect
and discriminate between the mechanical damage and the disease severity. This demonstrates that there is an avenue for passive plant physiological monitoring on AIAP species under bio-control using remote sensing. However, further research needs to be undertaken to investigate the operational potential of detecting chlorophyll fluorescence on AIAP under bio-control at the leaf, canopy, airborne and spaceborne level. Importantly, it is critical to determine the infestation period after which chlorophyll fluorescence can first be detected. This will assist environmental managers in determining the efficacy of bio-control agents at the initial stages of the infestation. In this manner bio-control strategies can be adapted accordingly by releasing more bio-control agents onto AIAP that require additional stress.

2.3. Remote sensing chlorophyll content

Leaf chlorophyll content is one of the most important bio-indicators of plant physiological state because of its direct role in photosynthetic processes. A number of studies have linked responses in leaf chlorophyll content to plant physiological stress (e.g. Curran et al., 1990; Carter and Knapp, 2001). Quantitative estimates of leaf chlorophyll content from remote sensing platforms can provide a useful means of assessing vegetation stress induced by bio-control agents (Zhang et al., 2008). Decades of research has been undertaken to identify biochemically sensitive regions that can be exploited to form vegetation indices (Main et al., 2011). Differences in reflectance between healthy and stressed vegetation due to changes in chlorophyll content primarily occur in the green region and red-edge region (Ustin et al., 2009). Stressed vegetation exhibit less overall absorption in the VR with the position of the red-edge position inflection point shifting towards shorter wavelengths. Consequently, chlorophyll indices have been developed from leaf optical properties that exploit differences in reflectance between healthy and stressed vegetation in the VR and red-edge region (Datt et al., 1999; Ustin et al., 2009).

NDVI is a popular vegetation index that has been used as a proxy for chlorophyll content estimation and vegetation stress assessment. A study conducted by Dennison et al. (2009) monitored tamarisk (Tamarix spp.) defoliation by applying an NDVI to multispectral Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. Results showed that NDVI values decreased within riparian areas between 2006 and 2007. The decrease in NDVI caused by defoliation was apparent despite partial defoliation within the study area. However, the application of broadband NDVIs are limited by their instability owing to variations in canopy structure, illumination and viewing angle (Main et al., 2011). Additionally, broadband NDVIs asymptotically approach a saturation level after a Leaf Area Index (LAI) of approximately four; therefore overestimating the vegetation condition (Seller, 1985; Mutanga and Skidmore, 2004). Despite these limitations, future studies should investigate the potential of using chlorophyll indices to estimate leaf chlorophyll content on AIAP species as no studies have attempted to do so thus far.
The significance of assessing chlorophyll content within physiological studies has prompted the development of numerous chlorophyll indices for chlorophyll content estimation. However, chlorophyll indices are generally species specific as the size, shape, surface and internal structure of the leaves vary from one plant species to another (Zhang et al., 2008). Ideally, the goal for researchers would be to develop vegetation indices that are not only as sensitive as possible to the desired parameter, but also robust across plant species and leaf structures (Main et al., 2011). In order to improve the robustness and generality of chlorophyll indices it is important to test their performance over a range of species (Zhang et al., 2008). Main et al. (2011) assessed the robustness of 73 published chlorophyll spectral indices using leaf level hyperspectral data collected from three crop species and a variety of savannah tree species. The authors identified two red-edge derivative based indices (red-edge position via linear extrapolation index and the modified red-edge inflection point index) that were the most consistent and robust. Despite their applicability, the performance of spectral indices would need to be assessed and calibrated when applied to specific AIAP species under bio-control at the leaf, canopy and airborne level (Zhang et al., 2008). Wu et al. (2010) evaluated the potential of chlorophyll content estimation using spaceborne EO-1 Hyperion hyperspectral data. Wavelengths in the red-edge region were selected to test a range of vegetation indices for chlorophyll content estimation in different canopy structures. Results showed that chlorophyll content can be successfully estimated by vegetation indices derived from Hyperion data with a Root Mean Square Error (RMSE) of 7.20–10.49 mg cm\(^{-2}\) for chlorophyll content. The (Modified Chlorophyll Absorption Ratio Index [MCARI]/Optimized Soil-Adjusted Vegetation Index [OSAVI\(705\)]) index provided the best estimation of the chlorophyll content (RMSE of 7.19 mg cm\(^{-2}\)). This is promising as environmental managers can produce regional maps of chlorophyll content thus determining the extent of physiological damage caused by bio-control agents on AIAP species. However, it is evident that there is a lack of studies that have investigated the potential of estimating chlorophyll content on AIAP species under bio-control stress. A suite of chlorophyll indices needs to be tested and compared to identify indices that perform well at the leaf, canopy and airborne level for specific AIAP species. Integrating chlorophyll content estimation with chlorophyll fluorescence estimation will form a critical component in providing a comprehensive physiological assessment of AIAP plants under bio-control to environmental managers.

3. The detection and mapping of bio-control damage

3.1. The detection and mapping of bio-control damage using multispectral remote sensing

Over the past few decades, great advances in sensor technologies have resulted in a suite of available data sources at a range of spectral, spatial and temporal resolutions deployed on either airborne or satellite platforms. Multispectral radiometers and multispectral imaging spectrometers (ASTER, Landsat 7 (ETM+), QuickBird, IKONOS, SPOT 5, WorldView, GeoEye), capture data between three to eight broad spectral bands from the VR, NIR and SWIR of the electromagnetic
spectrum. However, multispectral sensors average reflectance spectra over broad spectral bands lacking the detailed reflectance spectra required to accurately discriminate different severities of insect damage (Carson et al., 1995; Qin and Zhang, 2005). Multispectral remote sensing has been widely applied to detect insect damage as imagery is commercially available (Figure 2) (Franklin et al., 2003; Qin and Zhang, 2005; Meigs et al., 2011). However, few studies have applied the use of multispectral remote sensing to detect and map bio-control damage on AIAP species with studies mainly being focused on mapping bio-control damage on giant salvinia (Salvinia molesta) (Everitt et al., 2005).

Initially, studies that investigated the potential of mapping bio-control damage on AIAP species used multispectral radiometers coupled with aerial photographs (Everitt et al., 2005). Aerial photographs are expensive to purchase and incorporate a range of distortions from platform instability and variations in sun angle at different capture dates. However, the use of aerial photographs allow for small areas to be captured when required. Everitt et al. (2005) detected the bio-control damage on the AIAP species Salvinia molesta using multispectral field measurements and aerial photography. Field spectral measurements showed that moderately damaged and severely damaged plants had lower visible and near-infrared reflectance values than healthy plants (Everitt et al., 2005). In order to classify the aerial photographs, the image mode function in Adobe Photoshop was employed. Despite the basic nature of the image classification technique employed results showed that the three damage level classes (healthy, moderate and severely damaged) could be quantified and differentiated. A qualitative comparison of the computer classification to the original aerial photograph indicated that the image analysis technique was sufficient in identifying weevil
damage levels on giant salvinia. This illustrates the potential of detecting and mapping bio-control damage on AIAP species which can only be improved upon using multispectral satellite imagery.

AIAP species occur in large dense monocultures on the surface of freshwater bodies. The spatial resolution of satellite imagery used is highly dependent on the information needs of environmental managers. Moderate spatial resolution satellite imagery will be appropriate to conduct a general survey to ascertain the current infestation conditions. Moderate spatial resolution satellite imagery can be used to discriminate between healthy and damaged plants owing to distinctive variation in spectral reflectance. Chen et al. (2007) detected the “take-all” disease in wheat using moderate spatial resolution Landsat TM Imagery (30 m). An NDVI and Principal Components Analysis (PCA) were used to detect healthy wheat and take-all severely infected wheat canopies. Results showed that healthy wheat had a higher NDVI than infested wheat (healthy: 0.3864, infested: 0.2182). However, Chen et al., (2007) did not investigate the discrimination of different disease severities which may require remotely sensed imagery of a higher spatial or spectral resolution or the application of a more advanced and robust classification methodology. Mirik et al. (2013) monitored Wheat Streak Mosaic (WSM) progression using moderate spatial resolution Landsat 5 TM (30 m) imagery and a constrained energy minimization sub-pixel classifier for disease management in wheat. Even though moderate spatial resolution satellite imagery was used, the application of a sub-pixel classifier resulted in high classification accuracies. Results showed that the overall classification accuracies achieved were > 91% with kappa coefficients ranged between 0.80 and 0.94 for disease detection. Omission errors varied between 2% and 14%, while commission errors ranged from 1% to 21%. Importantly, this study detected and mapped healthy wheat and wheat with moderate to severe WSM using moderate spatial resolution imagery. The method employed can potentially be used to map different severities of bio-control damage using moderate spatial resolution imagery. The method employed was simple which would be advantageous to environmental and freshwater resource managers that want to quickly determine the overall extent and severity of bio-control damage across AIAP stands. The coarse level of information generated from these general surveys can be used to define management zones and identify areas requiring more detailed surveys.

Detecting low severities of bio-control damage at the initial stages of insect infestations is critical in assessing if insect populations are established and actively feeding. High spatial resolution satellite imagery could be better suited to detecting low severities of bio-control damage. Franke and Menz (2007) investigated multi-temporal wheat disease detection using high spatial resolution multispectral QuickBird images (2.5 m). A decision tree incorporating Mixture Tuned Matched Filtering (MTMF) and NDVI was employed which detected the early stage of crop infections. The overall classification accuracy of the first QuickBird scene was 56.8% whereas the
second scene was 88%. The use of high spatial resolution multispectral imagery was moderately suitable for the early detection of crop infections owing to the high misclassification rate observed in the early growth stages. It was recommended that the use of hyperspectral remotely sensed data be investigated. Qin and Zhang (2005) also concluded that higher spectral resolution is required in order to examine the capability of separating the light diseased plants from healthy plants. Qin and Zhang (2005) investigated the detection of rice sheath blight for in-season disease management using multispectral remotely sensed imagery. They applied a comprehensive field Disease Index (DI) to measure infection severity of the disease on broadband high spatial-resolution Airborne Data Acquisition and Registration (ADAR) (1 m) imagery. Results showed that it was difficult to discriminate healthy plants from light infected plants when DI < 20 because of their spectral similarities. It was clear that identification accuracy increased when infection reached medium-to-severe levels (DI > 35).

From the above, there is great potential to detect and map bio-control damage on AIAP species using moderate and high spatial resolution multispectral imagery. Spectral discrimination between healthy and severely damaged plants is evident; however, discriminating between healthy and early stage of infestations still poses a challenge when detecting low severities of bio-control damage. Advances in hyperspectral remote sensing, offers the high spectral resolution required which may assist in discriminating healthy and low severities of bio-control damage.

3.2. The detection and mapping of bio-control damage using hyperspectral remote sensing

Improvements in sensor technologies have resulted in hyperspectral sensors that capture data in hundreds of narrow contiguous (<10nm) bands in the VR, NIR and SWIR. The increased spectral dimensionality and sensitivity of hyperspectral data allow for the discrimination of spectrally similar but unique materials (Mutanga et al., 2009). Detecting bio-control damage using hyperspectral remote sensing would provide greater spectral discrimination of different severities of bio-control damage and infection especially between healthy and low damaged and infected plants. Hyperspectral field spectrometers can measure the reflectance of complex surfaces over small areas of known targets in situ either in the field or in a laboratory (Milton et al., 2009). Importantly, the main application of field spectroscopy is to perform a feasibility study to understand the spectral interactions of bio-control damage before upscaling to airborne or satellite remote sensing systems (Milton et al., 2009). Hyperspectral laboratory studies are central in determining key diagnostic features and spectral bands that will detect variable infestation levels for eventual inclusion on airborne and satellite remote sensors. Hyperspectral field spectrometers have been widely applied to capture and characterise the spectral profiles of different plant species under different stress conditions (Mirik et al., 2006; Song et al., 2011; Huang et al., 2012). However, few studies have investigated the use of handheld hyperspectral spectrometers to detect and discriminate bio-control damage on IAP species (Everitt et al., 2007; Everitt et al., 2012).
Everitt et al. (2012) used hyperspectral field reflectance data to assess bio-control damage (control, low, medium, high) on giant salvinia (*Salvinia molesta*). The analysis of variance and least significant difference comparison test results for both October and July revealed that generally the best bands for separating among treatments occurred in the green (505–595nm), red (605–635nm), red-near-infrared (695–745nm) edge and NIR (755–895nm) regions. Approximately 35 bands could be used to distinguish three to four treatments. Hyperspectral data allowed for the discrimination of healthy and low damaged giant salvinia using a relatively basic discrimination and band selection technique. This highlights the importance of utilizing high spectral resolution data for discriminating the early stages of bio-control infestations. Similarly Ge et al. (2011) carried out a multi-level defoliation assessment on saltcedar (*Tamarix ramosissima*) using hyperspectral data. The authors used red-edge positions and continuum-removed absorptions to discriminate four defoliation categories (healthy, newly defoliated, completely defoliated and refoliating). A classification tree was employed to integrate the red-edge positions and their derivatives with the central band depths of five continuum-removed absorptions. The overall classification accuracy was 87.5% with producer’s accuracy varying between 70 and 100% and user’s accuracy ranged between 77.78 and 100%. The authors concluded that single spectral dataset failed to separate the four stages but a combination of the two continuum-removed absorptions located in the red absorption (570-716nm) and the first water absorption (936-990nm) in the NIR improved the identification of defoliated canopies. This study combined the use of hyperspectral data with the implementation of a robust classification algorithm therefore achieving high classification accuracies. This study illustrates the potential of using hyperspectral field spectrometers to discriminate multi-level defoliation on AIAP species at the leaf level before upscaling.

With the inception of hyperspectral imaging spectrometers (HyMap, Hyperion) the extent of bio-control damage can potentially be mapped at a high spectral resolution. Hyperspectral imagery combines moderate to high spatial resolutions with the high spectral resolution required to map different severities of bio-control damage. It is evident that there is a paucity of research conducted on detecting and mapping bio-control damage on AIAP species using hyperspectral imagery. However, a study conducted by Fitzgerald et al. (2004) detected spider mite damage using AVIRIS hyperspectral imagery and a spectral mixture analysis. The procedure successfully distinguished between adjacent mite-free and mite-infested cotton fields. Similarly, Mewes et al. (2011) discriminated healthy and diseased wheat canopies using airborne HyMap imagery and a support vector machine classification algorithm. Results showed that with the original spectral resolution of the HyMap image, the highest classification accuracy could be achieved by using only 13 spectral bands with a Kappa coefficient of 0.59. This clearly illustrates the potential of mapping bio-control damage on AIAP species accurately using hyperspectral imagery. Further research needs to be undertaken that address information gaps in detecting different severities of bio-control damage and
in particular early stages of vegetation stress using hyperspectral imagery. Even though the utilization of hyperspectral imagery is advantageous, hyperspectral imagery are expensive to purchase on a regular basis and require highly skilled individuals to process; therefore imposing operational limitations. Peerbhay et al. (2015) looked at the semi-automated classification of alien invasive plants using the unsupervised random forest. Analysing hyperspectral data to extract meaningful information presents numerous challenges including data dimensionality and data redundancy. There is a need to investigate the advanced machine learning statistical techniques available to reduce data dimensionality so that bio-control damage classification procedures can be executed efficiently and repeatedly.

3.3. Potential classification techniques: machine learning algorithms

Classifying different severities of bio-control damage using hyperspectral data is a challenge owing to the high spectral dimensionality of the dataset. Over the past decade, machine learning algorithms have emerged as more accurate and efficient alternative to conventional parametric algorithms when processing high dimensional datasets. Machine learning algorithms are computationally efficient, offer higher classification accuracies and can process high dimensional datasets. Three popular machine learning algorithms namely Artificial Neural Networks (ANN), Random Forest (RF) and Support Vector Machines (SVM) algorithms have mainly been implemented for classifying vegetation stress (Table 1). However, to our knowledge no studies have investigated the potential of classifying different severities of bio-control damage on AIAP species using machine learning algorithms.

Table 1: Selected studies that have implemented machine learning algorithms for classifying insect induced stress and disease damage.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Algorithm</th>
<th>Data source</th>
<th>Type of stress</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singh et al.,  (2009)</td>
<td>Linear discriminant analysis, (LDA) Quadratic discriminant analysis (QDA), Malahanobis discriminant analysis</td>
<td>NIR hyper. imaging system</td>
<td>Insect damage: Sitophilus oryzae, Rhyzopertha dominica, Cryptolestes ferrugineus, Tribolium castaneum</td>
<td>Healthy wheat kernels and kernels damaged by S. oryzae, R. dominica, and C. ferrugineus were correctly identified with 91.7–100% accuracy using both LDA and QDA classifiers.</td>
</tr>
<tr>
<td>Liu et al., (2010b)</td>
<td>Artificial neural network (LVQ)</td>
<td>Analytical Spectral Device (ASD)</td>
<td>Disease damage: Rice glum blight disease, Rice false smut disease,</td>
<td>The overall accuracies of LVQ derived from the raw, inverse logarithmic, first, and second derivative reflectance spectra for the validation dataset were 91.6%, 86.4%, 95.5%, and 100% respectively.</td>
</tr>
<tr>
<td>Liu et al., (2010a)</td>
<td>Principal component analysis, Support vector machine</td>
<td>ASD</td>
<td>Disease damage: Nilaparvata lugens, Ustilaginoidea</td>
<td>Raw 96.55%, FDR 99.14%, SDR 96.55%</td>
</tr>
<tr>
<td>Authors</td>
<td>Techniques</td>
<td>Vectors</td>
<td>Damages</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------</td>
<td>------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Singh et al., (2010b)</td>
<td>Linear discriminant analysis, Quadratic</td>
<td>NIR hyper. imaging system, Digital color imaging</td>
<td>Insect damage: <em>Sitophilus oryzae,</em> <em>Rhyzopertha dominica,</em> <em>Cryptolestes ferrugineus,</em> <em>Tribolium castaneum</em></td>
<td>QDA: Had the highest accuracy. The accuracy for healthy kernels was 96.3%. The overall accuracy in classifying insect-damaged kernels using the top 10 features from 230 colour image features combined with NIR image features was 91.0–100.0%.</td>
</tr>
<tr>
<td>Singh et al., (2010a)</td>
<td>Linear discriminant analysis (LDA), Quadratic</td>
<td>NIR hyper. imaging system, Digital color imaging</td>
<td>Insect damage: Midge damage</td>
<td>LDA: The NIR hyperspectral image features and 10 colour image features gave the highest average accuracy of 95.3–99.3% in classifying healthy and midget damaged wheat kernels.</td>
</tr>
<tr>
<td>Ismail and Mutanga, (2011)</td>
<td>Random Forest</td>
<td>ASD</td>
<td>Insect damage: <em>Sirex noctilio</em></td>
<td>Random Forest classification using bands selected by the backward variable selection method had the lowest misclassification rate (6.14%).</td>
</tr>
<tr>
<td>Mewes et al., (2011)</td>
<td>Bhattacharyya distance forward feature search strategy, spectral angle mapper, support vector machine</td>
<td>HyMap</td>
<td>Disease damage: <em>Blumeria graminis</em></td>
<td>SVM: Using 13 bands at the original spectral resolution, the Kappa coefficient was 0.59.</td>
</tr>
<tr>
<td>Wang et al., (2012)</td>
<td>Linear discriminant analysis, Support vector machine</td>
<td>SWIR hyper imaging system</td>
<td>Disease damage: <em>Burkholderia cepacia</em></td>
<td>LDA: The overall classification accuracy was 80% when classifying healthy and sour skin-infected onions. SVM: The overall classification accuracy was 87.14% when classifying healthy and sour skin-infected onions.</td>
</tr>
<tr>
<td>Jin et al., (2013)</td>
<td>Discriminant analysis, Backward propagation neural network, Genetic backward propagation neural network (GA-BP), Support vector machine (SVM)</td>
<td>ASD</td>
<td>Disease damage: Cotton verticillium disease</td>
<td>GA-BP: The average accuracy was 85.7%. SVM: The average accuracy was 78.5%. The SVM was recommended as it is more stable and easier to implement.</td>
</tr>
<tr>
<td>Poona and Ismail, (2013)</td>
<td>Artificial neural network</td>
<td>Quickbird imagery</td>
<td>Disease damage: <em>Fusarium circinatum</em></td>
<td>The multilayer feed-forward neural network had an overall accuracy of 82.15%. Kappa of 0.65 optimal model Using the ANOVA as a pre-filtering technique the overall accuracy was 81.21%. Using the Random Forest variable importance as a pre-filtering technique the overall accuracy was 82.48%.</td>
</tr>
<tr>
<td>Adelabu et al., (2014)</td>
<td>ANOVA/backward feature elimination, Random Forest</td>
<td>ASD</td>
<td>Insect damage: Insect defoliation</td>
<td></td>
</tr>
</tbody>
</table>
ANN is a non-parametric classifier that makes no assumptions about the distribution of the data (Dixon and Candade, 2008; Song et al., 2012). ANN’s are computationally efficient, more resistant to noise and perform well with small training datasets (Song et al., 2012). However, the implementation of ANN’s is limited by the adjustment of the network parameters (network architecture, learning rate, and momentum). Different combinations of the model parameters lead to a large number of trials for network training with varying results (Shao and Lunetta, 2012). Poona and Ismail (2013) applied an ANN to detect *Fusarium circinatum* on *Pinus radiata* trees using QuickBird imagery. Several vegetation indices were calculated and incorporated into the neural network model. A multilayer feed-forward neural network showed high discriminatory power between healthy and infested tree crowns with an overall accuracy of 82.15% and Kappa of 0.65. Even though the study did not classify different stages of pitch canker the authors highlighted the importance of optimizing neural network model parameters with several runs in order for the model to attain good generalisation capabilities. Liu et al. (2010b) applied an ANN model to discriminate different fungal infection levels in rice panicles (*Oryza sativa* L.) using hyperspectral reflectance. A Learning Vector Quantization (LVQ) neural network classifier was employed to classify healthy, light, moderate and serious infection levels. Results showed that the overall accuracies of LVQ with PCA derived from the raw, inverse logarithmic, first, and second derivative reflectance spectra for the validation dataset were 91.6%, 86.4%, 95.5%, and 100% respectively, with corresponding Kappa coefficients of 0.887, 0.818, 0.939 and 1. The authors suggested that more studies need to explore the spectral response characteristics of rice crops under fungal stress of different levels in field conditions to develop a practical monitoring strategy. Despite this, the high overall accuracies illustrate the potential of discriminating different severities of bio-control damage using ANN.

RF is a tree based ensemble classifier that is capable of producing high classification accuracies (Gislason et al., 2006; Rodriguez-Galiano et al., 2012). RF is able to process high dimensional data sets efficiently and is robust against over-fitting the training data set (Breiman, 2001). Furthermore, RF is easier to implement than other ensemble classification methods as there are only two parameters to define. The classification results are not sensitive to the setting of the input parameters. In addition, RF provides a reliable estimate of error using the data that are randomly withheld from each iteration of tree development making it unnecessary to have an independent test data set (Breiman, 2001). It enables all the collected data to be used for training therefore potentially reducing the field work required (Lawrence et al., 2006). A study conducted by Ismail and Mutanga (2011) explored discriminating the early stages of *Sirex noctilio* infestation using classification tree ensembles and shortwave infrared bands. Results showed that the RF algorithm using bands selected by the backward variable selection method produced the lowest misclassification rate of 6.14%. Results from this study confirmed that RF is a robust and accurate method for classifying hyperspectral data in an application where (i) the number of samples is...
limited and (ii) classes have similar spectral characteristics. Further a study by Abdel-Rahman et al. (2014) explored detecting *Sirex noctilio* grey-attacked and lightning-struck pine trees using airborne hyperspectral data, RF and SVM classifiers. Results showed that when the most useful spectral bands as measured by RF were used, the overall accuracy improved from 74.50% to 78% for RF and from 73.50% to 76.50% for SVM. However the authors suggested the implementation of RF over SVM to classify hyperspectral data for several reasons including RF requires the optimization of two parameters only as well as the internal out of bag error rate of RF could be used for classification accuracy assessment when there are limited samples for independent accuracy assessments. In terms of bio-control damage detection, there is great potential in applying RF to discriminate spectrally similar damage classes owing to the high classification accuracies it achieves.

SVM’s is a popular technique for supervised data classification that provides higher classification accuracies than other pattern recognition techniques. SVM’s implements a classification strategy that exploits a margin based geometrical criterion rather than a purely statistical criterion (Melgani and Bruzzone, 2004). SVM’s adopts the method of structural risk minimization for class member discrimination which minimizes the classification error on unseen data without making prior assumptions on the probability distribution of the data (Mountrakis et al., 2011). This is advantageous as data acquired from remotely sensed imagery usually have unknown distributions (Mountrakis et al., 2011). SVM’s has high generalization capabilities with a relatively small number of training samples as well as robust against over-fitting and is less sensitive to the Hughes effect (Melgani and Bruzzone, 2004). However, the performance of the SVM algorithm is sensitive to the choice of kernel function and the setting of its associated parameters (Song et al., 2012). A study conducted by Liu et al. (2010a) discriminated rice panicles of different health conditions by applying a PCA and SVM classification on hyperspectral reflectance data. Results showed that the overall accuracies of the SVM classifications with principle components derived from the raw, first, and second order reflectance spectra for the testing dataset were 96.55%, 99.14%, and 96.55%, and the Kappa coefficients were 94.81, 98.71, and 94.82, respectively. Jin et al. (2013) investigated the application of hyperspectral remote sensing in identifying cotton verticillium disease severity. A wavelet transform was employed to extract the principal information and reduce the dimensions of the hyperspectral reflectance data. Subsequently, four identification models were built using discriminant analysis, back propagation (BP) neural network, genetic back propagation (GA-BP) neural network and SVM. Results showed that GA-BP produced the highest average accuracy of 85.7%. However, the authors recommended the use of the SVM model because it is more stable and better for practical application. In general the high overall accuracies when discriminating cotton verticillium disease severities illustrates the potential of discriminating different severities of bio-control damage using SVM.
From the above it is evident that there is great potential in implementing machine learning algorithms to classify hyperspectral data as they produce high classification accuracies. Further investigation would need to be conducted to assess and compare the performance of different machine learning algorithms when classifying different severities of bio-control damage on AIAP species. In addition, the optimization of input parameters and model efficiency would need to be assessed to determine the practicality of their regular application within operational procedures. Once operational procedures have been automated machine learning algorithms can be implemented efficiently to map bio-control damage.

4. Challenges mapping bio-control damage and future recommendations

Remote sensing is a powerful tool that can effectively replace traditional monitoring techniques to efficiently detect and map bio-control damage on AIAP species. Furthermore, remote sensing can potentially be utilised to ascertain the physiological status of AIAP species and the severity of bio-control damage on AIAP species. There are still numerous challenges present that would impede its widespread adoption by bio-control initiatives.

The cost associated with image acquisition and data processing is one of the key challenges that will limit the operational use of remote sensing within bio-control programmes in South Africa. Both multispectral and hyperspectral images are expensive to purchase and in some instances not feasible as a data collection method for some bio-control programmes. In addition, highly skilled personal are required to process the data which are expensive to employ permanently or on a contract basis. These costs will directly impact on the operational use of remote sensing to detect, map and monitor bio-control damage on AIAP species.

Assessing the photosynthetic capacity as a pre-visual indicator of bio-control damage on AIAP species is a major challenge. Chlorophyll fluorescence can be used as a proxy for photosynthetic capacity; however, it is difficult to detect owing to the complex nature of the reflectance spectrum. Future research should determine if chlorophyll fluorescence signal can be detected on AIAP species under bio-control and its relationship with reflectance spectra. Although various chlorophyll fluorescence indices have been developed using plants under heat and drought stress, their strength in quantifying chlorophyll fluorescence on AIAP species under bio-control needs to be determined. Importantly, the performance of spectral indices will have to be tested at the leaf and canopy level before upscaling to airborne and satellite sensors.

The utilization of multispectral data to detect and discriminate between healthy plants and low levels of bio-control damage is another research challenge. Detecting low levels of bio-control damage is critical in determining if bio-control agents are actively feeding. The low spectral and
spatial resolution of multispectral data, for example Landsat 7 (ETM+), is a limiting factor when discriminating between healthy plants and low levels of bio-control damage. However, multispectral imagery is available and accessible warranting its widespread use within bio-control initiatives in South Africa. Further research would need to be undertaken to investigate the potential of utilizing higher spatial resolution imagery and more robust classification algorithms to discriminate healthy and low bio-control damage in order to improve classification accuracies.

Hyperspectral data can be used to discriminate different severities of bio-control damage. However, the high data dimensionality and strong autocorrelation between adjacent bands introduces significant limitations when trying to extract meaningful information on bio-control damage. Future research should focus on identifying key bands that discriminate different severities of bio-control damage which is essential in selecting bands for inclusion on airborne/satellite sensors and developing spectral indices. The performance of numerous band selection algorithms would need to be compared to identify the most efficient and effective algorithm for implementation. Furthermore, developed spectral indices would need to be tested to determine its performance in classifying different severities of bio-control damage at the laboratory, field and airborne level.

Despite the challenges mentioned above, this paper has presented the potential of remote sensing to detect, map and monitor bio-control damage on AIAP species. It is evident that further research needs to be conducted to detect bio-control damage using both multispectral and hyperspectral remote sensing to bridge current technical and information gaps. Once remote sensing techniques have been fully developed and tested they can be operationally adopted to detect, map and monitor bio-control damage on AIAP species successfully.

References


Liu, ZY, Shi, JJ, Zhang, LW & Huang, JF 2010a, ‘Discrimination of rice panicles by hyperspectral reflectance data based on principal component analysis and support vector classification’, *Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)*, vol. 11, pp. 71-78.


Qin, Z & Zhang, M 2005, ‘Detection of rice sheath blight for in-season disease management using


