(Option I) A disease-specific spectral vegetation index to detect the invasive plant pathogen myrtle rust on economically important lemon myrtle trees. (Option II) New method to detect plant pathogens; a disease-specific spectral index to predict myrtle rust on lemon myrtle trees.

Rene Hans-Juergen Heim1, Ian Wright1, Jens Oldeland2, Andrew Allen1

Macquarie University1

Hamburg University2

18.07.2017

**Abstract: (Limit:300 words)** In 2010, the invasive fungal pathogen, *Austropuccinia psidii* was introduced in Australia. Native ecosystems are severely affected, with highly susceptible species likely to go extinct due to recurring infections with myrtle rust, the disease *A. psidii* causes. Also the Australian economy is affected, as the hardwood industry (e.g. Eucalyptus and Corymbia) is currently threatened by myrtle rust incursions. The lemon myrtle industry, producing food flavouring, essential oils and teas, is already directly impacteda and has been suffering yield losses up to 70 percent per year. These losses lead to fungicide application and increased costs and stopped the industry from expanding into lucrative new markets. Yet, detecting and monitoring disease outbreaks is currently only possible by eye, which is not just costly but also slow, labor intensive and often subject to human bias. Over the last 25 years, spectral vegetation indices (SDIs) have been designed to assess variation in biochemical or biophysical traits of vegetation. Application of SDIs in e.g. agriculture can provide an automatic and objective alternative to visual disease assessment. However, diagnosis of specific diseases based on common spectral vegetation indices is not possible so far, since these indices lack disease specificity. To meet the needs of modern plant disease detection, spectral indices must detect disease intensity and incidence before symptoms become visible, discriminate spectral patterns between different pathogens, host-species and abiotic stresses and estimate disease severity. Here we present a novel, specific spectral disease index (SDI) that can predict myrtle rust infections on lemon myrtle (*Backhousia citriodora*) plantations with an overall accuracy of 89%. We also compared our lemon myrtle-myrtle rust index (LMMR) performance to three other, in plant disease detection commonly applied indices (PRI, MCARI, NBNDVI) and found that they only could predict a myrtle rust infection with accuracies of 36%, 31%, 42%, respectively. If the LMMR can be positively validated on independent datasets from similar and different host-species, it could enable land managers to reduce their fungicide application by making swift decisions on which specific areas they have to treat on their plantation instead of applying their fungicides untargeted and wasteful.

**Journal: Remote Sensing of the Environment**

**HIGHLIGHTS** Include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). You can view example Highlights on our information site.

1. Disease-specific spectral index to detect myrtle rust on lemon myrtle trees.
2. Index enables plantation managers to spot local infection levels.
3. Fungicides can be applied targeted and therefore reduce costs and environmental impact.
4. Method can be adapted for other pathosystems.

# I. Introduction

Rust fungi and other plant pathogens are affecting humans and their environment by damaging plants and their products on which we depend for clothing, housing and, most importantly, food. Outbreaks of rust fungi may result in extensive damage to agricultural and forestry crops, as seen when a new, highly virulent strain of *Puccinia graminis* destroyed tens of thousands of hectares of wheat crops in southern Europe (Bhattacharya, 2017). This study focuses on the rust fungus *Austropuccinia psidii* (Sphaerophragmiaceae, Pucciniales). In Australia, *A. psidii* causes a disease commonly known as “myrtle rust” and is an obligate biotroph and pathogenic organism in the highly diverse phylum Basidiomycota (Helfer, 2014). In contrast to most other rust diseases, myrtle rust has the potential to infect hundreds of different species, escalating the potential consequences of infection. Several native species have already been severely damaged by myrtle rust in the wild (Carnegie et al. 2016). Consequently, the New South Wales scientific committee of the Department for Environment and Heritage has acknowledged myrtle rust to constitute a major threat to the native Australian environment and the Myrtaceae, listing it as a Key Threatening Process (NSW Scientific Committee, 2011).

However, the impact of myrtle rust in Australia has not been limited to native ecosystems, with the industries reliant on Myrtaceae also being affected. Loss of commercial varieties and trade restrictions, in addition to increased reliance on fungicides, have severely impacted upon the nursery and garden industry. The expanding lemon myrtle industry has also been significantly affected (Doran et al, Lee et al. (RIRDC report). Leaves are commercially harvested to produce lemon-flavoured herbal teas, culinary herbs or lemon-scented essential oils used for food flavouring and in personal care products. Cultivars of B. citriodora currently in use are moderately to highly susceptible to myrtle rust. Rust-affected leaves of B. citriodora are unsuitable for its main uses and the application of fungicides to control the disease is undesirable as the market demands a clean, organic product. A total farm gate value at between AUD7 million and AUD23 million for dried leaf and essential oil is estimated (RIRDC 2012). Therefore, the industries reliant on lemon myrtle are in urgent need of measures to reduce the use of fungicides or rust-resistant cultivars. Of particular concern, there are increasing reports of susceptibility within the eucalyptus sensu lato (Pegg et al., 2014; Potts et al., 2016). Seemingly, myrtle rust has the potential to impact on the chief genera used for hardwood forest industry in Australia. This was found after commercial plantations of *Eucalyptus globulus* in Brazil have suffered reduced growth and yield loss as a result of myrtle rust incursions (Alfenas et al., 2003). Until now, the detection of myrtle rust infections is only possible by human assessors, which is a slow, labour intensive process and results in costly and untargeted fungicide application. Fortunately, many researchers have shown the capabilities of remote sensing techniques and spectral sensors in the area of agriculture and crop production which has been revently reviewed by Mulla (2013). Also in the field of plant disease detection, Delalieux et al. (2007), Mahlein (2010, 2013), and Lopez&Lopez (2016) have proven the potential of spectral sensor systems for the detection of fungal diseases.

Spectral sensors take advantage of vegetation surfaces naturally reflecting, transmitting and absorbing fractions of the occurring light. Such interactions between light and vegetation are strongly determined by their structural, physiological and biochemical features. Chlorophyll pigments e.g. absorb specific wavebands of the electromagnetic spectrum to generate biomass via photosythesis. As some wavelengths are absorbed by pigments and others are transmitted though the leaf, only a fraction of the incoming light can be reflected. Also the cellular structure of a leaf's cell layers can influence the direction and intensity of reflected wavebands. Variation in those spectral reflectance patterns can provide insight into a plants nutritional and health condition and is often recorded using sensors with either moderate multispectral (>10 nm) or high hyperspectral (< 10nm) resolution.(Asner, G.P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment. 64: 234–253.). On the one hand, higher spectral resolution is useful to capture spectral information in greater detail but on the other it also results in an array of possibly redundant spectral wavebands.

For some specific applications, such as the assessments of plant pathogens in agriculture, not all of this information is relevant and redundant wavebands can be discarded (Mahlein, 2013). The remaining relevant wavebands can be summarized as e.g. ratios, sums or linear equations to simplify their interpretation (Jackson and Huete, 1991). When these equations are used to reflect vegetation features, they are referred to as spectral vegetation indices (SVIs). A typical example is the widely used Normalized Difference Vegetation Index (NDVI) which can be used to detect the presence of live green vegetation (e.g. J. R. G. TOWNSHEND & C. O. JUSTICE (1986) and Tucker, 2005). Other SVIs are used to inform on more complex situations than just assessing whether a plant is green or not. For instance, the physiological reflectance index (PRI) was developed by Peñuelas et al. (1994) to follow diurnal changes in xanthophyll pigments and photosynthetic rates. Others applied already existing indices to map plant communities (Oldeland et al. 2010) or assess agricultural crop characteristics (Thenkabail, Smith, and De Pauw 2000). Especially in agriculture, SVIs have been found useful to detect pathogens. Nowadays, spectral sensors and suitable indices are applied to develop automated and objective detection systems as an alternative to visual disease assessment (Bock et al. 2010). Visual estimation has become more accurate and reliable due to set guidelines and standards. However, an assessment will always be subject to a assessors experience and can be affected by temporal variation. This variation causes significant variability between assessors and makes repeated assessments inaccurate (Bock et al. 2010; Nutter et al. 2010). Automated methods, with high sensitivity, specificity and reliability are necessary to improve disease detection beyond that of visual estimation processes (Mahlein 2015).

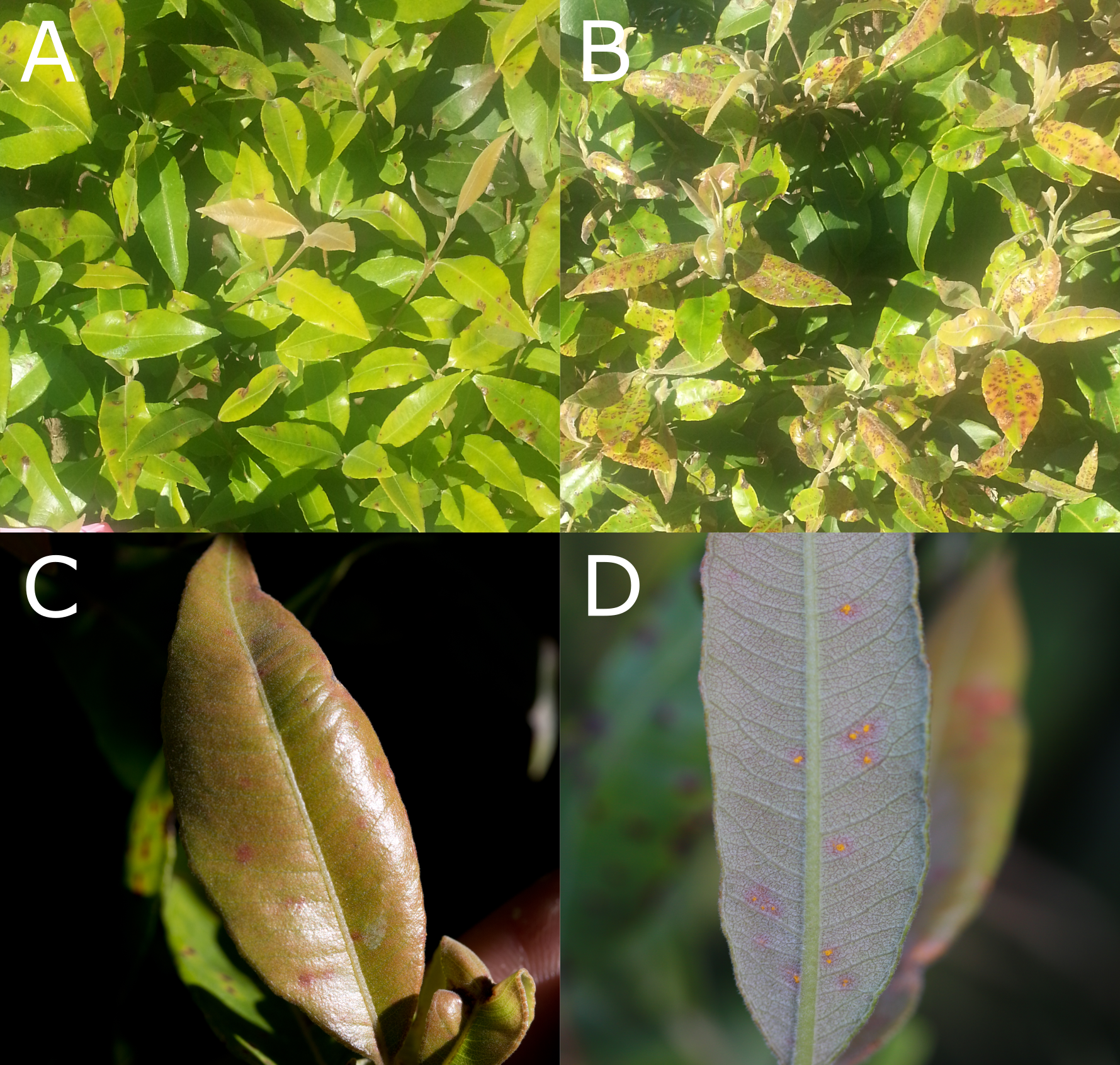
While detection systems are being developed, various challenges for spectral plant disease diagnostics have been discussed (Mahlein 2015). A system that can be acknowledged as an improvement over human estimates must (i) detect disease intensity and incidence (see (Bock et al. 2010) for definitions) before symptoms become visible, (ii) discriminate spectral patterns between different pathogens, host-species and abiotic stresses and (iii) estimate disease severity (Mahlein 2015). In an attempt to meet these criteria, Ashourloo, Mobasheri, and Huete (2014) compared the performance of multiple SVIs to classify wheat leaf rust on two cultivars of wheat. Another approach by Devadas et al. (2009), added a level of complexity by assessing the classification performance of ten SVIs between yellow (stripe), leaf and stem rust. Then, Mahlein (2010) compared three common SVI's on three sugar beet diseases under controlled conditions and noted that single SVIs lack the potential to differentiate among diseases. It was also found that the use of SVI combinations seems to be highly promising to improve disease detection and assignment versus biotic and abiotic plant stress. It was soon argued that SVIs were originally designed for other purposes (e.g. assessing canopy greenness or a plants content of anthocyanins and chlorophyll pigments) and therefore lack disease specificity (Mahlein, 2013). Consequently, Mahlein (2013) designed disease-specific spectral disease indices (SDI's) for the three sugar beet diseases she investigated earlier (Mahlein, 2010). The optimized disease indices were tested for their ability to detect and to classify healthy and diseased sugar beet leaves. With a high accuracy and sensitivity healthy sugar beet leaves and leaves, infected with Cercospora leaf spot, sugar beet rust and powdery mildew were classified (balanced classification accuracy: 89%, 92%, 87%, 85%, respectively).

The approach by Mahlein (2013) seems to be highly promising to also detect myrtle rust and no research has been undertaken to develop a SDI. To further advance research towards automated, objective detection system and to reduce costly and untargeted fungicide application, the aim of this study, was to test i) whether it is possible to select relevant wavebands and design a spectral disease index specific to *Backhousia citriodora* (lemon myrtle) trees infected with myrtle rust and ii) compare such a SDI with common SVIs for their accuracy to discriminate healthy and infected trees on a plantation.

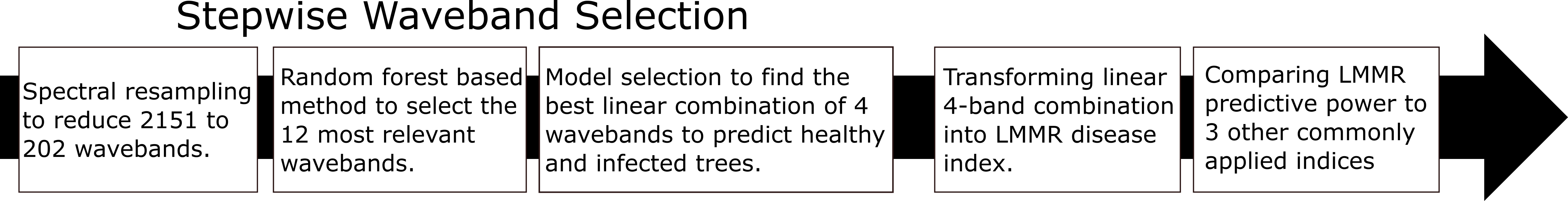
# II. Methods and Data

## Data Collection

Leaf spectral data was collected on a lemon myrtle (*Backhousia citriodora*) plantation in northern New South Wales, Australia (28.6911090 S, 153.295480 E) on myrtle rust (*Austropuccinina psidii*) "Infected" trees (Figure 1 B, C, D) and trees that had been "Treated" with fungicides (Figure 1 A). Spectral reflectance between 350 nm and 2500 nm, was measured using a portable non-imaging spectroradiometer (Spectral Evolution PSR+ 3500) with a spectral resolution of 3 nm up to 700 nm, 8 nm up to 1500 nm and 6 nm up to 2100 nm. A leaf clip holder with a 3-mm sample area, a built-in reflectance standard and a separate 5-watt light source (ILM-105) was used to take measurements. For more details, please refer to Heim *et al.* (2017). In this previous study, we showed that "Naive", "Treated" and "Untreated" *Backhousia citriodora* trees can be classified with high accuracy based on a broad set of 202 wavebands. Here, we refine this set of wavebands and develop a SDI that is easy to interpret and based on 4 wavebands only.



**Figure 1** A representation of fungicide treated (A) and untreated (B,C,D) *Backhousia citriodora* leaves that have been assessed on the plantation. Trees that had been treated with fungicides had only few necrotic lesions on older leaves (A). If young leaves were infected, the fungus could not spread far across the leaf surface as fungicide treatments had been done regulary. Leaves that were not treated were largely covered with dark necrotic lesions as the fungus was not contained (B). These lesions originated from fresh infection sites (C,D) that had been left untreated. A closeup of the adaxial (C) and abaxial (D) untreated leaf surfaces shows that myrtle rust is producing yellow urediniospores mostly on the abaxial surface (D). These fresh infection sites are also visible on the adaxial (C) surface and visible through pink/red flecking. On treated leaves, these infection spots were not present.



**Figure 2** Analysis workflow to visualize each waveband selection step (Box 1-3), the buidling step of the new lemon myrtle/myrtle rust index (LMMR) and a final step to compare the predictive performance of the LMMR to other suitable indices to predict myrtle rust infections on lemon myrtle trees.

## Analysis pipeline

Post data collection, all analyses were conducted using the R statistical platform (R Core Team, 2016) using several add-on packages (see below). The full analysis, including figures and tables, can be reproduced using code and data archived at <https://github.com/ReneHeim/>.... (doi to be provided on acceptance).

## Stepwise Waveband Selection

The spectral dataset, used in this study, contains spectral reflectance profiles of 236 fungicide treated and 228 untreated lemon myrtle trees and was collected in a previous study (Heim et al., 2017). Each profile was based on 2151 spectral wavebands, meaning that it contained more predictive variables than observations. This state is referred to as 'high dimensionality' and can cause multicollinearity effects (CITE). Usually, spectral vegetation indices, based on hyperspectral data, are generally designed using only two to four wavebands (A.-K. Mahlein et al. 2013), therefore it was necessary to remove redundant wavebands and account for multicollinearity effects. To design a lemon myrtle/myrtle rust index (LMMR) a three-step waveband selection was applied.

In a first step, spectral resampling was applied to reduce the original spectral resolution of 3 nm – 8 nm (2151 predictor variables) to a new spectral resolution of 10 nm (202 predictor variables). Spectral resampling was carried out using the 'prospectr' package (Stevens and Ramirez–Lopez 2014) in R (R Core Team 2016). One requirement for the LMMR was that output should be easy to interpret. Thus the original reflectance data was log transformed to make use of the following logarithm rule (Equation 1) to be able to denote the linear combination of wavebands as a ratio index. Therefore the LMMR estimates the probability of a tree being healthy or infected as 100% or 0% chance.

In a second step, a random forest based selection was applied to further reduce the 202 wavebands down to 12 wavbands. Here, the R package 'VSURF' was used (Genuer, Poggi, and Tuleau-Malot 2015) which is suitable for regular and high-dimensional datasets.

The third step was to use this refined set of twelve wavebands and perform an exhaustive model selection procedure that elimated inferior models in another stepwise approach. All models were calculated based on a subset of four wavebands, selected from the set of twelve bands. This stepwise selection was performed repeatedly until all possible waveband combination were built into a model. The R package 'glmulti' (Calcagno and Mazancourt 2010) was used because the number of wavebands to be included in the model can be tuned. Models were ranked according to the small-sample corrected Akaike Information Criterion (AICc). Burnham and Anderson (2002) suggested that if the ratio of sample size (n) and model parameters (k) is sufficiently large (e.g. ), then AIC and AICc are similar and will strongly tend to select the same model. Otherwise AICc will correct for small sample sizes. Thus, no negative effect is included when using AICc.

## Building the LMMR

After the most relevant four wavebands had been selected, they were used as input predictor variables in a binomial generalized linear model to determine optimal model parameters. Eventually, these parameters built the foundation to generate the LMMR index. To put the performance of the newly developed index into perspective, it is compared with other spectral vegetation indices (Table 1)commonly applied to detect pathogens in vegetation (Ashourloo, Mobasheri, and Huete 2014; Devadas et al. 2009; A.-K. Mahlein et al. 2013). Indices have been selected according to the biological processes they indicate and whether these processes could be linked with physiological changes caused by myrtle rust. It is known that urediniospores of rust fungi contain carotenoids and melanin-like pigments that change their colour to brown-orange-yellow (A.-K. Mahlein et al. 2013; Trocha, Daly, and Langenbach 1974). Changes in pigments can be detected, amongst others, with the physiological reflectance index (PRI) and the modified chlorophyll absorption in reflectance index (MCARI). Also the structural integrity of the mesophyll cells is reduced (Morin, Talbot, and Glen 2014) which is usually reflected in the near-infrared region (CITE). Therefore, the narrow-band normalised difference vegetation index (NBNDVI) could reflect this variation as it measures the ratio between the near-infrared and visual region. Each index (Table 1) was used as a predictor variable in a logistic regression and its performance to discriminate "Treated" and "Infected" trees was assessed based on four common accuracy measures, the overall accuracy (OA) and the Kappa statistic, sensitivity and specificity (Kuhn and Johnson 2013).

**Table 1** A brief summary of vegetation indices commonly used to detect plant stress caused by pathogens.

|  |  |  |  |
| --- | --- | --- | --- |
| SVI Name | Formula | Bioindicator | Reference |
| Physiological reflectance index | PRI=(R570-R531/R570+R531) | Stress indicator by indicating high levels of xanthophyll activity | (Penuelas et al., 1994) |
| Modified Chlorophyll Absorption in Reflectance Index | MCARI=(R701-R671)-0.2(R701-R549)]/(R701/R671) | ??? | (Daughtry et al., 2000) |
| Narrow-band normalised difference vegetation index | NBNDVI= (R850-R680/R850+R680) | Includes most key pigments.Similar to NDVI but suited to narrow-band sensors. Increases with leaf area index (LAI) and photosynthetically active radiation (PAR) or biomass (PAB). | (Thenkabail et al. 2000) |

## LMMR compared to other indices used for disease prediction

OA reflects the agreement between the reference and predicted classes and has the most direct interpretation. However, it does not provide information about the origin of an error (Kuhn and Johnson 2013). Here, an additional metric, the Kappa statistic (Cohen 1960) is useful. Kappa can take on values between –1 and 1; a value of 0 implies no agreement between the observed and predicted classes, while Kappa of 1 indicates perfect concordance of the model prediction and the observed classes. Landis and Koch (1977) first defined the following standards for the strength of agreement for Kappa coefficients: 0 = poor; 0.01–0.20 = slight; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = substantial; and 0.81–1 = almost perfect. A negative Kappa indicates that the prediction is in the opposite direction of the truth. Further insight into the index performance to separate infected and uninfected trees can be gained by evaluating the sensitivity and specificity of the classification model(Kuhn and Johnson 2013). The sensitivity of the model is the rate that the event of interest (e.g. infection) is predicted correctly for all samples experiencing the event (e.g. an infection). The specificity is defined as the rate that nonevent samples are predicted as nonevents.

# III. Results

This study designed a novel disease index to predict the probability of a *Backhousia citriodora* tree being infected with myrtle rust or being treated with fungicides and consequently uninfected. Therefore, a spectral dataset, containing 2151 wavebands was reduced in a 2-step waveband selection procedure to the 4 most relevant wavebands only. This new index could outperform other commonly applied indices with an OA of 89% (Table 2).

## Waveband Selection

The spectral dataset in this study originally contained 2151 wavebands. After a spectral resamling was performed to reduce the number of wavebands to 202, the total number of wavebands could be further reduced to the following twelve bands: 725 nm, 735 nm, 715 nm, 745 nm, 545 nm, 555 nm, 2145 nm, 1455 nm, 1485 nm, 1475 nm, 2175 nm, 2125 nm. In a next step, these twelve wavebands could be further reduced to the following four wavebands: 545 nm, 555 nm, 1485 nm, 2175 nm.

## Compiling the lemon myrtle-myrtle rust (LMMR) index

The LMMR utilisis these four wavebands that provide, when combined, the greatest predictive power to determine the probability of a lemon myrtle tree being "Infected" or fungicide "Treated". As the reflectance data was log-transformed before relevant wavebands were selected, the linear combination of wavebands and their modelled coefficients can be converted into a ratio index (Equation 2).

## Classification performance of the LMMR index

To put the predictive performance of the LMMR into context, its capability of predicting myrtle rust infections on lemon trees was compared to the performance of other commonly applied indices for the same task. It was found that the LMMR classification accuracy (OA=0.89, Kappa=0.78) exceeded the performances of the PRI (OA=0.64, Kappa=0.28), the MCARI (OA=0.69, Kappa=0.37) and also the NBNDVI (OA=0.58, Kappa=0.15).

**Table 2** Accuracy assessment of the LMMR index compared to other spectral indices that have been applied to detect pathogen related stress. Accuracy is assessed for the capability of an index to discriminate between fungicide Treated and Infected plants. Negative values do not represent that the according index has less than zero predictive power but that it classifies treated and untreated in the other direction of the LMMR.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | OA | sensitivity | specificity | Kappa |
| LMMR | 0.89 | 0.87 | 0.91 | 0.78 |
| PRI | 0.64 | 0.62 | 0.66 | 0.28 |
| MCARI | 0.69 | 0.68 | 0.69 | 0.37 |
| NBNDVI | 0.58 | 0.42 | 0.72 | 0.15 |

# IV. Discussion

**SECTION 1 - Report that we could create a new index that is capable of predicting the probability of a tree being infected.** This study designed a new disease-specific vegetation index, that predicts the presence of an invasive, fungal pathogen, myrtle rust (*Austropuccinia psidii*) on lemon myrtle trees (*Backhousia citriodora*) with higher accuracy than commonly applied vegetation indices. The lemon myrtle-myrtle rust index (LMMR) utilizes only four wavebands that have been selected from a hyperspectral dataset that originally contained 2151 wavebands. Such an array of wavebands often encompasses redudant spectral information and its processing requieres extensive computational effort. The LMMR can be swiftly applied on lemon myrtle plantations by feeding it the relevant wavebands and it is easy to interpret as it generates the disease probability being between 0% and 100%.

**SECTION 2 - Discuss how the LMMR performs compared to common indices used by others**

* The LMMR predicts the presence of myrtle rust with an overall accuracy of 0.89 and Kappa of 0.78. Comparing these accuracy metrics with other commonly applied vegetation indices for disease detection (Table 1) shows that it generates twice as accurate predictions (Table 2). This was expected, as we selected the most relevant wavebands for the classification of our pathosystem and therefore designed an index that is specific, as opposed to commonly applied and non-specific vegetation indices.
* That specific-disease indices have greater predictive power for specific disease classification problems was also found by Mahlein (2013). After Mahlein (2010) explored the potential of eight common vegetation indices (e.g. Normalised Difference Vegetation Index (NDVI), Anthocyanin Reflectance Index (ARI) and modified Chlorophyll Absorption Integral (mCAI)) to predict three sugar beet diseases (Cercospora leaf spot, powdery mildew and sugar beet rust) and found that each index alone might be able to predict a single disease but lacks the potential to differentiate among diseases, she concluded that the predictive power could be increased by combining the prediction of each individual index. Because common vegetation indices are neither disease- nor stress-specific (Mahlein, 2010) and only respond to variation of general physiological and structural plant processes (e.g. pigment content or leaf structural integrity), Mahlein (2013) developed disease-specific indices for the three sugar beet diseases and provided evidence that disease-specific indices will improve and simplify plant disease detection based on hyperspectral data. In a very similar approach to our study, Mahlein (2013) selected significant and most relevant wavelengths, describing the impact of cercospora leaf spot, powdery mildew and sugar beet rust on sugar beet leaves, from a hyperspectral dataset. To develop disease-specific indices for the detection of these diseases, the best weighted combination of a single wavelength and a normalized wavelength difference (WHAT IS THIS?? NEED TO READ, Jens??) was exhaustively searched and all possible combinations were tested. Mahlein's (2013) indices were optimized to classify healthy leaves and leaves infected with cercospora leaf spot, powdery mildew and sugar beet rust. The indices could do this with a high balanced classification accuracy of 89%, 92%, 87%, 85%, respectively. This time, it was possible to distinguish the three diseases and also estimate their severity because Mahlein's disease-specific indices were characterized by a high sensitivity and specificity for the detection and identification of the different foliar diseases of sugar beet.
* In a more recent study, similar to Mahlein (2010), Ashourloo, Mobasheri, and Huete (2014) tested another set of 22 common vegetation indices to detect disease symptoms of wheat leaf rust at leaf level. They added another level of complexity by staggering disease severity in multiple levels between 1% and 100%. Most indices could not satisfy their expectations, especially when disease severity was low. However, the indices NBNDVI, NDVI, PRI, GI, and RVSI showed accuracies of >60% even for a disease severity between 1% and 5%. Ashourloo, Mobasheri, and Huete (2014) concluded that spectral vegetation indices are valuable, as Mahlein (2010, 2013) did, because they reduce data dimensionality and increase the speed of data processing. However, it would also be necessary to consider more carefully the capability of spectral vegetation indices for the detection of plant diseases, especially for the diseases with different symptoms (Ashourloo, Mobasheri, and Huete 2014).
* According to the conclusions of these previous studies (Mahlein 2010, 2013 and Ashourloo, 2014) it is coherent to develop indices that are specific to certain pathogens and their hosts. Therefore, we designed the LMMR that is specific to the the pathogen **A. psidii** and, for now, to one of its hosts, **B. citriodora**. When comparing the LMMR we found that it outperformes common vegetation indices (Table 1), that had been applied in e.g. Ashourloo, Mobasheri, and Huete (2014) and Mahlein (2010) by far. Consequently, we can only recommend to generate more hyperspectral datasets and further research efforts to refine them and develop disease-specific spectral indices.

**SECTION 3 - Discuss why certain wavelength are selected**

* As can be inferred from a review about plant disease detection using sensor systems (Mahlein, 2015), the development of disease-specific indices requieres a fundamental understanding about the molecular, physiological and biochemical, pathogen-related processes that can be observed in diseased host plants. To be critical about the selected wavebands we need to understand the underlying biological processes that are linked to those wavebands (Mahlein, 2010). Our waveband selection found four (545 nm, 555 nm, 1485 nm, 2175 nm) wavebands that are most specific to myrtle rust infections on lemon myrtle trees. As one of Mahlein's (2013) investigated diseases, to discriminated whether sugar beet plants were generally healthy or infected, was sugar beet rust (caused by **Uromyces beticola**), it is reasonable to compare the relevant wavebands she found to those found for our rust pathogen. Mahlein (2013) found that the wavebands 513 nm, 520 nm, 534 nm, 570 nm, 584 nm, 698 nm, 704 nm and 734 nm to be important for her study. Specifically the response at 534 nm was of high relevance. According to Gamon et al. (1992) reflectance at 531 nm can detect the interconversion of the xanthophyll cycle pigments. Since xanthophyll cycle pigments are regulatory pigments linked to PSII light use efficiency, reflectance indices incorporating reflectance next to 531 nm could provide an indicator of photosynthetic function (Gamon et al., 1992; Rascher et al., 2010). Assuming that the waveband at 545 nm, found in our study, can be regarded as "close to 531 nm", it could indicate that **A. psidii** is causing changes in the xanthophyll cycle. Also the wavebands close to 550 nm are interesting to be compared to those found our study. Myrtle rust infections cause wounds that are visible as red/pink flecking, it is likely that wavebands at 545 nm and 555 nm respond to changes in anthocyanin content as has been found by Gitelson (2001). Anthocyanins are water-soluble vacuolar pigments of higher plants abundant in juvenile and senescing plants. Very often, in plant leaves, significant accumulation of anthocyanins is induced as a result of a number of environmental stresses such as strong light, UV-B irradiation, low temperature, drought, wounding, bacterial and fungal infections, nitrogen and phosphorus deficiencies, certain herbicides and pollutants. As the wavebands found in Mahlein (2013) and this study are in close proximity but not completely superposable, we assume that this is caused by specific interactions between incident light, pathogen and host.
* Another study by Delalieux et al. (2007), discriminated healthy from *Venturia inequalis* infected plants. At early stages of *V. inequalis* infections they found that wavebands in the SWIR region (1350 nm – 1750 nm and 2200 nm – 2500 nm) especially contributed to the discriminatory power of their models while in later stages wavebands in the VIS region gained importance.The shortwave-infrared region is linked to changes in water content (Peñuelas & Filella, 1998) caused by air humidity or water loss from lesions (Delalieux et al., 2007). In this study, many lesions were present on infected leaves, few on uninfected and almost no lesions on healthy leaves , providing a possible explanation for the selection of important features (1485 nm, 2175 nm) within the shortwave-infrared regions. While it is reassuring that the wavebands in our study have been found reasonable in others, there is still the need to use the LMMR index and compare myrtle rust infections to other symptoms on lemon myrtle trees caused by stressors.
* Further, it is known e.g. that urediniospores of rust fungi contain carotenoids and melanin-like pigments that change their colour to brown-orange-yellow (A.-K. Mahlein et al. 2013; Trocha, Daly, and Langenbach 1974). Also, it has been found that the structural integrity of the mesophyll cells is reduced due to entering of myrtle rust hyphae into the mesophyll layer of a leaf (Morin, Talbot, and Glen 2014). An array of research discussed which wavebands and their ratios might be the most effective to predict changes in e.g. pigments or leaf structure. Highlighting all of them would exceed the scope of this study but it can be mentioned that Thenkabail (2000) summarizes multiple findings in his book. For carotenoids it was shown that in different plant species, high photon flux induces small changes of relectance near 530 nm, attributable to the transformation of violaxanthin cycle xanthophylls.

**SECTION 4 - limitations and future directions**

* Disease-specific vegetation indices are regarded as a promising tool in plant disease detection as they are characterized by a high sensitivity and specificity for the detection and identification of the different foliar diseases (Mahlein, 2013). However, only few studies developed such indices (Mahlein, 2013) and none of them were applied in a plantation setting. Disease-specific vegetation indices requiere spectral reflectance data as input variables and therefore do not assess plant physiological parameters directly. They record a spectrum of signals where each signal is attributed to different biochemical, physiological or structural plant characteristics. Measurement quality can vary with varying leaf inclination, illumination, and be influenced by other abiotic or biotic stressors (Mahlein, 2015). Therefore, careful validation is necessary to confirm that the LMMR is specifically picking up the reflectance signal which is caused by infections through myrtle rust and not stress signals caused by other stressors. This can be achieved by testing the LMMR on independent spectral datasets from lemon myrtle trees. However, as hundreds of other plant species are jeopardized by myrtle rust, it could also be interesting to test the LMMR on spectral datsets from other host species because it is possible that it also detects myrtle rust infections on them.
* It further needs to be considered that the LMMR was developed based on a hyperspectral dataset that was recorded at leaf level. If the LMMR would be tested with data that was collected with aerial vehicles, the level of accuracy might change due to changes in illumination and recording geometry. To reduce the impact of different illumination, topography, crop variety or sensor specific effects, the wavelength differences could be normalized (Jackson, 1986;Lillesand & Kiefer, 2000; Lyon et al., 1998)

**SECTION 5 - Relevance and Conclusion**

Testing the non-invasive LMMR on our spectral dataset, resulted in a high specificity and sensitivity for detecting myrtle rust on lemon myrtle trees while discriminating them from lemon myrtle trees that had been treated with fungicide. The LMMR is the first disease-specific index to detect myrtle rust and, so far, only been tested on lemon myrtle trees. To be able to effectively utilize optical sensor data for diagnosis and disease detection, certain goals have been set: (i) the detection of a disease at early points in time, (ii) the differentiation among different diseases, (iii) the separation of diseases caused by abiotic stresses, and (iv) the quantification of disease severity (Mahlein, 2015). The LMMR needs to be developed according to these factors but already provides the advantage of delivering disease probability estimates between 0% and 100% which simplifies the interpretation for users. Based on the relevant wavelength found in this study, a simple and cheap sensor that only measures vegetation reflectance at these wavelength can be configured. This sensor could instantaneously calculate the LMMR value, and give the estimation of disease severity in seconds. It needs to be stressed that further work is needed to test a potential LMMR sensor in combination with different platforms (e.g. unmanned aerial vehicles or tractors) to turn its application into praxis. By doing such tests in close collaboration with land managers the LMMR could be a precise, reproducible and time saving method to detect myrtle rust on lemon myrtle trees on the leaf, canopy and field scale. CONCLUSION GEFAELLT MIR SEHR GUT!

**##############MORE IDEAS!!! NOT RELEVANT FOR REVIEW##################** \* Why might the four selected features be more important than the others of the set of twelve? + 720-750 nm RedEdge according to Gitelson in Thankabail (C6) found that indices based on relectance in the red edge (700–740 nm) region [18,19] were much better indicators of chlorophyll content than some of the more commonly used indices. The emission spectrum of ChlF is characterized by two broad peaks between 600 and 800 nm,centered at ∼685 nm (Red) and ∼740 nm (Far Red). For a given chlorophyll content, ChlF decline when plants are not under environmental stress; and, in general, ChlF and photosynthesis rate are negatively correlated.Relectance-based approaches examine the effects of ChlF on apparent relectance of plants in the red edge spectral region (∼650 to 800 nm) and have been developed from either relectance or derivative relectance spectra. These indices are usually combinations of narrowbands located at the ChlF signal peaks (685 or 740 nm) and unaffected reference bands. + Absorbed radiation (e.g., APAR) has three primary pathways: through the electron transport chain to the dark cycle photosynthetic processes; dissipation as heat via the xanthophyll system; and emission as ChlF. Plants continually adjust energy low to support photosynthesis, while avoiding photooxidation under stressful conditions. Thus, these three pathways are interrelated (Figure 12.4), and stress responses and photosynthetic function can be inferred from bioindicators related to the xanthophyll pigment cycle pigment and ChlF.

+ PRI selected because MR is entering the mesophyll where photosyntesis is happening. Thus it is likely that spectral changes in the PAR can be observed. PRI is a spectral bioindicator for changes in...

* The xanthophyll (Xanthophylls (originally phylloxanthins) are yellow pigments that occur widely in nature and form one of two major divisions of the carotenoid group; the other division is formed by the carotenes. ) pigment cycle serves as a photoprotective mechanism underlying relectance responses produced by physiologically induced spectral changes at 531 nm captured by a spectral bioindicator, referred to as the Photochemical Relectance Index (PRI). Probably red anthocyanins dominate but as MR is present, there must be increased levels of yellow pigments (urediniospores Mahlein)
  + According to Blackburn (2007) recent research for optimizing spectral vegetation indices resulted in the incorporation of more than two narrow (<10 nm) wavebands from the visible region (VIS, 400–700 nm), the red edge (700–740 nm) or the near infrared (NIR 700 nm – 1300 nm). Three band indices are often used at the leaf level while four band indices are applied at the canopy level (Gitelson et al., 2003; Sims & Gamon, 2002; Thenkabail et al., 2000). In the presented study the number of wavebands was stepwise reduced; from 2151, first down to 201 wavebands, then to 12 wavebands and finally to 4 wavebands that are most likely the ones of highest predictive power for myrtle rust on lemon myrtle trees. Four bands were chosen to be able to scale the LMMR up to the canopy level, as it already performs well on the leaf level. But why have these four wavebands been selected from amongst 2151? As every waveband is responding to variation of a plants biochemical or biophysical trait, there must be a combination of wavelengths that is specific to myrtle rust infections in lemon myrtle trees. As for the estimation of pigments, Richardson [29 Thenkabail] compared the performance of two commercially available handheld chlorophyll absorbance meters with that of several relectance indices for the estimation of leaf-level chlorophyll and found that indices based on relectance in the red edge (700–740 nm) region [18,19] were much better indicators of chlorophyll content than some of the more commonly used indices. According to Gamon et al. (1992) reflectance at 531 nm can detect the interconversion of the xanthophyll cycle pigments. Since xanthophyll cycle pigments are regulatory pigments linked to PSII light use efficiency, reflectance indices incorporating reflectance next to 531 nm could provide an indicator of photosynthetic function. Blackburn [27] e.g. suggested that the optimal individual waveband for carotenoids estimation is located at 470 nm. Yet another suggestion by Filella et al. [2] was made. They proposed a structure-insensitive pigment index in the form, (R800 − R445)/(R800 − R680). To retrieve the carotenoid/chlorophyll ratio for a range of individual leaves and conditions, Merzlyak et al. [5] found that the difference of relectances in the green and the blue ranges (R680 − R500) depends on the pigment composition. The index (R680 − R500)/R750 was found to be sensitive to carotenoid/chlorophyll ratio and used as a quantitative measure of plant senescence. For Anthocyanins, Gitelson (2001) found that leaf absorptance near 550 nm is linearly related with anthocyanins content. Anthocyanins are water-soluble vacuolar pigments of higher plants abundant in juvenile and senescing plants. Very often, in plant leaves, significant accumulation of anthocyanins is induced as a result of a number of environmental stresses such as strong light, UV-B irradiation, low temperature, drought, wounding, bacterial and fungal infections, nitrogen and phosphorus deficiencies, certain herbicides and pollutants. Finally, the NIR portion of the electromagnetic spectrum is highly sensitive to changes in biophysical quantities (e.g.??) and plant structure.
* As part of these validation tests, future studies could compare our approach with other waveband selection methods to further refine the set of relevant wavebands. As sensor technology is developed further, the amount of data that can be generated is increasing tremendously. With this, it becomes more challenging to choose suitable methods to reduce data dimensionality and select important wavebands. Our study provides one possible workflow and there are vast amounts of methods available that potentially lead to a similar outcome. DEN LETZTEN HALBEN SATZ MAG ICH NICHT, DER IST SO UNSPEZIFISCH. WAS WILLST DU DAMIT SAGEN?

Ashourloo, Davoud, Mohammad Mobasheri, and Alfredo Huete. 2014. “Evaluating the Effect of Different Wheat Rust Disease Symptoms on Vegetation Indices Using Hyperspectral Measurements.” *Remote Sensing* 6 (6): 5107–23. doi:[10.3390/rs6065107](https://doi.org/10.3390/rs6065107).

Bock, C. H., G. H. Poole, P. E. Parker, and T. R. Gottwald. 2010. “Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging.” *Crit Rev Plant Sci* 29. doi:[10.1080/07352681003617285](https://doi.org/10.1080/07352681003617285).

Burnham, Kenneth P., and David Raymond Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. New York: Springer.

Calcagno, Vincent, and Claire de Mazancourt. 2010. “Glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models.” *Journal of Statistical Software* 34 (12): 1–29. <https://core.ac.uk/download/files/153/6303128.pdf>.

Carnegie, Angus J., Amrit Kathuria, Geoff S. Pegg, Peter Entwistle, Matthew Nagel, and Fiona R. Giblin. 2016. “Impact of the Invasive Rust Puccinia Psidii (Myrtle Rust) on Native Myrtaceae in Natural Ecosystems in Australia.” *Biological Invasions* 18 (1): 127–44. doi:[10.1007/s10530-015-0996-y](https://doi.org/10.1007/s10530-015-0996-y).

Cohen, J. 1960. “A Coefficient of Agreement for Nominal Scales.” *Educational and Psychological Measurement* 20 (1): 37–46. doi:[10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104).

Devadas, R., D. W. Lamb, S. Simpfendorfer, and D. Backhouse. 2009. “Evaluating Ten Spectral Vegetation Indices for Identifying Rust Infection in Individual Wheat Leaves.” *Precision Agriculture* 10 (6): 459–70. doi:[10.1007/s11119-008-9100-2](https://doi.org/10.1007/s11119-008-9100-2).

Genuer, Robin, Jean-Michel Poggi, and Christine Tuleau-Malot. 2015. “VSURF: An R Package for Variable Selection Using Random Forests.” *The R Journal* 7 (2): 19–33. <https://journal.r-project.org/archive/2015-2/genuer-poggi-tuleaumalot.pdf>.

Kuhn, Max, and Kjell Johnson. 2013. *Applied Predictive Modeling*. New York, NY: Springer New York. <http://link.springer.com/10.1007/978-1-4614-6849-3>.

Landis, J. Richard, and Gary G. Koch. 1977. “The Measurement of Observer Agreement for Categorical Data.” *Biometrics* 33 (1): 159. doi:[10.2307/2529310](https://doi.org/10.2307/2529310).

Mahlein. 2015. “Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping.” *Plant Disease* 100 (2): 241–51. doi:[10.1094/PDIS-03-15-0340-FE](https://doi.org/10.1094/PDIS-03-15-0340-FE).

Mahlein, A.-K., T. Rumpf, P. Welke, H.-W. Dehne, L. Plümer, U. Steiner, and E.-C. Oerke. 2013. “Development of Spectral Indices for Detecting and Identifying Plant Diseases.” *Remote Sensing of Environment* 128: 21–30. doi:[10.1016/j.rse.2012.09.019](https://doi.org/10.1016/j.rse.2012.09.019).

Morin, Louise, Mark J. Talbot, and Morag Glen. 2014. “Quest to Elucidate the Life Cycle of Puccinia Psidii Sensu Lato.” *Fungal Biology* 118 (2): 253–63. doi:[10.1016/j.funbio.2013.12.004](https://doi.org/10.1016/j.funbio.2013.12.004).

Nutter, F., N. van Rij, S. K. Eggenberger, and N. Holah. 2010. “Spatial and Temporal Dynamics of Plant Pathogens.” In *Precision Crop Protection - the Challenge and Use of Heterogeneity*, edited by E. C. Oerke, R. Gerhards, G. Menz, and R. A. Sikora. Dordrecht, Netherlands: Springer. <http://dx.doi.org/10.1007/978-90-481-9277-9_3>.

Oldeland, Jens, Wouter Dorigo, Lena Lieckfeld, Arko Lucieer, and Norbert Jürgens. 2010. “Combining Vegetation Indices, Constrained Ordination and Fuzzy Classification for Mapping Semi-Natural Vegetation Units from Hyperspectral Imagery.” *Remote Sensing of Environment* 114 (6): 1155–66. doi:[10.1016/j.rse.2010.01.003](https://doi.org/10.1016/j.rse.2010.01.003).

Peñuelas, J, JA Gamon, AL Fredeen, J Merino, and CB Field. 1994. “Reflectance Indices Associated with Physiological Changes in Nitrogen-and Water-Limited Sunflower Leaves.” *Remote Sensing of Environment* 48 (2): 135–46.

R Core Team. 2016. “R: A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Stevens, Antoine, and Leonardo Ramirez–Lopez. 2014. “An Introduction to the Prospectr Package.” <ftp://200.236.31.2/CRAN/web/packages/prospectr/vignettes/prospectr-intro.pdf>.

Thenkabail, Prasad S., Ronald B. Smith, and Eddy De Pauw. 2000. “Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics.” *Remote Sensing of Environment* 71 (2): 158–82. <http://www.sciencedirect.com/science/article/pii/S003442579900067X>.

Trocha, Paul, J. M. Daly, and R. J. Langenbach. 1974. “Cell Walls of Germinating Uredospores.” *Plant Physiology* 53 (4): 519–26. doi:[10.1104/pp.53.4.519](https://doi.org/10.1104/pp.53.4.519).