An investigation into robust spectral indices for leaf chlorophyll estimation

Russell Main [a] *, Moses Azong Cho [a], Renaud Mathieu [a], Martha M. O'Kennedy [a], Abel Ramoelo[a], Susan Koch[b],

[a] The Council for Scientific and Industrial Research, Natural Resources and Environment, Ecosystems Earth Observation Group, P.O. Box 395, Pretoria 0001, South Africa

[b] Plant Protection Research Institute, Agricultural Research Council, Private Bag X134, Pretoria 0001, South Africa

* Corresponding author: Russell Main, rmain@csir.co.za, Tel: (+27)12 841 3840, Fax: (+27)12 841 3909

Abstract

Quantifying photosynthetic activity at the regional scale can provide important information to resource managers, planners and global ecosystem modelling efforts. With increasing availability of both hyperspectral and narrow band multispectral remote sensing data, new users are faced with a plethora of options when choosing an optical index to relate to their chosen leaf or canopy parameter. The literature base regarding optical indices (particularly chlorophyll indices) is wide ranging and extensive, however it is without much consensus regarding robust indices. The wider spectral community could benefit from studies that apply a variety of published indices to differing sets of species data. The consistency and robustness of 73 published chlorophyll spectral indices have been assessed, using leaf level hyperspectral data collected from three crop species and a variety of savanna tree species. Linear regression between total leaf chlorophyll content and bootstrapping were used to determine the predictive capabilities of the various indices. The indices were then ranked based on the prediction error (the average root mean square error (RMSE)) derived from the bootstrapping process involving 1000 iterative resampling with replacement. The results show two red-edge derivative based indices (Red-edge position via linear extrapolation index and the modified red-edge inflection point index) as the most consistent and robust, and that the majority of the top performing indices (in spite of species variability) were simple ratio or normalised difference indices that are based on off-chlorophyll absorption centre wavebands (690 – 730 nm).

Keywords: Leaf level reflectance, Leaf chlorophyll, Red-edge, Vegetation indices

1. Introduction

Leaf chlorophyll and nitrogen content have been shown to be important bio-indicators of plant physiological state, mainly due to their roles in photosynthesis (Carter, 1994b; Lichtenthaler, 1998). Being able to quantify leaf chlorophyll (chl) or nitrogen (N) contents, and by association photosynthetic activity, could provide useful information a) for precision agriculture at the stand/field scale, b) for improved resource use and planning at the protected area (e.g. National Parks) scale, c) for regional, and/or global, modelling of ecosystem services and productivity. Many years of published research surrounding the spectral changes experienced when vegetation chlorophyll and nitrogen contents change, has led to the inclusion of chlorophyll sensitive wavebands on a number of earth observation satellites. These advances being made in spaceborne multispectral, and hyperspectral, sensors should make the quantification of vegetation vitality increasingly possible for both novice and advanced users of remote sensing data. Leaf, and field, level measurements are often an important part in the development and calibration of vegetation indices (VIs) that are eventually used to quantify changes in vegetation productivity.

Decades of research has gone into finding biochemically sensitive regions within the vegetation spectrum that can be non-destructively extracted (i.e. quantified) using combinations of wavebands (i.e. vegetation indices) from remote sensing platforms. By far the most investigated part of the vegetation spectrum is the spectral red-edge, situated between 670 and 800 nanometres (nm) (Myneni and Asrar, 1994; Veroustraete *et al.*, 1996; Carter, 1998; Goetz *et al.*, 1999; Gupta *et al.*, 2003; Inoue *et al.*, 2008; Ustin *et al.*, 2009;). The red-edge region is characterised by an abrupt change in canopy reflectance between the red (670 nm) and near infrared (NIR) (800 nm), caused by the combined effects of strong chlorophyll absorption in the red wavelengths and high leaf structure-driven reflectance in the NIR (Gates *et al.*, 1965; Tucker, 1979; Horler *et al.*, 1983). The red-edge position (REP) has been shown to have good correlation to chlorophyll content, and is defined by the point of maximum slope between the red chlorophyll absorption region, and the region of high NIR reflectance (Horler, 1983). The shape and position of the red-edge are influenced by variations of chlorophyll content and leaf structure (Filella and Peñuelas, 1994). An increase in the amount of chlorophyll and/or water in leaves generally causes the red-edge to shift to longer wavelengths due to an expanding red absorption well (Gitelson, 1996; Mutanga *et al.*, 2003). Decreasing chlorophyll and water contents, usually associated with stress events or senescence, have been linked to a shift in the red-edge position towards shorter wavelengths (Rock, 1988). Quantitative hyperspectral remote sensing of terrestrial bio-chemistry therefore makes use of indices to monitor the position of the REP. Other chlorophyll, structural, and water-related indices are also used in order to better assess net primary production, environmental and nutritional stresses, and the effects of disease on vegetation vitality (Filella and Peñuelas, 1994; Gitelson and Merzlyak, 1997; Barry *et al.,* 2008; Delalieux *et al*., 2009).

The derivation of the most commonly used optical index for characterising canopy photosynthesis, the normalised difference vegetation index (NDVI), is based on the reflectance contrast between the red and the NIR (Rouse *et al.*, 1974; Tucker, 1979). Efforts in the remote sensing of canopy chlorophyll content via NDVI have however been hindered by the limitations in the spectral resolution of conventional broadband (> 10 nm) sensors such as Landsat TM (Curran, 2001; Gitelson and Merzlyak, 1997). There are major shortcomings with broadband NDVIs derived from red wavebands positioned in the chlorophyll absorption pit (at about 670-680 nm) and bands positioned in the NIR plateau (between 750-900 nm). Several studies have demonstrated the instability of broadband NDVI with varying soil brightness, canopy structure, illumination and viewing geometry, as well as atmospheric conditions (Baret and Guyot, 1991; Goward and Huemmrich, 1992; Huete *et al.*, 1992; Huete and Jackson, 1988; Kaufman and Tanré, 1992; Qi *et al.*, 1995). Furthermore, broadband NDVIs asymptotically approach a saturation level after a leaf area index (LAI) of approximately 4 (Seller, 1985; Mutanga and Skidmore, 2004; Cho *et al.*, 2007). Thus, broadband NDVIs are only effective in distinguishing broad differences in vegetation condition (e.g. greenness), but are not effective in providing a detailed quantitative assessment of canopy photosynthesis (Cho & Skidmore, 2006b).

Studies based on narrowband spectra (< 10 nm) have revealed a broadening of the major chlorophyll absorption feature centred around 670-680 nm with an increasing chlorophyll content (Carter, 1994b; Gitelson and Merzlyak, 1997; Yoder and Pettigrew-Crosby, 1995; Dawson *et al.*, 1999) causing a shift in the red-edge slope towards longer wavelengths (Cho *et al.*, 2008; Curran *et al.*, 1997; Horler *et al.*, 1983). The broadening of the absorption feature causes greater sensitivity of off-centre wavelengths (i.e. 690-730 nm) to subtle changes in chlorophyll content when compared to bands located in the centre of the absorption feature (Carter, 1994b). On the basis of this knowledge an increasing number of narrow waveband multispectral satellite sensors have been designed to include off-chlorophyll absorption centre wavebands, and have recently been launched into space (e.g. RapidEye, Worldview-2 and SumbandilaSAT). These sensors provide more sensitivity towards canopy biochemical constituents by including red-edge wavebands (i.e. 690-730 nm), while still having the characteristics of multispectral sensors (i.e. wider swaths, medium-high spatial resolution).

Prior to the above-mentioned developments, much of the research was focused on how best to relate the most sensitive bands in the red-edge to the vegetation biochemicals causing the spectral deviations. This often involved the development of VIs, which can take the form of normalised difference ratios (i.e. [Rx – Ry] / [Rx + Ry]), simple ratios (i.e. Rx / Ry), reflectance derivatives, or more complex band combinations. These VIs are initially developed at leaf level using hyperspectral data and empirically-derived relationships before being scaled up to canopy level and eventually applied to multispectral image data in order to produce regional maps. Up scaling to canopy level reflectance introduces a variety of "spectral noise" to the leaf reflectance spectra. Canopy reflectance is a combination of green and non-green plant parts (bark, flowers), and is influenced by plant structure (e.g. shadows and leaf orientation) and soil background (Blackburn, 1998). For this reason, vegetation index development often revolves around reducing unwanted reflectance effects while at the same time increasing the indices' sensitivity towards those biochemical (e.g. chlorophyll, stress pigments) and biophysical (e.g. LAI) parameters of interest. Not all VIs are developed to enhance the same parameters. Some, such as the Modified Chlorophyll Absorption Ratio Index have been shown to be sensitive to LAI, chlorophyll and chlorophyll-LAI interactions (Daughtry *et al.*, 2000). On the other hand, Datt (1999) developed an index that was more sensitive to pigments than it was to LAI or scattering influences. Ideally, the goal for researchers would be to develop VIs that are not only as sensitive as possible to the desired parameter, but also robust across species and leaf structures.

The performance of VIs to retrieve biochemical pigments (especially chlorophyll) has been the subject of several studies. Many of the published VIs, and subsequent comparative review studies, are based on one or only a few plant species (Vogelman *et al.*, 1993, Peñuelas *et al.*, 1994; Gitelson and Merzlyak, 1997; Stagakis *et al.*, 2010). While researchers do have the benefits of radiative transfer models at their disposal (i.e. PROSPECT and SAIL), which can model endless variations of leaf or canopy reflectance's by tweaking key input parameters within the models, these models cannot adequately capture the complexity of the interaction of light with all leaf or canopy types (Jacquemoud & Baret, 1990; Kuusk, 1991; le Maire *et al.*, 2004). Therefore, there is the continued need to establish the predictive capability of VIs (both narrow and broadband) for and across a range of species (both at leaf and canopy level) in various environments and ecosystems.

The paper aims to build on studies such as Sims & Gamon (2002) and le Maire *et al.* (2004) by testing the performance of a range of published (chlorophyll) indices in their ability to predict leaf level chlorophyll content (mg/m^2) for a variety of species datasets. The VIs in this paper were applied to leaf level reflectance data for three crop species (maize, tomato and cabbage), as well as a dataset containing eight savanna tree species. This study used linear regression and a bootstrapping technique in order to compare the estimation accuracy (root mean square error, RMSE) of each spectral index in determining chlorophyll content (mg/m^2). The RMSE performance of each VI was then ranked and summed on a per dataset basis, as well as for a combined species dataset, in order to gain insight into which of the VIs are more consistent across the species (i.e. datasets treated separately) and robust for the combined species dataset.

2. Methodology

2.1 Leaf spectral data

The leaf data used in this study were collected from garden crops and wild plants, which resulted in a wide range of structural differences. Leaf level spectral measurements were collected from maize plants (n = 73), cabbage plants (n = 35), tomato plants (n = 35), as well as from eight savanna tree species (n = 80, n = 10 per species), namely, *Combretum hereroense, Combretum molle, Combretum collinum, Euclea natalensis, Terminalia sericea, Sclerocarya birrea, Pterocarpus rotundifolius and Lannea discolor*. The maize plants had been grown under controlled conditions within a greenhouse and were being subjected to varying nutrient treatments. The savanna tree species were collected the summer of 2010 within the greater Kruger National Park, Mpumalanga Province, South Africa. The cabbage and tomato plants were of the garden variety and were growing in a common garden setting.

Spectral measurements were made using an ASD FieldSpec3(R) spectrometer (Analytical Spectral Devices, Boulder, CO, USA) and its associated leaf contact probe. The ASD collects data in the 350–2500 nm spectral region with a resampled spectral resolution of 1 nm. Two reflectance measurements were made of the adaxial leaf surface and then averaged. The contact probe has a diameter of 25 millimetres (mm), an instantaneous field of view of 10 mm, as well as its own halogen lamp light source. After each leaf level reflectance measurement, a leaf borer (diameter=18 mm) was used to clip the same area of the leaf that had just been measured. The collected leaf samples were kept cool and dry before being sent for chlorophyll content analysis, within 24 hours. The wet lab extraction technique was used to determine the chlorophyll concentration per unit area of leaf chlorophyll (Lichtenthaler & Wellburn, 1983). After recording the fresh weight of the leaf samples, the leaf pigments were extracted in 100% acetone. The extract was then spun in a micro centrifuge to precipitate the cell debris. The absorbance (A) of the samples was measured at 661.2 nm (for chlorophyll a) and 644.8 nm (for chlorophyll b) by the Ultra Violet to Visible spectrophotometer. Chlorophyll a, chlorophyll b and total chlorophyll content were computed using the following equations (Lichtenthaler & Buschmann, 2001):

$$\text{Chlorophyll a } (\mu g/ml) = 11.24A_{661.2} - 2.04A_{644.8} \qquad \text{Eq. 1}$$

$$\text{Chlorophyll b } (\mu g/ml) = 20.13A_{644.8} - 4.19A_{661.2} \qquad \text{Eq. 2}$$

$$\text{Total Chlorophyll} = \text{chlorophyll a} + \text{chlorophyll b} \qquad \text{Eq. 3}$$

The unit of the chlorophyll was subsequently converted to mg/m^2 using data on the volume of leaf pigment extract and the leaf disc area. Only the total chlorophyll was used in this study.

The four datasets represent a variety of leaf structures, leaf surfaces and leaf chlorophyll contents (Fig. 1), but every effort was made to also include leaves of different developmental stages and conditions for each individual dataset. No outliers were removed from the data as none had a consistent effect on all the indices, and as we discuss later, some indices appear to deal with them better than others.

(Figure 1)

2.2 Data analysis

Using the leaf reflectance data, we calculated 73 published chlorophyll indices (Table 1). The types of indices included simple ratio indices (e.g. R_{750nm} / R_{710nm}), normalised difference ratios (e.g. Normalised Difference Vegetation Index (NDVI) = $(R_{800} - R_{670}) / (R_{800} + R_{670})$), modified versions of these two types of indices (e.g. modified NDVI = $(R_{800} - R_{680}) / (R_{800} + R_{680} - 2R_{445})$), as well as REP based indices (Table 1). The indices included in this study vary widely in their original target parameters (i.e. chl a, chl b, chl total, stress or LAI), as well as the target levels (i.e. canopy or leaf) at which they were developed and/or intended. However, the majority of the indices included were developed at the leaf level and were intended to be related to chlorophyll parameters. A number of canopy level indices have been included out of interest and in preparation for future canopy level studies of a similar kind.

Linear regression and bootstrapping techniques were used to determine the performance of each index in predicting total chlorophyll content (mg/m^2) (Efron, 1983; Uraibi *et al.*, 2009). The bootstrapping technique iteratively (1000 iterations) resampled two-thirds of the dataset for model calibration and one-third of the dataset for validation, which makes it a good technique for assessing the model accuracy for datasets with a limited amount of samples (Verbyla and Litvaitis, 1989). Linear regressions between chlorophyll content and the spectral indices were used to compute the model coefficient of determination (R^2) and the prediction error (root mean square error (RMSE)) for leaf chlorophyll content. The techniques were implemented within Mathworks (2009), and the RMSE for each spectral index was calculated as an average of the RMSE generated from the 1000 iterations.

The consistency and robustness of the various VIs in estimating leaf chlorophyll content was assessed in two different ways, namely, for each dataset and for the combined data:

(i) In the first scenario, the RMSE values were computed for the linear regressions between the leaf chlorophyll content and the respective VI values for each leaf dataset (cabbage, tomato, maize and savanna trees) separately. Subsequently, the predictive performance of the 73 VIs was assessed by ranking the RMSE values in ascending order for each leaf species dataset. The overall performance of the indices across the four datasets was then evaluated by finding the sum of the ranks and then ordering the VIs according to increasing summed ranks, i.e., the best performing VI across the four datasets will have the lowest summed rank.

(ii) The second scenario involved combining the four datasets into one, for which the respective RMSE values were calculated. The VIs were then ranked in ascending order according to increasing RMSE value.

(Table 1)

3. Results

3.1 Performance of indices across datasets

In order to identify the consistently performing indices over the four datasets, we summed each index’s ranking position over the four datasets. In Table 2 the indices have been sorted according to their summed ranks, in ascending order. Looking at the rankings, one of the first observations is that many of the indices in the top quarter of the table make use of off-chlorophyll absorption centre wavebands, which lie between 690 nm and 730 nm (e.g. MTCI, Maccioni index, VOG3, and Datt1) (Table 2). These top indices make use of various off-chlorophyll absorption centre wavebands, in both derivative and raw reflectance form. The top indices are also calculated using a variety of methods. For instance some of the top indices include the modified red-edge inflection point index (mREIP) (Miller *et al.* 1990) that uses an inverted Gaussian fit on reflectance, the linearly extrapolated REP index (REP_LE) (Cho & Skidmore, 2006a) that utilises derivative values, the MERIS terrestrial chlorophyll index (MTCI) (Dash & Curren, 2004) that uses reflectance data in normalised difference ratios, and the Vogelmann index (Vogelmann *et al.*, 1993) that utilises two derivative wavebands in a simple ratio calculation. Out of all 73 indices there are four indices that seek to determine the red-edge position (REP), but only the REP_LE and the mREIP indices appeared in the top ten, while the red-edge inflection point (REIP) and linearly interpolated REP (REP_LI) indices appear in 16th and 39th position respectively. Of the top 25 indices there are at least eleven indices that usually have their focus on canopy level measurements (e.g. DDn, Boochs2, MCARI2/OSAVI2, TCARI/OSAVI, D2, MCARI2, OSAVI2, mSR2, D1, MTCI and mREIP). Two of these eleven indices perform well enough to appear in the top three of all the indices (i.e mREIP and MTCI). Indices that have their focus on carotenoids and stress related pigments (e.g. NPCI, SRPI, SIPI) include wavebands in the green and/or blue spectral regions (i.e. 450 – 550 nm), and therefore have poor relations to chlorophyll content, which results in their appearance towards the bottom of the rankings. Indices that are dominated by bands close to, on, or in, the chlorophyll absorption pit (i.e. 670nm to 680nm) and chlorophyll absorption plateau (i.e. 750nm to 900nm) also appear at the bottom of the rankings (e.g. mSR, mNDVI, RDVI).

(Table 2)

Table 2 also allows some insight into the leaf types and chlorophyll contents under which certain indices perform best. For instance, the D1 index (Zarco-Tejada *et al.*, 2003) performed well for both the cabbage and tomato datasets, but then struggled to deal with the low chlorophyll maize dataset (See Figure 1), as well as the variety of leaf structures in the savanna tree dataset. The maize data resulted in 8 of the top 15 indices recording their lowest ranks for all four datasets. The Maccioni and Datt indices experienced their lowest performance when applied to the high chlorophyll savanna tree dataset, but then had their highest performance with the low chlorophyll maize data (Table 2). It is also interesting to note the performance of the canopy-based, and soil adjusted, OSAVI2 index, which was the third highest performing index for the maize data (RMSE = 17.32 mg/m^2, Rank = 3). The mREIP index ranked above all other indices for both the savanna tree and maize datasets, but then experienced its lowest ranking with the medium chlorophyll tomato dataset.

3.2 Performance of indices for combined species dataset

To investigate the robustness of the indices across different species, we combined all the datasets and again looked into the relationships between each index and the combined chlorophyll content data (mg/m^2) (See Table 2). Once again the indices utilising off-chlorophyll absorption centre wavebands appear high in the rankings (e.g. Vogelmann3, Maccioni, MTCI, mND_{705}, Carter4). The derivative based REP_LE (Rank = 1, RMSE = 55.10 mg/m^2) and mREIP (Rank = 2, RMSE = 57.08 mg/m^2) showed their consistency in the previous scenario, and this time demonstrate their robustness by once again performing well and appearing at the top of the rankings.

In much the same vein as the previous scenario, the top placed indices are derived using an assortment of wavebands and methods, and also include a number of canopy-based indices. The OSAVI2 (Rank = 4, RMSE = 59.31 mg/m^2) and MTCI (Rank = 6, RMSE = 61.84 mg/m^2) indices are two such canopy-based models that produced low RMSE values. Indices such as the OSAVI2 and MCARI2 indices are modifications of the original indices (i.e. OSAVI and MCARI) in order to include off-chlorophyll absorption centre bands (e.g. 750 nm and 705 nm). Both of the modified indices perform considerably better than their predecessors (in both scenarios). The TCARI2 index is an exception to this though, as it is also modified to include off-chlorophyll absorption centre wavebands, but is outperformed by its predecessor (i.e.

TCARI) in both scenarios. Only three of the top 25 indices have wavebands in the blue or green region of the spectrum (e.g. MCARI2, G-NDVI, and mND_{705}), while many of the other indices that include these wavelengths appear towards the bottom of the rankings.

Fig. 2 shows a number of scatter graphs which depict the linear relationships of various indices that share common traits. Excluding the four indices that involve the REP or the REIP, Fig. 2 visualises the improved performance of the indices that utilise the off-chlorophyll absorption centre wavebands (e.g. MTCI, NDVI2, SR1 and OSAVI2), as opposed to those that don't (e.g. EVI, NDVI, SR and OSAVI). The indices with off-chlorophyll absorption centre wavebands have higher regression coefficients, greater linearity, and fewer signs of saturation at high chlorophyll values. Also visible in the scatter plots, are outliers that we assume were caused by low chlorophyll yellowish-brown leaves. These outliers were not removed as they did not appear to have a consistent effect on all the indices. In fact Fig. 2 illustrates how the indices with off-chlorophyll absorption centre wavebands better mitigate these low chlorophyll samples (i.e. EVI vs. MTCI, and SR vs. SR1).

(Figure 2)

4. Discussion

This study investigated the performance of 73 published indices using leaf spectra and chlorophyll content data from different species datasets. The aim was to understand which of the myriad of published VIs would be consistent and/or robust enough when applied to, and across, different species datasets. The indices varied greatly in terms of their original focus, and intended targets, but they were tested none-the-less and produced interesting results. We felt that the datasets that we applied the indices to would provide a more than adequate examination of their abilities, due to the variety of leaf structures, leaf surfaces, moisture contents and chlorophyll contents present. The maize data had low chlorophyll contents; the tomato and cabbage datasets had medium chlorophyll contents, while the savanna tree dataset consisted of a variety of leaf structures and leaf surfaces, and had the widest range and highest mean chlorophyll content.

A common observation in the study was that the indices using off-chlorophyll absorption centre wavebands (i.e. 690 – 730nm) appeared regularly in the top of the rankings for each of the scenarios. These bands form an integral part of the red-edge region, which has been shown to have a significant relationship with chlorophyll content and the physiological status of vegetation (Collins, 1978; Horler *et al.*, 1983). The performance of indices with these wavebands would support other literature that points to off-centre wavelengths having greater sensitivity to subtle changes in chlorophyll content when compared to bands in the absorption centre (Carter, 1994b; Zarco-Tejada *et al.*, 2003). For instance, the Maccioni (Maccioni et al, 2001) and Datt (Datt, 1999) indices were developed using high chlorophyll Eucalyptus leaves, and were meant to correct for leaf surface reflectance and scattering (Datt, 1999), yet they performed poorly in the high chlorophyll savanna tree dataset. Both indices include the 680 nm region, which is quick to saturate at low chlorophyll levels and therefore becomes insensitive to high chlorophyll contents (Sims and Gamon, 2002; Wu *et al*., 2008). The second Datt index (i.e. Datt2) is a simple ratio index that excludes the 680 nm region, and subsequently performs better for the high chlorophyll savanna dataset. The improved linearity, and resultant prediction power, of the off-chlorophyll absorption centre indices was evident in Fig. 2 of the results section.

The influence of the off-chlorophyll absorption centre wavebands could also be seen in the performance of the canopy indices. As pointed out in the results, the majority of the canopy indices that were modified to include off-chlorophyll absorption centre wavebands outperformed their predecessors that had bands in the 680 nm or 800 nm regions. These indices have well researched combinations of wavebands that have evidently been selected in order to minimise LAI interference and pick out any changes in canopy chlorophyll, most times at low concentrations and/or against soil background (Daughtry *et al.*, 2000; Haboudane *et al.*, 2002; Wu *et al.*, 2008). This is presumably part of the reason for impressive performances by canopy indices, such as the OSAVI2 index that performed particularly well on the low chlorophyll maize data, and also achieved the fourth highest rank in the combined dataset rankings. The poor performance of the TCARI2 index, compared to its predecessor TCARI is in contrast to what

Wu *et al.* (2008) found.

The phenological state of the leaves, as well as inherent differences between species, results in datasets with variable moisture contents, leaf surfaces, and leaf internal structures. The indices would have had different responses to these moisture and structure variations, which in turn could have influenced their ability to predict for chlorophyll content. The linear extrapolation REP index(i.e. REP_LE), which topped both sets of rankings for the two scenarios, was developed by Cho *et al* (2008) to be highly correlated to leaf chlorophyll content and less sensitive to leaf and canopy biophysical factors than other REP techniques. Sims and Gamon (2002) developed two indices (i.e. mND_{705} and mSR_{705}) that would be relatively insensitive to species and leaf structure variations. They showed that these indices could eliminate the effects of variability in surface reflectance and result in better chlorophyll content predictions across a wide variety of species and vegetation types. In this study their mND_{705} index dealt with the variety of species, and leaf structures, relatively well and was subsequently placed in the top 5 for both scenarios. We could assume that the best performing indices (in both scenarios) probably show a decreased sensitivity to varying leaf structures or moisture contents and can be considered more robust than indices that only did well for crop species.

Our results also have similarities to those reported in le Maire *et al.* (2004), where they used data from various deciduous trees species to test the performance of 60 published chlorophyll indices. Some of the same indices that performed well in the le Maire *et al.* (2004) study also perform well in this study (e.g. Maccioni index, Datt index and Vogelman indices). le Maire *et al.* (2004) intimated that there was little use for REIP like indices, partly due to the influence of the double-peak feature found in the derivative red-edge region and partly because there are computationally simpler and more effective indices. This is in contrast to what our results show in that our two most consistent and robust indices include the modified REIP (with inverted Gaussian fit) by Miller *et al.* (1990), and the linearly extrapolated REP index (REP_LE) that was specifically developed by Cho & Skidmore (2006a) to deal with the double-peak feature.

The question regarding whether there is one single index to use in order to estimate vegetation chlorophyll content has not been answered in this paper, and will continue to depend on the type of vegetation being measured, as well as local ecosystem conditions. The study has however pointed out that:

i) Given the varied datasets used in this study, we showcased indices that were robust and consistent, across datasets and species, and could therefore be seen as priority indices to be tested in any follow up work. For instance the modified REIP (mREIP) index by Miller *et al.* (1990) consistently performed well across the datasets, came second in terms of robustness across species, and was the best performer for the low chlorophyll maize data.

ii) We will be able to limit the number of indices used in follow up tests to narrowband indices, which use off-chlorophyll absorption centre wavebands, due to their prominence amongst the best performing indices in this study.

Further research is recommended regarding whether or not the results of this study would be any different should chlorophyll concentration, instead of chlorophyll content per unit area, be used. It also remains to be seen how the best performing indices in this study would perform using different species datasets from elsewhere in the world, but they have showcased their potential for being candidates in the search for more robust vegetation indices. Their ability to make the step up to canopy scale spectral measurements also needs to be further researched.

5. Conclusions

This study tested the chlorophyll content (mg/m^2) prediction ability of 73 published indices and showed that there are a number of indices that perform regularly well across different datasets and within combined species datasets, despite the varying moisture contents and leaf structures involved. With increasing availability of remote sensing data, especially hyperspectral data, the (potential) user base, be it novice or expert, for these indices is probably expanding. While novice users (e.g. farmers, resource managers) might not fully understand the science behind the indices, they will nonetheless want, or need, to know which of the plethora of published indices would be best suited to their conditions and data. We therefore believe studies such as this one would be useful, and should be encouraged in order to grow the knowledge base surrounding which indices work best for what kind of vegetation, growing where and in what kind of conditions. None of the indices used in this study can be touted as truly universal. However, if similar studies continue to compare published indices across a variety of species, then hopefully a level of consensus could be reached regarding an index's robustness under certain conditions or for particular vegetation types.

Acknowledgments

Our gratitude goes to the Council for Scientific and Industrial Research (CSIR), South Africa, for providing the funding for this study. The authors also wish to acknowledge Mulalo Nemutanzhela and Stella Manganye for technical assistance during the chlorophyll analysis.

6. References

Baret, F. and Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of Environment 35 (2-3), 161-173.

Barry, K.M., Stone,C., and Mohammed,C.L., 2008. Crown-scale evaluation of spectral indices for defoliated and discoloured eucalypts. International Journal of Remote Sensing 29 (1), 47-69

Blackburn, G., 1998. Quantifying Chlorophylls and Caroteniods at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. Remote Sensing of Environment 66 (3), 273-285.

Boochs, F., Kupfer, G., Dockter, K. and Kuhbauch, W., 1990. Shape of the red-edge as vitality indicator for plants. International Journal of Remote Sensing 11 (10), 1741-1753.

Broge, N.H. and Leblanc, E., 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sensing of Environment 76 (2), 156-172.

Carter, G.A., 1994a. Ratios of leaf reflectance in narrow wavebands as indicator of plant stress. International Journal of Remote Sensing 15 (3), 697-704.

Carter, G.A. and Miller, R.L., 1994b. Early detection of plant stress by digital imaging within narrow stress-sensitive wavebands. Remote Sensing of Environment 50 (3), 295-302.

Carter, G.A., 1998. Reflectance wavebands and indices for remote estimation of photosynthesis and stomatal conductance in pine canopies. Remote Sensing of Environment 63 (1), 61-72.

Chappelle, E.W., Kim, M. S., and McMurtrey, J. E. III., 1992. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b and carotenoids in soybean leaves. Remote Sensing of Environment 39 (3), 239–247.

Chen, J., 1996. Evaluation of vegetation indices and modified simple ratio for boreal applications. Canadian Journal of Remote Sensing 22 (3), 229–242.

Cho, M.A. and Skidmore, A.K., 2006a. A new technique for extracting the red-edge position from hyperspectral data: The linear extrapolation method. Remote Sensing of Environment 101 (2), 181-193.

Cho, M.A., Sobhan, I.M. and Skidmore, A.K., 2006b. Estimating fresh grass/herb biomass from HYMAP data using the red-edge position. Remote Sensing and Modeling of Ecosystems for Sustainability III, SPIE, San Diego, CA, USA, 629805-9.

Cho, M.A., Skidmore, A., Corsi, F., van Wieren, S.E. and Sobhan, I., 2007. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. International Journal of Applied Earth Observation and Geoinformation 9 (4), 414-424.

Cho, M.A., Skidmore, A.K. and Atzberger, C., 2008. Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using PROSPECT-SAILH simulated data. International Journal of Remote Sensing 29 (8), 2241-2255.

Collins,W., 1978. Remote sensing of crop type and maturity. Photogrammetric Engineering and Remote Sensing 44 (1), 43–55.

Curran, P.J., Kupiec, J.A. and Smith, G.M., 1997. Remote sensing the biochemical composition of a slash pine canopy. Geoscience and Remote Sensing. IEEE Transactions on 35 (2), 415-420.

Curran, P.J., 2001. Imaging spectrometry for ecological applications. International Journal of Applied Earth Observation, 3 (4), 305-312.

Dash, J., and Curran, P. J., 2004. The MERIS terrestrial chlorophyll index. International Journal of Remote Sensing 25 (23), 5403−5413.

Datt, B., 1998. Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+ b and total carotenoid content in eucalyptus leaves. Remote Sensing of Environment 66 (2), 111–121.

Datt, B., 1999. Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. International Journal of Remote Sensing 20 (14), 2741–2759.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., and McMurtrey, J. E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 74 (2), 229−239.

Dawson, T.P., Curran, P.J., North, P.R.J. and Plummer, S.E., 1999. The Propagation of Foliar Biochemical Absorption Features in Forest Canopy Reflectance: A Theoretical Analysis. Remote Sensing of Environment 67 (2), 147-159.

Delalieux, S., Somers, B., Verstraeten, W. W., van Aardt, J. A. N., Keulemans, W. and Coppin, P. 2009. Hyperspectral indices to diagnose leaf biotic stress of apple plants, considering leaf phenology. International Journal of Remote Sensing 30 (8), 1887-1912.

Efron, B., 1983. Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation. Journal of the American Statistical Association 78 (382), 316-331.

Elvidge, C. D., and Zhikang, C., 1995. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. Remote Sensing of Environment 54 (1), 38–48.

Filella, I., and Peñuelas, J., 1994. The red-edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. International Journal of Remote Sensing 15, 1459−1470.

Gandia, S., Fernández, G., García, J. C., and Moreno, J., 2004. Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign. ESA SP 578, 40−48.

Gates, D.M., Keegan, H.J., Schleter, J.C. and Weidner, V.R., 1965. Spectral properties of plants. Applied Optics 4 (1), 11-20.

Gitelson, A., and Merzlyak, M. N., 1994. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. Journal of Photochemistry and Photobiology B: Biology 22 (3), 247−252.

Gitelson, A. A., Kaufman, Y. J., and Merzlyak,M. N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 58 (3), 289–298.

Gitelson, A.A. and Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing 18 (12), 2691-2697.

Gitelson, A. A., Buschmann, C., and Lichtenthaler, H. K., 1999. The chlorophyll fluorescence ratio F735/F700 as an accurate measure of the chlorophyll content in plants. Remote Sensing of Environment 69 (3), 296–302.

Gitelson, A. A., Gritz, Y., and Merzlyak, M. N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal of Plant Physiology 160 (3), 271−282.

Goetz, S.J., Prince, S.D., Goward, S.N., hawley, M.M.T. and Small, J., 1999. Satellite remote sensing of primary production: An improved production efficiency modeling approach. Ecological Modeling 122 (3), 239-255.

Goward, S.N. and Huemmrich, K.F., 1992. Vegetation canopy PAR absorptance and the normalized difference vegetation index: An assessment using the SAIL model. Remote Sensing of Environment 39 (2), 119-140.

Gupta, R.K., Vijayan, D. and Prasad, T.S., 2003. Comparative analysis of red-edge hyperspectral indices. Advances in Space Research 32 (11), 2217-2222.

Guyot, G. and Baret, F., 1988. Utilisation de la haute résolution spectrale pour suivre l'état des couverts végétaux. In: Guyenne, T.D. and Hunt, J.J., Proc. Fourth International colloquium on spectral signatures of objects in remote sensing, ESA SP-287, Assois, France, 18-22 Jan., pp. 279-286.

Haboudane, D., John, R., Millera, J.R., Tremblay, N., Zarco- Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sensing of Environment 81 (2-3), 416–426.

Horler, D.N.H., Dockray, M. and Barber, J., 1983. The red-edge of plant leaf reflectance. International Journal of Remote Sensing 4 (2), 273-288.

Huete, A.R. and Jackson, R.D., 1988. Soil and atmosphere influences on the spectra of partial canopies. Remote Sensing of Environment 25 (1), 89-105.

Huete, A.R., Hua, G., Qi, J., Chehbouni, A. and van Leeuwen, W.J.D., 1992. Normalization of multidirectional red and NIR reflectances with the SAVI. Remote Sensing of Environment 41 (2-3), 143-154.

Huete, A. R., Liu, H. Q., Batchily, K., and vanLeeuwen, W., 1997. A comparison of vegetation indices global set of TM images for EOS-MODIS. Remote Sensing of Environment 59 (3), 440−451.

Inoue, Y., Peñuelas, J., Miyata, A. and Mano, M., 2008. Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice. Remote Sensing of Environment 112 (1), 156-172.

Jacquemoud, S. and Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment 34 (2), 75-91.

Jordan, C. F., 1969. Derivation of leaf area index from quality of light on the forest floor. Ecology 50 (4), 663−666.

Kaufman, Y.J. and Tanré, D., 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Transactions on Geoscience and Remote Sensing 30 (2), 261-270.

Kim, M.S., Daughtry, C.S.T., Chappelle, E.W., *et al.*, 1994. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (Apar). In: Proc. Sixth Symposium on Physical Measurements and Signatures in Remote Sensing, Val D'Isere, France, January 17–21, 299–306.

Kuusk, A., 1991. The angular distribution of reflectance and vegetation indices in barley and clover canopies . Remote Sensing of Environment 37 (2), 143-151.

le Maire, G., Francois, C., Dufrene, E., 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. Remote Sensing of Environment 89 (1), 1–28.

le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J. Y., Bréda, N., Genet, H., Davi, H., and Dufrêne, E., 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. Remote Sensing of Environment 112 (10), 3846−3864.

Lichtenthaler, H. K., and Wellburn, A. R., 1983. Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents. Biochemical Society Transactions

11 (5), 591−592.

Lichtenthaler, H. K., Lang, M., Sowinska, M., Heisel, F., and Miehe, J. A., 1996. Detection of vegetation stress via a new high resolution fluorescence imaging system. Journal of Plant Physiology 148 (5), 599−612.

Lichtenhaler, H.K., 1998. The Stress Concept in Plants: An Introduction. Annals of the New York Academy of Sciences 851, 187-198.

Lichtenthaler, H. K., and Buschmann, C. 2001. Current Protocols in Food Analytical Chemistry (Units: F4.3.1-F4.3.8), John Wiley & Sons Inc., New York .

Maccioni, A., Agati, G., and Mazzinghi, P., 2001. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. Journal of Photochemistry and Photobiology 61 (1–2), 52–61.

Mathworks, 2009. Numerical methods using Matlab, Prentice-Hall Inc., New Jersey, USA.

McMurtey III, J. E., Chappelle, E. W., Kim, M. S., Meisinger, J. J., and Corp, L. A., 1994. Distinguish nitrogen fertilization levels in field corns (Zea mays L.) with actively induced fluorescence and passive reflectance mesurements. Remote Sensing of Environment 47 (1), 36–44.

Miller, J. R., Hare, E. W., and Wu, J., 1990. Quantitative characterization of the vegetation red-edge reflectance. Remote Sensing 11 (10), 1755–1773.

Mutanga, O., Skidmore, A. K. and Wieren, S., 2003. Discriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry. ISPRS Journal of Photogrammetry and Remote Sensing 57 (4), 263-272.

Mutanga, O. and Skidmore, A.K., 2004. Hyperspectral band depth analysis for a better estimation of grass biomass (Cenchrus ciliaris) measured under controlled laboratory conditions. International Journal of Applied Earth Observation and Geoinformation 5 (2), 87-

96.

Myneni, R.B. and Asrar, G., 1994. Atmospheric effects and spectral vegetation indices. Remote Sensing of Environment 47 (3), 390-402.

Penuelas, J., Gamon, J.A., Fredeen, A.L., Merino, J. and Field, C.B., 1994. Reflectance Indices Associated with Physiological Changes in Nitrogen and Water Limited Sunflower Leaves. Remote Sensing of Environment 48 (2), 135-146.

Peñuelas, J., Filella, I., Lloret, P., Munoz, F., and Vilajeliu, M., 1995. Reflectance assessment of mite effects on apple trees. International Journal of Remote Sensing 16 (14), 2727−2733.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., and Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sensing of Environment 48 (2), 119−126.

Qi, J., Moran, M.S., Cabot, F. and Dedieu, G., 1995. Normalization of sun/view angle effects using spectral albedo-based vegetation indices. Remote Sensing of Environment 52 (3), 207-217.

Rondeaux, G., Steven,M., Baret, F., 1996. Optimization of soil- adjusted vegetation indices. Remote Sensing of Environment 55 (2), 95–107.

Roujean, J. L., and Breon, F. M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sensing of Environment 51 (3),375−384.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. and Harlan, J.C., 1974. Monitoring the vernal advancement and retrogradation of natural vegetation, NASA/GSFC, Type III Final Report, M.D., Greenbelt, 371.

Seller, P.J., 1985. Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing 6 (8), 1335-1372.

Sims, D. A., and Gamon, J. A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote

Sensing of Environment 81 (2-3), 337−354.

Smith, R. C. G., Adams, J., Stephens, D. J., and Hick, P. T., 1995. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. Australian Journal of Agricultural Research 46 (1), 113–125.

Stagakis, S., Markos, N., Sykioti, O., and Kyparissis, A., 2010. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. Remote Sensing of Environment 114 (5) 977-994.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8 (2), 127-150.

Uraibi, H.S., Midi, H., Talib, B.A., Yousif, J.B., 2009. Linear Regression Model Selection Based on Robust Bootstrapping Technique. American Journal of Applied Sciences 6 (6), 1191-1198

Ustin, S.L. Gitelson, A. A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J. A., Zarco-Tejada, P. J., 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment 113 (1),67-77.

Verbyla, D., L and Litvaitis, J., A., 1989. Resampling methods for evalutiing classification accuracy of wildlife habitat models. Environmental Management 13 (6), 783-787.

Vincini, M., Frazzi, E., and D'Alessio, P., 2006. Angular dependence of maize and sugar beet Vis from directional CHRIS/PROBA data. 4th ESA CHRIS PROBA Workshop, ESRIN, Frascati, Italy (pp. 19−21).

Veroustraete, F., Patyn, J. and Myneni, R.B., 1996. Estimating net ecosystem exchange of carbon using the normalized difference vegetation index and an ecosystem model. Remote Sensing of Environment 58 (1), 115-130.

Vogelman, J. E., Rock, B. N., and Moss, D. M., 1993. Red-edge spectral measurements from sugar maple leaves. International Journal of Remote Sensing 14 (8), 1563–1575.

Wu, C., Niu, Z., Tang, Q., Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. Agricultural and Forest Meteorology 148 (8-9), 1230-1241.

Yoder, B.J. and Pettigrew-Crosby, R.E., 1995. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. Remote Sensing of Environment 53 (3), 199-211.

Zarco-Tejada P.J., and Miller J.R., 1999, Land Cover Mapping at BOREAS using red-edge spectral parameters from CASI imagery. Journal of Geophysical Research 104 (D22),27,921-27,933.

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., and Sampson, P.H., 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 39 (7), 1491-1507.

Zarco-Tejada, P. J., Pushnik, J. C., Dobrowski, S., and Ustin, S. L., 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. Remote Sensing of Environment 84 (2), 283−294.

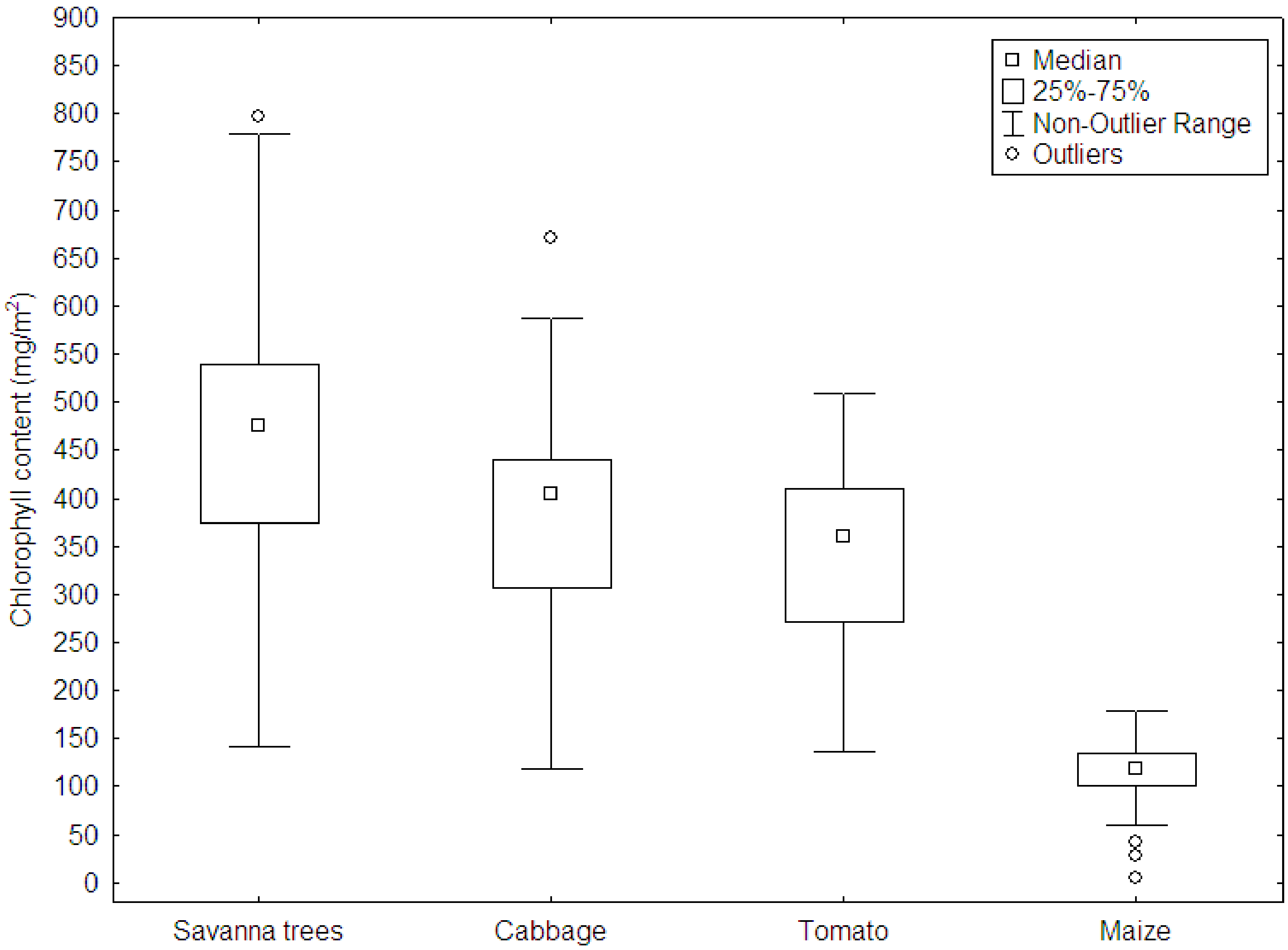
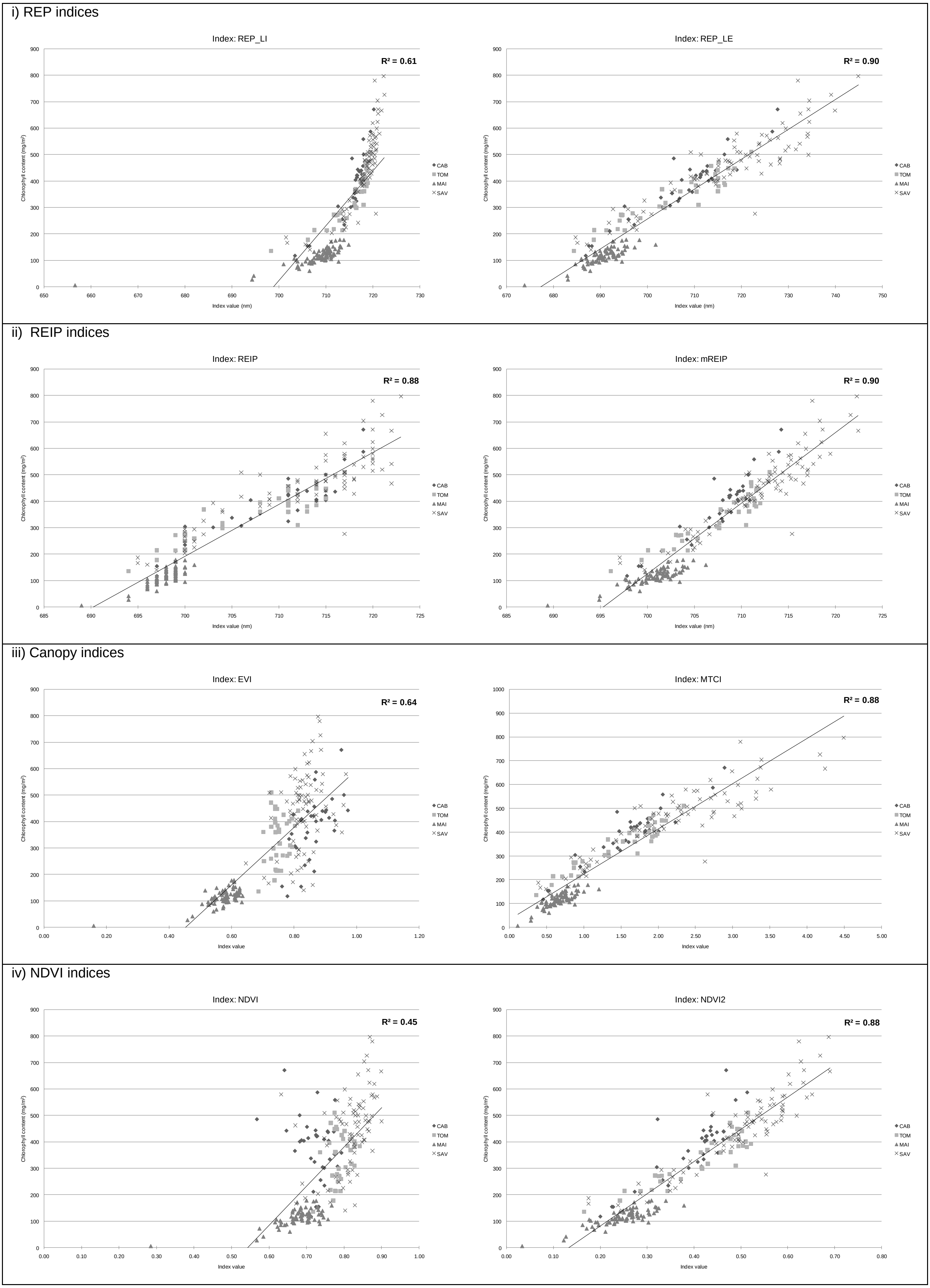
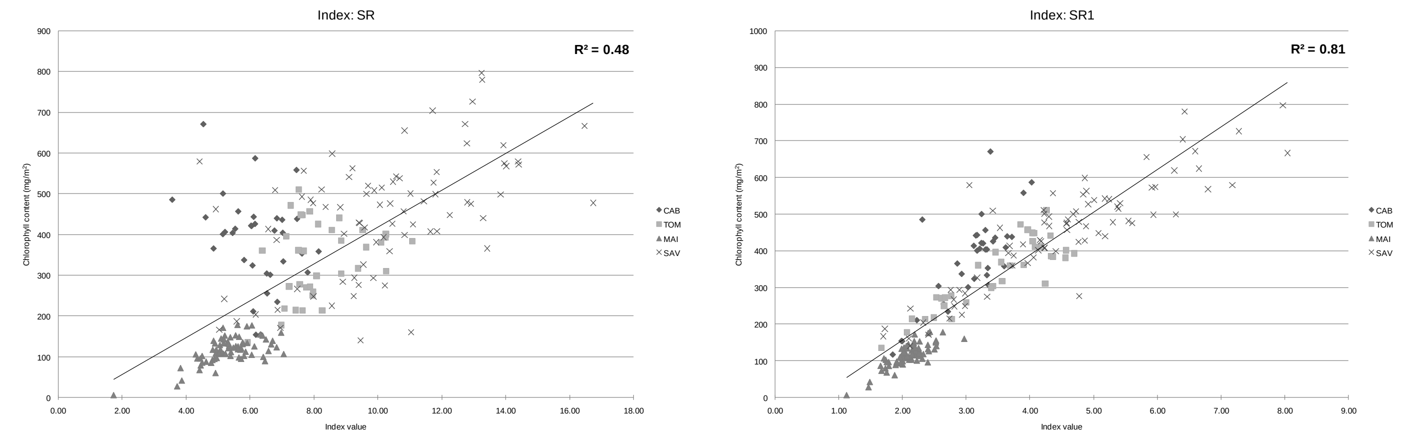


Figure 1: Box plots showing the variability of total chlorophyll content (mg/m^2) for the four species datasets.


i) REP indices
Index: REP_LI
R² = 0.61
Index: REP_LE
R² = 0.90
ii) REIP indices
Index: REIP
R² = 0.88
Index: mREIP
R² = 0.90
iii) Canopy indices
Index: EVI
R² = 0.64
Index: MTCI
R² = 0.88
iv) NDVI indices
Index: NDVI
R² = 0.45
Index: NDVI2
R² = 0.88
Chlorophyll content (mg/m²)
Index value (nm)
Index value
CAB
TOM
MAI
SAV

v) Simple Ratio indices



vi) Soil adjusted vegetation indices

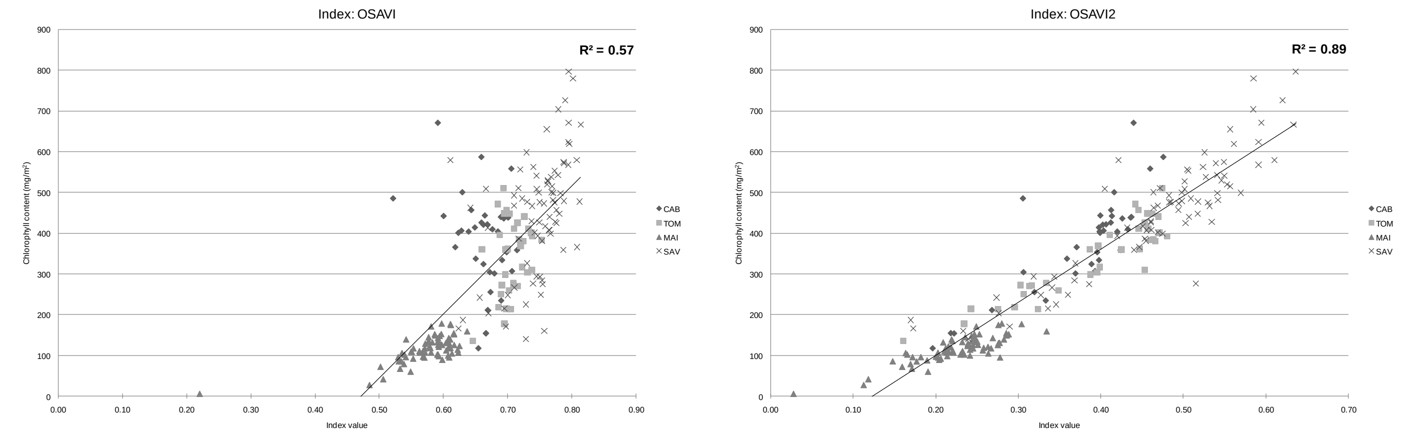


Figure 2: Scatter plots showing the improvements in the regression equations between the original and modified vegetation indices, which have similar attributes (i = REP indices, ii = REIP indices, iii = canopy indices, iv = NDVI indices, v = simple ratio indices, vi = soil adjusted indices). The regression was done using all four species datasets (CAB = cabbage, TOM = tomato, MAI = maize, SAV = savanna trees).

Table 1: Vegetation indices used in the study (adapted from Stagakis et al., 2010 and le Maire et al., 2004).

Index	Formulation	Scale	Related to	Reference
Boochs*	D_{703}	Canopy	chl *a*	Boochs *et al.* (1990)
Boochs2*	D_{720}	Canopy	chl *a*	
CARI (Chlorophyll Absorption Ratio Index)	$R_{700}*(SQRT((a*670 + R_{670} + b)^2)) / R_{670}*(a^2 + 1)^{0.5}$ [$a = (R_{700} - R_{550}) / 150$; $b = R_{550} - (a*550)$]	Leaf	chl	Kim *et al.* (1994)
CI *(Curvature Index)*	$R_{675} * R_{690} / R_{683}^2$	Canopy	chl *a*	Zarco-Tejada *et al.* (2003)
Carter*	R_{695} / R_{420}	Leaf	Stress	Carter (1994a)
Carter2*	R_{695} / R_{760}	Leaf	Stress	
Carter3*	R_{605} / R_{760}	Leaf	Stress	
Carter4*	R_{710} / R_{760}	Leaf	Stress	
Carter5*	R_{695} / R_{670}	Leaf	Stress	
Carter6*	R_{550}	Leaf	chl	
Datt*	$(R_{850} - R_{710}) / (R_{850} - R_{680})$	Leaf	chl	Datt (1999)
Datt2*	R_{850} / R_{710}	Leaf	chl	
Datt3*	D_{754} / D_{704}	Leaf	chl	
Datt4*	$R_{672} / (R_{550} \times R_{708})$	Leaf	chl *a,* chl total	Datt (1998)
Datt5*	R_{672} / R_{550}	Leaf	chl *b*	
Datt6*	$R_{860} / (R_{550} \times R_{708})$	Leaf		
DD (Double Difference Index)	$(R_{749} - R_{720}) - (R_{701} - R_{672})$	Leaf	chl total	le Maire *et al.* (2004)
DDn (new Double Difference Index)	$2*(R_{710} - R_{(710-50)} - R_{(710+50)})$	Canopy	chl total	le Maire *et al.* (2008)
DPI (Double Peak Index)	$(D_{688}*D_{710}) / D_{697}^2$	Canopy	chl fluorescence	Zarco-Tejada *et al.* (2003)
dRE	Maximum value of first derivative in red-edge region	Leaf	chl, stress	Filella and Peñuelas (1994)
D1*	D_{730} / D_{706}	Canopy	chl fluorescence	Zarco-Tejada *et al.* (2003)
D2*	D_{705} / D_{722}	Canopy		
EVI (Enhanced Vegetation Index)	$2.5*((R_{800} - R_{670}) / (R_{800} + (6*R_{670}) - (7.5*R_{475}) + 1))$	Canopy	chl	Huete *et al.* (1997)
EGFR (Ratio of first derivative maxima in red-edge and green regions)	dRE / dG	Leaf	chl, N	Peñuelas *et al.* (1994)
EGFN (Normalised ratio of first derivative maxima in red-edge and green regions)	(dRE – dG) / (dRE + dG)	Leaf	chl, N	
GI (Greenness Index)	R_{554} / R_{677}	Canopy	chl, LAI x chl	Smith *et al.* (1995)
Gitelson*	$1 / R_{700}$	Leaf	chl total	Gitelson et al. (1999),
Gitelson2*	$(R_{750-800} / R_{695-740}) - 1$	Leaf	chl	Gitelson et al. (2003)
Green NDVI	$(R_{800} - R_{550}) / (R_{800} + R_{550})$	Canopy	chl *a*	Gitelson *et al.* (1996)
MCARI (Modified Chlorophyll Absorption Ratio Index)	$((R_{700} - R_{670}) - 0.2*(R_{700} - R_{550}))*(R_{700} / R_{670})$	Canopy	chl, LAI	Daughtry *et al.* (2000)
MCARI / OSAVI	MCARI / OSAVI	Canopy	chl	
MCARI2	$((R_{750} - R_{705}) - 0.2*(R_{750} - R_{550}))*(R_{750} / R_{705})$	Canopy	chl	Wu *et al.* (2008)
MCARI2 / OSAVI2	MCARI2 / OSAVI2	Canopy	chl	
mNDVI (Modified NDVI)	$(R_{800} - R_{680}) / (R_{800} + R_{680} - 2R_{445})$	Leaf	chl total	Sims and Gamon (2002)
mND_{705}	$(R_{750} - R_{705}) / (R_{750} + R_{705} - 2R_{445})$	Leaf	chl total	
Maccioni*	$(R_{780} - R_{710}) / (R_{780} - R_{680})$	Leaf	chl	Maccioni *et al.* (2001)
mREIP (Modified Red-Edge Inflection Point)*	Modified REIP with inverted Gaussian fit on reflectance	Leaf + Canopy	chl	Miller *et al.* (1990)
MSAVI (Improved Soil Adjusted Vegetation Index)	$0.5*(2*R_{800} + 1 - SQRT((2*R_{800} + 1)^2 - 8*(R_{800} - R_{670})))$	Canopy	chl	Qi *et al.* (1994)
mSR (Modified Simple Ratio)	$(R_{800} - R_{445}) / (R_{680} - R_{445})$	Leaf	chl total	Sims and Gamon (2002)
mSR_{705}	$(R_{750} - R_{445}) / (R_{705} - R_{445})$	Leaf	chl total	
mSR2*	$(R_{750} / R_{705}) - 1 / SQRT((R_{750} / R_{705}) + 1)$	Canopy	chl + LAI	Chen (1996)
MTCI (MERIS Terrestrial chlorophyll index)	$(R_{754} - R_{709}) / (R_{709} - R_{681})$	Canopy	chl	Dash and Curran (2004)
NDVI (Normalised Difference Vegetation Index)	$(R_{800} - R_{670}) / (R_{800} + R_{670})$	Canopy	chl, LAI	Tucker (1979)
NDVI2 *	$(R_{750} - R_{705}) / (R_{750} + R_{705})$	Leaf	chl *a*	Gitelson and Merzlyak (1994)
NDVI3*	$(R_{682} - R_{553}) / (R_{682} + R_{553})$	Canopy	chl total	Gandia *et al.* (2004)
NPCI (Normalised Pigment chlorophyll Index)	$(R_{680} - R_{430}) / (R_{680} + R_{430})$	Leaf	(Total pigments) / chl, stress	Peñuelas *et al.* (1994)
OSAVI (Optimised Soil-Adjusted Vegetation Index)	$(1+0.16)*(R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)$	Canopy	chl	Rondeaux *et al.* (1996)
OSAVI2	$(1+0.16)*(R_{750} - R_{705}) / (R_{750} + R_{705} + 0.16)$	Canopy	chl	Wu *et al.* (2008)
RDVI (Renormalised Difference Vegetation Index)	$(R_{800} - R_{670}) / (SQRT(R_{800} + R_{670}))$	Canopy	chl, LAI	Roujean and Breon (1995)
REIP (Red-Edge Inflection Point)	Wavelength for maximum value of the first derivative in red-edge region	Leaf Canopy	chl, LAI chl x LAI, biomass	Collins (1978); Horler *et al.* (1983)
REP_LE* (Red-Edge Position linear extrapolation)	See Cho & Skidmore, 2006	Leaf	N, chl	Cho and Skidmore (2006)
REP_LI * (Red-Edge Position linear interpolation)	$700+40*((R_{670} + R_{780} / 2) / (R_{740} - R_{700}))$	Leaf	chl	Guyot *et al.* (1988)
SIPI (Structure Insensitive Pigment Index)	$(R_{800} - R_{445}) / (R_{800} - R_{680})$	Leaf	(pigments)/chl, stress	Peñuelas *et al.* (1995)
SPVI (Spectral Polygon Vegetation Index)	$0.4*(3.7*(R_{800} - R_{670}) - 1.2*SQRT((R_{530} - R_{670})^2)$	Canopy	chl x LAI	Vincini *et al.* (2006)
SR* (Simple Ratio Index)	R_{800} / R_{680}	Canopy	chl	Jordan (1969)
SR1*	R_{750} / R_{700}	Leaf	chl	Gitelson and Merzlyak (1997)
SR2*	R_{752} / R_{690}			
SR3*	R_{750} / R_{550}			

SR4*	R_{700} / R_{670}	Leaf	chl	McMurtey *et al.* (1994)
SR5*	R_{675} / R_{700}	Leaf	chl *a*	Chappelle *et al.* (1992)
SR6*	R_{750} / R_{710}	Leaf	chl	Zarco-Tejada & Miller (1999)
SR7*	R_{440} / R_{690}	Leaf	Stress	Lichtenthaler *et al.* (1996)
SRPI (Simple Ratio Pigment Index)	R_{430} / R_{680}	Leaf	(Total pigments)/chl, stress	Peñuelas *et al.* (1995)
Sum_Dr1*	Sum of first derivative reflectance between R_{625} and R_{795}	Canopy	chl	Elvidge and Zhikang, 1995
Sum_Dr2*	Sum of first derivative reflectance between R_{680} and R_{780}	Leaf	LAI, chl *a*, chl *b*, chl *a* + *b*	Filella and Peñuelas, 1994
TCARI (Transformed Chlorophyll Absorption Ratio Index)	$3*((R_{700} - R_{670}) - 0.2*(R_{700} - R_{550})*(R_{700} / R_{670}))$	Canopy	chl	Haboudane *et al.* (2002)
TCARI2	$3*((R_{750} - R_{705}) - 0.2*(R_{750} - R_{550})*(R_{750} / R_{705}))$	Canopy	chl	Wu *et al.* (2008)
TCARI / OSAVI	TCARI / OSAVI	Canopy	chl	Haboudane *et al.* (2002)
TCARI2 / OSAVI2	TCARI2 / OSAVI2	Canopy	chl	Wu *et al.* (2008)
TVI (Triangular Vegetation Index)	$0.5*(120*(R_{750} - R_{550}) - 200*(R_{670} - R_{550}))$	Canopy	LAI, canopy chlorophyll density	Broge and Leblanc (2000)
Vogelmann*	R_{740} / R_{720}	Leaf	chl	Vogelmann *et al.* (1993)
Vogelmann2*	$(R_{734} - R_{747}) / (R_{715} + R_{726})$	Leaf	chl	
Vogelmann3*	D_{715} / D_{705}	Leaf	chl	

R_x represents reflectance at wavelength x nm.

D_x represents the derivative of the reflectance spectrum at wavelength x nm.

dRE is the maximum value of the first derivative in red-edge region (670 – 800 nm)

dG is the maximum value of the first derivative in the visible green region (500 – 580 nm)

* No original index abbreviation found, so an appropriate one was inserted

Table 2: The ranking results of the performance (assessed using RMSE) of the 73 vegetation indices to predict total chlorophyll content (mg/m^2) according to the two scenarios, i) across all species datasets and ii) the combined species dataset.

i) Summed rank across datasets										**ii) Combined dataset rankings**		
Index	**Savanna Tree RMSE (mg/m^2)**	**Rank**	**Cabbage RMSE (mg/m^2)**	**Rank**	**Tomato RMSE (mg/m^2)**	**Rank**	**Maize RMSE (mg/m^2)**	**Rank**	**Summed Rank**	**Index**	**RMSE (mg/m^2)**	**Rank**
REP_LE	59.76	2	41.48	3	31.12	4	18.14	12	21	REP_LE	55.10	1
mREIP	59.48	1	46.24	8	34.56	11	17.19	2	22	mREIP	57.08	2
MTCI	62.62	6	40.92	2	29.99	3	18.48	13	24	Vogelmann3	57.85	3
mND_{705}	63.66	10	42.33	5	36.31	18	16.98	1	34	OSAVI2	59.31	4
Gitelson2	61.84	3	49.02	12	33.03	7	18.56	14	36	mND_{705}	59.82	5
Vogelmann3	63.58	9	41.61	4	28.50	1	19.45	25	39	Maccioni	61.30	6
Maccioni	67.17	17	47.68	10	33.89	10	17.35	4	41	MTCI	61.84	7
Vogelmann	62.04	4	49.33	13	32.30	6	18.77	18	41	Carter4	61.92	8
Datt	64.06	12	43.80	7	31.72	5	19.01	20	44	Datt	62.21	9
Vogelmann2	69.32	19	48.11	11	33.87	9	17.37	5	44	NDVI2	62.21	10
D1	70.20	20	39.45	1	29.99	2	20.63	31	54	Vogelmann	62.41	11
Datt2	63.15	8	56.34	21	35.26	13	18.57	15	57	REIP	62.82	12
SR6	62.93	7	56.18	20	35.44	15	18.59	16	58	MCARI2 / OSAVI2	63.09	13
Carter4	65.69	15	59.54	25	37.12	21	17.62	7	68	mSR2	63.58	14
DD index	65.10	13	51.05	15	35.29	14	21.61	35	77	Boochs2	63.70	15
mSR2	62.52	5	64.68	38	38.79	27	18.00	10	80	Gitelson2	64.41	16
REIP	72.12	24	50.50	14	38.20	24	19.44	24	86	DDn Index	65.03	17
OSAVI2	66.07	16	66.88	42	38.98	28	17.32	3	89	EGFN	65.17	18
mSR_{705}	63.79	11	86.63	55	35.45	16	17.88	8	90	DD index	65.41	19
NDVI2	65.44	14	66.13	41	39.17	29	17.62	6	90	SR6	66.74	20
MCARI2	67.80	18	63.77	37	38.65	26	18.12	11	92	MCARI2	68.47	21
D2	79.87	27	55.71	18	37.27	22	21.07	33	100	Datt2	68.65	22
TCARI / OSAVI	80.06	28	62.55	35	38.49	25	20.02	27	115	D1	69.64	23
MCARI2 / OSAVI2	71.05	22	65.03	39	39.19	30	19.85	26	117	Vogelmann2	70.50	24
Boochs2	87.03	33	71.86	47	40.42	34	17.94	9	123	Green NDVI	74.79	25
DDn Index	93.39	38	57.10	22	34.58	12	26.12	52	124	TCARI / OSAVI	76.04	26
Datt3	99.40	45	52.79	16	40.97	35	21.04	32	128	SR1	78.07	27
TCARI	81.61	29	56.08	19	36.61	19	30.43	62	129	EGFR	79.02	28
Datt4	99.32	43	42.83	6	33.73	8	35.15	73	130	SR3	81.78	29
SR1	70.53	21	78.95	52	46.36	45	18.66	17	135	D2	92.70	30
EGFN	90.31	34	53.31	17	39.43	31	26.76	55	137	Carter3	93.82	31
Green NDVI	86.12	31	98.35	59	39.97	32	19.17	21	143	Sum_Dr1	94.92	32
SR3	85.88	30	100.61	61	40.23	33	20.44	28	152	TCARI2 / OSAVI2	95.23	33
MCARI / OSAVI	91.75	35	63.61	36	43.59	42	23.65	40	153	TCARI	96.30	34
Boochs	96.08	41	47.44	9	60.58	54	26.21	53	157	Datt6	96.59	35
Carter	104.28	47	58.70	24	48.48	47	23.97	43	161	Carter2	96.74	36
Carter5	111.70	50	58.14	23	36.66	20	32.35	68	161	mSR_{705}	97.33	37
SR4	104.19	46	60.50	27	38.12	23	31.48	65	161	SPVI	99.77	38
Datt6	86.93	32	62.48	34	48.75	48	24.93	49	163	Datt3	101.84	39
REP_LI	79.20	26	80.45	53	42.52	39	24.28	45	163	Carter6	103.82	40
SR5	111.83	51	60.29	26	35.67	17	33.00	71	165	CI	104.90	41
MCARI	95.79	40	61.57	29	42.44	38	30.78	64	171	EVI	107.69	42
Carter2	92.88	37	99.28	60	55.02	52	19.42	23	172	SR2	107.92	43
Carter3	92.55	36	113.13	65	56.69	53	18.78	19	173	MSAVI	108.56	44
SR7	120.65	54	61.79	31	62.35	55	21.12	34	174	Sum_Dr2	109.93	45
EGFR	97.64	42	67.04	43	42.29	36	26.30	54	175	RDVI	110.54	46
Gitelson	77.11	25	77.77	51	50.11	50	25.42	50	176	Datt4	110.64	47
CI	99.33	44	62.28	33	44.32	44	28.16	56	177	Gitelson	112.19	48
TCARI2 / OSAVI2	71.60	23	82.17	54	42.39	37	32.66	69	183	MCARI / OSAVI	112.96	49
CARI	119.47	53	61.62	30	42.88	40	32.79	70	193	REP_LI	117.36	50
GI	124.69	56	62.10	32	43.04	41	31.57	66	195	OSAVI	118.06	51
Carter6	95.69	39	94.76	58	44.31	43	28.50	57	197	SR	128.95	52
NDVI3	135.94	64	60.80	28	49.63	49	30.56	63	204	GI	130.18	53
SR2	105.64	48	119.23	70	73.63	58	20.60	30	206	DPI	131.40	54
TCARI2	107.80	49	107.46	62	95.96	73	19.31	22	206	NDVI3	131.91	55
DPI	131.81	63	68.13	44	84.59	61	24.59	47	215	NDVI	133.43	56
Sum_Dr1	121.99	55	122.04	73	74.00	59	20.50	29	216	MCARI	136.04	57
NPCI	143.49	69	72.27	48	86.93	63	23.51	38	218	Carter5	139.60	58
SRPI	143.63	70	73.09	49	87.37	64	23.21	36	219	SR5	140.79	59
Datt5	152.53	73	65.66	40	47.88	46	29.80	61	220	Carter	146.10	60
SIPI	111.97	52	71.09	46	71.12	57	32.07	67	222	SR7	147.17	61
SPVI	128.86	59	121.16	72	80.92	60	23.34	37	228	TVI	147.48	62
MSAVI	127.13	58	114.20	66	90.95	68	23.52	39	231	SR4	154.82	63
dRE	138.30	66	69.85	45	54.04	51	33.32	72	234	Datt5	155.91	64
TVI	143.73	71	74.30	50	68.81	56	28.60	58	235	SIPI	163.31	65
OSAVI	129.07	60	114.51	67	90.12	67	23.77	42	236	mNDVI	163.99	66
NDVI	131.25	62	115.49	68	89.91	66	23.72	41	237	SRPI	170.71	67
EVI	138.08	65	94.45	57	92.31	70	24.53	46	238	dRE	171.12	68
RDVI	130.49	61	113.08	64	91.49	69	24.02	44	238	NPCI	171.62	69
SR	126.69	57	116.55	69	89.16	65	24.86	48	239	TCARI2	175.04	70
mNDVI	142.31	68	91.66	56	86.49	62	29.37	59	245	CARI	176.82	71

Sum_Dr2	140.62	67	111.04	63	93.32	71	25.56	51	252		Boochs	178.51	72
mSR	146.22	72	120.43	71	95.26	72	29.51	60	275		mSR	178.77	73

For index abbreviations and calculations, refer to Table 1.