Comparing the performance of multispectral vegetation indices and machine-learning algorithms for remote estimation of chlorophyll content: a case study in the Sundarbans mangrove forest

Hamed Gholizadeh*, Scott M. Robeson, and Abdullah F. Rahman

Department of Geography, Indiana University, Bloomington, IN 47405, USA; University of Texas Rio Grande Valley (UTRGV), South Padre Island, TX 78597, USA

(Received 3 July 2014; accepted 5 April 2015)

Optical vegetation indices (VIs) have been used to retrieve and assess biophysical variables from satellite reflectance data. These indices, however, also are sensitive to a number of confounding factors, such as canopy geometry, soil optical properties, and solar position. This suggests that VIs should be used cautiously for biophysical parameter estimation. Among biophysical variables, chlorophyll content is of particular importance as an indicator of photosynthetic activity. The goal of this study is to investigate the performance of multispectral optical VIs for chlorophyll content estimation in the world’s largest mangrove forest, the Sundarbans, and to compare these with machine-learning algorithms (MLAs). To this end, we have investigated the performance of 15 multispectral VIs and six state-of-the-art MLAs that are widely used for adaptive data fitting. The MLAs are Artificial Neural Networks (ANNs), Genetic Algorithm (GA), Gaussian Processes for Machine Learning (GPML), Kernel Ridge Regression (KRR), Locally Weighted Polynomials (LWP), and Multivariate Adaptive Regression Splines (MARS). We use an in situ data set of reflectance and chlorophyll measurements to develop and validate our models. Each MLA was evaluated 500 times with random partitions of training and validation data. Results showed that the weight optimization and term selection used within GA produce the most reliable chlorophyll content estimation. However, green normalized difference VI (GNDVI) is a simple and computationally efficient VI that produces results that are nearly as accurate as GA in terms of model fit and performance. Results also show that all methods except ANNs and MARS produce a quasi-linear relationship between spectral reflectance and chlorophyll content. Statistical transformations of GNDVI and chlorophyll content have the capability of further reducing model error.

1. Introduction

Mangrove forests are located in tropical coastlines. Despite their ecological importance, mangroves have been threatened by many factors such as sea level rise (Ellison 1993) and anthropogenic forcing (Walters et al. 2008). Due to their spatial extent and inaccessibility, studying mangroves using field surveys is time consuming and expensive. So, remote-sensing imagery of different types such as multispectral, hyperspectral, radar, and light detection and ranging (lidar) can be used for detailed mapping of mangrove forests (Green et al. 1998; Blasco, Aizpuru, and Gers 2001; Held et al. 2003; Heumann 2011). Another important application of remote sensing is providing information on vegetation biophysical parameters such as chlorophyll content (Gitelson et al. 2005; Jing et al. 2007; Zhang et al. 2012). Changes in chlorophyll content are indicators of vegetation growth,
environmental stress (Datt 1999), and vegetation gross primary production (Gitelson et al. 2006). One of the objectives of this paper is to investigate the potential of remote sensing in estimating chlorophyll content in mangrove ecosystems.

Spectral responses recorded by remote-sensing sensors can be used to monitor vegetation condition. At a higher level, combinations of spectral bands in different forms (e.g., linear and nonlinear) have been developed to highlight characteristics of vegetation such as productivity and leaf area index (LAI) (Baret, Guyot, and Major 1989). These combinations, hereafter called vegetation indices (VIs), are, however, sensitive to external factors such as sun angle, soil background, and atmosphere (Baret and Guyot 1991). Consequently, it is not clear which VI should be used to monitor vegetation status. In addition, these indices use the same calibration coefficients for all vegetation types.

For non-destructive estimation of chlorophyll content, several remote-sensing techniques, most of which employ VIs, have been developed (Gitelson and Merzlyak 1994; Lichtenthaler, Gitelson, and Lang 1996; Gitelson and Merzlyak 1997). However, as stated, the VIs, as a family of indicators, have some general limitations, and each VI has specific limitations. These limitations can be summed up as follows.

- External factors such as illumination condition, soil background, atmosphere, and canopy structure affect remotely sensed data and their products (Baret and Guyot 1991).
- VIs have not been calibrated for all species. In other words, the fact that the same coefficients are used for different species calls VIs into question. In addition, there is a higher performance when methods that incorporate the VIs are applied to the images that were used to develop the methods (Verrelst et al. 2010). Hence, as characteristics of leaves and canopy structure may vary across species, the application of these indices to a different vegetation type usually requires the development of a new empirical relationship.

As an alternative to VIs, machine-learning algorithms (MLAs) have the capability of establishing adaptive and robust relationships between input and output data (Hastie, Tibshirani, and Friedman 2009). Knudby, LeDrew, and Brenning (2010), for instance, concluded that MLAs have better performance than linear models. Verrelst et al. (2012) also tested the capability of four state-of-the-art MLAs for biophysical parameter retrieval: Artificial Neural Networks (ANNs), Support Vector Regression, Kernel Ridge Regression (KRR), and Gaussian Processes Machine Learning (GPML). The main conclusion of Verrelst et al. (2012) was that, in general, GPML yielded the best results. In this article, we compare VI-based approaches to six state-of-the-art MLAs including (1) ANNs (Hecht-Nielsen 1987; Haykin 1999; Huang 2003), (2) Genetic Algorithm (GA) (Holland 1975), (3) GPML (Rasmussen 2004), (4) KRR (Shawe-Taylor and Cristianini 2004), (5) Locally Weighted Polynomials (LWP) approximation (Cleveland and Devlin 1988), and (6) Multivariate Adaptive Regression Splines (MARS) (Friedman 1991). It is an important endeavour to identify whether MLAs show an appreciably better performance than VIs in estimating vegetation chlorophyll content. Using 500 random partitions of training and validation data sets, we compare the performance of 15 commonly used multispectral VIs and six state-of-the-art MLAs in terms of mean absolute error (MAE) and correlation coefficient. To investigate MLA performance, we consider two scenarios: a three-band scenario, which uses the green (G), red (R), and near-infrared (NIR) bands as input variables, and a two-band scenario in which bands G and NIR are used. Overall, the specific objectives of this article are: (1) investigating the performance of 15 multispectral
VIs in estimating chlorophyll content in mangrove ecosystems, (2) evaluating the ability of six MLAs to assess chlorophyll content for two different band-combination scenarios, and (3) identifying the nature of the relationship between chlorophyll content and reflectance.

2. Materials and methods

2.1. Spectral measurements and chlorophyll data

For this study, chlorophyll and reflectance data were collected in the Ganges-Brahmaputra delta region of coastal Bangladesh, which is home to the world’s largest single patch of mangroves known as the Sundarbans mangrove forest (Blasco, Aizpuru, and Gers 2001). The Sundarbans is a UNESCO World Heritage Site and covers approximately 140,000 ha on the Bay of Bengal. Important tree species in the study area are presented in Table 1 (Siddiqi 2001; Iftekhar and Islam 2004). The data collection protocol was as follows: first we located a small or medium-sized sunlit branch of a tree of one species in our plot. Using a tree pruner, we cut the branch at the base. On the ground, we quickly took CCM-300 Chlorophyll Content Meter and UniSpec measurements of at least five representative leaves of the branch. The CCM-300 Chlorophyll Content Meter measures chlorophyll and UniSpec measures the spectral reflectance factor. The representative leaves were selected visually. This process was repeated for all species of trees in a given plot and for all plots included in the study area.

A total of 745 reflectance measurements (before preprocessing) were collected at leaf level with a UniSpec spectrometer from nine different species during December 2012. Sampling was carried out at 15 plots distributed evenly in the study area. Outliers were excluded from the analyses and remaining spectral measurements were averaged for each species in each plot. Following the spectral measurements, the chlorophyll contents of the same species were measured a minimum of five times using the CCM-300 Chlorophyll Content Meter. The measurements include 880 CCM readings (before preprocessing). Similar to reflectance measurements, chlorophyll contents were averaged for each species in each plot after preprocessing. The CCM-300 Chlorophyll Content Meter device was calibrated in the laboratory and the CCM measurements were converted into leaf chlorophyll content expressed as the amount of chlorophyll in micrograms per square centimetre ($\mu$g cm$^{-2}$) using a laboratory-based calibration obtained by ordinary least squares (OLS). The chlorophyll levels as measured by the CCM-300 Chlorophyll Content Meter showed an excellent agreement with laboratory-measured chlorophyll content in terms of

<table>
<thead>
<tr>
<th>Local name</th>
<th>Scientific name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gewa</td>
<td><em>Excoecaria agallocha</em></td>
</tr>
<tr>
<td>Sundri</td>
<td><em>Heritiera fomes</em></td>
</tr>
<tr>
<td>Hodo (Tiger Fern)</td>
<td><em>Acrostichum aureum</em></td>
</tr>
<tr>
<td>Passur</td>
<td><em>Xylocarpus mekongensis</em></td>
</tr>
<tr>
<td>Amoor</td>
<td><em>Amoora cucullata</em></td>
</tr>
<tr>
<td>Shingra</td>
<td><em>Cynometra ramiflora</em></td>
</tr>
<tr>
<td>Hantal</td>
<td><em>Phoenix paludosa</em></td>
</tr>
<tr>
<td>Goran</td>
<td><em>Ceriops decandra</em></td>
</tr>
<tr>
<td>Keora</td>
<td><em>Sonneratia apetala</em></td>
</tr>
</tbody>
</table>
coefficient of determination \( R^2 = 0.93 \). After removing the outliers from the reflectance and chlorophyll content data and averaging the remaining measurements for each species in each plot, 67 reflectance and chlorophyll content data points for nine species from 15 plots were available. These spectral measurements coupled with chlorophyll content measurements are used to develop and test models to estimate chlorophyll content. To generate distributions of error, each MLA was evaluated 500 times with random subsets of training and validation data. For fair comparisons, the same training/validation points were used for all methods. It should be noted that the combined data set (i.e. reflectance spectra and chlorophyll content measurements from leaves) is divided into model development and model evaluation data sets: 60\% of the data points are used to train the MLAs while 40\% of the data points are used for model evaluation (i.e. the points that are not used in model development).

This study aims at utilizing multispectral data to estimate chlorophyll content at large scales, so reflectance measurements were resampled to those of Landsat Thematic Mapper 5 using the relative spectral response (RSR) of the sensor. We then investigated the performances of the widely used multispectral VIs (Table 2).

2.2. Machine-learning algorithms

The relationship between the input (e.g. reflectance) and output (e.g. chlorophyll) can be established using MLAs. Six MLAs are employed: (1) ANN, (2) GA, (3) GPML, (4) KRR, (5) LWP, and (6) MARS. These methods are chosen to cover different paradigms of MLAs, ranging from kernel-based methods (e.g. KRR) to evolutionary algorithms (e.g. GAs). It is noteworthy that the learning style of these methods is supervised.

2.2.1. Artificial neural networks

ANNs consist of units called neurons. The neurons receive an input signal and then send the processed signal to other units in the network. A comparison of neural networks with different network architectures can be found in Stathakis (2009). In the present work, we use feed-forward neural networks (Svozil, Kvasnicka, and Pospichal 1997). Neurons in the feed-forward neural networks can be grouped into three layers: (1) input, (2) hidden, and (3) output layers. The main advantage of multi-layer neural networks (including the feed-forward neural network) is their ability to model nonlinear relationships. Among the factors affecting the performance of an ANN are the number of layers and the number of neurons in each layer. So, we consider two feed-forward ANNs with different architecture: (1) a one-hidden layer neural network, \( \text{ANN}(1) \), with \( 2p + 1 \) nodes as per Hecht-Nielsen (1987), where \( p \) is the number of input nodes (i.e. spectral bands); and (2) a two-hidden-layer neural network, \( \text{ANN}(2) \), with \( \sqrt{(q + 2)N} + 2\sqrt{N/(q + 2)} \) nodes in the first layer and \( q\sqrt{N/(q + 2)} \) nodes in the second layer according to Huang (2003), where \( N \) is the number of input samples and \( q \) is the number of output neurons (i.e. 1). In this study, we apply the Levenberg–Marquardt training algorithm. To calculate the output of a neuron, a hyperbolic tangent sigmoid transfer function is used. More background on neural networks can be found in Bishop (1995), Haykin (1999), Simpson (1990), and Atkinson and Tatnall (1997).
Table 2. Definition of vegetation indices.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Name</th>
<th>Vegetation index</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNDVI</td>
<td>Green normalized difference vegetation index</td>
<td>$\frac{\rho_{NIR} - \rho_{G}}{\rho_{NIR} + \rho_{G}}$</td>
<td>Gitelson, Kaufman, and Merzlyak (1996)</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalized difference vegetation index</td>
<td>$\frac{\rho_{NIR} - \rho_{R}}{\rho_{NIR} + \rho_{R}}$</td>
<td>Rouse et al. (1974)</td>
</tr>
<tr>
<td>RVI</td>
<td>Ratio vegetation index</td>
<td>$\rho_{R}$</td>
<td>Pearson and Miller (1972)</td>
</tr>
<tr>
<td>SAVI*</td>
<td>Soil-adjusted vegetation index</td>
<td>$\rho_{NIR} - \rho_{G} + L$ $(1 + L)$</td>
<td>Huete (1988)</td>
</tr>
<tr>
<td>GCI</td>
<td>Green chlorophyll index</td>
<td>$\frac{\rho_{NIR}}{\rho_{R}} - 1$</td>
<td>Gitelson et al. (2005)</td>
</tr>
<tr>
<td>RGV1</td>
<td>Red–green ratio vegetation index</td>
<td>$\rho_{R}$</td>
<td>Jordan (1969)</td>
</tr>
<tr>
<td>DVI</td>
<td>Difference vegetation index</td>
<td>$\rho_{NIR} - \rho_{R}$</td>
<td>Jordan (1969)</td>
</tr>
<tr>
<td>MSAVI</td>
<td>Modified soil-adjusted vegetation index</td>
<td>$\frac{\rho_{NIR} - \rho_{R}}{\rho_{NIR} + \rho_{G} + L}$ $(1 + L)$</td>
<td>Qi et al. (1994)</td>
</tr>
<tr>
<td>RDVI</td>
<td>Renormalized difference vegetation index</td>
<td>$\sqrt{\frac{\rho_{NIR} - \rho_{R}}{\rho_{NIR} + \rho_{R}}}$</td>
<td>Roujean and Breon (1995)</td>
</tr>
<tr>
<td>TVI</td>
<td>Triangular vegetation index</td>
<td>$60(\rho_{NIR} - \rho_{G}) - 100(\rho_{R} - \rho_{G})$</td>
<td>Broge and Leblanc (2001)</td>
</tr>
<tr>
<td>PVI</td>
<td>Perpendicular vegetation index</td>
<td>$\rho_{NIR} - \rho_{R}$</td>
<td>Richardson and Weigand (1977)</td>
</tr>
<tr>
<td>SAVI2</td>
<td>Second soil-adjusted vegetation index</td>
<td>$\rho_{NIR} / (\rho_{R} - \frac{b}{a})$</td>
<td>Major, Baret, and Guyot (1990)</td>
</tr>
<tr>
<td>TSAVI</td>
<td>Transformed soil-adjusted vegetation index</td>
<td>$\frac{a(\rho_{NIR} - ap_{R} - b)}{ap_{NIR} + \rho_{R} - ab}$</td>
<td>Baret, Guyot, and Major (1989)</td>
</tr>
<tr>
<td>ATSAVI**</td>
<td>Adjusted transformed soil-adjusted vegetation index</td>
<td>$\frac{a(-ap_{R} - b)}{(ap_{NIR} + \rho_{R} - ab + X(1 + a^{2}))}$</td>
<td>Baret and Guyot (1991)</td>
</tr>
<tr>
<td>GEMI</td>
<td>Global environmental monitoring index</td>
<td>$\eta(1 - 0.25\eta - (\rho_{R} - 0.125)/(1 - \rho_{R})$</td>
<td>Pinty and Verstraete (1992)</td>
</tr>
</tbody>
</table>

Notes: *$L$ is used to minimize soil background effect, which depends on the amount of vegetation cover. For dense vegetation $L$ is 0 (i.e. SAVI = NDVI), and for no vegetation cover $L$ is 1. Huete (1988) suggested $L = 0.5$.

**$X$ is an adjustment factor used to reduce background effects.

$\rho$ denotes reflectance. $a$ and $b$ are the coefficients of the soil line. These values were proposed by Huete (1988) ($L = 0.5$), Baret and Guyot (1991) ($X = 0.08$), and Richardson and Everitt (1992) ($a = 0.96916$, $b = 0.084726$).
2.2.2. GA-based approach

Each VI uses a different combination of spectral bands (Datt 1998, 1999; Gitelson et al. 2005; Zhang et al. 2008) and there is no global relationship with chlorophyll content. In other words, as leaf characteristics vary for different species, the relationship of chlorophyll to a VI needs to be re-established (usually using least-squares regression). To attempt to improve on this approach, we propose a mathematical model based on GAs that determines the optimal relationship between spectral bands and chlorophyll content. Since most of the conventional indices are in the form of the reflectance ratios and use combinations of G, R, and NIR bands, the most general form of our proposed mathematical model is in the form of a rational function (RF) with 13 terms, as shown in Equation (1). To capture differential responses of the dependant variable (i.e. chlorophyll), interaction terms have also been considered:

\[
\text{Chlorophyll content} = \frac{A_0 + A_1 \rho_G + A_2 \rho_R + A_3 \rho_{NIR} + A_4 \rho_G \rho_R + A_5 \rho_G \rho_{NIR} + A_6 \rho_R \rho_{NIR}}{1 + A_7 \rho_G + A_8 \rho_R + A_9 \rho_{NIR} + A_{10} \rho_G \rho_R + A_{11} \rho_G \rho_{NIR} + A_{12} \rho_R \rho_{NIR}},
\]

where \(\rho_G\), \(\rho_R\), and \(\rho_{NIR}\) denote reflectance in the green, red, and NIR bands, respectively. The ‘\(A\)’ coefficients in Equation (1) represent weights of the spectral bands. Not all 13 terms in Equation (1) are used, as GA selects the optimal form of Equation (1). Like other VIs, the coefficients of this model are not necessarily physically meaningful but provide an optimal weighting for each band. Making decisions about the optimum combination of spectral bands to estimate chlorophyll content is a search and optimization problem. Since the number of possible combinations of RF terms is large and not all should necessarily be incorporated into the final solution, GA is a natural choice to solve this problem. Introduced by Holland (1975), GAs are adaptive methods inspired by evolution. The GA solution leads to a RF that establishes the optimal relationship between reflectance and chlorophyll content. The coefficients of the RF are determined using nonlinear least squares, and fitness of the RF is evaluated by MAE of the estimated chlorophyll content. Results of our study show that choosing different GA fitness functions including MAE, correlation coefficient, refined index of agreement (\(d_r\)) (Willmott, Robeson, and Matsuura 2012), and root mean square error (RMSE) does not affect the results. Willmott and Matsuura (2005) concluded that MAE is a more natural measure of average error, and unlike RMSE is insensitive to error distribution or sample size, so MAE is used as our GA cost function. In our case, the chromosome length of GA is equal to the number of RF terms (i.e. 13). Initial populations were randomly generated. For parent selection, binary tournament selection is used. Moreover, we use two-point crossover. The crossover probability and mutation rate are 0.8 and 0.05, respectively. To establish a trade-off between accuracy and computational cost of the method, the number of generations chosen is 200. The algorithm stops before 200 generations if there is no improvement in the fitness function for 50 consecutive generations. Grefenstette (1986) concluded that GA is so robust that the values of GA parameters are not particularly critical, but the fitness function and coding scheme play an important role in the convergence of GA. Goldberg and Holland (1988) provide more detailed background information on GA.

2.2.3. GPMLs

GPMLs, which use a Bayesian approach, can be considered as a generalization of the Gaussian distribution where instead of fitting distributions to data vectors, distributions
are fitted to mean and covariance functions. For some data sets, using linear models for regression may result in a poor prediction performance. In such cases, using GPML can be a more reliable choice (Rasmussen 2004). A Gaussian process (GP) consists of a finite number of variables and any subset of these variables has a joint Gaussian distribution. A GP can be written as \( f \sim \text{GP}(m, k) \); the random function \( f \) is distributed as a Gaussian process with mean function \( m \) and covariance function \( k \) (Rasmussen 2004). The MATLAB toolbox used in this study provides several options for mean and covariance functions. The parameters of these functions (called hyperparameters) are optimized in the training stage (Rasmussen and Nickisch 2010). In this paper, the hyperparameters are optimized via a conjugate gradient (CG) optimizer. A good introduction to GPML can be found in Williams (1998) and Rasmussen and Williams (2006). Here, we use the GPML MATLAB toolbox available from http://gaussianprocess.org/gpml/code/matlab/.

2.2.4. KRR

Least squares is a popular solution to a regression problem. This solution, however, is vulnerable to overfitting issues, which can be detected by parameters with large values. To address this issue, a regularization term is added to the cost function, which controls the values of the estimated parameters. The KRR method (Saunders, Gammerman, and Vovk 1998) solves this cost function by mapping the data to a new feature space with higher dimensions. The advantage of KRR is that it has a closed-form solution (Cawley, Talbot, and Chapelle 2006). In this article, the KRR parameters are optimized via cross-validation. We refer the interested reader to Cristianini and Shawe-Taylor (2000) and Scholkopf and Smola (2002) for a more detailed introduction to kernel methods.

2.2.5. LWP

Despite their simplicity, linear models cannot be used to solve the regression problem for data of a nonlinear nature. Nonparametric models such as LWP approximation (Cleveland and Devlin 1988) can be applied in such cases. Let’s assume the following model:

\[
y = F(x) + \epsilon,
\]

where \( y \) is the dependent variable, \( F \) is the regression function, \( x \) is the input variable vector, and \( \epsilon \) is random error. The goal is to approximate the regression function using the points within the neighbourhood of the query point. The points are weighted using a weighting function which is usually in the form of a Gaussian function. In LWP, the regression is done via fitting lower-degree polynomials to the data. Consider a first-degree polynomial:

\[
F(x) = a_0 + \sum_{j=1}^{d} a_j x_j,
\]

where \( d \) is the size of the input variable vector (i.e. number of input bands) and \( x_j \) is the \( j \)th element of vector \( x \). Using \( n \) training observations, the coefficient vector \( a \), which contains the coefficients of the polynomial in Equation (3), is calculated by minimizing
\[ a = \arg \min_a \sum_{i=1}^{n} W(x_{\text{query}}, x_i) (F(x_i) - y_i)^2, \]  

(4)

where \( W \) is a weight function, \( x_i \) is the reflectance vector of the \( i \)th point of the training data, \( y_i \) is the chlorophyll content of the \( i \)th point of the training data, and \( x_{\text{query}} \) is the query point (which is also a reflectance vector). In this article, the degree of the polynomials is set equal to 1. The MATLAB toolbox for LWP is available from http://www.cs.rtu.lv/jekabsons/. More detailed introduction to LWP regression can be found in Moore, Schneider, and Deng (1997) and Ruppert and Wand (1994).

2.2.6. MARS

MARS (Friedman 1991) is another nonparametric regression model. MARS divides the data into smaller portions, and fits basis functions (BFs) to each of these portions (endpoints of each of these portions are called knots). The relationship between the independent variables and dependent variable is established using the BFs and the corresponding coefficient for each BF. MARS consists of two steps: forward and backward. In the forward step, MARS defines large number of knots to maximize the model fit (i.e. overfitting). In the second step, the BFs with the least contribution to the model fit are removed. The generalized cross-validation (GCV) criterion is used to assess the model fit. One of the main advantages of MARS is that local interactions between the variables can be considered in the model (Leathwick, Elith, and Hastie 2006). In this study, the maximal number of BFs included in the model in the forward step is set equal to 20 and only one-level interactions are specified. The MATLAB toolbox is available from http://www.cs.rtu.lv/jekabsons/.

3. Results and discussion

We first focus on evaluating the relationship between chlorophyll content and multispectral VIs using correlation analysis. Then, we evaluate MLAs for two- and three-band scenarios and compare these results to the optimal VI in terms of the correlation coefficient and MAE. Independent predictions are evaluated on the validation data sets. Finally, the sensitivity of different values of spectral reflectance and estimated chlorophyll content, as well as the impact of data transformation, are investigated.

3.1. Multispectral VIs

The results presented in Table 3 show the mean correlation coefficient of 500 runs between VIs and chlorophyll content, as well as 95% confidence intervals (CIs). For each run, 40 sample points are chosen randomly and then the correlation coefficient is calculated. The results show that green NDVI (GNDVI) is the most highly correlated with chlorophyll content measurements, a finding that is in accordance with Shanahan et al. (2003). As GNDVI is the VI that consistently has the strongest relationship with chlorophyll, we compare its results to those of the MLAs.

3.2. Two- and three-band scenarios

Here, we estimate the relationships between chlorophyll content and spectral reflectance for two different scenarios using (1) G and NIR bands and (2) G, R, and NIR bands.
<table>
<thead>
<tr>
<th>VI</th>
<th>GNDVI</th>
<th>NDVI</th>
<th>RVI</th>
<th>SAVI</th>
<th>GCI</th>
<th>RGVI</th>
<th>DVI</th>
<th>MSAVI</th>
<th>RDVI</th>
<th>TVI</th>
<th>PVI</th>
<th>SAVI2</th>
<th>TSAVI</th>
<th>ATSAVI</th>
<th>GEMI</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>r</em></td>
<td>0.720</td>
<td>0.499</td>
<td>0.456</td>
<td>0.433</td>
<td>0.671</td>
<td>0.297</td>
<td>0.500</td>
<td>0.399</td>
<td>0.221</td>
<td>0.295</td>
<td>0.485</td>
<td>0.516</td>
<td>0.500</td>
<td>0.422</td>
<td></td>
</tr>
<tr>
<td>Lower bound</td>
<td>0.614</td>
<td>0.303</td>
<td>0.276</td>
<td>0.233</td>
<td>0.560</td>
<td>0.096</td>
<td>0.314</td>
<td>0.197</td>
<td>0.017</td>
<td>0.094</td>
<td>0.289</td>
<td>0.331</td>
<td>0.311</td>
<td>0.230</td>
<td></td>
</tr>
<tr>
<td>Upper bound</td>
<td>0.826</td>
<td>0.695</td>
<td>0.636</td>
<td>0.632</td>
<td>0.782</td>
<td>0.498</td>
<td>0.686</td>
<td>0.601</td>
<td>0.425</td>
<td>0.495</td>
<td>0.681</td>
<td>0.702</td>
<td>0.690</td>
<td>0.614</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Correlation coefficients (*r*) between chlorophyll content and vegetation indices, and the corresponding 95% confidence intervals.
Results for GNDVI (which, of course, uses just two bands) and the MLAs for both training and validation sets are shown in Figures 1–4. It is important to note that we use training data to establish the relationship between GNDVI and chlorophyll content using OLS. In other words, in this article ‘GNDVI’ refers to the relationship between GNDVI and chlorophyll content established via OLS. As a predefined VI, GNDVI itself does not need training but a different OLS relationship must be estimated for each training data set.

Figure 1. Correlation coefficients between measured and estimated chlorophyll content for two- and three-band scenarios using training data set. Error bars: mean ± 1 standard deviation (500 runs).

Notes: *The results of the optimized RF are shown for both scenarios. **‘GNDVI’ here refers to the relationship between GNDVI and chlorophyll content established via OLS. GNDVI results are the same for both scenarios, as GNDVI uses just two bands.

Figure 2. Correlation coefficients between measured and estimated chlorophyll content for two- and three-band scenarios using validation data set. Error bars: mean ± 1 standard deviation (500 runs).
Correlation coefficients between measured chlorophyll and estimated chlorophyll and MAE of estimated chlorophyll show that the three-band scenario minimally improves the training performance (i.e. correlations are slightly higher and MAEs are slightly lower for the three-band scenario). But what is important is the performance of the models once validated. On the whole, results of the two-band scenario on the validation data set are better than the three-band scenario. As the bands in the two-band scenario are the same as those bands used in GDVI (i.e. G and NIR), we are able to investigate directly the capability of the MLAs in comparison to GNDVI.

Figure 3. MAE of estimated chlorophyll content for two- and three-band scenarios using training data set. Error bars: mean ± 1 standard deviation (500 runs).

Figure 4. MAE of estimated chlorophyll content for two- and three-band scenarios using validation data set. Error bars: mean ± 1 standard deviation (500 runs).

Correlation coefficients between measured chlorophyll and estimated chlorophyll and MAE of estimated chlorophyll show that the three-band scenario minimally improves the training performance (i.e. correlations are slightly higher and MAEs are slightly lower for the three-band scenario). But what is important is the performance of the models once validated. On the whole, results of the two-band scenario on the validation data set are better than the three-band scenario. As the bands in the two-band scenario are the same as those bands used in GDVI (i.e. G and NIR), we are able to investigate directly the capability of the MLAs in comparison to GNDVI.
The error bars in the bar plots show the stability of different methods over 500 runs. The length of the error bars (mean ± 1 standard deviation) is larger once validation data are used. All methods except ANNs performed stably when applied to both validation and training data sets. The poor performance of ANNs is due to the fact that several parameters (including network architecture parameters) have to be tuned, which may greatly impact ANN prediction performance. This result is in agreement with Verrelst et al. (2012). Given the validation results, the GA method gives the best results in terms of correlation coefficient and MAE. This is likely due to the flexible optimization procedure of GA, which removes redundant terms. LWP, GPML, and KRR rendered validation results that are close to those of GNDVI. Mean correlation coefficients and MAEs of 500 runs are tabulated in Table 4. Overall, GA and LWP provide the best results on the validation data set, with mean correlation coefficients of 0.752 and 0.727, respectively. Inspection of mean MAE of the validation data set also reveals that the mean MAE of GA is 2.767 µg cm$^{-2}$, whereas that of LWP is 2.899 µg cm$^{-2}$. These validation results show the marginally superior performance of GA. GNDVI yielded a correlation coefficient of 0.726 with a MAE of 2.964 µg cm$^{-2}$ and, given the very low computation cost of GNDVI, these results demonstrate the effectiveness of GNDVI in estimating chlorophyll content.

For each of the 500 validation data sets used, we show scatterplots of the correlation coefficients between the estimated chlorophyll content via MLAs and field measurements and correlation coefficients between GNDVI and field measurements (Figure 5). Scatterplots for ANN(1) and ANN(2) were similar, so the scatterplot of ANN(1) is presented here. As can be seen in Figure 5, GA has a performance that is typically better than GNDVI (i.e. its correlations are usually above the 1:1 line). Despite its better performance, GA is much more computationally complex than GNDVI. The long computational time of GA is due to a multi-part optimization step during training. Similar to previous results, ANN and MARS perform poorly here in the estimation of chlorophyll content. It should be noted that the limited number of data points for model development and validation may explain the highly variable performance of many of the models runs.

### 3.3. Sensitivity analysis

In this section, the sensitivity of each estimated model to a wide range of G and NIR values is analysed using response surfaces (and contour plots). The analysis allows us to use the models for chlorophyll content estimation across the full range of reflectance factor values and, therefore, to visualize the linearity/nonlinearity of each model. In other words,

<table>
<thead>
<tr>
<th></th>
<th>ANN(1)</th>
<th>ANN(2)</th>
<th>GNDVI</th>
<th>GA</th>
<th>GPML</th>
<th>KRR</th>
<th>LWP</th>
<th>MARS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean correlation coefficient</td>
<td>Training data</td>
<td>0.755</td>
<td>0.736</td>
<td>0.720</td>
<td>0.774</td>
<td>0.747</td>
<td>0.747</td>
<td>0.742</td>
</tr>
<tr>
<td></td>
<td>Validation data</td>
<td>0.615</td>
<td>0.516</td>
<td>0.726</td>
<td>0.752</td>
<td>0.721</td>
<td>0.716</td>
<td>0.727</td>
</tr>
<tr>
<td>Mean MAE (µg cm$^{-2}$)</td>
<td>Training data</td>
<td>2.676</td>
<td>2.737</td>
<td>2.853</td>
<td>2.632</td>
<td>2.700</td>
<td>2.680</td>
<td>2.680</td>
</tr>
<tr>
<td></td>
<td>Validation data</td>
<td>3.685</td>
<td>4.030</td>
<td>2.964</td>
<td>2.767</td>
<td>2.947</td>
<td>2.968</td>
<td>2.899</td>
</tr>
</tbody>
</table>
sensitivity analysis shows the nature of the relationship between different values of reflectance factors and chlorophyll content. Figure 6 shows the sensitivity of chlorophyll to selected values of G and NIR reflectance factors. Since each method has been run 500 times, and it is not possible to show all 500 realizations here, the realization with the correlation coefficient that is closest to the median correlation coefficient of the 500 runs is selected and shown here for each method. In other words, the ‘typical’ model corresponding to the median correlation coefficient is used to estimate chlorophyll content for each method (Figure 6). Interestingly, only ANNs and MARS show a response surface that is clearly nonlinear. The absolute difference of the estimated chlorophyll content using GNDVI (the best VI) and GA (the best MLA) is shown in Figure 7. As expected, for those regions covered with training data, the difference is low (1–2 µg cm\(^{-2}\)). Outside the training space, the two methods diverge. As a nonlinear optimization method, GA should be used cautiously when far outside the data envelope, which further reinforces the robustness of the results using GNDVI.

3.4. Impact of data transformation on the relationship between GNDVI and chlorophyll content

The relationship between GNDVI and chlorophyll content is generally assumed to be linear (Shanahan et al. 2003). Consequently, it is worth considering whether the performance of GNDVI can be further improved by transforming the data. To this end, we applied the Tukey ladder-of-powers transformation (Tukey 1977) to both GNDVI and chlorophyll content for a
wide range of transformation parameters ($\lambda = [-6, 6]$) and then re-estimated the correlation coefficient (see Figure 8). The Tukey transformation is defined as

$$
Y = \begin{cases} 
X^\lambda & \text{if } \lambda > 0, \\
\ln(X) & \text{if } \lambda = 0, \\
-(X^\lambda) & \text{if } \lambda < 0,
\end{cases}
$$

(5)

where $X$ denotes the original data and $Y$ denotes the transformed data.

The contour plot in Figure 8 shows that there is an optimal transformation of both GNDVI and chlorophyll content that leads to an even higher correlation coefficient ($r = 0.782$) than the untransformed data ($r = 0.726$). The same is true for a simple ln–ln transformation ($r = 0.756$). Scatter plots of these transformations, however, show that extreme transformations, such as the optimal one illustrated here, may cause the data to become more non-normally distributed (Figure 9). Overall, transformations have the potential to further improve chlorophyll content estimation using GNDVI, but careful examination of the implications of the transformation is warranted.
4. Conclusions

In this study, we investigated the application of multispectral data for estimating chlorophyll content in coastal mangroves. We used six MLAs and compared their performance...
to various VIs. Results showed that, among the selected VIs, GNDVI had the highest correlation with chlorophyll content. It also was observed that weight optimization and term selection via a machine-learning approach, GA, slightly improved chlorophyll content estimation on the validation data set compared with GNDVI. LWP provided the second best result on the validation data set, but was only slightly better that GNDVI. While machine-learning approaches such as GA show considerable promise, GNDVI is a simple yet computationally efficient VI that has a consistently high correlation with chlorophyll content. In addition, data transformation might further improve the

Figure 9. Scatterplots showing the impact of transformations on the relationship between chlorophyll and GNDVI: (a) original data, (b) ln(GNDVI) vs. ln(chlorophyll), and (c) transformed data using Tukey transformation with $\lambda = \arg \max_\lambda \text{correlation}(\text{GNDVI}, \text{measured chlorophyll})$. 
performance of GNDVI in estimating chlorophyll content to the point where it performs even better than the machine-learning approaches. In sum, our findings suggest that GNDVI is an accurate, simple, and robust approach for estimating chlorophyll content using multispectral data. The fact that it had comparable or better performance to some very complex MLAs leads us to conclude that GNDVI is a very reliable and efficient proxy for chlorophyll content in mangroves. Nonetheless, there still may be opportunities for improving on conventional VIs in the estimation of chlorophyll content in other environments.

Acknowledgements
The authors are indebted to Dr Jonathan Karty (Mass Spectrometry Facility at Indiana University) for kindly calibrating the CCM-300 Chlorophyll Content Meter. We wish to thank Dr Chunfeng Huang (Department of Statistics at Indiana University) for providing helpful comments.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the National Aeronautics and Space Administration (NASA) [grant number NNX11AF50G] to AFR.

References


Jing, L., J. Jinbao, C. Yunhao, W. Yuanyuan, S. Wei, and H. Wenjiang. 2007. “Using Hyperspectral Indices to Estimate Foliar Chlorophyll a Concentrations of Winter Wheat under Yellow Rust


