Int J Appl Earth Obs Geoinformation 73 (2018) 616–626

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Informative bands used by efficient hyperspectral indices to predict leaf biochemical contents are determined by their relative absorptions

Jia Jin[a], Quan Wang[a,b,*]

[a] Faculty of Agriculture, Shizuoka University, Shizuoka 422-8529, Japan
[b] Research Institute of Green Science and Technology, Shizuoka University, Shizuoka 422-8529, Japan

ARTICLE INFO

Keywords:
Hyperspectral indices
Robustness
Informative bands
Mechanism
Biochemical
RTM

ABSTRACT

Hyperspectral indices have proven to be useful for quantifying various plant parameters including a number of biochemical components. However, it remains a great challenge to unveil the physical and physiological mechanisms of wavelengths used by these statistically/empirically identified indices, and this unveiling is a critical step towards developing generally applicable indices. In this study, we have introduced a dummy variable to the well-accepted leaf scale radiative transfer model (PROSPECT-4) and have conducted a series of virtual experiments by tweaking different absorption features (intensity, peak wavelength, and half-width) of the dummy variable for investigating the underlying mechanisms of those wavelengths used by efficient indices. Results clearly indicated that the informative bands used in efficient indices, instead of the target parameter's absorption peaks, are generally concentrated among wavelengths that the targeted parameter has relatively high specific absorption coefficients compared with other biochemical components. This finding has also been validated by replacing the dummy variable with carotenoids based on PROSPECT-5B simulated dataset. Further analysis reveals that although the concentrations of chlorophyll are higher than other pigments in most leaves, the commonly recommended wavelength of 550 nm may be disturbed by other pigments and may only be applicable when anthocyanins are minor. Results obtained in this study have largely explained why specific wavelengths are used in efficient indices, and thus this should lay a basis for understanding the underlying mechanisms and help to develop robust indices for estimating vegetation biochemical parameters.

1. Introduction

Biochemical components, including pigments, water, other nitrogen-rich compounds (e.g. proteins), and structural materials (lignin and cellulose) of foliage are important indicators of a series of ecosystem processes (Roberts et al., 2016; Wang et al., 2009). Their spatiotemporal patterns are highly needed by a variety of ecological applications (Verrelst et al., 2015). Studies in the past have shown that remote sensing provides valuable insights into ecosystem monitoring and offers probably the only practical solution through which foliar biochemical parameters can be monitored over a high spatial-temporal resolution (Thenkabail et al., 2012; Wang et al., 2009). Especially, hyperspectral remote sensing is receiving increasing attention for its usefulness in quantifying plant parameters (Ben-Dor et al., 2013).

Traditionally, three approaches have been attempted to retrieve terrestrial vegetation biochemical components from hyperspectral remote sensing data: (a) statistically/empirically based methods searching for statistical relationships between spectral information and vegetation biochemical parameters; (b) physically-based methods via Radiative Transfer Model (RTM) inversion; and (c) hybrid methods using physically-based models' outputs to establish statistical relationships (Dorigo et al., 2007; Verrelst et al., 2015). Among them, statistically/empirically based methods are the simplest and most popular approach being applied in hyperspectral remote sensing so far (Li and Wang, 2012; Liang, 2005), ranging from simple/multiple linear regression, stepwise regression, partial least squares regression, to artificial neural network and so on, by linking *in situ* measurements of vegetation properties (e.g. pigments/nutrients/water contents) with reflectance values or vegetation indices (Blackburn, 1998; Claudio et al., 2006; Hansen and Schjoerring, 2003; Huang et al., 2004; Inoue et al., 2012; Madeira et al., 2000; Nguyen and Lee, 2006; Ramoelo et al., 2013; Ryu et al., 2011; Sims and Gamon, 2003; Sonobe and Wang, 2017; Wang and Li, 2012; Yi et al., 2014). By far, statistical regressions based on vegetation indices (VIs) are the oldest, most widely studied and employed approaches in plant biochemical properties for their easy manipulation (Glenn et al., 2008; Verrelst et al.,

* Corresponding author.
E-mail address: wang.quan@shizuoka.ac.jp (Q. Wang).

https://doi.org/10.1016/j.jag.2018.08.002
Received 17 May 2018; Received in revised form 19 July 2018; Accepted 2 August 2018

2015; Yue et al., 2018).

Vegetation indices usually consist of several reflected or transformed spectral bands to reduce disturbance while enhancing spectral features sensitive to a vegetation property (Bannari et al., 1995; Glenn et al., 2008; Verrelst et al., 2015). During the last 40 years, numerous indices under different formulae and band combinations have been proposed (Haboudane et al., 2004; le Maire et al., 2004; Main et al., 2011; Roberts et al., 2016; Wang and Li, 2012; Zarco-Tejada et al., 2005; Zhao et al., 2005). Despite the fact that they estimated the targeted parameters effectively, most of them were calibrated or tested on one or a few species on the basis of statistics (le Maire et al., 2008), resulting in their poor performances when applied to a wide variety of plant leaves and conditions, and leading to a difficulty of reaching a consensus on a "best" index (Féret et al., 2011).

Tremendous efforts have been invested towards generating "universal" or "robust" indices, by combining experimental data of various species as well as synthetic datasets simulated with physical models, for addressing the problem of low reliability of spectral indices generated from small-sized datasets (Féret et al., 2011; le Maire et al., 2004, 2008; Main et al., 2011; Sonobe and Wang, 2017; Wang and Li, 2012). However, development of universal (widely applicable) and robust (with physical and biological mechanisms) vegetation indices for leaf biochemical parameters still remains a big challenge. Various results showed that there was a high bias between simulations and measurements, resulting in a difficulty in selecting the "best" index (le Maire et al., 2008). Furthermore, wavelengths used by these indices are purely selected on the basis of statistics, while their underlying physical and biological mechanisms have been studied far less, making it difficult to arrive at a consensus on their robustness.

Two pathways might reveal the underlying mechanisms of a given vegetation index, via either control experiments or radiative transfer models. In the past, diverse control experiments have been designed for physically and physiological interpreting sensitive wavelengths used by vegetation indices, such as the classical experiments presented in Gamon et al. (1990). Based on the results, one of the most studied "pigment-related" spectral index, Photochemical Reflectance Index (PRI) (Gamon et al., 1992), was developed (Féret et al., 2017). The success of the index is due preliminary to the spectral band of 531 nm used, which is related to leaf xanthophyll cycle proved through inhibitor dithiothreitol (DTT) treatment experiments (Gamon et al., 1990). Using 531 nm and 570 nm enables the index to capture the physiological response of the vegetation to a short-term environmental stress inducing slight changes in photosynthetic activity (Féret et al., 2017; Gamon et al., 1997). Unfortunately, such laboratory experiments are not only insufficient currently but are also costly and time-consuming. Alternatively, the physical-law based radiative transfer models such as PROSPECT (Feret et al., 2008; Féret et al., 2017; Jacquemoud and Baret, 1990), which can model endless variations of reflectance by tweaking key input parameters, allow a better understanding of the influence of vegetation biochemical parameters on reflectance (Féret et al., 2017; Main et al., 2011). Therefore, these models offer a chance to reveal the underlying mechanisms of varying spectral information and shall improve our understanding of the relationships between the vegetation biochemical parameters and the reflectance derived spectral indices in a more systematic way (Feret et al., 2008; Féret et al., 2017). However, even though the sensitivity of reflectance bands to pigments and the performance of VIs have been the major subject of previous studies, no comprehensive study is available yet to fully discuss the wavelengths used by those indices, which is a critical step towards developing sensitive and robust indices.

To date, most studies devoted to leaf pigments have focused on chlorophyll, since the pigment has the strongest effect on reflectance in the visible spectral regions and can be estimated with relatively good accuracy using spectral indices (Féret et al., 2017). A number of published chlorophyll indices were reviewed in le Maire et al. (2004) and Main et al. (2011). Interestingly, most wavelengths used are situated between 670 and 800 nm (Main et al., 2011), instead of the maximum absorption wavelengths of around 450 and 680 nm. The reason for such incompatibility remains unclear. In addition, why so many wavelengths are being used has received less attention, even though it has been projected that locally accompanying biochemical components play important roles. Further, specific sensitive wavelengths for other pigments are far less investigated because of the overlaps between their absorptions with that of chlorophyll (Gitelson et al., 2002; Sims and Gamon, 2002). The underlying mechanisms behind wavelengths for estimating other biochemical components are thus far seldom being investigated.

In this study, we attempt to address the aforementioned questions through an analysis of the results obtained from a series of virtual experiments, based on the well-accepted leaf scale radiative transfer model PROSPECT (Jacquemoud and Baret, 1990). By introducing a dummy variable with convertible specific absorption coefficients into the original PROSPECT-4 model (Feret et al., 2008), we have produced plentiful simulations by tweaking different absorption features (intensities, peak wavelengths, and half-widths) of the dummy variable. We aim to reveal the wavelengths used by those efficient hyperspectral indices to predict different leaf biochemical parameters and to explore the relationships between these wavelengths with absorption features for a better understanding of the underlying mechanisms of those efficient indices.

2. Materials and methods

2.1. Experiment design and data simulation

In order to investigate the relationships between the wavelengths used by hyperspectral indices and the absorption features of the biochemical components, we have designed virtual experiments based on PROSPECT-4 by adding a dummy input variable (C_{dummy}) with convertible specific absorption coefficients. PROSPECT is a radiative transfer model for simulating the leaf reflectance and transmittance from 400 nm to 2500 nm (Feret et al., 2008; Jacquemoud and Baret, 1990). In its PROSPECT-4 version, the reflectance is calculated as a function of leaf structure index (N), leaf chlorophyll content (C_{ab}, μg/cm^2), leaf water content (C_w, g/cm^2), leaf mass area (C_m, g/cm^2), and the specific absorption coefficients K of each component (le Maire et al., 2004). The total absorption coefficient at wavelength λ ($k(\lambda)$) for one layer is expressed as:

$$k(\lambda) = \frac{C_{ab}}{N}\cdot K_{ab}(\lambda) + \frac{C_w}{N}\cdot K_w(\lambda) + \frac{C_m}{N}\cdot K_m(\lambda) \quad (1)$$

where $K_{ab}(\lambda)$, $K_w(\lambda)$, and $K_m(\lambda)$ are the specific absorption coefficients at wavelength λ of total chlorophyll, water, and dry matter, respectively.

In this study, following Feret et al. (2008), we generated the specific absorption coefficients of the intentionally added dummy variable (K_{dummy}) using Gaussian function :

$$K_{dummy}(\lambda) = a\cdot e^{-\frac{(\lambda-b)^2}{2\cdot c^2}} \quad (2)$$

where a is the height of the absorption peak, b is the wavelength of the absorption peak, and c is the standard deviation which controls the width of absorption bands, whose value was generated from the predetermined half-width (W) (Tsai and Philpot, 1998):

$$c = \frac{W}{2\cdot(2\cdot ln2)^{\frac{1}{2}}} \quad (3)$$

Thus, the total absorption coefficient at wavelength λ ($k(\lambda)$) for one layer in our modified PROSPECT-4 is expressed as: