BIOSYSTEMS ENGINEERING 160 (2017) 69–83

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Comparison of different uni- and multi-variate techniques for monitoring leaf water status as an indicator of water-deficit stress in wheat through spectroscopy

Bappa Das [a,*], *Rabi N. Sahoo* [a], *Sourabh Pargal* [a], *Gopal Krishna* [a], *Rakesh Verma* [b], *Viswanathan Chinnusamy* [b], *Vinay K. Sehgal* [a], *Vinod K. Gupta* [a]

[a] *Division of Agricultural Physics, Indian Council of Agricultural Research — Indian Agricultural Research Institute, New Delhi 110012, India*

[b] *Division of Plant Physiology, Indian Council of Agricultural Research — Indian Agricultural Research Institute, New Delhi 110012, India*

ARTICLE INFO

Article history:
Received 30 December 2016
Received in revised form
22 May 2017
Accepted 27 May 2017

Keywords:
Hyperspectral reflectance
Spectral indices
Relative water content
Water deficit stress
Continuum removal
Multivariate models

Ten different wheat genotypes were studied for understanding their differential behaviour to different water-deficit stress levels. Hyperspectral data (350–2500 nm) and relative water content (RWC) of plants were measured at different stress level for identifying optimal spectral bands, indices and multivariate models to develop non-invasive phenotyping protocols. Evaluation of water sensitive existing spectral indices, proposed indices and band depth analysis at selected wavelengths was done with respect to RWC and prediction models were developed. The prediction models developed were efficient in predicting RWC with the R^2 values ranging from 0.73 to 0.88 for spectral indices and 0.74–0.85 with continuum depth. Then, the ratio spectral indices (RSI) and normalised difference spectral indices (NDSI) were obtained in all possible combinations within 350–2500 nm and their correlations with RWC were quantified to identify the best indices. The best spectral indices for estimating RWC in wheat were RSI (R_{1391}, R_{1830}) and NDSI (R_{1391}, R_{1830}) with R^2 of 0.86 and 0.81, respectively. Spectral reflectance data were also used to develop partial least squares regression (PLSR) followed by multiple linear regression (MLR), support vector machine regression (SVR), multivariate adaptive regression spline (MARS) and random forest (RF) model to calculate plant RWC. Among these multivariate models, PLSR was the best model for prediction of RWC (R^2 and RMSE of 0.96 and 3.88%; 0.91 and 6.52% for calibration and validation, respectively). The methodology developed would help for its further use in high-throughput phenomics of different crops for drought stress.

* *Corresponding author*. Present address: Section of Natural Resource Management, Indian Council of Agricultural Research — Central Coastal Agricultural Research Institute, Old Goa, Goa 403403, India. Fax: +91 832 2285649.
E-mail address: bappa.das@icar.gov.in (B. Das).

http://dx.doi.org/10.1016/j.biosystemseng.2017.05.007

1. Introduction

Production of winter wheat (*Triticum aestivum* L.) is often limited by water-deficit stress across many parts of the world, due to non-availability of timely rainfall or irrigation water. It has been, therefore, a continuous search for wheat cultivars that can withstand the stress, and perform well under scarce water supply condition. One of the prerequisites for the development of stress-tolerant cultivars is accurate phenotyping of the germplasm and the breeding population. Crop water status could be effectively employed as a phenotyping parameter for selection of genotypes in breeding programmes (Munjal & Dhanda, 2005), monitoring crop growth under water-deficit stress condition (Peñuelas, Gamon, Griffin, & Field, 1993; Tucker, 1980) and advocating water management to prevent the stress (Köksal, 2008). Crop water status is represented by the leaf water per unit mass of fresh (Garnier & Laurent, 1994) or dry (Chuvieco, Riaño, Aguado, & Cocero, 2002) leaves. This may also be expressed as leaf water per unit of leaf area (g cm^{-2}), designated as equivalent water thickness (EWT) (Datt, 1999). The former is usually preferred as leaf area is difficult to measure, particularly in conifer needles (Cheng, Rivard, & Sánchez-Azofeifa, 2011; Ullah et al., 2014). While conventional phenotyping methods for water-deficit stress are time and labour intensive, and often destructive, remote sensing techniques could be used (Hunt, Rock, & Nobel, 1987; Peñuelas et al., 1993; Sepulcre-Cantó et al., 2006). Literature indicates that the use of multispectral satellite sensor for the detection of crop water status have serious limitations like low spatial and spectral resolution and larger revisit time (Berni, Zarco-Tejada, Suarez, & Fereres, 2009; Hunt & Rock, 1989; Pierce, Running, & Riggs, 1990). Aerial or ground-based hyperspectral sensors could be effective (Eitel, Gessler, Smith, & Robberecht, 2006; Goetz, Vane, Solomon, & Rock, 1985), considering that these sensors are able to capture the information in narrow contiguously spaced electromagnetic bands. This allows in deriving the spectral indices using specific wavelengths that are sensitive to minor fluctuations in crop water status (Gao, 1996; Horler, Docray, & Barber, 1983; Peñuelas et al., 1993). For instance, electromagnetic energy is weakly absorbed in the near infrared (NIR) (720–1000 nm) (Curcio & Petty, 1951; Palmer & Williams, 1974) and strongly absorb in the short-wave infrared (SWIR) (1400–1900 nm) (Datt, 1999; Thomas, Namken, Oerther, & Brown, 1971) region by water molecules present in the leaves. So the retrieval of leaf water content using NIR region of the electromagnetic spectrum is less effective compared to the SWIR (Datt, 1999). Several researchers reported that the water absorption features between 1400 and 1900 nm wavebands are strongly correlated to leaf water content (Bowman, 1989; Ceccato, Flasse, & Grégoire, 2002; Ceccato, Flasse, Tarantola, Jacquemoud, & Grégoire, 2001; Hunt & Rock, 1989).

A continuum removal (CR) approach was used to understand the leaf biochemistry using laboratory spectroscopy, which was further extended to remotely sensed observation (Clark et al., 2003; Kokaly, 1999). Specific absorption characteristics can be isolated from the continuum (Cheng et al., 2011). The continuum, or background, is the overall albedo of the reflectance curve (van der Meer, 2004) and provides an estimation of the materials causing absorption in a sample. When the continuum is removed the sections where the spectral curve approaches the continuum are scaled to 100% reflectance. The continuum removed spectra gives an indication of the shape, depth and width of the absorption feature. The band depths (BDs) and the area under the curve can then be correlated to various biophysical and biochemical characteristics of the plants.

Larger information on crop water status can be extracted only by investigating the entire spectrum. However, use of an entire spectrum of hyperspectral data has a problem of high dimensionality and high degree of multi-collinearity in the associated data (Vaiphasa, Ongsomwang, Vaiphasa, & Skidmore, 2005). Multiple linear regression (MLR) models developed from hyperspectral data generally suffers from multi-collinearity and are often over-fitting as in numbers of observations could be equal or lesser than the predictors (Curran, 1989). On the contrary, partial least squares regression (PLSR), is a combination of principal component analysis (PCA) and MLR, and can effectively reduce the multi-collinearity (Cho, Skidmore, Corsi, van Wieren, & Sobhan, 2007). The PLSR also reduces the problem of over-fitting. This is a robust technique for predicting leaf biochemical traits from hyperspectral data (Thomas & Haaland, 1990), and can be used to understand the relationship between canopy reflectance and foliar properties under varying canopy structures (Asner & Martin, 2008). PLSR was also successfully used with spectral data for plant moisture content prediction (Johansson, Hagman, & Fjellner, 2003), estimation of chlorophyll and water content in wheat leaves (Ji, Wang, & Yan, 2007), estimation of leaf area index and chlorophyll content (Darvishzadeh et al., 2008), soil salinity at laboratory-, field- and airborne-scales (Farifteh, Van der Meer, Atzberger, & Carranza, 2007). But very few studies are there where a comparison of various multivariate models has been done for retrieving relative water content (RWC) taking a diverse set of genotypes of wheat. In the present investigation, PLSR along with MLR, support vector machine regression (SVR), multivariate adaptive regression spline (MARS) and random forest (RF) is used for quantitative estimation of leaf water content from the hyperspectral reflectance spectra. SVR, MARS and RF are more suitable when the relationship between spectra and properties to be modelled is not linear (Bao et al., 2014; Ding & Peng, 2012). On this backdrop, an effort is being made to develop a high throughput non-destructive hyperspectral spectroscopy based phenotyping method to access plant water status in wheat under water-deficit stress with the objectives to (i) evaluate the performance of spectral indices for high throughput phenotyping of leaf water status in comparison to conventional methods; (ii) assess and compare the retrieval accuracy of leaf water status when using five different multivariate models.

2. Materials and methods

Thirty leaf samples each belonging to 10 different wheat genotypes viz. Ajanta, CBW12, Datatine, HD2402, HD2009, KRL14, KRL19, Wilgoyne, WR544 and WR704 (each with