Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 247 (2021) 119104

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

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

Evaluation of different water absorption bands, indices and multivariate models for water-deficit stress monitoring in rice using visible-near infrared spectroscopy

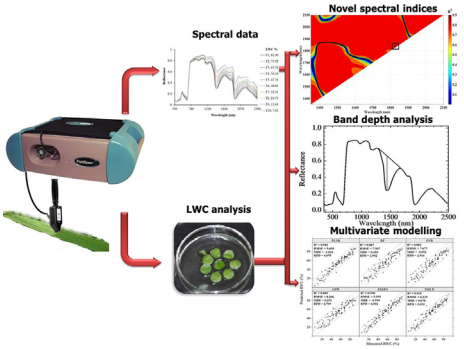
Bappa Das [a,*], Rabi N. Sahoo [a], Sourabh Pargal [a], Gopal Krishna [a], Rakesh Verma [b], Chinnusamy Viswanathan [b], Vinay K. Sehgal [a], Vinod K. Gupta [a]

[a] Division of Agricultural Physics, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India
[b] Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India

HIGHLIGHTS

- The novel spectral indices accurately estimated the changes in LWC due to water-deficit stress.
- Bands centered on 1400 nm revealed the best relationship with LWC at different stress levels.
- Among the six multivariate models, PLSR was found to be the best model for prediction of LWC.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 August 2020
Received in revised form 6 October 2020
Accepted 13 October 2020
Available online 24 October 2020

Keywords:
VNIR spectroscopy
Phenotyping
Spectral indices
Leaf water content
Water-deficit stress
Multivariate models

ABSTRACT

Accurate estimation of plant water status is a major factor in the decision-making process regarding general land use, crop water management and drought assessment. Visible-near infrared (VNIR) spectroscopy can provide an effective means for real-time and non-invasive monitoring of leaf water content (LWC) in crop plants. The current study aims to identify water absorption bands, indices and multivariate models for development of non-destructive water-deficit stress phenotyping protocols using VNIR spectroscopy and LWC estimated from 10 different rice genotypes. Existing spectral indices and band depths at water absorption regions were evaluated for LWC estimation. The developed models were found efficient in predicting LWC of the samples kept in the same environment with the ratio of performance to deviation (RPD) values varying from 1.49 to 3.05 and 1.66 to 2.63 for indices and band depths, respectively during validation. For identification of novel indices, ratio spectral indices (RSI) and normalised difference spectral indices (NDSI) were calculated in every possible band combination and correlated with LWC. The best spectral indices for estimating LWC of rice were RSI (R_{1830}, R_{1834}) and NDSI (R_{1830}, R_{1834}) with R^2 greater than 0.90 during training and validation, respectively. Among the multivariate models, partial least squares regression (PLSR) provided the best results for prediction of LWC (RPD = 6.33 and 4.06 for training and validation, respectively). The approach developed in this study will also be helpful for high-throughput water-deficit stress phenotyping of other crops.

* Corresponding author at: Section of Natural Resource Management, ICAR-Central Coastal Agricultural Research Institute, Old Goa, Goa 403403, India.
E-mail address: bappa.das@icar.gov.in (B. Das).

https://doi.org/10.1016/j.saa.2020.119104
1386-1425/

1. Introduction

Plant water-deficit stress is one of the major limiting factors for final crop yield and is usually detected only after it becomes visually apparent. The detections are often too late to escape crop yield loss. Till date, majority of the scientists have used soil moisture sensors to detect water-deficit stress in different crop species. Though these sensors may fairly assess the imposed water-deficit stress, they fail to represent in most cases the level of water-deficit stress originally experienced by the plants [1]. Moreover, such measurements are time-consuming and not that easy to employ at field level, thus limiting their wide scale applicability. Contrarily, assessing the leaf water content (LWC) may provide a better picture of plant's physiological response to varying level of soil moisture scarcity and consequent water-deficit stress actually experienced by plants. In agriculture, it will be useful for the crop irrigation scheduling and crop yields estimation at regional scale [2]. Field scale estimation of LWC is time-consuming, laborious and highly location specific. Remote sensing, on the other hand, provides an effective non-destructive alternative for field sampling of plants' leaf for water status retrieval with the continuous coverage on a large area [3]. However, previous reports revealed the considerable limitations of multispectral satellite sensor images to monitor the variations in plant water content [4]. Airborne or ground spectrometers have greater spatial and spectral resolution to detect the spectral changes corresponding to changes in plant water status with more accuracy [5]. Arial hyperspectral sensors and field spectrometers are able to capture the electromagnetic spectrum at very narrow contiguous wavebands. This feature permits the remote estimation of plant water status from canopy reflectance as the latter changes with the former [6].

Retrieving LWC based on visible-near infrared (VNIR) spectroscopy has been mostly studied using narrow band indices vegetation indices (VIs) [7–12]. Krishna et al. [13] reported very good accuracy of hyperspectral normalized difference and ratio vegetation indices for estimation of LWC in rice with coefficient of determination (R^2) greater than 0.80 during validation. In addition to the simple ratio and normalized difference indices, Cao et al. [14] evaluated single wavelength, difference and double difference hyperspectral indices for estimation of leaf water status in deciduous broadleaved species from leaf dehydration experiment. They reported that first order derivative reflectance based simple ratio and normalized difference indices performed better than other indices. Das et al. [15] reported superior performance of water sensitive spectral indices containing adjustable wavebands as compared to the indices with fixed reference wavebands. Lausch et al. [16] compared greenness, pigment sensitive, red edge, water sensitive, soil and atmosphere adjusted indices for estimation of leaf water content in barley. They found that normalized pigments reflectance index (NPCI) performed the best for leaf water content estimation which could be due to the indirect effect of water stress on chlorophyll content. But the main weakness of VIs is that they only utilise very little information (two or three narrow bands) available in the full spectrum of hyperspectral reflectance with the loss of significant information [17]. Alternatively, several studies in the recent past have focused on use of multi-variate models, such as stepwise multiple linear regression (SMLR), for the retrieval of biophysical and biochemical variables [18]. However, SMLR suffers from multi-collinearity when used for spectroscopic data analysis [19,20].

The partial least square regression (PLSR) is a powerful regression technique as compared to SMLR, as the technique reduces the problem of multi-collinearity [21]. The PLSR is also superior to principle component regression (PCR), because PCR only accounts for the variance of explanatory variables (e.g., reflectance) without considering the response variable (e.g., biophysical or biochemical parameters), while PLSR accounts for the both [22,23]. However, all the above mentioned chemometrics techniques assume the response variable to be normally distributed. Support vector machine regression (SVR) with kernel trick, gaussian process regression (GPR), multivariate adaptive spline regression (MARS) and random forest (RF) are non-parametric techniques and do not require such assumptions. But very few studies are conducted to compare various spectral indices based uni-variate techniques as well as full spectrum based parametric and non-parametric chemometrics techniques for retrieving the LWC with a diverse set of rice genotypes. In the present investigation, the performance of PLSR, SMLR, GPR, SVR, MARS and RF were assessed for quantitative retrieval of LWC from the leaf reflectance spectra. On this background, the present study was undertaken with the objectives to evaluate the prediction performance of indices, bands and six chemometric techniques (PLSR, SMLR, GPR, SVR, MARS and RF) for modeling the LWC in rice as compared to the conventional methods for developing a spectroscopy-based high-throughput non-destructive water-deficit stress monitoring protocol.

2. Materials and methods

For the present study, thirty leaf samples from 10 different rice genotypes (3 replicates each) were collected from well-watered experimental fields of the Division of Plant Physiology at the ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India in the month of September. The genotypes namely Vandana, Nagina 22, Sahbhagidhan, MTU 1010, Nerica-L-44, CR262-4, PS 5, IR 64, Pusa 44 and PB 6 were transplanted in the month of July 2014. Leaves from different rice genotypes were collected and kept in plastic sampling bags to avoid dehydration after harvesting. The samples were immediately transferred to the laboratory for estimation of LWC and spectral measurements. Around 3 g leaves were kept in glass tubes with distilled water at room temperature to attain full turgidity.

2.1. Collection of LWC and leaf spectra

The LWC estimation and spectral measurements were made at fully turgid condition, within two hours after the samples were harvested. The leaves were allowed to dry naturally under normal room conditions and spectral observations along with LWC estimation were made at 15 min interval. The LWC was computed using the following expression

$$LWC(\%) = \frac{(FW - DW)}{(TW - FW)} \times 100 \quad (1)$$

where FW is the fresh weight of the leaf samples taken after each 15 min time interval, DW is the dry weight taken after keeping the leaves in hot air oven at 60 °C until constant weight was obtained and TW is turgid weight taken after keeping the leaves in distilled water for 2 h and is expressed in percentage. For every genotype, ten readings were obtained at an equal time interval. The spectra at leaf level were captured in laboratory (i.e., dark room, ±25 °C) using contact probe with an internal light source attached to a spectroradiometer (Analytical Spectral Devices, Boulder, CO) in the spectral range of 350–2500 nm. Each leaf sample consisted of an overlapping pile of 2–3 leaves on spectrally black surface to eliminate the background interferences on the spectrum.

2.2. Spectral indices

For the quantification of the variations in LWC, a number of different spectral indices have been developed. The spectral indices considered in this study included the simple ratio [24], the normal-

ized difference vegetation index [25], the water index [26], the normalized water index-1, 2, [27] 3, 4 [28], the 1650/2220 nm ratio [29], the floating water band index [26], and the maximum difference water index [30]. The different water sensitive spectral indices used in this study are listed in Table 1.

2.3. Narrow band indices

The two narrow band indices namely ratio spectral index (RSI) and normalised difference spectral index (NDSI) with the following formulation were computed in all possible two-pair combinations of wavelengths.

$$RSI = \frac{R_{\lambda 1}}{R_{\lambda 2}} \quad (2)$$

$$NDSI = \frac{(R_{\lambda 1} - R_{\lambda 2})}{(R_{\lambda 1} + R_{\lambda 2})} \quad (3)$$

The calculated indices were then correlated with LWC and the square of correlation coefficient values was plotted in a matrix plot to identify areas with relatively high values. The indices with the highest coefficient of determination values were selected.

2.4. Band-depth (BD) normalization

Band depth at water absorption features centred around 970 nm (BD_{970}), 1200 nm (BD_{1200}), 1400 nm (BD_{1400}) and 1900 nm (BD_{1900}) [31] were calculated as:

$$BD_{\lambda i} = 1 - R_{\lambda i}$$

where *i* represents the 970, 1200, 1400 and 1900 nm wavelengths.

2.5. Multivariate techniques

Different multivariate techniques viz. PLSR, SMLR, GPR, SVR, MARS and RF were adopted to model LWC. The multivariate models were developed and validated using randomly divided calibration (70%) and validation (30%) dataset. The optimum number of latent variables (LV) for PLSR model was selected using leave one out cross-validation. Wavebands with the highest and lowest absolute loading weight values were identified for each LVs. These wavebands were used as input for SMLR model to identify the most important ones. MARS which is a nonparametric regression procedure was used to develop relationship between LWC and spectral data. SVR is the implementation of the support vector machine (SVM) method for regression and function approximation [32].

Table 1
Spectral indices for water-deficit stress monitoring.

Spectral index	Definitions	References
SR	$SR = \frac{R_{800}}{R_{680}}$	[24]
NDVI	$NDVI = \frac{(R_{800} - R_{670})}{(R_{800} + R_{670})}$	[25]
Water Index	$WI = \frac{R_{970}}{R_{900}}$	[26]
NWI-1	$NWI = \frac{(R_{970} - R_{900})}{(R_{970} + R_{900})}$	[27]
NWI-2	$NWI = \frac{(R_{970} - R_{850})}{(R_{970} + R_{850})}$	[27]
NWI-3	$NWI = \frac{(R_{970} - R_{880})}{(R_{970} + R_{880})}$	[28]
NWI-4	$NWI = \frac{(R_{970} - R_{920})}{(R_{970} + R_{920})}$	[28]
1650/2220 nm ratio	$1650/2220nmratio = \frac{R_{1650}}{R_{2220}}$	[29]
Floating Water Band Index (fWBI)	$fWBI = \frac{R_{900}}{minR_{920-980}}$	[26]
Maximum Difference Water Index	$MDWI = \frac{(Rmax_{1500-1700} - Rmin_{1500-1700})}{(Rmax_{1500-1700} + Rmin_{1500-1700})}$	[30]

The radial basis kernel was used. RF is a non-parametric regression technique based on the ensemble of classification and regression tree (CART) algorithm. The hyperparameters of RF like mtry, splitrule and min.node.size were optimized while the number of trees (ntree) was set to the default value of 500. The GPR is a non-parametric probabilistic kernel-based model. The advantage of GPR model is that they can provide reliable estimate of their own prediction uncertainty. In this study, PLSR, SVR, RF, GPR were implemented using "pls" [33], "kernlab" [34], "ranger" [35] and "kernlab" [34] packages in R software environment version 3.3.1 [36]. The hyperparameters of the machine learning models were optimized using tenfold cross validation and five repetitions with "caret" package [37].

The performance of the models was evaluated by the coefficient of determination (R^2), root mean square error of calibration (RMSEC), validation (RMSEP), mean bias error (MBE) and RPD (ratio of the standard deviation of the measured data (SD) to RMSEP).

3. Results and discussion

3.1. Variation in leaf water content

Substantial variations in the LWC were observed among the different rice genotypes. The distribution and mean of LWC for ten rice genotypes are shown in Fig. 1. The violin plot embedded with boxplots revealed that the LWC data was almost normally distributed. Among the rice genotypes, IR64 and PUSA44 showed the least variations in LWC over time. The descriptive statistics revealed that LWC in the calibration subset ranged between 6.93% and 91.81% with mean of 53.36% and standard deviation of 21.91% (Table 2); while in validation, the subset varied between 4.91% and 91.30% with the mean of 52.92% and standard deviation of 22.50%. These kinds of broad ranges are required to assess the predictive ability of the vegetation indices and multivariate models [38]. The normality of LWC for both calibration and validation dataset was analysed using Lilliefors test of normality, which revealed that the LWC was normally distributed (p-value > 0.05). The equality of mean, median and distribution of calibration and validation dataset was tested using *t*-test, Mann-Whitney and Kolmogorov-Smirnov analysis. The results showed that the datasets were having equal mean, median and distribution (p-value > 0.05).

3.2. Response of leaf reflectance to variations in leaf water status

In general, the dynamic changes in leaf reflectance were witnessed with varying LWC under different water-deficit stress levels (Fig. 2). Different stress levels significantly affected the spectral responses of rice leaves with differential spectral responses at various regions electromagnetic spectrum. The sensitive spectral range for LWC extends from 800 to 2500 nm. Minor variations in visible ranges are due to change in pigment content with different water stress levels. Fresh leaves exhibited the lowest reflectance throughout this spectral range and the reflectance increases with reduction in LWC (Fig. 2). As LWC decreases from the highest (82.49%) to the lowest (7.01%) value, the strong water absorption spectral characteristics at 1400 nm and 1900 nm become feebler, and the reflectance in the SWIR region intensifies. In addition, the 1670–1830 nm and 2000–2200 nm absorption features corresponding to leaf dry matter (e.g., protein, lignin and cellulose) [39], become more dominant with reducing LWC. The dynamism of spectral reflectance under different stress levels in various rice genotypes makes it possible to develop spectroscopy-based quantitative models for estimation of LWC.

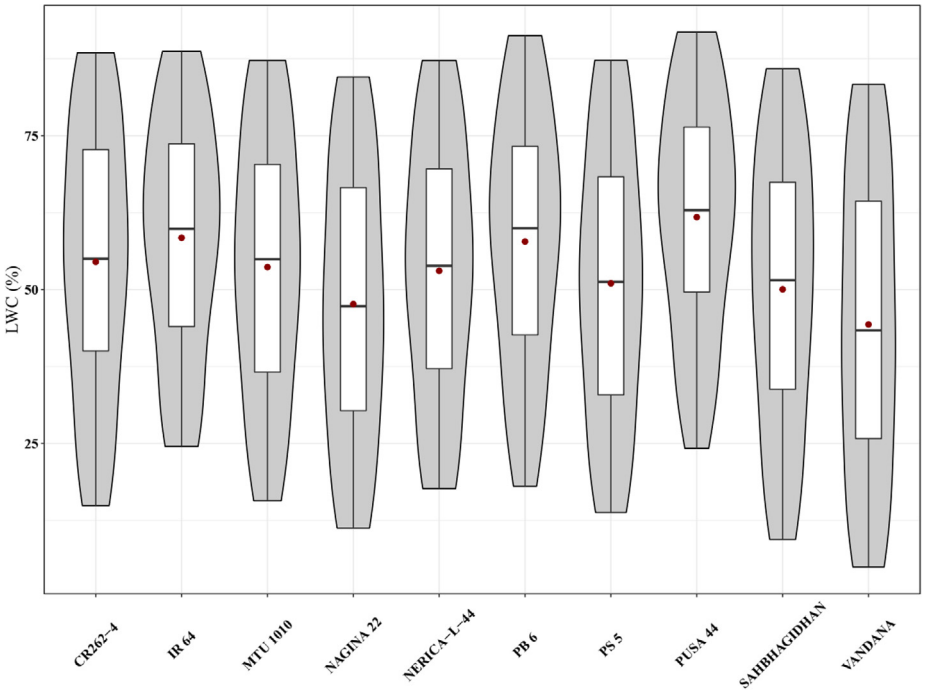


Fig. 1. Violin plot showing the means and spreads of leaf water content in different rice genotypes.

Table 2
Descriptive statistics of the LWC (%) of calibration and validation datasets.

	Sample size	Mean	Standard deviation	Min	Max	Skewness	Kurtosis	Lilliefors test of normality (p-value)
Calibration	200	53.36	21.91	6.93	91.81	−0.16	−0.99	0.11
Validation	100	52.92	22.50	4.91	91.30	−0.18	−1.01	0.29
Test					p-value			
Mann-Whitney					0.90			
t-test					0.87			
Kolmogorov-Smirnov					0.99			

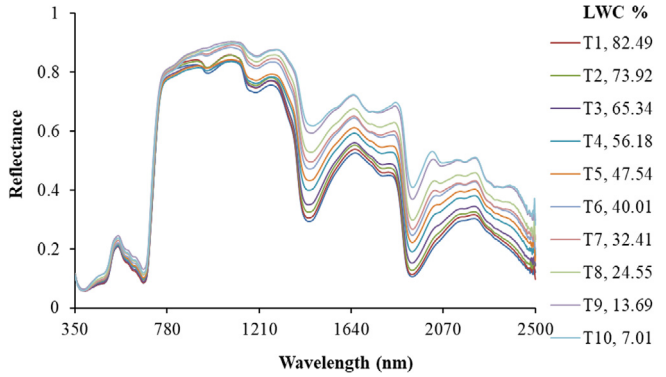


Fig. 2. Spectral reflectance of rice leaves under different water-deficit stress levels.

3.3. Leaf water content and spectral indices

The best performing regression models between various indices and water content expressed as LWC are given in Table 3. The strongest correlation of LWC was observed with MDWI with R^2 of 0.89 and RPD of 3.05 for validation. The lowest correlation was recorded with SR with R^2 0.55 and RPD of 1.49 during validation. Our result confirmed the poor correlation of traditional greenness indices such as NDVI and SR with leaf EWT while estimating the plant water content at leaf and landscape level [40,41]. The vegetation indices with stronger absorption bands in SWIR region are more responsive to variations in LWC than the indices having weaker water absorption bands from the NIR region at leaf level [40,42]. The better performance of MDWI is in line with the previous findings of Eitel et al. [30], Das et al. [15] and Krishna et al. [13]. However, this finding differs with Zhang and Zhou [43] who reported a better sensitivity of chlorophyll indices to the changes in canopy water status in maize.

3.4. Proposed new spectral index: RSI and NDSI ($R_{\lambda1}$, $R_{\lambda2}$)

The reduced sampling technique with the calibration dataset (70% samples) was adopted [44], aiming to analyze the relationship of LWC with Ratio Spectral Index (RSI) and Normalized Difference Spectral Index (NDSI) from 350 nm to 2500 nm at 10 nm intervals to identify the sensitive bands with greater R^2 values. A number of "hotspots" with high correlation coefficients of LWC with RSI and NDSI were found in the NIR and SWIR region as shown in Figs. 3 and 4(a). Similar kind of approach was successfully applied for LWC [15,45], nitrogen [44] and leaf area index [38] modeling in wheat.

Furthermore, through precise sampling of these sensitive spectral regions, more detailed contour maps of R^2 values between LWC and RSI; LWC and NDSI at 1 nm intervals were obtained (Figs. 3 and 4b). Based on R^2, the best two wavelengths at 1830 and 1834 nm were selected. Though this kind spectral indices formulation using close wavebands is not desirable due to high collinearity, were reported in many previous studies [46–49].

Table 3
Relationships of LWC to available spectral indices.

Hyperspectral indices	Model	$R^{2\ (Cal)}$	$R^{2\ (Val)}$	RMSEP	RPD
SR	y = 8.3x – 32.4	0.52	0.55	15.09	1.49
NDVI	y = 448.9x – 305.9	0.51	0.58	14.71	1.53
WI	y = −1207.8x + 1251.4	0.86	0.80	10.27	2.19
MDWI	y = 258.3x – 3.9	0.87	0.89	7.38	3.05
NWI-1	y = −2396.9x + 43.9	0.85	0.79	10.36	2.17
NWI-2	y = −1939.8x + 62.9	0.84	0.79	10.36	2.17
NWI-3	y = −2269.3x + 50.4	0.85	0.79	10.48	2.15
NWI-4	y = −2663.6x + 37.2	0.85	0.79	10.41	2.16
fWBI	y = 1373.4x – 1335	0.81	0.74	11.53	1.95
1650/2220 nm ratio	y = 139.9x – 176	0.82	0.84	9.13	2.45

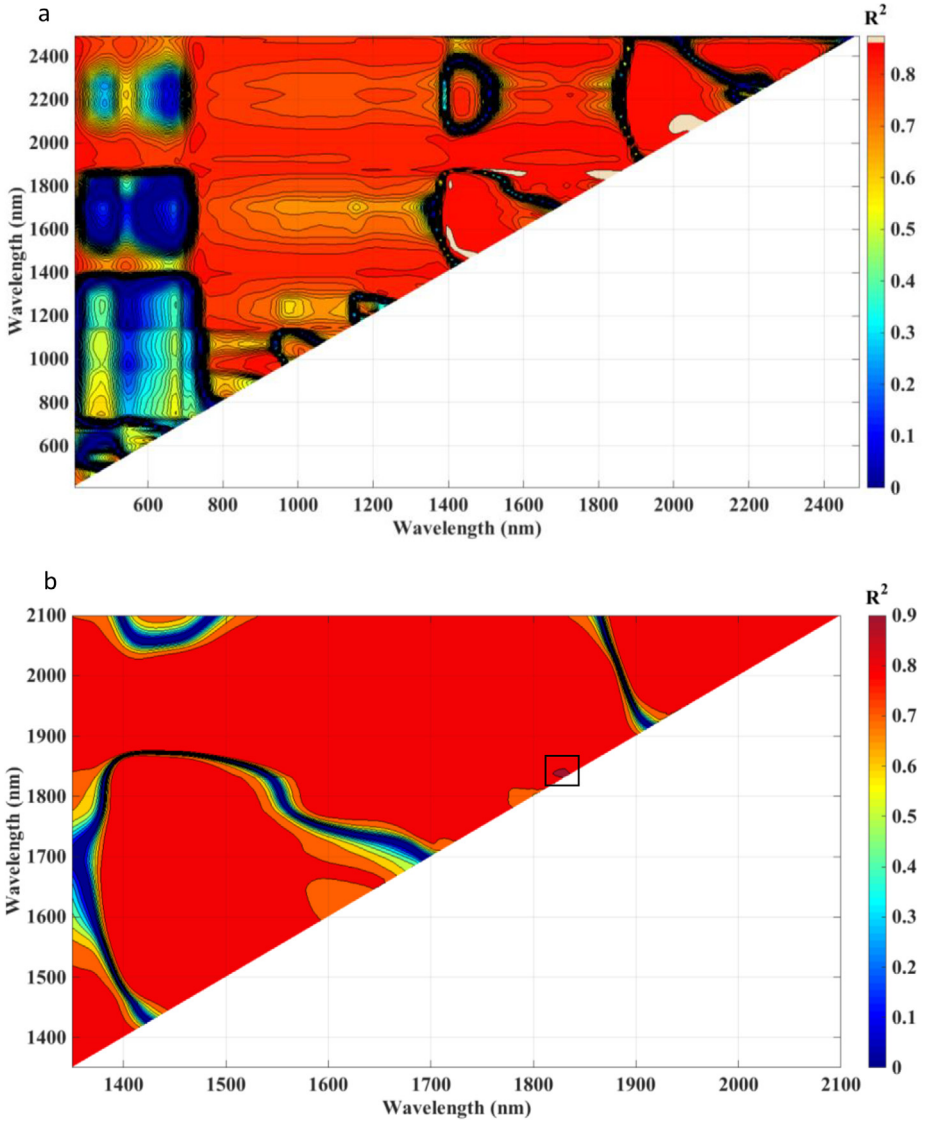


Fig. 3. Contour maps of coefficients of determination (R^2) for linear relationship between RSI and LWC, of wheat leaves under different water-deficit stress levels. (a) 10 nm sampling interval at 350–2500 nm; (b) 1 nm sampling interval. Black box indicates hotspot.

Empirical spectral models were calibrated based on the RSI and NDSI derived from the above identified wavelengths in the linear, power, exponential, logarithmic and 2nd order polynomial functional forms as shown in Table 4. Our results indicated the developed spectral indices and their spectral models were performing well irrespective of kind empirical relations whether derived from linear or non-linear equations. Among them, the 2nd order polynomial model for both RSI and NDSI were the best (R^2 = 0.93, 0.93, RMSEC = 5.81, 5.81; at $p < 0.01$ for RSI and NDSI, respectively) (Table 4).

3.5. Band depth analysis

Simple linear regression analysis was implemented to develop the relationship between LWC of different genotypes recorded at different stress levels and band depth at the water absorption

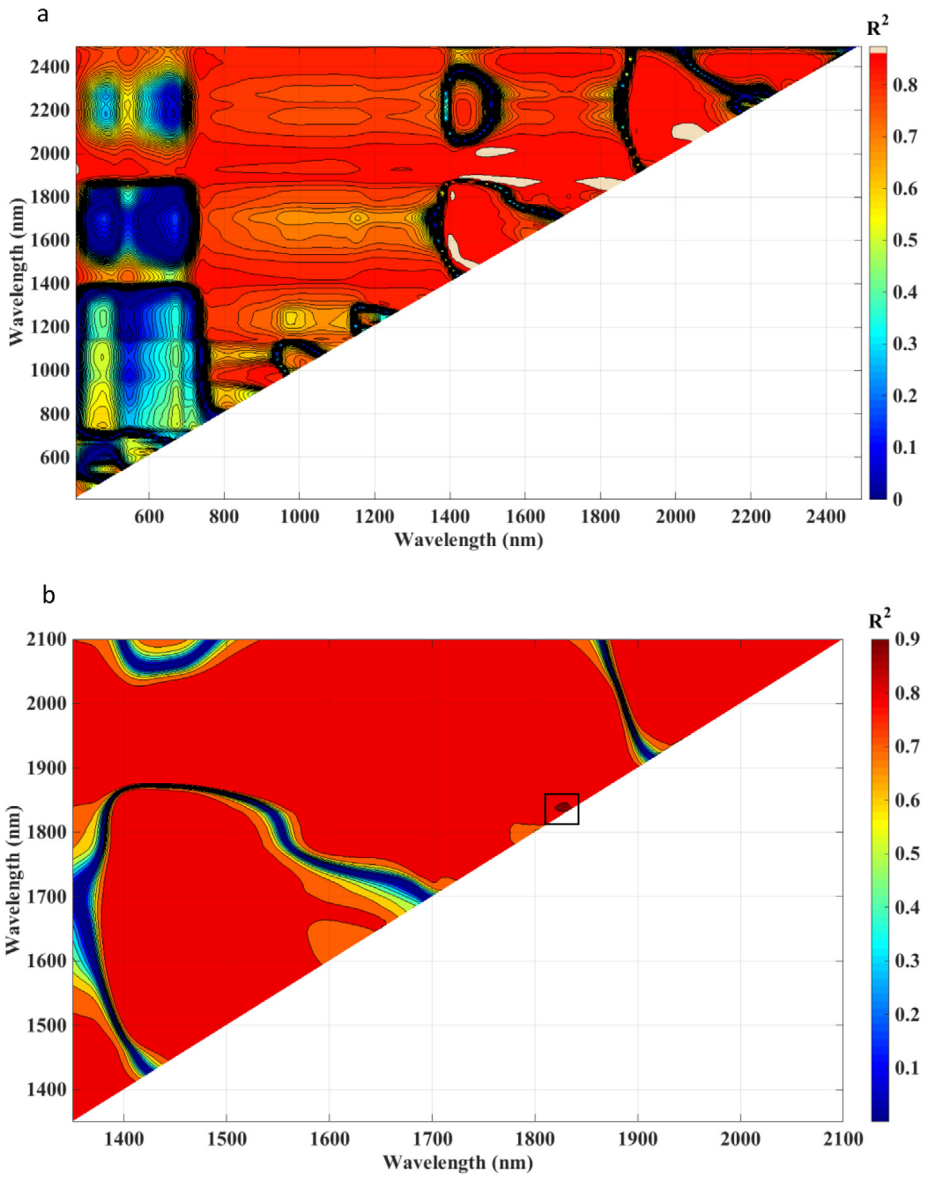


Fig. 4. Contour maps of coefficients of determination (R^2) for linear relationship between NDSI and LWC, of wheat leaves under different water-deficit stress levels. (a) 10 nm sampling interval at 350–2500 nm; (b) 1 nm sampling interval. Black box indicates hotspot.

Table 4
Relationships of LWC to the best spectral indices of RSI and NDSI (n = 200).

Spectral indices	Regression equations	R^2	RMSEC (%)
RSI (R_{1834}, R_{1830})	$y = 7445.5x - 7396.8$	0.93	5.82
	$y = 42.82x^{172.8}$	0.87	6.77
	$y = 4E\text{-}74e^{172.71x}$	0.87	6.78
	$y = 7446.9\ln(x) + 48.79$	0.93	5.84
	$y = 62851x^2 - 118284x + 55481$	0.93	5.81
NDSI (R_{1834}, R_{1830})	$y = -148.94x + 0.48$	0.92	5.82
	$y = 0.43e^{-345.6x}$	0.87	6.77
	$y = 2660.6x^2 - 148.38x + 0.48$	0.93	5.81

regions i.e. between 950 and 980 nm, 1130–1250 nm, 1300–1600 nm and 1830–2250 nm. It was found that out of several bands identified for water stress studies in rice, best R^2 values were obtained for bands at 1400 nm followed by at 1900 nm. The results of regression analysis between band depth and LWC emphasized that the correlation was the highest at major water absorption bands i.e. 1400 and 1900 nm as depicted in Fig. 5(a–d). Our results corroborate well with the previous findings where the band depths centred around 1400 and 1900 nm were reported as a splendid indicator of plant water status [41,50].

3.6. Evaluation of spectral models for LWC estimation

The reliability the developed models based on indices and band depth for LWC estimation in rice were analysed using 30% of the total data set. The performance of the established models was evaluated by statistical parameters such R^2, RMSEP, and RPD as shown in Fig. 6(a and b). It was found that the models derived from RSI (R_{1830}, R_{1834}) and NDSI (R_{1830}, R_{1834}) for LWC gave good accuracy, with R^2 of 0.95 and 0.95, RMSEP of 5.18 and 5.19, RPD of 4.35 and 4.33, respectively. Compared to the performance of the spectral models developed from previously reported spectral indices, the new proposed indices gave more stable and reliable results. Thus, the new spectral indices can be reliably used for the accurate estimation of changes in LWC caused by water-deficit stress.

The correlation between the measured and predicted LWC for band depths at different wavelengths, i.e. 970, 1200, 1400 and

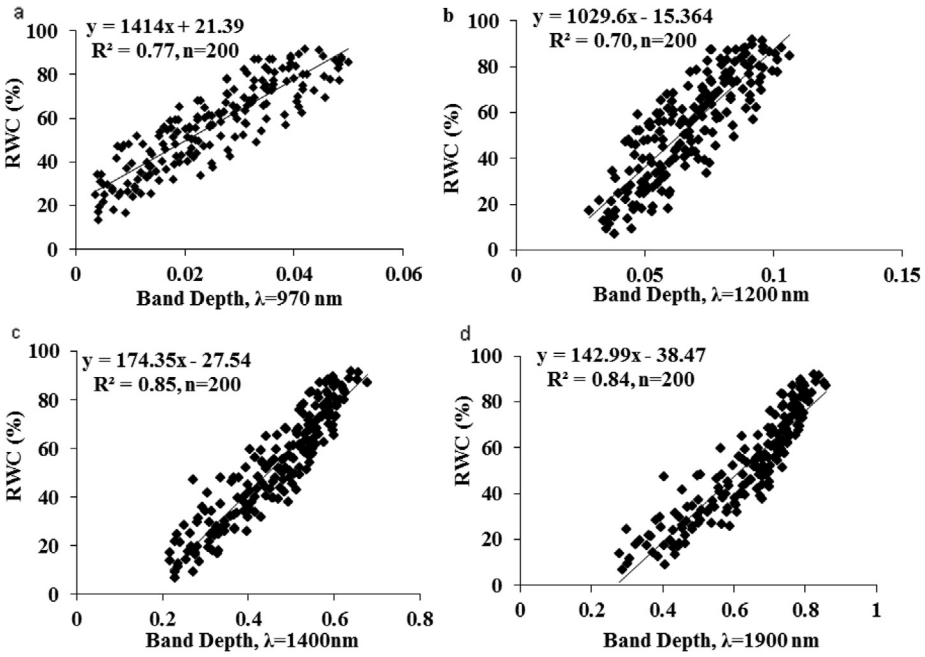


Fig. 5. The relationship between band depths and measured relative water content LWC (%) based on the model developed from band depth analysis; (a) 970 nm (b) 1200 nm (c) 1400 and (d) 1900 nm.

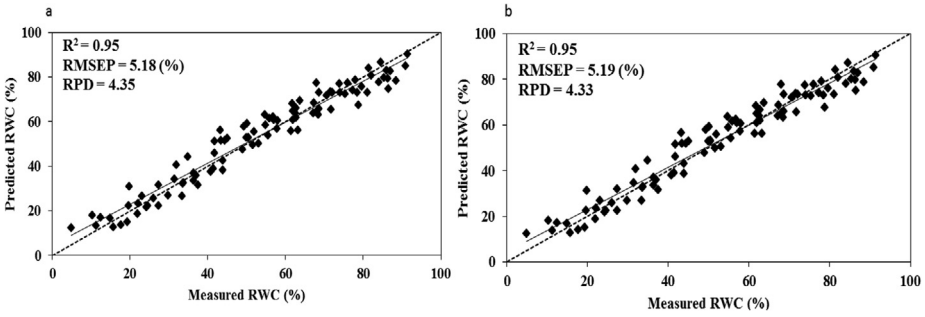


Fig. 6. The relationship between measured and predicted relative water content LWC (%) based on the model developed from (a) RSI and (b) NDSI. The solid line shows the regression line, while the dotted line is the 1:1 line.

1900 nm are depicted in Fig. 7(a–d). The regression model using band depth at 1400 nm performed the best compared to other water absorption features. This band provides an estimation of LWC of plant leaves that is very close to the measured LWC. Model developed from regressing band depth at 1400 nm with LWC was the most efficient for quantification of the LWC in rice.

Previous studies relating to retrieval of LWC through water sensitive indices indicated that retrieval accuracy reduces and becomes asymptomatic as the LWC exceeds 40% [39]. But in the present study using newly developed indices and band depths, the asymptomatic trend was observed above 70% LWC. Our results are in better agreement with Ullah et al. [51] who evaluated the three mid-wave infrared indices for retrieval of LWC with nine different plant species

3.7. Multivariate analysis of spectral model for LWC estimation

Partial least squares regression (PLSR) model with leave one out cross validation was performed using the leaf reflectance values and the measured LWC. The optimum number latent variables (LVs) was found to be 22 as indicated by the global minimum of root mean squared error of cross validation (RMSECV) for LWC modelling in rice. In the calibration, the R^2 and RPD value obtained for this study were 0.975 and 6.355, respectively (Table 5) and while validating, the R^2 value was 0.943 with the RPD of 4.078 (Fig. 8).

Later, the highest and lowest loaded variables from each LV were chosen as an important spectral variable for building SMLR model. The SMLR was not applied with the whole spectral dataset because under situation of larger number predictor variables than the sample size, SMLR suffers from over-fitting [52]. The results of the SMLR analysis for both calibration and validation are depicted in Table 5 and Fig. 8. The SMLR models for LWC estimation of rice is given below

$$y = 56.98 + 1792.3 \times R_{350} - 1355.2 \times R_{400} + 582.04 \times R_{441} - 693.53 \times R_{504} + 283.91 \times R_{685} - 469.09 \times R_{1359} + 698.76 \times R_{1590} - 605.79 \times R_{1869} + 470.82 \times R_{1883} - 303.9 \times R_{2391}$$

The model yielded higher R^2 and lower RMSE values for both calibration and validation (R^2 = 0.928, 0.925; RMSEC = 5.865 and RMSEP = 6.129, respectively). As expected, the SMLR model being having few important wavelengths was inferior to that of PLSR, which used the whole spectral ranges; however, the RPD obtained during model evaluation yielded reliable accuracy for retrieval LWC (3.653). This is in accordance with several previous studies

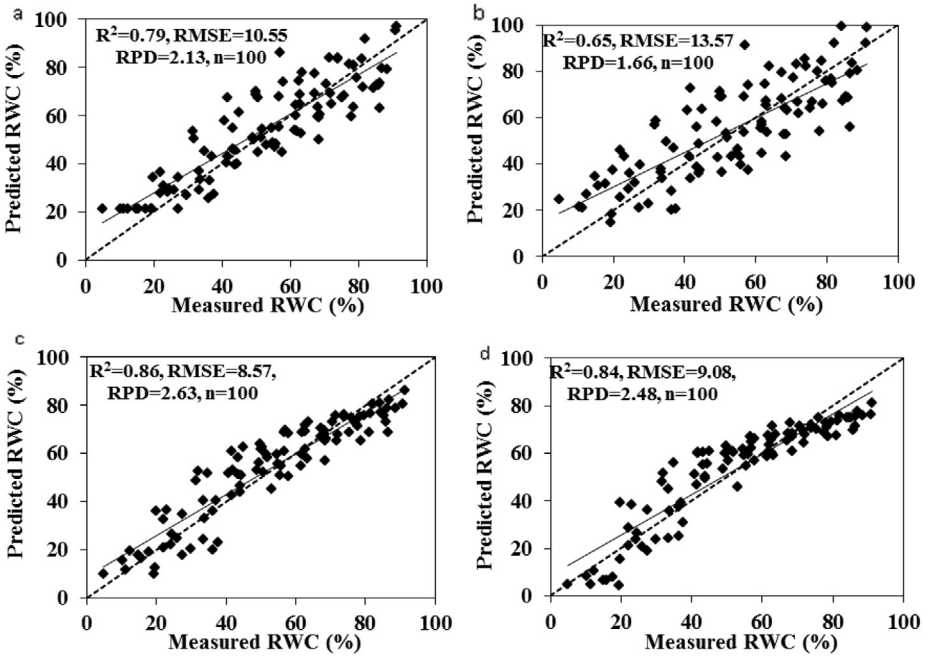


Fig. 7. The relationship between measured and predicted LWC (%) based on the model developed from band depth analysis; (a) 970 nm (b) 1200 nm (c) 1400 and (d) 1900 nm. The solid line shows the regression line, while the dotted line is the 1:1 line.

Table 5
Performance of different multivariate models for LWC prediction during calibration.

Model	Hyperparameters	R^2	MBE	RMSE	RPD
SVR	sigma = 0.0008, C = 1	0.946	0.131	5.121	4.280
GPR	sigma = 0.0008	0.924	0.200	6.214	3.527
MARS	nprune = 17 and degree = 1	0.957	0.000	4.519	4.851
PLSR	LV = 22	0.975	0.000	3.449	6.355
RF	mtry = 2151, splitrule = extratrees, min.node.size = 5	0.983	−0.035	2.947	7.438
SMLR	–	0.928	0.000	5.865	3.737

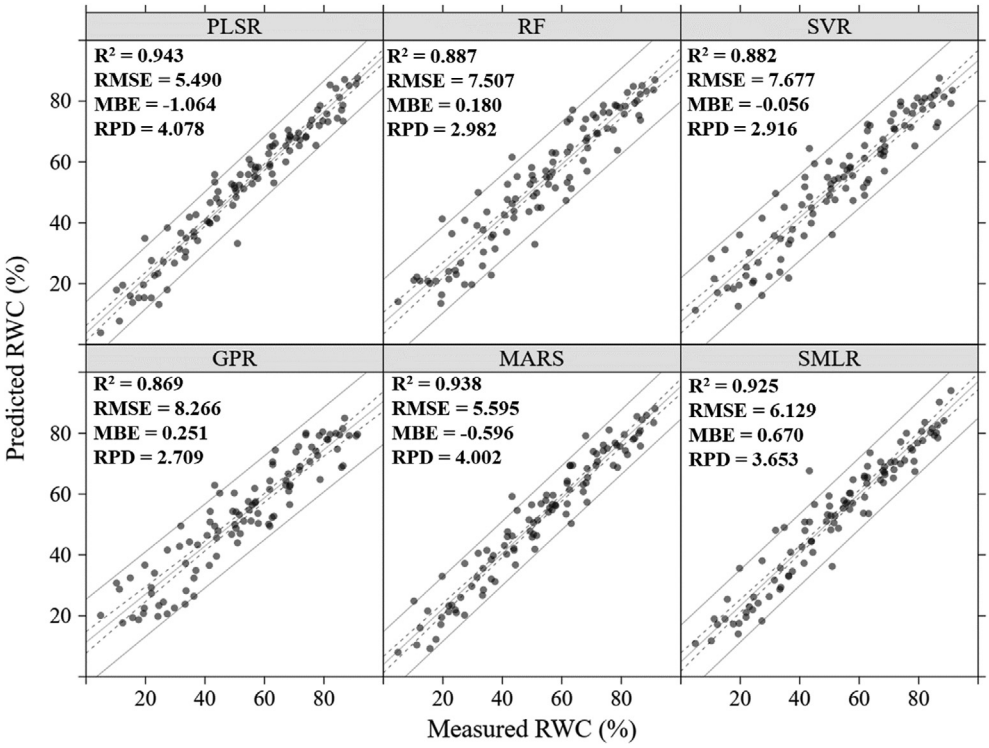


Fig. 8. Relationship between predicted and observed values of LWC (%) for rice leaves using six multivariate models during validation. The innermost solid line shows the regression line, while the dotted and outermost lines represent the confidence interval of regression and prediction at 95%, respectively.

reporting improved prediction ability of SMLR with prior selection of important wavelengths [17,53].

The wavelength 1359, 1907 and 2391 nm are associated with water absorption bands [54]. Wavelengths 350, 400, 441, 504 and 685 are relevant to plant pigment content that gradually changes during water-deficit stress [55,56]. The absorption at 1590, 1830 and 1869 nm is related to leaf dry matter constituents that become prominent as water content decreases (e.g., starch, protein, lignin and cellulose) [39,57]. Several authors in the past have also reported the potential of PLSR for accurate estimation of LWC [12,13,15,58,59]. Among the various multivariate models, PLSR yielded the most accurate estimates of LWC with respect to RPD of validation (RPD = 4.078). GPR showed the lowest accuracy among the multivariate models for estimation of LWC (RPD = 2.709) (Fig. 8). RWC was overestimated by all the models except RF (MBE = −0.035) during calibration while during validation overestimation was observed for SMLR, GPR and RF (MBE > 0). The order of performance of the multivariate models with respect to R^2 during validation was as follows: PLSR (0.943) > MARS (0.938) > SMLR (0.925) > RF (0.887) > SVR (0.882) > GPR (0.869) (Fig. 8). Our results differs with previous findings of Thissen et al. [60], Coen et al. [61], Shi et al. [62] and Tange et al. [63] reporting superior performance of SVR over PLSR. In the present study, performance of SMLR was found superior to SVR which is in disaggrement with Wang et al. [64] while estimating nitrogen concentration in rape.

4. Conclusions

In the current study, various methods were tested to determine the most accurate empirical model for LWC estimation in rice. The hyperspectral indices (RSI and NDSI) selected by precise analysis using the wavelengths in SWIR region produced the best accuracy and the least errors for LWC estimation in rice. Among the water absorption features, 1400 nm was found to be the best. Our results showed that PLSR is the most accurate method for estimation of LWC among the multivariate models. The study showed the potential of the proposed techniques for high-throughput non-invasive phenotyping of different crop genotypes for water-deficit stress tolerance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

First author acknowledges the fellowship received from CSIR: 18-12/2011(ii)EU-V, New Delhi, India to undertake this research work as part of Ph.D. Authors also acknowledge National Agriculture Science Fund (NASF: NASF(ICAR)- Phen 2015/2011-12), ICAR for conducting the experiment under its project on Phenomics of moisture deficit and low temperature stress tolerance in rice at IARI.

References

[1] T.R. Sinclair, M.M. Ludlow, Who taught plants thermodynamics? The unfulfilled potential of plant water potential, Funct. Plant Biol. 12 (1985) 213–217.
[2] R. Ranjan, R.N. Sahoo, U.K. Chopra, M. Pramanik, A.K. Singh, S. Pradhan, Assessment of water status in wheat (Triticum aestivum L.) using ground based hyperspectral reflectance, Proc. Natl. Acad. Sci. India Sect. B Biol. Sci. (2015) 1–12, https://doi.org/10.1007/s40011-015-0618-6.
[3] S. Ullah, M. Schlerf, A.K. Skidmore, C. Hecker, Identifying plant species using mid-wave infrared (2.5–6μm) and thermal infrared (8–14μm) emissivity spectra, Remote Sens. Environ. 118 (2012) 95–102, https://doi.org/10.1016/j.rse.2011.11.008.
[4] L.L. Pierce, S.W. Running, G.A. Riggs, Remote detection of canopy water stress in coniferous forests using the NS001 thematic mapper simulator and the thermal infrared multispectral scanner, PE&RS Photogramm. Eng. Remote Sens. 56 (1990) 579–586.
[5] B. Das, G.R. Mahajan, R. Singh, Hyperspectral remote sensing: use in detecting abiotic stresses in agriculture, in: Adv. Crop Environ. Interact., Springer Singapore, Singapore, 2018, pp. 317–335. https://doi.org/10.1007/978-981-13-1861-0_12.
[6] H.C. Stimson, D.D. Breshears, S.L. Ustin, S.C. Kefauver, Spectral sensing of foliar water conditions in two co-occurring conifer species: Pinus edulis and Juniperus monosperma, Remote Sens. Environ. 96 (2005) 108–118, https://doi.org/10.1016/j.rse.2004.12.007.
[7] D.A. Sims, J.A. Gamon, Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages, Remote Sens. Environ. 81 (2002) 337–354, https://doi.org/10.1016/S0034-4257(02)00010-X.
[8] T.J. Jackson, D. Chen, M. Cosh, F. Li, M. Anderson, C. Walthall, P. Doriaswamy, E. R. Hunt, Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans, Remote Sens. Environ. 92 (2004) 475–482. http://linkinghub.elsevier.com/retrieve/pii/S0034425703003353.
[9] M. Maki, M. Ishiahra, M. Tamura, Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data, Remote Sens. Environ. 90 (2004) 441–450, https://doi.org/10.1016/j.rse.2004.02.002.
[10] D. Chen, J. Huang, T.J. Jackson, Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands, Remote Sens. Environ. 98 (2005) 225–236, https://doi.org/10.1016/j.rse.2005.07.008.
[11] J.G.P.W. Clevers, L. Kooistra, M.E. Schaepman, Estimating canopy water content using hyperspectral remote sensing data, Int. J. Appl. Earth Obs. Geoinf. 12 (2010) 119–125, https://doi.org/10.1016/j.jag.2010.01.007.
[12] M. Mirzaie, R. Darvishzadeh, A. Shakiba, A.A. Matkan, C. Atzberger, A. Skidmore, Comparative analysis of different uni- and multi-variate methods for estimation of vegetation water content using hyper-spectral measurements, Int. J. Appl. Earth Obs. Geoinf. 26 (2014) 1–11, https://doi.org/10.1016/j.jag.2013.04.004.
[13] G. Krishna, R.N. Sahoo, P. Singh, V. Bajpai, H. Patra, S. Kumar, R. Dandapani, V.K. Gupta, C. Viswanathan, T. Ahmad, P.M. Sahoo, Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing, Agric. Water Manag. 213 (2019) 231–244, https://doi.org/10.1016/j.agwat.2018.08.029.
[14] Z. Cao, Q. Wang, C. Zheng, Best hyperspectral indices for tracing leaf water status as determined from leaf dehydration experiments, Ecol. Indic. 54 (2015) 96–107, https://doi.org/10.1016/j.ecolind.2015.02.027.
[15] B. Das, R.N. Sahoo, S. Pargal, G. Krishna, R. Verma, V. Chinnusamy, V.K. Sehgal, V.K. Gupta, Comparison of different uni- and multi-variate techniques for monitoring leaf water status as an indicator of water-deficit stress in wheat through spectroscopy, Biosyst. Eng. 160 (2017) 69–83, https://doi.org/10.1016/j.biosystemseng.2017.05.007.
[16] A. Lausch, M. Pause, A. Schmidt, C. Salbach, S. Gwillym-Margianto, I. Merbach, Temporal hyperspectral monitoring of chlorophyll, LAI, and water content of barley during a growing season, Can. J. Remote Sens. 39 (2013) 191–207, https://doi.org/10.5589/m13-028.
[17] S.E. El-Hendawy, M. Alotaibi, N. Al-Suhaibani, K. Al-Gaadi, W. Hassan, Y.H. Dewir, M.A.E.G. Emam, S. Elsayed, U. Schmidhalter, Comparative performance of spectral reflectance indices and multivariate modeling for assessing agronomic parameters in advanced spring wheat lines under two contrasting irrigation regimes, Front. Plant Sci. 10 (2019) 1–20, https://doi.org/10.3389/fpls.2019.01537.
[18] Z. Liu, W. Zhou, J. Shen, S. Li, G. Liang, X. Wang, J. Sun, C. Ai, Soil quality assessment of acid sulfate paddy soils with different productivities in Guangdong Province, China, J. Integr. Agric. 13 (2014) 177–186, https://doi.org/10.1016/S2095-3119(13)60594-8.
[19] P.J. Curran, Remote sensing of foliar chemistry, Remote Sens. Environ. 30 (1989) 271–278, https://doi.org/10.1016/0034-4257(89)90069-2.
[20] Y.L. Grossman, S.L. Ustin, S. Jacquemoud, E.W. Sanderson, G. Schmuck, J. Verdebout, Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data, Remote Sens. Environ. 56 (1996) 182–193, https://doi.org/10.1016/0034-4257(95)00235-9.
[21] H.T. Nguyen, B.-W. Lee, Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression, Eur. J. Agron. 24 (2006) 349–356, https://doi.org/10.1016/j.eja.2006.01.001.
[22] M. Schlerf, C. Atzberger, J. Hill, H. Buddenbaum, W. Werner, G. Schüler, Retrieval of chlorophyll and nitrogen in Norway spruce (Picea abies L. Karst.) using imaging spectroscopy, Int. J. Appl. Earth Obs. Geoinf. 12 (2010) 17–26.
[23] A.J. Steidle Neto, J.V. Toledo, S. Zolnier, D. de C. Lopes, C.V. Pires, T.G.F. da Silva, Prediction of mineral contents in sugarcane cultivated under saline conditions based on stalk scanning by Vis/NIR spectral reflectance, Biosyst. Eng. 156 (2017) 17–26, https://doi.org/10.1016/j.biosystemseng.2017.01.003.
[24] C.F. Jordan, Derivation of leaf-area index from quality of light on the forest floor, Ecology. 50 (1969) 663–666, https://doi.org/10.2307/1936256.
[25] J.W. Rouse, R.H. Haas, J.A. Schell, D.W. Deering, J.C. Harlan, Monitoring the Vernal Advancements and Retrogradation, Texas A M Univ., Texas, 1974.

[26] J. Peñuelas, J.A. Gamon, K.L. Griffin, C.B. Field, Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance, Remote Sens. Environ. 46 (1993) 110–118, https://doi.org/10.1016/0034-4257(93)90088-F.
[27] M.A. Babar, M.P. Reynolds, M. Van Ginkel, A.R. Klatt, W.R. Raun, M.L. Stone, Spectral reflectance indices as a potential indirect selection criteria for wheat yield under irrigation, Crop Sci. 46 (2006) 578–588, https://doi.org/10.2135/cropsci2005.0059.
[28] B. Prasad, B.F. Carver, M.L. Stone, M.A. Babar, W.R. Raun, A.R. Klatt, Genetic analysis of indirect selection for winter wheat grain yield using spectral reflectance indices, Crop Sci. 47 (2007) 1416–1425, https://doi.org/10.2135/cropsci2006.08.0546.
[29] C.D. Elvidge, R.J.P. Lyon, Estimation of the vegetation contribution to the 1·65/2·22 μm ratio in airborne thematic-mapper imagery of the Virginia Range, Nevada, Int. J. Remote Sens. 6 (1985) 75–88, https://doi.org/10.1080/01431168508948425.
[30] J.U.H. Eitel, P.E. Gessler, A.M.S. Smith, R. Robberecht, Suitability of existing and novel spectral indices to remotely detect water stress in Populus spp., For. Ecol. Manage. 229 (2006) 170–182, https://doi.org/10.1016/j.foreco.2006.03.027.
[31] R. Pu, S. Ge, N.M. Kelly, P. Gong, Spectral absorption features as indicators of water status in coast live oak (Quercus agrifolia) leaves, Int. J. Remote Sens. 24 (2003) 1799–1810, https://doi.org/10.1080/01431160210155965.
[32] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222.
[33] B.-H. Mevik, R. Wehrens, The pls package: principal component and partial least squares regression in R, J. Stat. Softw. 18 (2007) 1–24.
[34] A. Karatzoglou, A. Smola, K. Hornik, A. Zeileis, kernlab – An S4 Package for Kernel Methods in {R}, J. Stat. Softw. 11 (2004) 1–20. http://www.jstatsoft.org/v11/i09/.
[35] M.N. Wright, A. Ziegler, Ranger: a fast implementation of random forests for high dimensional data in C++ and R, J. Stat. Softw. 77 (2017). https://doi.org/10.18637/jss.v077.i01.
[36] R Core Team, R: A Language and Environment for Statistical Computing, 2016. https://www.r-project.org/.
[37] M. Kuhn, caret: Classification and Regression Training, 2020. https://cran.r-project.org/package=caret.
[38] S. Tanaka, K. Kawamura, M. Maki, Y. Muramoto, K. Yoshida, T. Akiyama, Spectral index for quantifying leaf area index of winter wheat by field hyperspectral measurements: a case study in Gifu prefecture, Central Japan, Remote Sens. 7 (2015) 5329–5346, https://doi.org/10.3390/rs70505329.
[39] T. Cheng, B. Rivard, A. Sánchez-Azofeifa, Spectroscopic determination of leaf water content using continuous wavelet analysis, Remote Sens. Environ. 115 (2011) 659–670, https://doi.org/10.1016/j.rse.2010.11.001.
[40] R. Colombo, M. Meroni, A. Marchesi, L. Busetto, M. Rossini, C. Giardino, C. Panigada, Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling, Remote Sens. Environ. 112 (2008) 1820–1834, https://doi.org/10.1016/j.rse.2007.09.005.
[41] A.B. González-Fernández, J.R. Rodríguez-Pérez, M. Marabel, F. Álvarez-Taboada, Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression, Sci. Hortic. (Amsterdam) 188 (2015) 15–22, https://doi.org/10.1016/j.scienta.2015.03.012.
[42] F.M. Danson, P. Bowyer, Estimating live fuel moisture content from remotely sensed reflectance, Remote Sens. Environ. 92 (2004) 309–321, https://doi.org/10.1016/j.rse.2004.03.017.
[43] F. Zhang, G. Zhou, Estimation of canopy water content by means of hyperspectral indices based on drought stress gradient experiments of maize in the North Plain China, Remote Sens. 7 (2015) 15203–15223, https://doi.org/10.3390/rs71115203.
[44] X. Yao, Y. Zhu, Y. Tian, W. Feng, W. Cao, Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat, Int. J. Appl. Earth Obs. Geoinf. 12 (2010) 89–100, https://doi.org/10.1016/j.jag.2009.11.008.
[45] S.E. El-hendawy, N.A. Al-suhaibani, S. Elsayed, W.M. Hassan, Potential of the existing and novel spectral reflectance indices for estimating the leaf water status and grain yield of spring wheat exposed to different irrigation rates, Agric. Water Manag. 217 (2019) 356–373, https://doi.org/10.1016/j.agwat.2019.03.006.
[46] Y. Zhu, W. Wang, X. Yao, Estimating leaf nitrogen concentration (LNC) of cereal crops with hyperspectral data, Hyperspectral Remote Sens. Veg. (2012) 187–206.
[47] X. Chu, Y. Guo, J. He, X. Yao, Y. Zhu, W. Cao, T. Cheng, Y. Tian, Comparison of different hyperspectral vegetation indices for estimating canopy leaf nitrogen accumulation in rice, Agron. J. 106 (2014) 1911–1920, https://doi.org/10.2134/agronj14.0084.
[48] B. Das, K.K. Manohara, G.R. Mahajan, R.N. Sahoo, Spectroscopy based novel spectral indices, PCA- and PLSR-coupled machine learning models for salinity stress phenotyping of rice, Spectrochim Acta Part A Mol. Biomol. Spectrosc. 229 (2020), https://doi.org/10.1016/j.saa.2019.117983 117983.
[49] M.M. Lungaria, B.I. Karande, K.I. Patel, V. Pandey, Determination of optimal narrow bands for vegetation indices to discriminate nitrogen status in wheat crop, J. Agrometeorol. 17 (2015).
[50] L. Zhang, Z. Zhou, G. Zhang, Y. Meng, B. Chen, Y. Wang, Monitoring the leaf water content and specific leaf weight of cotton (Gossypium hirsutum L.) in saline soil using leaf spectral reflectance, Eur. J. Agron. 41 (2012) 103–117, https://doi.org/10.1016/j.eja.2012.04.003.
[51] S. Ullah, A.K. Skidmore, T.A. Groen, M. Schlerf, Evaluation of three proposed indices for the retrieval of leaf water content from the mid-wave infrared (2–6μm) spectra, Agric. For. Meteorol. 171–172 (2013) 65–71, https://doi.org/10.1016/j.agrformet.2012.11.014.
[52] K. Zhao, D. Valle, S. Popescu, X. Zhang, B. Mallick, Hyperspectral remote sensing of plant biochemistry using Bayesian model averaging with variable and band selection, Remote Sens. Environ. 132 (2013) 102–119, https://doi.org/10.1016/j.rse.2012.12.026.
[53] D. Wu, J. Chen, B. Lu, L. Xiong, Y. He, Y. Zhang, Application of near infrared spectroscopy for the rapid determination of antioxidant activity of bamboo leaf extract, Food Chem. 135 (2012) 2147–2156, https://doi.org/10.1016/j.foodchem.2012.07.011.
[54] J. Zhang, J. Wu, L. Zhou, Deriving vegetation leaf water content from spectrophotometric data with orthogonal signal correction-partial least square regression, Int. J. Remote Sens. 32 (2011) 7557–7574, https://doi.org/10.1080/01431161.2010.524677.
[55] M.F. Dreccer, L.R. Barnes, R. Meder, Quantitative dynamics of stem water soluble carbohydrates in wheat can be monitored in the field using hyperspectral reflectance, F. Crop. Res. 159 (2014) 70–80, https://doi.org/10.1016/j.fcr.2014.01.001.
[56] B. Das, R.N. Sahoo, S. Pargal, G. Krishna, R. Verma, V. Chinnusamy, V.K. Sehgal, V.K. Gupta, S.K. Dash, P. Swain, Quantitative monitoring of sucrose, reducing sugar and total sugar dynamics for phenotyping of water-deficit stress tolerance in rice through spectroscopy and chemometrics, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 192 (2018) 41–51, https://doi.org/10.1016/j.saa.2017.10.076.
[57] K. Kawamura, N. Watanabe, S. Sakanoue, Y. Inoue, Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection, Grassl. Sci. 54 (2008) 131–145, https://doi.org/10.1111/j.1744-697x.2008.00116.x.
[58] S. Ullah, A.K. Skidmore, A. Ramoelo, T.A. Groen, M. Naeem, A. Ali, Retrieval of leaf water content spanning the visible to thermal infrared spectra, ISPRS J. Photogramm. Remote Sens. 93 (2014) 56–64, https://doi.org/10.1016/j.isprsjprs.2014.04.005.
[59] S. Elsayed, M. Elhoweity, H.H. Ibrahim, Y. Hassan, H.M. Migdadi, U. Schmidhalter, Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes, Agric. Water Manag. 189 (2017) 98–110, https://doi.org/10.1016/j.agwat.2017.05.001.
[60] U. Thissen, M. Pepers, B. Üstün, W.J. Melssen, L.M.C. Buydens, Comparing support vector machines to PLS for spectral regression applications, Chemom. Intell. Lab. Syst. 73 (2004) 169–179, https://doi.org/10.1016/j.chemolab.2004.01.002.
[61] T. Coen, W. Saeys, H. Ramon, J. De Baerdemaeker, Optimizing the tuning parameters of least squares support vector machines regression for NIR spectra, J. Chemom. 20 (2006) 184–192, https://doi.org/10.1002/cem.989.
[62] Q. Shi, G. Niu, Q. Lin, T. Xu, F. Li, Y. Duan, Quantitative analysis of sedimentary rocks using laser-induced breakdown spectroscopy: comparison of support vector regression and partial least squares regression chemometric methods, J. Anal. At. Spectrom. 30 (2015) 2384–2393, https://doi.org/10.1039/c5ja00255a.
[63] R. Tange, M. Rasmusson, E. Taira, R. Bro, Application of support vector regression for simultaneous modelling of near infrared spectra from multiple process steps, J. Near Infrared Spectrosc. 23 (2015) 75, https://doi.org/10.1255/jnirs.1149.
[64] Q. Wang, P. Li, X. Chen, Modeling salinity effects on soil reflectance under various moisture conditions and its inverse application: a laboratory experiment, Geoderma. 170 (2012) 103–111, https://doi.org/10.1016/j.geoderma.2011.10.015.