Agricultural Water Management 213 (2019) 231–244

Contents lists available at ScienceDirect

Agricultural Water Management

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

Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing

Gopal Krishna[a,b], Rabi N. Sahoo[c,*], Prafull Singh[b], Vaishangi Bajpai[c], Himesh Patra[c], Sudhir Kumar[d], Raju Dandapani[d], Vinod K. Gupta[c], C. Viswanathan[d], Tauqueer Ahmad[a], Prachi M. Sahoo[a]

[a] Division of Sample Surveys, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
[b] Amity Institute of Geoinformatics and Remote Sensing, Amity University, Noida, U.P., India
[c] Division of Agricultural Physics, ICAR-Indian Agricultural Research Institute, New Delhi, India
[d] Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, India

ARTICLE INFO

Keywords:
Hyperspectral reflectance
Water deficit stress
Relative water content (RWC)
Multivariate analysis
ANN

ABSTRACT

This study was conducted to understand the behaviour of ten rice genotypes for different water deficit stress levels. The spectroscopic hyperspectral reflectance data in the range of 350–2500 nm was recorded and relative water content (RWC) of plants was measured at different stress levels. The optimal wavebands were identified through spectral indices, multivariate techniques and neural network technique, and prediction models were developed. The new water sensitive spectral indices were developed and existing water band spectral indices were also evaluated with respect to RWC. These indices based models were efficient in predicting RWC with R^2 values ranging from 0.73 to 0.94. The contour plotting using the ratio spectral indices (RSI) and normalized difference spectral indices (NDSI) was done in all possible combinations within 350–2500 nm and their correlations with RWC were quantified to identify the best index. Spectral reflectance data was also used to develop partial least squares regression (PLSR) followed by multiple linear regression (MLR) and Artificial Neural Networks (ANN), support vector machine regression (SVR) and random forest (RF) models to calculate plant RWC. Among these multivariate models, PLSR-MLR was found to be the best model for prediction of RWC with R^2 as 0.98 and 0.97 for calibration and validation respectively and Root mean square error of prediction (RMSEP) as 5.06. The results indicate that PLSR is a robust technique for identification of water deficit stress in the crop. Although the PLSR is robust technique, if PLSR extracted optimum wavebands are fed into MLR, the results are found to be improved significantly. The ANN model was developed with all spectral reflectance bands. The 43 developed model didn't produce satisfactory results. Therefore, the model was developed 44 with PLSR selected optimum wavebands as independent x variables and PLSR-ANN model 45 was found better than the ANN model alone. The study successfully conducts a comparative 46 analysis among various modelling approaches to quantify water deficit stress. The methodology developed would help to identify water deficit stress more accurately by predicting RWC in the crops.

1. Introduction

Quantification of leaf biochemical and canopy biophysical variables is a key element for the successful deployment of remote sensing in crop condition monitoring. Accurate estimation of biophysical parameters from remote sensing can assist in the determination of vegetation physiological status (Carter, 1994). Estimation of one of the most important biochemical constituent, crop water content through remote sensing has important significances in agriculture and forestry (Zarco-Tejada et al., 2003; Gao and Goetz, 1995). Determination of plant water status plays a significant role in assessing drought stress, predicting susceptibility to wildfire (Ustin et al., 1998; Pyne et al., 1996) and monitoring the general physiological status of crops (Datt, 1999; Cheng et al., 2011). The determination of water content in plants is very crucial for drought assessment because the insufficient amount of water in crop hampers the production of the food grains negatively. The remote sensing is very widely used for accurate retrieval of leaf water content (Hunt and Rock, 1989; Peñuelas et al., 1997). The leaf water

* Corresponding author at: Division of Agricultural Physics, ICAR-Indian Agricultural Research Institute, New Delhi 110012, India.
E-mail address: rnsahoo.iari@gmail.com (R.N. Sahoo).

https://doi.org/10.1016/j.agwat.2018.08.029
Received 9 October 2017; Received in revised form 19 August 2018; Accepted 21 August 2018
Available online 30 October 2018
0378-3774/

content is commonly expressed as equivalent water thickness (EWT), gravimetric water content (GWC) and relative water content (RWC) (Datt, 1999; Cheng et al., 2010). The EWT is mass per unit leaf area (g/cm2) whereas the GWC expresses leaf water content as the gravimetric proportions of water relative to other plant material. The RWC can be expressed as the ratio of the difference between fresh weight and dry weight to that of the difference of turgid weight and dry weight. The RWC serves as a key leaf parameter to determine leaf water content (Ullah et al., 2014; Das et al., 2017). Although the remote sensing technique is widely used for timely detection of variations in the spectral response of plants to changing levels of plant water status over large areas (Peñuelas et al., 1997; Ustin et al., 1998; Pu et al., 2003; Stimson et al., 2005; Eitel et al., 2006), The multispectral satellite remote sensors exhibit serious limitations to accurately detect changes in plant water status due to coarse spectral resolution and larger revisit time. Therefore, the need of high spectral and spatial resolution remote sensing instruments and sensors was experienced. This contributed for the advent of highly precise spectroradiometers for detection of spectral changes. The field spectroradiometers and hyperspectral sensors has the capability to detect the electromagnetic spectrum in very narrow contiguous bands which allows the development of spectral indices using minor fluctuations of wavelengths due to change in water status (Horler et al., 1983; Gao, 1996; Peñuelas et al., 1993; Eitel et al., 2006).

Several previous studies have demonstrated the utilization of spectral reflectance in 350–2500 nm range to assess water content in plants through spectral indices, regression analysis and radiative transfer modeling (Féret et al., 2011; Zarco-Tejada et al., 2003). In the earlier studies, the primary and secondary effects of water content on the spectral response of leaf were evaluated by Carter (1994) and it was concluded that 1450 nm, 1940 nm, and 2500 nm are the most optimal wavebands showing sensitivity to water content. The wavelength 400 nm and 700 nm (red edge position) were also found to be sensitive to plant water content (Filella and Peñuelas, 1994). Roberts et al., 1997 reported the NDVI as a water content sensitive index. Several studies demonstrated a good relationship between spectral indices developed through NIR region (700–1300 nm) and plant water content (Peñuelas et al., 1997; Serrano et al., 2000; Ceccato et al., 2002; Asner et al., 2003; Imanishi et al., 2004; Stimson et al., 2005). Few studies have also indicated that NIR region is the less sensitive region of the spectrum compared to SWIR (1300–2500 nm) to establish a relationship between indices and water content (Danson et al.,1992; Ceccato et al., 2002; Eitel et al., 2006). Most of the indices are two band simple ratio indices, utilizing two spectral wavebands. Mostly one of the wavelengths is found within strong absorption region of water and another is found outside the absorption region of water (Sims and Gamon, 2003; Eitel et al., 2006).

To extract larger information on crop water status, investigation of entire spectrum is essential. Use of multivariate regression techniques, machine learning methods, and artificial neural network approach can utilize the entire spectrum for detection of crop water stress. However the high dimensionality and contiguity of hyperspectral data is a problem (Vaiphasa et al., 2005) when utilizing entire spectrum (350–2500 nm range). The reason is that the regression techniques like multiple linear regression (MLR) may suffer from multi-collinearity and are often prone to over-fitting as numbers of observations could be equal or lesser than the predictors (Curran, 1989). Contrary to MLR, the partial least square regression technique (PLSR) is a robust technique for development of prediction models. The PLSR is a combination of principal component analysis (PCA) & MLR techniques. The concept behind PLS is to find a few eigenvectors of spectral matrices that will produce score values that both summarize the variance of spectral reflectance well and are highly correlated with response variables (Li et al., 2007). Several researches indicate that PLSR can effectively decrease complexity and the multi-collinearity of spectral responses by performing simple projection operations in a vector space (Araújo et al., 2001; Galvão et al., 2001, 2008; Mahmood et al., 2012) consequently reducing the over-fitting. The PLSR combines the most useful information from hundreds of contiguous spectral bands into several principal components to develop a calibration model. Several studies have highlighted that PLSR is a robust prediction model development technique and researchers have used PLSR successfully to establish a relationship between spectral reflectance and leaf biochemical and biophysical properties under varying canopy structures (Asner and Martin, 2008). The PLS regression has been successfully used with spectral data to predict chlorophyll content (Zhao et al., 2016; Ji et al., 2012), estimation of carotenoid content (Zhao et al., 2015), estimation of relative water content (Ullah et al., 2014), estimation of protein, lignin and cellulose (Thulin et al., 2014),estimation of leaf nitrogen content (Ecarnot et al., 2013), estimation of leaf area index and chlorophyll content (Darvishzadeh et al., 2008), estimation of soil organic carbon (Peng et al., 2014), prediction of soil properties (Mahmood et al.,2012) and retrieval of leaf fuel moisture content (Li et al., 2007). Though PLSR is the most robust technique for prediction model development, few researchers have reported that there is a possibility of over-fitting that would lead to inaccurate results when testing the developed model on a very different dataset to the calibration one (Féret et al., 2011). Therefore, optimum wavebands extracted from PLSR were fed into MLR and ANN techniques separately to check whether the outcome of the combined models is better or not. Neural networks technique has also been evaluated for development of water content prediction models. Dawson et al. (1998) developed the ANN model for prediction of leaf water content and reported a satisfactory coefficient of determination as 0.86 with low RMSE (1.3%).There are several researches which evaluate multivariate techniques for estimation of crop biochemical and biophysical parameters using spectral reflectance data but very few studies have demonstrated the comparison among efficiency and accuracy of various multivariate models to estimate water content of crop from hyperspectral observations. This study bridges this gap by comparing models developed from PLSR, MLR, RF and SVR multivariate techniques and ANN too. The present investigation was done with the following objectives (i) Evaluation of existing water bands indices as well as development of new efficacious water band indices (ii) Identification of the most optimum wavebands sensitive to predict RWC in crops (iii) Development of various RWC prediction models using multivariate techniques and neural networks, and their comparison with each other. (iv) Evaluation of PLSR-MLR model to test its efficacy over model developed through only PLS regression.

2. Materials and methods

2.1. Study area

The study of the research study was ICAR-Indian Agricultural Research Institute (IARI), New Delhi research farms (28°38′28.59″N, 77° 9′28.09″E). This study area was selected to conduct the research study because it has all the ideal conditions required for the experiment and the adjoining labs have plentiful facilities. The study area has an average elevation of 230 m above sea level. The soil is mostly well-drained sandy loam. The minimum temperature is recorded between 0 °C to 7 °C during the winter season and the maximum temperature ranged between 41 °C to 46 °C. The average annual rainfall is about 750 mm. The relative humidity (RH) is found to be the highest during the monsoon season. In the summer months, the RH is observed between 40 to 45%. Ten rice genotypes were grown in the farms of the division of plant pathology, ICAR-IARI, New Delhi. Five genotypes were Drought Sensitive - MTU 1010, Patchaiperumal, Pusa Basmati-1, Pusa Sugandha-5, IR 64 and five were Drought Tolerant - Sahbhagidhan, CR-143, Nerica L44, Moroberekan, APO.

Table 1
Spectral indices related to water status and their respective definition.

Spectral Indices related to water status	Definition (Wavelengths in nm)	References
Water Band Index (WBI)	R_{900}/R_{970}	Peñuelas et al. (1997)
Moisture Stress Index (MSI)	R_{1600}/R_{820}	Hunt and Rock (1989)
Hyperspectral Normalized Difference Vegetation Index (hNDVI)	$(R_{900}-R_{685})/(R_{900}+R_{685})$	Rouse et al. (1974)
Normalized Difference Water Index (NDWI)	$(R_{820}-R_{1240})/(R_{820}+R_{1240})$	Gao (1996)
Normalized Difference Infrared Index (NDII)	$(R_{820}-R_{1649})/(R_{820}+R_{1649})$	Hardisky et al. (1983)
Maximum Difference Water Index (MDWI)	$(R_{max1500-1750})$-$(R_{min1500-1750})/(R_{max1500-1750})+(R_{min1500-1750})$	Eitel et al. (2006)
Ratio Index	(R_{1650}/R_{2220})	Elvidge and Lyon (1985)
Simple Ratio Water Index (SRWI)	R_{800}/R_{1200}	Zarco-Tejada and Ustin (2001)

2.2. Data used

Four leaves sample per genotype for above mentioned 10 genotypes were collected from the field experiment site. The plots were in randomized block design and were well irrigated. Leaves were quickly placed in plastic bags in an airtight container and immediately transferred to the laboratory for spectroscopic measurements at predetermined time intervals. In the laboratory, the spectroscopic data of above mentioned 10 genotypes were collected using an ASD Field Spec 3 spectroradiometer. This instrument collects data into 350 to 2500 nm wavelength at resampled wavelength interval of 1 nm. Approximately 3 g of fresh leaves for each genotype were put into capped glass tubes filled with distilled water and kept at room temperature to attain full turgidity.

2.3. Collection of spectroscopic data from leaves

The spectral measurements of fresh leaves were recorded in the lab immediately. After first spectral reading leaves were allowed to dry at room temperature and spectral measurements were again recorded after 2, 3, 4, 5, 6, 8 and 10 h from the time of first spectral observation collection. For all 10 genotypes, 8 spectral observations were recorded. For each genotype four spectral observations were recorded for above mentioned hours, therefore, total 320 spectral observations (10 genotypes x4 replication in observations x8 different hours) were recorded. The spectral observations were recorded in a dark room having ± 25 °C by using an ASD contact probe (Analytical Spectral Devices, Boulder, CO). This contact probe touches the surface of the leaf and has its own constant light source inside it for illumination; a black surface has been given which comes underside of the leaf while collecting spectra to minimise the electromagnetic radiation transmitted through the leaf. This contact probe is calibrated using a spectralon. This contact probe is an accessory of ASD Field Spec 3 spectroradiometer which records spectral reflectance in the 350 to 2500 nm range at sampling intervals of 1.4 nm in the 350–1050 nm range and of 2 nm in the 1000–2500 nm and It provides data after resampling at the 1 nm interval. The spectral observations were taken from leaf sample consisted of an overlapping pile of 3–4 leaves to eliminate the background effect.

2.4. Relative water content (RWC) computation

The water content in the leaves was analyzed using RWC computation. For RWC computation, the Fresh Weight (FW), Turgid Weight (TW) and Dry Weight (DW) were determined for all genotypes. Turgid weight was determined after placing the leaves in deionised water for 2 h. To obtain dry weight, leaves were oven dried at 70 °C temperature for 3 days until constant weight was obtained. The RWC was calculated using following equation –

$$RWC(\%) = \frac{(FW - DW)}{(TW - DW)} \times 100$$

2.5. Spectral indices computation

The plant water status spectral indices utilize simple ratios between the reflectance of a wavelength located within an range of the electromagnetic spectrum strongly absorption by water, described as water absorption bands, and another wavelength located outside the water absorption band typically used as a control (Sims and Gamon, 2003; Eitel et al., 2006). In this study indices related to plant water status only were evaluated. Spectral indices evaluated are given in Table 1.

2.6. Correlation analysis between narrow band indices and RWC through contour plotting

Two narrow band indices were computed and the correlation between computed indices with RWC was determined. The coefficient of determination (R^2) was plotted with wavelengths by a predefined matrix scheme. This plotting (the contour plotting - lambda versus lambda plotting approach) exhibits a specific pattern where highest R^2 can be seen as hot spots. Many studies have reported this plotting as the best approach for identification of wavelength having maximum R^2 (Sahoo et al., 2015). The highest R^2 value was extracted from the hot spot area. The optimal indices were selected by choosing the wavelength combination that portrayed the highest R^2 value in the contour plot. For the implementation of contour plotting, a program was written in Matlab.

2.7. Multivariate analysis

To perform multivariate analysis, the data was split into the training set and the test set for calibration and validation respectively. The training set of data was 2/3 sample and test data was 1/3 sample of the whole dataset. The overall performance and robustness of the models were appraised by the coefficient of determination (R^2), root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP), and ratio of prediction deviation (RPD) and upper & lower confidence intervals of regression at 95% confidence level. The RPD is computed as the ratio between standard deviation and RMSE. Excellent calibrations were those with $R^2 > 0.95$, RPD > 4 (Nduwamungu et al., 2009b). The ratio of prediction deviation (RPD) is considered as a parameter of strength for the prediction model. A model having RPD value 0–2.3 is considered as very poor, 2.4–3.0 is considered as poor, 3.1–4.9 is considered as fair and prediction are considered as reliable, 5.0–6.4 is considered as good, 6.5–8.0 is considered as very good with very reliable predictions and model with RPD above 8.1 is considered as excellent for prediction (Williams and Sobering, 1993). The detailed schematic diagram of methodology is given in Fig. 1.

2.7.1. Multivariate techniques evaluated

Support vector regression (SVR), Artificial neural networks (ANN), random forest (RF) and the partial least square regression (PLSR), PLSR followed by multiple linear regression (MLR) and PLSR followed by ANN were evaluated to determine the best suitable multivariate model

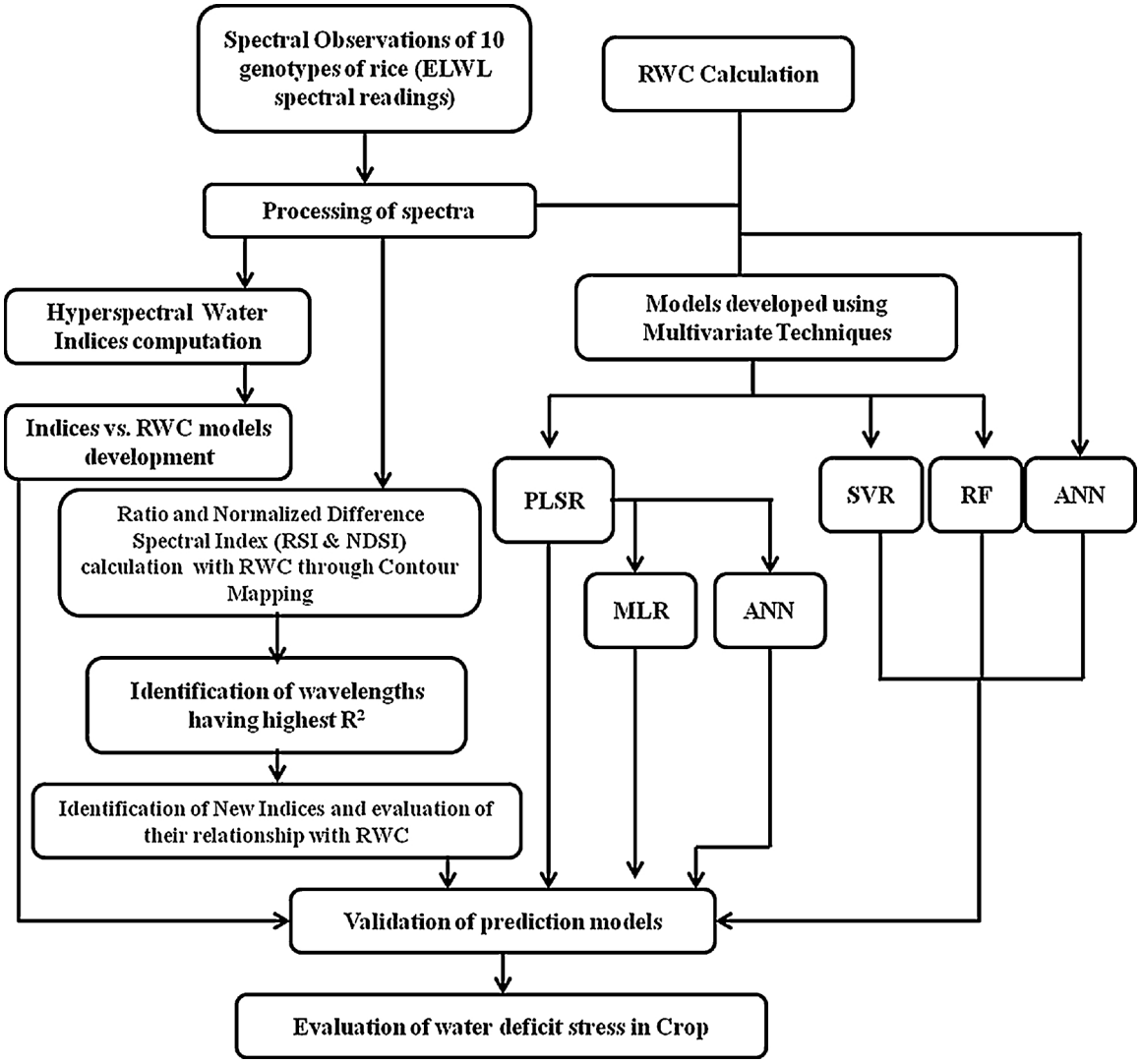

Fig. 1. The schematic diagram of the methodology.

for regression between spectral reflectance and RWC.

2.7.2. The partial least square regression (PLSR)

The PLSR multivariate analysis was performed on spectral reflectance data and RWC. Other multivariate regression models based on hyperspectral data shows a high degree of collinearity especially when the numbers of predictors are equal or higher in number than sample observations and the input data lead to a high R^2 (Curran, 1989). The PLSR has proved as the robust technique which can handle high dimensionality of hyperspectral data. Many researchers have successfully used PLSR for estimation of various leaf biochemicals (Asner and Martin, 2008; Huang et al., 2004; Ramoelo et al., 2011) and leaf water status (Ullah et al., 2014). PLSR is very popular and has been extensively used in Remote Sensing (Asner and Martin, 2008; Darvishzadeh et al., 2008; Li et al., 2007; Ramoelo et al., 2011). The reason behind its extensive use is the fact that PLSR has the capability to process multi-collinear hyperspectral data by inputting all spectral bands simultaneously and select uncorrelated variables from a matrix of explanatory variables (Geladi and Kowalski, 1986). The PLSR analysis was implemented through a program written using 'pls' library (Mevik and Wehrens, 2007) in R studio. The PLSR analysis selected 30 optimum wavebands which were highly sensitive to water deficit stress. The selected wavebands were then fed into multiple linear regression (MLR) model.

2.7.3. The multiple linear regression (MLR)

Multiple linear regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data and every value of the independent variable x is associated with a value of the dependent variable y (Lattin et al., 2003; Krishna et al., 2014). The Multiple Linear Regression (MLR) model was used to account for the relationship between Rice crops' reflectance and RWC data. The band used as input were retrieved from PLSR selected optimum wavebands. This approach of using PLSR selected optimum wavebands was applied because previous studies show that MLR has several shortcomings such as leading to negative and extremely large estimates (Zhu et al., 2017).

2.7.4. The support vector regression (SVR)

Support Vector Regression system is based on Support Vector Machines (Cortes and Vapnik, 1995) that is derived from statistical learning theory. SVM separates the classes with a decision surface that maximizes the margin between the classes. The surface is called the optimal hyperplane, and the data points closest to the hyperplane are called support vectors. Among the separating hyperplanes, the one for which the distance to the closest point is maximal is called optimal separating hyperplane (Chapelle et al., 1999). The support vectors are the critical elements of the training set. The key idea of using SVM is map points with a mapping function to a space of sufficiently high

dimension so that they will be separable by a hyperplane. SVR is the implementation of the SVM method for regression and function approximation (Smola and Schölkopf, 2004; Das et al., 2017). In this study, the SVM regression was performed using package 'e1071' (Meyer et al., 2015) in R language.

2.7.5. The artificial neural networks (ANN)

The neural networks are based on backpropagation algorithm and structure is inspired by the brain. The backpropagation is a fast algorithm and at the heart of backpropagation is an expression for the partial derivative $\partial C/\partial w$ of the cost function C with respect to any weight w (or bias b) in the network (Nielsen, 2015). For predicting nonlinear system problems, a nonlinear neural network with additional intermediate or hidden processing layers is very much useful to handle the nonlinearity and complexity problems (Subasi and Erçelebi, 2005). A model with very few nodes would be incapable of differentiating between complex patterns while too many nodes may lead to over parameterization. The determination of hidden intermediate layers is by trial and error. Too many hidden layers make the process very much time-consuming. The neural network regression was performed in R language with 'neuralnet' package (Fritsch and Guenther, 2016), using the 'neuralnet' function.

2.7.6. The random forest (RF)

The random forest regression technique is an addition to the bagging (Breiman, 1994) of classification trees. The classification using bagging is different from the boosting because in bagging, successive trees do not depend on earlier trees and each is independently constructed using a bootstrap sample of the data set (Liaw and Wiener, 2002). The final result is predicted using a simple majority vote. In the process of random forest, each node is split using the best among a subset of predictors randomly chosen at that node. This process of somewhat immoderate splitting of node provides very good results compared to other regression and classification techniques like support vector regression, discriminant analysis, and neural networks, and is robust against overfitting (Breiman, 2001). This regression technique was implemented using 'randomForest' (Breiman, 2001) package of R language.

3. Results and discussions

3.1. Changes in spectral reflectance pattern due to water deficit stress

Normally the plants of a particular crop show a similar pattern of reflectance spectra. But water deficit stress conditions bring noticeable changes in reflectance spectra. The study shows the reflectance patterns of plants with different water deficit stress conditions i.e. decline in relative water content. The water content varies from 96.5% to 0.7%. The reflectance of the fresh plant was less whereas the reflectance of the dry plant was high. The reflectance in SWIR region increases as the RWC decreases from the highest to lowest. The reason behind the increase in reflectance is weakening of the water absorption features at 1400 nm and 1900 nm A similar pattern of increasing reflectance with a decrease in water content was observed at 350 to 700 nm wavelength region. The spectrum in the blue and red region (chlorophyll *a* & b absorption ranges) was showing a trend of higher reflectance with decreasing water content due to loss of chlorophyll. A shift of 1400–1925 nm wavelength range towards shorter wavelengths was observed with the drying of leaves and increase in spectral reflectance is also visible. With the decrease in relative water content, the absorption features in 1400 to 1500 nm and 1850 to 1900 nm were seen as becoming shallow. The reason behind the decrease in absorption is weakening of water absorption features due to the decrease in water content. The scattering in spongy mesophyll at 810 to 1350 nm was also reflected a similar trend of increasing reflectance with the decrease in water content. In addition, absorption at the middle infrared (1100–2500 nm) is also a zone of strong absorption, primarily by water in a fresh leaf and secondarily by dry matter (e.g., protein, lignin and cellulose) when the leaf wilts (Jacquemoud and Ustin, 2001), become more visible with decrease in RWC.

3.2. Change in relative water content (RWC)

The genotypes showed a significant variation over time in RWC. The calibration data shows variation of RWC between 95.4% to 1.0% whereas validation subset data shows 97.0%–2.0%. The standard deviation for calibration subset was 27.5% whereas 29.8%. The MTU 1010 (Fig. 2) genotype showed the highest variation and

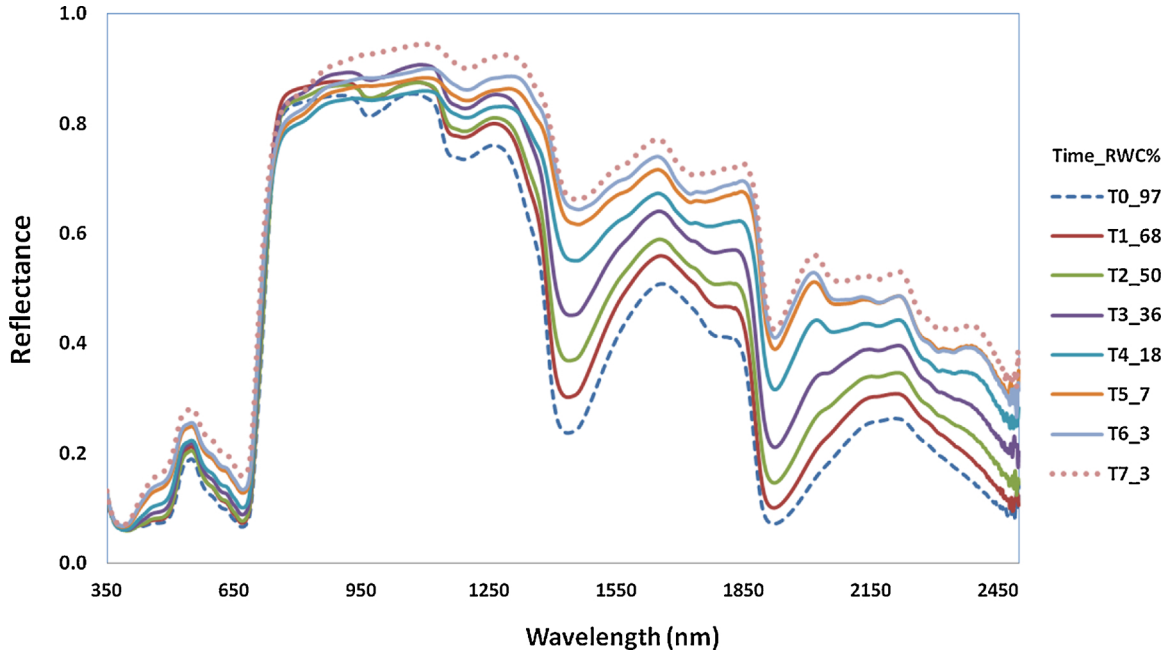

Fig. 2. Representative mean spectral reflectance observations of the genotypes with decreasing RWC (%) in leaves of rice, showing percentage of RWC and corresponding spectra at different time intervals.

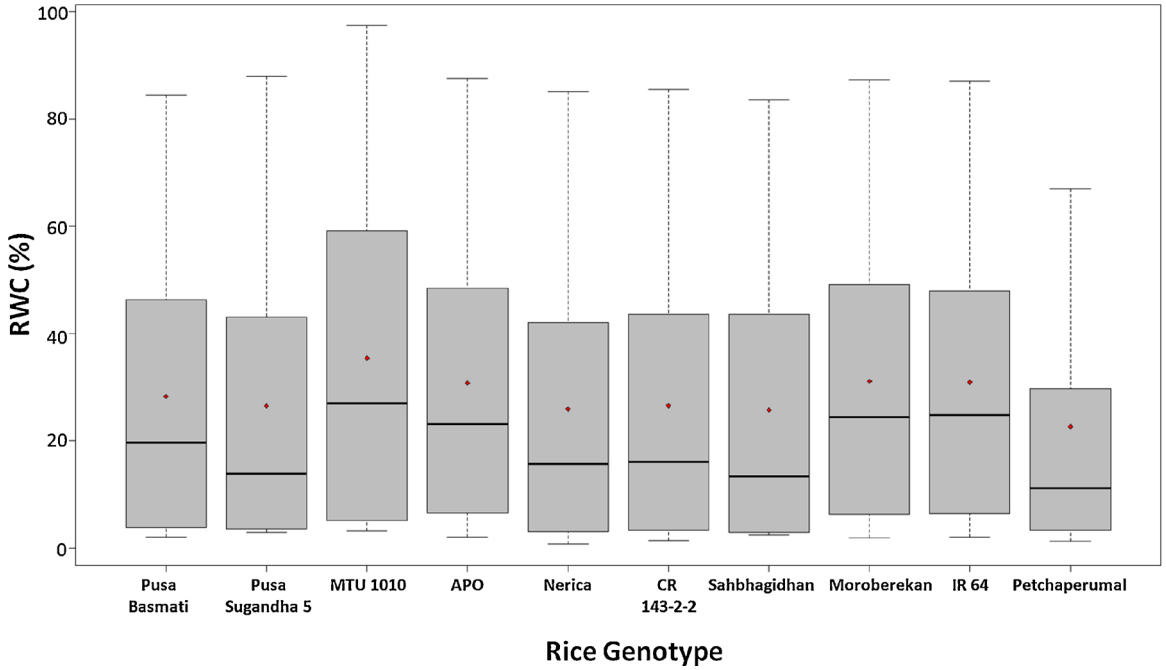

Fig. 3. Boxplots showing the means and spreads of relative water content (RWC) in different Rice genotypes.

Table 2
Relationships between Relative Water Content and Spectral Indices.

Index	Model equation	R^2 Cal.	R^2 Val.	RMSEP	RPD
WBI (Water Band Index)	$11{,}709.54x^2 - 24{,}484.34x + 12{,}785.86$	0.88	0.90	6.59	4.35
MSI (Moisture Stress Index)	$384.09x^2 - 815.07x + 420.02$	0.89	0.92	5.51	5.21
hNDVI (Hyperspectral NDVI)	$2675.03x^2 - 3526.23x + 1163.95$	0.85	0.89	7.67	3.74
NDWI (Normalized Difference Water Index) (R_{820} & R_{1240} nm)	$5703.28x^2 + 857.87x + 17.08$	0.86	0.89	7.06	4.06
NDWI (Normalized Difference Water Index) (R_{820} & R_{1640})	$598.54x^2 + 240.64x - 13.93$	0.89	0.89	9.93	2.89
NDII (Normalized Difference Infra Red Index) (R_{820} & R_{1649} nm)	$618.39x^2 + 243.26x - 13.91$	0.89	0.92	5.48	5.23
NDII (Normalized Difference Infra Red Index) (R_{819} & R_{1600} nm)	$484.87x^2 + 220.29x - 16.31$	0.89	0.92	5.44	5.27
MDWI (Max Difference Water Index)	$-149.08x^2 + 473.70x - 21.27$	0.92	0.92	5.23	5.49
Ratio Index (R_{1650}/R_{2220} nm)	$-61.97x^2 + 376.51x - 411.61$	0.88	0.89	6.84	4.19
SRWI (Simple Ratio Water Index) (R_{820}/R_{1200} nm)	$876.84x^2 - 1388.30x + 525.01$	0.87	0.89	7.07	4.06
Normalized Multi Band Drought Index	$61.44x^2 - 316.61x + 380.50$	0.86	0.85	9.36	3.06
WBI/NDVI	$573.96x^2 - 1746.51x + 1329.82$	0.87	0.91	6.62	4.33
Simple ratio (R_{895}/R_{675})	$0.29x^2 + 4.32x - 31.30$	0.73	0.80	7.90	3.63
Proposed Ratio Index (R_{1233}/R_{1305} nm)	$5213.38x^2 - 8594.07x + 3408.03$	**0.94**	**0.93**	**4.27**	**6.99**
Proposed Normalized Difference Ration index ($R_{1233} - R_{1305}$)/($R_{1233} + R_{1305}$ nm)	$24455x^2 + 3671.2x + 27.356$	**0.94**	**0.93**	**4.28**	**6.98**

Petchaperumal showed the least variation in RWC. The boxplots show the distribution of measured RWC where median values are depicted by horizontal dark lines (Fig. 3). The length of boxes indicates spread of water content and corresponds to interquartile range (Q3(75%) – Q1(25%)). The lines attached to the dotted line and situated above & below boxes represent the upper and lower limit of RWC for a particular genotype (Fig. 2). The points indicate the mean values.

The relationship between conventional water band indices with RWC was evaluated (Table 2). The MDWI exhibits the strongest correlation with R^2 as 0.92 for both calibration and validation sets (Fig. 4). The Moisture Stress Index (MSI) and Normalized Difference Infra Red Index (NDII) also showed a strong correlation. The MDWI is computed using the maximum reflectance value from $max_{1500–1750\ nm}$ and minimum reflectance value from $min_{1500–1750\ nm}$ located at the atmospheric window between 1500 and 1750 nm. Both MSI and NDII performed the correlation with R^2 as 0.89 and 0.92 for correlation and validation respectively. The lowest correlation was observed for simple ratio index with R^2 as 0.73 (calibration) and 0.80 (validation). The MDWI performed well because it allows the best combination of numerator and denominator from 1500 and 1750 nm wavelength range. This dynamism of choosing better absorption feature, under varying plant water-deficit stress conditions provides better results ((Eitel et al., 2006; Peñuelas et al., 1997).

3.3. Contour mapping approach for exploring new useful water band spectral indices

The contour mapping approach has the advantage of providing an efficient selection of the optimal combination of wavebands for development of the effective spectral indices. The contour maps of R^2 values from linear regression between RWC and all possible combinations of RSI (Ratio Spectral Index -ratio approach) and NDSI (Normalized

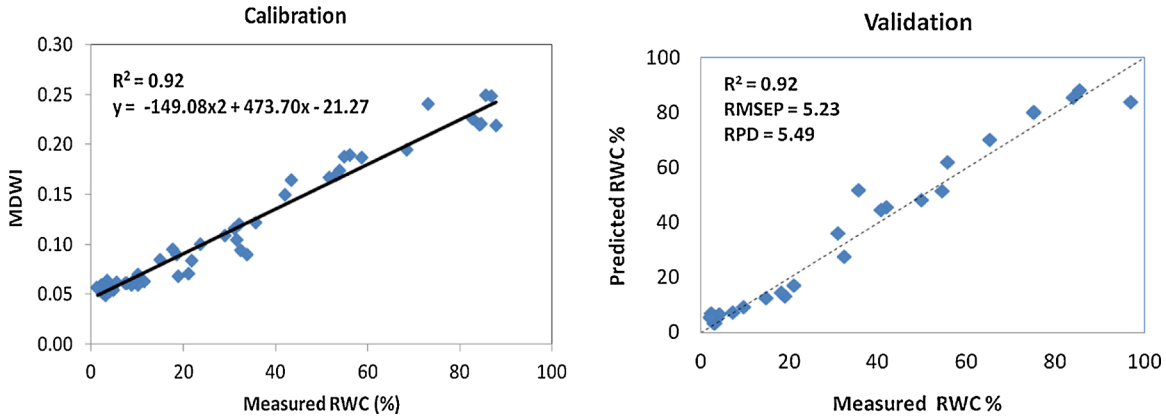

Fig. 4. The Calibration model developed through the relationship between MDWI and Measured RWC (%) and its validation. (Calibration −N = 55 & validation −N = 25). The solid black line is regression line and dotted line is 1:1 line.

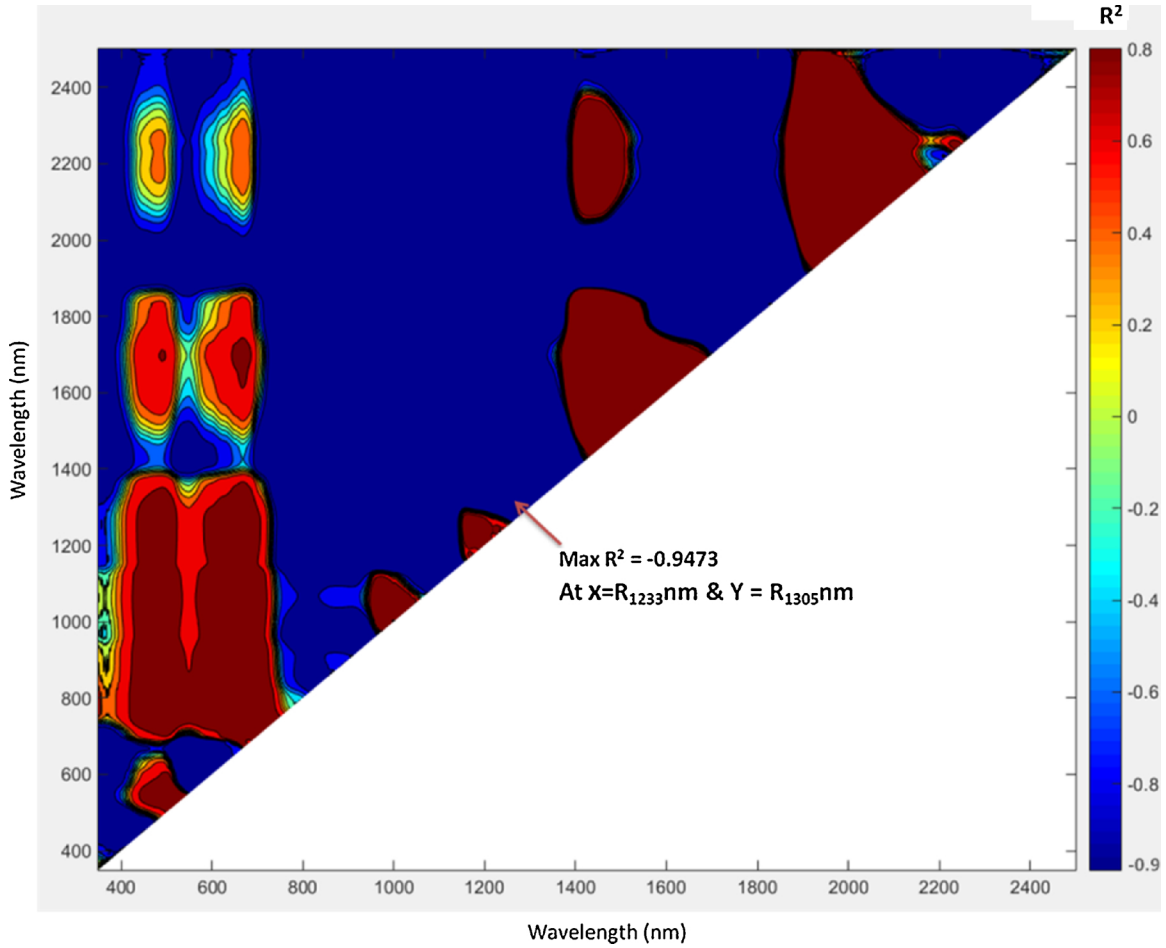

Fig. 5. The Contour plot (lambda by lambda) showing different combinations of RSI (Ratio Spectral Index -ratio approach) and NDSI (Normalized Difference Spectral Index -normalized difference approach). The arrow indicates the wavelength where max R^2 was observed.

Difference Spectral Index -normalized difference approach) reveal hotspot positions that have high correlation values (Fig. 5). The contour mapping was performed at 1 nm interval and all of the hotspots were analyzed. Consequently, one highest R^2 value each for RSI and NDSI was extracted from the hotspots which were found at 1233 and 1305 nm combination. Therefore, on the basis of highest R^2, the best combinations selected were Ratio Index (R_{1233}, R_{1305}) and Normalized Difference Ratio Index (R_{1233}, R_{1305}) for RWC. The linear, polynomial, exponential and logarithmic regression functions were evaluated for establishing regression equation between RWC- Ratio Index and RWC-Normalized Difference Ratio Index (Table 3). The 2nd order polynomial equation was found to be the best in predicting RWC with both Ratio Index and Normalized Difference Ratio Index (R^2_{Cal} = 0.94, RMSEP = 4.27; R^2_{Cal} = 0.94, RMSEP = 4.28, respectively) (Figs. 6 and 7).

3.4. Validation of the RSI and NDSI models

The validation results of regression models from Ratio Index and Normalized Difference Ratio Index to predicted RWC exhibit the R^2 as 0.93 for both indices. The RMSEP was 4.27 and 4.28 for Ratio Index and

Table 3
The regression equations and related statistics of model for proposed indices (Ratio Index and Normalized Difference Ratio Index).

Spectral Index	Regression Equation	R^2	RMSEP	RPD
Proposed Ratio Index (R_{1233}, R_{1305})	$y = 1899.5x - 1871.2$	0.941		
	$y = 1911\ln(x) + 28.475$	0.941		
	$y = 2E\text{-}33e^{77.826x}$	0.756		
	$y = 15.62x^{78.5}$	0.760		
	$y = 5213.4x^2 - 8594.1x + 3408$	0.942	4.27	6.99
Proposed Normalized Difference Index (R_{1233}, R_{1305})	$y = 3822.18x + 28.48$	0.941		
	$y = 24{,}454.84x^2 + 3671.22x + 27.36$	0.942		
	$y = 15.62e^{157.01}$	0.760	4.28	6.98

Note: Power and Logarithmic regression equation were not computed for Proposed Normalized Difference Index because there were negative values in it.

Normalized Difference Ratio Index respectively. The newly proposed indices yield better results compared to previous conventional indices. The RMSEP was found low compared to RMSEP of other indices. Thus the newly proposed indices can be reliably used for accurate estimation of changes in RWC caused by water deficit stress in plants. The RPD values of both the proposed indices were found significantly reliable compared to existing indices.

3.5. Multivariate models

3.5.1. The PLSR

The PLSR model provides reasonable explanations for independent variables using fewer latent variables compared to principal component regression. PLS regression was computed considering independent X variables as spectral reflectance observations and relative water content as dependent y variable. Increasing the number of latent variables (LV) in the PLS regression model tended to decrease the RMSE. However, the inclusion of too many latent variables led to over-fitting (Ecarnot et al., 2013). Therefore, the model with 3 components was considered as optimum. The number of components was determined using percent variation explained by components and cross-validated RMSECV. The component one explained 94.4% variation; second component explained 2.7% whereas component 3 explained 0.2% variation. The optimum wavebands were selected from the peaks and troughs of loading weight values (latent variables) in the spectral region 350–2500 nm. These optimum wavebands were: 357, 415, 511, 549, 691, 713, 766, 770, 815, 960, 1053, 1057, 1154, 1155, 1244, 1255, 1402, 1404, 1690, 1705, 1870, 1885, 1930, 1996, 2042, 2219, 2222, 2261, 2267 and 2411 nm (Fig. 8).

The model was both cross validated and validated with separate set of test data. The cross validation was performed with 'LOO' (leave one out) method. In the calibration model, the R^2 was 0.96 with RMSE as 5.63 and RPD as 4.89 and in the validation, the R^2 was 0.96 with RMSE as 5.37 and RPD as 5.55 (Fig. 9).

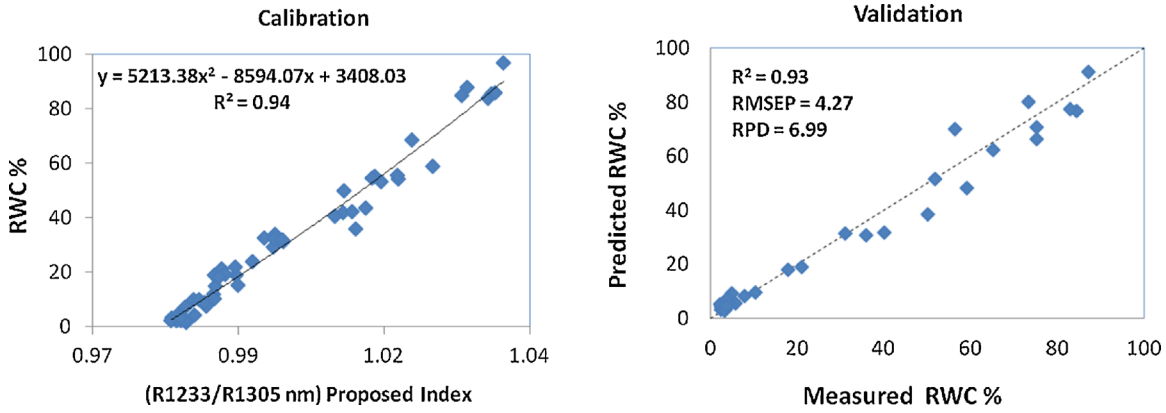

Fig. 6. The proposed Ratio Index (R_{1233}–R_{1305}) for prediction of RWC. (Calibration −N = 55 & validation −N = 25). The solid black line is regression line and dotted line is 1:1 line.

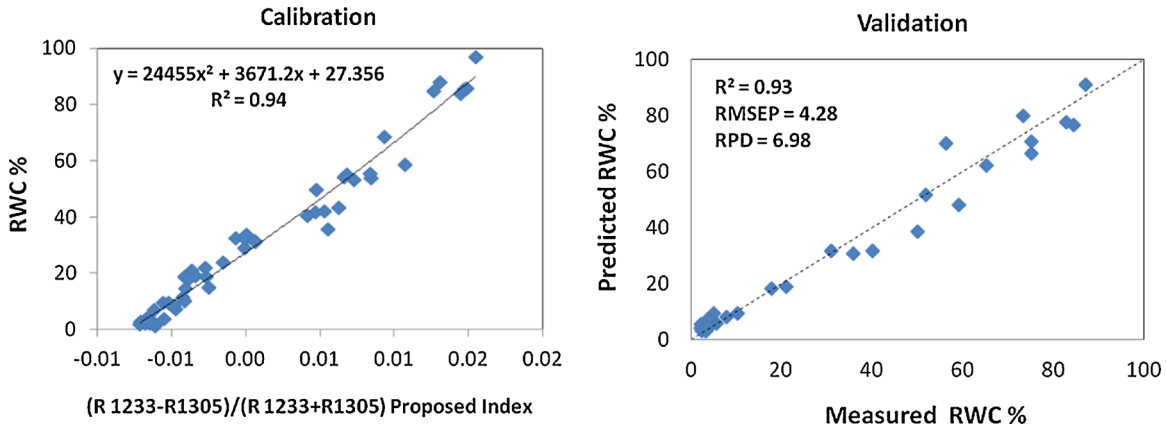

Fig. 7. The proposed Normalized Difference Ratio Index (R_{1233}–R_{1305})/(R_{1233} + R_{1305}) for prediction of RWC. (Calibration −N = 55 & validation −N = 25). The solid black line is regression line and dotted line is 1:1 line.

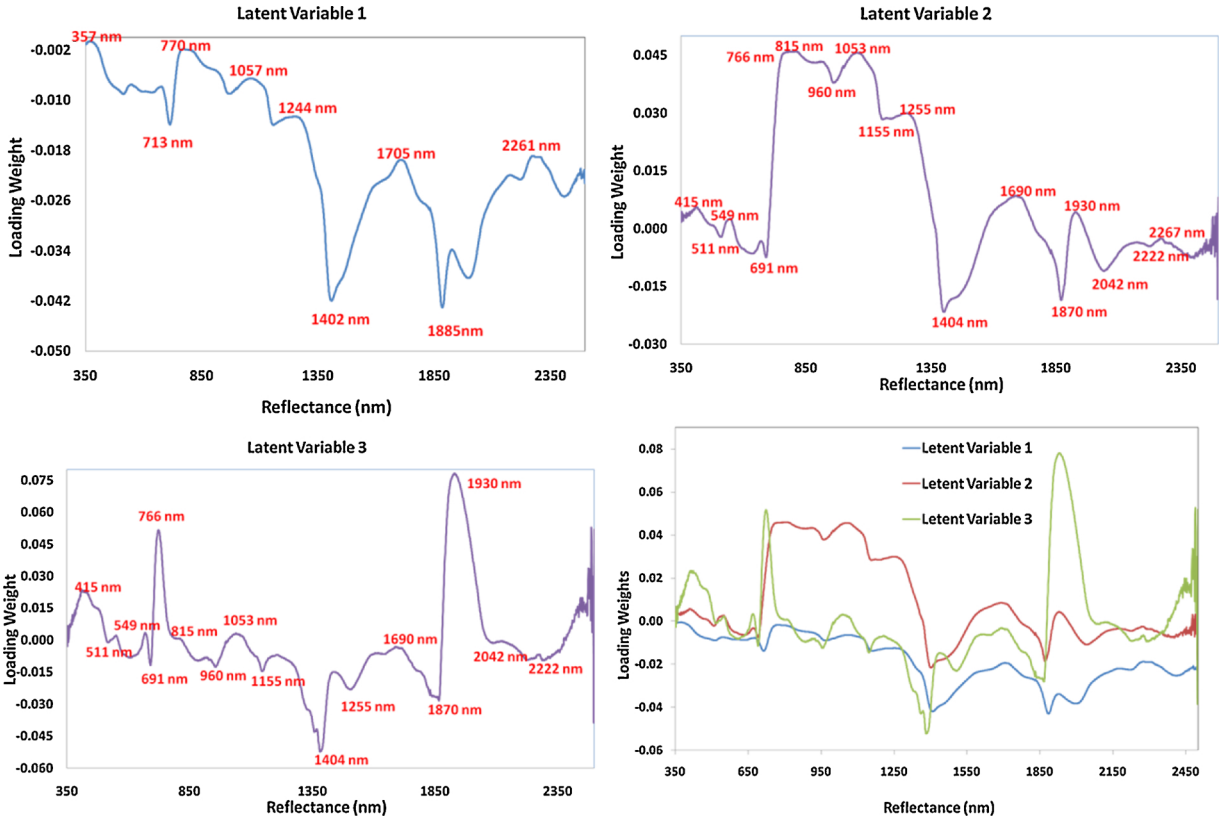

Fig. 8. Latent variables extracted from PLS regression model. The peaks and troughs of spectra are the optimum wavebands for RWC prediction. The lower right plot shows all three latent variables overlaid.

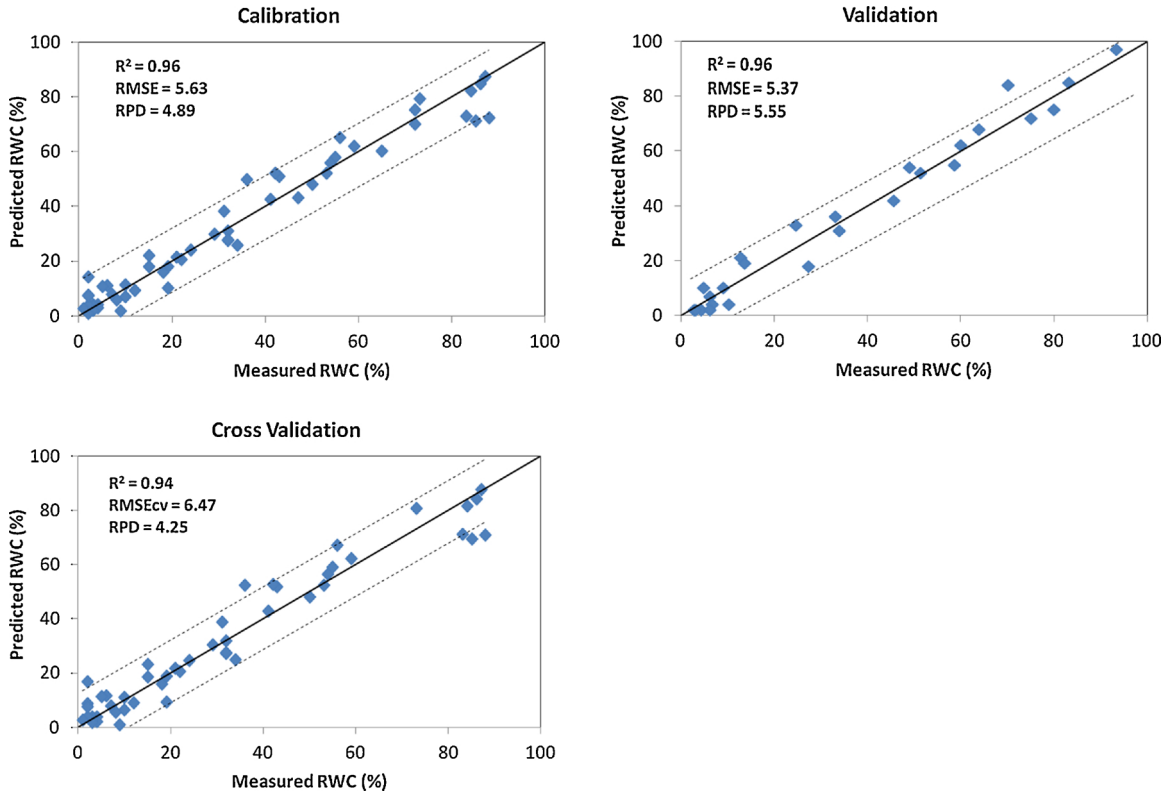

Fig. 9. The PLSR model calibration (N = 55), validation (N = 25) and cross validation ((N = 55) plots with respect to RWC of rice crop. The dotted lines are upper and lower confidence interval lines at 5% confidence interval; the black line is 1:1 line.

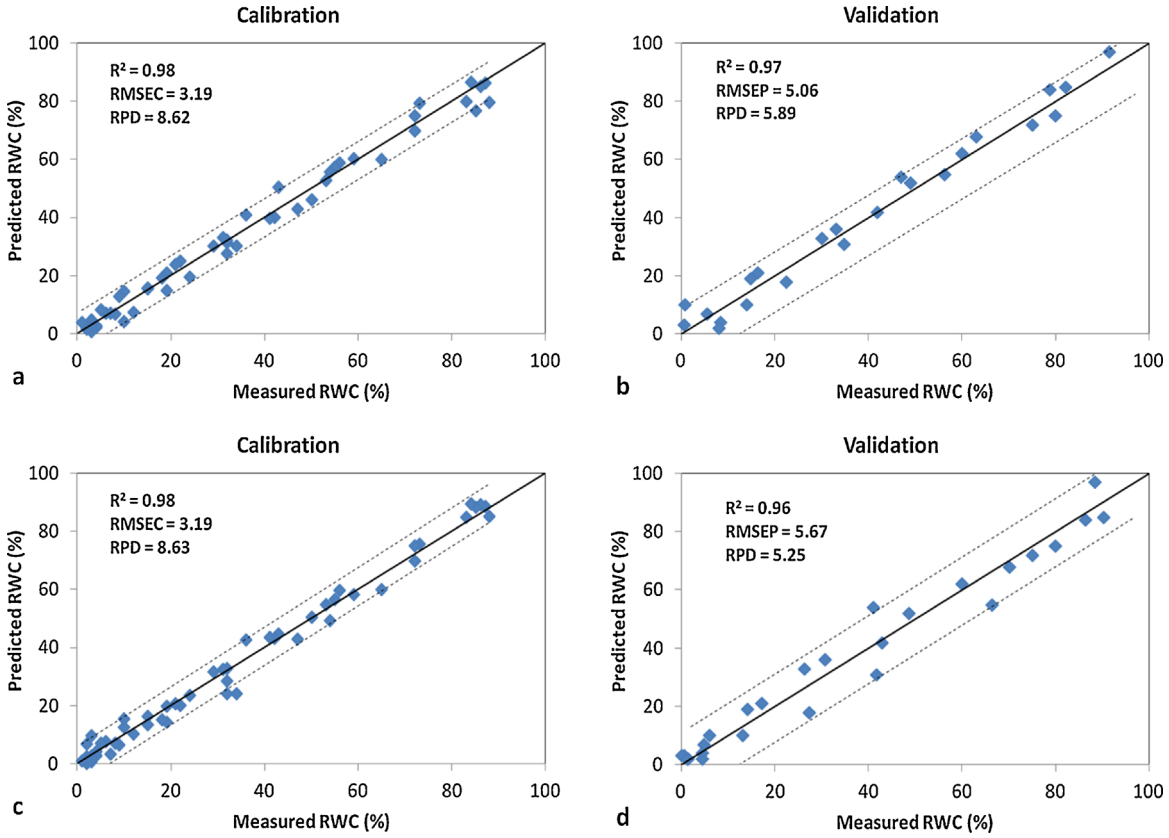

Fig. 10. The PLSR-MLR model calibration & validation plots (a, b), and the PLSR-ANN model calibration & validation plots (c, d) (Calibration −N = 55 & validation −N = 25). The dotted lines are upper and lower confidence interval lines at 5% confidence interval; the black line is 1:1 line.

3.5.2. The MLR

The PLSR is an extension of MLR technique with improved and robust regression approach but in PLSR equation, every coefficient has a RMSE associated with it which makes it more susceptible to the deviation. Therefore, the optimum wavebands extracted from PLSR were used as independent x variables in a stepwise MLR model. The MLR model equation is given below-

$$y = 80.47 - 351*R_{357} - 241*R_{511} + 1395*R_{770} - 1791*R_{815} + 2225*R_{1154} - 1447*R_{1255} - 14{,}612*R_{1402} + 13{,}988*R_{1404} + 3069*R_{1690} - 2475*R_{1705} - 367*R_{1930} + 472*R_{1996} - 12{,}005*R_{2261} + 11{,}584*R_{2267}$$

This model was evaluated as the best one among all the techniques evaluated in this study. The developed MLR model demonstrated the highest R^2 values, lower RMSEP values and the highest RPD values for both calibration and prediction data sets (R^2 = 0.98, RMSEC = 3.19 and RPD = 8.62 for calibration and in validation, R^2 = 0.97, RMSEP = 5.06 and RPD = 5.89 (Fig. 10a, b). This combination of two multivariate techniques proved the best one because the MLR model used the PLSR selected optimum reflectance wavebands rather than the whole 2151 spectral reflectance wavebands. Use of optimum wavebands as independent variables removed data redundancy and minimized the susceptibility to the deviation, therefore, provided the best results.

The wavelengths used by MLR model equation are the most prominent wavelengths for prediction of relative water content in plants. The shorter wavelengths of visible region 356 and 511 nm are related to chlorophyll and other pigment contents of the plant which exhibits changes during the water deficit stress condition. The 510 to 530 nm shows absorption for zeaxanthin pigment which modulates chlorophyll for photosynthesis (Dall'Osto et al., 2012) and shows changes during water deficit stress condition. The 770 nm is related to red edge position. The red edge position starts from 710 nm in healthy plants and gets shifted towards 800 nm if water stress is prevalent in the plant. The 1154 and 1255 nm are related with cell structure of leaf and canopy which show higher reflectance if the plant is facing water deficit stress. The wavelengths 1402, 1404, 1930 and 1996 nm are related to water absorption in the spectrum and are therefore directly related to water deficit stress. The selected wavebands in the SWIR region (near 1400 nm and 1600 nm) are related to the absorption features associated with moisture, cellulose, and starch in plant leaves (Curran, 1989; Thenkabail et al., 2004; Ullah et al., 2014). The 2261 and 2267 nm are sensitive to leaf biochemicals, protein, cellulose, lignin, etc which tend to be in higher proportion in the condition of water deficit stress (Thulin et al., 2014; Kokaly, 1999; Elvidge, 1990).

3.5.3. The ANN

The ANN model was developed with all spectral reflectance bands. The developed model didn't produce satisfactory results; therefore, the model was developed with PLSR selected optimum wavebands as independent x variables.

The ANN model with all spectral reflectance bands was developed with 1 hidden layer. Use of two or more hidden layers produced a large mean square error (MSE) compared to one hidden layer. In calibration, R^2 was 0.97, RMSEC was 5.62 and RPD was also 5.62 whereas in validation R^2 was 0.85, RMSEP was 13.06 and RPD was 2.28 (Fig. 12e, f). The ANN model predicted the RWC values poorly compared to other techniques because the model has a RMSE value associated with every coefficient which makes it more susceptible to the deviation. Another reason is that the accuracy of ANN technique is affected by the outliers in the data set compared to least-squares-based regression methods (Clrovic, 1997).

The ANN model developed with PLSR selected optimum wavebands as x variables produced better results compared to above ANN model.

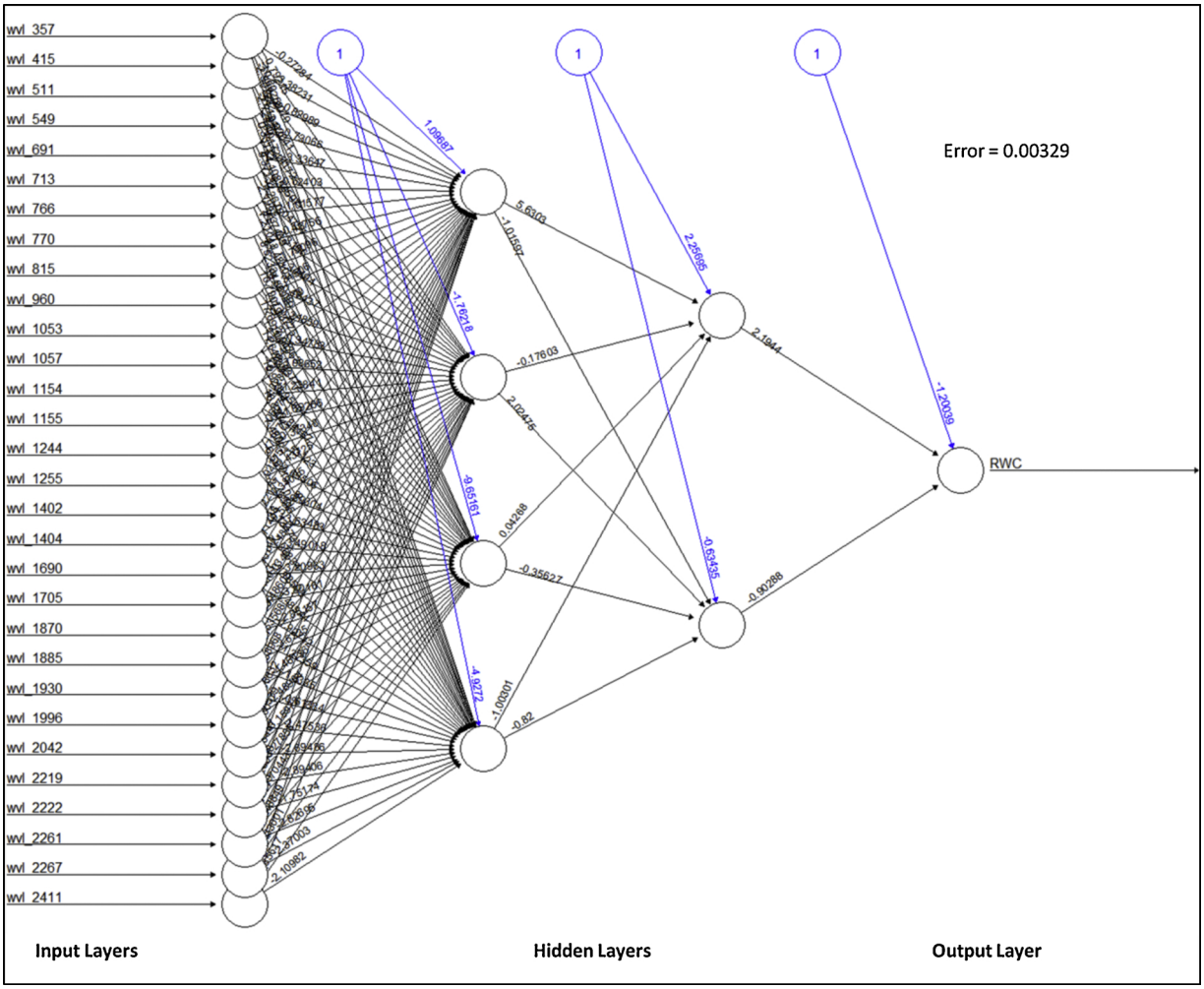

Fig. 11. The Architecture of prediction model developed through ANN technique.

The architecture of this ANN model is given in Fig. 11. For this model, two hidden layers were considered as sufficient on the basis of MSE. This ANN model displayed the R^2 as 0.98, RMSEC as 3.19 and RPD as 8.61 for calibration data set whereas in validation, the R^2 was 0.96, RMSEP was 5.67 and RPD was also 5.25 (Fig. 10c, d). This model was found to be the second best multivariate model as evident from model accuracy statistics. Use of PLSR selected optimum wavebands as x variables enabled to use two hidden layers; consequently, the prediction ability of the model was improved. Apart from the use of two hidden layers, in this ANN model, the data redundancy and outliers were already removed by PLSR technique. Therefore, the model was able to perform better.

3.5.4. The SVR

The SVR technique was also evaluated to develop a RWC prediction model. The model performed well with all independent variables. The model displayed a strong combination of higher R^2 and low RMSEC with excellent level of RPD (R^2 = 0.98, RMSEC = 3.53, RPD = 7.79 for calibration, in validation R^2 = 0.97, RMSEP = 5.75 and RPD = 5.18) (Fig. 12a, b).

3.5.5. The RF

The ensemble regression technique random forest provided intermediate results with R^2 = 0.97, RMSEC = 5.05 and RPD = 5.67. For validation data set the R^2 was 0.96, RMSEP = 5.26 and RPD was 5.45 (Fig. 12c, d).

The PLSR followed by MLR was proved as the best technique for RWC prediction model development, out of all multivariate techniques evaluated through this study. The model equation developed through PLSR-MLR techniques is also useful in monitoring water content in plants. All the wavelengths included in the model equation are highly relevant with respect to water stress prediction. The second best model developed was the combination of PLSR and ANN. The support vector regression was also proved to be a useful technique with satisfactory results. The SVR determines maximum-margin hyperplane; therefore, it reduces the prediction error. The ANN is vulnerable to outliers, therefore when applied on the whole dataset; its prediction was very poor. The random forest is an ensemble tree classifier and has the goodness of decision tree system. The RF proved as an intermediate classifier compared to others. It was proved slightly better over PLSR in this study. In the PLSR equation, every coefficient has a RMSE error associated with it which makes it more susceptible to deviation (Krishna et al., 2014), therefore PLSR model developed through all of the x variables produced intermediate results compared to PLSR-MLR combination. The order of performance of the multivariate models with respect to R^2 and RMSEP is as follows: PLSR-MLR > PLSR-ANN > SVR > RF > PLSR > ANN (Fig. 13). This order of performance is also supported by the value of RPD for all models.

This study evaluated multivariate techniques and indices based approach including contour plotting. The comparison of results clearly reflects that use of multivariate techniques enhances the prediction capability of models significantly. The multivariate techniques have many positive approaches compared to conventional indices based approach like self- identification and removal of outliers, use of

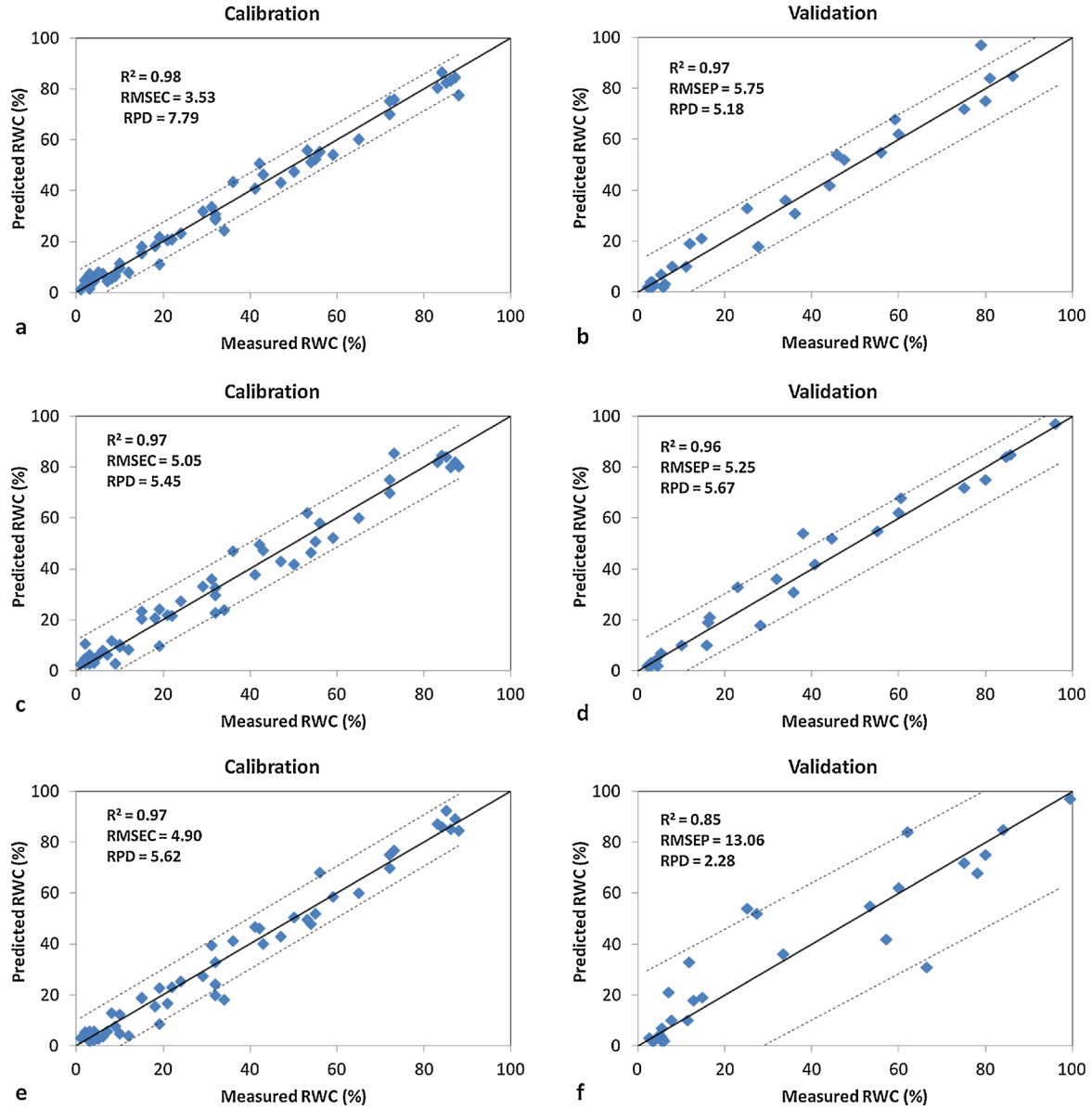

Fig. 12. The SVR model calibration & validation plots (a, b), the RF model calibration & validation plots (c, d) and the ANN model calibration & validation plots (e, f), (Calibration −N = 55 & validation −N = 25). The dotted lines are upper and lower confidence interval lines at 5% confidence interval; the black line is 1:1 line.

principal components, ability to deal with multi-collinearity, use of decision tree approach etc. Multivariate techniques all utilize all the water absorption related bands which increase model's accuracy considerably by unveiling improved sensitivity to changes in the RWC whereas index-based approaches use only two or three prominent water absorption bands. Several researches in the past have used multivariate techniques for determination of various plant biochemical contents i.e. chlorophyll (Schlerf et al., 2010; Daughtry et al., 2000; Atzberger et al., 2010; Zhao et al., 2016), carotenoids (Zhao et al., 2016) and Nitrogen (Ecarnot et al., 2013; Schlerf et al., 2010; Atzberger et al., 2010; Ryu et al., 2011) as well as RWC (Ullah et al., 2014) and leaf EWT (Colombo et al., 2008). Ullah et al. (2014) utilized various parts of the spectrum using PLSR to predict RWC. The leaf nitrogen content and leaf mass per unit area of wheat were also assessed using PLS regression technique (Ecarnot et al., 2013). Zhang and Zhou (2015), estimated the canopy water content using indices based approach and successfully developed a model for estimation of canopy water content and leaf equivalent water thickness for maize crop. Colombo et al. (2008) evaluated the performance of different hyperspectral indices for estimation of leaf equivalent water thickness and leaf water content using the PLSR model. The PLSR also displayed considerably good results in this study. This study has successfully applied the MLR and ANN models on PLSR selected optimum wavebands which increased the accuracy of model significantly. Use of PLSR selected optimum wavebands as input removed the multi-collinearity problem in MLR, and provided outliers free x variables to ANN; consequently, improving the efficiency of the PLSR model.

4. Conclusion

This study successfully evaluates the indices based, multivariate techniques based and neural networks based approaches to predict relative water content (RWC) under water deficit stress condition of rice genotypes with significant accuracy. Existing water band indices were evaluated and new water band indices sensitive to water stress were proposed. The MDWI was found to be the best index among all conventional existing indices. The newly proposed indices outperformed all other indices. The multivariate model developed through PLSR and

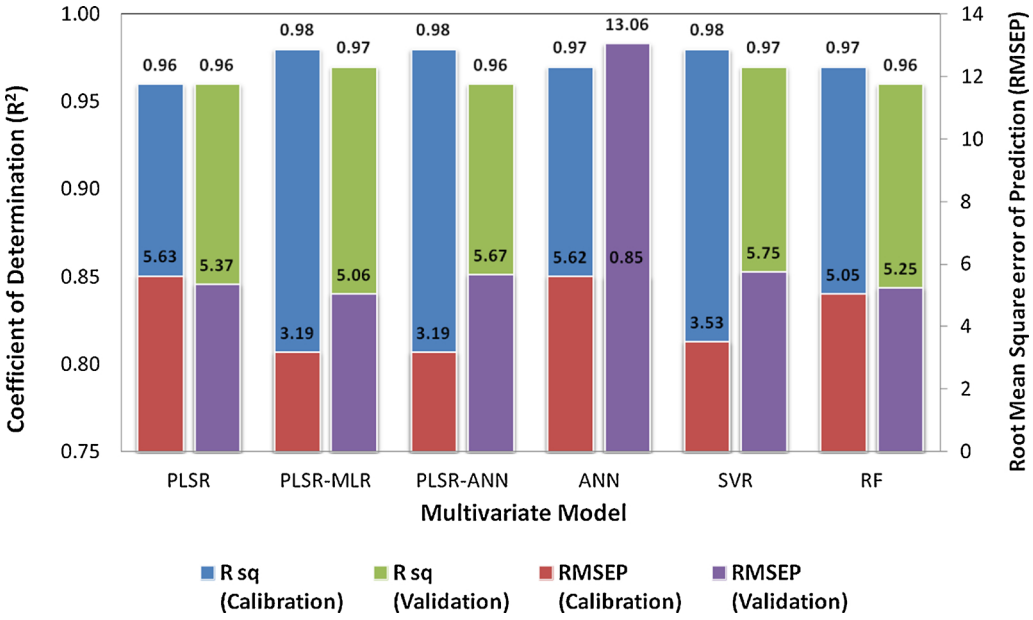

Fig. 13. Performance assessment of multivariate models as well as neural network models using R^2 and RMSEP of calibration and validation.

MLR techniques (PLSR-MLR model) proved to be the best ((yielded high R^2 and low RMSEP) followed by the model developed through PLSR and ANN techniques (PLSR-ANN model) for estimation of RWC in rice crop. Thus from this study it may be concluded that timely detection of water deficit stress is quite important for precision agriculture. The model and indices developed through this study can be effectively used to detect water deficit induced stress. Measurement of the relative water content (RWC) at different stages of crop using hyperspectral reflectance may provide timely detection of the water deficit stress. Use of hyperspectral images may provide large area coverage and will be more suitable compared to ground based spectroradiometer data. Unavailability of hyperspectral images over the study area poses a limitation to assess water deficit stress at regional scale. Use of airborne/satellite-borne hyperspectral data in future studies may considerably enhance the utility of such research studies. The methodology developed for prediction of RWC would help to identify water deficit stress more accurately using crop reflectance spectra and may prove useful in developing drought resistant varieties.

Funding

ICAR- National Agricultural Science Fund; Grant Code : NASF/Phen-6005 /2016-17.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The first author acknowledges the ICAR-Indian Agricultural Research Institute, New Delhi and ICAR-Indian Agricultural Statistical Research Institute, New Delhi for providing resources to conduct this research. Authors also acknowledge Dr. Sourabh Pargal for providing contour plotting program.

References

Araújo, M.C.U., Saldanha, T.C.B., Galvão, R.K.H., Yoneyama, T., Chame, H.C., Visani, V., 2001. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. Chemometr. Intell. Lab. Syst. 57, 65–73.
Asner, G.P., Martin, R.E., 2008. Spectral and chemical analysis of tropical forests: scaling from leaf to canopy levels. Remote Sens. Environ. 112, 3958–3970.
Asner, G.A., Nepstad, D., Cardinot, G., Moutinho, P., Harris, T., Ray, D., 2003. EO-1 hyperion measures canopy drought stress in Amazonia. AVIRIS Proceedings.
Atzberger, C., Guerif, M., Baret, F., Werner, W., 2010. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. Comput. Electron. Agric. 73 (2), 165–173. https://doi.org/10.1016/j.compag.2010.05.006.
Breiman, L., 1994. Bagging predictors, Technical Report 421. Department of Statistics, UC Berkeley.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. Int. J. Remote Sens. 15, 697–704.
Ceccato, P., Flasse, S., Grégoire, J.M., 2002. Designing a spectral index to estimate vegetation water content from remote sensing data: part 2- validation and applications. Remote Sens. Environ. 82, 198–207.
Chapelle, O., Haffner, P., Vapnik, V., 1999. SVMs for histogram-based image classification. IEEE Trans. Neural Netw. 10 (5), 1055–1064.
Cheng, T., Rivard, B., Sánchez-Azofeifa, G.A., Feng, J., Calvo-Polanco, M., 2010. Continuous wavelet analysis for the detection of green attack due to mountain pine beetle infestation. Remote Sens. Environ. 114, 899–910.
Cheng, T., Rivard, B., Sanchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sens. Environ. 115, 659–670.
Clrovic, D.A., 1997. Feed-forward artificial neural networks: applications to spectroscopy. Trends Analyt. Chem. 16, 148–155.
Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., Panigada, C., 2008. Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. Remote Sens. Environ. 112, 1820–1834.
Cortes, C., Vapnik, V., 1995. Support vector networks. Mach. Learn. 20, 273–297.
Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. 30 (3), 271–278. https://doi.org/10.1016/0034-4257(89)90069-2.
Dall'osto, L., Holt, N.E., Kaligotla, S., Fuciman, M., Cazzaniga, S., Carbonera, D., Frank, H.A., Alric, J., Bassi, R., 2012. Zeaxanthin protects plant photosynthesis by modulating chlorophyll triplet yield in specific light-harvesting antenna subunits. J. Biol. Chem. 287, 41820–41834.
Danson, F.M., Steven, M.D., Malthus, T.J., Clark, J.A., 1992. High-spectral resolution data for determining leaf water content. Int. J. Remote Sens. 13, 461–470.
Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., Cho, M., 2008. LAI and chlorophyll estimation for heterogeneous grassland using hyperspectral measurements. Isprs J. Photogramm. Remote. Sens. 63 (4), 409–426. https://doi.org/10.1016/j.isprsjprs.2008.01.001.
Das, B., Sahoo, R.N., Pargal, S., Krishna, G., Verma, R., Chinnusamy, V., Sehgal, V.K., Gupta, V.K., 2017. 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, 69–83.
Datt, B., 1999. Remote sensing of water content in Eucalyptus leaves. Aust. J. Bot. 47, 909–923.
Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown de Colstoun, E., McMurtrey III, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74, 229–239.
Dawson, T.P., Curran, P.J., Plummer, S.E., 1998. LIBERTY—modelling the effects of leaf biochemical concentration on reflectance spectra. Remote Sens. Environ. 65, 50–60.
Ecarnot, M., Compan, F., Roumet, P., 2013. Assessing leaf nitrogen content and leaf mass per unit area of wheat in the field throughout plant cycle with a portable

spectrometer. Field Crops Res. 140, 44–50.
Eitel, J.U.H., Gessler, P.E., Smith, A.M.S., Robberecht, R., 2006. Suitability of existing and novel spectral indices to remotely detect water stress in Populus spp. For. Ecol. Manage. 229 (1–3), 170–182. https://doi.org/10.1016/ j.foreco.2006.03.027.
Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. Int. J. Remote Sens. 11 (10), 1775–1795. https://doi.org/10.1080/01431169008955129.
Elvidge, C.D., Lyon, R.J.P., 1985. Estimation of the vegetation contribution to the 1.65/2.22 mm ratio in airborne thematicmapper imagery of the Virginia Range, Nevada. Int. J. Remote Sens. 6 (1), 75–88.
Féret, J.-B., François, C., Gitelson, A.A., Barry, K.M., Panigada, C., Richardson, A.D., Jacquemoud, S., 2011. Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. Remote Sens. Environ. 115, 2742–2750.
Filella, I., Peñuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. Int. J. Remote Sens. 15–17, 1459–1470.
Fritsch, S., Guenther, F., 2016. Neuralnet: Training of Neural Networks. R Package Version 1.33. https://CRAN.R-project.org/package=neuralnet.
Galvão, R.K.H., Pimentel, M.F., Araújo, U.M.C., Yoneyamaa, T., Visani, V., 2001. Aspects of the successive projections algorithm for variable selection in multivariate calibration applied to plasma emission spectrometry. Anal. Chim. Acta 443, 107–115. https://doi.org/10.1016/S0003-2670(01)01182-5.
Galvão, R.K.H., Araújo, M.C.U., Fragoso, W.D., Silva, E.C., José, G.E., Soares, S.F.C., Paiva, H.M., 2008. A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm. Chemometr. Intell. Lab. Syst. 92, 83–91.
Gao, B.-C., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 322–331.
Gao, B.-C., Goetz, A.F.H., 1995. Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data. Remote Sens. Environ. 52, 155–162.
Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Anal. Chim. Acta 185, 1–17.
Hardisky, M.A., Klemas, V., Smart, R.M., 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogramm. Eng. Remote Sens. 49, 77–83.
Horler, D.N.H., Docray, M., Barber, J., 1983. The red edge of plant leaf reflectance. Int. J. Remote Sens. 4 (2), 273–288. https://doi.org/10.1080/01431168308948546.
Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R., Foley, W.J., 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. Remote Sens. Environ. 93, 18–29.
Hunt Jr., E.R., Rock, B.N., 1989. Detection of changes in leaf water content using near- and middle-infrared reflectance. Remote Sens. Environ. 30 (1), 43–54.
Imanishi, J., Sugimoto, K., Morimoto, Y., 2004. Detecting drought status and LAI of two Quercus species canopies using derivative spectra. Comput. Electron. Agric. 43, 109–129.
Jacquemoud, S., Ustin, S.L., 2001. Leaf optical properties: a state of the art. Proceedings of the Eighth International Symposium Physical Measurements & Signatures in Remote Sensing, CNES.
Ji, L., Wylie, B.K., Nossov, D.R., Peterson, B., Waldrop, M.P., McFarland, J.W., Rover, J., Hollingsworth, T.N., 2012. Estimating aboveground biomass in interior Alaska with Landsat data and field measurements. Int. J. Appl. Earth Obs. Geoinf. 18, 451–461.
Kokaly, R., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. Remote Sens. Environ. 67 (3), 267–287. https://doi.org/10.1016/S0034-4257(98)00084-4.
Krishna, G., Sahoo, R.N., Pargal, S., Gupta, V.K., Sinha, P., Bhagat, S., Saharan, M.S., Singh, R., Chattopadhyay, C., 2014. Assessing wheat yellow rust disease through hyperspectral remote sensing. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 40 (8), 1413–1416.
Lattin, J.M., Douglas, J., Green, P.E., 2003. Analyzing Multivariate Data. Machine Press, China, pp. 38–40.
Li, L., Ustin, S.L., Riano, D., 2007. Retrieval of fresh leaf fuel moisture content using genetic algorithm partial least squares (GA-PLS) modeling. Geosci. Remote Sens. Lett. IEEE 4, 216–220.
Liaw, A., Wiener, M., 2002. Classification and regression by random forest. R News 2 (3), 18–22.
Mahmood, H., Hoogmoed, W., Henten, E., 2012. Sensor data fusion to predict multiple soil properties. Precision Agric. 13, 628–645.
Mevik, B.-H., Wehrens, R., 2007. The pls package: principal component and partial least squares regression in r. J. Stat. Softw. 18 (2), 1–24.
Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2015. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R Package Version 1.6-6. http://CRAN.R-project.org/package=e1071.
Nduwamungu, C., Ziadi, N., Tremblay, G.F., Parent, L.-E., 2009b. Near-infrared reflectance spectroscopy prediction of soil properties: effects of sample cups and preparation. Soil Sci. Soc. Am. J. 73, 1896–1903.
Nielsen, M.A., 2015. Neural Networks and Deep Learning. Determination Press.
Peng, X., Shi, T., Song, A., Chen, Y., Gao, W., 2014. Estimating soil organic carbon using Vis/NIR spectroscopy with SVMR and SPA methods. Remote Sens. (Basel) 6, 2699–2717.
Peñuelas, J., Gamon, J.A., Griffin, K.L., Field, C.B., 1993. Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. Remote Sens. Environ. 46 (2), 110–118. https://doi.org/10.1016/0034-4257(93)90088-F.
Peñuelas, J., Piñol, J., Ogaya, R., Filella, I., 1997. Estimation of plant water concentration by the reflectance water index WI (R900/R970). Int. J. Remote Sens. 18, 2869–2875.
Pu, R., Ge, S., Kelly, N.M., Gong, P., 2003. Spectral absorption features as indicators of water status in coast live oak (Quercus agrifolia) leaves. Int. J. Remote Sens. 24, 1799–1810.
Pyne, S.J., Andrews, P.L., Laven, R.D., 1996. Introduction to Wildland Fire, 2nd ed. Wiley, New York.
Ramoelo, A., Skidmore, A.K., Schlerf, M., Mathieu, R., Heitkönig, I.M.A., 2011. Water removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. ISPRS J. Photogramm. Remote Sens. 66, 408–417.
Roberts, D.A., Green, R.O., Adams, J.B., 1997. Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. Remote Sens. Environ. 62, 223–240.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Freden, S.C., Mercanti, E.P., Becker, M. (Eds.), Third Earth Resources Technology Satellite–1 Syposium. Volume I: Technical Presentations. NASA SP-351, NASA, Washington, D.C, pp. 309–317.
Ryu, C., Suguri, M., Umeda, M., 2011. Multivariate analysis of nitrogen content for rice at the heading stage using reflectance of airborne hyperspectral remote sensing. Field Crops Res. 122, 214–224.
Sahoo, R.N., Ray, S.S., Manjunath, K.R., 2015. Hyperspectral remote sensing of agriculture. Curr. Sci. 108 (5), 848–859.
Schlerf, M., Atzberger, C., Hill, J., Buddenbaum, H., Werner, W., Schueler, G., 2010. Retrieval of chlorophyll and nitrogen in Norway spruce (Picea abies L. Karst.) using imaging spectroscopy. Int. J. Appl. Earth Obs. Geoinf. 12, 17–26.
Serrano, L., Ustin, S.L., Roberts, D.A., Gamon, J.A., Peñuelas, J., 2000. Deriving water content of chaparral vegetation from AVIRIS data. Remote Sens. Environ. 74, 570–581.
Sims, D.A., Gamon, J.A., 2003. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. Remote Sens. Environ. 84, 526–537.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14 (3), 199–222.
Stimson, H.C., Breshears, D.D., Ustin, S.L., Kefauver, S.C., 2005. Spectral sensing of foliar water conditions in two co-occurring conifer species: Pinus edulis and Juniperus monosperma. Remote Sens. Environ. 96, 108–118.
Subasi, A., Erçelebi, E., 2005. Classification of EEG signals using neural network and logistic regression. Comput. Methods Programs Biomed. 78, 87–99.
Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Legg, C., De Dieu, M.J., 2004. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. Remote Sens. Environ. 90, 23–43.
Thulin, S., Hill, M., Held, A., Jones, S., Woodgate, P., 2014. Predicting levels of crude protein, digestibility, lignin and cellulose in temperate pastures using hyperspectral image data. Am. J. Plant Sci. 5 (7), 997–1019.
Ullah, S., Skidmore, A.K., Ramoelo, A., Groen, T.A., Naeem, M., Ali, A., 2014. Retrieval of leaf water content spanning the visible to thermal infrared spectra. ISPRS J. Photogramm. Remote Sens. 93, 56–64.
Ustin, S.L., Roberts, D.A., Pinzon, J., Jacquemoud, S., Gardner, M., Scheer, G., Castañeda, C.M., Palacios-Orueta, A., 1998. Estimating canopy water content of chaparral shrubs using optical methods. Remote Sens. Environ. 65, 280–291.
Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., Skidmore, A.K., 2005. Tropical mangrove species discrimination using hyperspectral data: a laboratory study. Estuar. Coast. Shelf Sci. 65, 371–379.
Williams, P.C., Sobering, D.C., 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Spectrosc. 1, 25–32.
Zarco-Tejada, P.J., Ustin, S.L., 2001. Modeling canopy water content for carbon estimates from MODIS data at land EOS validation sites. Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium 342–344.
Zarco-Tejada, P.J., Rueda, C.A., Ustin, S.L., 2003. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. Remote Sens. Environ. 85, 109–124.
Zhang, F., Zhou, G., 2015. 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. (Basel) 7, 15203–15223.
Zhao, L., Xiang, W., Li, J., Lei, P., Deng, X., Fang, X., Peng, C., 2015. Effects of topographic and soil factors on woody species assembly in a Chinese subtropical evergreen broadleaved forest. Forests 6, 650–669.
Zhao, Y.R., Li, X., Yu, K.Q., Cheng, F., He, Y., 2016. Hyperspectral imaging for determining pigment contents in cucumber leaves in response to angular leaf spot disease. Sci. Rep. Ist. Super. Sanita 6.
Zhu, J., Huang, Z., Sun, H., Wang, G., 2017. Mapping forest ecosystem biomass density for Xiangjiang River Basin by combining plot and remote sensing data and comparing spatial extrapolation methods. Remote Sens. (Basel) 9 (241), 1–23.