J. Agr. Sci. Tech. (2016) Vol. 18: 1705-1718

Using Leaf Based Hyperspectral Models for Monitoring Biochemical Constituents and Plant Phenotyping in Maize

F. Kahriman[1*], K. Demirel[2], M. Inalpulat[2], C. O. Egesel[3], and L. Genc[2]

ABSTRACT

The aim of this study was to develop and validate qualitative and quantitative models to discriminate different types of maize and also estimate biochemical constituents. Spectral data were taken from the central leaf of randomly-chosen plants grown in field trials in 2011 and 2012. Leaf chlorophyll and protein content and stalk protein content were determined in the same plants. Four different Support Vector Machine (SVM) models were generated and validated in this study. In qualitative models, maize type was designated as dependent variable while Full Spectral (FS) data (400-1,000 nm) and Spectral Indices (SI) data (34 indices/bands) were independent variables. In the two quantitative models (SVMR-FS and SVMR-SI), independent variables were the same, whereas dependent variables were assigned as the quantitatively measured traits. Results showed the qualitative models to be a robust method of classification for distinguishing different maize types, such as High Oil Maize (HOM), High Protein Maize (HPM) and standard (NORMAL) maize genotypes. The SVMC-FS model was superior to SVMC-SI in terms of the genotypic classification of maize plants. Quantitative models with full spectral data gave more robust prediction than the others. The best prediction result (RMSEC= 222.4 µg g^{-1}, R^2 for Cal= 0.739, SEP= 213.3 µg g^{-1}; RPD= 2.04 and r= 0.877) was obtained from the SVMR-FS model developed for chlorophyll content. Indirect estimation models, based on relationships between leaf-based spectral measurements and leaf and stalk protein content, were less satisfactory.

Keywords: Genotypic classification, Support Vector Machine, *Zea mays.*

INTRODUCTION

Hyperspectral techniques are a novel method widely used in different areas nowadays. Remote sensing and ground-based spectral techniques are utilized in various fields such as agriculture, geology, and mining. The application of hyperspectral techniques in agriculture particularly has been made possible by research into yield prediction (Weber *et al.,* 2012; Demirel *et al.,* 2014), detection of plant stress (Genc *et al.,* 2011), measurement of pigment content (Sims and Gamon, 2002), analysis of biochemical components (Özyiğit and Bilgen, 2013; Kaufman *et al.,* 2010).

Previous studies aimed to associate spectral data and the biochemical components of plant leaves by developing hyperspectral models obtained from canopy level measurements. One well-known example is the PROSPECT radiative transfer model, aimed at determining protein and chlorophyll content in plant leaves (Botha *et al.,* 2005). Another study showed that it was

[1] Department of Field Crops, Faculty of Agriculture, Çanakkale Onsekiz Mart University, Çanakkale, Turkey.
[*] Corresponding author; e-mail: fkahriman@hotmail.com
[2] Agricultural Sensor and Remote Sensing Laboratory, Faculty of Agriculture, Çanakkale Onsekiz Mart University, Çanakkale, Turkey.
[3] Department of Agricultural Biotechnology, Faculty of Agriculture, Çanakkale Onsekiz Mart University, Çanakkale, Turkey.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

possible to determine pigment content in wheat leaves using spectral indices generated by spectral measurements (Jin *et al.,* 2012). Spectral prediction models were also developed to determine the chlorophyll content in maize (Haboudane *et al.,* 2002). In recent years, there has been considerable effort to develop innovative models or instruments utilizing hyperspectral techniques, such as HyperART, which enables the user to determine disease symptoms and pigment content in leaves (Bergsträsser *et al.,* 2015). Most studies using these methods have so far focused on determining pigment content and nutrient deficiency in plants, as summarized above. Comparative studies involving the new approaches could also make a significant contribution to the literature and open up new areas of application. In the case of using this innovative method to characterize and distinguish genetic materials in maize breeding experiments, it would become a powerful tool for relatively easy use by researchers.

The genetic resources of maize vary highly in their morphological and physiological features. Specialty maize genotypes (e.g. high oil, high protein) differ not only in their kernel structure but also leaf characteristics (leaf thickness, pigment content, etc.) (Kahrıman, 2013). Discriminant models, which have the potential to discern between normal and specialty maize genotypes, could be developed based on these differences. Indeed, previous studies have shown that it is possible to discriminate different species with different leaf characteristics using hyperspectral measurements (Slaton *et al.,* 2001). In recent years, research has been carried out to distinguish between weeds in agricultural crops based on hyperspectral measurements at canopy level (Martin *et al.,* 2011; De Castro *et al.,* 2012; Herrman *et al.,* 2013). Additionally, a recent paper illustrated how it was possible to monitor crop status at canopy level based on hyperspectral measurements (Rundquist *et al.,* 2014). However, although there have been approaches aimed at discerning different species, to our knowledge there has been no/limited research aimed at discerning genotypes within the same species.

The discriminatory power of hyperspectral techniques is further increased by analyzing data using modern statistical methods, such as Support Vector Machine (SVM), which is one of the best methods for non-linear classification and solving regression problems. SVM is applicable to small number of samples, it is not affected by outliers and is a user friendly method compared to other multivariate data analyses. Also, it has high generalization ability because it uses the maximum margin of hyperplane for discrimination and has several non-linear discriminant functions (Abu-Kalaf, 2015). The different functions in SVM are linear, polynomial, and radial basis functions. Among these, radial basis function is the most commonly used thanks to its attractive features in preserving the structure of the data analyzed (Liu *et al.,* 2012). SVM was used for accurate disease identification (Abu-Kalaf and Salman, 2014). Previous studies have shown that using full spectral data may give better results for discrimination of samples and/or quantification of constituents rather than using data for spectral indices. However, in the main, previous studies addressed the relationship between spectral indices and the investigated traits. There is a need for more extensive research using both methods in order to reveal which method is more effective in its discrimination/quantification potential.

From this standpoint, the aim of this study was to generate and compare spectral models based on data from full spectral measurements and spectral indices. We hypothesized that there was potential to discriminate special maize genotypes from normal maize genotypes based on differences in their leaf structure. Additionally, we wanted to show the possibility of indirect estimation of protein content in other plant parts based on leaf-based spectral measurement.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

MATERIALS AND METHODS

Plant Material and Experimental Design

We used a set consisting of 8 genotypes, Including normal (B73 and Mo17), High Oil (IHO, IHOxB73, IHOxMo17) and High Protein (IHP, IHPxB73, IHPxMo17) maize genotypes. IHO and IHP were obtained from North Central Regional Plant Introduction Station, Ames, USA. Experimental hybrids were generated by crossing B73 and Mo17 with high oil and high protein parents in 2010. Field experiments were conducted at the Dardanos Agricultural Research Station of Çanakkale Onsekiz Mart University in 2011-2012. The sowing dates were May 18 and May 13 in 2011 and 2012, respectively. Randomized complete block design with three replicates was used as the experimental design. Genotypes were planted in two-row plots, 70×20 cm apart. Fertilization was made with 180 kg ha^{-1} pure nitrogen. Drip irrigation was used for the water management. Plots were given 422.6 mm and 420.2 mm of irrigation water in the first and second year of the experiment, respectively. Monthly data showed that mean temperature values were higher in both years than the long term average. The second year was hotter, especially in the growing period (May to September), compared to the first year of the experiment. However, precipitation data indicates that the second year received more rainfall than the first (Table 1).

Spectral Measurements

Hyperspectral data were collected with an ASD Field Spectroradiometer (Analytical Spectral Devices Inc, Boulder, CO, USA) from the central leaf of each sampled plant. The measured leaves were tagged with a permanent marker pencil for further use. In each of the two experimental years, spectral measurements were made within 325-1075 nm intervals on five sampling dates designated as Days After Sowing (DAS): DAS40, DAS60, DAS82, DAS100, DAS122. We measured 720 randomly-selected plants (360 plants per year) in this study. Five scans were taken from each of the examined plants using a lens with 1° field-of-view. These scans were simultaneously recorded onto a laptop computer serially connected to the spectroradiometer. All measurements were made on cloudless days between the hours of 10:00-14:00. Sunlight was used as the light source and photosynthetically active radiation (PAR) and temperature values were recorded (Figure 1) on a microclimate station (Onset Computer Corporation, USA) at 2-minute intervals. The spectroradiometer

Table 1. Monthly temperature and precipitation in Canakkale Province.[a]

	Mean temperature (°C)			Total precipitation (mm)		
Month	2011	2012	Long Term	2011	2012	Long Term
January	7.01	4.79	6.3	35.4	72.4	93.3
February	6.44	4.76	6.7	16.7	74.8	71.5
March	8.48	9.16	8.3	52.9	25.4	68.4
April	11.45	15.04	12.6	58.8	59.8	46.5
May	18.18	19.57	17.5	14	89.4	32.2
June	24.05	26.38	22.3	44.6	6.0	21.8
July	27.78	30.09	25.1	0.2	0.0	11.9
August	26.95	28.66	24.9	1.6	45.5	6.5
September	25.00	23.97	20.8	2.4	3.8	23.6
October	15.41	20.51	16.0	121	95.5	56.2
November	9.40	14.94	11.8	4.6	42.7	86.7
December	9.54	8.37	8.5	157.3	222.7	109.8

[a] Source: Turkish Meteorological Office. Long term data is mean from 1950 to 2014.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

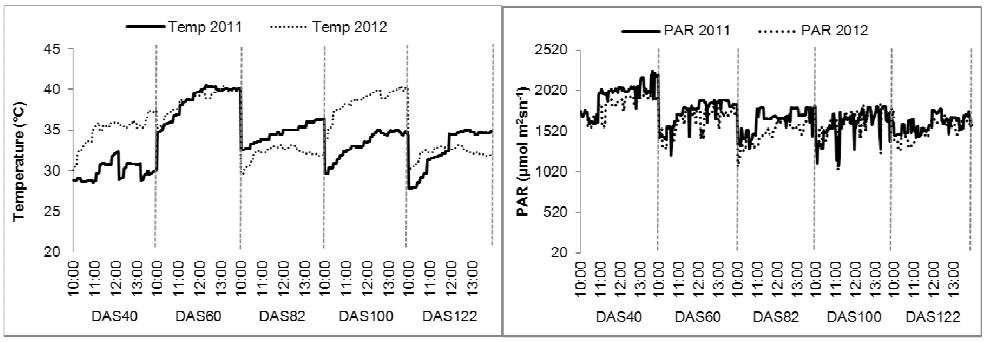

Figure 1. Temperature and Photosynthetically Active Radiation (PAR) values within 10:00-14:00 time interval on measurement days.

was calibrated with a plate of barium sulphate prior to five plant measurements. After completion of the field experiment, the spectral data were converted to text files using the device's own software. Then, five scans of each sample were averaged to find the mean spectral reflectance of each sample. Collected data were used to calculate several spectral indices (30 indices) utilized in the current study (Table 2). Twenty-two of these indices were related to leaf chlorophyll content and nine were water indices, which were commonly used in previous studies (Jordan, 1969; Tucker, 1979; Penuelas *et al.,* 1994; Carter *et al.*, 1996; Gamon and Surfus, 1999; Penuelas *et al.,* 1995; Daughtry *et al.,* 2000; Barton 2001; Strachan *et al.*, 2002; Sims and Gamon, 2002; Gitelson, 2004; Gitelson *et al.*, 2005; Babar *et al.*, 2006; Gitelson *et al.*, 2006; Guang and Liu, 2009; Kaufman *et al*., 2010; Bergsträsser *et al*., 2015). We used the water indices to further discriminate the experimental material. Also, we used two band values (940 and 970 nm) and their ratios (970/940, 940/970), related to pigment concentration, for spectral model development.

Determination of Biochemical Constituents

Ten leaf discs were taken from the tagged leaves used in spectral measurements using a cork borer. These ten samples were bagged in striped bags and saved in dry ice. Next, in the chlorophyll analysis, the chlorophyll a, chlorophyll b, and chlorophyll a+b content of the samples was determined according to the Dimethyl Sulphoxide (DMSO) method (Hiscox and Israelstam, 1979). The measured plants were harvested at soil level, and then separated by plant parts (stalk, leaves, and ears). The separated parts of the samples were dried at 70°C. The dried samples were ground in a laboratory mill (Fritsch, pulverisette 14, Germany) with a 0.5 mm-size sieve. Protein content in the leaf and stalk samples was determined using a Near Infrared Spectrophotometer (Spectrastar 2400D, Unity Scientific, USA). For this purpose, scans were taken at 1,200-2,400 nm intervals using the appropriate sample cup (stationary or rotated) of the instrument according to sample quantity. These scans were subjected to an INGOT (International NIR Global Operating Technologies) calibration model, namely, Grass Silage and Forage, to estimate the protein content of the samples.

Development of Calibration Models

The calibration models were developed using Unscrambler 10X software (CAMO Software, Oslo, Norway). Averaged scans (raw spectra) were converted to reflectance values (1/R) before model development. To reduce the effect of noise at the beginning (375-400 nm) and end (1,000-1,075 nm) of the spectral range, only measurements between 400-1,000 nm were used for the

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

Table 2. Spectral indices used in this study.

Spectral Indices	Equations
Water Band Index (WBI); Normalized Pigment Chlorophyll Index (NPCI)[a]; Photochemical Reflectance Index (PRI)[a]; Red Edge Normalized Difference Vegetation Index ($NDVI_{705}$); Red/Green Index (RGI)[a]; Normalized Difference Vegetation Index (NDVI)[a]; Infrared-Red Difference Index (IR-RED); Pigment Specific Simple Ratio (PSSR)[a]; Simple Ratio (705) [a]; Structure Intensitive Pigment Index (SIPIpen)[a]; Green Normalized Difference Vegetation Index (GNDVI)[a]; REPht; Wide Dynamic Range Vegetation Index (WDRVI) [a]; Green Chlorophyll Indices (GCI)[a]; Ratcart ratio[a]; Plant Senescence Reflectance Index (PSRI)[a]; Red Edge Model (REM); Green Model (GM)[a]; Vogelman (VOG)[a]; Zarco and Miller Index (ZM)[a]; Enhanced Vegetation Index (EVI)[a]; Floating-position water band index (fWBI); Normalized Water Indices (NWI-1); Normalized Water Indices (NWI-2); Simple Ratio Index (SR)[a]; Simple Ratio Pigment Index (SRPI)[a]; Structure Intensive Pigment Index (SIPI)[a]; Modified Chlorophyll Absorption in Reflectance Index (MCARI)[a]; Transformed Chlorophyll Absorption in Reflectance Index (TCARI)[a]; Modified Triangle Vegetation Index (MTVI) [a]	$WBI = \frac{R_{900}}{R_{970}}; NPCI = \frac{(R_{680} - R_{430})}{(R_{680-} R_{430})};$ $PRI = \frac{(R_{531} - R_{570})}{(R_{531} + R_{570})}; NDVI_{705} = \frac{R_{750} - R_{705}}{R_{750} + R_{705}}; RGI = \frac{R_{690}}{R_{550}}$ $; NDVI = \frac{R_{800} - R_{680}}{R_{800} + R_{680}}; IR - RED = R_{789} - R_{663};$ $PSSR = \frac{R_{775}}{R_{747}}; SR_{705} = \frac{R_{750}}{R_{705}}; SIPI_{PEN} = \frac{(R_{800} - R_{445})}{(R_{800} + R_{445})};$ $GNDVI = \frac{(R_{800} - R_{445})}{(R_{800} + R_{445})}; REPht = \frac{R_{600}}{R_{800}};$ $WDRVI = \frac{0.1x(R_{800} - R_{680})}{0.1x(R_{800} + R_{680})}; GCI = \frac{R_{800}}{R_{550}} - 1;$ $Ratcart = \frac{R_{695}}{R_{760}}; PSRI = \frac{R_{680} - R_{500}}{R_{750}}; REM = \frac{R_{800}}{R_{700}} - 1;$ $GM = \frac{R_{800}}{R_{550}} - 1; VOG = \frac{R_{740}}{R_{720}};$ $ZM = \frac{R_{750}}{R_{710}}; EVI = \frac{2.5(R_{800} - R_{675})}{(R_{800} + 6xR_{675} - 7.5xR_{450} + 1)};$ $fWBI = \frac{R_{900}}{\min(R_{930} - R_{980})}; NWI - 1 = \frac{(R_{970} - R_{900})}{(R_{970} + R_{900})};$ $NWI - 2 = \frac{(R_{970} - R_{850})}{(R_{970} + R_{850})}; SR = \frac{R_{800}}{R_{680}};$ $SRPI = \frac{R_{430}}{R_{680}}; SIPI = \frac{(R_{800} - R_{450})}{(R_{800} - R_{650})}$ $MCARI = [(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})]x\left(\frac{R_{700}}{R_{670}}\right)$ $TCARI = 3((R_{700} - R_{670}) - 0.2(R_{700} - R_{550})x\left(\frac{R_{700}}{R_{670}}\right))$ $MTVI = 1.5x[1.2x(R_{712} - R_{550}) - 2.1x(R_{670} - R_{550})]$

[a] Chlorophyll sensitive index.

model development. Data were randomly partitioned into two sets, namely, calibration sets (n= 240 per year) and validation sets (n= 120 per year). Pre-treatment of the spectral data was carried out to improve the prediction performance of the spectral models. Standard Normal Variate (SNV) transformation was used for light scatter correction (Dhanoa *et al.,* 1994), and first derivative (Segment size= 5, Gap size= 2) was applied to correct the baseline shift in spectral data (Chu *et al.,* 2004). These pre-treatments were applied only to Full Spectral (FS) data. To generate the Spectral Indices (SI) data, we used untreated spectral scans

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

from the calibration and validation sets independently.

In the current study, we compared the two main models and two sub-models, which were generated by use of full spectral data (400-1,000 nm) and spectral indices (34 indices/bands), Qualitative and quantitative models were developed for the classification of genotypes according to their groups (NORMAL, HOM, HPM) and for quantification of biochemical constituents (leaf chlorophyll and protein content, stalk protein content), respectively. The qualitative models (n= 240 for each year) were generated by Support Vector Machine Classification (SVMC) while the quantitative models were generated by Support Vector Machine Regression (SVMR) methods. Radial Basis Function (RBF) was used as kernel function with default parameter for model development (CAMO Software, Oslo, Norway). The dependent variable in the qualitative models (SVMC-FS and SVMC-SI) was designated as the maize type (HOM, HPM, and NORMAL). The independent variables were full spectral data for the SVMC-FS model and spectral indices data for the SVMC-SI model. Independent variables were the same for the quantitative models (SVMR-FS and SVMR-SI), but dependent variables were assigned as the quantitatively measured traits: chlorophyll a, chlorophyll b, chlorophyll a+b, and leaf and stalk protein content of the whole plant. We construct two different models regarding the abovementioned statistical methods. First type of the models was direct models, for which chlorophyll content was determined in the same area of the measured leaf as spectral measurements. Second type of the models was indirect estimation models where leaf-based (central leaf) spectral measurements were associated with the protein content of different parts of the whole plant such as leaves and stalk. External validation was undertaken to learn the estimation power of the generated models using validation datasets (n= 120) generated by independent samples from calibration sets. True classification rate was used in the evaluation of qualitative calibration models. Quantitative models were evaluated according to parameters such as Root Mean Squares Error of Calibration (RMSEC), Standard Error of Estimation (SEE), determination coefficient (R^2), Standard Error of Prediction (SEP), Relative Performance to Deviation (RPD), and multiple correlation coefficient (r).

RESULTS AND DISCUSSION

Changes in Measured Traits

Descriptive statistics of dependent variables in the quantitative models are presented here for the calibration (n= 240 for each year) and validation (n= 120 for each year) sets (Table 3). The mean values for quantitative traits were higher in 2012 than

Table 3. Descriptive statistics for measured traits in calibration and validation sets by year.

		Calibration				Validation			
	Year	Mean	Min	Max	STD[a]	Mean	Min	Max	STD
Chlorophyll a	2011	707.2	7.51	2205.3	429.9	690.1	5.24	1765.7	434.3
	2012	982.9	1.21	3057.7	528.2	929.6	8.96	2558.4	554.9
Chlorophyll b	2011	132.7	5.78	360.7	79.74	129.9	16.1	332.9	76.9
	2012	185.1	2.18	888.4	120.5	174.7	14.8	609.5	105.7
Chlorophyll a+b	2011	850.4	23.2	2243.8	509.9	830.4	23.2	2125.0	515.2
	2012	1182.8	3.40	3915.2	642.7	1118.2	30.7	3024.3	656.7
Leaf protein	2011	12.21	5.79	19.85	2.74	12.07	5.18	17.24	2.68
	2012	12.58	7.58	17.58	2.06	12.53	7.83	17.37	2.23
Stalk protein	2011	6.70	3.25	13.9	2.25	6.89	3.35	12.43	2.23
	2012	7.14	3.31	12.5	1.71	7.22	3.91	12.26	1.61

[a] Standard Deviation.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

2011. This shows that climatological differences in the second year caused significant changes in the observed traits. The measured traits herein are quantitatively inherited and they interact with environmental factors. Previous studies indicated that pigment concentration in maize leaves decreases under stress conditions (Homayoun *et al.,* 2011). Our monthly data indicated that the second year of the experiment had higher temperature values; however, they were not high enough to cause temperature stress (Table 1). Thus, we expected that the chlorophyll content would be less in 2012. However, previous studies showed that a moderate increase in temperature might enhance photosynthesis performance in plants, depending on geographical region and species (Xu *et al.,* 2011). Additionally, high temperatures in 2012 might have promoted an increase in leaf biochemical traits such as chlorophyll and protein content.

Evaluation of Qualitative Models for Plant Phenotyping

The True Classification Rates (TCR) of the qualitative models, developed for discrimination of different maize types (HOM, HPM, NORMAL), are summarized in Table 4. The SVMC-FS model had greater success in classification of different types of maize (TCR> 70-80%) than the SVMC-SI model. *TCR* values of the SVMC-FS model were similar in both years (2011= 83.1; 2012= 83.4), while the SVMC-SI model had a lower *TCR* value (68.9%) in the first than the second year (Table 4). It is noticeable that the NORMAL genotypes had lower true classification rate than the others in the SVMC-FS model. External validation results also showed that SVMC-FS had better classification potential than the SVMC-SI model. Overall, we can say that this model may have the potential to discriminate different maize types (Table 4). The literature lacks data on the discriminatory potential of spectral techniques for different maize types. Therefore, comparing our results with previous studies is difficult. However, there have been a few studies of cultivar identification for different agricultural products. Kong *et al.* (2013) found that using full spectra was better for cultivar identification in rice than optimal wavelength selection procedures. They could distinguish four cultivars using SVM model with 100% accuracy. Our results are in agreement with this study indicating that using full spectral data is appropriate for discrimination studies. He *et al.* (2007) also developed Artificial Neural Network (ANN) models to discriminate 8 groups of tea varieties. They classified the tea varieties with 100% accuracy in their study. However, we should note that the abovementioned studies were conducted in controlled environments and relied on seed or processed samples, favoring the rate of

Table 4. True Classification Rates (TCR) in SVMC-FS and SVMC-SI models for distinguishing maize types.

		True classification rate					
		Calibration			Validation		
Year	TYPE	N	SVMC-FS	SVMC-SI	N	SVMC-FS	SVMC-SI
2011	NOR	60	78.3	75.5	30	66.7	66.7
	HOM	90	83.3	65.6	45	57.8	44.4
	HPM	90	87.8	65.6	45	53.3	44.4
2012	NOR	60	76.7	70.0	30	63.3	63.3
	HOM	90	86.7	76.7	45	48.9	44.4
	HPM	90	86.7	76.7	45	60.0	51.3
2011		240	83.1	68.9	120	59.3	51.8
2012		240	83.4	74.5	120	57.4	53.0

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

success. To our best knowledge, our study is the first for genotypic classification using hyperspectral techniques in field conditions. There may be numerous environmental factors affecting the model performance in field studies as in the current research. Thus, classification powers of the models developed here were lower than those reported in literature.

Evaluation of Quantitative Models for Monitoring Biochemical Constituents

The evaluation parameters of the models revealed that the SVMR-FS model gives more robust results than the SVMR-SI model (Table 5). Both models gave more accurate estimations in the first year of the experiment (Table 5, Figure 2). External validation also showed that the SVMR-SI model made a more accurate prediction for quantification of leaf chlorophyll content. Concerning models for chlorophyll content, the SVR-FS model for chlorophyll a content, developed by using 2011 data, was the best model, giving robust estimations (RMSEC= 222.4 µg g^{-1}, R^2 for Cal= 0.739, SEP= 213.3 µg g^{-1}; RPD= 2.04, r= 0.877). The SVMR-SI model for total chlorophyll content followed this model in terms of prediction performance (Table 5, Figure 2). When considering model evaluation parameters, R^2 values of the models for total chlorophyll content varied between 0.672 and 0.739, and these values were higher (R^2 changes 0.14 to 0.66) than in some previous studies (Botha *et al.,* 2005) yet lower (R^2 varied between 0.868 to 0.964) than others (Jin *et al.,* 2012).

Indirect estimation models had relatively lower robustness than the models for pigment concentration. However, it was understood that the spectral measurements taken from the central leaf of a maize plant could give an idea of the protein content in the whole leaf and stalk of the plant. Leaf and stalk protein content were more accurately predicted based on the first year data, as seen in models related to chlorophyll content. The SVMR-FS model had an obvious advantage over the SVMR-SI model in determination of both leaf and stalk protein content (Figure 2). The best prediction result (RMSEC= 1.666%, R^2=

Table 5. Evaluation parameters in calibration and validation sets for quantitative estimation models.

Trait	Model	Year	Calibration (n= 240)		Validation (n= 120)	
			RMSEC	R^2	SEP	RPD
Chlorophyll a	SVMR-FS	2011	222.4	0.739	213.3	2.04
		2012	292.0	0.685	266.4	1.96
Chlorophyll a	SVMR-SI	2011	234.6	0.709	216.4	2.01
		2012	286.4	0.688	262.6	1.99
Chlorophyll b	SVMR-FS	2011	46.5	0.666	41.8	1.84
		2012	64.5	0.565	54.3	1.69
Chlorophyll b	SVMR-SI	2011	50.9	0.592	42.7	1.80
		2012	63.3	0.572	54.2	1.70
Chlorophyll a+b	SVMR-FS	2011	263.8	0.739	256.1	2.01
		2012	355.0	0.672	320.6	1.93
Chlorophyll a+b	SVMR-SI	2011	280.8	0.704	260.2	1.98
		2012	347.8	0.688	315.8	1.96
Leaf protein	SVMR-FS	2011	1.666	0.649	1.615	1.66
		2012	1.456	0.504	1.690	1.32
Leaf protein	SVMR-SI	2011	1.795	0.574	1.711	1.57
		2012	1.423	0.518	1.656	1.35
Stalk protein	SVMR-FS	2011	1.645	0.546	1.674	1.33
		2012	1.449	0.318	1.411	1.13
Stalk protein	SVMR-SI	2011	1.622	0.513	1.642	1.36
		2012	1.417	0.325	1.388	1.17

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

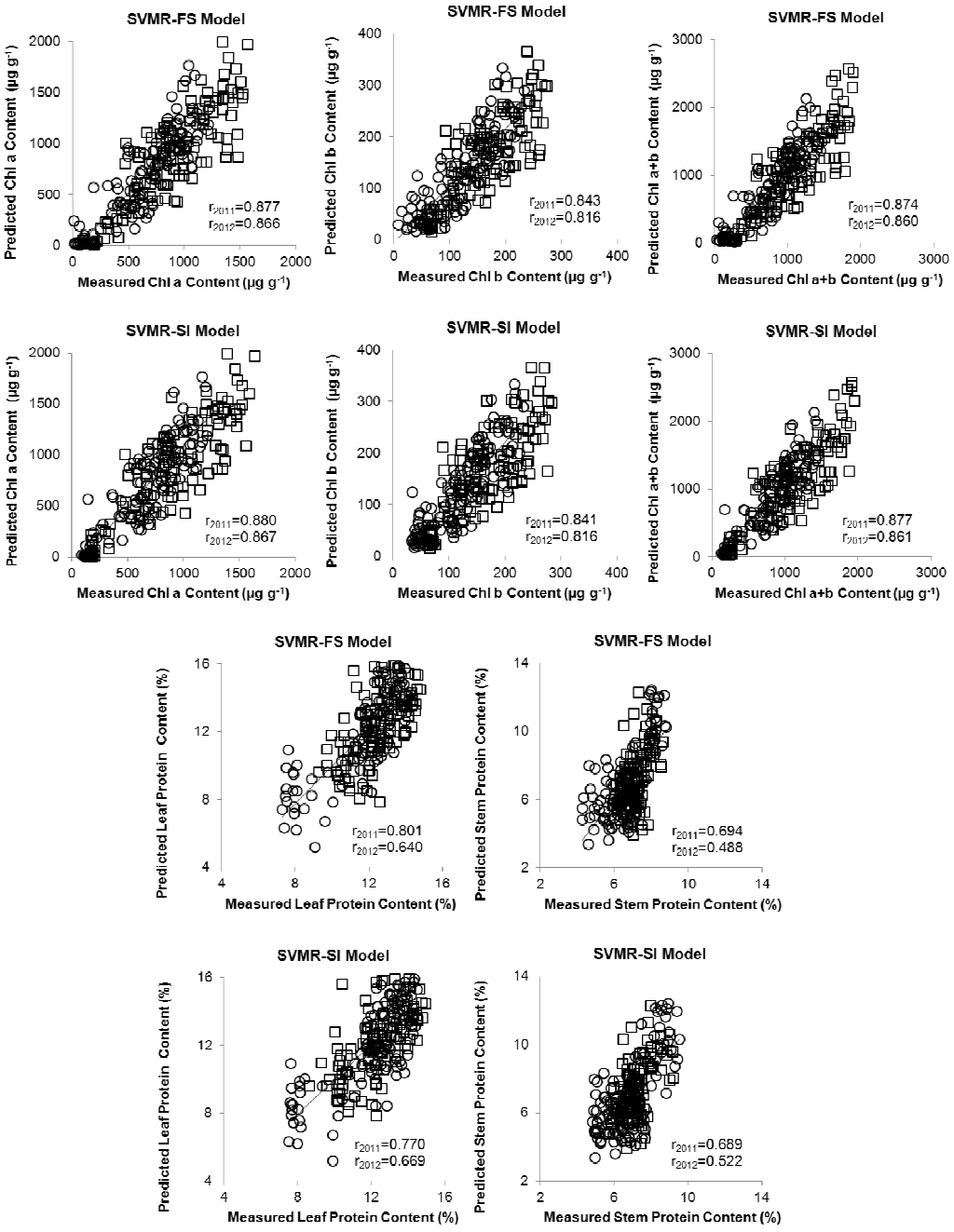

Figure 2. Observed and predicted values in calibration sets of SVMR-SI and SVMR-FS models for chlorophyll a, chlorophyll b, and chlorophyll a+b, leaf protein content, and stem protein content, respectively. Null circles indicate the 2011 while null squares indicate 2012 data.

0.649, SEP= 1.615, r= 0.801) was obtained from the SVMR-FS model developed using first year data for prediction of leaf protein content. This finding was confirmed by external validation (Table 5). Although the results obtained demonstrate that the SVMR-FS model was superior, neither this nor the SVR-SI model showed sufficient success regarding indirect estimation of leaf and stalk protein content. Indeed, the *RPD*

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

values of models for leaf and stalk protein content verified this finding. Diller (2002) reported that when *RPD* value was below 2, calibration model is not suitable for use. We should say that both SVMR-FS and SVMR-SI models need to be improved for estimation power, because *RPD* values were below 2 for all of them.

On the other hand, it should be considered that the R^2 and *RPD* values of the SVMR-SI model were notably lower compared to SVMR-FS models. Although there have been some studies aimed at determination of leaf biochemical constituents in the literature (Yi *et al.,* 2008; Wang and Li, 2012), they did not investigate the relationship of spectral measurement to biochemical constituents in other parts of the plant, such as leaves and stalk. In this regard, the current study may well be notable for proposing a new approach to the use of spectral techniques. More detailed investigation into the relationship of leaf-based spectral measurement with other biochemical constituents in different parts of the maize plant, and testing the models under different conditions, could provide a better understanding of the potential of spectral techniques in maize research.

The estimation power of quantitative models was skewed by the effect of climatological changes during the two experimental years. Models based on second year data appeared to be less reliable than models for the first year. This was due to the second year being hotter, and also having irregular light conditions, as seen in Figure 1. We used sunlight as the light source in the current study. Spectral reflectance is significantly affected by climatological conditions such as air temperature, air moisture, dust, aerosols and cloudlessness (Pfitzner *et al.,* 2011). In our study, the temperature and *PAR* values changed year to year during measuring (Figure 1), which in this case resulted in differences in the estimation power of the generated models based on yearly data. Our data showed, interestingly, that air temperature was higher in the second year while *PAR* was lower. Although it was expected that an increase in air temperature would result in an increase in *PAR* values, some studies have shown that this relationship could in fact show differences depending on climate (Chang and Root, 1975). Indeed, Furuuchi *et al.* (2013) found that temperature increase over time resulted in a global radiation decrease. Our results were in line with this finding and this phenomenon should be clarified. The adverse effect of high temperature and irregular light conditions on the power of spectral models was observed in our study. Some advice is given in the next section for improving the prediction accuracy of generated models.

CONCLUSIONS

Hyperspectral techniques can be used for different purposes in maize research. The qualitative models developed herein have the potential to discriminate maize plants by their genotypic specialties (HOM, HPM or NORMAL). We found that leaf-based spectral measurements can also be used for quantitative determination of leaf pigment and protein content. Our findings revealed that climatological changes had an important effect on the discrimination/estimation power of both qualitative and quantitative calibration models. This study was conducted using a destructive sampling procedure resulting in the taking of spectral measurements from different plants. In future research, it may be possible to develop more robust models using non-destructive sampling methods. Individual plant selection and characterization could be practiced with the support of spectral techniques. New instruments which have a wide scanning interval (750-2,500 nm) and own-light source may provide sensitive results for development of qualitative and quantitative prediction models. Additionally, remote sensing techniques can be used for genotype characterization in maize breeding studies with the use of canopy level measurements.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

REFERENCES

1. Abu-Kalaf, N. and Salman, M. 2014. Visible/Near Infrared (VIS/NIR) Spectroscopy and Multivariate Data Analysis (MVDA) for Identification and Quantification of Olive Leaf Spot (OLS) disease. *Palest. Tech. Univ. Res. J.,* **2:** 1-8.
2. Abu-Kalaf, M. 2015. Sensing Tomato's Pathogen Using Visible/Near Infrared (VIS/NIR) Spectroscopy and Multivariate Data Analysis (MVDA). *Palest. Tech. Univ. Res. J.,* **3:** 12-22.
3. Babar, M. A., Reynolds, M. P., Van Ginkel, M., Klatt, A. R., Raun, W. R. and Stone, M.L. 2006. Spectral Reflectance Indices as Potential Indirect Selection Criteria for Wheat Yield under Irrigation. *Crop Sci.,* **46:** 578–588.
4. Barton, C. V. M. 2001. A Theoretical Analysis of the Influence of Heterogeneity in Chlorophyll Distribution on Leaf Reflectance. *Tree Physiol.,* **21:** 789–795.
5. Bergsträsser, S., Fanourakis, D., Schmittgen, S., Cendrero-Mateo, M. P., Jansen, M., Scharr, H. and Rascher, U. 2015. HyperART: Non-invasive Quantification of Leaf Traits Using Hyperspectral Absorption-reflectance-transmittance Imaging. *Plant Method.,* **11:** 1. doi:10.1186/s13007-015-0043-0.
6. Botha, E. J., Zebarth, B. J. and Lebon, B. 2005. Non-destructive Estimation of Potato Leaf Chlorophyll and Protein Contents from Hyperspectral Measurements Using the PROSPECT Radiative Transfer Model. *Can. J. Plant Sci.,* **86:** 279-291.
7. Carter, G. A., Cibula, W. G. and Miller, R. L. 1996. Narrow-band Reflectance Imagery Compared with Thermal Imagery for Early Detection of Plant Stress. *J. Plant Physiol.,* **148:** 515–522.
8. Chang, J. H. and Root, B. 1975. On the Relationship between Mean Monthly Global Radiation and Air Temperature. *Arch. Meteorol. Geophys. Bioclimatol. B.,* **23:**13–30.
9. Chu, X. L., Yuan, H. F. and Lu, W. Z. 2004. Progress and Application of Spectral Data Pretreatment and Wavelength Selection Methods in NIR Analytical Technique. *Progress in Chemistry*, **16:** 528-542.
10. Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B. and McMurtrey, J. E. 2000. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.,* **74:** 229-239.
11. De Castro, A. I., Jurado-Expósito, M., Gómez-Casero, M. T. and López-Granados, F. 2012. Applying Neural Networks to Hyperspectral and Multispectral Field Data for Discrimination of Cruciferous Weeds in Winter Crops. *Scientific World J.,* 1–11. doi:10.1100/2012/630390.
12. Demirel, K., Genc, L., Bahar, E., Inalpulat, M., Smith, S. and Kizil, U. 2014. Yield Estimate Using Spectral Indices in Eggplant and Bell Pepper Grown under Deficit Irrigation. *Fresenius Environ. Bull.,* **23:** 1232-1237.
13. Dhanoa, M. S., Lister, S. J., Sanderson, R. and Barnes, R. 1994. The Link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR Spectra. *J. Near Infrared Spectrosc.,* **2:** 43–47.
14. Diller, M. 2002. Investigations for the Development of a NIRS-method for Potatoes in Organic Farming with Special Reference to the Influence of the Year and the Potato Pine. PhD Thesis, Rheinische Friedrich-Wilhelms-Universitat, Bonn, Germany. (in German)
15. Furuuchi, H., Jenkins, M. W., Senock, R. S., Houpis, J. L. C. and Pushnik, J. C. 2013. Estimating Plant Crown Transpiration and Water Use Efficiency by Vegetative Reflectance Indices Associated with Chlorophyll Fluorescence. *Open Ecol. J.,* **3:** 122-132.
16. Gamon, J. A. and Surfus, J. S. 1999. Assessing Leaf Pigment Content and Activity with a Reflectometer. *New Phytol.,* **143:** 105–117.
17. Genc, L., Demirel, K., Camoglu, G., Asik, S. and Smith, S. 2011. Determination of Plant Water Stress Using Spectral Reflectance Measurements in Watermelon (*Citrullus vulgaris*). *Amer-Eur. J. Agri. Environ. Sci.,* **11:** 296-304.
18. Gitelson, A. A. 2004. Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *J. Plant Physiol.,* **161:** 165–173.
19. Gitelson, A. A., Keydan, G. P. and Merzlyak, M. N. 2006. Three-band Model for Noninvasive Estimation of Chlorophyll, Carotenoids, and Anthocyanin Contents in

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

Higher Plant Leaves. *J. Geophys. Res. Letter,* **33(L11402):** 5. DOI: 10.1029/2006GL026457.
20. Gitelson, A. A., Vina, A., Ciganda, V., Rundquist, D. C. and Arkebauer, T. J. 2005. Remote Estimation of Canopy cChlorophyll Content in Crops. *J. Geophys. Res. Lett.,* **32:** L08403. http://dx.doi.org/10.1029/2005 GL022688.
21. Guang, L. and Liu, X. N. 2009. Two Kinds of Modified Spectral Indices for Retrieval of Crop Canopy Chlorophyll Content. *Adv. Earth Sci.,* **29:** 548–554.
22. Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J. and Dextraze, L. 2002. Integrated Narrow-band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sens. Environ.*, **81**: 416-426.
23. He, Y., Li, X. and Deng, X. 2007. Discrimination of Varieties of Tea Using Near Infrared Spectroscopy by Principal Component Analysis and BP Model. *J. Food Eng.*, **79:** 1238–1242.
24. Herrman, I., Shapira, U., Kinast, S., Karnielli, A. and Bonfil, D. J. 2013. Ground-level Hyperspectral Imagery for Detecting Weeds in Wheat Fields. *Precis. Agric.*, **14:** 637-659.
25. Hiscox, J. D. and Israelstam, G. F. 1979. A Method for the Extraction of Chlorophyll from Leaf Tissue without Maceration. *Can. J. Bot.,* **57:** 1332-1334.
26. Homayoun, H., Daliri, M. S. and Mehrabi, P. 2011. Effect of Drought Stress on Leaf Chlorophyll in Corn Cultivars (*Zea mays*). *Middle East J. Sci. Res.,* **9:** 418-420.
27. Jin, X., Wang, K., Xiao, C., Diao, W., Wang, F., Chen, B. and Li. S. 2012. Comparison of Two Methods for Estimation of Leaf Total Chlorophyll Content Using Remote Sensing in Wheat. *Field Crop Res.,* **135:** 24-29.
28. Jordan, C. F. 1969. Derivation of Leaf Area Index from Quality of Light on the Forest Floor. *Ecol.,* **50:** 663–666.
29. Kahrıman, F. 2013. Investigation of Yield and Quality Traits in Maize by Source Sink Releationship and Physiological Parameters. Çanakkale Onsekiz Mart University, Çanakkale, Turkey. (in Turkish)
30. Kaufmann, H., Segl, K., Itzerott, S., Bach, H., Wagner, A., Hill, J., Heim, B., Oppermann, K., Heldens, W., Stein, E., Müller, A., Van der Linden, S., Leitão, P. J., Rabe, A. and Hostert, P. 2010. *Hyperspectral Algorithms: Report of the Frame of EnMAP Preparation Activities*. 268 S. Scientific Technical Report, STR 10/08, Deutsches GeoForschungsZentrum GFZ, Potsdam. DOI: 10.2312/GFZ.b103-10089.
31. Liu, Z., Zuo, M. J. and Xu, H. 2012. Parameter Selection for Gaussian Radial Basis Function in Support Vector Machine Classification. In: "*Proceedings of International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*", IEEE, 15–18 June 2012, Piscataway, NJ, Chengdu, China.
32. Kong, W., Zhang, C., Liu, F., Nie, P. and He, Y. 2013. Rice Seed Cultivar Identification Using Near-infrared Hyperspectral Imaging and Multivariate Data Analysis. *Sensor.,* **13:** 8916-8927.
33. Martin, M.P., Barreto, L., Riaño, D., Fernandez-Quintanilla, C. and Vaughan, P. 2011. Assessing the Potential of Hyperspectral Remote Sensing for the Discrimination of Grassweeds in Winter Cereal Crops. *Int. J. Remote Sens.,* **32:** 49-67.
34. Özyiğit, Y. and Bilgen, M., 2013. Use of Spectral Reflectance Values for Determining Nitrogen, Phosphorus, and Potassium Contents of Rangeland Plants. *J. Agr. Sci. Tech.,* **15:** 1537-1545.
35. Penuelas, J., Baret, F. and Filella, I. 1995. Semi-Empirical Indices to Assess Carotenoids/Chlorophyll a Ratio from Leaf Spectral Reflectance. *Photosynthetica,* **31**: 221–230.
36. Penuelas, J., Gamon, J.A., Fredeen, A. L., Merino, J. and Field, C. 1994. Reflectance Indices Associated with Physiological Changes in Nitrogen and Water-limited Sunflower Leaves. *Remote Sens. Environ.,* **48:** 135-146.
37. Pfitzner, K., Bartolo, R., Carr, G., Esparon, A. and Bollhöfer, A. 2011. *Standards for Reflectance Spectral Measurement of Temporal Vegetation Plots.* Supervising Scientist Report 195, Supervising Scientist, Darwin NT.
38. Rundquist, D., Gitelson, A., Leavitt, B., Zygielbaum, A., Perk, R. and Keydan, G. 2014. Elements of an Integrated Phenotyping System for Monitoring Crop Status at Canopy Level. *Agron.,* **4:** 108-123.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

39. Sims, D. A. and Gamon, J. A. 2002. Relationship between Leaf Pigment Content and Spectral Reflectance across a Wide Range Species, Leaf Structures and Development Stages. *Remote Sens. Environ.,* **81:** 337–354.
40. Slaton, M. R., Hunt, E. R. and Smith, W. K. 2001. Estimating Near-infrared Leaf Reflectance from Leaf Structural Characteristics. *Am. J. Bot.,* **88:** 278-284.
41. Strachan, I.B., Pattey, E. and Boisvert, J.B. 2002. Impact of Nitrogen and Environmental Conditions on Corn as Detected by Hyperspectral Reflectance. *Remote Sens. Environ.,* **80:** 213–224.
42. Tucker, C. J. 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.,* **8:** 127-150.
43. Wang, Q. and Li, P. 2012. Hyperspectral Indices for Estimating Leaf Biochemical Properties in Temperate Deciduous Forests: Comparison of Simulated and Measured Reflectance Data Sets. *Ecol. Indic.,* **14:** 56–65.
44. Weber, V. S., Araus, J. L., Cairns, J. E., Sanchez, C., Melchinger, A. E., and Orsini, E. 2012. Prediction of Grain Yield Using Reflectance Spectra of Canopy and Leaves in Maize Plants Grown under Different Water Regimes. *Field Crop Res.,* **128:** 82–90.
45. Xu, Z., Zhou, G., Han, G. and Li, Y. 2011. Photosynthetic Potential and Its Association with Lipid Peroxidation in Response to High Temperature at Different Leaf Ages in Maize. *J. Plant Growth Regul.,* **30:** 41-50.
46. Yi, Q., Huang, J., Wang, F. and Wang, X. 2008. Quantifying Biochemical Variables of Corn by Hyperspectral Reflectance at Leaf Scale. *J. Zhejiang Univ-Sci. B,* **9:** 378-384.

استفاده از مدل های فرا طیفی مبتنی بر برگ برای پایش ترکیبات بیوشیمیایی و شناسایی فنوتیپ گیاه ذرت

ف. قهرمان، ک. دمیرل، م. اینالپولات، ز. ا. اگسل، و ل. جنک

چکیده

هدف این پژوهش تهیه و راستی آزمایی مدل های کمّی و کیفی برای شناسایی ژنوتیپ های مختلف ذرت و همچنین برآورد ترکیبات بیو شیمیایی آنها بود. به این منظور، در سال های ۲۰۱۱ و ۲۰۱۲ در شرایط مزرعه از برگ های مرکزی گیاهانی که به طور تصادفی انتخاب شده بودند داده های طیف سنجی جمع آوری شد. محتوای کلروفیل و پروتئین برگ و پروتئین ساقه در هر گیاه تعیین شد. در ادامه پژوهش، چهار مدل Support Vector Machine (SVM) به دست آمد و راست آزمایی شد . درمدل های کیفی ، تیپ ذرت به عنوان متغییر و.ابسته در نظر گرفته شد در حالی که داده های طیف کامل (FS) در محدوده (400-1000 nm) و شاخص های طیفی (SI) به تعداد ۳۴ شاخص به ازای باندها به عنوان متغیر های مستقل بودند. در دو مدل کمّی (SVMR-FSو SVMR-SI) ، متغیر های مستقل همان های قبلی بودند ولی متغیرهای وابسته صفاتی بودند که به صورت کمّی اندازه گیری شده بودند. نتایج نشان داد که برای دسته بندی به منظور متمایز کردن تیپ های مختلف ذرت از قبیل ژنوتیپ های ذرت دارای روغن زیاد (HOM) ، دارای پروتئین زیاد(HPM)، و روغن استاندار

[DOR: 20.1001.1.16807073.2016.18.6.3.1]

(NORMAL)، مدل های کیفی روشی محکم و مطمئنی بود. همچنین، برای دسته بندی ژنوتیپ های ذرت، مدل SVMC-FS بر مدل SVMC-SI برتری داشت. نیز، مدل های کمّی با داده های طیف سنجی کامل، پیش بینی و برآورد مطمئن تری از دیگر روش ها به دست داد. بهترین نتایج پیش بینی ($\mu g\ g^{-1}$ RMSEC=222.4، R^2 برحسب $\mu g\ g^{-1}$ برای Cal=0.739 و SEP=213.3، RPD=2.04، و r=۰/۸۷۷) از مدل SVMR-FS برای کلروفیل به دست آمد. نتیجه مدل های برآورد غیر مستقیم مبتنی بر روابط بین اندازه گیری های طیف سنجی مبتنی بر برگ و مقدار پروتئین برگ و ساقه کمتر رضایتبخش بود.

[DOR: 20.1001.1.16807073.2016.18.6.3.1]