Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations

Y. Karimi[1]*, S.O. Prasher[1], A. Madani[2] and S. Kim[3]

*[1]Department of Bioresource Engineering, McGill University, Macdonald Campus, 21111 Lakeshore Road, Ste-Anne-de-Bellevue, Quebec H9X 3V9, Canada; [2]Department of Engineering, Nova Scotia University, P.O. Box 550, Truro, Nova Scotia B2N 5E3, Canada; and [3]Department of Environmental Engineering, Yeungnam University, Kyongsan 712-749, South Korea. *Email:yousef.karimizindashty@mail.mcgill.ca*

Karimi, Y., Prasher, S.O., Madani, A. and Kim, S. 2008. **Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations.** Canadian Biosystems Engineering/Le génie des biosystèmes au Canada **50***:* 7.13 - 7.20. This study investigated the potential of support vector machine (SVM) methodology to extract information on crop growth and various biophysical parameters from airborne hyperspectral remote sensing data. The study was carried out in a corn field, consisting of a main plot with four weed control strategies (no weed control, broadleaf control, grass control, and full weed control) and a sub-plot with three nitrogen (N) fertilization rates (60, 120, 250 N kg/ha), all replicated four times. Hyperspectral data were taken in 72 narrow wavebands (409 to 947 nm) using a compact airborne spectrographic imager (CASI) sensor. Various crop physiological parameters were measured concurrently including: leaf greenness (SPAD readings), plant height, leaf nitrogen content, and leaf chlorophyll content. The objective of the study was to evaluate the ability of SVM regression models to extract continuous vegetation variables using aerial hyperspectral observations. Several SVM models were developed to estimate the crop biophysical parameters and the grain yield. The study showed that by using reflectance data collected at the tasseling stage, crop parameters can be estimated with reasonable accuracy. Generally speaking, the coefficients of determination (r^2) were greater than 0.9 for biomass, yield, plant height, and SPAD with the training data set. The r^2 was slightly lower for the test data set (0.51, 0.82, 0.91, and 0.86, respectively), which is acceptable given the small size of the data set used in the study. The results of the five fold cross validation procedure indicated that the SVM results were consistent. The results were also compared with those obtained with a stepwise approach, and the SVM results were found to be superior. **Keywords:** hyperspectral, remote sensing, corn, nitrogen, weeds, crop parameters, support vector machine.

Cette étude chercha à évaluer le potentiel d'une méthode machine exemple support (MES) pour extraire des information sur la croissance d'une culture de maïs et divers paramètres biophysiques à partir de données provenant de aérienne hyperspectrale. L'étude situé dans un champ de maïs, comprena des parcelles principales avec quatre stratégies de contrôle des mauvaises herbes (aucun, contrôle des mauvaises herbes à feuilles larges, contrôle des graminées, et contrôle complet des mauvaises herbes) et des sous parcelles recevant trois taux de fertilisation en azote (60, 120, 250 N kg/ha), le tout répété quatre fois. Les données hyperspectrales furent acquises dans 72 gammes d'ondes étroites (409 à 947 nm) grâce au capteur d'un imageur spectrographique compact aéroporté (CASI). Divers paramètres physiologiques de la culture furent mesurés en même temps, soit l'indice de verdure (mesure SPAD), la taille des plantes, ainsi que la teneur en N et en chlorophylle de leurs feuilles. L'étude tenta d'évaluer l'habilité de modèles de régression MES à extraire de façon continue des variables liées à la végétation à partir d'observations aériennes hyperspectrales. Plusieurs modèles MES furent développés afin d'estimer les paramètres biophysiques de la culture et le rendement en grain. L'étude démontra qu'en utilisant les données de réflectance acquises à la floraison mâle l'on pouvait estimer avec une bonne exactitude les paramètres voulues pour la culture. En général, les cœfficients of détermination (r^2) dépassèrent 0.9 pour la biomasse, le rendement, la taille des plantes et les lectures SPAD avec les données de calibration. Les valeurs de r^2 avec les données de validation furent quelque peu moins élevées (0.51, 0.82, 0.91, et 0.86, respectivement), ce qui, étant donné que la banque de données utilisée fut relativement petite, est acceptable. Les résultats d'une validation croisée répété à vingt reprises indiquèrent que les résultats de la SEM étaient cohérents. Lorsque les résultats de la SEM furent comparés à ceux d'un procédé de régression pas à pas, cette dernière méthode se prouva inférieur à celle de SEM. **Mots clefs:** hyperspectral, télédétection, maïs, azote, mauvaises herbes, paramètres de la culture, machine exemple support.

INTRODUCTION

Producing food in a cost-effective way is the main goal of every farmer and agricultural manager. To get more production, traditionally more agricultural inputs (e.g., pesticides, herbicides, fertilizer) are applied, which consequently have more environmental impacts. In this respect, precision farming with site-specific application of agricultural inputs can lead to a reduction in the application of inputs, without affecting agricultural production (Christensen et al. 1998; Tomer et al. 1997). However, in implementing this strategy, the availability of reliable and relevant spatial information in a timely way is imperative. Hyperspectral observations can prove to be very useful in such cases. However, hyperspectral data sets can be quite complex and very large to analyze and interpret using traditional methods.

Approaches based on artificial intelligence, such as artificial neural networks (ANN) and support vector machines (SVM),

can be used to processes such data. They possess the capability of developing relationships between inputs and outputs using a set of examples, and do not require a priori knowledge of the governing processes (Anderson and Rosenfeld 1988; Wasserman 1989; Zornetzer et al. 1990; Gunn 1998). While ANN models have been used widely in processing remote sensing information (Smith 1993; Pierce et al. 1994; Jin and Liu 1997; Kimes et al. 1997, 1998; Panda and Panigrahi 2000), the use of SVM technology in agriculture has not been fully explored as yet.

The support vector machine algorithm, a supervised learning tool based on the statistical learning theory (Vapnik 1995), is a new and innovative method in the field of artificial intelligence for data mining (Boser et al. 1992; Cortes and Vapnik 1995). Using a training data set, a classification/regression function is setup in SVM. Applying structural risk minimization (SRM) principle, SVM focuses on minimizing a bound on the risk function, rather than minimizing the error in training data (which is usually the case in ANNs). In this way, the over-fitting problem, more common to most ANN models, is prevented. Furthermore, comparing the SVM regression with multiple regression models, although multiple regression models have the ability of selecting specific wavebands, they are more useful when there is a linear relationship between the dependent and independent variables. Also, unlike in SVM, the stepwise regression method has a limitation on the selection of the number of wavebands used in model development. For example, for stepwise regression with the maximum r^2 method, it is suggested that the ratio of the number of wavebands to the total field samples should fall between 0.15 and 0.2 (Thenkabail et al. 2000; Goel et al. 2003). No such restrictions are known to exist for SVM methods. Furthermore, collinearity exists in hyperspectral data and it can lead to problems if the estimation methods depend on the order in which input variables are presented. However, in the case of projection methods, like SVM, where the input data are first projected on to a higher dimensional space before they are employed in the estimation process, the results are not affected by collinearity (Morlini 2006).

Support vector machines employ linear functions for learning. For nonlinear cases, SVMs use a so-called kernel technique to plot the data into a higher dimensional feature space, where linear functions can be applied. The simplicity of the method is one of its main advantages over other data mining techniques, such as ANNs. Thus, only a few parameters need to be adjusted by the users to optimize the model. Support vector machines have emerged in recent times as a popular technique for data mining, such as tissue classification (Furey et al. 2000; Pavlidis et al. 2004), shape extraction and classification (Cai et al. 2001; Du and Sun 2004), protein recognition (Zien et al. 2000), bakery process data (Rousu et al. 2003), hyperspectral data (Gualtieri and Cromp 1998), crop classification (Camps-Valls et al. 2003), and regression problems (Mukherjee et al. 1997; Gunn 1998; Pontil et al. 1998; Sivapragasam et al. 2001; Gao et al. 2003; Bray and Han 2004).

The objective of this study was to examine the applicability of the SVM method in analyzing aerial hyperspectral observations taken over a corn field. More specifically, crop yield, biomass, plant height, and leaf-greenness were predicted. The efficiency of the new technique was evaluated by comparing the results with those obtained by Goel et al. (2003) in a stepwise regression method.

MATERIALS and METHODS

SVM method

This section provides a brief explanation of the basic theory of SVM for regression problems. SVM is based on Vapnik's statistical learning theory (Vapnik 1995).

The theory is explained using a simple problem where the data set has a linear relationship with M observations. Each observation consists of a pair: a vector $x_i \in \boldsymbol{R}^n$; $i = 1, . ., M$ and the corresponding response variable y_i. The final objective of SVM regression is to develop a linear function that can make the best approximation of the dependent response variable. The function can be formulated as:

$$y = f(x) = < w \bullet x > + b \tag{1}$$

where:

w, b = regression parameters, and
<w•x> = dot product of **w** and **x**.

The optimal regression function can be obtained, according to Gunn (1998) and Cristianini and Shawe-Taylor (2000), by minimizing a function, Ψ, as:

Minimize

$$\Psi(w, \lambda) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\left(\lambda_i^- + \lambda_i^+\right) \tag{2}$$

such that

$$\left(w \bullet x_i + b - y_i\right) + \lambda_i^+ \geq \varepsilon$$
$$\left(w \bullet x_i + b - y_i\right) - \lambda_i^- \leq -\varepsilon$$

where:

C = regularization constant, and
λ_i^-, λ_i^+ = slack variables that represent upper and lower constraints on the regression function.

To optimize this function, SVM regression uses a loss function, which shows the maximum allowed deviation of the predicted values from the measured one. Some of the commonly used loss functions are quadratic, Laplace, Huber, and ε-insensitive (Gunn 1998). Among these, the ε-insensitive loss function was proposed by Vapnik (1995) as a robust loss function to reduce sensitivity to the outliers. Therefore, for this study, the ε-insensitive loss function was selected. An SVM regression model based on this function,calculates the difference between the predicted and the actual values, and if the differences are less than the ε, the regression function is considered to be most desirable and accurate (Smola and Scholkopf 1998). Most AI techniques find the best fit between the observed and predicted values, however, SVM's ε-insensitive loss function focuses on optimizing a bound around the regression function, thereby making it more robust against the outliers.

The explicit solution of Eq. 2 is rather difficult. Using Lagrangian multipliers, the solution to this optimization problem can be written as:

Minimize:

$$\varepsilon\sum_{i=1}^{M}\left(\alpha_i + \alpha_i^*\right) + \frac{1}{2}\sum_{i,j=1}^{M}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)\left(x_i \bullet x_j\right)$$

$$+ \sum_{i=1}^{M} y_i\left(\alpha_i - \alpha_i^*\right) \quad (3)$$

subject to:

$$\sum_{i=1}^{M}\left(\alpha_i - \alpha_i^*\right) = 0 \quad and \quad 0 \le \alpha_i, \alpha_i^* \le C$$

where: α_i, α_i^* =Lagrange multipliers.

To handle non-linear regression cases, the data are first linearized by mapping into a higher dimensional space, called "feature space", by incorporating kernel functions, so that linear regression functions can be applied. The commonly-used kernels are the radial basis function (RBF) kernels, sigmoid kernels, and polynomial kernels (Gunn 1998; Chang and Lin 2001). The RBF kernel, most commonly used in SVM approaches, is defined as:

$$K\left(x, y\right) = e^{-y\left(x-y\right)^2} \quad (4)$$

where: γ = a kernel parameter.

Using the kernel function, the optimization function Eq. 3 can be rewritten as:

Minimize:

$$\varepsilon\sum_{i=1}^{M}\left(\alpha_i + \alpha_i^*\right) + \frac{1}{2}\sum_{i,j=1}^{M}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)K\left(x_i, x_j\right) + \sum_{i=1}^{M} y_i\left(\alpha_i - \alpha_i^*\right) \quad (5)$$

subject to:

$$\sum_{i=1}^{M}\left(\alpha_i - \alpha_i^*\right) = 0 \quad and \quad 0 \le \alpha_i, \alpha_i^* \le C$$

The solution of this problem will yield α_i and α_i^* for all i = 1, 2, .., M. It should be mentioned that all the training points within the ε-sensitive zone will yield α_i and α_i^* equal to zero. The type of kernel function to be used is selected by the user. The user also needs to adjust the kernel-specific parameters, the values of parameters γ, C, and ε. The selection of the optimal values of these parameters determines the success of the SVM approach for a given problem. For more detail information, readers are referred to Vapnik (1995), Burges (1998), and Cristianini and Shawe-Taylor (2000).

The SVM regression models were developed using the LIBSVM software (Chang and Lin 2001). The SVM model is formulated using a portion of the data set (e.g., 50%), containing both the dependent and independent variables. The remaining 50% of the data (unseen data) are used for testing the predictive accuracy or performance of the model. In order to find the optimum value of parameters, γ, ε, and C, the model is run with different sets of values of γ, ε, and C, using the training data set. To identify the optimum parameter set, LIBSVM employs an *n*-fold cross-validation technique, where the training data are divided into *n* subsets. The model is trained with *n-1* subsets, and the unseen subset is used for testing. The optimal set is determined on the basis of minimum root mean square error (RMSE) criterion.

Experimental details

A field experimentation was carried out in the summer of 2000 at the Lods Agronomy Research Centre of Macdonald Campus, McGill University, Ste. Anne-de-Bellevue, Québec (45º25´45" N, 73º56´00" W). Corn (hybrid DK389Bty) was planted under different weed control strategies and nitrogen application rates. The experiment was laid out as a split-plot design with weed as the main treatment. The weed treatments were: no weed control (W1), control of grasses (W2), control of broadleaves (W3), and full weed control (W4). The nitrogen fertilizer treatments consisted of low nitrogen (60 N kg/ha, N_{60}), normal nitrogen (120 N kg/ha, N_{120}), and high nitrogen (250 N kg/ha, N_{250}) plots. These combinations represented twelve different growth conditions. Each experimental plot was 20 x 20 m in size. All treatments were replicated four times, thus making 48 plots in total. On the sowing day (May 30), fertilizer was initially applied at 10-20-50 (N-P_2O_5-K_2O) kg/ha with the planter. Subsequently, the rest of the required N fertilizer was broadcasted manually. To establish weed treatments, the herbicides were sprayed on June 26. For more detailed information on cultural practices, applied fertilizers, and herbicides, readers are referred to Goel et al. (2003).

Spectral data acquisition

Hyperspectral images were acquired using a compact airborne spectrographic imager (CASI) in 72 narrow wavebands (408.73 to 947.07 nm) in the visible and NIR region, at a spatial resolution of 2 m. Two observations were made during the growing period, corresponding to two different critical growth stages of the crop: the first at the early growth stage (30 days after planting on June 30) and the second at the tasseling stage (66 days after planting on August 5). Specifications for the CASI sensor and the various radiometric, atmospheric, and geometric correction procedures used to correct the images are given in Goel et al. (2003). Using the ENVI software (ENVI 3.1, Research System, Inc., Boulder, CO), the average reflectance values for each plot on different wavebands were extracted from the imageries. The average spectral response of corn under different nitrogen levels and weed control treatments is illustrated in Fig. 1 for the early growth and tasseling stage observations.

Plant parameters

Various crop canopy and other parameters, including plant height, leaf greenness, leaf area index (LAI), leaf chlorophyll content, leaf nitrogen content, and biomass were measured during the entire growth season. Crop yield and final biomass were also measured at the end of the cropping season. For plant height, ten representative plants were randomly selected and measured in each plot, and then averaged. LAI values (cm^2 foliage area per cm^2 ground area) were recorded using a LAI-2000 Plant Canopy Analyzer (Li-Cor, Inc., Lincoln, NE). SPAD chlorophyll meter (Minolta Camera Ltd., Osaka, Japan) was used to measure the greenness or the amount of chlorophyll in plant leaves, which is the most obvious indicator of plant condition and stress level. Since the chlorophyll molecules contain most of the leaf nitrogen (N), this measure can also be used as an indicator of the N status of the plant. Biomass was estimated on the basis of five to ten plants harvested and weighed from each plot. Crop yield was calculated using ten representative plants from each plot. Detailed information on various measurements made on crop parameters at different flight times can be found in Goel et al. (2003).

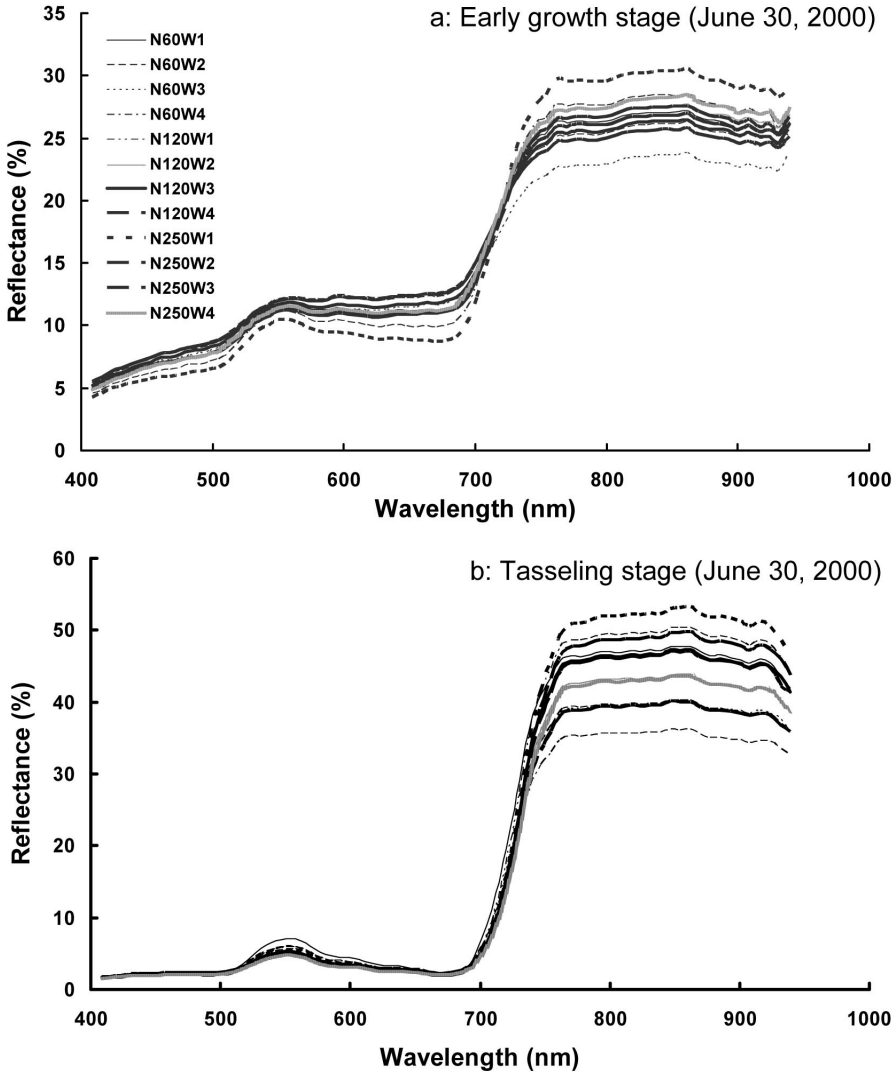

Fig. 1. Measured reflectance response curves of corn at early growth and tasseling stages under different nitrogen application rates and weed control conditions.

Data analysis

A pre-analysis of the reflectance data (Fig. 1) showed that the trend of data for the 72nd waveband was abnormal. While the reflectance values of the other wavebands were generally less than 40%, in the case of the 72nd waveband, the reflectance values appeared to be more than 100% (about 121%). Therefore, the reflectance values for this waveband were considered to be erroneous and too noisy and were eliminated from the analysis.

The spectral data sets for the two flights were analyzed separately. In the beginning, the whole data set for each flight was divided into two subsets of 75 and 25% for training and testing purposes, respectively. Better and consistent results were obtained with the second flight and, therefore, the data from the second flight were analyzed in more detail. Considering the smaller size of this data set (48 observations), the following 5-fold cross-validation approach was employed. The data set was randomly divided into two separate sets for training (80%) and testing (20%), and this process was repeated five times.

For each training data set, a five-fold cross-validation is used internally by the LIBSVM software to determine an optimal SVM parameter set, and the generalization ability and predictive accuracy of the model are determined with the help of unseen test data sets.

The crop biophysical parameters were considered as dependent variables and the spectral responses of the crop (i.e., the reflectance values recorded in various wavebands) were considered as independent variables

By comparing the measured and estimated values of crop parameters, the performance of the developed regression models was evaluated. This was done by regression analyses, where the intercepts and slopes of regression lines were determined and compared with their ideal values of 0 and 1, respectively. Root mean square error (RMSE), relative RMSE (RRMSE), and mean bias error (MBE) were also used for comparison purposes. They were determined using a statistical software, called IRENE (Integrated Resources for Evaluating Numerical Estimates) (Fila et al. 2003). The intercept and slope of the linear regression line is indicative of bias in the data, and RMSE can be used as an indicator of the mean difference between the measured and estimated values (Kobayashi and Salam 2000). Since RMSE values can be affected by the data and the units of measurement, it is more beneficial to use RRMSE (Bellocchi et al. 2002). The RRMSE is calculated by dividing the RMSE by the mean of measured data, and its value can vary from 0 to infinity, with the lower RRMSE values indicating better model performance.

RESULTS and DISCUSSION

The estimated and measured crop parameters (biomass, yield, plant height, SPAD reading, plant nitrogen content, and plant chlorophyll content) were plotted against each other. The calculated regression parameters are presented in Table 1 for both training (75% of the total data set at each growth stage) and testing data sets (25% of the total data set at each growth stage). In enerally, for all crop parameters, better results were obtained for the tasseling stage, as compared to the early growth stage.

For the early growth stage data, the low values of RMSE, RRMSE, and MBE (Table 2), along with a moderate r^2 between the observed and simulated values of plant height and SPAD (Table 1) are generally good indicators of closer agreement between the observed and simulated values. However, the r^2 obtained between the observed and simulated values yield and biomass were very low (Table 1). In most treatments at this

Table 1. Relationship between observed and estimated crop parameters from aerial hyperspectral data using SVM method.

	Training			Testing		
	r^2	Slope	Intercept	r^2	Slope	Intercept
a) Early growth stage (June 30, 2000)						
Biomass	0.579	0.576*	0.554*	0.309	0.609	0.502
Yield	0.589	0.458*	3.337*	0.364	0.321	4.355
Plant height	0.643	0.509*	9.468*	0.836	0.374*	12.148*
SPAD	0.730	0.576*	15.466*	0.672	0.550*	16.774*
Nitrogen	0.272	0.001*	47.67*	0.030	-0.001*	47.750*
Chlorophyll	0.136	0.103*	0.010*	0.009	0.380*	0.010*
b) Tasseling stage (August 5, 2000)						
Biomass	0.933	0.935	0.096	0.508	0.484*	0.652
Yield	0.908	0.862	0.763	0.822	0.802	1.155
Plant height	0.963	0.961	7.42	0.913	0.887	23.720
SPAD	0.874	0.835	6.85*	0.856	0.733	11.610*
Nitrogen	0.802	0.661*	23.859*	0.673	0.578*	31.544*
Chlorophyll	0.282	0.225*	0.012*	0.549	0.307*	0.011*

*Significantly different at 5% level

Table 2. Statistical parameters calculated from measured and estimated corn crop parameters.

		Early growth stage (June 30, 2000)			Tasseling stage (August 5, 2000)		
		RMSE*	RRMSE	MBE	RMSE	RRMSE	MBE
Yield	Train	1.0153	0.1785	0.2536	0.4430	0.0763	-0.0414
	Test	1.2466	0.2000	-0.3218	0.6921	0.1248	0.0579
Biomass	Train	0.1779	0.1435	0.0285	0.0619	0.0487	0.0128
	Test	0.1957	0.1504	-0.0069	0.2168	0.1796	0.0293
Plant height	Train	1.0867	0.0556	-0.1318	5.0702	0.0274	0.2270
	Test	1.4800	0.0731	-0.5339	10.4874	0.0593	3.6688
SPAD	Train	2.1964	0.0605	0.0523	2.2234	0.0505	-0.4988
	Test	2.4246	0.0645	-0.1288	2.3147	0.0538	0.1482
Nitrogen	Train	9.9017	0.2138	1.3976	4.9787	0.0747	1.2765
	Test	7.7173	0.1457	-5.2458	6.9270	0.1095	4.8363
Chlorophyll	Train	0.0023	0.2124	-0.0002	0.0019	0.1291	0.0003
	Test	0.0033	0.2450	-0.0025	0.0022	0.1575	0.0009

*RMSE = root mean square error; RRMSE = relative RMSE; MBE = mean bias error

stage, the crop canopy was small and soil surface was clearly evident which could have led to less desirable values of these parameters. Moreover, both the slope and intercept are differing from their ideal values of 1 and 0 ($P \leq 0.05$), respectively, which implies that the reflectance data, collected at the early growth stage, could not provide substantial information towards the prediction of the above-mentioned crop biophysical parameters. These results are in agreement with those obtained from a stepwise regression method by Goel et al. (2003). They had also attributed poorer results at this growth stage to the smaller size of crop canopy and the interference of soil reflectance. Also, unacceptable values of slope and intercept between the observed and simulated values of plant nitrogen and chlorophyll content imply that no relationship could be established between them and the reflectance data.

At the tasseling stage, the values of RMSE, RRMSE, and MBE (Table 2) were lower than those at the early growth stage, which indicates that the reflectance data can provide more useful information for precision farming. Furthermore, the r^2 between the observed and simulated values of all crop parameters were much higher than those calculated for the early growth stage (Table 1). Both the slope and intercept were not differing from their ideal values ($P \leq 0.05$) for yield and plant height, which indicates excellent model performance. Also, for biomass, the intercept, and for SPAD, the slope, were not different from their ideal values ($P \leq 0.05$). This demonstrates that the reflectance data collected at the tasseling growth stage should be more suitable for the estimation of crop biophysical parameter values. Similar results were also reported by Goel et al. (2003) using stepwise regression. As for the early growth stage period, poorer regression parameter values between the observed and simulated values of chlorophyll content and plant nitrogen indicate no relationship between them and crop reflectance.

To test the generalization ability of the SVM method, a five-fold cross-validation procedure was used for the tasseling stage only as the results of initial analysis of spectral data with SVM were superior for this stage (Tables 1 and 2). Table 3 reports on the r^2, slope, and intercept of the regression lines between the measured and estimated crop parameters for the five-fold cross validation procedure. Results show that, for all five folds, the slope and the intercept of regression lines were in agreement with their ideal values of 1.0 and 0.0 ($P \leq 0.05$); r^2 values were above 0.9 for plant height and yield; for SPAD and biomass data, they varied from 0.85 to 0.94 and 0.67 to 0.91, respectively. This clearly demonstrates the ability of SVM models to estimate crop biophysical parameters.

Table 3. Relationship between observed and estimated crop parameters from aerial hyperspectral data taken on tasseling growth stage (August 5) using SVM method with five fold.

	Training			Testing		
	r^2	Slope	Intercept	r^2	Slope	Intercept
a) First fold						
Biomass	0.746	0.691	0.412	0.705	0.583	0.477
Yield	0.958	0.948	0.319	0.680	0.620	2.397
Plant height	0.930	0.897	19.097	0.868	0.889	22.414
SPAD	0.850	0.818	8.054	0.808	0.813	9.058
b) Second fold						
Biomass	0.776	0.712	0.377	0.419	0.602	0.544
Yield	0.901	0.843	0.870	0.886	0.697	1.796
Plant height	0.966	0.955	7.514	0.955	0.908	20.763
SPAD	0.940	0.923	2.978	0.649	0.945	3.028
c) Third fold						
Biomass	0.807	0.744	0.335	0.563	0.482	0.667
Yield	0.984	0.959	0.249	0.704	0.952	0.289
Plant height	1.000	0.996	0.775	0.712	0.975	4.498
SPAD	0.940	0.923	2.978	0.649	0.945	3.028
d) Fourth fold						
Biomass	0.911	0.848	0.190	0.418	0.662	0.416
Yield	0.939	0.887	0.643	0.815	0.933	0.260
Plant height	0.978	0.974	5.717	0.794	0.962	7.038
SPAD	0.917	0.890	4.600	0.922	1.036	-2.462
e) Fifth fold						
Biomass	0.665	0.582	0.507	0.767	0.666	0.314
Yield	0.911	0.852	0.850	0.840	0.846	0.914
Plant height	0.929	0.912	16.330	0.980	0.938	11.726
SPAD	0.931	0.915	3.584	0.779	0.814	6.499

Table 4. Relationship between observed and estimated crop parameters using regression equations developed by stepwise regression (Geol et al. 2003) with aerial hyperspectral data collected at tasseling time.

	Training			Testing		
	r^2	Slope	Intercept	r^2	Slope	Intercept
Biomass	0.731	0.733	0.340	0.557	0.682	0.489
Yield	0.933	0.931	0.403	0.849	0.766	1.274
Plant height	0.918	0.979	40.830	0.792	0.859	61.657
SPAD	0.932	0.931	3.031	0.569	0.592	19.268

Table 5. Relationship between observed and estimated crop parameters using regression equations developed by SVM method with aerial hyperspectral data collected at tasseling stage (data radndomization was the same used in Table 4).

	Training			Testing		
	r^2	Slope	Intercept	r^2	Slope	Intercept
Biomass	0.847	0.845	0.186	0.692	0.580	0.579
Yield	0.945	0.881	0.569	0.858	0.745	1.319
Plant height	0.846	0.753	43.660	0.872	0.749	42.460
SPAD	0.879	0.839	6.556	0.593	0,536	20.286

Furthermore, to test the predictive ability of the SVM method, the results of SVM models were compared with the results obtained from a stepwise approach (Goel et al. 2003). For both methods, the same data set was used for training and testing. The statistics of the results from both methods are given in Tables 4 and 5, respectively, for the tasseling stage. In almost all cases, the numerical values of statistical parameters obtained with SVM are better, especially for the testing data set. The regression models are generally more useful when there is a linear relationship between the dependent and independent variables. By mapping the data on to a higher dimensional feature space, SVM models appear to be more reliable, especially for the prediction of nonlinear data.

Many other studies have also shown SVM models to perform better than other prediction methods. Li et al. (2007) compared SVM and partial least squares (PLS) methods for the prediction of the protein *N*-glycosylation and reported that higher accuracy was found with the SVM method. In another study, SVM, PLS, and multiple linear (MLR) methods were compared by Liao et al. (2006) and they also found SVM modeling to be superior to the others for the prediction of the logarithm of the partition coefficient between *n*-octanol and water. Kovalenko et al. (2006) applied PLS, SVM, and ANN modeling to near-infrared spectroscopy data in order to determine the amino acid composition of soybean and reported that PLS and SVM results were significantly better than ANN results.

The results obtained in this study are in-line with the above results. SVM modeling produced better results than those obtained with stepwise regression. Therefore, it can be concluded that SVM modeling is another (better?) way of predictive machine-learning based modeling, especially for hyperspectral data analysis. Collinearity in hyperspectral data can lead to problems if the estimation methods depend on the order in which input variables are presented. However, in the case of projection methods, like SVM, where the input data are first projected on to a higher dimensional space before they are employed in the estimation process, such methods are not affected by collinearity (Morlini 2006).

CONCLUSIONS

The capability of the SVM method in analyzing hyperspectral data, collected at the early growth and tasseling stages in a corn field was investigated for the estimation of crop physiological parameters. For the tasseling stage data, the low values of the RMSE, RRMSE, MBE, and the high r^2 values between the estimated and measured crop parameters (biomass, yield, plant height, and SPAD) indicated that the SVM regression models could provide good estimations for such parameters. To test the consistency of the results, a five-fold cross-validation scheme was used, which also confirmed that the SVM results were consistent. The SVM results were compared to those obtained with a stepwise regression method, and the results with SVM were better.

REFERENCES

Anderson, J.A. and E. Rosenfeld (editors). 1988. *Neurocomputing: Foundations of Research*. Cambridge, MA: MIT Press.

Bellocchi, G., M. Acutis, G. Fila and M. Donatelli. 2002. An indicator of solar radiation model performance based on a fuzzy expert system. *Agronomy Journal* 94:1222-1233.

Boser, B.E., I.M. Guyon and V.N. Vapnik. 1992. A training algorithm for optimal margin classifiers. In *Procedings of the Fifth Annual Workshop on Computational Learning Theory*, 144-152. New York, NY: ACM.

Bray, M. and D. Han. 2004. Identification of support vector machines for runoff modeling. *Journal of Hydroinformatics* 6(4): 265-280.

Burges, C.J.C. 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2:121–167.

Cai, Y.-D., X.J. Liu, X. Xu and G.P. Zhou. 2001. Support vector machines for predicting protein structural class. *Bioinformatics* 2:3.

Camps-Valls, G., L. Gómez-Chova, J. Calpe-Maravilla, E. Soria-Olivas, J.D. Martín-Guerrero and J. Moreno. 2003. Support vector machines for crop classification using hyperspectral data. In *Pattern Recognition and Image Analysis*, 134-141. Berlin, Germany: Springer.

Chang, C. and C. Lin. 2001. LIBSVM: A library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm/ (2005/09/20).

Christensen, S., E. Nordbo, T. Heisel and A.M. Walter. 1998. Overview of development in precision weed management, issues and future directions being considered in Europe. In *Precision Weed Management in Crops and Pasture. Proceedings of a Workshop,* eds. R.W. Medd and J.E. Pratley. Adelaide, NSW, Australia: NSW Department of Primary Industries.

Cortes, C. and V.N. Vapnik. 1995. Support vector networks. *Machine Learning* 20: 273-297.

Cristianini, N. and J. Shawe-Taylor. 2000. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. New York, NY: Cambridge University Press.

Du, C.J. and D.W. Sun. 2004. Shape extraction and classification of pizza base using computer vision. *Journal of Food Engineering* 64: 489–496.

Fila, G., G. Bellocchi, M. Acutis and M. Donatelli. 2003. IRENE: A software to evaluate model performance. *European Journal of Agronomy* 18: 369-372.

Furey, T.S., N. Cristianini, N. Duffy, D.W. Bednarski, M. Schummer and D. Haussler. 2000. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16(10): 906-914.

Gao, J.B., S.R. Gunn and C.J. Harris. 2003. SVM regression through variational methods and its sequential implementation. *Neurocomputing* 55: 151– 167.

Goel, P.K., S.O. Prasher, J.A. Landry, R.M. Patel, A.A. Viau and J.R. Miller. 2003. Estimation of crop biophysical parameters through airborne and field hyperspectral remote sensing. *Transactions of the ASAE* 46(4):1235-1246.

Gualtieri, J.A. and R.F. Cromp. 1998. Support vector machines for hyperspectral remote sensing classification. *Proceedings of the SPIE* 3584: 221–232.

Gunn, S. 1998. Support vector machines for classification and regression. Technical report. Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science, University of Southampton, Southampton, England.

Jin, Y.-Q. and C. Liu. 1997. Biomass retrieval from high-dimensional active/passive remote sensing data by using artificial neural networks. *International Journal of Remote Sensing* 18(4): 971-979.

Kimes, D.S., K.J. Ranson and G. Sun. 1997. Inversion of a forest backscatter model using neural networks. *International Journal of Remote Sensing* 18(10): 2181-2199.

Kimes, D.S., R.F. Nelson, M.T. Manry and A.K. Fung. 1998. Review article: Attribute of neural networks for extracting continuous vegetation variables from optical and radar measurements. *International Journal of Remote Sensing* 19(14): 2639-2663.

Kobayashi, K. and M.U. Salam. 2000. Comparing simulated and measured values using mean squared deviation and its components. *Agronomy Journal* 92: 345-352.

Kovalenko, I.V., G.R Rippke and C.R. Hurburgh. 2006. Determination of amino acid composition of soybeans (Glycine max) by near-infrared spectroscopy. *Journal of Agricultural & Food Chemistry* 54(10): 3485-3491.

Liao, Q., J. Yao and S. Yuan. 2006. SVM approach for predicting LogP. *Molecular Diversity* 10(3): 301-309.

Li, S., B. Liu, Y.R. Cai and Y. Li. 2007. Predicting protein N-glycosylation by combining functional domain and secretion information. *Journal of Biomolecular Structure & Dynamics* 25(1) 49-54.

Morlini I. 2006. On multicollinearity and concurvity in some nonlinear multivariate models. *Statistical Methods and Applications* 15: 3–26.

Mukherjee, S., E. Osuna and F. Girosi. 1997. Nonlinear prediction of chaotic time series using support vector machines. In *Proceeding of the 1997 IEEE Workshop*, 511-520.

Panda, S.S. and S. Panigrahi. 2000. Analysis of remotely sensed aerial images for precision farming. ASAE Paper No. 003055. St. Joseph's, MI: ASABE.

Pavlidis, P., I. Wapinski and W.S. Noble. 2004. Support vector machine classification on the web. *Bioinformatics* 20 (4): 586–587.

Pierce, L.E., K. Sarabandi and F.T. Ulaby. 1994. Application of an artificial neural network in canopy scattering inversion. *International Journal of Remote Sensing* 15(16): 3263-3270.

Pontil, M., S. Mukherjee and F. Girosi. 1998. On the noise model of support vector machine regression. In *Proceedings of the 11th International Conference on Algorithmic Learning Theory*, 316-324. London, UK: Springer-Verlag.

Rousu, J., L. Flander, M. Suutarinen, K. Autio, P. Kontkanen and A. Rantanen. 2003. Novel computational tools in bakery process data analysis: A comparative study. *Journal of Food Engineering* 57 (1): 45–56.

Sivapragasam, C.; S.-Y. Liong and MF.K. Pasha. 2001. Rainfall and runoff forecasting with SSA–SVM approach. *Journal of Hydroinformatics* 3: 141-152.

Smith, J.A. 1993. LAI inversion using a back-propagation neural network trained with multiple scattering model. *IEEE Transactions on Geoscience and Remote Sensing* 31(5):1102-1106.

Smola, A.J. and A. Scholkopf. 1998. A tutorial on support vector regression. NeuroCOLT2 Technical Report NC2-TR-1998-030.

Thenkabail, P.S., R.B. Smith and E.D. Pauw. 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing Environ*ment 71(2): 158–182.

Tomer, M.D., J.L. Anderson and J.A. Lamb. 1997. Assessing corn yield and nitrogen uptake variability with digitized aerial infrared photographs. *Photogrammetric Engineering & Remote Sensing* 63(3): 299-306.

Vapnik, V. 1995. *The Nature of Statistical Learning Theory*. New York, NY: Springer-Verlag.

Wasserman, P.D. 1989. *Neural Computing: Theory and Practice.* New York, NY: Van Norstrand Reinhold.

Zien, A., G. Rätsch, S. Mika, B. Schölkopf, T. Lengauer and K.R. Müller. 2000. Engineering support vector machine kernels that recognize translation initiation sites. *Bioinformatics* 16 (9): 799-807.

Zornetzer, S.F., J.L. Davis and C. Lau. 1990. *An Introduction to Neural and Electronic Networks.* New York, NY: Academic Press.