Journal of Remote Sensing & GIS

Research Article   **Open Access**

Retrieving Leaf Area Index from Remotely Sensed Data Using Advanced Statistical Approaches

Gowda PH[1],*, Oommen T[2], Misra D[3], Schwartz RC[4], Howell TA[4] and Wagle P[5]

[1]USDA-ARS Grazinglands Research Laboratory, El Reno, OK 73036, USA
[2]Department of Geological and Mining Engineering and Sciences, Michigan Technological University, 1400 Townsend Drive, Houghton, MI, USA
[3]Department of Mining and Geological Engineering, University of Alaska Fairbanks, P.O. Box 755800, Fairbanks, AK, USA
[4]USDA-ARS Conservation and Production Research Laboratory, P.O. Drawer 10, Bushland, TX, USA
[5]Department of Microbiology and Plant Biology, and Center for Spatial Analysis, University of Oklahoma, Norman, OK 73019, USA

Abstract

Mapping and monitoring leaf area index (LAI) is critical to model surface energy balance, evapotranspiration, and vegetation productivity. Remote sensing helps in rapid collection of LAI on individual fields over large areas, in a time and cost-effective manner using empirical regression between LAI and spectral vegetation indices (SVI). However, these relationships may be ineffective when sun-surface sensor geometry, background reflectance and atmosphere-induced variations on canopy reflectance are larger than variations in the canopy itself. This requires development of superior and region-specific LAI-SVI models. In recent years, statistical learning methods such as support vector machines (SVM) and relevant vector machines (RVM) have been successful over the ordinary least square (OLS) regression models for complex processes. The objective of this study is to develop and compare OLS, SVM, and RVM-based reflectance models to estimate LAI for major summer crops in the Texas High Plains. The LAI was measured in 47 randomly selected commercial fields in Moore and Ochiltree counties. Data collection was made to coincide with Landsat 5 satellite overpasses on the study area. Numerous derivations of SVIs were examined for estimating LAI using OLS, SVM, and RVM models. Analysis of the results indicated that the SVI-LAI models based on the ratio of TM bands 4 and 3, and normalized difference vegetation index (NDVI) are most sensitive to LAI. The R^2 values for selected models varied from 0.79 to 0.96 with the SVM model producing the best results. However, accuracy of reported LAI models needs further evaluation that accounts for in-field spatial variability in the LAI for wider applicability.

Keywords: Semi-arid; Ogallala aquifer region; ET modeling; Texas high plains; Statistical learning method

Introduction

Leaf area index (LAI) is a measure of foliage density that plays a major role in photosynthesis, groundwater-surface water interactions through evapotranspiration (ET) [1], atmospheric gas exchange [2], nutrient uptake [3], and crop productivity [4]. Obviously, it is one of the sensitive input parameters to plant growth, atmospheric circulation [5,6], energy balance [7], terrestrial ecology [8], global climate change [9], and water quality simulation models [10] at field to landscape to global scales. Accurate estimates of LAI are also useful in estimating soil water content from microwave remote sensing data by subtracting the effects of crop water content on reflectance [11].

Traditional *in-situ* techniques to measure LAI involve destructive sampling of leaves and are time intensive. However, numerous indirect in-situ methods have been developed to measure LAI including the scanner method [12], electronic leaf area meter (Delta-T Devices Ltd, Cambridge[1], UK) [13] and LAI-2000, Plant Canopy Analyzer [14]. Although these *in-situ* techniques can be accurate, they are not practical to monitor LAI spatially and temporally over large geographic areas. An alternative approach is to employ satellite remote sensing techniques to estimate LAI. However, the use of satellite data can be a practical and alternative to *in-situ* measurements, provided spectral vegetation indices (SVI) based LAI models are available.

In the last three decades, numerous SVIs have been developed to estimate LAI from Landsat Thematic Mapper (TM) data for corn and soybean [4] and from Enhanced Thematic Mapper Plus (ETM+) data for mixed crops mainly sugarcane [15] and for corn and soybean [16] crops. These studies and others have shown that there is a strong contrasting relationship between spectral reflectance measurements of canopy cover in red and infrared wavelengths. Consequently, a simple ratio (SR), normalized difference vegetation index (NDVI), and soil adjusted vegetation index (SAVI) [17] are some of the commonly used SVIs to estimate LAI. However, these indices are not sensitive at LAI values greater than 3.0 m² m⁻² [11]. The normalized difference water index (NDWI; [11]) uses normalized difference between NIR and shortwave infrared (SWIR) reflectance and the green index (GI; [18]) uses green in place of red reflectance. This green index appears to remain sensitive for LAI values between 3.0 and 6.0 [16].

Most of LAI-SVI statistical relationships reported in the literature is based on ordinary least square (OLS) regressions. In recent years, numerous artificial neural networks (ANN), support vector machines (SVM), and relevant vector machines (RVM) based models have been developed to estimate biophysical characteristics of both agricultural and forest canopies [19]. These models have proved to be superior over the OLS regression models [20]. The remote sensing of LAI often suffers due to its coarse spatial resolution resulting in complex observation with large noise. A key advantage of the advanced statistical algorithms such as SVM and RVM over OLS is the possibility of using a loss function for handling noise in the data together with the ability to obtain a sparse solution to the regression problem. A detailed review on application of

SVM in remote sensing can be found in the literature [21]. Therefore, the applicability of these advanced statistical models for LAI prediction could be valuable in approximating the complex relationship and to deal with the noise in the data.

While satellite remote sensing based SVIs have been used for mapping and monitoring LAI, the strengths and transferability of empirical LAI-SVI relationships to other regions may potentially be affected by exogenous factors such as sun-surface sensor geometry, background reflectance, cultural practices and atmosphere-induced variations on canopy reflectance [22-25]. There have been few tests to compensate for exogenous effects on LAI-SVI relationships and results are mixed. Further, most studies in the past considered one vegetation type. Moreover, comparisons across studies have been limited [22].

Moreover, canopy cover reflectance is a multiple function of many variables that are different over time and space. Therefore, a single SVI-based LAI model for one region may not be a viable option for application to different regions. Overall, spectrally based LAI models for agricultural crops in semi-arid regions are scarce. Further, there are a few OLS-based models that have been developed and validated for the Texas High Plains and but there are no SVM-based LAI models that exist for this region. Development of region-specific LAI models will minimize errors in estimating LAI for use as input in the operational remote sensing-based evapotranspiration mapping programs [26], agronomic [4], ecological [8] and climatic [9]. The main objective of this study is to develop and compare a set of OLS, SVM, and RVM-based reflectance models to estimate LAI for major summer crops in the Texas High Plains.

Materials and Methods

Study area

This study was conducted with LAI data collected from 47 randomly selected irrigated fields (23 in Ochiltree County and 24 in Moore County) in the Texas High Plains underlain by the Ogallala aquifer, which is being depleted by irrigated agriculture. Corn, sorghum, cotton and soybean are the major summer crops in both Ochiltree and Moore counties. Annual average precipitation is about 481 and 562 mm for Moore and Ochiltree counties, respectively. Crop water needs are supplemented with groundwater from the underlying Ogallala aquifer. Most of the croplands in both Moore and Ochiltree counties have nearly level to gently sloping fields with clay loam soils of the Sherm series (fine, mixed, superactive, mesic Torrertic Paleustolls) [27].

Field data collection

Two Landsat TM scenes used in this study were acquired on June 27, 2005 for Ochiltree County and the other on July 4, 2005 for Moore County, for developing LAI models. Development of LAI models consists of three steps: 1) ground-truth (field) data collection, 2) atmospheric correction of Landsat TM imagery for deriving surface reflectance values for ground-truth location, and 3) development of SVI-LAI statistical relationships using OLS and SVM and/or RVM techniques. On the day of the Landsat 5 satellite overpass, ground-truth data were collected from 47 randomly selected commercial fields in the study area. Ground-truth data included geographic coordinates using a handheld Global Positioning System (GARMIN GPSMAP 76, Garmin Ltd), plant type and density, width of plant rows and extraction of one representative plant for LAI measurement in the laboratory. The LAI was measured using the electronic leaf area meter (Delta-T Devices Ltd, Cambridge, UK; [13]).

Each pixel in the Landsat 5 image has a spatial resolution of 30 m after resampling. Ground truth pixel locations on the image were identified using the GPS coordinates. For model development, mean reflectance data from 9-pixels (ground-truth pixel and surrounding 8 pixels) were used, as it was difficult to precisely identify the ground truth location in the field, without proper reflectors installed to validate the GPS coordinates with those of the satellite pixels. Atmospherically corrected surface reflectance (ρ_{SUR}) for each TM band was used to develop SVIs. As a first step to compute ρ_{SUR}, the digital numbers in each TM band image except thermal band 6 were converted into spectral radiance (L_b), using the equation: L_b=Gain x DN+Bias, where Gain and Bias were extracted from the image header files from the satellite data provider. The top-of-the-atmosphere reflectance (ρ_{TOA}) was calculated for each pixel in the image using the procedure outlined in [28]. In this procedure, the ρ_{TOA} for each pixel was calculated by dividing spectral radiance by the incoming energy (radiance) in the same short-wave band. The incoming irradiance is a function of mean solar exo-atmospheric irradiance, solar incidence angle, and square of the relative earth-to-sun distance. The ρ_{SUR} was computed after applying atmospheric interference corrections for short-wave absorption and scattering using narrow band transmittance calibrated for each band with MODTRAN Version 4, a radioactive transfer model [29].

The LAI-SVI relationships were evaluated using OLS, SVM, and RVM techniques with measured LAI as the independent variable. A set of SVIs evaluated to develop LAI-SVI statistical relationships included difference indices, sum indices, product indices, ratio indices, and normalized difference indices including NDWI [11]. In addition, LAI-SAVI was evaluated with L value varied from .05 to 1 at 0.05 intervals. The OLS models used to evaluate each of the SVIs were linear, exponential, power, and quadratic. Finally, the best model in each category was identified and reported for the study area.

Regression using support vector machine (SVM)

Regression using SVM is referred to as the Support Vector Regression (SVR). Here we provide a brief overview of the theoretical concepts of SVR. A detailed depiction of SVR is beyond the scope of this paper and can be obtained from [30-33]. SVM or RVM involves three steps in the regression processes. These are: (1) a random subset of the entire dataset is used as training data to develop the model. The model is developed, just like any other regression model, with some dependent variables that predict an independent variable such as the LAI; (2) a hyper-weighted convolution process where a kernel function is used weigh the data to reside within a specified goal (ε) of a regression boundary (see e.g., [30]) and a bias function is added to the data that falls outside of this goal in order to account for the deviation from the specified goal; and (3) once the training model is developed, it is applied to the rest of the data to test and assess how well the model confers or justifies the broader information.

In machine learning, the data used for developing the model is referred to as the training data. Suppose we have training data $(x_{11}, x_{12}, \ldots . x_{1n}, y_1), \ldots \ldots, (x_{l1}, x_{l2}, \ldots . x_{ln}, y_l) \subset X \times R$, where $(x_{l1} \ldots \ldots x_h)$ represent the predictor variables and y_l represents observed LAI. The goal in SVR is to find a function $f(x)$ that has the most ε deviation from the observed LAI y_l, for all the training data, and at the same time, is as flat as possible. Mathematically,

$$f(x) = \langle w, x \rangle + b \text{ with } w \in X, b \in R \qquad (1)$$

where $\langle w, x \rangle$ denotes the dot product in X. Flatness in equation 1 means a small value of w, and it can be obtained by minimizing the

Euclidean norm, i.e., $\|w\|^2$. Thus, the SVR problem can be formulated as shown in equation 2.

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 \qquad (2)$$
$$\text{subject to} \quad \begin{cases} (\langle w, x_i \rangle + b) - y_i \le \varepsilon \\ y_i - (\langle w, x_i \rangle + b) \le \varepsilon \end{cases}$$

However, in some cases having a function $f(x)$ that is flat with errors less than ε is not feasible. To deal with these infeasible situations a constant C and slack variables ξ_i^-, ξ_i^+ are introduced which leads to the formulation (equation 3) as stated in [30].

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i^- + \xi_i^+) \qquad (3)$$
$$\text{subject to} \quad \begin{cases} (\langle w, x_i \rangle + b) \le \varepsilon + \xi_i^- \\ y_i - (\langle w, x_i \rangle + b) \le \varepsilon + \xi_i^+ \\ \xi_i^-, \xi_i^+ \ge 0 \end{cases}$$

where C is the pre-specified term that controls the magnitude of penalty associated with errors outside the error margin, and ξ_i^-, ξ_i^+ are slack variables representing upper and lower constraints on the output system. The constant C>0 determines the tradeoff between the flatness of the function and the amount to which deviations larger than ε are tolerated [34].

Lagrange multipliers that employ the Kharush-Kuhn-Tucker (KKT) method are then used to solve the optimization problem in equation 3. The KKT method converts the inequality constraint into an equation of the form h(x)=0 by adding or subtracting slack variables and then solving the corresponding equality constrained quadratic optimization problem. Solution of the optimization problem results in a dual pair variable Lagrangian $L_d(\alpha_i, \alpha_i^*)$, one for each of the training patterns. The pairs that result in non-zero α_i or α_i^* are termed as the support vectors. During the training phase of the SVR model the support vectors are used to define the hyper-plane (regression line) and to constrain the width of the optimal margin. Any training data that is outside the optimal margin, does not contribute to the definition of the regression line.

Often in complex nonlinear problems the original input space (predictor variable) is non-linearly related to the predicted variable (LAI). This limits a linear formulation of the problem as shown in equation 3. In SVR, this limitation is overcome by mapping the input space onto some higher dimensional space (feature space) using a nonlinear mapping function (kernel function). Commonly used kernel functions include the linear, polynomial, Gaussian radial basis, and sigmoid kernel functions. It was demonstrated that a linear kernel is a special case of the Gaussian radial basis kernel and that the sigmoid kernel behaves like a Gaussian radial basis kernel for certain parameters [35]. Therefore, in this study we use a Gaussian radial basis kernel function (equation 4).

$$K(x, x) = \exp\left(-\frac{\|x - x\|}{2\gamma^2}\right) \qquad (4)$$

In any SVR problem, when we use the Gaussian radial basis function as the kernel function, we have three parameters to optimize during training: they are the Gaussian radial basis function parameter γ, magnitude of penalty term C, and the width/deviation of the error margin ε. To avoid overfitting, the SVR model parameters were optimized using k-fold cross validation [36]. In k-fold cross validation, for each of the k splits, we use (k-1) folds for training and the remaining one fold for testing. The advantage of k-fold cross validation is that

all the examples in the data set are eventually used for both training and testing. Finally, the optimal model parameter is determined by evaluating which model has the best generalizability. It was demonstrated that a k=-10 or k=-5 fold cross validation is appropriate to evaluate non-linear models [36]. Therefore, in this study, the SVR model was optimized using a -5 fold cross-validation.

Regression using relevance vector machine (RVM)

RVM, introduced by [33], provides a regression method in a Bayesian framework. RVM is a special case of a sparse linear model, where the kernel function Ö centered at the different training points forms the basis function given by:

$$y(x) = \sum_{i=1}^{N} w_i \Phi_i(x) \quad (5)$$

where the output is a linearly-weighted sum of N, generally nonlinear and fixed, basis functions $Ö(x) = (Ö_1(x)........Ö_N(x))^T$ [33]. For the RVM learning, we use a set of input vectors $\{x_i\}_{i=1}^{N}$ along with corresponding set of targets $\{t_i\}_{i=1}$, which in this case is a real value (or LAI). The RVM utilizes the kernel function given in eq. 5 to learn the relationship between the input vector and the targets. By assuming the targets are independent, the likelihood of the dataset can be written as

$$p(t | w, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\{-\frac{1}{2\sigma^2}\|t - \Phi w\|^2\} \qquad (6)$$

where t, w, and Ö is the N × (N+1) training matrix. Further, the solution of Eq. 6 is constrained to avoid over-fitting to the noise by defining an explicit prior probability distribution over them.

A popular choice of prior probability distribution for the RVM is the zero-mean Gaussian prior distribution with a distinct hyper-parameter, α_i, for each weight,

$$p(w | \alpha) = \Pi_{i=0}^{N} N(w_i | 0, \alpha_i^{-1}) \qquad (7)$$

The optimal parameters of the RVM model are then derived by minimizing the penalized negative log-likelihood,

$$\log\{P(t/w)p(w/\alpha)\} = \sum_{i=1}^{N}\left[t_i \log y_i + (1-t_i)\log(1-y_i)\right] - \frac{1}{2}w^T A w \quad (8)$$

where $y_i = ó\{y(x_i; w)\}$, and $A = diag(á)$ is a diagonal matrix with non-zero elements given by the vector of hyper-parameters. The outcome of this optimization is that many elements of α go to infinity such that will have only a few nonzero weights that will be considered as relevant vector. Similar to SVM, the RVM model was also optimized using a -5 fold cross-validation.

Results and Discussion

Ground truth LAI measurements were made from 16 commercial fields planted with corn, 12 with cotton, 11 with sorghum, and 8 with soybean. Measured LAI values varied from 0.09 to 6.21 with a standard deviation of 1.91 m² m⁻². Table 1 presents range, mean, and standard deviation of measured LAIs for major summer crops in the study area. The large variability in the measured LAI could be attributed to various stages of crop growth in different fields. Table 2 presents a best LAI-SVI model in each of the linear, quadratic, exponential, and power, SVM and RVM categories and associated R^2, adjusted R^2, root mean square error (RMSE), and F statistic for major summer crops in the Texas High Plains. The optimal R^2 values with SVM and RVM modes were obtained using the NDVI43 as input.

Figure 1a-d illustrates the best statistically significant linear, quadratic, exponential, and power relationships between LAI and SVIs (see Table 2 for the models used). All four SVIs contain TM band 4 (NIR)

Crop Type	Number of Fields	Measured LAI (m^2 m^{-2})		
		Range	Mean	Standard Deviation
Corn	16	0.45–6.21	4.08	1.3
Cotton	12	0.1–0.83	0.42	0.22
Sorghum	11	0.09–4.58	1.56	1.37
Soybean	8	0.56–2.3	1.16	0.59

**Table 1:** Measured leaf area index (LAI) in 47 commercial fields in the Texas High Plains.

	Model[1]	R^2	Adj. R^2	RMSE	F
Linear					
	LAI=-0.306 + 0.811 * R47	0.79		0.90	160.77
Quadratic					
	LAI=-1.195 + 0.775 * R43 -0.023 * $(R43)^2$	0.80	0.79	0.89	83.12
Exponential					
	LAI=0.029 $e^{6.137 \, ^* \, NDVI43}$	0.81		0.55	182.41
Power					
	LAI=8.768 * $(NDVI43)^{3.616}$	0.84		0.50	228.10
Support Vector Machine (Optimal model parameters)					
kernel=radial basis (C=8.08, γ=212.94, ε=0.08)		0.96		0.42	
Relevant Vector Machine (Optimal model parameter)					
kernel=radial basis (γ=538.48)		0.88		0.65	

[1]R43, and R47=Ratio of bands 4 and 3, and 4 and 7, respectively; NDVI43=normalized difference between bands 4 and 3.
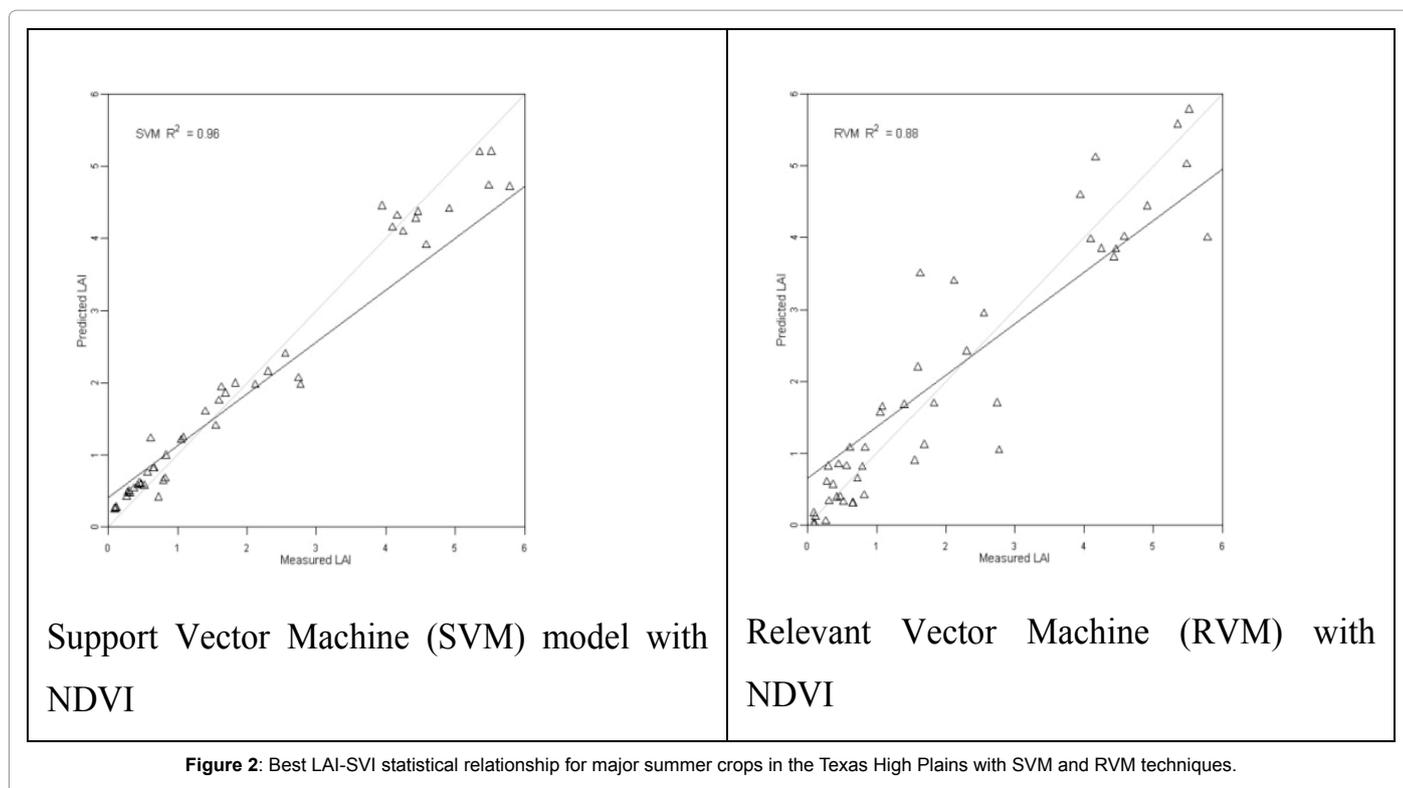
Table 2: Landsat 5 Thematic Mapper (TM) based leaf area index (LAI) models for major summer crops in the Texas High Plains.

as one of the variables indicating that the NIR is more sensitive to LAI. Their R^2 values varied from 0.79 to 0.84. Among the four OLS models, the model that used NDVI (NDVI43) i.e., normalized difference of TM bands 4 and 3 was found to be relatively more sensitive to lower LAI values than that with ratio indices. The LAI-SVI relationship that uses the ratio of TM bands 4 and 7 (R47) gave the highest R^2 value (0.79) with relatively lower RMSE (0.90) and high F statistic (160.77) in the linear category (Figure 1a). Figure 1b shows the best statistically significant quadratic relationship between LAI and SVIs. The relationship pattern was similar to linear model; however, the performance statistics were slightly improved with relatively high R^2 values (0.80) and low RMSE values (0.89). Figure 1c-d shows the best exponential and power relationships between LAI and SVIs, respectively. Both relationships use NDVI (NDVI43) as a variable and gave best performance statistics with R^2 value of 0.81 for exponential and 0.84 for power relationships. The RMSE values (0.50-0.55) were much lower than that for linear and quadratic LAI-SVI relationships. The highest LAI estimated from this study was 6.21 m^2. However, saturation in the LAI-SVI relationships was not observed. This is consistent with the observation made by [15] for mixed land use data.

The LAI-SAVI relationships (not reported here) were not better than the LAI-NDVI relationship in linear, quadratic, exponential or power forms of regression models, and LAI values were not sensitive to L value. However, it provided relatively comparable results with R^2 value of 0.76 for linear, 0.78 for quadratic, 0.80 for exponential, and 0.82 for power models. These results are consistent with the L value of 0.16
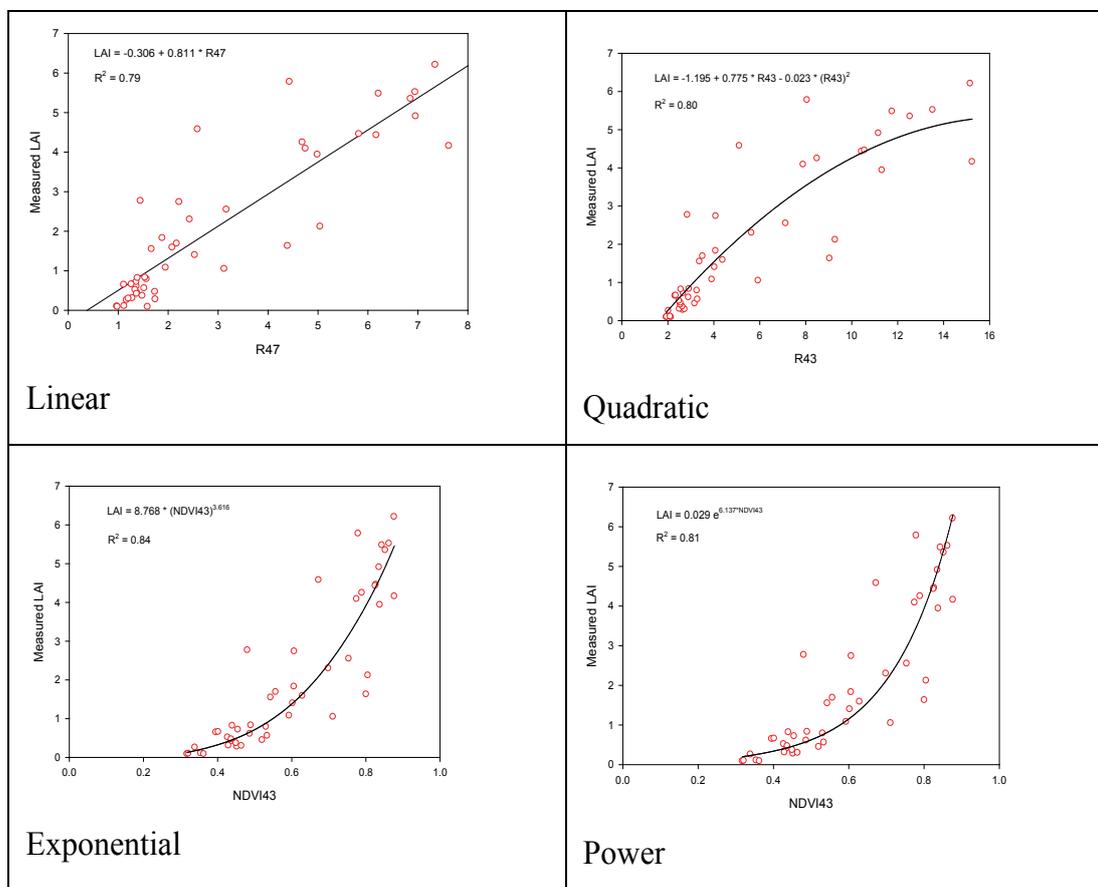
Figure 1: Best LAI-SVI OLS regression relationship in (a) Linear, (b) quadratic, (c) exponential, and (d) power categories.

Support Vector Machine (SVM) model with NDVI

Relevant Vector Machine (RVM) with NDVI

Figure 2: Best LAI-SVI statistical relationship for major summer crops in the Texas High Plains with SVM and RVM techniques.

reported in [37]. The reduced sensitivity of L may be due to the fact that there was not much variation in the soil background as clay loam soils of the Sherm series occupy nearly all of the cropland in both Moore and Ochiltree counties.

Figure 2a-b illustrates the best LAI-SVI relationships with SVM and RVM techniques, respectively. The best models were associated with the NDVI (NDVI43) in both categories. The LAI model with SVM was significantly better than OLS and RVM models with an R^2 value of 0.96 and a RMSE value of 0.42. The RVM model was slightly comparable to the exponential model with a slightly better R^2 (0.88) and a slightly poor RMSE (0.65).

Five of the six reported OLS models (Table 2) use either SR (R43=TM band 4/TM band 3) or NDVI. The remaining linear model use the ratio of TM band 3 (R) or 7 (MIR). This indicates that the R and NIR bandwidths are sensitive to the LAI and is consistent with those observed by previous studies [4,16,17]. The difference in reflectance of R and NIR from plants is attributed to the chlorophyll pigments in plant leaf absorbing energy in the red band of the electromagnetic spectrum resulting in relatively low red reflectance and transmittance while lignin in plant cell walls causing scattering of near-infrared energy resulting in relatively high near-infrared reflectance and transmittance [38]. Based upon the RMSE and F statistic, exponential and power models (Table 2) were found to the best for estimating LAI. However, based on the normalized difference (NDVI) between the TM bands 4 and 3, SVM and RVM were found to be the best models for estimating LAI. The power model that uses NDVI43 accounted for 84% of the variability in the measured LAI data with a relatively small RMSE of 0.50.

Conclusion

Leaf Area Index (LAI) is important for spatially distributed modeling of surface energy balance, evapotranspiration and vegetation productivity. The Landsat 5 Thematic Mapper (TM)-based spectral vegetation indices (SVIs) were evaluated using the ordinary least square (OLS), support vector machine (SVM), and Relevant Vector Machine (RVM) regression techniques for their ability to estimate LAI for major crops in the Texas High Plains. The R^2 values for the selected models varied from 0.79 to 0.96 with the SVM model providing the least RMSE, followed by the power model. Both the SVM and the power model used NDVI as input variable and resulted in producing the best regression with field-measured LAI. Analysis of the results indicated that SR and NDVI were sensitive to LAI. Overall, the remote sensing approach is promising for the rapid collection of LAI information on individual fields over large areas in the Northern High Plains of Texas in a time and cost-effective manner. These are preliminary modeling analysis and we are pursuing further evaluation of the models to establish the utility of either the OLS or the machine learning algorithm such as SVM and RVM.

References

1. Sumner DM, Jacobs JM (2005) Utility of Penman–Monteith, reference evapotranspiration, and pan evaporation methods to estimate pasture evapotranspiration. Journal of Hydrology 308: 81-104.

2. Burkart S, Manderscheid R, Weigel HJ (2004) Interactive effects of elevated atmospheric CO_2 concentrations and plant available soil water content on canopy evapotranspiration and conductance of spring wheat. European journal of agronomy 21: 401-417.

3. Leonard RA, Knisel W, Still D (1986) GLEAMS: Groundwater loading effects of agricultural management systems. ASAE.

4. Thenkabail PS, Ward AD, Lyon JG, Merry CJ (1994) Thematic Mapper

vegetation indices for determining soybean and corn growth parameters. Photogrammetric Engineering and Remote Sensing 60: 437-442.

5. Sellers P, Dickinson R, Randall D, Betts A, Hall F, et al. (1997) Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. Science 275: 502-509.

6. Ruimy A, Saugier B, Dedieu G (1994) Methodology for the estimation of terrestrial net primary production from remotely sensed data. Journal of Geophysical Research: Atmospheres 99: 5263-5283.

7. Gowda PH, Chavez JL, Colaizzi PD, Evett SR, Howell TA, et al. (2008) ET mapping for agricultural water management: present status and challenges. Irrigation science 26: 223-237.

8. Parton WJ, Haxeltine A, Thornton P, Anne R, Hartman M (1996) Ecosystem sensitivity to land-surface models and leaf area index. Global and Planetary Change 13: 89-98.

9. Chase TN, Pielke RA, Kittel TG, Nemani R, Running SW (1996) Sensitivity of a general circulation model to global changes in leaf area index. Journal of Geophysical Research: Atmospheres (1984–2012) 101: 7393-7408.

10. Gowda PH, Ward AD, White DA, Baker DB, Lyon JG (1999) Using field scale models to predict peak flows on agricultural watersheds. JAWRA Journal of the American Water Resources Association 35: 1223-1232.

11. Anderson M, Neale C, Li F, Norman J, Kustas W, et al. (2004) Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. Remote sensing of environment 92: 447-464.

12. Yang J, Chen Q, Han X, Li X, Lie H (2002) Measurement of vegetable leaf area using digital image processing techniques. Trans. CASE 18: 155-158.

13. Rouphael Y, Colla G (2005) Radiation and water use efficiencies of greenhouse zucchini squash in relation to different climate parameters. European Journal of Agronomy 23: 183-194.

14. Welles JM, Norman J (1991) Instrument for indirect measurement of canopy architecture. Agronomy journal 83: 818-825.

15. Xavier AC, Vettorazzi C (2004) Mapping leaf area index through spectral vegetation indices in a subtropical watershed. International Journal of Remote Sensing 25: 1661-1672.

16. Walthall C, Dulaney W, Anderson M, Norman J, Fang H, et al. (2004) A comparison of empirical and neural network approaches for estimating corn and soybean leaf area index from Landsat ETM+imagery. Remote Sensing of Environment 92: 465-474.

17. Huete AR (1988) A soil-adjusted vegetation index (SAVI). Remote sensing of environment 25: 295-309.

18. Gitelson AA, Viña A, Arkebauer TJ, Rundquist DC, Keydan G, et al. (2003) Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophysical Research Letters 30.

19. Jensen RR, Binford M (2004) Measurement and comparison of Leaf Area Index estimators derived from satellite remote sensing techniques. International Journal of Remote Sensing 25: 4251-4265.

20. Samui P, Gowda PH, Oommen T, Howell TA, Marek TH, et al. (2012) Statistical learning algorithms for identifying contrasting tillage practices with Landsat Thematic Mapper data. International journal of remote sensing 33: 5732-5745.

21. Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66: 247-259.

22. Turner DP, Cohen WB, Kennedy RE, Fassnacht KS, Briggs JM (1999) Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. Remote sensing of environment 70: 52-68.

23. Chen JM, Cihlar J (1996) Retrieving leaf area index of boreal conifer forests using Landsat TM images. Remote sensing of Environment 55: 153-162.

24. Jacquemoud S, Baret F, Andrieu B, Danson F, Jaggard K (1995) Extraction of vegetation biophysical parameters by inversion of the PROSPECT+SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. Remote sensing of environment 52: 163-172.

25. Meyer P, Itten KI, Kellenberger T, Sandmeier S, Sandmeier R (1993) Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment. ISPRS Journal of Photogrammetry and Remote Sensing 48: 17-28.

26. Chavez J, Gowda P, Howell T, Copeland K (2009) Radiometric surface temperature calibration effects on satellite based evapotranspiration estimation. International journal of Remote sensing 30: 2337-2354.

27. Unger PW, Pringle FB (1986) Sherm Soils-Distribution, Importance, Variability, and Management.

28. Chander G, Markham B (2003) Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. Geoscience and Remote Sensing, IEEE Transactions 41: 2674-2677.

29. Berk G, Anderson P, Acharya M, Hoke J, Chetwynd L, Bernstein, et al. (2003) Modtran 4 v3 Rev 1 User's Manual.

30. Vapnik V (2013) The nature of statistical learning theory. Springer Science & Business Media.

31. Kecman V (2001) Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. MIT press.

32. Hastie T, Tibshirani R, Friedman J, Franklin J (2005) The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer 27: 83-85.

33. Tipping M (2000) The Relevance Vector Machine. Advances in Neural Information Proceeding Systems, 12, in, MIT Press, Cambridge, MA, pp. 652À658.

34. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Statistics and computing 14: 199-222.

35. Keerthi SS, Lin CJ (2003) Asymptotic behaviors of support vector machines with Gaussian kernel. Neural computation 15: 1667-1689.

36. Breiman L, Spector P (1992) Submodel selection and evaluation in regression. The X-random case. International statistical review/revue internationale de Statistique 291-319.

37. Rondeaux G, Steven M, Baret F (1996) Optimization of soil-adjusted vegetation indices. Remote sensing of environment 55: 95-107.

38. Gates DM, Keegan HJ, Schleter JC, Weidner VR (1965) Spectral properties of plants. Applied optics 4: 11-20.