Estimating leaf area index of salt marsh vegetation using airborne hyperspectral data

Asim Banskota
February, 2006
Estimating leaf area index of salt marsh vegetation using airborne hyperspectral data

by

Asim Banskota

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: (Environment System Analysis and Management)

Thesis Assessment Board

Dr. Jan de Leeuw, Chairman
Associate Professor, Department of Natural Resources
ITC, Enschede, The Netherlands

Prof. Dr. Freek D. van der Meer, External Examiner
Professor of Imaging Spectrometry
ITC, Enschede and Delft University of Technology, The Netherlands

Dr. Mark van der Meijde, Internal Examiner
Assistant Professor, Department of Earth System Analysis
ITC, Enschede, The Netherlands

Prof. Dr. Andrew K. Skidmore, Supervisor
Professor of Land Cover Mapping and Monitoring
ITC, Enschede and Wageningen University, The Netherlands
Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.
Abstract

Information on Leaf area index (LAI) can be important for an improved understanding of the ecological processes of the salt marsh ecosystem. It is very difficult to estimate LAI given the heterogeneity and large spatial gradients within salt marshes. Hyperspectral remote sensing has demonstrated wide applicability in the area of estimating and mapping LAI of forest and agricultural ecosystems. This study attempts to evaluate different empirical models in estimating the LAI of salt marsh area from a hyperspectral remote sensing data.

The empirical models were derived from three different regression techniques: simple linear regression, partial least square regression and neural network. Narrow band optimal normalized difference vegetation index (optimal NDVI) and optimal modified soil adjusted vegetation index (optimal MSAVI) derived from all possible two-band combinations were used in simple linear regression. The full spectrum data (455nm-1622nm) and the useful channels selected by a genetic algorithm were used in partial least square regressions. Feature extractions for neural networks were carried out using principle component analysis and minimum noise fraction. Cross validation procedure was used to assess the prediction power of the regression models. Analyses were performed on the entire data set (subset A, n=78) or on subsets stratified as high marsh (subset H, n=45), low marsh (subset L, n=33) and elymus vegetation type (subset E, n=15). Partial least square regression coupled with genetic algorithm provided the highest accuracy in predicting the LAI of the study area. Neural networks performed inconsistently with different subsets. Neural network provided higher accuracy than simple linear regression when entire dataset (n=78) was used in analysis. For subsets, H, L and E, simple linear regressions provided higher accuracies than neural networks.

The study showed that the partial least square regression can be a useful tool for the estimation of LAI of the study area using hyperspectral data. Partial least square regression can utilize a greater number of channels of hyperspectral data than simple linear regression; whereas it was found to work well with minimum data size compared to neural network. The highest accuracy obtained for subset E emphasized the need of building vegetation type specific model to estimate LAI of the study area with higher accuracy.
Acknowledgements

I would like to extend my sincere appreciation to some special people, who supported me with their valuable knowledge, time and resource to make this research possible.

I am very grateful to Professor Andrew Skidmore, my supervisor, for teaching me how to conduct a research in a scientific framework. I received plenty of valuable comments and words of encouragement from him. I was never hesitant to discuss with him openly mainly because of his very ‘accommodating’ and ‘cheerful’ outlook.

I am also grateful to Dr. Iris C. van Duren for being supportive and providing valuable comments. I would also like to thank my previous supervisor, Dr. Clement Atzberger, who really helped me in engineering the concept of this research.

My sincere gratitude goes to Dr. Michael Weir, my Programme Director, who has been very kind and supportive throughout my research period. I must say, he is a very generous person and a wonderful Programme Director.

I would like to extend earnest appreciation to Ms. Roos Veeneklaas, PhD student from University of Groningen, who provided me a ‘crash course’ in the identification of salt marsh vegetation species. It was unbelievable how she acquainted me with more than 20 dominant species within less than three hours of time. Besides, she also helped me in arranging accommodation and materials during the field work. I would also like to thank University of Groningen in this regards to allow me staying in their field station during my entire field work period. I can not forget the contribution of Mr. Valentijn Venus and my colleague Mr. Habtamu Mulatu for their considerable support during field data collection. Thanks to Ms. Roshanak Darvishzadeh for providing orientation with field instruments.

Dr. David Rossitter was another special person, who encouraged me with his few inspiring words during the MSc day of ITC. I am indebted to you David; I really got extra gear from there on to conclude my thesis.

I am obliged to Mr. Chudamani Joshi, PhD student in ITC, for advising me and helping me in finding some useful literatures related to my work. I would also like to thank Dr. Jan de leew, Dr Patrick van Larker, Mr. Boudewijn van Leewuen, Mr. Jelger Kooistra, Mr. Henk van Oosten, Mr. Istiyak Sobhan, for giving their valuable advice during my research.

I extend my sincere gratitude to Netherlands government and ITC for providing me opportunity to pursue my Master study. I would also take this opportunity to thank my ‘special friends’ who tried their best to provide me a pleasant environment here far away from my home.

With gratitude I remember my family and friends who contributed to my overall well being.
# Table of contents

1. Introduction ......................................................................................................................... 1  
   1.1. Background ..................................................................................................................... 1  
   1.2. Research Objective......................................................................................................... 3  
   1.3. Specific Objectives.......................................................................................................... 3  
   1.4. Research Questions ....................................................................................................... 3  
2. Methods ............................................................................................................................... 4  
   2.1. Study area....................................................................................................................... 4  
   2.2. Data collection ............................................................................................................... 4  
   2.3. Pre-processing ............................................................................................................... 5  
   2.4. Modeling ...................................................................................................................... 8  
3. Results ............................................................................................................................... 12  
   3.1. Band combinations for narrow band vegetation indices .............................................. 12  
   3.2. Optimal narrow band vegetation index relationship with LAI .................................... 13  
   3.3. Partial least square regression .................................................................................... 14  
   3.4. Models selection using neural networks ...................................................................... 15  
   3.5. Accuracies associated with different models ............................................................. 17  
4. Discussion ......................................................................................................................... 19  
   4.1. Comparison of the regression techniques in predicting LAI ....................................... 19  
   4.2. Comparison of optimal narrow band indices .............................................................. 20  
   4.3. Performance of feature extraction techniques for neural network............................ 21  
   4.4. Band combinations for optimal vegetation indices ..................................................... 21  
   4.5. Comparison of subsets ............................................................................................... 22  
   4.6. Variation of LAI at plot level ...................................................................................... 22  
5. Conclusion ......................................................................................................................... 23  
6. Recommendation .............................................................................................................. 24  
   References ........................................................................................................................ 25  
   Appendix ........................................................................................................................... 29
List of figures

Figure 1: Principal component analysis .........................................................8
Figure 2: Minimum noise fraction .................................................................9
Figure 3: 2D correlation plot for vegetation indices ......................................13
Figure 4: Fitness plot from genetic algorithm ..............................................16
Figure 5: Model selection of neural network (1) .........................................17
Figure 6: Model selection of neural network (2) .........................................17
Figure 7: Bar diagram of the cross validation errors of all models ...............19
List of tables

Table 1: Stratification of field observations.................................................................7
Table 2: Narrow band vegetation indices.................................................................7
Table 3: Parameters setting for genetic algorithm..................................................11
Table 4: Band combinations for narrow band vegetation indices..........................14
Table 5: Result of simple linear regressions............................................................14
Table 6: Model selection for partial least square regression.................................15
Table 7: Result of partial least square regression (1)...............................................15
Table 8: Result of partial least square regression (2)...............................................16
Table 9: Result of regressions with neural networks.............................................18
1. Introduction

1.1. Background

Salt marshes, which are characteristics subsystems of coastal areas, have been subjected to intense pressure because of unprecedented human activities, increased input of atmospheric nitrogen and sea level rise (Bockelmann, 1999). This has generated great concern because salt marshes provide essential functions in the buffering and transferring of nutrients and maintenance of coastal water quality. Besides, coastal salt marshes are important as natural ecosystem remnants offering wildlife habitat and tourist destinations. Information on the quantity of biophysical parameters is critical for an improved understanding of ecological processes and better management and monitoring of salt marsh. Leaf area index (LAI) is one of the fundamental biophysical parameters of vegetation. It is defined as one-half of the total surface area of leaves per unit ground area. Information on LAI is important for quantifying the energy and mass exchange characteristics of salt marsh ecosystem such as photosynthesis, respiration, transpiration, carbon and nutrient cycle (Gong et al., 1995). However it is difficult to estimate LAI accurately given the heterogeneity and large spatial gradients within salt marshes. Because of extreme edaphic heterogeneity within salt marshes, a few centimetres of elevation change creates a different vegetation composition and hence variations in LAI. The inherent heterogeneity within the system and difficulty in direct access favour the need for remote sensing technologies (Zhang et al., 2005).

Remote sensing techniques offer practical means to measure LAI at the landscape scale (Gong et al., 1995). Because LAI is functionally linked to the canopy spectral reflectance, its retrieval from remote sensing data has prompted many researches in recent years. These researches were mainly conducted on the estimation of forest and crop LAI from remote sensing data (e.g., Schlerf et al., 2005; Haboudane et al., 2003; Brodge and Leblanc, 2000). The advents of hyperspectral remote sensing have offered new opportunities to estimate LAI even for the heterogeneous (mixed species) ecosystems like salt marshes. Hyperspectral sensors uniquely measure reflected sunlight in hundreds of contiguous, narrow spectral bands. Recent literatures have shown that these narrow bands are crucial with significant improvements over broad bands in quantifying LAI (Treitz and Howarth, 1999).

Estimation of LAI from spectral reflectance measurements can be derived using two types of analytical techniques: deterministic or stochastic canopy radiation models and empirical models. Stochastic techniques model the radiative transfer process between the land surface and the sensor to invert reflectance measurements to LAI. Though radiation models are precise in their treatment of radiative transfer in vegetation canopies, they are difficult to parameterize and are often developed for relatively homogeneous vegetation cover types (Treitz and Howarth, 1999). Consequently, these are more suited to agricultural canopies as opposed to natural vegetation consisting of species mixtures with variations in leaf optical and structural properties.
Hence, the majority of studies which estimated LAI from remotely sensed data have used empirical modelling techniques to relate spectral data or their various derivatives. Several ways of empirically modelling LAI from hyperspectral data exist; among which one of the most widely used approaches is the univariate regression analysis with vegetation indices (e.g., Schlerf et al., 2005; Pu and Gong, 2004). However, vegetation indices have unavoidable limitations as they are not single and unique measure of LAI (Haboudane et al., 2003) Most of the indices such as normalized difference vegetation index (NDVI) and simple ratio (SR), which are frequently correlated with LAI, are derived from the bands of visible and near infrared region. Canopy reflectance in these regions are not only affected by LAI but also by inherent factors like pigment concentration, canopy architecture and extrinsic factors like soil background, illumination and viewing geometry (Broge and Leblanc, 2000). In light of this, several other new vegetation indices have been proposed which attempt to reduce the effect of some of these extraneous factors. For instance, soil adjusted vegetation index (SAVI) and modified soil adjusted vegetation index (MSAVI) have been developed to reduce the soil background and brightness influences (Gong et al., 2003).

All the above-mentioned vegetation indices utilize only two bands out of many available bands of hyperspectral data. Theses indices derived from narrow bands perform better than their corresponding broad-band indices in estimating biophysical variables like LAI (Gong et al., 2003; Thenkabail et al., 1999; Elvidge and Chen, 1995). However, the strength of the hyperspectral data, i.e. availability of large number of bands, is not exploited if only two bands from red and infra-red region are used in formulating the indices. Recently, few studies (e.g., Schlerf et al., 2005; Hansen and Schjoerring, 2003; Thenkabil et al., 1999) have shown that the utilization of the entire spectrum of data in search of the optimal combination of bands for vegetation indices help in estimating the biophysical variables like LAI with higher accuracy.

Multiple linear regression (MLR) is another popularly used empirical approach which utilizes several bands or band ratios that are usually correlated to measure LAI. The disadvantage of multiple linear regression is that it suffers from overfitting (Coops et al., 2003) and singularity (Geladi and Kowalski, 1986) problems in case of redundant and large number of predicting variables as in the case of hyperspectral data. One of the solutions of the problems related to multiple regression is partial least square regression, widely used in chemometry, and which has recently found attention in remote sensing application. It is particularly useful for constructing predictive models when the variables are many and highly collinear (Coops et al., 2003; Townsend et al., 2003). It fully takes advantage of the high dimensionality of the spectral data by building predictive models through the extraction factors from the original data which best explains both the response and predictor variables (Townsend et al., 2003).

Though partial least square regression can utilize the full spectrum of the high dimensional data, the least square solution is derived assuming linear estimates (Frank and Friedman, 1993). Many biophysical variables including LAI tend to exhibit non-linear behaviour with the spectral data (Uno et al., 2005; Mutanga and Skidmore, 2004; Schmidt, 2003). There exists another popular solution for the non linear function approximation problems known as neural network. The most popular neural network currently in use is the multi-layer perceptron (MLP), which uses the back propagation training algorithm (Murnion and Kubo, 1998; Skidmore et al., 1997). It works well with training data
sets that are smaller in size than those required by the standard non-linear statistical procedures (Wilkinson, 1997). On the other hand, network training times can be lengthy, while choice of the design of network architecture (in terms of numbers of hidden layers and neurons per layer) and the values of the learning rate parameters is not straightforward (Wilkinson, 1997; Skidmore et al., 1997). The most important problem associated with neural network is that, full spectrum of data can not be utilized because of the redundant information in hyperspectral bands (Ozkan and Erbek, 2005). Again, as the number of bands increases, the number of training samples required also increases nonlinearly. Therefore, it is necessary to select or extract features before training the algorithm.

One of the most common algorithms for the feature extraction is principal component analysis (PCA). Although, PCA can dramatically reduce the number of dimension, there is always the possibility of some information loss if the higher order principle components are omitted. Similar to PCA, there is another dimensionality reduction technique called minimum noise fraction (MNF) which always transforms images with decreasing image quality with increasing component number. This desired optimal ordering is not always achieved with PCA (Green et al., 1988).

1.2. Research Objective

The main objective of this research is to compare different empirical modelling techniques in estimating the leaf area index of salt marsh vegetation using airborne hyperspectral data.

1.3. Specific Objectives

- To investigate the usefulness of different combination of bands in selected hyperspectral vegetation indices in estimating LAI.

- To develop univariate regression models between LAI and selected vegetation indices.

- To evaluate the usefulness of partial least square regression in estimating LAI.

- To evaluate the usefulness of neural network in estimating LAI.

- To compare the performance of minimum noise fraction and principle component analysis in feature extraction for the neural network model.

1.4. Research Questions

- Which regression technique; simple regression, partial least square regression or neural network; can estimate LAI with higher accuracy?

- Which of the selected vegetation indices calculated from all possible combinations of narrow bands can estimate LAI with higher accuracy?

- Which feature extraction method, principle component analysis or minimum noise fraction, extracts features for the neural network model to estimate LAI with higher accuracy?
2. Methods

2.1. Study area

The study area comprises a coastal salt marsh on the southern side of the island of Schiermonnikoog, which is a part of the Dutch Waddenzee ecosystem. The area comprises 120 vegetation species, of which about 15 are dominant. The canopy never exceeds 1.5 m, but most canopies are on average 25 cm high. The canopy is mainly erectophile, or circular (random leaf angles) with varying degree of bare soil and litter between vegetation communities. The salt marsh gradually rises in elevation towards the north, ending in natural and man-made dunes, which comprises the highest topography of the island.

2.2. Data collection

2.2.1. Acquisition of Airborne Image Data (AHS)

A flight was undertaken on 15th June, 2005 by Flemish Institute for Technological Research (VITO) in order to acquire the airborne hyperspectral images of the study area. The airborne hyperspectral sensor (AHS) was used in the acquisition. AHS measures a total of 80 images in the wavelength range from 455 nm to 12,227 nm. The sensor has a spatial resolution of 3.7 X 3.5 meter.

2.2.2. Field data collection

The field visit was carried out during the month of August, 2005. In order to build a meaningful relationship between phenologically varying vegetation and remotely sensed data, ground measurements should preferably be contemporaneous with the image acquisition time. Because of some unavoidable practical reasons, the date discrepancy between the ground truthing and image acquisition was six weeks. As this study focuses on building the empirical relationship between the image data and LAI, it was assumed that the date inconsistency will not considerably affect the relationship.

A total of 85 vegetation plots were described in the field. Each vegetation plot was of 2.5 m by 2.5 m. The plot area was defined to cover homogenous parts of each dominant vegetation type. Careful attention was given to select homogenous areas which were roughly at least 10 m by 10 m. The central location of each plot was determined with an accuracy of about +/- 5 m using a global positioning system (GPS). The full species composition description, with their percentage cover estimates, was made for every plot.

15-20 random LAI were estimated in each plot using LI-COR LAI-2000 Plant Canopy Analyzer. This portable lightweight instrument can estimate LAI in a very rapid manner using diffuse radiation. A sealed fisheye lens containing five concentric light-detecting silicon rings measures the diffuse radiation at five known zenith angles. LAI is determined from the canopy's gap fraction which is based on ratios of above and below (A and B) canopy readings (LI-COR Inc., 1992). Several factors
such as sky conditions, foliage clumping, woody materials and plant phenology all affect LAI estimates (Schlerf et al., 2005).

The LAI-2000 was only operated under overcast sky conditions. For few operations carried out during clear sky, a big umbrella was used to make shadow to avoid the direct sunlight. A 270° view restrictor was used on the sensor. For each random measurement, 5 B readings and 1 A reading were taken.

2.3. Pre-processing

2.3.1. Image data

The AHS images were preprocessed by the data supplier VITO, the institute that undertook the flight in the study area. The images were corrected for atmospheric and geometric distortions. Georeferencing was done by the direct georeferencing methodology. Direct georeferencing is the direct measurement of the position and orientation parameters of a sensor and it is a technique increasingly used in airborne mapping applications because of its economical advantages (Honkavaara, 2004). For the validation of the georeferencing methods, five whiskbroom and pushbroom scanner images were used. The steps used in atmospheric correction involved the transformation of at-sensor radiance to surface radiance, atmospheric correction using the MODTRAN4 radiative transfer code and finally the retrieval of surface reflectance of each pixel.

In their pre-processing report obtained in the last week of December 2005, VITO pointed out that there was a linear degradation of optics during the flight day due to accumulation of ice and dirt. Hence the reflectance values in the image did not directly correspond to the surface reflectance. They mentioned that they would be recalibrating and reprocessing the dataset to be ready in the last week of January, 2006. Because of the time limitation, it was not feasible to wait for the newly processed dataset.

Pixels had strange values within the range of 2031 nm to 2500 nm and were noisy in this range. Hence, all those wavebands starting from 2031 nm were dropped and only 21 remaining bands were retained for this study. The reflectance values were dramatically different in the overlapping areas in two different strips. It provided difficulties in mosaicing the two images. Hence one of the strips was dropped, necessitating the removal of seven ground truth observations.

2.3.2. Field data

Pre-processing of LAI 2000 data was carried out using LAI-2000 data file viewer software. 15-20 LAI estimates were averaged to get a single LAI for each plot. Together with LAI, mean tip angles (MTA) estimated by LAI 2000 PCA were also averaged per plot. Before averaging, measurements of each plot were carefully analyzed for the bad readings. None of the outer rings were eliminated in the gap fraction inversion. No correction for the foliage clumping was applied.

Several authors have suggested deriving species specific relationship of spectral data and LAI (Schlerf et al., 2005). Given the heterogeneity of the study area and time limitation of this study, it was not possible to collect significant number of LAI data based on the vegetation types. However, there were reasonable numbers of samples to be stratified into two categories i.e. high/middle marsh vegetation
type and low/pioneer marsh vegetation type. Examining the vegetation composition of all the sampling plots, it was found that species belonging to genus *Elymus* was predominantly occurring in 15 plots. Accordingly, the field data was reorganized into four different subsets as shown in the table 1. The division of dataset was carried out following the guidelines of TMAP salt marsh topology key, version 2004. The codes for the subsets in this study were adopted differently for the clarity purpose.

Table 1: Resulting subsets after stratification of the field observations

<table>
<thead>
<tr>
<th>subset</th>
<th>Designation</th>
<th>Code</th>
<th>No. of observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>All observations</td>
<td>A</td>
<td>78</td>
</tr>
<tr>
<td>2</td>
<td>High/middle salt marsh</td>
<td>H</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Low/pioneer salt marsh</td>
<td>L</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>Elymus vegetation type</td>
<td>E</td>
<td>15</td>
</tr>
</tbody>
</table>

2.3.3. **Image grass spectra extraction**

In order to utilise the image data in regression analysis, vegetation spectra for each sampling unit had to be extracted. ENVI 4.2 Software was used in the extraction of spectra from the AHS image. The point map of field sampling plots was overlaid on the airborne image. For each sampling plot, a continuous spectrum of reflectance values for the range 455nm to 1622nm was extracted in ASCII format by defining the whole sampling plots as regions of interest. After extraction, it was found that most of the XY coordinates of the extracted spectra were slightly different (up to +/- 3 m) than their corresponding point map coordinates. This difference occurred because of the pixel size. LAI attributes were hence assigned to the extracted spectra by visually comparing their coordinates with the point map coordinates.

2.3.4. **Calculation of hyperspectral vegetation indices**

In this study, the availability of hyperspectral data in 21 discrete narrow bands allowed computation of all possible 21 X 20 = 420 combination of narrow bands. The narrow bands were calculated to find out suitable pair of wavebands for the normalized difference vegetation index (NDVI) and modified soil adjusted vegetation index (MSAVI). NDVI was selected because it is the most known and widely used vegetation index in estimating LAI. Despite its extensive use, NDVI has been found to be highly sensitive to optical properties of soil background (Baret and Guyot, 1991). Hence MSAVI was also selected as it attempts to minimize the soil brightness effect (Qi et al., 1994).

Table 2: Narrow band indices investigated in this study

<table>
<thead>
<tr>
<th>Index</th>
<th>Abbreviation</th>
<th>Computation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized difference vegetation index</td>
<td>NDVI</td>
<td>$\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}$</td>
<td>Rouse et al. (1974)</td>
</tr>
<tr>
<td>Modified soil adjusted vegetation index</td>
<td>MSAVI</td>
<td>$\frac{1}{2}[2\lambda_1+1-\sqrt{(2\lambda_1+1)^2-8(\lambda_1-\lambda_2)^2}]$</td>
<td>Qi et al. (1994)</td>
</tr>
</tbody>
</table>
2.3.5. Feature extraction for neural network

The neural network used in this study (i.e. multilayer perceptron, which will be described later) can learn to weight significant variables and ignore less important ones (Mutanga and Skidmore, 2004; Boyd, 2002). However, selection of important bands (feature selection) was necessary because; i) there is the presence of redundant information in hyperspectral bands (Ozkan and Erbek, 2005), ii) it requires large number of training samples for neural network with the use of increased number of bands (Statsoft, 2005), and iii) more the bands used, increased processing time is required. Feature selection is usually performed in transformed space, which is also called feature extraction. The original feature vectors are transformed into new set of orthogonal and hence uncorrelated variables, and the selection of appropriate features is done based upon the amount of explained variance by each feature. In this study, feature extraction was done by principle component analysis and minimum noise fraction method.

I) Principle component analysis

Principal component analysis (PCA) is one of the oldest and best-known techniques of multivariate analysis. In PCA, an orthogonal subspace projection is performed on the high-dimensional data to produce a new sequence of uncorrelated data, finally retaining only the significant components for further processing without much loss of information. The objective of PCA is to find a lower-dimensional representation that accounts for the major variance of the original dataset.

In this study, PCA was performed on the subset of the image defined by the sample plots using ENVI 4.2 software. The background value of the image was masked for the calculation of statistics and subsequent transformation of the data. PCA result showed that the first four components contained the cumulative variance of 99.72%. However, all the 21 components were retained considering that the relevant number of components would be tested during the model selection process for neural network.

![Figure 1: Percentage variation expressed by the principle components. Only the first 10 components are shown for the clarity purpose. First 4 components contain 99.7% of the variation](image)

II) Minimum noise fraction

With PCA, there is always some risk of information loss particularly with large number of variables, if the components with less variation are discarded (Green et al., 1988). On the other hand, minimum noise fraction (MNF) transform, as described by Green et al. (1988), transforms data with optimal ordering of high noise to signal ratio in the last few components. It is essentially a two cascaded principle component transformation. The first transformation based on estimated covariance matrix,
decorelates and rescales the noise in the data. It results into a set of transformed data in which the noise has unit variance and no-band to band correlations. The second step is a standard principle component transformation of the noise-whitened data. The advantage of MNF over PCA is that it accounts for the fact that the noise in some bands may be larger than the signal in other bands.

The eigenvalues produced by MNF are scaled in sigma noise units analogous to signal to noise ratio so that the number of eigenvalues greater than unity gives an estimate of the number of dimensions with variance larger than the amplitude of the noise estimate (ENVI, 2003). In this study, MNF was performed on the same subset of image as in PCA. The background value of the image was masked for the calculation of statistics and subsequent transformation of the data. The inherent dimensionality of the data was determined by the examination of final eigenvalues. Those coherent components of the transformed spectral data which had eigenvalues greater than unity were selected to be input to the neural network. This led to the dimension reduction from 21 to 20 after MNF transformation.

![Figure 2: Eigen values and percentage variation associated with the MNF components](image)

2.4. **Modeling**

Three regression techniques; simple linear, partial least square and neural network; were used for modelling the relationship between ground observed LAI and spectral data. Leave-one-out cross validation technique was applied to validate all the regression models. With n runs in the data set, the model equation was fitted to n-1 runs and a prediction was made from this model for the remaining one. Cross validation root mean square errors (CV-RMSE) were calculated for all the models to assess the generalization capability of the models. Because the predicted samples are not the same as samples used to build the model, CV-RMSE statistic gives a good indication of the predictive power of the model (Schlerf et al., 2005).

2.4.1. **Simple linear regression**

Simple linear regression is a common procedure to establish an empirical relationship between the vegetation index and LAI by statistically fitting measured LAI values and the corresponding values of the vegetation index. Vegetation index is often related to LAI through a linear or exponential regression model, depending upon the saturation effects. Visual examination of scatter plot showed no distinct saturation effect. Therefore, for this study, the data were only fitted with linear models. The regressions were fitted to all the four subsets of the data with narrow band NDVI and MSAVI. These narrow band indices were derived from the combination of optimum bands out of all possible 420 band combinations.
2.4.2. Partial Least square regression

I) Introduction

Partial regression is a relatively new technique which generalizes and combines features from principle component regression and multiple linear regression (Geldi & Kowalski, 1986). It is particularly useful for constructing predictive models when the variables are many and highly collinear as in the case of hyperspectral data. Partial least square regression finds components from both predictor(X) and response(Y) variable by searching for a set of components (called latent variables) that performs a simultaneous decomposition of X and Y with the constraint that these components explain as much as possible of the covariance between X and Y. It is followed by the regression step where the decomposition of X is used to predict Y.

Partial least square can benefit from the large number of predictor variables. In some cases, however, some variables may contain noise or interfering signal which may actually deteriorate the accuracy of a regression model (Wise et al., 2005). In these cases, it can be advantageous to discard some variables. Genetic algorithm for variable selection is a technique which helps to identify the subsets of the measured variables that are, for a given problem, the most useful for regression model. Given an X-block of predictor data and a Y-block of values to be predicted, a random subset of variables from X can be used to determine the cross validation error (CV-RMSE) by using only that subset of variables in a regression model. Genetic algorithm uses this approach iteratively to locate the variable subset (or subsets) which gives the lowest CV-RMSE. For a detail description of genetic algorithm for variable selection, please refer to the Wise et al. (2005).

II) Partial least square regression using full spectrum

Partial least square regressions were initially fitted utilizing the full spectrum of data (21 bands). Nonlinear iterative partial least squares (NIPLS) algorithm in PLS toolbox 3.5 was used for the purpose as this is the most complete and elegant partial least square algorithm when prediction is desired (Geldi & Kowalski, 1986). Before the models were developed, the values for the independent variables were mean centred. The full spectrum data was utilized to find the partial least square latent variables. To avoid overfitting, CV-RMSE statistics for each latent variable were calculated based on leave-one-out cross validation procedure. The numbers of significant latent variables to form a final simplest model were chosen based on the following rules of thumb as suggested by Wise et al. (2005).

- Do not include additional latent variable unless they improve the CV-RMSE by at least 2%.
- Also take consideration of the variance captured in making a final determination.
- When in doubt, choose the model with fewer latent variables.

III) Partial least square regression coupled with genetic algorithm for variables selection

The selections of useful variables using genetic algorithm were done for all four subsets using PLS toolbox 3.5 in Matlab software. The table 3 shows the setting of input parameters that were used when carrying out genetic algorithm. Small descriptions of each parameter are also provided in the table. The population was kept intermediate (60) to constrain the iteration time. Window width was set to one because the total variable numbers was not high. Penalty slope was set to zero not to guide the algorithm towards a solution with the specified number of variables. Max generation and convergence
were kept medium (100 and 50 respectively) to avoid overfitting. Mutation rate was kept minimum
(0.005) to exclude underrepresented and include overrepresented variables in the mutation process.

‘Fitness versus variable usage plots’ were used to locate the suitable number of variables. The
variables were selected on the basis of their contribution in reducing CV-RMSE of different models
generated by genetic algorithm. After finding the suitable variables, partial regressions were fitted for
all subsets following the same procedure as described above with full spectrum.

Table 3: Parameters for the variable selection procedure using genetic algorithm

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value/type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Total number of subsets</td>
<td>60</td>
</tr>
<tr>
<td>Window width</td>
<td>Number of adjacent variables to be grouped in initial subsets</td>
<td>1</td>
</tr>
<tr>
<td>Initial terms (in %)</td>
<td>Approximate number of variables in initial subset</td>
<td>30</td>
</tr>
<tr>
<td>Penalty slope</td>
<td>Required if needed to guide the algorithms towards specified number of target variables (not specified here)</td>
<td>0</td>
</tr>
<tr>
<td>Max generation</td>
<td>Finite number of iteration (ending criteria)</td>
<td>100</td>
</tr>
<tr>
<td>Convergence (in %)</td>
<td>Percentage of the subsets having identical variables (ending criteria)</td>
<td>50</td>
</tr>
<tr>
<td>Mutation rate</td>
<td>Rate of exchange of variables among subsets</td>
<td>0.005</td>
</tr>
<tr>
<td>Cross over</td>
<td>Method of breeding the subsets</td>
<td>Single</td>
</tr>
<tr>
<td>Cross validation</td>
<td>Splitting of variables within subset into training and test set for regression</td>
<td>Random</td>
</tr>
<tr>
<td>Replicate runs</td>
<td>Number of times of algorithm running</td>
<td>7</td>
</tr>
</tbody>
</table>

2.4.3. Neural network

I) General introduction

Neural network algorithm has emerged in the remote sensing field as a promising technique for a
number of situations such as non-normality, complex feature spaces and multivariate data types,
where traditional methods fail to give accurate results (Atkinson and Tatnall, 1997). The most widely
used neural network model in remote sensing applications is the backpropagating multi-layer
perceptron (Skidmore et al., 1997). It is a fully connected feed-forward neural network, which
generally consists of three layers of simple processing units. Each unit in each layer is fully connected
to each of the units in the next layer.

In backpropagating multi-layer perceptron, each unit performs a biased weighted sum of their inputs
and passes this activation level through a transfer function to produce its output. The units are
arranged in a layered feed forwarded topology. The network thus has a simple interpretation as a form
of input-output model, with the weights and thresholds (biases) the free parameters of the model. Such
networks can model function of almost any arbitrary complexity, with the number of layers, and the
number of units in each layer, determining the function complexity. Important issues in network
design include specification of the number of hidden layers and the number of units in these layers.

One of the standard backpropagation is a gradient descent algorithm, in which the network weights
are moved along the negative of the gradient of the error function. A sequence of such moves (slowing
when we near the bottom) will eventually find a minimum of some sort. The difficult part is to decide
how large the steps should be. Large steps may converge quickly, but may also overstep the solution,
or proceed in the wrong direction. In contrast, very small steps may proceed in the right direction, but
they also require a large iteration time. In practice, the step size is proportional to the slope, and to a
special constant: the learning rate. The algorithm can be modified by the inclusion of a momentum
term which provides faster convergence. It encourages movement in a fixed direction, so that if several steps are taken in the same direction, the algorithm picks up speed. Acting like a low pass filter, momentum allows the network to ignore small features in the error surface. Without momentum a network may get stuck in a shallow local minimum. A detailed discussion of backpropagating MLP and several other neural networks can be found in Demuth and Beale (2004).

Neural network toolbox in Matlab software was used to carry out the modelling with neural network in this study. The procedures followed are described in the following sections.

II) Training the network
The gradient descent backpropagation with momentum as above described was used to train the network. The network was trained using batch training. A two layer network with tan-sigmoid transfer function in the hidden layer and a linear transfer function in the output layer was used, which is considered to be a useful structure for regression problems (Demuth and Beale, 2004). Number of epochs, learning rate and momentum was set to 5000, 0.3 and 0.1 respectively by conducting several trial and error experiments while holding the other network parameters constant. Network inputs were scaled by normalizing the input data to improve the performance of the network. At first, all the extracted principal components were used as inputs to the network. The relevant numbers of input components were further tested during the model selection. Once the model selection was over with the principal components, same process was repeated for the minimum noise fraction components.

III) Models Selection
Leave-one-out cross validation was used for model selection by choosing one of several models that has the smallest prediction errors. The procedure described by Anders and Korn (1998) was adopted for the selection of a model.

The procedure started by moving bottom-up (starting from one hidden unit) in order to determine the appropriate number of hidden units. In each step, the cross validation errors of two models were compared, one of which contained an additional hidden unit. Each cross validated model was run three times and the average errors of the three cross validated models were reported as the cross validated RMSE (CV-RMSE). If the more complex network turned out to show a smaller CV-RMSE, the additional hidden unit was accepted and the network enlarged. The procedure stopped when no further hidden unit was able to reduce the CV-RMSE of the previous model.

Subsequently, the top-down strategy followed which started from the fully connected network obtained in the first step and tried to detect irrelevant input units. All sub-models with one of the input units removed were estimated and compared with the full network by means of CV-RMSE. In the finally chosen network model, no input unit could be removed without increasing the CV-RMSE. With the finally chosen input units, the appropriate number of hidden units was further tested to see whether removing or adding extra hidden unit would decrease the CV-RMSE.

Model selection by the abovementioned procedure was carried out with subset A. The procedure for individual subset could not be repeated because of the time constraint in this study. Hence, the finally chosen network model with subset A was applied to all four subsets.
3. Results

The data analyses were carried out with four different subsets, full dataset (subset A), subset with high/middle marsh observations (subset H), subset with low/pioneer marsh observations (subset L) and subset with elymus vegetation types (subset E). The results of data analyses are presented as follows.

3.1. Band combinations for narrow band vegetation indices

To determine the optimal narrow band vegetation index, the square of the coefficients of determination ($R^2$) between all possible two-band vegetation indices and ground observed LAI were computed. One of the results (subset A) is illustrated in 2D-correlation plots (figure 3). Each box corresponds to the combination of two bands, representing the $R^2$ value between LAI and MSAVI calculated from the reflectance values in those two wave bands. Similar correlation plots were computed for all subsets with both NDVI and MSAVI (see appendix 2). Based on the $R^2$ values in the 2D correlation plots, band combinations that formed the best indices were determined. These were considered optimal indices and hence named optimal NDVI and optimal MSAVI. For example, in figure 3, the combination of bands 862nm and 804 nm were optimal for MSAVI. Band positions thus identified for both optimal NDVI and optimal MSAVI are tabulated in table 4.

![2D-correlation plot](image)

Figure 3: 2D-correlation plot that shows the square of the correlation coefficient ($R^2$) between LAI and narrow band MSAVI values (subset A) for every band combination.
The results showed that different band combinations were utilized in formulating the optimal NDVI and optimal MSAVI for different subsets. For subset A, both vegetation indices utilized bands from near-infrared region. For other subsets, the optimal band combinations were different for two vegetation indices. Optimal MSAVI for subset H and Optimal NDVI for subset L also utilized both bands from near-infrared region. Optimal MSAVI utilized bands from green and blue region, whereas optimal NDVI utilized bands from red edge and green region for subset E.

### 3.2. Optimal narrow band vegetation index relationship with LAI

Simple linear regressions were fitted to the ground observed LAI with narrow band optimal NDVI and optimal MSAVI for all subsets. The results of all the linear models are reported in table 5. The first column illustrates the type of the vegetation index used. The second column refers to the subsets used in regression analysis. The third column shows the linear regression equation between the dependent variable (LAI) and independent variable. Regression coefficients and intercepts of the equations were calculated by averaging the corresponding coefficients and intercepts of all cross validation models. The fourth column reports the cross validated error (CV-RMSE) of each regression equation.

The results show that simple linear regressions using optimal MSAVI provided higher accuracy than the optimal NDVI in predicting LAI for subset A. For other three subsets, optimal NDVI provided higher accuracy than optimal MSAVI. Among four subsets, highest accuracy was obtained for subset E and lowest accuracy was obtained for subset L with both narrow band optimal vegetation indices.

### Table 5: Simple linear regressions of LAI with optimal narrow band indices. Cross validated RMSE is reported for each regression model for each subset.

<table>
<thead>
<tr>
<th>VI</th>
<th>Subset</th>
<th>Regression equation</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal MSAVI</td>
<td>A</td>
<td>$Y = 0.74 + 93.49 \times X$</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>$Y = -1.56 + 188.7 \times X$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>$Y = 6.37 - 76.59 \times X$</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>$Y = 0.31 - 100.49 \times X$</td>
<td>0.69</td>
</tr>
<tr>
<td>Optimal NDVI</td>
<td>A</td>
<td>$Y = 0.82 + 78 \times X$</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>$Y = -0.49 + 18.15 \times X$</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>$Y = -1.47 - 50.99 \times X$</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>$Y = 20.26 - 38.51 \times X$</td>
<td>0.63</td>
</tr>
</tbody>
</table>
3.3. Partial least square regression

I) Partial least square regression with full spectrum

Partial least square regressions were fitted to LAI with full spectrum (21 bands) for all subsets. The results of the analysis for subset A are presented in table 6. The first column shows the number of latent variables included in the models. The second and the fourth columns show the variation of the predictor and predicted variables respectively captured by the particular latent variable. Similarly, the third and the fifth columns relate to the total variance of the predictor and predicted variables respectively captured by the models. The final column depicts the cross validation errors related with the number of latent variables in the models.

Table 6: Cross validated RMSE and captured variance of predictor and predicted variables associated with different numbers of latent variables (up to ten) in partial least square regression models (subset A). The latent variable (LV) at which the lowest CV-RMSE achieved and its associated attributes are highlighted with bold fonts.

<table>
<thead>
<tr>
<th>LVs</th>
<th>Captured variance (X)</th>
<th>Captured variance (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current</td>
<td>Total</td>
</tr>
<tr>
<td>1</td>
<td>92.67</td>
<td>92.67</td>
</tr>
<tr>
<td>2</td>
<td>3.54</td>
<td>96.2</td>
</tr>
<tr>
<td>3</td>
<td><strong>3.25</strong></td>
<td><strong>99.45</strong></td>
</tr>
<tr>
<td>4</td>
<td>0.29</td>
<td>99.74</td>
</tr>
<tr>
<td>5</td>
<td>0.07</td>
<td>99.81</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td>99.86</td>
</tr>
<tr>
<td>7</td>
<td>0.07</td>
<td>99.93</td>
</tr>
<tr>
<td>8</td>
<td>0.02</td>
<td>99.95</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>99.95</td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>99.96</td>
</tr>
</tbody>
</table>

Using the first rule of thumb (see section 2.4.2), to only choose additional factors when the CV-RMSE improves by at least 2%, a model with three latent variables made a good final choice. This choice was also supported by the variance of the predictor variables captured by the model. Similar analyses were done for all subsets of the data (see appendix 4) and selection of the final models was done. The number of latent variables retained and the corresponding cross validation errors of the final models of all subsets are tabulated in table 7.

Table 7: Cross validated RMSE and number of latent variables associated with final partial least square models.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Number of latent variables</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>1.05</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>0.78</td>
</tr>
<tr>
<td>L</td>
<td>2</td>
<td>1.40</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>0.81</td>
</tr>
</tbody>
</table>

II) Partial least square regression with variables selected from genetic algorithm

‘Fitness versus variable usage plots’ were produced with genetic algorithm. One of those plots for subset A is presented in figure 4. Only those models with CV-RMSE better than the average fitness of the models of final population are shown in the plot. Each model constituted three to nine variables. The inclusion marks for the variables, which improved CV-RMSE, appear towards the bottom of the figure. There is no mark for the variables which did not appear in any model and hence are not useful. The variables which appeared in all models are more useful. Hence the variables (wavebands in nm)
455, 542, 689, 746, 862, 948 and 1004 were finally selected as they together provided the least cross validation errors. Similar analyses were done for all the subsets and the useful variables were thus identified (see Appendix 3).

Partial least square regressions were then fitted to all subsets by using the most useful variables selected by genetic algorithm. The results of the analyses are presented in table 8. The comparison with the results obtained from partial least square regression using full spectrum (table 7) show that variable selection by genetic algorithm improved the accuracy for each subset.

Table 8: Cross validated RMSE of the partial least square models after variables selection by genetic algorithm. The number of variables selected and the number of the latent variables used in the final models are also reported.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Number of Variables</th>
<th>No of latent variables</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7</td>
<td>6</td>
<td>0.97</td>
</tr>
<tr>
<td>H</td>
<td>9</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>L</td>
<td>5</td>
<td>5</td>
<td>1.21</td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>6</td>
<td>0.54</td>
</tr>
</tbody>
</table>

3.4. Models selection using neural networks

Predictive modelling using neural network involved a search of system parameters (input and hidden units) that increased the accuracy of the model. Cross validated RMSE was used to find the appropriate number of hidden units using all the principal components (PCs) and minimum noise fraction (MNF) components. This process was followed by the search of significant number of input units while holding the number of hidden units constant.
The results show that four hidden units provided the lowest cross validation error by using all the principal components and MNF components as input units (see figure 5). Addition of further hidden unit could not decrease the cross validation error and hence the network with four hidden units was finally accepted.

Figure 5: Cross validation errors of the network models with increasing hidden units. The number of PCs was fixed at twenty one whereas number of MNF components was fixed at twenty. Lowest cross validation errors are obtained with four hidden units for both PCs and MNF components.

By holding the number of hidden units fixed at four, CV-RMSE started decreasing steadily with reduced number of principle components until the number of input components reached six (see figure 6). The cross validation error could not be decreased further by either adding or removing the number of hidden units while holding the number of components fixed at six. Hence the network with four hidden units and six principle components was retained as final model. Similarly eleven MNF components provided the lowest cross validation error and were hence retained in the final model.

Figure 6: Cross validation errors of the network models with decreasing input units (PCs and MNF components). The number of hidden units was fixed at four. The lowest cross validation errors were obtained with six PCs and eleven MNF components.
The results of the finally retained network models for all subsets are presented in table 9. Average calibration RMSE (C-RMSE) and cross validation RMSE (CV-RMSE) of the final models for all subsets are reported. The results show that the network models with PCs provided slightly higher accuracy than the MNF components for subsets A, H and E. The difference between CV-RMSE and C-RMSE was high for subset L, H and E, whereas, comparatively minimal difference between these two statistics was observed for subset A.

Table 9: Calibration and cross validation errors of finally chosen network models for all subsets. The differences between calibration and cross validation errors are provided to see the extent of overfitting. Six PCs were used as the input units. Similarly, the networks were run with eleven MNF components later. The number of input hidden units was fixed at four.

<table>
<thead>
<tr>
<th>Input</th>
<th>Subset</th>
<th>No of components</th>
<th>Hidden units</th>
<th>C-RMSE</th>
<th>CV-RMSE</th>
<th>Difference (Δ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCs</td>
<td>A</td>
<td>6</td>
<td>4</td>
<td>0.75</td>
<td>0.97</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td></td>
<td></td>
<td>0.24</td>
<td>1.27</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td></td>
<td></td>
<td>0.43</td>
<td>2.12</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td></td>
<td>0.32</td>
<td>1.24</td>
<td>0.92</td>
</tr>
<tr>
<td>MNF components</td>
<td>A</td>
<td>11</td>
<td>4</td>
<td>0.71</td>
<td>1.03</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td></td>
<td></td>
<td>0.17</td>
<td>1.58</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td></td>
<td></td>
<td>0.51</td>
<td>2.01</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td></td>
<td>0.11</td>
<td>1.46</td>
<td>1.35</td>
</tr>
</tbody>
</table>

3.5. **Accuracies associated with different models**

CV-RMSE of finally selected models derived from different empirical techniques are re-grouped together in bar diagram (figure 7). In the diagram, CV-RMSE of simple linear regression with optimal NDVI and optimal MSAVI are reported as SLR-NDVI and SLR-MSAVI respectively. Similarly, the errors associated with partial least square regression with full spectrum and with selected variables from genetic algorithm are reported as PLSR-FS and PLSR-GA respectively. Errors associated with neural network integrated with PCA and MNF are reported as NN-PCA and NN-MNF respectively.

The results show that partial least square regressions coupled with genetic algorithm provided the lowest CV-RMSE for all subsets, while partial least square regression utilizing full spectrum was a close second. Neural network with PCA also provided high accuracy for subset A followed by neural network with MNF, whereas both of their accuracies were very low for the rest of the subsets. All the models performed poorly for subset L.
Figure 7: CV-RMSE between ground observed and estimated LAI of four subsets for different empirical regression techniques analyzed. Blue and yellow bars represent the CV-RMSE associated with simple linear regression models with optimal NDVI and optimal MSAVI respectively. Pink and green bars correspond to the errors of partial least square models with full spectrum and partial least square regressions with selected variables respectively. Brown and black bars correspond to the neural network models with PCA and MNF respectively.
4. Discussion

4.1. Comparison of the regression techniques in predicting LAI

The main objective of this study was to evaluate which modelling technique; partial least square regression, simple linear regression or neural network; would predict the LAI of the study area with higher accuracy. The LAI of the study area was predicted with highest accuracy using the partial least square regression. The superior performance of partial least square regression in this study can be attributed to its ability to utilize the multivariate information available with high dimensional data (Hansen and Schjoerring, 2003). When coupled with a genetic algorithm, partial least square regression was further able to choose informative channels thus overcoming the effect of overfitting.

The ability of neural network to predict LAI in this study was inconsistent. It provided comparable accuracy with partial least square regression for the whole dataset (subset A) and lowest accuracies for all other subsets. The poor performance of neural network in predicting the LAI of the study area is linked to the well known underlearning and overfitting problems (Skidmore et al., 1997). Underlearning occurred when the network failed to detect the complexity of the data with few hidden units and input components. While increasing the number of hidden units and input components, network overfitted. The problem of overfitting became worse when the size of the training samples were reduced from the full dataset (subset A, n=78) to the observations related to high/middle marsh (subset H, n=45), low/pioneer marsh (subset L, n=33) and elymus vegetation types (subset E, n=15). This overfitting problem is evident, if we look at the difference between calibration and prediction errors of the final models (see table 9). For subset A, the differences (Δ=0.22 and Δ=0.32) were not great, compared to other three subsets. However, the high calibration errors (C-RMSE=0.75 and C-RMSE=0.71) suggest that, both the models with principle components and minimum noise fraction components could not sufficiently learn due to the complexity of the data. Though calibration errors were minimized for other three subsets, the high differences with cross validation errors (ranging from Δ=0.92 to Δ=1.69) clearly suggest that the networks were highly overfitted.

The overfitting problem of neural network associated with subsets A, H and E may also be accounted by the network architecture used, in addition to the small training samples size. The entire model selection processes were not carried out with these subsets as done with subset A. The finally selected model for subset A was applied to all subsets because of time limitation of this study. Had the model selection been done for each subset, the cross validation errors might have been decreased.

The results obtained with simple linear regressions indicated that information contained in two-narrow bands does not suffice to build a good predictive model to estimate the LAI of the study area. Genetic algorithm selected most of those bands (if not all) used by narrow band optimal indices. Hence, in addition to those bands used by simple regressions, partial least square regressions also utilized useful information from other bands and hence provided higher accuracy.
Contrary to the results of this study, several investigators have found neural network as well as simple regression with vegetation indices useful in predicting LAI of forest and agricultural areas (e.g., Schlerf et al., 2005; Gong et al., 2003; Danson et al., 2003; Fang and Liang, 2003). There is no literature available related to the prediction of LAI of salt marsh area using remote sensing. The extreme heterogeneity within salt marshes is what makes it different from agriculture and forest areas in the context of remote sensing. Salt marsh vegetation is characterized by great variation in leaf angles, canopy cover, canopy height, standing litter etc. Leeuwen and Heute (1995) and Asner (1998) have shown that, these factors contribute significantly to canopy reflectance both within and between species. The contribution of these factors can even exceed the contribution of LAI. The highest accuracy obtained with subset E shows that, interference of these factors can be reduced if vegetation type specific relationships are built. However, the collection of sufficient number of samples to build vegetation type specific relationship is difficult. Hence, it requires a regression technique, which can work well with limited data size and at the same time deal with the complexity of the salt marsh vegetation types by utilizing all the useful information available from the hyperspectral data. The study has shown that, partial least square regression can be a useful tool for estimating the LAI of salt marshes.

4.2. **Comparision of optimal narrow band indices**

Optimal MSAVI predicted LAI with slightly higher accuracy (1.8%) than optimal NDVI for subset A. For all other subsets, optimal NDVI provided higher accuracy than optimal MSAVI, the greater difference observed for subset E (12%). These results are contrary to the findings of Haboudane et al. (2004), who found MSAVI to be the best estimator of LAI in terms of sensitivity to canopy parameters. The standard deviation of mean tip angle (or leaf inclination angle) was highest for subset L and lowest for subset E (see Appendix 1). But in both cases, optimal NDVI provided higher accuracy than optimal MSAVI. The results obtained in this study are comparable with the results obtained by Lawrence and Ripple (1998). They also found NDVI performing better than other soil adjusted vegetation indices in determining LAI of a heterogeneous landscape. The possible explanation can be drawn from Thenkabail et al. (2000). Soil adjusted indices provide higher accuracy when vegetation are studied on widely varying soils. When other factors (e.g. canopy architecture) become prominent, the soil adjustment can reduce the accuracy of the prediction. Though the soil characteristics were not investigated in this study, the greater variation in soil conditions (in terms of moisture and nutrients) can be expected within the sampling plots of subset A. This is because the subset A corresponds to entire observations taken from the study area. This might be the reason for the lower accuracy of optimal NDVI compared to optimal MSAVI for this subset. The stratification of all observations into high salt marsh (subset A), low salt marsh (Subset L) and elymus vegetation type (subset E) might have resulted in subsets of observations belonging to relatively homogenous soil conditions. This assumption of homogenous soil condition within these three subsets is reasonable as Bakker et al. (1993) stated that there is the different degree of sedimentation, nitrogen mineralization and exposure to saline water along the gradient of lower to higher salt marshes. Hence, because the variations in canopy factors dominated the variation in soil type, optimal NDVI predicted the LAI with higher accuracy than optimal MSAVI for these three subsets.
4.3. **Performance of feature extraction techniques for neural network**

Principal component analysis (PCA) performed better than Minimum Noise Fraction (MNF) for feature extraction in predicting LAI of subset A using neural network. Though PCA still provided higher accuracy for subset H and subset L, the cross validation errors were too high to be compared.

In this study, six Principal components as input units provided the lowest cross validation error for subset A. Using MNF as feature extraction technique, eleven components needed to obtain the lowest cross validation error. This is because six principal components carried almost all the variation (> 99%) of the original signal, whereas eleven MNF components together could only explain less than ninety percent of variation (see figure 2). The final eigen values (see figure 2) showed that MNF could obtain the desired optimal ordering of decreasing data quality with increasing component number as illustrated by Green et al. (1988). However, the variation of the original data could not be compressed in first few components as in the case of PCA. Hence, the network had to use a greater number of MNF components in the model. This might be the reason for the poor predictably of the network using MNF components; as the greater number of input components used, the greater number of training samples are required (Statsoft, 2005).

4.4. **Band combinations for optimal vegetation indices**

The results showed that different band combinations were utilized in formulating the optimal NDVI and optimal MSAVI. Even for the same vegetation index type, different optimal band combinations resulted for the four different subsets. These results demonstrated that the widely used red and near-infrared (NIR) band combinations are not necessarily useful for two-band NDVI and MSAVI type indices to predict LAI of the study area. For example, optimal NDVI and optimal MSAVI for subset A utilized both bands located at NIR region (862 nm & 804 nm), whereas, optimal NDVI for subset E utilized bands located at red edge (718 nm) and green region (571 nm). Other optimal indices for the rest of the subsets either utilized both bands at NIR region or the bands located at green and red regions. One of the reasons for not utilizing the red and NIR band combinations as used in conventional broad band indices may be related to the problems of saturation. With the increase of LAI, NIR reflectance will continue to rise, but red reflectance will only show modest decrease, resulting in only slight changes in the ratio (Thenkabail et al., 2000).

Schlerf et al. (2005) found that the vegetation indices based on the wavelength positions related to the water absorption features in near infrared provide highest accuracy in predicting forest LAI. Similar results were obtained by Gong et al. (2003) with the bands located around the water absorption features in the NIR and short wave infrared (SWIR) region. Lee et al. (2004) found spectral channels in the red-edge and SWIR regions more important than those in the NIR for predicting LAI. The AHS data utilized in this study did not have contiguous narrow wavebands in NIR region (bandwidth ≈ 71 nm) and hence did not provide the opportunity to investigate the usefulness of these absorption features. Similarly, the SWIR parts of the spectrum were noisy and could not be included in the analysis. Hence, the band positions identified in this study could not be compared to the findings of the abovementioned researchers. The results, however, demonstrated the usefulness of all possible band combinations in formulating the vegetation indices in predicting LAI of the study area. The
reasonable explanation for the different band combinations used by different optimal vegetation indices is beyond the scope of this study and hence not further investigated.

4.5. Comparison of subsets

Stratification of entire dataset (subset A) into high/middle marsh (subset H) and low/pioneer marsh (subset L) did not improve the ability of the models to predict LAI. The improvements in accuracies were observed for subset H but not for subset L. The high variation in both LAI and leaf angles of subset L may be one of the reasons for poor accuracies of the models (see Appendix 1). Furthermore, low and pioneer marsh areas are frequently inundated by the tidal floods; hence the vegetation may be completely or partially covered by water depending on the low tide/high tide period. This may be other possible reason associated with the poor predictability of the regression models for this subset.

Among all subsets, highest accuracies of the models were obtained for subset E. Subset E was built with the field observations belonging to similar vegetation types (dominated by *Elymus*). The higher degree of homogeneity in terms of biophysical (e.g. canopy characteristics), ecological (e.g. species composition) and environmental factors (e.g. soil characteristics) compared to other subsets may be the possible reason for the highest accuracies obtained for subset E.

4.6. Variation of LAI at plot level

The summary statistics of observed LAI of all dataset (subset A) showed a high variation in LAI ranging from 0.93 to 5.9 (see Appendix 1). Even for the same vegetation type (subset E), LAI varied widely from 2.41 to 5.03. Different species composition within the same vegetation type may be one of the possible reasons. To our knowledge, there is no literature available which explains the observed heterogeneity of the LAI of salt marsh. However, there have been ample researches carried out pertaining to the different sources of variation in biomass of the salt marsh. The differences in biomass have been attributed to many factors like water logging and salinity (Rozema and Blom, 1997), frequency of inundation (Bakker et al., 1993) grazing (Bakker, 1989), addition and removal of nitrogen (Kiehl et al., 1997; Schotten et al., 1987) etc. Biomass and LAI are usually correlated (Visconti, 2001) and hence, the possible causes of variation in LAI observed in this study may be linked to the findings of the earlier ecological studies.

The LAI distribution map of the salt marsh area could not be produced because of the lower accuracy obtained with the prediction models. However, this study showed that it is possible to achieve higher accuracy in estimating the LAI of the study area provided that the vegetation type specific models are built. A LAI distribution map produced based on vegetation types may provide an overall picture of the response of salt marsh vegetation towards different environmental and ecological factors. Furthermore, salt marsh species are known to respond differently to exposure to elevated CO$_2$ in atmosphere (Arp et al., 1993). The exposed area of plant leaves plays a key role in plant CO$_2$ exchange with the atmosphere (Haboudane et al., 2004). Hence, in the context of global warming as well, a LAI distribution map may be important in the monitoring of salt marsh vegetation in the future years.
5. Conclusion

The main objective of this research was to evaluate three different regression techniques; (i) simple linear regression, (ii) partial least square regression, and (iii) neural network; in estimating the LAI of a salt marsh area of Schiermonnikoog island of Netherlands. The other objectives were to compare the optimal NDVI and optimal MSAVI calculated from all possible combinations of narrow bands in estimating the LAI of the study area; and to compare principal component analysis (PCA) and minimum noise fraction (MNF) as feature extraction techniques for neural network to estimate the LAI of the study area. The main outcomes of this research are summarized as follows.

- Partial least square regression, when integrated with genetic algorithm for variable selection, provided the highest accuracy in estimating the LAI of the study area. Partial least square regression utilized more useful information in building the predictive model than simple linear regression with vegetation indices, and hence provided higher accuracy. Limited size of the training data was found to be the main reason of poor performance of the neural network.

- Partial least square regression was found to be noise sensitive. Utilizing the full spectrum data, it overfitted the noise and provided less accuracy. Selection of useful variables using genetic algorithm helped in overcoming the effect of overfitting, and hence, improved the accuracies of models.

- Optimal MSAVI did not perform better than optimal NDVI in all subsets. The greater variation in canopy parameters than the variation in soil can be attributed to the lower accuracy of optimal MSAVI in predicting the LAI of the study area.

- Widely used red and near infrared band combinations were not found to be optimal band combinations for NDVI and MSAVI to predict the LAI of the study area. Optimal NDVI and optimal MSAVI utilized different band combinations ranging from blue to near infrared for different subsets.

- Minimum noise fraction was not found to be better than principle component analysis as feature extraction technique for neural network. Though the data was transformed in an optimal ordering of signal to noise ratio, the optimal compression of the data in first few components as in principle components analysis could not be achieved.

- The predictive models apart from neural network provided higher accuracies when the vegetation type specific relationships were built. The wider variation in canopy parameters across the vegetation types deteriorated the accuracies of the prediction models built with all observations and with observations belonging to high marsh and low marsh vegetation types.
6. Recommendation

This study showed that it is possible to estimate LAI of a salt marsh area using airborne hyperspectral remote sensing data. Partial least square regression integrated with genetic algorithm provided the highest accuracies among the studied modelling techniques. However, this study did not investigate the usefulness of new indices which utilize more than two narrow band combinations. Hence, it is recommended for the future studies to test the ability of those indices as well. Because of the poor accuracy obtained with neural network in this study and the difficulties faced in adjustment and fine tuning of network parameters, this study does not recommend the neural network to be used for estimating the LAI of the salt marsh area.

For the future studies, it is recommended to build vegetation type specific predictive models to map the LAI of the area. Once the LAI distribution map is produced with higher accuracy, it broadens the scope of understanding the interaction of vegetation types of the salt marsh area with different environmental and ecological factors in a spatial context.
References


Appendix

1. Summary statistics

The summary statistics of the observed data are presented in the following tables. The first table shows the summary statistics of leaf area index (LAI) data whereas second table shows the summary statistics of mean tip angle or leaf angle (MTA) data for each subset.

<table>
<thead>
<tr>
<th>Subset</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>78</td>
<td>0.93</td>
<td>5.9</td>
<td>3.33</td>
<td>1.24</td>
</tr>
<tr>
<td>H</td>
<td>45</td>
<td>1.33</td>
<td>5.74</td>
<td>3.61</td>
<td>1.11</td>
</tr>
<tr>
<td>L</td>
<td>33</td>
<td>0.93</td>
<td>5.90</td>
<td>2.94</td>
<td>1.32</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>2.41</td>
<td>5.03</td>
<td>3.42</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Summary statistics of observed MTA data for all subsets

<table>
<thead>
<tr>
<th>Subset</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>78</td>
<td>28</td>
<td>79</td>
<td>57.37</td>
<td>9.41</td>
</tr>
<tr>
<td>H</td>
<td>45</td>
<td>28</td>
<td>67</td>
<td>53.56</td>
<td>6.87</td>
</tr>
<tr>
<td>L</td>
<td>33</td>
<td>45</td>
<td>79</td>
<td>62.67</td>
<td>9.99</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>47</td>
<td>66</td>
<td>54.4</td>
<td>5.62</td>
</tr>
</tbody>
</table>

2. 2D correlation plots between all possible band combinations and observed LAI

2D correlation plot for MSAVI (Subset H)

2D correlation plot for MSAVI (Subset L)
ESTIMATING LEAF AREA INDEX OF SALT MARSH VEGETATION USING AIRBORNE HYPERSPECTRAL DATA

2D correlation plot for MSAVI (Subset E)

2D correlation plot for NDVI (Subset A)

2D correlation plot for NDVI (Subset H)

2D correlation plot for NDVI (Subset L)

2D correlation plot for NDVI (Subset E)
3. Results of genetic algorithm for variable selection

The fitness versus variable usage plots for three subsets are shown in the following figures. Please refer section 3.3 for the details about these plots.

‘Fitness versus variable usage plot’ for subset H

Selected variables (in nm): 513, 571, 659, 689, 718, 833, 862, 891 and 1622

‘Fitness versus variable usage plot’ for subset L

Selected variables (in nm): 542, 630, 774, 948, 975

‘Fitness versus variable usage plot’ for subset E

Selected variables (in nm): 484, 513, 630, 659, 746, 774 and 804
4. Selection of latent variables in partial least square regression

The results of the partial least square regression utilizing full spectrum and utilizing only those variables selected by genetic algorithm are shown in the following tables. The first column shows the number of latent variables included in the models. Second and fourth columns show the variation of the predictor and predicted variables respectively captured by the particular latent variable. Similarly, third and fifth columns relate to the total variance of the predictor and predicted variables respectively captured by the partial least square models. The final column shows the cross validation error related with the number of latent variables in the PLSR models. The number of finally included latent variables and the associated attributes are highlighted with bold fonts.

CV-RMSE associated with different number of latent variables with full spectrum data (subset H).

<table>
<thead>
<tr>
<th>Current</th>
<th>Captured variance (X)</th>
<th>Captured variance (Y)</th>
<th>Current Total</th>
<th>Total CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.27</td>
<td>13.27</td>
<td>30.5</td>
<td>30.5</td>
</tr>
<tr>
<td>2</td>
<td>78.35</td>
<td>91.62</td>
<td>2.31</td>
<td>32.81</td>
</tr>
<tr>
<td>3</td>
<td>7.13</td>
<td>98.75</td>
<td>8.33</td>
<td>41.14</td>
</tr>
<tr>
<td>4</td>
<td>0.74</td>
<td>99.49</td>
<td>8.51</td>
<td>49.65</td>
</tr>
<tr>
<td>5</td>
<td><strong>0.18</strong></td>
<td><strong>99.67</strong></td>
<td><strong>14.27</strong></td>
<td><strong>63.92</strong></td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td>99.72</td>
<td>3.04</td>
<td>66.95</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>99.77</td>
<td>2.85</td>
<td>69.8</td>
</tr>
<tr>
<td>8</td>
<td>0.09</td>
<td>99.86</td>
<td>0.86</td>
<td>70.66</td>
</tr>
<tr>
<td>9</td>
<td>0.04</td>
<td>99.9</td>
<td>1.01</td>
<td>71.68</td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>99.91</td>
<td>1.52</td>
<td>73.19</td>
</tr>
</tbody>
</table>

CV-RMSE associated with different number of latent variables with full spectrum data (subset L).

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Captured variance (X)</th>
<th>Captured variance (Y)</th>
<th>Current Total</th>
<th>Total CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>96.39</strong></td>
<td><strong>96.39</strong></td>
<td><strong>0.67</strong></td>
<td><strong>0.67</strong></td>
</tr>
<tr>
<td>2</td>
<td>2.28</td>
<td>98.67</td>
<td>7.82</td>
<td>8.49</td>
</tr>
<tr>
<td>3</td>
<td>0.92</td>
<td>99.59</td>
<td>4.7</td>
<td>13.2</td>
</tr>
<tr>
<td>4</td>
<td>0.12</td>
<td>99.71</td>
<td>15.06</td>
<td>28.25</td>
</tr>
<tr>
<td>5</td>
<td>0.13</td>
<td>99.85</td>
<td>5.14</td>
<td>33.4</td>
</tr>
<tr>
<td>6</td>
<td>0.09</td>
<td>99.93</td>
<td>8.21</td>
<td>41.6</td>
</tr>
<tr>
<td>7</td>
<td>0.02</td>
<td>99.96</td>
<td>14.16</td>
<td>55.77</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>99.96</td>
<td>8.07</td>
<td>63.84</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>99.97</td>
<td>3.91</td>
<td>67.75</td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>99.97</td>
<td>3.73</td>
<td>71.47</td>
</tr>
</tbody>
</table>

CV-RMSE associated with different number of latent variables using selected variables from genetic algorithm (subset A)
### CV-RMSE associated with different number of latent variables using selected variables from genetic algorithm (subset H)

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Captured variance (X) Current</th>
<th>Captured variance (X) Total</th>
<th>Captured variance (Y) Current</th>
<th>Captured variance (Y) Total</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.54</td>
<td>27.54</td>
<td>10.08</td>
<td>10.08</td>
<td>1.12</td>
</tr>
<tr>
<td>2</td>
<td>67.63</td>
<td>95.17</td>
<td>1.96</td>
<td>12.04</td>
<td>0.94</td>
</tr>
<tr>
<td>3</td>
<td>4.43</td>
<td>99.6</td>
<td>10.75</td>
<td>22.79</td>
<td>0.99</td>
</tr>
<tr>
<td>4</td>
<td><strong>0.17</strong></td>
<td><strong>99.77</strong></td>
<td><strong>40.57</strong></td>
<td><strong>63.36</strong></td>
<td><strong>0.7</strong></td>
</tr>
<tr>
<td>5</td>
<td>0.13</td>
<td>99.9</td>
<td>2.09</td>
<td>65.45</td>
<td>0.77</td>
</tr>
<tr>
<td>6</td>
<td>0.03</td>
<td>99.92</td>
<td>0.82</td>
<td>66.28</td>
<td>0.77</td>
</tr>
<tr>
<td>7</td>
<td>0.04</td>
<td>99.97</td>
<td>0.15</td>
<td>66.43</td>
<td>0.77</td>
</tr>
<tr>
<td>8</td>
<td>0.02</td>
<td>99.99</td>
<td>0</td>
<td>66.44</td>
<td>0.79</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
<td>100</td>
<td>0</td>
<td>66.44</td>
<td>0.76</td>
</tr>
</tbody>
</table>

### CV-RMSE associated with different number of latent variables using selected variables from genetic algorithm (subset L)

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Captured variance (X) Current</th>
<th>Captured variance (X) Total</th>
<th>Captured variance (Y) Current</th>
<th>Captured variance (Y) Total</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>97.61</td>
<td>97.61</td>
<td>0.65</td>
<td>0.65</td>
<td>1.29</td>
</tr>
<tr>
<td>2</td>
<td>1.4</td>
<td>99.01</td>
<td>11.55</td>
<td>12.2</td>
<td>1.34</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>99.75</td>
<td>5.69</td>
<td>17.89</td>
<td>1.38</td>
</tr>
<tr>
<td>4</td>
<td>0.23</td>
<td>99.99</td>
<td>10.21</td>
<td>28.09</td>
<td>1.3</td>
</tr>
<tr>
<td>5</td>
<td><strong>0.01</strong></td>
<td><strong>100</strong></td>
<td><strong>10.82</strong></td>
<td><strong>38.91</strong></td>
<td><strong>1.21</strong></td>
</tr>
</tbody>
</table>

### CV-RMSE associated with different number of latent variables using selected variables from genetic algorithm (subset E)

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Captured variance (X) Current</th>
<th>Captured variance (X) Total</th>
<th>Captured variance (Y) Current</th>
<th>Captured variance (Y) Total</th>
<th>CV-RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92.98</td>
<td>92.98</td>
<td>13.13</td>
<td>13.13</td>
<td>0.86</td>
</tr>
<tr>
<td>2</td>
<td>6.69</td>
<td>99.66</td>
<td>26.78</td>
<td>39.91</td>
<td>0.75</td>
</tr>
<tr>
<td>3</td>
<td>0.06</td>
<td>99.72</td>
<td>40.44</td>
<td>80.35</td>
<td>0.8</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
<td>99.8</td>
<td>8.27</td>
<td>88.62</td>
<td>0.59</td>
</tr>
<tr>
<td>5</td>
<td>0.08</td>
<td>99.88</td>
<td>2.14</td>
<td>90.75</td>
<td>0.56</td>
</tr>
<tr>
<td>6</td>
<td><strong>0.08</strong></td>
<td><strong>99.96</strong></td>
<td><strong>0.09</strong></td>
<td><strong>90.85</strong></td>
<td><strong>0.54</strong></td>
</tr>
<tr>
<td>7</td>
<td>0.04</td>
<td>100</td>
<td>0.01</td>
<td>90.86</td>
<td>0.58</td>
</tr>
</tbody>
</table>
5. Scatter plots of predicted versus observed LAI

The scatter plots of predicted LAI versus observed LAI for all the predicting models are presented below. The coefficients of determination ($R^2$) between observed and predicted LAI and cross validated RMSE (CV-RMSE) of each model are also provided in the scatter plot.

Narrow band optimal MSAVI

\[
\begin{align*}
\text{Subset E} & \quad R^2 = 0.38 \\
& \quad \text{CV-RMSE} = 0.69 \\
\text{Subset H} & \quad R^2 = 0.35 \\
& \quad \text{CV-RMSE} = 0.91 \\
\text{Subset A} & \quad R^2 = 0.32 \\
& \quad \text{CV-RMSE} = 1.13 \\
\text{Subset L} & \quad R^2 = 0.05 \\
& \quad \text{CV-RMSE} = 1.36
\end{align*}
\]
Narrow band optimal NDVI

Subset A

\[ R^2 = 0.33 \]
\[ CV-RMSE = 1.14 \]

Subset H

\[ R^2 = 0.35 \]
\[ CV-RMSE = 0.86 \]

Subset L

\[ R^2 = 0.11 \]
\[ CV-RMSE = 1.24 \]

Subset E

\[ R^2 = 0.41 \]
\[ CV-RMSE = 0.63 \]

Partial least square regression with full spectrum

Subset A

\[ R^2 = 0.32 \]
\[ CV-RMSE = 1.05 \]

Subset H

\[ R^2 = 0.44 \]
\[ CV-RMSE = 0.78 \]
Partial least square regression with selected variables
Neural network with MNF components

Subset A

$R^2 = 0.35$
CV-RMSE = 1.03

Subset H

$R^2 = 0.0907$
CV-RMSE = 1.58

Subset L

$R^2 = 0.01$
CV-RMSE = 2.01

Subset E

$R^2 = 0.006$
CV-RMSE = 1.46

Neural network with principle components

Subset A

$R^2 = 0.37$
CV-RMSE = 0.97

Subset H

$R^2 = 0.12$
CV-RMSE = 1.27
ESTIMATING LEAF AREA INDEX OF SALT MARSH VEGETATION USING AIRBORNE HYPERSPECTRAL DATA

Subset L

\[ R^2 = 0.02 \]
\[ CV-RMSE = 2.12 \]

Subset E

\[ R^2 = 0.06 \]
\[ CV-RMSE = 1.24 \]