

## Comparative evaluation of inversion approaches of the radiative transfer model for estimation of crop biophysical parameters\*\*

Nilimesh Mridha<sup>1</sup>, Rabi N. Sahoo<sup>1\*</sup>, Vinay K. Sehgal<sup>1</sup>, Gopal Krishna<sup>1</sup>, Sourabh Pargal<sup>1</sup>, Sanatan Pradhan<sup>1</sup>, Vinod K. Gupta<sup>1</sup>, and Dasika Nagesh Kumar<sup>2</sup>

<sup>1</sup>Division of Agricultural Physics, Indian Agricultural Research Institute, New Delhi-110012, India

<sup>2</sup>Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, India

Received July 23, 2014; accepted January 26, 2015

**A b s t r a c t.** The inversion of canopy reflectance models is widely used for the retrieval of vegetation properties from remote sensing. This study evaluates the retrieval of soybean biophysical variables of leaf area index, leaf chlorophyll content, canopy chlorophyll content, and equivalent leaf water thickness from proximal reflectance data integrated broadbands corresponding to moderate resolution imaging spectroradiometer, thematic mapper, and linear imaging self scanning sensors through inversion of the canopy radiative transfer model, PROSAIL. Three different inversion approaches namely the look-up table, genetic algorithm, and artificial neural network were used and performances were evaluated. Application of the genetic algorithm for crop parameter retrieval is a new attempt among the variety of optimization problems in remote sensing which have been successfully demonstrated in the present study. Its performance was as good as that of the look-up table approach and the artificial neural network was a poor performer. The general order of estimation accuracy for parameters irrespective of inversion approaches was leaf area index > canopy chlorophyll content > leaf chlorophyll content > equivalent leaf water thickness. Performance of inversion was comparable for broadband reflectances of all three sensors in the optical region with insignificant differences in estimation accuracy among them.

**K e y w o r d s:** genetic algorithm, neural network, PROSAIL, leaf area index

### INTRODUCTION

Vegetation is a fundamental element of the earth surface and has a major influence on the exchange of energy between the atmosphere and the earth surface (Bacour *et al.*, 2002). Accurate quantitative estimation of vegetation biochemical and biophysical variables is useful for a large variety of agricultural, ecological, and meteorological applications (Houborg *et al.*, 2007). The direct measurement

of these characteristics is labour-intensive and costly, and is thus only practically possible on experimental plots of limited size. Because of its global coverage, repetitiveness, and non-destructive and relatively cheap characterization of land surfaces, remote sensing has been recognized as a reliable method and a practical means of estimating various biophysical and biochemical vegetation variables (Cohen *et al.*, 2003). Among the many vegetation characteristics, leaf area index (LAI) and leaf chlorophyll content ( $C_{ab}$ ) are of prime importance (Darvishzadeh *et al.*, 2008). Researchers (Rasmussen, 1997) showed that LAI is one of the key parameters widely used in crop monitoring, yield estimations, and ecosystem productivity models from regional to global scales. At the same time, monitoring spatial patterns in the biochemical composition of plant foliage, in particular chlorophyll and water content, is required for understanding growth dynamics in plant communities (Hilker *et al.*, 2011) and serve as bio-indicators of vegetation stress (Zarco-Tejada *et al.*, 2001).

Remote sensing measurement of plant biophysical parameters can broadly be classified into empirical and analytical/physical approaches (Hilker *et al.*, 2011). Both approaches have their advantages and disadvantages. Simplicity and computational efficiency of the empirical approach makes it highly desirable for large-scale remote sensing applications. However, the limits of this approach are obvious, such as the limited amount of spectral information, the diversified empirical equations used and their sensitivity to non-vegetation factors, and lack of generality (Fang *et al.*, 2005). Since canopy reflectance depends on a complex interaction of several internal and external factors that may

\*Corresponding author e-mail: rnsahoo.iari@gmail.com

\*\*This work was financial supported by Department of Science and Technology, Government of India under Project No. NRDMS/11/1669/2010.

vary significantly in time and space and from one crop type to another, spectral reflectance relationships will be site-, time-, and crop-specific, making the use of a single relationship for an entire region unfeasible (Colombo *et al.*, 2003).

Alternately, the analytical/physically-based models have proven to be a promising alternative as they describe the transfer and interaction of radiation inside the canopy based on physical laws and thus provide an explicit connection between the biophysical variables and the canopy reflectance. The earth and plant surface features reflect radiation anisotropically. Knowledge that the canopy biophysical characteristics are related to surface reflectance anisotropy (Bacour *et al.*, 2002) provides strong scientific basis for the application of these models (Roman *et al.*, 2011). The Bidirectional Reflectance Distribution Function (BRDF) approach taking care of non-isotropic radiation distribution is found to be more reliable in characterizing vegetation as well as retrieval of parameters. A number of canopy radiative transfer models (RTMs) of different complexities have been reported in literature, which simulate the bi-directional reflectance as a function of canopy characteristics (Verhoef, 1984). Among all RTMs, the PROSAIL model is the most popular one and widely applied, which describes both the spectral and directional variations of canopy reflectance as a function of leaf biochemistry and canopy architecture (Jacquemoud *et al.*, 2009). Therefore, the inversion of bi-directional canopy reflectance models emerged as a promising alternative for retrieval of biophysical parameters (Kimes *et al.*, 1998).

Model inversion, however, requires significant computational resources, which are slow on large data sets. This problem results from both the complex description of the radiative field within the canopy and the inversion method itself. Different inversion techniques have been proposed for physical models, including numerical optimization methods (Jacquemoud *et al.*, 1995), look-up table (LUT) approaches (Tripathi *et al.*, 2012), artificial neural networks (Walthall *et al.*, 2004), genetic algorithm (Fang *et al.*, 2003), and, very recently, support vector machines regression (Durbha *et al.*, 2007). In the iterative optimization approach, a stable and optimum inversion is not guaranteed, as the search algorithm may get trapped in local minima before reaching the global minimum. Moreover, the technique is computationally intensive, in particular when using complex radiative transfer models. This makes the retrieval of biophysical variables unfeasible for large geographic areas (Houborg *et al.*, 2007). The look-up tables and neural network methods are computationally more efficient than the traditional optimization approach, can be applied on a per pixel basis of satellite images, and they do not require any initial guesses. Moreover, they can be applied to the most sophisticated models without any simplifications. Although look-up table techniques may provide an efficient alternative, the definition of the cost function to be minimized still remains an open question when the uncertainties and their

structure are not very well known (Verger *et al.*, 2011). A limitation shared by all of the physically based models is the ill-posed nature of model inversion (Combal *et al.*, 2002), the fact that different combinations of canopy parameters may correspond to almost similar spectra. Therefore, there is a strong need for developing new advanced inversion techniques. While much work exists in the realm of retrieving biophysical parameters from spectral indices and from other model inversion methods, applications of genetic algorithms (GA) to a variety of optimization problems in remote sensing have only been demonstrated in recent years. The fundamental concept of GA is based on the concept of natural selection in the evolutionary process, which is accomplished by genetic recombination and mutation (Goldberg, 1989). Genetic algorithms have been developed for retrieval of land surface roughness and soil moisture (Wang and Jin, 2000). The most significant advantages of the GA are that it avoids the initial guess selection problem and provides a systematic scanning of the whole population and several acceptable local solutions such that a global optimum solution could be identified. Although the GA has been applied to various disciplines, there has been no work so far on applying GA for estimating crop biophysical parameters from either field measured reflectance or remotely sensed surface reflectance using the PROSAIL model. At the same time, there is still dearth of ample information on rigorous comparison of the various inversion methods in terms of accuracy and stability, computational time and number of variables obtainable (Liang, 2004).

Our objective was to perform a comparison of PROSAIL model inversion by look-up table (LUT), genetic algorithm (GA), and neural network (ANN) approaches and to simultaneously derive biophysical parameters such as leaf chlorophyll content ( $C_{ab}$ ), equivalent leaf water thickness ( $C_w$ ) and leaf area index (LAI) of soybean crop from field measured hyperspectral BRDF data.

## MATERIALS AND METHODS

A field experiment was conducted in the experimental farm of Indian Agricultural Research Institute, New Delhi, located at 28°38'23" North latitude and 77°09'27" East longitude with altitude of 228.6 meter above mean sea level. The climate is subtropical and semiarid characterized by hot dry summer and cold winter. The mean maximum temperature in the kharif season (June to September) ranges from 28.3 to 39.5°C and the mean minimum temperature from 12 to 28°C. Mean annual rainfall (30 years average) is 769.3 mm, of which 75% is received during the south west monsoon season between July to September and very little rain is received in rabi season.

Cultivar Pusa 9814 of soybean (*Glycine max* L.) which is of 125-day-duration was raised during the kharif season of 2011 in four plots each of size 6 x 6 m<sup>2</sup> with sufficient margins for irrigation channels and bunds under optimum

growing condition. Sowing of soybean was done on 2nd July 2011 by seed drill with row spacing of 45 cm. Soybean being a leguminous crop the nitrogen requirement was very low *ie* only 20 kg ha<sup>-1</sup> and being a kharif season (rainfed) crop only one extra irrigation of 3 cm depth was given at the vegetative growth stage. All the recommended doses of fertilizers were applied as basal. All standard plant protection measures and cultural practices were followed as recommended for the region.

The most popular canopy radiative transfer model PROSAIL (Jacquemoud *et al.*, 2009) was used in the present study to retrieve plant parameters. It is a combination of two models *ie* the PROSPECT model (Jacquemoud and Baret, 1990), which describes leaf optical properties, and the SAIL model (Verhoef, 1984), which characterizes canopy optical properties. The PROSAIL model considers the detailed information on leaf optical properties and also accounts for relative geometry of the source and sensor. By inverting the PROSAIL model, both the leaf and canopy parameters could be estimated.

The PROSAIL model needs measurement of different plant parameters like leaf area index (LAI), average leaf angle (LAD), leaf chlorophyll content ( $C_{ab}$ ), specific leaf weight ( $C_m$ ), equivalent leaf water ( $C_w$ ), leaf length, canopy height, and dry biomass to simulate soybean canopy reflectances as well as inversion of biophysical parameters. These parameters were measured following standard procedures synchronizing with the spectral observations taken at 52 DAS, 66 DAS, 84 DAS, and 100 DAS corresponding to the beginning bloom stage (R1), beginning pod stage (R3), beginning seed stage (R5), and full seed stage (R6) of soybean crop, respectively.

The bi-directional reflectance measurements at different relative azimuth and view zenith angles were taken using ASDI FieldSpec-3 hand held spectroradiometer with 10° Field of View (FOV) along with a field goniometer at four dates corresponding to the above mentioned growth stages of soybean crop. In case of soybean, the reflectance was measured in the spectral range of 350-2 500 nm at 36 relative azimuthal angles (relative to the azimuth angle of the sun) (0 to 350° at 10° interval) and in six zenith angles (20, 30, 40, 50, 60°, and nadir). In the principal plane (which is aligned to sun azimuth), 0° relative azimuth refers to backward scattering direction of light while 180° relative azimuth refers to forward scattering direction of light. A detail study was done to find the hotspot position among all possible combination of view zenith and azimuth angles as model inversion performed best at hotspot position.

The reflectance measurements at a 1 nm interval from spectro-radiometric were integrated to sensor broadband reflectances corresponding to optical bands of Terra MODIS, LandSat TM, and IRS LISS-3 sensors by using their respective band-wise relative spectral response (RSR)

curves. The broadband reflectances were used to the inversion approaches of LUT, GA, and ANN to retrieve corresponding soybean biophysical parameters.

In this study, mainly three inversion approaches of look-up table (LUT), genetic algorithm (GA), and artificial neural network (ANN) were used to retrieve biophysical parameters.

The simplest method of solving of a radiative transfer model is by the use of a LUT. The LUT was built in advance of the actual inversion through forward running of the PROSAIL model from a priori knowledge of the variation of the biophysical parameters. For the inversion, only search operations are needed to identify the parameter combinations that yield the best fit between measured and LUT spectra. However, to achieve high accuracy for the estimated parameters, the dimension of the LUT must be sufficiently large (Combal *et al.*, 2002). The range of free variables was defined by a priori knowledge from the field observation carried out during the experiment and as reported in literature. Out of total 14 variables, only three free variables of equivalent leaf water thickness, chlorophyll content, and leaf area index were varied to generate LUT. For retrieval of soybean biophysical parameters,  $C_w$  was varied from 0.01-0.05 cm at an interval of 0.001, chlorophyll from 30-100 µg cm<sup>-2</sup> at an interval of 1 µg cm<sup>-2</sup>, and LAI from 0.5-6.5 at an interval of 0.2. A total of 74124 combinations were generated. Parameters which did not vary much and were difficult to measure (Car, skyl) were fixed at model default values (Houborg *et al.*, 2007). The values of the other fixed parameter  $C_{brown}$  was taken as 0.15, N (structural parameter) as 1.6,  $\rho_{soil}$  as 0.2, and  $C_m$  as 0.008 g cm<sup>-2</sup> as per measurements. The average leaf angle (LAD) was set as measured in the field for each date given in Table 1. The view zenith angle was set at hotspot position *ie* 30° in the backscattering direction of the principal plane (*ie* relative azimuth of 0°). As during the course of the crop season the sun illumination geometry varied, four different LUTs were built, each corresponding to four specific dates of observation (Koetz *et al.*, 2005). The sun zenith angles were 20, 27, 30, and 37° corresponding to 52, 66, 84, and 100 DAS during the growing season of soybean crop.

In the LUT approach, mean of 10% solutions was used in parameter retrieval and mean squared error (MSE) was used as a cost function. The simple mechanism is that the calculated MSE between observed spectra and the LUT simulated spectra was sorted in ascending order *ie* lowest to highest. Then best (smallest MSE) 10% LUT entries were chosen and the mean of the parameter corresponding to this set was the final solution. The formula of MSE is given below:

$$MSE = \sum_{i=1}^n \frac{(R_{obs} - R_{sim})^2}{n}, \quad (1)$$

where:  $R_{obs}$  – observed reflectance,  $R_{sim}$  – simulated reflectance,  $n$  – number of observations.

**Table 1.** Soybean biophysical parameters at different dates of crop growth

Parameters	Soybean			
	52 DAS	66 DAS	84 DAS	100 DAS
Chlorophyll content ( $\mu\text{g cm}^{-2}$ )	52	60	65	60
Leaf area index ( $\text{m}^2 \text{m}^{-2}$ )	1.8	2.8	3.15	2.85
Average leaf angle ( $^\circ$ )	64	59	57	55
Leaf length/plant height	0.31	0.26	0.24	0.21
Specific leaf weight ( $\text{g cm}^{-2}$ )	0.007	0.008	0.009	0.007
Equivalent leaf water thickness (cm)	0.014	0.017	0.019	0.015

The genetic algorithm method is an iterative search algorithm based on an analogy with the process of natural selection (Darwinism) and evolutionary genetics. The search aims to optimize a user-defined function (the function to be optimized) called the fitness function, which calculates differences between user-supplied measured reflectance and model-simulated reflectances using statistical parameters like MSE. To perform this task, GA maintains a ‘population’ of candidate points, called ‘individuals’, over the entire search space. At each iteration, called a ‘generation’, a new population is created using ‘genetic operators’ such as mutation and crossover, which, from an algorithmic point of view, can be considered respectively as a means to change locally the current solutions and to combine them. This new generation generally consists of individuals that fit better than the previous ones on the basis of their fitness score into the external environment as represented by the fitness function. As the population iterates through successive generations, the individuals will in a general tend toward the optimum of the fitness function. What makes GA attractive is its ability to accumulate information about an initially unknown search space and to exploit this knowledge to guide subsequent search into useful sub-spaces. The fundamental implicit mechanism underlying this search consists of the combination of high-performance ‘building blocks’ discovered during past trials. Three important features distinguish the GA approach:

- GA works in parallel on a number of search points (potential solutions) and not on a unique solution, which means that the search method is not local in the scope but rather global over the search space;
- GA requires from the environment only a fitness function (objective function) measuring the fitness score of each individual and no other information nor assumptions such as derivatives and differentiability;
- both selection and recombination steps are performed by using probability rules rather than deterministic ones; this aims to maintain the global explorative properties of the search.

Separate programs for different growth stages of soybean crop were written in MATLAB for a fitness function for broadband reflectance (MODIS, TM, and LISS3). As a fitness function, mean squared error (MSE) and nash-sutcliffe efficiency (NSE) were used. NSE was successfully used in the generalized likelihood uncertainty estimation (GLUE) optimization method for different hydrological experiments (Jin *et al.*, 2010). Mathematical expression of nash-sutcliffe efficiency (NSE) is given below:

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}, \quad (2)$$

where:  $O_i$  – observed reflectance at specific wavelength band  $i$ ,  $P_i$  – simulated reflectance in the LUT at that wavelength band  $i$ ,  $\bar{O}$  – mean of observed reflectance,  $n$  – number of wavelength bands.

To carry out this technique, GA optimization toolbox in MATLAB was used for retrieval of soybean biophysical parameter. Among various GA parameters, ‘creation function’ was set as ‘feasible population’, ‘mutation function’ as ‘adaptive feasible’, ‘crossover function’ as ‘scattered’, and ‘selection type’ as ‘stochastic uniform’. During trials with GA parameters, it was observed that different GA parameters had a different degree of influence on the final result.

The value of best fitness decreases with the increasing number of generations up to 15, after which it was found to be almost constant. However, with the increase in the population size, the fitness value is reduced insignificantly. Trials with a crossover fraction did not show any specific trend, so a default value of 0.8 was used. In the GA, the major computational time was used in the GA optimization process although reducing the number of free parameters helps. In the GA, the space of initial conditions has to be scanned and a large number of iterations are needed to converge toward appropriate solutions. To solve this problem, the GA optimization process was integrated to the PROSAIL model to obtain desired solutions.

Artificial neural nets (ANN) are able to approximate very complex non-linear relationships (Kimes *et al.*, 1998). Once trained, ANN is very fast to handle a high volume of data taking very little computer storage and time. In this study, a feed-forward back propagation neural network with three layers (input, hidden, and output layer) was used. The topology of ANN was four or five nodes in the input layer for LISS-3 and ETM+ band reflectances, twenty nodes in the hidden layer and three nodes in the output layer corresponding to three biophysical parameters of  $C_{ab}$ ,  $C_w$ , and LAI. In the neurons of the hidden layer, a trans-sigmoid (TRANSIG) transfer function and in the neurons of the output layer, a linear transfer function were used as it is recognized as capable of fitting any type of function (Verger *et al.*, 2011). The Levenberg-Marquardt optimization was used for training the network because of its efficient convergence performance (Ngia and Sjoberg, 2000) and the adaptation learning function was LEARNGDM (gradient descent with momentum weight and bias learning function). There were enough neurons in the hidden layer because these are known to be universal function approximators (Atkinson and Tantnall, 1997). To perform the inversion of the canopy reflectance model using ANN, it was first trained by the simulated datasets of the PROSAIL model. The schematic flow diagram describing the steps involved in the LUT, GA, and ANN inversion approaches is shown in Fig. 1.

The look-up table of simulated values generated earlier with 74 124 cases was used for ANN training and validation. In this study, 70% of the total simulated set of entries was used for training *ie* adjusting of weights and the rest 30% was used for the validation purpose. A maximum of 1 000 epochs were allowed though ANN generally took around 200 epochs for learning. After completion of the training process, the sought biophysical parameters were calculated with the validation datasets. The validation was characterized on the basis of determination of coefficient ( $R^2$ ) and MSE (mean square error value) to evaluate the performance of training. Neural network toolbox available in MATLAB™ software was used to implement the inversion approach.

The biophysical parameters retrieved were LAI,  $C_{ab}$ ,  $C_w$ , and CCC *ie* canopy chlorophyll content, to picturize the chlorophyll content of the whole canopy. The CCC was calculated as a product of  $C_{ab}$  and LAI. These parameters retrieved through the LUT (mean of best 10% solution), GA, and ANN inversion approaches were compared with the measured one to check the performance of the inversion methods. Statistics of determination of coefficient ( $R^2$ ), root mean square error (RMSE), normalized RMSE (nRMSE), and ratio of deviation to prediction (RDP) between the measured and model-retrieved values were computed and used to evaluate the inversion methods. RDP is a simple ratio of standard deviation to RMSE, which shows performance of the model prediction. Model prediction is considered to be good when the RDP value is greater than 1 and poor when it is less than 1.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Pobs - Psim)^2}{n}}, \quad (3)$$

where: *Pobs* – observed parameter value, *Psim* – estimated parameter value, *n* – number of observations, nRMSE – RMSE/mean measured value, RDP – standard deviation/RMSE.

RESULTS

The soybean biophysical parameters measured at four dates respectively, which are given in Table 1, were used as input to the PROSAIL model for simulating bidirectional reflectance at those corresponding dates. The measurements of bi-directional measurements were taken at 52, 66, 84, and 100 DAS corresponding to the beginning bloom stage (R1), beginning pod stage (R3), beginning seed stage (R5), and full seed stage (R6) of soybean crop. This set of observed parameters could thus be a representative set to use for PROSAIL model validation accounting for a range of soybean growth conditions.

The performance of three inversion approaches of the look-up table, genetic algorithm, and artificial neural network were evaluated and compared in terms of their

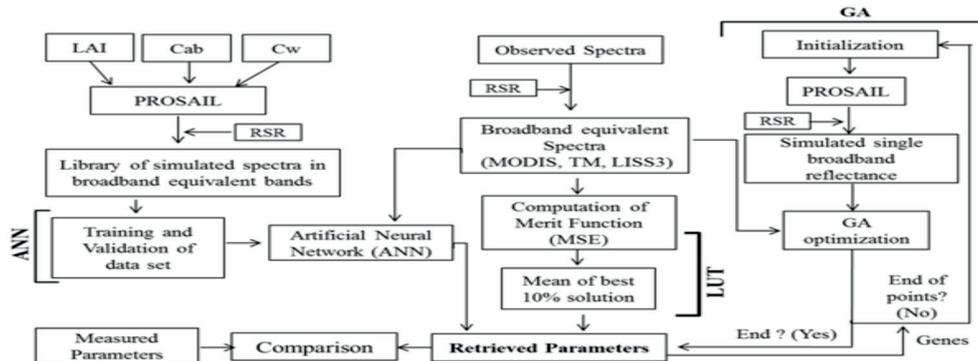


Fig. 1. Flow chart showing steps involved in the look-up table (LUT), genetic algorithm (GA), and artificial neural network (ANN) approaches.

performance to retrieve LAI, total chlorophyll content of leaf, and leaf water thickness of soybean crop. All the three approaches were tried for soybean crop on MODIS, TM, and LISS-3 measured reflectance and the results are presented and summarized showing comparison of all the three inversion approaches in Figs 2-4 and in Table 2.

In this approach, mean of 10% solutions was used in parameter retrieval. Corresponding to each observed spectra, the LUT was sorted on MSE in ascending order *ie* lowest to highest. Then best (smallest MSE) 10% LUT entries were chosen and the mean of the parameter corresponding to this set was the final solution. The results showed different levels of accuracy for different biophysical parameters (Table 2).

The observed values of LAI varied between 1.8- 3.15 and the estimated values varied between 2.1-3.07 for MODIS, 2.21-3.21 for TM, and 2.01-3.05 for LISS-3. There was overestimation at low LAI values and underestimation of high values. Retrieval of LAI was comparable for all three sensors with a correlation  $R^2=0.95$  and  $RMSE=0.1615$  for MODIS,  $R^2=0.94$  and  $RMSE=0.2174$  for TM, and  $R^2=0.94$  and  $RMSE=0.1448$  for LISS3 showing a strong relation between the observed and estimated values. The observed values of  $C_{ab}$  varied between 52-65  $\mu\text{g cm}^{-2}$  and the estimated values varied between: 54-68, 46-69, 48-69  $\mu\text{g cm}^{-2}$  for MODIS, TM, and LISS-3, respectively. The correlation was comparable for all three sensors.  $RMSE$  varied from 3.9 for LISS3 to 6 for TM (6.5 and 10% of the measured mean value, respectively). The observed values of CCC varied between 0.96-2.05  $\text{g m}^{-2}$  and the estimated values varied between: 1.18-2.05, 1.01-2.21, and 0.98-2.08  $\text{g m}^{-2}$  for MODIS, TM, and LISS-3, respectively. Among all the retrieved parameters, the highest correlation of 0.99 (0.0001) was observed for CCC. The observed values of  $C_w$  varied between 0.014-0.018 cm and the estimated values varied between: 0.015-0.02 cm for MODIS, 0.014-0.021 cm for TM, and 0.015-0.022 cm for LISS-3. The estimation of parameter  $C_w$  was poor, as the LUT approach could not capture the variations in the observed values. In this approach, PROSAIL predicted very well for LAI having  $RDP=3.5$ , which describes the performance of model prediction and very poorly for  $C_w$  with  $RDP$  of 0.4. CCC accounting for greater variability than  $C_{ab}$  showing better retrieval with a higher correlation and  $RDP$ . An overestimation was found for almost all values of LAI and  $C_w$ . In the case of  $C_{ab}$ , low values were underestimated and high values were overestimated.

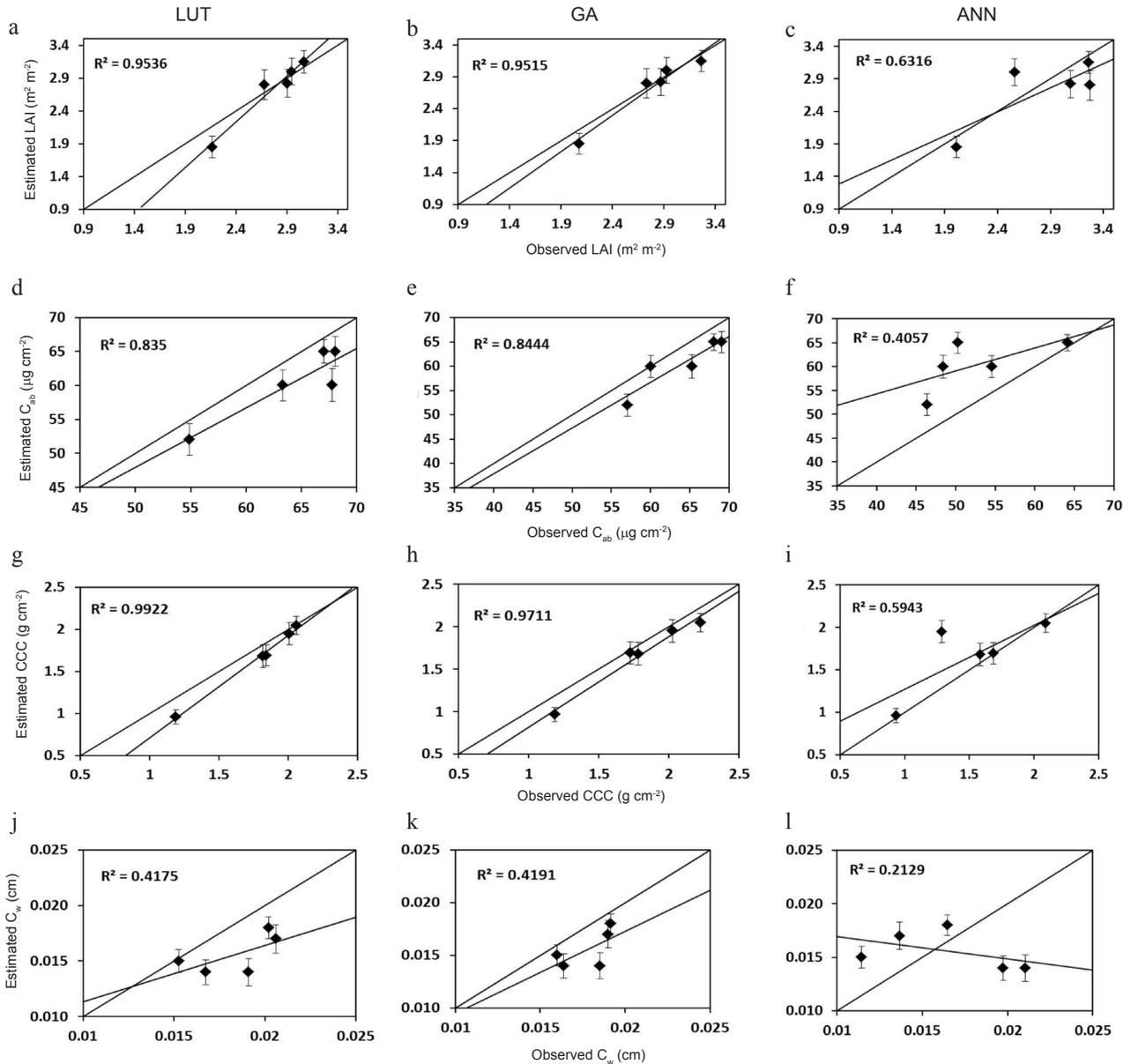
Overall, in the LUT approach, the order of accuracy estimation was  $LAI > CCC > C_{ab} > C_w$  for all the three sensors. The level of retrieval accuracy varied among the sensors without showing any significant differences in the accuracy.

In GA, one of the most important steps to have higher retrieval accuracy is the GA optimization process. The GA optimization process is a time consuming step, but inversion took very little time. The LAI estimation showed the

highest correlation ( $R^2$ ) of 0.96 ( $p=0.0009$ ) with an  $RMSE$  of 0.1448, which was 5% of the mean measured value for LISS3. The highest correlation  $R^2$  of  $C_{ab}$  was 0.86 ( $p=0.01$ ) with an  $RMSE$  of 6.1, which was 10% of the mean observed value for MODIS. CCC showed the highest correlation ( $R^2$ ) of 0.97 ( $p=0.0006$ ) among all the parameters, as found in the LUT approach. GA showed better retrieval values of  $C_w$  than the other two approaches for all the three sensors with correlation ( $R^2$ ) of 0.42 ( $p=0.21$ ) for MODIS, 0.52 ( $p=0.14$ ) for TM, and 0.51 ( $p=0.14$ ) for LISS3. The  $RDP$  value for model prediction varied from 4.1 for LAI estimation to 0.45 for  $C_w$  estimation, which showed that the GA approach could predict LAI,  $C_{ab}$ , and CCC well enough with a high level of accuracy like LUT. Like LUT, in the GA approach, the estimated values for all the parameters were closer to a 1:1 line (Figs 2-4). In the case of  $C_{ab}$ , low values were underestimated and high values were overestimated. An overestimation was found almost for all values of LAI and  $C_w$ . In the GA approach, the order of accuracy in estimation was  $LAI > CCC > C_{ab} > C_w$  for all the three sensors. The level of retrieval accuracy varied among the sensors without showing any significant differences in the accuracy.

The ANN took considerably more time in learning than the other two approaches, but its inversion time was quite low. The correlation coefficient of LAI was 0.63 ( $p=0.08$ ) with an  $RMSE$  of 0.32, which was 12% of the mean observed value. The  $C_{ab}$  estimation showed a correlation of 0.47 ( $p=0.18$ ) with an  $RMSE$  of 9, which was 15% of the observed mean value. ANN underestimated all the values of  $C_{ab}$  and  $C_w$  and overestimated all the values of LAI. The ANN approach performed very poorly for  $C_w$  as indicated by the insignificant value of the correlation coefficient. Unlike LUT and GA, the ANN could not capture variations in the observed CCC as indicated by a correlation of 0.59 ( $p=0.1$ ) with a high  $RMSE$  value, which was 20% of the mean observed value. The  $RDP$  value showed that ANN could predict LAI with considerable error (high  $RMSE$ ) but could not predict  $C_{ab}$  and  $C_w$ . In this approach, the estimated values for all the parameters were scattered to 1:1 line (Figs 2-4), which showed a poor performance of this technique to retrieve the biophysical parameters of soybean.

The use of ANN as an inversion technique resulted in relatively poorer results than LUT and GA. Overall, in all the three approaches, the order of estimation accuracy was  $LAI > CCC > C_{ab} > C_w$ . Performance of all the three inversion methods for all the three sensors using  $R^2$ ,  $RMSE$ ,  $nRMSE$ , and  $RDP$  is shown in Table 2 and a comparison among the three inversion approaches are graphically shown through the scatter plot of the observed versus estimated values of LAI,  $C_{ab}$ , CCC, and  $C_w$  in Fig. 2 for MODIS, in Fig. 3 for TM, and in Fig. 4 for LISS3.



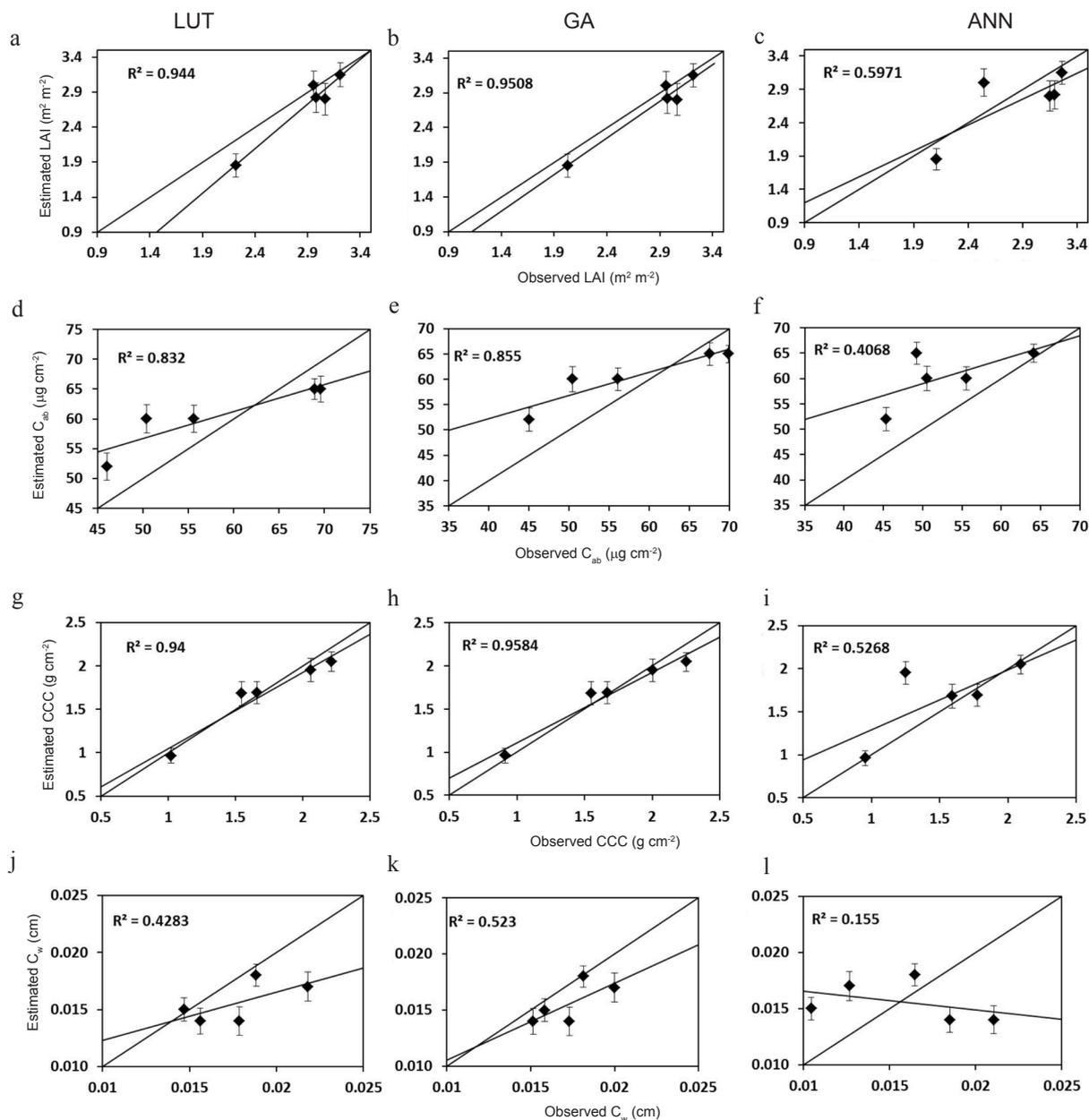
**Fig. 2.** Scatter plot of observed versus retrieved values by: LUT (a, d, g, j); GA (b, e, h, k); and ANN (c, f, i, l) inversions for MODIS in soybean crop.

## DISCUSSION

Inversion of physics-based radiative transfer models is an area of rapid development in remote sensing of earth surface to derive various biophysical parameters (Liang, 2007). Among the canopy radiative transfer models, PROSAIL is one of the most popular models that have been inverted to derive canopy biophysical variables from hyperspectral and multispectral remote sensing observations (Darvishzadeh *et al.*, 2008; Vohland *et al.*, 2010; Verger *et al.*, 2011). A rigorous comparison of the accuracy of various inversion methods for physically based models is very rare in the literature. There are several comparisons of the accuracy and computational efficiencies of various traditional inversion methods (Renders *et al.*, 1992).

However, comparisons between the look-up table methods, genetic algorithm, and neural network methods and particularly in soybean crop are absent in the literature. Consequently, the relative accuracy that can be obtained with each method cannot be stated until more comparisons are made. Therefore, this study was undertaken to evaluate the performance of PROSAIL inversions by LUT, GA, and ANN for MODIS, TM, and IRS LISS-3 broadband reflectances for soybean crop.

The general order of estimation accuracy among the parameters in all the inversion approaches was  $LAI > CCC > C_{ab} > C_w$ . The higher accuracy of LAI estimation by all the approaches may be due to the fact that structural variables (*eg* LAI, LAD) determine the total canopy reflectance of

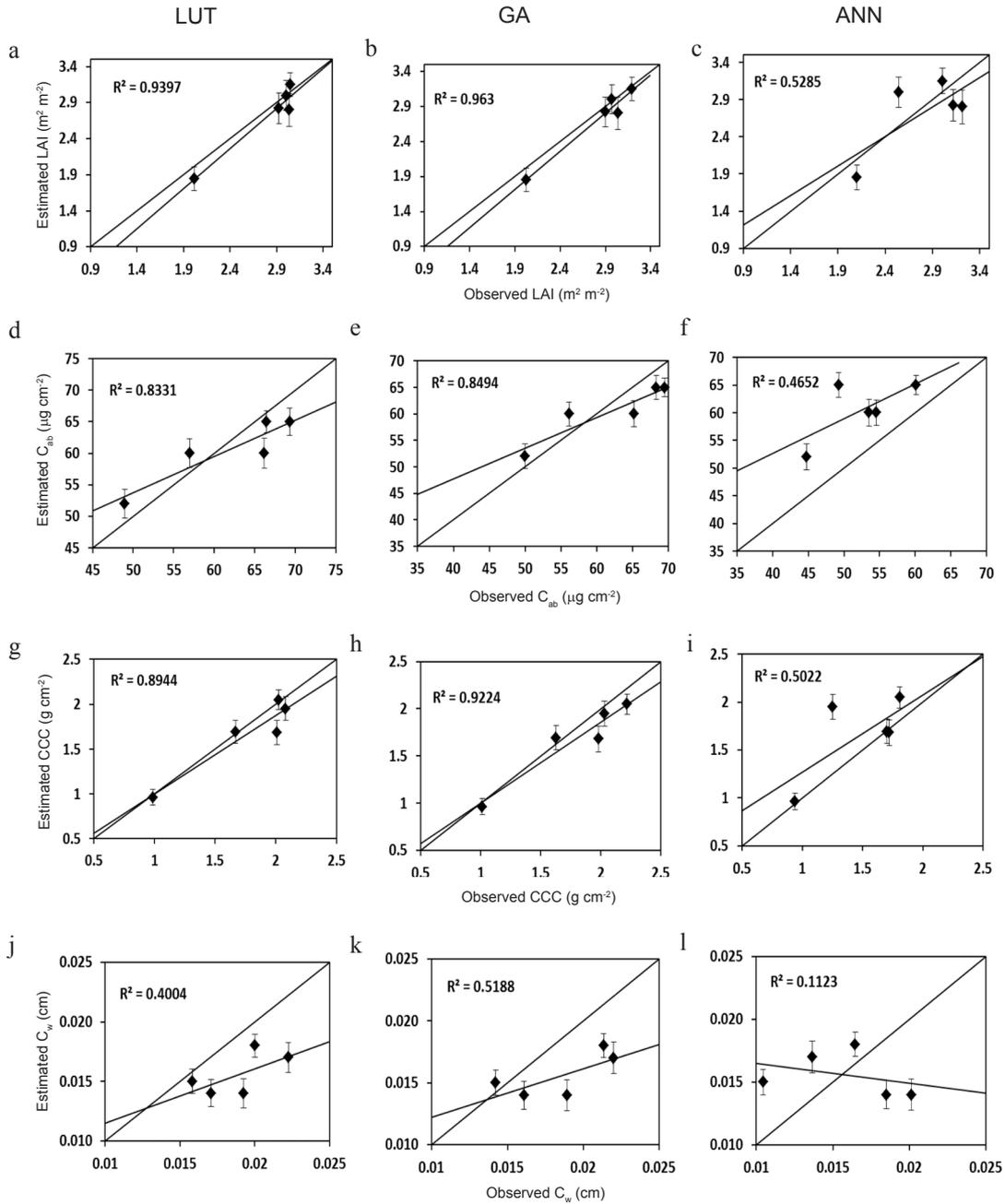


**Fig. 3.** Scatter plot of observed versus retrieved values by LUT (a, d, g, j), GA (b, e, h, k), and ANN (c, f, i, l) inversions for TM in soybean crop.

crops much more significantly than biochemical variables. The high sensitivity of Red and NIR bands and the moderate sensitivity of Green bands to changes in LAI (Jacquemoud *et al.*, 2009) may have resulted in better simulation of the reflectance spectra by PROSAIL leading to relatively accurate inversion.

The relationships between the measured and estimated leaf chlorophyll  $C_{ab}$  were poorer than for LAI in all the inversion processes as indicated by the lower  $R^2$  values and higher nRMSE. This result is in line with results of previous studies (Darvishzadeh *et al.*, 2008). It is argued that there

is always poor signal propagation from the leaf to canopy scale resulting in poor estimation of leaf biochemical parameters by canopy reflectance (Jacquemoud *et al.*, 1996). Moreover, only the VIS band is sensitive to leaf chlorophyll variation and this band reflectance has a very low dynamic range due to dominance of absorption. Therefore, there is more chance of error in PROSAIL simulation of reflectance in the VIS band, leading to poorer estimate of leaf chlorophyll. A number of studies showed increased robustness and accuracy when estimates of biochemical variables were integrated at the canopy level (*eg* canopy



**Fig. 4.** Scatter plot of observed versus retrieved values by LUT (a, d, g, j), GA ( b, e, h, k), and ANN (c, f, i, l) inversions for LISS 3 in soybean crop.

chlorophyll,  $C_{ab} \times LAI$ ) rather than at the leaf level (Vohland *et al.*, 2010). This means that leaf scale results for chlorophyll are generally inferior to those at the canopy level. We have found similar results, as the accuracy of CCC was better to that of  $C_{ab}$ . In the case of the leaf water content, all the three inversion approaches failed as indicated by the non-significant  $R^2$  values though GA performed better in  $C_w$  retrieval among the three.

GA performed well in retrieval like LUT. The ANN approach underperformed as compared to LUT and GA, even though the ANN training time was considerably

large and the process computationally intensive. The ANN showed underestimation in all the parameters. These results are in conformity with the results reported by Vohland *et al.* (2010). They also found that ANN performance was poor as compared to numerical optimization and LUT.

Comparing the results of MODIS, TM, and LISS-3, very small differences in the level of retrieval accuracy of the biophysical parameters were observed. This may be due to the fact that the parameters have negligible sensitivity to those extra bands included in MODIS and TM.

**Table 2.** Comparison of different model inversion techniques for retrieval of soybean biophysical parameters

Parameters	LAI ( $m^2 m^{-2}$ )			$C_{ab}$ ( $\mu g cm^{-2}$ )			CCC ( $g m^{-2}$ )			$C_w$ ( $g cm^{-2}$ )		
	MODIS	TM	LISS3	MODIS	TM	LISS3	MODIS	TM	LISS3	MODIS	TM	LISS3
LUT	0.9536	0.9440	0.9397	0.8350	0.8320	0.8331	0.9922	0.9400	0.8944	0.4175	0.4283	0.4004
GA	0.9515	0.9508	0.9630	0.8444	0.8550	0.8494	0.9711	0.9584	0.9224	0.4191	0.5230	0.5188
ANN	0.6316	0.5971	0.5285	0.4057	0.4068	0.4652	0.5943	0.5268	0.5022	0.2129	0.1550	0.1123
LUT	0.0014	0.0021	0.0024	0.0176	0.0182	0.0180	0.0000	0.0024	0.0073	0.2165	0.2082	0.2299
GA	0.0016	0.0016	0.0009	0.0157	0.0136	0.0147	0.0006	0.0011	0.0039	0.2152	0.1440	0.1465
ANN	0.0859	0.1026	0.1406	0.2257	0.2248	0.1815	0.1041	0.1416	0.1569	0.4186	0.4994	0.5711
LUT	0.1615	0.2174	0.1448	4.3453	6.0468	3.9297	0.1382	0.1114	0.1596	0.0032	0.0029	0.0037
GA	0.1242	0.1621	0.1384	3.9905	6.1035	3.9225	0.1394	0.1137	0.1622	0.0026	0.0021	0.0036
ANN	0.3255	0.3316	0.3326	9.1029	8.9688	8.8984	0.3007	0.3171	0.3304	0.0047	0.0047	0.0043
LUT	0.0593	0.0798	0.0531	0.0717	0.0998	0.0648	0.0829	0.0668	0.0958	0.2026	0.1824	0.2354
GA	0.0456	0.0595	0.0508	0.0658	0.1007	0.0647	0.0837	0.0682	0.0973	0.1615	0.1326	0.2286
ANN	0.1195	0.1217	0.1221	0.1502	0.1480	0.1468	0.1805	0.1903	0.1983	0.2954	0.2993	0.2713
LUT	3.1517	2.3418	3.5163	1.2243	0.8798	1.3537	3.0762	3.8167	2.6629	0.5134	0.5702	0.4418
GA	4.0988	3.1404	3.6791	1.3331	0.8716	1.3562	3.0489	3.7399	2.6204	0.6439	0.7842	0.4550
ANN	1.5638	1.5352	1.5307	0.5844	0.5931	0.5978	1.4135	1.3404	1.2865	0.3521	0.3475	0.3833

## CONCLUSIONS

1. This work evaluated three inversion approaches of the look-up table (mean of 10% solutions), genetic algorithm, and artificial neural network for concurrent retrieval of biophysical parameters of leaf area index, total chlorophyll content of leaf, canopy chlorophyll content, and leaf water thickness of soybean.

2. All approaches could capture the variability in the measured soybean biophysical parameters though the accuracy of estimation varied among the parameters. The order of estimation accuracy among the parameters in all the three approaches was leaf area index > canopy chlorophyll content > leaf chlorophyll content > leaf water thickness.

3. The performance of the genetic algorithm was similar to the look-up table approach, and the artificial neural network showed poor performance. None of the approaches could estimate variation in leaf water thickness significantly though the correlation and estimation accuracy for leaf water thickness was better in the genetic algorithm than in the look-up table and artificial neural network. Thus, the genetic algorithm optimization method may provide an alternative to invert the radiative transfer models in remote sensing with better understanding of coupling of the genetic algorithm to radiative transfer models.

4. The advantage of the genetic algorithm is twofold. First, it scans all the initial conditions and provides several possible solutions for the detailed examination of the global optimum solution, thus it avoids the inaccuracies introduced by traditional minimization algorithms. Second, it only runs the forward radiative transfer model with constrained parameter space and is straightforward in the optimization process. Experiments are needed to test this method in more complicated areas.

## REFERENCES

- Atkinson P.M. and Tatnall A.R.L., 1997. Neural networks in remote sensing. *Int. J. Remote Sensing*, 18(4), 699-709.
- Bacour C., Jacquemoud S., Leroy M., Hauteceur O., Weiss M., and Prevot L., 2002. Reliability of the estimation of vegetation characteristics by inversion of three canopy reflectance models on airborne polder data. *Agronomie: Agric. Environ.*, 22, 555-565.
- Cohen W.B., Maieringer T.K., Gower S.T. and Turner D.P., 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing Environ.*, 84(4), 561-571.
- Colombo R., Bellingeri D., Fasolini D., and Marino C.M., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing Environ.*, 86(1), 120-131.
- Combal B., Baret F., Weiss M., Trubull A., Macc D., and Pragnere A., 2002. Retrieval of canopy biophysical variables from bi-directional reflectance using prior information to solve the ill-posed inverse problems. *Remote Sensing Environ.*, 84, 1-15.
- Darvishzadeh R., Skidmore A., Schlerf M., and Atzberger C., 2008. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in heterogeneous grassland. *Remote Sensing Environ.*, 112, 2592-2604.
- Durbha S.S., King R.L., and Younan N.H., 2007. Support vector machines regression for retrieval of leaf area index from multi-angle imaging spectroradiometer. *Remote Sensing Environ.*, 107, 348-361.
- Fang H. and Liang S., 2005. A hybrid inversion method for mapping leaf area index from MODIS data: Experiments and application to broadleaf and needle leaf canopies. *Remote Sensing Environ.*, 94, 405-424.
- Fang H., Liang S., and Kuusk A., 2003. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing Environ.*, 85, 257-270.
- Goldberg D.E., 1989. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison-Wesley.
- Hilker T., Lepine L., Coops N.C. Jassal R.S., Black T.A., Wulder M.A., Ollinger S., Tsui O., and Day M., 2011. Assessing the impact of N-fertilization on biochemical composition and biomass of a Douglas-fir canopy – A remote sensing approach. *Agric. Forest Meteorol. Agric. Forest Meteorol.*, 153,124-133.
- Houborg R., Soegaard H., and Boegh E., 2007. Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data. *Remote Sensing Environ.*, 106(1), 39-58.
- Jacquemoud S., Bacour C., Poilve H., and Frangi J.P., 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing Environ.*, 74(3), 471-481.
- Jacquemoud S. and Baret F., 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing Environ.*, 34(2), 75-91.
- Jacquemoud S., Baret F., Andrieu B., Danson F.M., and 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 Sens. Environ.*, 52(3), 163-172.
- Jacquemoud S., Ustin S.L., Verdebout J., Schmuck G., Andreoli G., and Hosgood B., 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing Environ.*, 56, 194-202.
- Jacquemoud S., Verhoef W., Baret F., Bacour C., Zarco-Tejada P.J., Asner G.P., François C., and Ustin S.L., 2009. PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing Environ.*, 113, 56-66.
- Jin X., Xu C., Zhang Q., and Singh, V.P., 2010. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *J. Hydrol.*, 383, 147-155.
- Kimes D.S., Nelson R.F., Manry M.T., and Fung A.K., 1998. Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements. *Int. J. Remote Sens.*, 19(14), 2639-2662.
- Koetz B., Baret F., Poilve H., and Hill J., 2005. Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics. *Remote Sensing Environ.*, 95, 115-124.

- Liang S., 2007.** Recent developments in estimating land surface bio geophysical variables from optical remote sensing. *Progress Physical Geography*, 31(5), 501-516.
- Ngia L.S.H. and Sjoberg J., 2000.** Efficient training of neural nets for nonlinear a adaptive filtering using a recursive Levenberg-Marquardt algorithm. *IEEE Trans. Signal Proc.*, 48, 1915-1927.
- Rasmussen M., 1997.** Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. *Int. J. Remote Sensing*, 18(5), 1059-1077.
- Renders J-M., Flasse S.P., Verstraete M.M., and Nordvik J-P., 1992.** A comparative study of optimization methods for the retrieval of quantitative information from satellite data. Joint Research Center Report EUR 14851, Brussels.
- Roman M.O., Gatebe C.K., Schaaf C.B., Poudyal R., Wang Z., and King M.D., 2011.** Variability in surface BRDF at different spatial scales (30 m – 500 m) over a mixed agricultural landscape as retrieved from airborne and satellite spectral measurements. *Remote Sensing Environ.*, 115, 2184-2203.
- Tripathi R., Sahoo R.N., Sehgal V.K., and Tomar R.K., and Chakraborty Nagarajan S., 2012.** Inversion of PROSAIL Model for Retrieval of Plant Biophysical Parameters *J. Indian Soc. Remote Sensing*, 40(1), 19-28.
- Verger A., Baret F., and Camacho F., 2011.** Optimal modalities for radiative transfer neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations. *Remote Sensing Environ.*, 115, 415-426.
- Verhoef W., 1984.** Light scattering by leaf layers with application to canopy reflectance modelling: The SAIL model. *Remote Sensing Environ.*, 16(2), 125-141.
- Vohland M., Mader S., and Dorigo W., 2010.** Applying different inversion techniques to retrieve stand variables of summer barley with PROSPECT + SAIL. *Int. J. Appl. Earth Observation Geoinformation*, 12, 71-80.
- Walthall C., Dulaney W., Anderson M., Norman J., Fang H., and Liang S., 2004.** A comparison of empirical and neural network approaches for estimating corn and soybean leaf area index from Landsat ETM+ imagery. *Remote Sensing Environ.*, 92(4), 465-474.
- Wang Y. and Jin Y., 2000.** A genetic algorithm to simultaneously retrieve land surface roughness and soil moisture. *J. Remote Sensing (Chinese)*, 4(2), 90-94.
- Zarco-Tejada P.J., Miller J.R., Mohammed G.H., Noland T.L., and Simpsom P.H., 2001.** Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Trans. Geoscience Remote Sensing*, 39(7), 1491-1507.