SIMULTANEOUS RETRIEVAL OF SURFACE PARAMETERS BY MODEL INVERSION

Wout Verhoef
National Aerospace Laboratory NLR
P.O. Box 153
8300 AD Emmeloord
The Netherlands
verhoef@nrl.nl

Working Group VII/1

KEY WORDS: Optical, Reflectance models, Model inversion, Parameter retrieval, Red edge, Hyperspectral data, Imaging spectrometers

ABSTRACT

The estimation of biophysical parameters of vegetation canopies from optical remote sensing data is very important for the study of land surface processes. Traditionally, this has been realised mostly by means of spectral indices, e.g. the normalised difference vegetation index (NDVI), which is derived from spectral reflectances or radiances in the near infrared and red parts of the spectrum. Spectral indices can be applied very easily and they can have a fairly high correlation with a surface parameter such as fractional vegetation cover. However, the spectral reflectance of vegetation canopies is influenced by many factors simultaneously, so a good relationship between a spectral index and a single variable is not very likely to be found when other parameters vary at the same time. Therefore, it is expected that simultaneous retrieval of all parameters that are known to have an effect will be more accurate.

In a numerical modelling case study it has been investigated whether it would be possible to retrieve from optical remote sensing data all the (bio)physical parameters of the coupled soil-vegetation-atmosphere system that are expected to have an effect on spectral radiances detected by advanced hyperspectral spaceborne sensors. For this, optical spectral data on single leaves generated by means of the PROSPECT model (Jacquemoud & Baret, 1990) have been applied in the integrated optical soil-canopy-atmosphere radiation model OSCAR (Verhoef, 1998). The influences of two soil parameters, two leaf parameters, four canopy parameters and three atmospheric parameters on hyperspectral directional planetary reflectances have been simulated in a numerical model inversion experiment. The most important parameters to be retrieved were soil brightness, leaf chlorophyll content, canopy leaf area index and atmospheric visibility. The simultaneous retrieval of the eleven parameters has been investigated using classical model inversion by means of the Gauss-Newton method of non-linear least squares parameter estimation. The results indicate that this approach has potential, as in a number of widely differing cases the retrieval of all model parameters from 10 nm resolution hyperspectral red edge planetary reflectance data under five directions was successful.

1 INTRODUCTION

The estimation of biophysical parameters of vegetation canopies from optical remote sensing data is very important for the study of land surface processes. Traditionally, this has been realised mostly by means of spectral indices, e.g. the normalised difference vegetation index (NDVI), which is derived from spectral reflectances or radiances in the near infrared and red parts of the spectrum. Spectral indices have the advantage that they can be applied very easily and that they sometimes have a fairly high correlation with a surface parameter such as fractional vegetation cover. However, the spectral reflectance of vegetation canopies is influenced by many factors simultaneously, so it is not very likely that one can obtain a good relationship between a spectral index and a single variable when other parameters vary at the same time.

In the previous SASSSIS study (Verhoef & Menenti, 1998) it was shown by means of model simulations that canopy LAI might be retrieved from multiple linear regression equations based on TOA multispectral (i.e. 7 bands) radiance observations under 5 different viewing directions at an accuracy of about 12%. In this study, most variations found in reality were included, such as variation of soil type, leaf type, canopy structure, solar zenith angle and atmospheric conditions. However, it is known that the relations between surface parameters and spectral radiances are non-linear, so better retrieval results should be possible when taking these non-linear relationships into account.

The reason that the red and the near infrared are used so often to map green vegetation is that the absorption of radiation in a green leaf canopy is so different in both wavelength regions. In the red there is a strong absorption due to
chlorophyll in the leaves, whereas in the near infrared there is almost no absorption at all. When the leaves are green,
this great difference in absorption leads to a strong spectral response due to variation of LAI over a spectrally nearly flat
soil background. This behaviour is the basis under any red – near infrared vegetation index. However, what would
happen if the leaves turned yellow due to chlorophyll demolition? In that case there would be a strong spectral response
as well, but it might easily be mistaken for a large decrease in canopy LAI.

From radiative transfer theory it is known that the relation between a quantity called infinite reflectance and the single
scattering albedo of the medium is highly non-linear. The single scattering albedo is one minus the fraction of the
intercepted radiation that is absorbed, so the absorption coefficient of the material also has a non-linear influence on the
reflectance. This also holds when the optical thickness of the medium is less than infinite. For weak absorption we see
an almost linear relationship between optical thickness and reflectance, whereas for strong absorption this relationship
is strongly (negative-)exponential. Altogether, this means that in a region where the absorption coefficient varies
significantly with wavelength, the shape of the spectral reflectance curve will change with the optical thickness. In other
words, when the reflectance at two wavelengths is given, the reflectance at a wavelength halfway between the first two
will contain additional information, because the shape of the spectrum changes with optical thickness, even if the
absorption coefficient varies linearly in the interval. This is a strong motivation to the need for high spectral resolution
in certain wavelength regions. For leaf canopies the above effects come to expression as follows:

- When the leaf chlorophyll concentration decreases, this is first noticed at the wavelengths where chlorophyll
absorption is weakest, i.e. in the red edge close to the near infrared and in the green. Only when it decreases much
further, an effect in the red becomes visible. In the actual near infrared nothing will change in the canopy
reflectance, however, as there is no sensitivity to leaf chlorophyll in that part of the spectrum.

- When the canopy LAI decreases from a high value at high leaf chlorophyll concentration, it is first noticed in the
near infrared, because leaf absorption is minimum there. When LAI is low, sensitivity to LAI changes is greatest in
the red.

This is also demonstrated in Fig.1, which shows the modelled canopy reflectance times 10000 in the red edge for a
series of combinations of LAI and leaf chlorophyll concentration. The LAI varies according to the series 0, 0.25, 0.5, 1,
2, 4 and 8. Chlorophyll concentration is given by the series 15, 20, 25, …, 60, 70 and 80 µg / cm². The spectral bands
represent the wavelength range from 670 to 800 nm at 10 nm intervals.

![Spectral Profile](image)

Figure1. Effects of leaf chlorophyll and LAI on canopy reflectance in the red edge

The position of the red edge point clearly moves to longer wavelengths for higher chlorophyll contents, but the canopy
LAI complicates this simple relationship, and incorporation of other parameters would complicate it even more. It can
be concluded that leaf chlorophyll concentration and canopy LAI act in different ways on spectra of the vegetation
canopy reflectance in the red edge region. Due to this, it should also be possible to retrieve both parameters from high
resolution spectral reflectance data in this region. Note, that with spectral reflectance data from red and near infrared
alone this is not possible: several combinations of leaf chlorophyll concentration and canopy LAI can give the same
combination of red and near infrared reflectance, so it would be impossible to retrieve both parameters from spectral

Verhoef, Wout

data in the red and near infrared alone. However, hyperspectral data from the red edge region not only tell us the reflectance at both ends, but also the shape of the curve in between, and this enables us to retrieve both parameters. There have been numerous studies on the use of the red edge point (the wavelength where the first spectral derivative is maximum) as an indicator for LAI and other vegetation parameters and on how the red edge point can be derived best from hyperspectral data [Hare et al. (1984), Leprieur (1989), Guyot & Baret (1988), Bach (1995), Bach & Mauser (1997)]. However, the red edge position has the same disadvantage as the NDVI, namely that it responds to several parameters simultaneously, so that no single one can be retrieved accurately, unless all other ones are constant, which of course is not very likely.

The red edge spectral region (670 – 800 nm) is very suitable for information on leaf chlorophyll, as in this region the chlorophyll absorption coefficient varies from very high to zero, so a large dynamic range is traversed in a relatively short wavelength interval. Also, few other optical parameters vary in this range, and the ones that do, vary only little. Therefore this region was chosen as the spectral interval of the model inversion experiment.

While leaf chlorophyll concentration and canopy LAI may be the most interesting biophysical parameters to be retrieved from remote sensing data, in reality other canopy parameters, leaf parameters, soil parameters, and atmospheric properties are expected to vary as well and are having an influence on spectral reflectances, so for some degree of realism these should also be considered in the numerical experiment.

2 METHODS

Model inversion experiments with several canopy reflectance models have already been carried out since the work of Goel and Thompson [Goel & Thompson (1984), Goel (1988)]. In most of these experiments the model is called iteratively and a merit function is defined which indicates the squared distance to the solution and in which sometimes a penalty function is incorporated to avoid the exceeding of parameter boundaries. The iteration stops when the model outputs match with measured spectral data.

Models can be calibrated if measured observables and parameters are both available. In the case of many output parameters, such as with hyperspectral and multidirectional data, the single quadratic merit function is formed out of information from a large number of “channels”. In this approach, the model inversion algorithm has to get its information on how to change the parameters in the direction of a solution from the Hessian, the matrix of second order partial derivatives of the merit function with respect to the parameters. As in this method all squared deviations from the target pattern are added up to form the merit function, information on how to improve the fit in a subspace of the patterns is lost. Therefore, in this section an alternative is proposed, the well-known Gauss-Newton method. Here the information on how to change the parameters in the right direction is derived from the Jacobian, which is the matrix of first partial derivatives of all model output variables with respect to the input parameters. It will be clear that in this case much more information on how to change the parameters is directly available. Therefore, this method is preferred. This method proceeds as follows:

Let the Jacobian matrix $J$ be defined by

$$
\Delta r = J \Delta p,
$$

where $\Delta p$ is a small change in the parameter vector, and $\Delta r$ is the resulting change in the vector of spectral-directional reflectances, then multiplication by the transposed of $J$ gives

$$
J^T \Delta r = J^T J \Delta p, \quad \text{or} \quad \Delta p = (J^T J)^{-1} J^T \Delta r.
$$

The numerical stability of the Gauss-Newton method can greatly be improved by means of the Levenberg-Marquardt algorithm, which is based on the modified equation

$$
\Delta p = (J^T J + \mu I)^{-1} J^T \Delta r,
$$

where $\mu$ is a scalar parameter that is used to control the numerical behaviour of the algorithm. When this parameter is high, the conversion to the solution follows the steepest descent direction and therefore is secure but slow. When it is low, the Gauss-Newton direction is approximately followed, which is fast in the neighbourhood of the final solution, but unstable and possibly slow otherwise. Therefore one usually starts with a high value of the parameter, and when good progress is made, as evidenced by a decreased distance to the solution, the parameter is lowered by a certain
factor. If the distance to the solution turns out to have increased, this may lead to instability, and in that case the parameter is increased in order to regain control via the steepest descent direction.

Regarding the model simulations, the following procedure was followed:

Spectra of the single leaf reflectance and transmittance in the red edge region have been simulated for a range of chlorophyll concentrations \( C_{ab} \) and 3 values of the leaf mesophyll parameter \( N \) by means of the PROSPECT model (Jacquemoud & Baret, 1990). The simulations have been carried out for the wavelength range 670 – 800 nm with 10 nm steps, thus resulting in optical data at 14 wavelengths. Chlorophyll concentrations used were 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70 and 80 \( \mu g/cm^2 \) (12 values) and the 3 values of the mesophyll parameter were 1.0, 1.5, and 2.0.

In a first numerical experiment, the data generated by PROSPECT were used in combination with a very simple Kubelka-Munk type canopy reflectance model in order to verify whether canopy LAI and leaf chlorophyll indeed gave different spectral responses in the red edge region. In this experiment also the soil brightness and the soil’s spectral slope were varied. This enabled us to evaluate methods of compensating for soil background effects by means of first and second spectral derivatives.

In the next step the simple KM-type model was replaced by the SAILH model and a model inversion experiment was carried out in order to verify whether the most important parameters could be retrieved from hyperspectral bidirectional reflectance data. The SAILH model (Verhoef, 1998) is an extended version of the SAIL model (Verhoef, 1984) to include the hot spot effect. Here the parameters to be retrieved in the groups soil, canopy and leaf were:

- **Soil brightness**
- **Soil spectral slope**

- **Canopy LAI**
- **Average leaf slope**
- **Bimodality parameter of the LIDF**
- **Canopy hot spot parameter**

- **Leaf chlorophyll concentration**
- **Leaf mesophyll parameter**

In this experiment the solar zenith angle, the viewing zenith angle and the relative azimuth were all held fixed and no diffuse incident light on the canopy top was assumed. During the development of the model inversion procedure several enhancements were introduced, such as transformations to the input parameters in order to avoid problems due to strong non-linearity and due to the interactions between parameters and their valid ranges.

In the last experiment the SAILH model was replaced by the OSCAR model in order to further enhance the degree of realism.

The model OSCAR (Optical Soil-Canopy-Atmosphere Radiance) (Verhoef, 1998) is an integrated model in which SAILH has been interfaced with the soil’s reflectance and scattering in the atmosphere. The number of parameters was increased from 8 to 11 by incorporation of 3 atmospheric parameters, namely:

- **Visibility at sea level in km**
- **Aerosol Ångström coefficient**
- **Aerosol single scattering albedo**

Also, in order to simulate TOA observations, the planetary reflectance was used as the remote sensing observable on which model inversion was based. As in the previous experiment it already became clear that retrieval of the LIDF parameters was the most difficult if only one direction was applied, it was decided to include more directions. For this, the five best directions for estimation of LAI from planetary reflectance data, as resulting from the SASSSIS study, were chosen. These are:

<table>
<thead>
<tr>
<th>Viewing zenith</th>
<th>Azimuth</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>120</td>
</tr>
<tr>
<td>69</td>
<td>120</td>
</tr>
<tr>
<td>52</td>
<td>90</td>
</tr>
<tr>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>69</td>
<td>180</td>
</tr>
</tbody>
</table>
Multiplied by the number of spectral bands this gives 70 hyperspectral-directional data points on which a model inversion can be based. The solar zenith angle was fixed at a value of 45 degrees. The OSCAR model allows also variation of the target surroundings for modelling of the adjacency effect (Verhoef, 1998), but in this experiment the optical properties of the surroundings were held constant in order to avoid too much complications.

The model input parameters and possibly their transformations are described below:

It is assumed that the soil reflectance changes linearly with wavelength in this short interval. The soil’s reflectance in the interval is described by means of the parameters P₁ and P₂, called brightness and spectral slope. They are defined by

\[ P₁ = \frac{(R_{670} + R_{800})}{2}, \text{ and} \]
\[ P₂ = \frac{R_{800}}{R_{670}}. \]

where \( R_{670} \) and \( R_{800} \) are the soil’s reflectance at 670 and 800 nm, respectively. From these parameters the soil’s reflectance at any wavelength can always be reconstructed.

The ranges for the parameters considered during model inversion were 0 – 0.6 for \( P₁ \) and 1.0 – 1.5 for \( P₂ \).

LAI is transformed to \( P₃ = \exp(-0.2 \times \text{LAI}) \). This reduces non-linearity, especially in the near infrared, and the valid range becomes 0 – 1.

The LIDF parameters \( a \) and \( b \) of the SAILH model are related to average leaf slope and bimodality of the leaf inclination distribution function, respectively (Verhoef, 1998). Both parameters can vary from −1 to 1, but not independently, as the sum of their absolute values should stay less than or equal to unity. Therefore, these parameters are transformed into:

\[ P₄ = 0.5 (a + b + 1) \]
\[ P₅ = 0.5 (a - b + 1) \]

These parameters both range (independently) from 0 to 1 when \( a \) and \( b \) are in the valid range.

Parameter \( P₆ \) is the hot spot parameter and it is supposed to vary between 0 and 0.5.

Parameters \( P₇ \) and \( P₈ \) are the leaf parameters, describing chlorophyll concentration and the leaf mesophyll parameter. During model inversion, care must be taken to ensure that the leaf parameters actually are integer numbers in this case, as the PROSPECT model had not been integrated into OSCAR, but rather use was made of simulation results obtained with PROSPECT for discrete cases of these parameters in an earlier stage. Therefore, \( P₇ \) is an integer from 1 to 12, and \( P₈ \) is an integer from 1 to 3. In the model inversion results discussed later, these parameters are presented as real numbers, because there the predictions of the correct parameters are shown, not the values actually used in the simulations.

The atmospheric parameter visibility \( V \) has a strongly non-linear effect on the optical properties of the atmosphere. A change from 6 to 5 km may have more effect than a change from 50 to 30 km visibility. Therefore, this parameter is replaced by

\[ P₉ = \exp(-0.1 \times V). \]

This reduces non-linearity for the atmospheric effects and the valid range becomes 0 – 1.

The other atmospheric parameters are \( P₁₀ \), the aerosol Ångström coefficient, which is assumed to vary from −0.6 to −1.3, and \( P₁₁ \), the aerosol single scattering albedo. This parameter is assumed to vary from 0.6 to 1.0.

3 RESULTS

3.1 Simulated hyperspectral planetary reflectance data

In this section some results of model inversion based on simulated hyperspectral data are discussed. Theoretically one should always find the exact solution in this case, because the artificial data were generated by the same model as the one used in the model inversion. In the Gauss-Newton method it can be very important to have good starting guesses for
the parameters. However, when there is no information available, the best starting guess is the centre of parameter space, so this was taken as a starting point for the model inversion.

From several numerical experiments it has become clear that in 8-70 iterations an accurate solution can be obtained in this way and it has been demonstrated that in principle it is possible to retrieve all parameters, useful biophysical parameters as well as ones that can be considered disturbing factors.

3.2 Ground based hyperspectral reflectance measurements

The classical model inversion method was also applied to measured hyperspectral reflectance data from sugar beet, obtained in an Anglo-French collaboration field experiment (Maltus, 1990). The measurements used for model inversion were collected in Thiverval-Grignon, France in July 1990. Some of the main characteristics for these measurements are the following:

- LAI variation created by thinning
- Chlorophyll variation created by herbicide treatment
- Artificial soil background variation
- Vertical hyperspectral measurements from IRIS instrument under natural diurnal illumination from sun and sky

In order to create similar conditions to the previous case of inverting simulated data and to PRISM characteristics, the hyperspectral data were resampled spectrally to 10 nm resolution and the same red edge wavelength range of 670-800 nm was used. Because in this case no directional information could be used in the model inversion, also the number of parameters to be retrieved had to be reduced. The hot spot parameter was held fixed at a value of 0.5, as it was known to be a sugar beet crop. The LIDF bimodality was slaved to the mean leaf slope by assuming a fixed relationship between them. Finally, the fraction diffuse sky irradiance was taken from measured data, so this parameter did not have to be retrieved from model inversion. This left as the six parameters to be retrieved:

- Soil brightness in the red edge
- Soil spectral slope in the red edge
- LAI
- Mean leaf slope
- Leaf chlorophyll concentration
- Leaf mesophyll parameter

Applying the automatic model inversion procedure on real measurement data turned out unsuccessful with the present software because of failures to converge. Therefore, a “manual” procedure was used as an alternative. This was achieved by trying to find a minimum error varying only one parameter at a time, and then moving to the next parameter. Such a cycle through all parameters was repeated until no further improvements were possible. Figure 2 shows an example of the result of a manual inversion of a single measurement on the basis of measured and modelled hyperspectral reflectances in the red edge region. It appears that the modelled spectrum matches the measured one fairly closely.

![Figure 2](image-url)

Figure 2. Measured and modelled hyperspectral reflectance data for a model inversion.

However, what is more interesting is of course the correspondence between measured and retrieved parameters. Table 1 shows the results of this comparison for the same measurement as for which the spectral data were shown in Fig. 3. Note that in the measured parameters LAI appears twice: the first entry was measured, whereas the second was
estimated from optical soil coverage. The parameter VAI means void area index and is sometimes used to describe single leaf mesophyll structure. It is believed to be equivalent to the mesophyll parameter $N$ of the PROSPECT model. The retrieval results for this measurement are rather good for the parameters soil brightness, mesophyll parameter and leaf chlorophyll concentration. However, rather poor results are obtained for the canopy structure parameters LAI and mean leaf slope. There are a few reasons for this that can be identified:

- Thinning is not a normal cause of LAI reduction in crops and creates large discrepancies with model assumptions, as thinning makes the crop even more heterogeneous than already was the case due to crop rows.
- Directional information was not used in the experiment, while this is essential for accurate retrieval of canopy structural parameters such as LAI and mean leaf slope.

Table 1 Measured and retrieved parameters

<table>
<thead>
<tr>
<th>Measured parameter</th>
<th>Value</th>
<th>Retrieved parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil brightness</td>
<td>0.188</td>
<td>Soil brightness</td>
<td>0.2</td>
</tr>
<tr>
<td>LAI measured</td>
<td>1.44</td>
<td>Soil spectral slope</td>
<td>1</td>
</tr>
<tr>
<td>LAI estim. from cov.</td>
<td>2.09</td>
<td>LAI</td>
<td>2.81</td>
</tr>
<tr>
<td>Mean leaf slope</td>
<td>34.6</td>
<td>Mean leaf slope</td>
<td>60</td>
</tr>
<tr>
<td>Chlorophyll conc.</td>
<td>23.8</td>
<td>Chlorophyll conc.</td>
<td>20</td>
</tr>
<tr>
<td>VAI</td>
<td>1.53</td>
<td>Mesophyll par.</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**4 CONCLUSIONS**

It has been demonstrated that, at least theoretically, it is possible to retrieve important (bio)physical parameters from red edge hyperspectral remote sensing data acquired under different directions. The most important of these parameters are soil brightness in the red edge region, leaf chlorophyll concentration, canopy leaf area index and atmospheric visibility. This can be very relevant for the study of land surface processes.

Spectral-directional TOA radiance data appear to be unique with respect to the possible combinations of parameters, provided these do not lie on a boundary of parameter space. When modelled data are used as input for the model inversion procedure, the correct input parameters are accurately retrieved. Retrieval results obtained from actual hyperspectral reflectance data are disappointing. Especially canopy structural parameters are poorly estimated, but this might be improved substantially if directional hyperspectral data were used as input.

Simultaneous retrieval of all radiometrically relevant parameters is aimed at the estimation of biophysically interesting parameters such as leaf chlorophyll and canopy LAI, while at the same time the factors influencing the relationships between these parameters and remotely sensed data are also estimated. If the retrieval of all parameters is correct, it means that also the disturbing influences are known, so that the accuracy of the retrieved parameters of interest will be higher than when these factors are completely ignored.

Classical model inversion of hyperspectral multidirectional planetary reflectance data requires enormous processing efforts, so for application in practice it is recommended to investigate accelerations in numerical modelling, possibilities for improving convergence, and alternatives such as artificial neural networks or multivariate look-up tables. Also several practical problems associated with parameter retrieval by inversion of complex models need to be investigated in more detail than was possible in the current study. For instance mixing continuous and stepwise varying parameters (classes) in the same model creates several poorly understood problems when one tries to invert this model. The results obtained are still very preliminary. A more comprehensive investigation should also consider the influence of instrumental noise and a more realistic modelling of the adjacency effects and of gaseous absorption in the atmosphere. From an operational point of view it would be important to investigate how a less dense spectral and directional sampling would influence the success of parameter retrieval.

**ACKNOWLEDGEMENTS**

This study has been carried out as part of a Requirements Study towards a Land Surface Processes and Interactions Mission (LSPIM), under ESA contract no. 13176/98/NL/GD.
REFERENCES


