Application to MISR land products of an RPV model inversion package using adjoint and Hessian codes

Thomas Lavergne a, Thomas Kaminski b, Bernard Pinty a,*, Malcolm Taberner a, Nadine Gobron a, Michel M. Verstraete a, Michael Vossbeck b, Jean-Luc Widlowski a, Ralf Giering b

a Global Environment Monitoring Unit, IES, EC Joint Research Centre, TP 440, via E. Fermi, I-21020 Ispra (VA), Italy
b FastOpt, Schanzenstrasse 36, D-20357 Hamburg, Germany

Received 8 February 2006; received in revised form 28 April 2006; accepted 7 May 2006

Abstract

The capability of the non-linear Rahman–Pinty–Verstraete (RPV) model to 1) accurately fit a large variety of Bidirectional Reflectance Factor (BRF) fields and 2) return parameter values of interest for land surface applications motivate the development of a computer efficient inversion package. The present paper describes such a package based on the 3 and 4 parameter versions of the RPV model. This software environment implements the adjoint code, generated using automatic differentiation techniques, of the cost function. This cost function itself balances two main contributions reflecting 1) the \textit{a priori} knowledge on the model parameter values and, 2) BRF uncertainties together with the requirement to minimize the mismatch between the measurements and the RPV simulations. The individual weights of these contributions is specified notably via covariance matrices of the uncertainties in the \textit{a priori} knowledge on the model parameters and the observations. This package also reports on the probability density functions of the retrieved model parameter values that thus permit the user to evaluate the \textit{a posteriori} uncertainties on these retrievals. This is achieved by evaluating the Hessian of the cost function at its minimum. Results from a variety of tests are shown in order to document and analyze software performance against complex synthetic BRF fields simulated by radiation transfer models as well as against actual MISR-derived surface BRF products.

© 2006 Elsevier Inc. All rights reserved.

Keywords: RPV; Inversion; MISR; Adjoint code; Hessian code; Posterior uncertainties; Radiation transfer

1. Introduction

The Rahman–Pinty–Verstraete (RPV) model (Rahman et al., 1993) approximates the Bidirectional Reflectance Factor (BRF) of an arbitrary surface as a function of the geometry of illumination (radiation arriving from the single direction $\Omega_0(\theta_0, \phi_0)$) and observation (radiation leaving into the single direction $\Omega(\theta, \phi)$), as well as a small set of parameters meant to explain the anisotropy of the surface. This model belongs to a class of so-called semi-empirical models (see, for instance, the reviews by Lucht and Roujean, 2000; Roberts, 2001; Verger et al., 2005). This model is not derived from basic physical laws but recognizes that the BRF shape and amplitude of an arbitrary target can be usefully decomposed into a limited set of well identified contributions such as bowl versus bell-shape, backward versus forward scattering regimes or absence of shadows in the exact backscattering direction. This category of models represents and approximates BRF fields at very limited computer cost and has proved relevant for purposes such as 1) the specification of a lower boundary condition for solving an atmospheric (aerosol and/or cloud) radiation transfer problem (e.g., Martonchik et al., 1998a,b), 2) the representation of clear-sky BRF fields at the top of the atmosphere (e.g., Gobron et al., 2000) which allows, for instance, the detection of cloud and cloud shadow occurrences from a BRF time-sequence measured at the top of the atmosphere (e.g., Pinty et al., 2000) as well as the extrapolation of BRF values corresponding to incoming and exiting directions other than those measured (e.g., Lattanzio et al., 2005), 3) the estimation of angularly integrated quantities like the Directional Hemispherical Reflectance (Black Sky Albedo) and the BiHemispherical Reflectance (White Sky Albedo) (e.g., Lucht et al., 2000;
Martonchik et al., 1998a,b; Schaaf et al., 2002), 4) the identification of some surface characteristics (e.g., Cabot and Dedieu, 1997; Duchemin, 1999; Lewis and Vives Ruiz de Lope, 1997; Rahman et al., 1993), including those related to the presence and density of vertical structures (e.g., Gao et al., 2003; Pinty et al., 2002; Widlowski et al., 2001). Although all these applications except the last could, in principle, be satisfied with the help of powerful enough mathematical formulae, e.g., orthogonal polynomials, the fourth type of applications cited above suggests that such models as RPV could be helpful to retrieve, for instance, some vegetation structural properties.

The RPV formulation splits a BRF field into a scalar amplitude component and the associated angular field describing the anisotropy of the surface, that is:

$$\rho_{sfc}(z_0, \Omega_0, \Omega; \rho_0, k, \Theta, \rho_c) = \rho_0 \hat{\rho}_{sfc}(z_0, \Omega_0, \Omega; k, \Theta, \rho_c)$$

(1)

where $\rho_0$ gives the overall reflectance level and where the angular field of the surface BRF, $\hat{\rho}_{sfc}(z_0, \Omega_0, \Omega; k, \Theta, \rho_c)$, is formulated as follows:

$$\hat{\rho}_{sfc}(z_0, \Omega_0, \Omega; k, \Theta, \rho_c) = M_I(\theta_0, \theta; k) F_{HG}(g; \Theta) H(\rho_c; G)$$

(2)

where

$$M_I(\theta_0, \theta; k) = \frac{\cos^k \theta_0 \cos^k \theta}{(\cos \theta_0 + \cos \theta)^{n-k}}$$

(3)

$$F_{HG}(g; \Theta) = \frac{1 - g^2}{[1 + 2g \cos \Theta + \Theta^2]^{3/2}}$$

(4)

$$H(\rho_c; G) = 1 + \frac{1 - \rho_c}{1 + G}$$

(5)

with

$$\cos g = \cos \theta \cos \theta_0 + \sin \theta \sin \theta_0 \cos \phi_r$$

(6)

$$G = [\tan^2 \theta_0 + \tan^2 \theta - 2 \tan \theta_0 \tan \theta \cos \phi_r]^{1/2}$$

(7)

$$\phi_r = \phi_0 - \phi$$

(8)

where $\phi_0$ denotes the relative azimuth between the illumination and viewing conditions specified by $\Omega_0$ and $\Omega$ which are directional vectors at a target on the surface pointing to the Sun and to the detector, respectively.

The parameter $k$ in the modified version of the Minnaert’s function (Eq. (3)) (Minnaert, 1941) controls the bowl or bell-shape of the BRF field, the parameter $\Theta$ establishes the degree of forward versus backward scattering, depending on its sign, following the Henyey–Greenstein formulation (Henyey and Greenstein, 1941) and the parameter $\rho_c$ accounts for the hot spot effect especially significant in the backscattering region. When the surface scattering properties follow a Lambert law, meaning that the upward intensity field is scattered isotropically, the angular function, $\rho_{sfc}(z_0, \Omega_0; \Omega; k, \Theta, \rho_c)$ is constant and equal to unity, a condition fulfilled numerically by setting the $\Theta$, $k$ and $\rho_c$ parameter values equal to 0, 1 and 1, respectively.

In summary, the RPV model uses four parameters for representing the BRF field and describing the anisotropy of an arbitrary target. In some instances, the parameter accounting for the hot spot effect $\rho_c$ can be set at a fixed value or forced to be equal to the scalar amplitude parameter $\rho_0$, in which cases only 3 instead of 4 model parameters are needed by RPV in order to represent the angular field of the BRF (see Engelsen et al., 1996). In the following of this paper, the 3 parameter version thus refers to an RPV implementation using, $\rho_0$, $k$ and $\Theta$ parameters only, with $\rho_c$ equal to $\rho_0$.

Applications of this category of semi-empirical models require fitting/inverting the selected model against a set of BRF measurements or simulations for a given scattering target. Such a procedure typically yields a range of model parameter values (expressing the non-uniqueness of solutions) that are providing statistically acceptable fits between measured and modeled BRF values, together with a documentation of their associated uncertainties, as a result of either uncertainties in the input measurements or intrinsic limitations in the semi-empirical model, or both. In this context, linear models such as the kernel-driven models (e.g., Lucht and Lewis, 2000) appear easier to handle, thanks to the least-square theory and related statistical analysis, than the non-linear formulation proposed by RPV. This motivated the substitution of the original Henyey–Greenstein expression (Eq. (4)) by a simpler exponential function, yielding the MRPV model which, after a logarithmic transformation, shows a quasi-linearity (Engelsen et al., 1996) and thus also becomes easily applicable in operational contexts (Martonchik et al., 1998a,b).

Nonetheless, the RPV model has been shown to be quite powerful in representing a whole range of complex BRF fields (see, for instance, Engelsen et al., 1996; Gobron and Lajas, 2002; Lucht, 1998; Privette et al., 1997; Rahman et al., 1993; Roberts, 2001). To efficiently invert the non-linear RPV model, Gobron and Lajas (2002) have already proposed a simpler mathematical scheme than those traditionally adopted in the case of non-linear systems. This scheme implements the solution to a $\chi^2$ estimator defined with a pre-specified accuracy level for fitting the data with the model. It requires the analytical solution of a second order polynomial equation using an ensemble of pre-computed functions discretized along a range of those RPV model parameter values that control the BRF shape. This scheme has proved to be computer efficient and numerically robust in delivering the most representative set of solutions minimizing the quadratic distance between the measurements and the model simulations. Its computer performance and accuracy in actual retrievals remain, however, constrained by the pre-specified range and discretization of the space of the RPV parameters controlling the shape of the BRFs.

The present paper revisits the issue of the RPV model inversion by devising and documenting a methodology based on inverse theory as promoted, for instance, by Tarantola (1987) and Enting (2002). This methodology presents a number of
advantages including the complete assessment of the measurement-model mismatch covariance matrix and the engineering possibility to operate adjoint software codes themselves derived from automatic differentiation techniques (Griewank, 2000). This leads to a software package allowing the user to perform, under the classical Bayesian approach, the inversion of the non-linear RPV model in a numerically and computationally very efficient manner, while at the same time generating a rather complete and mostly unbiased estimation of the Probability Density Functions (PDFs) for the retrieved parameters. A series of applications based on BRF fields simulated by radiation transfer models and derived from data acquired by the Multiangle Imaging SpectroRadiometer (MISR) instrument are conducted to evaluate the performances of the RPV inversion package.

2. Inversion methodology

The formulation of an inverse problem aims to optimize the use of available information specified through a priori knowledge on the values of the model parameters $X$, the measurement, and the constraint provided by the model. This a priori knowledge stabilizes the inversion procedure and compensates for an incomplete or non-exhaustive set of measurements which helps solving ill-conditioned problems. It is appropriate to formulate the inverse problem in terms of Probability Distribution Functions (PDFs) within the spaces of model parameters and measurements. Finding solutions to the inverse problem thus means estimating the a posteriori PDFs for the model parameters.

The associated algebra is particularly convenient (see Tarantola, 1987), if the measurements and the a priori knowledge on the model parameters can be approximated by Gaussian probability distributions that is, if they can be represented by their mean values $d$ and $X_{po}$ and associated covariance matrices denoted by $C_d$ and $C_{X_{po}}$, respectively. If, in addition, the local linearisation provides a good approximation for the non-linear model $M(X)$, then the a posteriori probability distribution $P(X)$ is close to Gaussian as well. Hence, $P(X)$ can be approximated by its mean value and covariance matrix denoted $X_{po}$ and $C_{X_{po}}$ respectively (where the superscript $^T$ represents the matrix transpose):

$$P(X) \approx \exp \left( -\frac{1}{2} (X-X_{po})^T C_{X_{po}}^{-1} (X-X_{po}) \right)$$

(9)

$X_{po}$ represents the maximum likelihood estimator of the mean of $P(X)$ and minimises the following cost function $J(X)$ expressed as follows:

$$J(X) = \frac{1}{2} \left[ (M(X)-d)^T C_d^{-1} (M(X)-d) \right.$$

$$+ (X-X_{po})^T C_{X_{po}}^{-1} (X-X_{po}) \left. \right]$$

(10)

The first term of the right-hand side of Eq. (10) quantifies the mismatch between the model simulations and the measurements while the second term expresses the constraint given by a priori knowledge on the model parameters $X$. The covariance matrices expressing the uncertainties in the measurements and in the a priori knowledge on model parameters can be conceived of as weighting factors in Eq. (10). If the covariance matrices are diagonal (implying zero correlation between the matrix elements), then Eq. (10) reduces to a least squares formulation and each component of $d$ and $X_{po}$ is weighted in inverse proportion to the uncertainty level that is associated with it. Under conditions where a large set of accurate independent measurements is available, the relative weight of the a priori knowledge on the $X$ decreases, and the solution to Eq. (10) is essentially controlled by the first term on the right-hand side of this equation. This latter term ensures that the retrieved model parameter values generate BRF values that are close to the measurements.

The inverse problem is solved with a gradient algorithm minimizing Eq. (10) which iteratively evaluates both $J(X)$ and its gradient $\nabla J(X)$ with respect to $X$. The exact evaluation of this gradient can be achieved via the adjoint model of $J(X)$ which also saves significant computer time by contrast to finite-difference techniques (Giering and Kaminski, 1998). An additional term, which drastically increases the cost function in case of unrealistic combinations of model parameters (not shown but entering Eq. (10)), essentially excludes the corresponding regions in the model parameter space. This is somewhat similar to adding bounds on the parameter space.

Under the regularity conditions mentioned above, the Hessian $\nabla^2 J(X)$ at the minimum of $J$ approximates the inverse of the a posteriori covariance matrix $C_{X_{po}}$. Comparison with the a priori uncertainty indicates how well individual directions in the space of model parameters are observed through the measurements. Eigenvectors of $C_{X_{po}}$ indicate the independent directions in the parameter space, and the corresponding eigenvalues quantify the uncertainties associated with the retrievals.

When seeking the minimum of a cost function such as $J(X)$, the crux of the problem lies in the generation of a computationally efficient software package while preserving the numerical accuracy of the various terms contributing to this cost function. Now popular inverse modeling and data assimilation methods have promoted the development of advanced software tools, including those based on automatic differentiation techniques (Griewank, 2000). In the present RPV model application, the derivative codes have been generated by the compiler tool Transformation of Algorithms in Fortran (TAF) (Giering and Kaminski, 1998; Giering et al., 2005) available from FastOpt (http://www.FastOpt.com/). The software routines for achieving minimization (dfmin) and matrix inversion (performed by computing the eigenvalues and eigenvectors of this matrix (jacobi)) are provided by the scientific library available from Press et al. (1986).

The performance of this inversion package is illustrated in this section against a model-simulated BRF scenario taken from the RAdition transfer Model Intercomparison (RAMI) exercise (Pinty et al., 2004a,b). This particular scenario corresponds to an idealized coniferous forest scene with a flat bright underlying background meant to render a snow-covered ground typical of winter conditions, albeit represented by a simple Lambertian
scattering law. A detailed description of this scene (labelled as HET03 with no topography in the RAMI environment) is given in Pinty et al. (2004a,b) and can also be found at http://rami-benchmark.jrc.it/. The BRF fields analyzed here were calculated in the principal and orthogonal planes using the Raytran ray tracing Monte-Carlo model of Govaerts and Verstraete (1998), operated with the RAMI conditions set up to address the radiation transfer regime in the near-infrared spectral domain.

The RPV model inversion package described here was applied separately to the BRF values generated once in the principal plane and once in the orthogonal plane. Each application was performed with the RPV model using 3 parameters, as well as with the 4 parameter version. In all cases analyzed in this section, and elsewhere in this paper unless otherwise stated, the a priori knowledge on parameters X was specified in the following way: 1) the a priori values for parameters \( \rho_0, k, \Theta, \) and \( \rho_c \) are set at 0.01, 1.00, 0.00, and 0.01, respectively, 2) large uncertainty values (standard deviations equal to \( 10^0 \)), are associated with these parameter values and, 3) the covariance matrix, \( C_{X_a} \), is diagonal. As mentioned earlier, the setting of large uncertainties on the a priori knowledge on model parameters gives a very small weight to the second contributing term on the right-hand side of the cost function (Eq. (10)). The covariance matrix associated with the measurements, \( C_d \), is assumed diagonal as well with a uniform standard deviation value, corresponding to 10% of the mean BRF field in each separate viewing plane, that is equal to \( 3.9 \times 10^{-2} \) and \( 3.4 \times 10^{-2} \) for the principal and orthogonal planes, respectively.

Fig. 1 displays the BRF values reconstructed from the RPV model given the 3 (red line) or 4 (green line) best fit parameters, respectively. The square symbol represents the BRF values given as input to the RPV model inversion package. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2 provides the evolution of the values of \( J(X) \) (normalized by the number of measurements and plotted on a logarithmic scale) as a function of the iteration number. The full (dashed) lines corresponds to the inversion of the RPV package against BRF values in the principal (orthogonal) plane. Results with the 3 (4) parameter version are indicated in red (green) color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
of the number of iterations required to find its minimum, for the four studied cases, namely the principal (full lines) and orthogonal (dashed lines) planes and using the version implementing 3 (in red) and 4 (in green) model parameters in the RPV package. In both viewing planes, the 4 parameter version finds lower minima of $J(X)$ than is possible with the 3 parameter version, hence a slightly better fit, but at the cost of a larger number of iterations, as can be expected. In the studied cases, the number of iterations required to identify an acceptable minimum is quite small given the non-linearity of the RPV model (the decrease in the values of $J(X)$ is very fast) and illustrates the excellent performance of the RPV inversion package used here. This number of iterations is determined by the gradient of the cost function. In the current setup, the minimization is stopped once the norm of this gradient is less than a prescribed conservative threshold value equal to $10^{-6}$. A higher threshold value would thus reduce the number of iterations and hence speed up the inversion procedure.

Table 1 provides basic statistical information about the RPV model parameter values delivered by the inversion package. The estimated values of the standard deviation (corresponding to the values located along the diagonal of the *a posteriori* covariance matrix $C_{\lambda_k}$) associated with the mean values of the retrieved parameters are systematically larger in the orthogonal than in the principal plane. In other words, the number of iterations required to identify an acceptable minimum is quite small given the non-linearity of the RPV model (the decrease in the values of $J(X)$ is very fast) and illustrates the excellent performance of the RPV inversion package used here. This number of iterations is determined by the gradient of the cost function. In the current setup, the minimization is stopped once the norm of this gradient is less than a prescribed conservative threshold value equal to $10^{-6}$. A higher threshold value would thus reduce the number of iterations and hence speed up the inversion procedure.

<table>
<thead>
<tr>
<th>Viewing conditions</th>
<th>Moment</th>
<th>$\rho_0$</th>
<th>$k$</th>
<th>$\Theta$</th>
<th>$\rho_u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal plane with 3 parameters</td>
<td>Mean</td>
<td>0.239</td>
<td>0.893</td>
<td>−0.084</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>0.58 × 10^{-2}</td>
<td>2.26 × 10^{-2}</td>
<td>0.85 × 10^{-2}</td>
<td>−0.152</td>
</tr>
<tr>
<td>Orthogonal plane with 3 parameters</td>
<td>Mean</td>
<td>0.199</td>
<td>0.834</td>
<td>−0.152</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>2.92 × 10^{-2}</td>
<td>12.13 × 10^{-2}</td>
<td>9.08 × 10^{-2}</td>
<td>2.167</td>
</tr>
<tr>
<td>Principal plane with 4 parameters</td>
<td>Mean</td>
<td>0.221</td>
<td>0.875</td>
<td>−0.069</td>
<td>−0.041</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>1.75 × 10^{-2}</td>
<td>2.82 × 10^{-2}</td>
<td>1.61 × 10^{-2}</td>
<td>28.30 × 10^{-2}</td>
</tr>
<tr>
<td>Orthogonal plane with 4 parameters</td>
<td>Mean</td>
<td>0.514</td>
<td>1.063</td>
<td>−0.361</td>
<td>2.167</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>71.69 × 10^{-2}</td>
<td>12.13 × 10^{-2}</td>
<td>63.50 × 10^{-2}</td>
<td>149.4 × 10^{-2}</td>
</tr>
</tbody>
</table>

3. Inversion results

This section documents results delivered by the RPV model inversion package when operated against 1) a large range of highly anisotropic conditions as generated by radiation transfer model simulations on three-dimensional (3-D) scenes and 2) actual surface BRF values derived operationally by the Multangle Imaging SpectroRadiometer (MISR) instrument on board NASA’s Terra platform. These two series of experiments are intended to test the performance of the inversion package with respect to its robustness and computer speed under a variety of BRF amplitudes and shapes as well as sampling conditions.

3.1. Ensemble of model-based simulations

We have opted for 3-D model-simulated vegetation scenes to establish the performance of the RPV model inversion package. The various scenarios considered here are all based on realistic forest system properties (birch trees in this case) as reported in

Table 2

Correlation of a posteriori uncertainties of the retrieved RPV parameter values

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\rho_0$</th>
<th>$k$</th>
<th>$\Theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0$</td>
<td>1.000</td>
<td>0.400</td>
<td>0.832</td>
</tr>
<tr>
<td>$k$</td>
<td>0.985</td>
<td>1.000</td>
<td>0.222</td>
</tr>
<tr>
<td>$\Theta$</td>
<td>0.996</td>
<td>0.983</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Values obtained in the principal (orthogonal) plane are reported in the upper-right (italicized in the lower-left) corner, respectively.
the literature (Widlowski et al., 2003) and span a large range of ecological situations. In particular, the inclusion of three widely different background brightness conditions (from dark to very bright) allows us to test the importance of the contribution of the uncollided radiation to the total signal; that is, the fraction of radiation that has travelled through the gaps of the canopy layer and has been scattered by the background only (Pinty et al., 2004a,b). The resulting overall BRF anisotropy may vary dramatically from strongly bowl-(with dark background) to bell-(with bright background) shaped patterns offering thus an appropriate ensemble of conditions to evaluate the RPV fitting performance. The simulations of the radiation transfer regimes characterizing this ensemble of forest systems are conducted in the red and near-infrared spectral domains for three different values of the Sun zenith angle, 25°, 45° and 65°, using the Rayspread model, a 3-D Monte Carlo ray-tracing model that implements the local estimator variance reduction technique for the BRF simulations (Widlowski et al., 2006).

Fig. 3. Scatterplots of the 3-D model-simulated BRF values against those reconstructed by the RPV model using the mean retrieved parameter values. The precise location of the original and reconstructed BRF values in the angular domain (Ω₀, Ω₁) is given in Fig. 4. Results from the 3 (4) parameter versions are shown on the top (bottom) panel. Labels Rbgd and Fc refer to the lambertian albedo of the underlying background and the vegetation fractional cover used in the simulations, respectively.

Fig. 4. Geometry of the viewing angles constraining (★) the inversion procedure on 3-D model-based simulations. The agreement between the reconstructed and original BRF values is assessed along different viewing directions (◊). The three inner circles (dashed lines) correspond to zenith angles of 25°, 45° and 65°, respectively. The backscattering regime occurs on the right hand side of the graph.

The 3 and 4 parameter versions of the RPV model inversion package were both applied on this ensemble of 3-D model-based simulations. The inversion procedure was activated on 756 different scenarios using 25 view angles distributed in both the principal and orthogonal planes, between 0° and 75°. Fig. 3 shows the scatterplots between the 3-D model-based simulations and those BRF values calculated with the forward RPV model at azimuthal and zenithal angles different from those originally used as input. The precise location of these exiting directions is sketched on Fig. 4. The very good unbiased agreement with quite limited scatter between the two BRF datasets illustrates the capability of both 3 (top panel) and 4 (bottom panel) parameter versions of the RPV model to provide accurate fits of BRF fields. As expected from results highlighted in the previous section, the RPV version using 4 parameters performs slightly better in terms of its fitting performance. This feature is, however, somewhat counterbalanced by the larger uncertainties on the retrieved mean parameter values as documented in Table 1, thus rendering the subsequent analysis of the mean retrieved values more difficult.

The histograms of the retrieved parameters obtained when using the 3 parameter version are shown in Fig. 5 for both red (red line) and near-infrared (black line) spectral bands. The
distributions of \( \rho_0 \) values (top left panel) are analogous in both spectral bands due to the large range in the background conditions that was specified in the 3-D model-based simulations. By contrast, the \( k \) parameter (top right panel) exhibits distinct (with limited overlap) PDFs in the red, where large absorption by leaves embedded in the vertical tree structures are simulated yielding a most frequent occurrence of bell-shaped BRFs (\( k>1 \) values) conditions, and in the near-infrared band, where scattering by vegetation elements is significant. This spectral difference confirms earlier conceptual findings by Pinty et al. (2002) and Widlowski et al. (2004) and suggests the possibility to derive vegetation structure information from multi-angle data, especially under bright background conditions (Nolin, 2004). The PDFs of the retrieved \( \Theta \) parameter (bottom left panel) in both spectral bands are slightly skewed towards negative values, thus indicating the predominance of backscattering regimes due to the decrease in the relative contribution of darkening caused by shadowing. The identification of the \( J(\mathbf{X}) \) function minima for these cases was achieved after a number of iterations ranging from about 5 to 40, as shown in Fig. 5 (bottom right panel).

Each of these 756 inversion cases (42 forest scenes \( \times \) 2 spectral bands \( \times \) 3 background brightness values \( \times \) 3 Sun zenith angle values) required \( 3.9 \times 10^{-3} \) seconds on average for the 3 parameter version and \( 7.7 \times 10^{-3} \) for the 4 parameter version. These small numbers correspond to the user time (defined here as the total number of CPU-seconds that the process spends in user mode) on a PC (dual Xeon 2.4 GHz) with 2 GB RAM, running under the Linux 2.6 kernel. The software was compiled using the GNU gfortran compiler (gcc-4.0.1) with no optimization flag. The average number of iterations required to find the minima of the cost function (Eq. (10)) are equal to 12 and 15 in the red and near-infrared spectral bands, respectively, when using the same \textit{a priori} knowledge, that is \( \mathbf{X}_{pr} \) for the

---

Fig. 5. Histograms of the RPV model parameters retrieved by the inversion package using 3 parameters against the ensemble of 3-D model-simulated scenarios. The top left, top right and bottom left panels display results for the \( \rho_0 \), \( k \) and \( \Theta \) parameter, respectively. The bottom right panel illustrates the histogram of the number of iterations. The red (black) lines correspond to retrievals obtained in the red (near-infrared) spectral band. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
parameter values and $C_d$ and $C_{X_p}$ for the covariance matrices, as those specified in Section 2. Since the values specified in $C_{X_p}$ do not add any constraints on the range of the parameter values to be retrieved, the 3 parameter inversion procedure returned a small set (3% of the cases) of “unrealistic” values whose combination however provides a good fit to the BRFs. This situation can be avoided easily by constraining the range of allowed parameter values.

3.2. MISR derived Surface BRFs

The Multiangle Imaging SpectroRadiometer (MISR) instrument on board NASA’s Terra platform acquires reflectance data from any Earth target in four solar spectral bands, from nine different directions, in at most seven minutes. In the default mode of operation, MISR data are downloaded from the platform at a spatial resolution of 275 m in all bands of the nadir camera and in the red band of the off-nadir cameras, and at 1.1 km in the remainder (see Diner et al., 1998). The MISR instrument thus samples the information related to the anisotropic patterns of the spectral radiation scattered by various types of land surfaces. As shown in Diner et al. (2005), these angular signatures can be used to obtain unique information about the geometric and physical properties of the environment. Specifically, under favorable conditions of illumination and background brightness, the BRF field in the red spectral domain can assume a bell-shape pattern. When this occurs, the $k$ parameter of the RPV model takes on values larger than unity, and this has been interpreted as indicating the presence of dark vertical structures such as trees over a relatively bright background at the MISR subpixel scale (see for instance Nolin, 2004; Pinty et al., 2002; Widlowski et al., 2004). Separately, the RPV model parameter $\Theta$ has been used to assess the presence of wetlands and shallow free standing water areas when the spectral signature is not sufficient (Pinty et al., 2003).

The RPV model inversion package (3 parameters version) was applied to block 103 (path 175) of MISR level 2 surface-derived BRFs (terrain-projected) showing contrasted surface conditions Fig. 6. Comparisons between the MISR Land BRFs (path 175, block 103, orbit 8650, August 3, 2001) and those simulated using the optimal set of mean RPV parameter values delivered by the inversion package using 3 parameters. The top (bottom) left and right panels corresponds to results obtained in the blue (red) and green (near-infrared) bands, respectively.
and encompassing the Mongu site in Zambia (15.44° S; 23.25° E). This site was instrumented in the context of the SAFARI 2000 project experiment1 and the characteristics of the vegetation cover over this area adjacent to the Zambezi river basin have been extensively documented during this field campaign (Shugart et al., 2004). The predominant vegetation belongs to a class of broad-leaf deciduous woodland with sparse woody vegetation understory. In this Kalahari woodland, the trees making up the overstory range from 8 to 12 m in height and cover about 50% of the understory which itself covers close to 30% of the ground (see for instance Privette et al., 2004; Scholes et al., 2004). The strongly seasonal rainfall induces changes in the Leaf Area Index (LAI) from about 0.5 (in August) to 2.0 (in January) (Huemmrich et al., 2005). The surface BRF values used in this study were generated by the MISR operational processor (version 17) on data acquired during the year 2001. Over the selected block, the widest spatial coverage with valid surface BRF estimates was found on August 3, 2001 (orbit 8650) and we thus focused first on the analysis of data acquired on this date. The inversions were performed by specifying a uniform standard deviation value corresponding to 5% of the mean MISR land BRF field in the diagonal covariance matrix associated with the measurements, C_d.

The comparison between the surface BRF values derived from MISR data and those reconstructed using the RPV model with the selected ensemble of parameter values delivered by the inversion package is shown on Fig. 6. These two datasets appear to agree to a large degree in all spectral bands, e.g., the root mean square error is equal to $5.6 \times 10^{-3}$ for the near-infrared band (bottom right panel) and close to $3 \times 10^{-3}$ for the cameras operating at the shorter wavelengths (top and bottom left panels). These values correspond to approximately 2% (in the near-infrared) to 5% (in the blue) relative error of the mean BRF values.

Fig. 7 (top panel) displays a false color composite of the mean values of the retrieved RPV $\rho_0$ parameter in the red, green and blue bands. This map exhibits strong spatial and spectral contrasts over this region, both primarily related to vegetation density, i.e., low (high) vegetation density are associated with high (low) $\rho_0$ values in the red, green and blue bands. Indeed, the spatial patterns of $\rho_0$ are visually very similar to those exhibited by the Fraction of Photosynthetically Active Radiation (FAPAR) (bottom panel) delivered at 275 m spatial resolution by the VEGetation Activity and Structure (VEGAS) algorithm which uses top of the atmosphere L1B2 MISR data (Gobron et al., 2002, Appendix I). Incidentally, these FAPAR values are in the same range as those reported from ground-based estimations at locations close to the tower (see Huemmrich et al., 2005).

The maps of the mean values of the parameters controlling the angular shape of the BRFs, namely $k$ and $\Theta$, are given on Fig. 8 for both the red and near-infrared spectral bands. As illustrated in previous studies (see for instance, Diner et al., 2005; Gobron et al., 2002; Gobron and Lajas, 2002) the mean values of these parameters exhibit large spatially coherent variations probably associated with specific surface cover features. The $k$ parameter in

1 http://daac.ornl.gov/S2K/.
the red band is rather contrasted with values both significantly larger and smaller than unity. These can probably be attributed to changes in vegetation density and related gappiness at the sub-pixel resolution. By comparison, the $k$ values in the near-infrared band vary within a more limited numerical range, although the signature of spatially organized surface features is still detectable. The maps of $\Theta$ exhibit predominantly a backscatter regime in both wavelengths, although the occasional detection of forward scattering areas suggests the presence of water bodies, such as lakes and shallow water bodies in the Zambezi river basin (blue color code in right panels of Fig. 8). The detection of such features, especially on the $\Theta$ maps, is sensitive to the MISR aerosol retrieval algorithm, which is also responsible for the occasional blockiness at 17.6 km (16 MISR pixels at 1.1 km resolution). As can be anticipated, this contamination is much more significant for parameter values retrieved from the blue band (not shown) than at any other wavelength.

This study underscores the complementarity of the traditional spectral analysis and of the anisotropy investigation made possible with MISR. For instance, the largest FAPAR values are associated, in the red band, either with bowl-shape (lowest $k$ values in red color) or bell-shape patterns (highest $k$ values in blue color) therefore suggesting the presence of different surface features most probably related to vegetation spatial heterogeneity. By the same token, areas exhibiting spatially uniform low FAPAR values (due to low vegetation density but also quasi bare soils, roads and water bodies) are associated with different features in the $k$ and $\Theta$ parameters, the latter being particularly relevant for shallow water bodies detection. Fig. 9 exhibits some of the information contained in the retrieved parameter space by

Fig. 8. Maps of the mean values of the $k$ and $\Theta$ parameters retrieved in the MISR red (top panels) and near-infrared (bottom panels) spectral band from data acquired on August 3, 2001 (path 175, block 103, orbit 8650). North is up in the figure and the horizontal extent of the image is the MISR swath, i.e., about 380 km. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 9. False color composite map of parameter $k$ retrieved in the MISR red band (coded in red color), FAPAR from L1B2 MISR data (coded in green color) and parameter $\Theta$ retrieved in the MISR near-infrared band (coded in blue color) from data acquired on August 3, 2001 (path 175, block 103, orbit 8650). Greenish hues imply a marginally higher vegetation density, reddish hues suggest a relatively structured environment while blue spots indicate free standing water. North is up in the figure and the horizontal extent of the image is the MISR swath, i.e., about 380 km. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
combining these elements into a false color composite image, where the parameter $k$, derived from MISR in the red spectral band is coded in red, the FAPAR estimated from MISR L1B2 is coded in green, and the parameter $\Theta$ in the near-infrared spectral band is represented in blue. The various hues can thus be easily related to different associations of vegetation density and greenness (greenish hues imply a marginally higher vegetation density), gap occurrence within the canopy cover (redish hues suggest a relatively structured environment) while blue spots indicate free standing water.

The PDFs of standard deviation values on the three RPV model parameters over that same MISR block are reported in Fig. 10. The large majority of the values taken by these standard deviations, all spectral bands included, are smaller than $0.6 \times 10^{-2}$, $3.5 \times 10^{-2}$ and $1.4 \times 10^{-2}$ for the $\rho_0$ (top left panel), $k$ (top right panel) and $\Theta$ (bottom left panel) parameters, respectively. Since these values are much smaller than the variations exhibited in the spatial patterns of Figs. 7 and 8, the spatial organization suggested by these figures is statistically significant. The effect of blockiness mentioned previously translates into relatively large values of the cost functions (bottom right panel) associated with these retrievals in the blue spectral band. As can be logically expected, these values are decreasing with an increase in the spectral location of the bands, i.e., the best retrievals are thus achieved for the near-infrared band. This same spectral trend seems to affect, but only slightly, the fields in the $k$ and $\Theta$ parameters as well over this particular region.

This analysis was repeated throughout the year 2001, whenever MISR land BRF products were available for that same MISR block. Fig. 11 shows time profiles of the $\rho_0$ and $k$ parameters together with those delivered by the operational

![Fig. 10. Probability density functions of the standard deviation values of the RPV model parameters retrieved by the inversion package using 3 parameters against the MISR land BRFs (path 175, block 103, orbit 8650, August 3, 2001). The top left, top right and bottom left panels display results for the $\rho_0$, $k$ and $\Theta$ parameter, respectively. The bottom right panel illustrates the PDF of the cost function. The blue, green, red and black lines correspond to results obtained with the MISR blue, green, red and near-infrared spectral bands, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)](image-url)
processor (version 17) in the red and near-infrared spectral bands. These profiles are relative to the four MISR pixels (1.1 km resolution) located within the vicinity of the Mongu tower site. Results from June 31, 2001 were excluded from these figures because very large numerical values of the cost function $J(X)$ (Eq. (10)) indicate unreliable retrievals. Further inspection of the MISR products suggests that rather unfavorable but unflagged atmospheric conditions were prevailing over the Mongu tower site on that particular date.

Fig. 11 illustrates that both sets of parameters are in excellent agreement and are both able to account for the seasonal variations in these parameters. The spatial variability between results obtained over these 4 MISR pixels is also very limited and with low values of the cost function, thus adding a good level of confidence on the observed patterns. The $\rho_0$ parameter in the red band gradually and smoothly increases from about day 100 up to approximately day 250 and thereafter drops down to lower values. This evolution, in full agreement with ground-based observations of vegetation activity (see for instance Privette et al., 2004), is not however observable in the near-infrared band. This suggests that the decrease of solar absorption in the photosynthetic range is not accompanied by a change in the multiple scattering regime. In the mean time, the $k$ parameter values retrieved in the red (near-infrared) exhibit a decreasing (increasing) trend over the same period of time. These patterns, when considered together, can be interpreted as a consequence of the leaf loss at the onset and during the dry season (see for instance Shugart et al., 2004, Fig. 1). Indeed, at the end of the wet season (close to the beginning of our time series) green healthy (absorbing in the photosynthetic domain) leaves are clumped into trees covering only a fraction of the background. This situation yields low $\rho_0$ and large (eventually larger than unity) $k$ values in the red band associated with low values (less than 0.8) of $k$ in the near-infrared band. As the dry season proceeds, a smaller amount of green leaves is present and thus their obscuring effect of the uncollided radiation (scattered by the background) that exits the canopy layer at large angles is lessened. This logically translates into an increase (decrease) of the bowl-shape signature in the red (near-infrared) bands (Pinty et al., 2002). It is noteworthy that the last MISR sample in our time series suggests that the opposite trend occurs at the end of the dry season. These limited examples of maps and time series illustrate the added value provided by the thorough analysis and interpretation of the BRF angular shape.

4. Conclusions

This paper describes a new inversion package for the non-linear RPV model which capitalizes on recent computer software techniques (such as automatic differentiation used to generate the adjoint and Hessian code of a cost function) and evaluates its performances. This package inverts the RPV model in a numerically accurate and computer efficient manner and delivers extensive statistical information on the results, allowing the user to appreciate the quality of the retrievals and the performance of the procedure. This information is extremely valuable to properly interpret the RPV model parameter values.

The current package is designed to address the inversion of both the 3 and 4 parameter versions of the RPV model, i.e., including the optional hot spot parameter. It was shown that, as
can be expected from logical considerations, the 4 parameter version provides slightly better fits to a large ensemble of BRF fields, but at the cost of higher uncertainties on the retrieved parameter values. This study also confirms that when multi-angular observations are plentiful, measurements taken in the principal plane are more constraining on the inversion procedure and, therefore, yield more precise results than when they are acquired in the orthogonal plane. The gain in knowledge and the improvement in product accuracy and precision are documented and quantified by the analysis of the values populating the \textit{a posteriori} covariance matrix. Since this covariance matrix is estimated optimally from the Hessian code, this gain in knowledge is not penalized by a significant increase in computing resources as it is generally the case with traditional inversion techniques. The current inversion package thus removes the necessity of linearizing/simplifying the RPV formulation for the sole purpose of improving the speed of its inversion.

A series of tests has been conducted first against an extensive set of 3-D model simulated BRFs exhibiting somewhat complex angular shapes. These applications of the RPV inversion package confirm the excellent fitting capabilities of the RPV model and the robustness of the inversion procedure over a wide range of BRF conditions. A second series of inversion exercises were performed against MISR land BRF products generated by the operational processor (version 17). The selected MISR-derived data sets correspond to products available during year 2001 over the particular MISR block encompassing the SAFARI 2000 Mongu area located in Zambia. The main conclusions from this latter application can be summarized as follows: 1) the 3 parameter RPV version provides good fits to MISR land BRF products with an unequivocal information about the uncertainty levels in both the quality of the fit and the probability density functions of the model parameters, 2) even though the current version has not yet been optimized with respect to computing performance criteria, the package already operates in a quite computer efficient manner as can be assessed from a large number of sampling and scene scenarios, 3) the RPV parameter values representing the amplitude and the bowl to bell shape of the angular BRF fields, namely $\rho_0$ and $k$ are in excellent agreement with the analogous products generated by version 17 of the MISR operational processor and, 4) the inversion delivers spatially coherent fields of the RPV parameters representing the amplitude and controlling the anisotropy of the surface. These fields, when associated with limited uncertainties, allow the user to analyze further the surface anisotropy patterns. Such analyses confirm earlier findings regarding the benefits from multiaangular data acquisition in particular, the potential to elaborate on the processes leading to vegetation seasonal cycles.

Acknowledgements

The authors would like to thank the NASA Langley Research Center Atmospheric Sciences Data Center for providing them with the MISR products. This research would not have been possible without the support of the Global Environment Monitoring unit of the Institute for Environment and Sustainability at the Joint Research Centre. The RPV inversion package (software written in FORTRAN 90 language) implementing the procedure presented in this paper is available upon request by sending a message to thomas.lavergne@jrc.it and bernard.pinty@jrc.it.

References


