Development Of An Operational Procedure To Estimate Surface Albedo From The SEVIRI/MSG Observing System In Using Polder BRDF Measurements

Igor Pokrovsky and Oleg Pokrovsky

Voeikov Main Geophysical. Observatory, 7 Karbyshev Str., St.-Petersburg, 194021, Russia

Abstract— A statistical inversion method is presented in support to the application of kernel-based BRDF (Bi-directional Reflectance Distribution Function) models for the calculation of the surface albedo. We present an operational procedure for the inversion of kernel-driven BRDF model and further albedo retrieval to be applicable to the SEVIRI/MSG reflectance measurements. The processing steps applied to space-borne POLDER sensor data were as follows: 1) quality control, 2) accumulation of a priori information on model coefficients of directional hemispherical reflectance, 3) implementation of the BRDF model inversion methods based on the biased estimation instead of usual non-biased least solution, which has too big variance.

Keywords: ***BRDF model inversion; albedo retrieval; a priori information; data control procedure; Fisher and T statistics ; biased estimation***

1. INTRODUCTION

The determination of a surface albedo product requires the implementation of a number of data processing steps in which the cloud screening and the removal of atmospheric effects, merely water vapor content and aerosols loading, must concentrate our efforts. Furthermore, satellite systems provide a sparse angular sampling of the Bi-directional Reflectance Distribution Function (BRDF) when a hemispherical knowledge of this latter is necessary to derive an albedo variable. Embarked on a sun-synchronous platform, the AVHRR sensor scans everyday the same target under varying viewing conditions and similar sun geometry. After a few days, it provides a sampling of the BRDF close to the principal plane – the plane containing the sun and the target – in both forward and backscattering directions. On the other hand, the SEVIRI/MSG mission, as it is the characteristic of geo-stationary sensor systems, will provide a large variety of solar angular measurements but at fixed view zenith angle and for various sets of relative azimuths. In this latter case, the BRDF sampling will be a warping of the perpendicular plane in the backscattering area, away from the tropical belt and at the exception of the summer season. It is anticipated that the lack of having data in the principal plane – where angular effects are amplified – will lead to biased estimates of the BRDF and thereby surface albedo. Since only restricted angular sampling for land surface reflectance measurements is available, it yields a difficulty to calibrate BRDF models. In many cases the ill-conditioned index (ICI - the ratio of minimum to maximum eigen-values of matrix to be inverted) lies in the interval of small values. Therefore, the retrieved BRDF model coefficients are sensitive to small perturbations in the reflectance values registered by the satellite sensor. As a result there is a possibility to obtain unphysical solutions. In some respect this situation might be improved by the selection of optimal angular subsets providing higher ICI values than its initial values (Pokrovsky and Roujean, 2002 a). But, nonetheless, inevitable contamination of measured signal coming from the land surface through the atmosphere becomes a question of a serious concern. Any atmospheric correction module cannot solve this problem principally because of frequent impact of partial cloudiness, which is not properly detected by existing remote sensing methods (Bicheron and Leroy, 2000) and cannot be described by radiation transfer methods. This paper is aimed to consider above problem with application to SEVIRI/MSG geometry case.

2. DATA QUALITY CONTROL

2.1 Measurement Data

We note **b** the vector standing for coefficients $k_{i=0,1,2}$ of the BRDF parameter model. The traditional Gauss-Markov regression model reads:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{b} + \boldsymbol{\varepsilon} \quad (1)$$

Having **y** as an *n*-by-1 vector of measurements, **A** as a given *n*-by-*p* matrix of predictors, **b** as a *p*-by-1 vector of model coefficients to be estimated, $\boldsymbol{\varepsilon}$ as an *n*-by-1 vector of random disturbances having unknown variance matrix of standard form, $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = \sigma^2 \mathbf{I}$. Usually, the model (1) is related to the minimization of the quadratic norm of residual vector

$$\mathbf{r} = \mathbf{y} - \mathbf{A} \cdot \hat{\mathbf{b}} = \mathbf{y} - \hat{\mathbf{y}} \quad (2)$$

In (2), $\hat{\mathbf{y}}$ stands for vector of the model predictor. Then, a possible estimate of **b** is:

$$\hat{\mathbf{b}} = (\mathbf{A}^T . \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{y} \quad (3)$$

('T' is a sign of transpose and '-1' a sign of inverse matrix). However, this definition has poor numerical properties. Particularly dubious is the computation of $(\mathbf{A}^T \cdot \mathbf{A})^{-1}$, which is both costly and imprecise. To avoid those difficulties, most prevalent methods are the singular value decomposition (SVD) and QR decomposition of **A**, **A=Q·R,** where **Q** and **R** are the orthogonal and triangular matrices, respectively (Pokrovsky and Roujean, 2002 b). The residue in (2) might be rewritten in the following form:

$$\mathbf{r} = (\mathbf{I} - \mathbf{A} \cdot (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T) \cdot \boldsymbol{\varepsilon} \quad (4)$$

We consider now the residual in more details and begin the demonstration with the simple case of having a single angular measurement y_i. The posterior probability density function (pdf) for a single element of residual, say $r_i = y_i - \mathbf{a}_i^T \cdot \mathbf{b}, (i = 1,2,..., n)$, where $\mathbf{a}_i^T$ is a *i*-th row of **A,** is in the form of a univariate Student-t pdf with a posterior mean

$$\hat{r}_i = y_i - \mathbf{a}_i^T \cdot \hat{\mathbf{b}} (i = 1,2,..., n) \quad (5)$$

and variance $\sigma_i^2 = \mathbf{a}_i^T \cdot (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{a}_i \nu s^2 / (\nu - 2), (\nu > 2)$. That is,

$$t_\nu = (r_i - \hat{r}_i) / s[\mathbf{a}_i^T \cdot (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{a}_i]^{1/2}, (\nu > 2) \quad (6)$$

an univariate Student-t pdf with ν degrees of freedom. Note that the posterior mean of (5), which is a least square residual, is its optimal estimate relatively to a quadratic loss function appeared in method of least squares.

We assume that in (5)-(6) the estimate for σ^2 is used in form of $s^2 = \hat{\mathbf{r}}^T \cdot \hat{\mathbf{r}} / (n - p)$. In other words, the estimated standard deviation of the *i*-th residual r_i is s_i, where s_i^2 is the *i*-th diagonal element of the covariance matrix for (4).

The statistical tests to detect outliers in linear models were the focus of a number of studies. Most of them were based on obtaining residuals standardized by their individual standard deviations (Draper, and Smith, 1981). A main advantage of this test procedure is its simplicity in regard to its general degree of application to any linear models without the necessity of a re-analysis of the suspected outlier either omitted or treated as missing. Determination of the exact percentage points of test statistics is based on the standardized residual $\left(r_i / s_i\right)$. Tietjen et al (1973) proposed a test procedure for the detection of a single outlier in a simple linear regression model and determined critical values of test statistics:

$$R_n = \max_i \left| r_i / s_i \right| \quad (7)$$

using large-sized samples. Critical values of R_n for significance levels $\alpha = 0.1, 0.05, 0.01$ were given for a range of sample size up to *n*=100. We consider also here alternative statistics for the present case of BRDF linear model:

.

2.2 Output Coefficients

The classical work in field of computerized data screening related to census data has been carried out by O'Reagan (1969). Later, there has been an expansion in the use of bases consisting of vectors of data following multivariate distribution (Gnanadesikan, and Kettenring, 1972). This and other following papers were based on implementation of the generalized likelihood ratio *T* statistics of Hotelling (Anderson, 1958). We used here a similar statistical procedure to reveal strange BRDF model coefficient patterns. Let the vector **b** represent a set of observations distributed in accordance to the normal law $N(\bar{\mathbf{b}}, \mathbf{\Sigma}_\mathbf{b})$. As we obtained some sample of BRDF coefficient sets, it is easy to compute the estimates for both of them: vector of means $\bar{\mathbf{b}}$ and covariance matrix $\mathbf{\Sigma}_\mathbf{b}$. So, we may assume that $\bar{\mathbf{b}}$ and $\mathbf{\Sigma}_\mathbf{b}$ are known. We also suppose that after simple transformation $\bar{\mathbf{b}}$ and $\mathbf{\Sigma}_\mathbf{b}$ have been scaled to zero means and unit variances. Let $\tilde{\mathbf{b}} = \mathbf{b} + \mathbf{e}$ be the input vector to be tested. Here $\mathbf{e}$ is a vector of data capture errors. Then the screening of $\tilde{\mathbf{b}}$ consists of a test of null H_o hypothesis $H_o : \mathbf{e} = \mathbf{0}$. One possible alternative hypothesis to is $H_1 : \mathbf{e} \neq \mathbf{0}$: it means that an arbitrary error vector is present. In this case the test statistic indicated is the generalized likelihood ratio *T* statistics:

$$T = \tilde{\mathbf{b}}^T \cdot \mathbf{\Sigma}_\mathbf{b}^{-1} \cdot \tilde{\mathbf{b}} \quad (8)$$

which has a good performance for a wide class of error vectors $\mathbf{e}$ (Anderson, 1958). The preceding discussion is based on the implicit assumption that $\bar{\mathbf{b}}$ and $\mathbf{\Sigma}_\mathbf{b}$ are known. In practice, it will be rarely the case. Currently, we have available some a priori sample of N vectors, which have been verified manually and assumed to be accurate. From this sample we compute estimates $\breve{\mathbf{b}}$ and $\mathbf{S}_\mathbf{b}$ and use these estimates in place of $\bar{\mathbf{b}}$ and $\mathbf{\Sigma}_\mathbf{b}$. A serious question then arises as to how large *N* must be for this substitution to yield acceptable accuracy in the distribution of test statistics. In the case of *T*, the distributional problem is well investigated (Anderson, 1958). In fact, if $\breve{\mathbf{b}}$ and $\mathbf{S}_\mathbf{b}$ are used in place of $\bar{\mathbf{b}}$ and $\mathbf{\Sigma}_\mathbf{b}$, then *T* is computed in accordance to (8) as $T = (\mathbf{b} - \breve{\mathbf{b}})^T \cdot \mathbf{S}_\mathbf{b}^{-1} \cdot (\mathbf{b} - \breve{\mathbf{b}})$ for an arbitrary incoming data vector $\mathbf{b}$.

3. COMPARISON OF INVERSION TECHNIQUES

The Least Square (LS) solution of (1) can be obtained by minimization of cost function $\min_\mathbf{b} (\mathbf{A} \cdot \mathbf{b} - \mathbf{y})^T \cdot \mathbf{\Sigma}_\mathbf{\varepsilon}^{-1} \cdot (\mathbf{A} \cdot \mathbf{b} - \mathbf{y})$ (see (Pokrovsky, O.M., and J.L. Roujean, 2002 b)), with $\mathbf{\Sigma}_\mathbf{\varepsilon} = \sigma^2 \mathbf{I}$. Bayes formulas for conditional probability permit to come to the generalized form for the cost function when we dispose of a priori statistics:

$$\min_\mathbf{b} \{(\mathbf{A} \cdot \mathbf{b} - \mathbf{y})^T \cdot \mathbf{\Sigma}_\mathbf{\varepsilon}^{-1} \cdot (\mathbf{A} \cdot \mathbf{b} - \mathbf{y}) + (\mathbf{b} - \bar{\mathbf{b}})^T \cdot \mathbf{\Sigma}_\mathbf{b}^{-1} \cdot (\mathbf{b} - \bar{\mathbf{b}})\} \quad (9)$$

Minimum of (9) is a regularized solution of problem (1). Details on numerical algorithms related to (9) can be found in (Pokrovsky O.M., 1984).

Hoerl and Kennard (1970) suggested to use the ridge regression, as a procedure for dealing in the case of multi-collinear columns of matrix **A** in a regression model like (1). Their ridge estimator of the standardized (Anderson, 1958) regression model coefficient vector is

$$\hat{\mathbf{b}}(\beta) = (\mathbf{A}^T \cdot \mathbf{A} + \beta \cdot \mathbf{I})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{y} \quad (10)$$

where β is a positive constant and $\mathbf{A}^T \cdot \mathbf{A}$ is in correlation form (Anderson, 1958). The selection of parameter β yields the major issue. Most studies on this subject suggested more or less empirical rules, according to the problem to be solved. Here, we use the acceptability conception as a background for more substantiated selection rule. One example of inversion technique comparison is presented at fig.1. It showed that ordinary LS solution (statistical inversion) provides unphysical solution for albedo retrieval and alternative methods: statistical regularization and ridge regression permit to obtain solutions, which are in reasonable agree with land surface measurement data.

4. CONCLUDING REMARKS

An innovative work has been presented in which routines were developed for control procedure to detect outliers at input and output stages of the BRDF model inversion based on both statistical and physical criteria. Two principal statistics frequently used in multiple regressions were considered: the model explained variance and the ratio Fisher statistics. Another important step was the accumulation of BRDF model coefficients and albedo a priori statistics. This statistical information permit us to sort out a more sophisticated inversion technique. In fact, the filtering procedure for outliers is more specific to sensor characteristics than models.

The ill-posed problem of kernel matrix inversion could be solved based on the ridge regression and other biased estimates in order to increase the reliability of biophysical parameters. It is worth noting that ill-conditioning does not only rely on restricted angular sampling but also occurs when co-linear kernels are considered. Further improvements in albedo retrieval will be to perform an optimal selection of kernels for composite model and to adapt SR method to composite model.

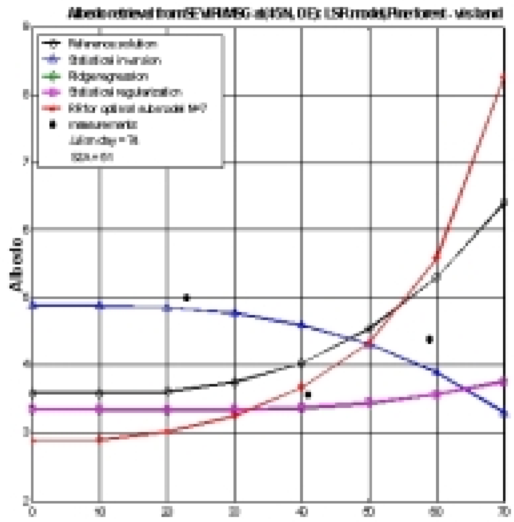

Figure1. Example of albedo retrieval from MSG/SEVIRI data

5. Acknowledgments

This work has been performed in the framework of the Satellite Application Facilities (SAF) program, which is supported by the European Meteorological and Satellite (EUMETSAT) agency and by the National Meteorological Services. We also thank Météo-France for technical assistance. Oleg Pokrovsky and Igor Pokrovsky received a financial support from EUMETSAT as a Visiting Scientists at CNRM. Jean-Louis Roujean is with the Centre National de la Recherche Scientifique. This study was also supported by Russian Fond of Background Research grants 01-05-65283 and 02-05-64 757.

References

Anderson, T.W., 1958, An Introduction to Multivariate Statistical Analysis, N.Y., John Wiley and Sons Inc., 548 p.

Bicheron, P., and M.Leroy, 2000. BRDF signatures of major biomes observed from space. J. Geophys. Res.,v. 105, p.26669 - 26681.

Draper, N. and H. Smith, 1981, Applied regression analysis, second edition, N.Y., John Wiley and Sons Inc., 1981, 467 p.

Gnanadesikan, R. and J.R. Kettenring, 1972, Robust estimates, residuals and outliers detection in multi-response data., Biometrics, v.28, p.81-124.

Hoerl A.E., and R.W. Kennard, 1970, Ridge regression: biased estimation for non-orthogonal problems., Technometrics, v.12, p.55-67

Pokrovsky O.M., 1984. Meteorological Remote Sensing of the Atmosphere from Satellites.-Leningrad, Hydrometeoizdat , 287 p.

Pokrovsky, O.M., and J.L. Roujean, 2002 a, Land surface albedo retrieval via kernel-based BRDF modeling: II. An optimal design scheme for the angular sampling , Remote Sens. Environ., 84, 120-142.

Pokrovsky, O.M., and J.L. Roujean, 2002 b, Land surface albedo retrieval via kernel-based BRDF modeling: I. Statistical inversion method and model comparison., Remote Sens. Environ., 84, 100-119..

Tietjen, G.L., Moore R.H. and Beckman R.J., 1973. Testing for a single outlier in simple linear regression. Technometrics, v.15, p.717-721

O'Reagan, R.T., 1969, Relative cost of computerized error inspection plans., Journal of American Statist.Assoc., v.64, p.1245-1255.