A Priori Knowledge Accumulation and Its Application to Linear BRDF Model Inversion

Xiaowen Li[1,2], Feng Gao[2,3], Jindi Wang[1] and Alan Strahler[2]

1 Research Center for Remote Sensing, Beijing Normal Univ., Beijing 100875, China

2 Dept. Geography and Center for Remote Sensing, Boston Univ., 725 Commonwealth Ave., Boston, MA 02215, USA

3 Nanjing Institute of Geography and Limnology, CAS, Nanjing 210008, China

Abstract. *A priori* knowledge can significantly improve the retrieval of surface bidirectional reflectance and spectral albedo from satellite observations. Here, *a priori* knowledge takes the form of field measurements of bidirectional reflectance factors for various surface cover types in red and near−infrared bands. Bidirectional reflectance and albedo retrieval refers to inversion of a kernel−driven bidirectional reflectance distribution function (BRDF) model using surface reflectance observations derived from orbiting spacecraft. *A priori* knowledge is applied when noise and poor angular sampling reduce the accuracy of model inversion given a limited number of observations. In such cases, *a priori* knowledge can indicate when retrieved kernel weights or albedos are outside expected bounds, leading to a closer examination of data. If data are noisy, *a priori* knowledge can be used to smooth the data. If the data exhibit poor angular sampling, *a priori* knowledge can be used according to Bayesian inference theory to yield *a posteriori* estimates of unknown kernel weights. In the latter application, Bayes theory is applied in data space rather than in parameter space. Extensive study and simulation using 73 sets of field observations and 395 spaceborne observation sets from the POLDER instrument validates the importance of *a priori* information in improving inversions and BRDF retrievals.

Keywords: BRDF, linear model, inversion, *a priori* knowledge

1. Introduction

As advances in the field of multiangle remote sensing progress, it becomes increasingly likely that BRDF (bidirectional reflectance distribution function) models can be inverted to estimate structural parameters and spectral component signatures of earth surface cover types[1]. Therefore, quantitative remote sensing seems to be on the near horizon. However, inversion of geoscience models is a very difficult problem that still needs study from the view−point of information theory [2] and the comprehensive practice of model inversion [3−5]. The real physical system that couples the atmosphere and land surface is sufficiently complex that it requires many, many parameters to describe it faithfully. Any practical physical model can only be an approximation of this real system, and a good model will have limited number of the most important parameters that capture the major variation of the real system. We denote such a forward model as

$$z = f(X, S) \tag{1}$$

where z is single measurement; X is a vector of controllable measurement conditions such as wave band, viewing direction, time, sun position, polarization, etc.; and S is a vector of m parameters of the system approximation.

Now, with ability of the satellite sensors to acquire multiple bands, multiple viewing directions, etc., while keeping S essentially the same, we can obtain the set of simultaneous equations

$$D = f(X, S) + e, \tag{2}$$

where D is now a vector in M−dimensional measurement space, i.e., with M values corresponding to M different measurement conditions; and e is the vector of random noise. It has long been hoped that if M is large enough, S can be determined from observations. However, remotely sensed observations are usually more or less correlated. Moreover increasing M (such as adding spectral bands) may bring in more unknown parameters (such as spectral reflectance of surface elements in new bands). Thus, the remotely sensed signal, no matter how fine its spectral and angular resolution, has only limited information. Therefore, BRDF inversion problems, like those in geoscience generally, are usually underdetermined, at least for some parameters of S, and the utilization of *a priori* knowledge is necessary.

As the ancient philosopher Confucius pointed out, "Our knowledge consists of two parts − what we know, and what we know we don’t know." In the case that remote sensing signals contain limited but valuable information, it is important to extract information about what we don’t know or what is uncertain, rather than to invert for all, pretending that we know nothing. Using this principle in earlier work, we expressed *a priori* knowledge of model parameters as best guesses with associated uncertainties. The results seemed encouraging [5], and thus we tried to formalize the approach [2].

The ideal expression of *a priori* knowledge of model parameters is a joint probability density $p_S(S_1, S_2, \ldots S_m)$; and the expression for *a priori* knowledge of the model accuracy and measurement noise is a conditional joint probability density $p_D(d_{obs}|S)$. The *a priori* probability density function (PDF) of the observations is then

$$p_D(d_{obs}) = p_D(d_{obs}|S) * p_S(S) \tag{3}$$

The formal way to apply *a priori* knowledge in our model inversion is based on Bayes' Theorem, which states that because

$$P(x, y) = P(x|y)P(y) = P(y|x)P(x) \tag{4}$$

we have

$$P(x|y) = P(y|x)P(x)/P(y) \tag{5}$$

When we interpret the $P(y|x)$ as the prior knowledge of model prediction of d_{obs}, giving parameters in parameter space, and *P(x)* as the prior knowledge of parameters, and *P(y)* as prior knowledge of marginal density of observations, then (6) can be used as an inversion

$$p(S|d_{obs}) = p_D(d_{obs}|S) * p_S(S) / P_D(d_{obs}) \qquad (6)$$

where the marginal probability of observation $P_D(d_{obs})$ is

$$P_D(d_{obs}) = \int_S p_D(d_{obs}|S) * p_S(S) * dV_s \qquad (7)$$

where dV_s is a differential volume element in the parameter space. Eq. (6) is usually called Bayesian Inference in theory of nonlinear parameter estimation. It has also been rather widely applied in geophysical inversion [6] and atmospheric remote sensing [7], probably because in those fields the inverse problems are even more ill−posed than in land surface remote sensing.

An important feature of Bayesian inversion is that there is no prerequisite number of independent observations required for a successful inversion. So long as new observations are acquired, *a priori* probability density in parameter space can be modified to obtain posterior density, and thus knowledge can be accumulated.

In practice, we may classify the *a priori* knowledge into different levels: general knowledge about the land surface ("global knowledge"), land cover type−related knowledge, target−specific knowledge etc. The means to accumulate knowledge at these different levels may be different, but include the following:

1) Applicable forward model(s);

2) Physical limits and probability density in model parameter space;

3) Statistics of model accuracy and noise in remote sensing signals;

4) Seasonal change associated with land cover types or targets;

5) Confidence of the above knowledge.

Note that even a single observation can change the *a priori* probability density function (PDF) of more than one parameter significantly. In [2], we provided an example how this is achieved in parameter space. However, the required numerical integration in parameter space is time−consuming whenever *m* is large, as in our application to retrieval of surface BRDF and albedo from satellite data.

In this paper, we first present the way we accumulate *a priori* knowledge of land surface BRDF, then introduce our approach to application of this knowledge for the BRDF/Albedo products in the MODIS data stream [1, 24].

2. Kernel−Driven BRDF Models

Linear kernel−driven BRDF models were designed to ease the difficulties of inverting nonlinear physical models, at the expense of some approximation of the original physics. A linear kernel−driven BRDF model has the following form

$$BRDF = f_{iso} + f_{vol} * k_{vol}(t_i, t_v, \phi) + f_{geo} * k_{geo}(t_i, t_v, \phi) \tag{8}$$

where k_{vol} and k_{geo} are "kernels", i.e., known functions of illumination and viewing geometry that describe volume and geometric scattering respectively; t_i is the zenith angle of the solar direction; t_v is the zenith angle of the view direction; ϕ is the relative azimuth of sun and view directions; and f_{iso}, f_{vol} and f_{geo} are three unknown coefficients to be adjusted to fit observations. With more than three uncorrelated multiangular observations, a regression method can provide estimates of the three parameters much more easily than least squares error fitting of a nonlinear model.

The original design of our inversion algorithm [1] included a collection of different kernels for various land cover types. However, the user community prefers a single combination of kernels. After an extensive validation effort [8, 9], it was determined that the combination of RossThick (k_{vol}) and LiSparse (k_{geo}) kernels has the best overall ability to fit BRDF measurements and to extrapolate for BRDF and albedo [10]. Generally speaking, the RossThick kernel characterizes a weak hotspot with a "bowl shape" [9] at large zenith angles, while the LiSparse kernel provides a dome−shaped BRDF [9] with a stronger hotspot. More recent work [11] suggests that the LiTransit kernel, which combines LiSparse and LiDense in a transition, may give yet better fitting. For details please refer to [10, 11]. We use the RossThick−LiTransit combination for the applications presented in this paper.

Because of the linearity of the kernel−driven BRDF model, the inversion is very straightforward and fast, if the number of looks (NOL) is large enough and the directions of the looks are well distributed. The inversion employs a typical LSE (least squared error) approach, i.e., when $Y_{obs}[M]=BRDF$ is the observation vector, $A[M][3]$ is the kernel matrix for the given M geometries, the vector of unknown parameters

$X[3]=[f_{iso}, f_{vol}, f_{geo}]'$ is solved for by

$$X=[A'A]^{-1}A'Y_{obs} \tag{9}$$

which is a well−known solution for minimizing the sum of squared fitting error (SSE)

$$SSE=Cost(X)=(A*X-Y_{obs})'(A*X-Y_{obs}) \tag{10}$$

In the case that the noise covariance matrix C_d in the observation Y_{obs} is known, normalization should be taken so that the solution is obtained to minimize

$$Cost(X)=(A*X-Y_{obs})'C_d^{-1}(A*X-Y_{obs}) \tag{11}$$

In present practice, the noise covariance matrix is rarely known. At best, we can sometimes make intelligent guesses about the noise levels (variances) for different bands and viewing directions and normalize the observations accordingly. Such knowledge may be accumulated with future practice in BRDF inversions.

After a successful inversion for the three parameters, the results are used to obtain the pixel's white−sky albedo (WSA) and the black−sky albedo (BSA) at solar zenith angles 0, 30, 45, 60° and [4]. The former is the BRDF integrated over all view and sun positions, and so it corresponds to the albedo under conditions of isotropically downwelling radiance (perfectly white sky). The latter is the BRDF integrated over all viewing positions, so corresponds to the albedo under conditions of unidirectional collimated beam irradiance (perfectly black sky).

Note that the condition that $[A'A]^{-1}$ exists (i.e., $A'A$ is far from singular) can not be easily met for satellite remote sensing because of restrictions in view and illumination geometries. As a result, we have studied the application of *a priori* knowledge from the beginning of algorithm development for the MODIS/MISR BRDF/Albedo product[1]. The at−launch algorithm applies *a priori* knowledge in an approach called "magnitude inversion" when the inversion is ill−posed. In future enhanced versions, the application of *a priori* knowledge will be further developed and tested with the accumulation of *a priori* knowledge.

3. Initial Knowledge Base in the Form of Kernel Model Parameters

A total of 29 field measured BRDF data sets were used to initialize our land surface BRDF knowledge base. These data sets were collected by Kimes *et al.* [12], Ranson *et al.* [13], Deering and Leone [14], Irons *et al.* [15], Leroy *et al.* [16], and made available to us by the courtesy of these investigators. Among these 29 data sets, 27 were previously used for algorithm tests [8]. More details on these data sets are summarized in [8]. The data provide some ideas about what a land surface BRDF should look like. However, BRDF as a function of four variables $(t_i, t_v, \phi,$ and wavelength) is not easy to display or comprehend or to use as *a priori* knowledge.

According to Li et al. [2], *a priori* knowledge can be represented in different forms: 1) (wide−bound form) hard−bounded ranges of parameter values, usually applicable for physical limits; 2) (δ−bound form) some parameters (or relations among them) are accurately known, or we know they are insensitive to the given observation geometry, and can be fixed; 3) (soft−bound form) unknown parameters have *a priori* known joint probability density function (JPDF), which we call a soft−bound range. All three kinds of *a priori* knowledge can be used in inversion depending upon the situation. In this application, what we need is to accumulate soft−bound knowledge for the three kernel coefficients f_{iso}, f_{vol} and f_{geo}. Because of the semiempirical nature of the model, we don't have physical limits for them. On the other hand, without soft−bound knowledge, it is unlikely we can have any δ−bound knowledge. Therefore, we decided to construct our initial knowledge base in form of statistics of inverted model parameters from these 29 data sets. One problem is the difference in instruments used to acquire the data. Since we have no basis for cross−calibrating them, the knowledge must be very raw − but still better than nothing to start. Table 1 lists the statistics of the 29 sets of three RossThick−LiTransit coefficients. We'll refer these data as the 29−collection.

Table 1. The statistics of the 29 sets of three RossThick−LiTransit coefficients

Strugnell and Lucht collected 68 BRDF data sets [17]. Among these 68 sets, those from Kimes, Ranson, Irons, and some from Deering are the same as those in the initial 29−collection. Other contributions include additional sets from Deering, as well as data of Eidenshink and Faundeed [18] and Tsay et al. [19]. More details (including names of sets) are provided in [17].

As an initial experiment, we applied our knowledge base from the 29−collection to the Strugnell and Lucht collection for identifying problems in data compatibility. We used the RossThick−LiTransit model to invert all 46 new datasets and then checked them for f values more than 2 standard deviations from the mean. We treated the datasets beyond these bounds as "strange," and examined them with results as follows.

(1) 1988/chuck.site (Deering): This data set has only one directional observation. Apparently, there were unrecoverable errors in data transition, so this set is deleted.

(2) SCAR/forest (Tsay): This data set has very fine viewing zenith (VZN) angular resolution (1 degree in both viewing zenith from 0 to 90, and azimuth from 1 to 360), but only one solar zenith (SZN) position (56.67 degrees). There is a very strong forward scattering peak from 70 to 90 degrees, possibly due to smoke aerosols. The current model kernel combination does not model this feature well. In addition, no other data sets have measurements at such large viewing zenith angles. So we decided to delete measurements beyond 70 degrees of viewing zenith angle. (We may include the full data set later, when specular features are added to the model and more measurements at such large viewing zenith have been acquired.)

(3) 1988/dune−flat (Deering): Like the chuck.site, there were errors in data transfer, and thus we delete this data set.

(4) 1987/ifc1−site2 (Deering): There was error in band information processing. We found and corrected the error and then the inversion results are no longer strange.

(5) 1988/snow−on−lake.ice (Deering): This is new BRDF that is not represented in the initial knowledge base, and should be added to the knowledge base.

(6) 1989/silt.playa (Deering): New BRDF, results should be added to the knowledge base.

Following this analysis, we combined these data with the original 29 data sets to provide an expanded knowledge base for a total of 73 data sets. The global knowledge is given in Table 2. The covariances between the three parameters are listed in Table 3. As compared to the statistics of the 29−collection, the means change very little, while variances are increased because the new extremes are now in the knowledge base. In later text, we term this new data base the "73−collection" and its statistics "73−statistics." The above analysis shows that our BRDF knowledge base can be used to identify strange inversion results. However, we have only two definitely failed inversions (1 and 3 above). In order to make our partitioning of parameter space more confident, we need to have more surely failed inversions and find their locations in parameter space. During the development of MODIS BRDF inversion algorithms, daily AVHRR HRPT data of New England from a fairly clear 16−day period (14−29 August, 1995) were used to test different algorithms [24]. Without any *a priori* knowledge, inversion for some pixels will fail. Here a failed inversion is defined as one yielding albedo, either WSA or any of four BSAs, that is negative or greater than one [10]. Comparison of inversion results of the 73−collection with that of failed inversions using the AVHRR data show they distribute in distinctively different regions in parameter space (Fig. 1).

Table 2. Statistics of the 73 sets of three RossThick−LiTransit coefficients

Table 3. Covariances of the 73 sets among the three parameters

Fig. 1

4. Expansion of the Knowledge Base

The POLDER BRDF data base of P. Bicheron et al. [20] provides another opportunity to expand our BRDF knowledge base. The data sets are classified by IGBP land cover types. The 395 POLDER BRDF observations are similar to those of the 73−collection when plotted in parameter space (Fig. 2), even though these spaceborne observations have coarse spatial resolution. Their statistics are presented in Table 4 and 5. Compared to the statistics of the 73−collection, the POLDER means are well within the other's means +/− one standard deviation (σ) and are also distinctively separated from the subspace where AVHRR failed inversions scattered (Fig. 3). We checked the POLDER sets to see which data sets are "strange," that is, with at least one parameter beyond +/− 2 σ of the 73−statistics. In total, 66 sets among 395 POLDER sets are strange. Considering there are 6 parameters, this ratio is not a surprise. Among these 66 strange sets, only one has all six parameters strange: sequence number 331: C14.africa.dec.997.3241. The site name means an IGBP category 14 (crop and natural vegetation mosaic) in Africa during December, located at the POLDER coordinate 997.3241 (North Algeria). In Fig. 3a, it is the dark dot in the upper left overlapped with failed (grey) ones. The original data show very little sign of vegetation (winter) (NDVI smaller than 0.15) but a very strong hotspot (at the hotspot, reflectance in red is greater than 0.423 and NIR greater than 0.528) and a dome−shaped BRDF. So this data set may actually be new to the 73−statistics, or it may be an artifact of data noise and a rather narrow solar zenith angle (SZN) range (56.8−61.1°). We call this inversion problematic, and in need of further investigation.

Table 4. The statistics of the 395 POLDER BRDF data sets of three RossThick−LiTransit coefficients

Table 5. The covariances of the 395 sets between the three parameters

Fig. 2

Fig. 3

Five out of total six BRDF’s in IGBP land cover category 15 (ice and snow) look strange with 4 parameters out of range in the 73−statistics. This is not a surprise, since the 73−collection contains only one ice/snow sample, which thus has little weight in the global statistics. On the contrary, it is interesting to ask why the only other category 15 data set has no parameter out of the range. The set is 363: C15.n.america.june.904.1612. The original data show consistently lower reflectance (0.2−0.3) in both red and NIR than the other five ice/snow sites (all around 0.9). Drs. P. Bicheron and M. Leroy have verified that this site should actually be category 16 (barren).

The cases of the other 60 data sets can be summarized as in Table 6. In Table 6, the type of problem includes: 1) Failure, which is judged by negative albedo, or an unrealistic drop in BSA from high values to near zero; 2) Problematic, which is identified by a very large bowlshape index $(BI = f_{vol} - f_{geo})$ or a very large domeshape index (DI = $-BI$), which makes BSA change sharply with SZN, but not become negative, and 3) New feature. The reason of failure is further explained as: 1) noisy, indicating an unusually large fitting error; 2) noisy and narrow, indicting a large fitting error and narrow SZN range. Note that unusually large noise alone can make inversion a failure, and that a narrow SZN range always yields failed inversions when significant noise is present. Some sites have SZN ranges as small as 4°, but the inversion is still not a failure when the RMSE is small. Also, the boundary between failed and problematic inversion is not very clear. For example, site 331 can be classified as failure owing to noise and a narrow SZN range, which yields its strong domeshape −but it happens to have very high reflectances and thus its BSA can afford a big drop and be still far from zero. Moreover, we don’t know what the ground surface type is, so we have to treat it as problematic inversion for further investigation.

Table 6. Reasons of strange for the cases of the other 60 data sets

The typical (small, large) values in the evaluation mentioned above are listed in Table 7. The noise seems basically from daily or orbital variation in reflectance. For those sites with large RMSEs, almost all have notable daily variation. Plausibly some (dusty) days may have heavy tropospheric aerosol loading, which is not corrected according to the authors [20]. Moreover, thin subpixel clouds are not well cleared for "dusty" days (personal communication of Dr. Leroy).

Table 7. The typical (small, large) values in the evaluation of strange cases

Some inversions that yield parameters within acceptable limits may also be problematic. For example, site 352 has a DI of 0.06 in the red band, which is small, but is still large enough to bring BSA from 0.037 at nadir to 0.009 at 60 degrees. The problem may be again caused by a "dusty" day − day 6, which has samples basically with VZN values less than 45 degrees with blue and red reflectance 5−10 times higher than in other orbits, thus resulting in a dome shape.

Based on the above observations and analysis, we may conclude that the most of the POLDER BRDF data base is compatible with the 73−collection or new to it, providing more land cover features. However, eight inversions are clear failures and about 10 are problematic. Considering the POLDER data base is acquired from orbit, the spatial resolution is 6 km, and tropospheric aerosol correction has not been done yet, the consistency with field measurements is still very high. The database seemingly can be further improved if daily tropospheric aerosol correction can be done. A simple remedy may be to drop dusty days, since usually there are still enough looks from clear days.

Pooling the 73−collection and POLDER data base together, we calculated the covariance matrix of all six kernel weight parameters and its eigenvector matrix and eigenvalues (Table 8). The first three components contain 89.6% of the information.

However, examination of the corresponding eigenvectors shows the first two components are analogous to broadband albedo and spectral vegetation index (VI) respectively, though now the directional effects are corrected.

Table 8 Eigenvalues and eigen vector matrix of covariance matrix for six parameters

Considering that the datasets now contain ice, snow, urban, playa, etc., it’s no surprise that albedo and spectral VI take the most of variance. Given that a site is vegetation, the 3rd, 4th, and 5th components are more likely to contain the structural information from BRDF (23 percent of total variance). Their physical meanings may be approximately explained as NIR bowlshape index, red bowlshape index, and broadband geometric factor, respectively. Note now the bowlshape index is a weighted difference rather than simple difference we used previously. We will leave it for further study to decide whether it is more convenient to use the parameters directly, or a simple index, or a principal components−derived index in later applications. Also, with POLDER database added into our BRDF knowledge base, we now have a better potential to build up a raw observation −dependent knowledge base, but here we first illustrate how *a priori* knowledge is to be used in linear model inversion.

5. **Making Use of the Knowledge Base**

With the BRDF knowledge base, we are able to tell whether an inversion result indicates something unusual or unlikely, as demonstrated above. By mapping the *a priori* JPDF in the parameter space into the measurement space, we can further tell if a specific directional observation is unlikely or unusual when the inversion fails or is problematic. This enables us to identify possible outliers in a set of directional observations, and to smooth a noisy set if necessary (Section 5.1). Moreover, when we don't have enough directional observations or the directions are poorly located, that is, when the information from the observations is not sufficient for a full inversion of three unknowns, we can inject our *a priori* information in the inversion for a unique, most likely, posterior solution. Based on the Bayes Theorem, we developed a simple algorithm specially for regression of linear kernel models (Section 5.2).

5.1. **Using A *Priori* Knowledge to Find and Smooth Noisy Samples**

As mentioned previously, the reasons for a failed inversion may be either poor sampling, which includes a small NOL or poor directional range, and noisy sample(s), or both. For example, one of the NIR inversion failures in Fig. 1 has 8 observations. The original data are listed in Table 9. Inversion results for NIR are: f_{iso}= 0.61703, f_{vol}= −0.76090, and f_{geo} = 0.39594. All three parameters are very strange, and this yields a WSA = −0.004808, which is clearly a failure. Based on the *a priori* knowledge in parameter space, it is easy to map the parameter JPDF into data space assuming every PDF is normal and JPDF is multivariate normal. Thus, we can obtain the *a priori* PDF of each observation in each viewing direction. The comparison of true observations and corresponding prior mean estimates is given in Table 10; where Dist = (Est.ref − NIR.ref) / sqrt(Est.var). We then decide to remove the observations with the largest Dist till the WSA > 0, and three looks were removed. Inversion using the remaining 5 looks gave the results: f_{iso} = 0.5353, f_{vol} = −0.3399, and f_{geo} = 0.2921. This yielded a WSA = 0.118.

Table 9. Selected original data from the failed NIR inversion

Table 10. The comparison of true observations and corresponding prior mean estimates

Another failed inversion has the similar patterns with NOL = 7. The inversion fails because of a negative BSA at SZN=60 degrees. After removing the two least likely looks, the inversion results are: f_{iso} = 0.5397, f_{vol} = −0.3532, and f_{geo} = 0.2827. This yields a WSA = 0.132, and a decreasing BSA from SZN =0, 30 45, to 60 degrees: 0.314, 0.249, 0.168, 0.052.

The above two inversions after removing the least likely looks are no longer failures, but since the NOL is only 5 in both examples, the result is still problematic − unusually negative f_{vol}, and large f_{geo}, thus providing an extremely large DI.

Knowing some looks may be noisy, we may smooth them rather than drop them from inversion. Smoothing noisy data as well as dropping outliers are both widely accepted in practice. However, in hemispherical directions coupled with an anisotropic BRDF, it is very hard to determine how to take an average and in which directions. Thus, we tried to smooth noisy looks by forward simulations, using the average of NIR.ref and Est.ref in previous tables for noisy looks. Adding these three smoothed values to the unchanged looks, the inversion results are now: f_{iso} = 0.424, f_{vol} = −.0054, and f_{geo} = 0.172. This yields a WSA = 0.215 and BSA values of: 0.282, 0.254, 0.222, and 0.184.

Similarly, after smoothing the two noisy looks for the above Example 2, the inverted posterior estimates are: f_{iso} = 0.437, f_{vol} = −.0511, and f_{geo} = 0.173. This yields a WSA = 0.218 and BSA values of: 0.295, 0.264, 0.228, and 0.183.

Now both inversion results look reasonable, but 3 and 2 simulated looks are added into the original 8 and 7 respectively. Therefore, we must add remarks into the quality flag for the inversion, such as "smoothed," or "3/8 or 2/7 *a priori* information added for smoothing." Later in the text we will call this ratio "*a priori* information ratio."

5.2. Using Poorly Sampled Data to Obtain Posterior Estimates of Unknowns

When sampling is poor, i.e., there are too few looks, or directions are poorly located, the inversion matrix may be singular or have a large condition number or noise propagation factor [21], the inversion is underdetermined and there exists no least squares solution. In such a case, application of *a priori* knowledge is a must. Bayes inference theory is the best way to make use of *a priori* knowledge, but conventional Bayes inversion requires the integral of p_D and p_S in parameter space (7). A simpler approach uses a maximum likelihood estimate that minimizes a cost function [6, 7]

$$Cost(X) = (A*X - Y_{obs})' C_d^{-1} (A*X - Y_{obs}) + (X - X_0)' C_p^{-1} (X - X_0) \qquad (12)$$

where Y_{obs} is the vector of BRDF observations, A is the kernel matrix, X is parameter vector f_*, X_0 is the *a priori* best−guess of the vector X; C_d is noise covariance matrix of data noise and model inaccuracy, and C_p is the covariance matrix of *a priori* knowledge of X. The total *Cost(X)* thus consists of two parts: the cost of data misfitting and the cost of parameter deviation from the *a priori* best guess.

For the purpose of illustration, let's assume now we have only a single look. Then the cost of data misfitting is zero, if and only if

$$f_{iso} + f_{vol} * k_{vol}(t_i, t_v, \phi) + f_{geo} * k_{geo}(t_i, t_v, \phi) - y_{obs} = 0 \qquad (13)$$

The solution of course is a plane in the 3−parameter space. Everywhere on this plane, the misfitting cost is exactly zero, thus we can never find a unique solution for three unknowns on this plane without *a priori* knowledge. Now assume the parameters have much larger variances (i.e., large uncertainty of *a priori* knowledge) than noise in the single look. What Bayes inversion does in this extreme case is to get the posterior JPDF over the solution plane of eq (6). And, minimizing eq (5) is actually seeking the peak of this JPDF over the solution plane. For timely global MODIS BRDF products, evaluating the integral in 3−D space is not practically possible, and even minimizing eq (12) will cost too much computing time. Therefore, we suggest a mapping of the *a priori* JPDF in 3−D parameter space into data space to localize the operations, similar to the smoothing we did in the last subsection.

To inject *a priori* information into poorly sampled data, we may simulate more directional BRDF data as:

$$Y_{simu} = A_{simu} * X_0 \qquad (14)$$

Pooling these simulation data and real observations together, it can be guaranteed that it will be an overdetermined inversion and thus linear regression can be used to get a least squares solution. The problem is how much new information from observation and how much *a priori* guess is contained in such a solution, and how to sample the bihemispherical directions so that invertibility of *A'A* can be guaranteed.

We may explicitly write the cost function for this case as:

$$Cost(X)=[A*X-Y_{obs}]'[A*X-Y_{obs}]+(X-X_0)'[A_{simu}'A_{simu}]*(X-X_0) \qquad (15)$$

where *A* is the kernel matrix for real observations in eq (12). As before, the total cost function consists of two parts − the cost of data misfitting, and the cost of parameter deviation. Note that if A_{simu} is so selected that $A_{simu}'A_{simu}$ is equal to C_p^{-1} , then eq (15) will be equivalent to the eq (12). Since the covariance matrix C_p of the *a priori* PDF of three unknowns is positive−determined, we can write it in the form:

$$C_p=E*\wedge*E' \qquad (16)$$

where $\wedge$ is diagonal matrix of eigenvalues and E is corresponding matrix of column eigenvectors. Then, we can simulate three data points by:

$$A_{simu}=\wedge^{-1/2}*E' \qquad (17)$$

and

$$Y_{simu}=\wedge^{-1/2}*E'*X_0 \qquad (18)$$

Note that A_{simu} need only be calculated once and can be stored together with X_0 and C_p.

This simulation does not employ the forward model, but directly generates three data points that make their corresponding cost function numerically equivalent to the cost function of parameter deviation. So actually, they are an abstract of a large number of random simulations based on *a priori* JPDF in parameter space.

Now adding the single−look and simulated observations together, the conventional least squares approach will minimize eq (15) and get a maximum likelihood estimator of X that is somewhere between X_0 and the solution plane of eq (13). For the purpose of illustration, let's imagine that C_p is spherical and has very large variance. Then the above estimator is the perpendicular projection of X_0 onto the solution plane of eq (13) (Fig. 4a). Such an estimator is better than no unique solution and should be closer to the true value of the unknown than X_0.

But the cost of misfitting should be weighted by the inverse of the noise covariance matrix C_d^{-1}. In the case that we don't have knowledge of C_d as an alternative, we may add an artificial weight to the real observation part of the cost function

$$Cost(X) = n * [A * X - Y_{obs}]'[A * X - Y_{obs}] + (X - X_0)'[A_{simu}' A_{simu}] * (X - X_0) \qquad (19)$$

where the weight n can be easily introduced into the standard linear regression. This weight actually depends on how much we trust the new observation and how much we trust the *a priori* knowledge. The larger the n, the closer the solution X will be drawn to the solution plane of (6), and the farther it will be from X_0. But it will be less stable, since the solution plane of (6) depends on the geometry of the single look and its noise level (Fig. 4c). Similar to the case for smoothing, we will term the ratio 3/*n* the "*a priori* information ratio."

Fig. 4

In order to initially determine what weight *n* should be used, we can subset a single look measurement from one well−sampled set of the 73 sets and apply our algorithm. Their ’true values’ of parameters and WSA’s are obtained from full dataset. Then we can compare their one−look inversion results with their ’true values’ for different *a priori* information ratios. Then we subset another single look and repeat. Finally, we repeat the above procedure for another well−sampled data set and go on. The same global prior knowledge is used in all cases. The results show that, for single look inversion problem, *a priori* information ratio 3/4 can get a rather stable posterior estimates. Note this approach to the use of *a priori* information is in essence the same as that suggested by Jackson [22], but has been adapted for the use together with the standard linear inversion algorithm.

To test this approach to incorporating prior knowledge in the inversion process, we applied it to the the New England AVHRR NIR data and calculated a new white−sky albedo map (Fig. 5a). Compared with the map using the same AVHRR data without *a priori* knowledge (Fig. 5b), the improvement is obvious. In the retrieval of BRDF parameters, we inverted all observations with at least 3 looks in Fig. 5b (since 3 looks is the minimum requirement for 3−parameter BRDF model inversion) and at least 1 look in Fig. 5a. The pixels inverted with prior knowledge are shown on Fig. 5c with the red color. The green color in Fig. 5c identifies the good inversions which did not require *a priori* knowledge, and are those the same in Figs. 5A and 5b. The red pixels in Fig. 5b are the obviously failed inversions with WSA > 1.0. There are no failed inversions in Fig. 5a.

Fig. 5

6. Conclusion

Satellite remote sensing deals with a complex system of light absorption, emission, and scattering by coupled atmosphere and surface media. Any physical model describing the system with reasonable precision needs several to tens of parameters. Without *a priori* knowledge of these parameters, model inversion requires the number of independent observations to be greater than the number of unknown parameters. In practice, this requirement can hardly be satisfied, even with the high volumes of measurements to come in the EOS area. Therefore, we have developed 3−4 parameter semiempirical models to retrieve surface BRDF. Even though the number of parameters is quite reduced, much of the land surface cannot have enough clear looks to overdetermine the inversion because of limited directional sampling ability, cloud, and variable aerosol loading. In order to effectively extract information from remote sensing observations, *a priori* knowledge of model parameters has to be accumulated and used in the inversion process. By using a Bayesian approach, there is no prerequirement on how many independent observations must be made, and the knowledge gained merely depends on the information content of the new observations. In this paper, we have exploited both field measurements and satellite observations to provide prior knowledge of a general form. As data from POLDER, MISR, and MODIS are acquired in the near future, we can anticipate building a smarter *a priori* knowledge base with finer grouping of data, based on raw spectral observations of landuse classifications.

The advantage of the data−space Bayes algorithm using simulated data is its simplicity for injecting *a priori* knowledge into linear regression. However, the general approach of Bayes linear regression should be applicable to all underdetermined linear inversion problems. Of course, whenever we inject *a priori* knowledge into inversion for better invertibility of [*A'A*] of eq. (9), inevitably the uncertainty of such *a priori* knowledge will be introduced into the inversion as an additional noise source and will propagate into the results. Both the gain (better invertibility) and cost (additional noise) can be quantitatively characterized (e.g. [23]), so the automatic decision of whether *a priori* knowledge should apply should not be a big problem if we know the uncertainty of our *a priori* knowledge. However, when we decide to smooth an "outlier," which may present some rare but real pattern, we may lose the chance of making an interesting discovery. Such a decision can only be better made when we have enough knowledge of land surface BRDF's and enough confidence in that knowledge.

Acknowledgments. The work is supported partly by NASA grants NAS5−31369, NAG5−7217; and partly by China's National Key Basic Research Plan 95−Y−38, and the Major State Basic Research Development Program G2000077900. The first author would like to thank Mr. LIU Qiang for productive discussions.

References

[1] Strahler, A.H., W. Lucht, C. B. Schaaf, T. Tsang, F. Gao, X. Li, J.−P. Muller, P. Lewis and M. J. Barnsley, MODIS BRDF/Albedo product: Algorithm Theoretical Basis Document, NASA EOS−MODIS Doc., v5.0, 53pp, 1999.
(URL: http://modarch.gsfc.nasa.gov/MODIS/LAND/#albedo−BRDF)

[2] Li, X., J. Wang, B. Hu, and A. Strahler, On utilization of prior knowledge in inversion of remote sensing models, Science in China, (Series D), 41(6):580−586, 1998.

[3] Privette, J. L., Eck, T. F., and Deering, D. W., Estimating spectral albedo and nadir reflectance through inversion of simple BRDF models with AVHRR/MODIS−like data, J. Geophys. Res., 102 (D24): 29,529 − 29,542, 1997.

[4] Wanner, W., Strahler, A. H., Hu, B., Lewis, P., Muller, J. −P., Li, X., Barker Schaaf, C. L., Barnsley, M. J., Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm, J. Geophys. Res. 102 (D24): 17,143 − 17,162, 1997.

[5] Li, X., Yan, G., Liu, Y., Wang, J., Zhu, C., Uncertainty and sensitivity matrix of parameters in inversion of physical BRDF model, Journal of Remote Sensing, 1 (Suppl.): 113−122, 1997.

[6] Tarantola A., Inverse Problem Theory −− Methods for Data Fitting and Model Parameter Estimation, Elsevier, 1987, 613pp.

[7] Rodgers, C. D., Retrieval of atmospheric temperature and composition from remote measurement of thermal radiation, Reviews of Geophysics and Space Physics, 14 (4): 609 − 624, 1976.

[8] Hu, B., W. Lucht, X. Li, and A. Strahler, Validation of kernel−driven models for the BRDF of land surfaces, Remote. Sens. of Environ., 62: 201−214, 1997.

[9] Wanner, W., X. Li, and A. Strahler, On the derivation of kernels and kernel−driven models of bidirectional reflectance, J. Geophys. Res., 100 (D10): 21077−21089, 1995.

[10] Li, X., F. Gao, L. Chen, and A. Strahler, Derivation and validation of a new kernel for kernel−driven BRDF models, in Remote Sensing for Earth Science, Ocean and Sea Ice Applications, SPIE Proc. Series, Vol. 3868, pp.368−379, Florence, Italy, Sept. 1999.

[11] Gao, F., X. Li, A. H. Strahler, C. B. Schaaf, Comparison and validation of the new Li−Transit kernel, Second International Workshop on Multiangular Measurements and Models, 15−17 September 1999, Ispra, Italy.

[12] Kimes, D.S., W.W. Newcomb, R.F. Nelson, and J.B. Schutt, Directional reflectance

distributions of hardwood and pine forest canopy, IEEE Transactions on Geoscience and Remote Sensing, 24: 281−293, 1986.

[13] Ranson, K. J., Biehl, L. L., and Bauer, M. E., Variation in spectral response of soybeans with respect to illumination, view and canopy geometry, Int. J. of Remote Sensing, 6 (12): 1827 − 1842, 1985.

[14] Deering, D., P. Leone, A sphere−scanning radiometer for rapid directional measurements of sky and ground radiance, Remote Sens. of Environ.,19 (1): 1 − 24, 1986.

[15] Irons, J. R., G. S. Campbell, J. M. Norman, D. W. Graham, and W. M. Kovalick, Prediction and measurement of soil bidirectional reflectance, IEEE Transactions on Geoscience and Remote Sensing, 30 (2): 249 − 260, 1992.

[16] Leroy, M., and F. M. Breon, Surface reflectance directional signatures from airborne POLDER data, Col. on Phys. Meas. and Signatures in Rem. Sens., Val d’Isere, France, 1994, 699−706.

[17] Strugnell, N., W. Lucht, Continental−scale albedo inferred from AVHRR data, land cover class and field observations of typical BRDFs, submitted to J. Climate, 1999.

[18] Eidenshink, J. C., and Faundeed, J. L., The 1−km AVHRR global land data set: First stages in implementation, Int. J. of Remote Sensing, 17: 3443 − 3462, 1994.

[19] Tsay, S.C., M. D. King, G. T. Arnold and J. Y. Li, Airborne spectral measurements of surface anisotropy during SCAR−B, J. Geophys. Res., 103 (D24): 31,943−31,953, 1998.

[20] P. Bicheron and M. Leroy, BRDF signatures of major biomes observed from space, submitted to J. Geophys. Res., 1999.

[21] Gao, F., A. H. Strahler, W. Lucht, Z. Xia, X. Li, 1998, Retrieving albedo in small sample size, 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings: Volume V, p.2411−2413.

[22] Jackson, D., The use of *a priori* data to solve non−uniqueness in linear inversion, Geophys. J. R., astr. Soc., 57: 137−157, 1979.

[23] Li, X., F. Gao, J. Wang and A. Strahler, Parameter error propagation in BRDF derived by fitting multiple angular observations at single sun position, IGARSS’2000, accepted.

[24] Lucht, W., C. B. Schaaf and A. H. Strahler, An algorithm for the retrieval of albedo from space using semiempirical BRDF models, IEEE Trans. Geosci. Remote Sens., 38:

977−998, 2000.

Figure Captions

Figure 1 Comparison of inversion results of the 73−collection (dark) with that of failed inversions (gray) using sequential New England AVHRR images (NIR band as examples): a) Top−view (f_{vol}, f_{geo}); b) Side−view (f_{iso}, f_{vol}).

Figure 2 Comparison of parameter distributions of 73−collections (cross) and POLDER BRDF (dot) database (NIR): a) Top−view (f_{vol}, f_{geo}); b) Side−view (f_{iso}, f_{vol}).

Figure 3 Comparison of inversion results of the all 73−collection and POLDER BRDF (dark) database with that of failed inversions (gray) using sequential New England AVHRR images (NIR):

a) Top−view (f_{vol}, f_{geo}); the most top−left dark point in the failed region (gray points) is POLDER data set 331 (see text);

b) Side−view (f_{iso}, f_{vol}); the most bottom−right dark points are five ice and snow, and the set 331.

Figure 4 Illustration of "single−look" BRDF inversion in 2−D parameter space.

a) The best guess of the unknown $X_0 = (p_1, p_2)$, with uncertainty very large, and isotropic. Then the MLE is the projection of X_0 to the underdetermined solution of the single look, a straight line. Three different looks are depicted. The distance of the MLE of any look to the true unknown is shorter than or equal to the distance from X_0 to the true unknown.

b) Similar to a), but there is *a priori* knowledge of covariance of two parameters.

c) Similar to b), but now the observation contains noise comparable to the uncertainty of *a priori* knowledge.

d) Costs of misfitting to the single look and of parameter deviation from best guess. The algorithm actually minimizes the sum of the two.

Figure 5 New England NIR white−sky albedo derived from sequential AVHRR images using *a priori* knowledge (5a) and without *a priori* knowledge (5b). The pixels inverted with *a priori* knowledge are shown on 5c with red color. Red color in 5b identifies the failed inversions with WSA > 1.0.

Table 1. Statistics of the 29 sets of three RossThick−LiTransit coefficients

	Red			Near−Infrared		
	f_{iso}	f_{vol}	f_{geo}	f_{iso}	f_{vol}	f_{geo}
Mean	0.142	0.044	0.049	0.400	0.189	0.083
σ	0.126	0.045	0.052	0.108	0.158	0.080

Table 2. Statistics of the 73 sets of three RossThick−LiTransit coefficients

	Red			Near−Infrared		
	f_{iso}	f_{vol}	f_{geo}	f_{iso}	f_{vol}	f_{geo}
Mean	0.153	0.041	0.043	0.393	0.162	0.079
σ	0.144	0.043	0.054	0.126	0.120	0.087

Table 3. Covariances of the 73 sets between the three parameters

Red			Near−Infrared		
f_{iso} , f_{vol}	f_{iso} , f_{geo}	f_{vol} , f_{geo}	f_{iso} , f_{vol}	f_{iso} , f_{geo}	f_{vol} , f_{geo}
0.00012	−0.00029	0.00403	−0.00556	0.00493	−0.00713

Table 4. Statistics of the three parameters of 395 POLDER BRDF data sets

	Red			Near−Infrared		
	f_{iso}	f_{vol}	f_{geo}	f_{iso}	f_{vol}	f_{geo}
Mean	0.154	0.038	0.035	0.340	0.111	0.082
σ	0.138	0.063	0.042	0.101	0.078	0.052

Table 5. Covariances of the 395 POLDER BRDF sets between the three parameters

Red			Near−Infrared		
f_{iso} , f_{vol}	f_{iso} , f_{geo}	f_{vol} , f_{geo}	f_{iso} , f_{vol}	f_{iso} , f_{geo}	f_{vol} , f_{geo}
−0.00220	0.00273	−0.00092	−0.00267	0.00208	−0.00148

Table 6. Breakdown of cases with strange parameter values

No. of strange parameters	No. of cases	Breakdown by type
3	3	Failed: noisy (1), Problematic (1), New (1)
2	4	Failed: noisy (1), noisy and narrow (1), New (2)
1	53	Failed: noisy (3), noisy and narrow (2), Problematic (6), New (42)

Table 7. Typical ranges in the evaluation of strange cases

Fitting error (RMSE)		Solar zenith angle		Bowlshape index	
Small	Large	Small	Large	Small	Large
0.005	0.03	5	20	−0.1	0.25

Table 8. Eigenvalues and eigenvector matrix of covariance matrix for six parameters

Component	Eigenvalues	$f_{iso,red}$	$f_{vol,red}$	$f_{geo,red}$	$f_{iso,nir}$	$f_{vol,nir}$	$f_{geo.nir}$
1st	0.0304	0.616	−0.102	0.089	0.736	−0.231	0.081
2nd	0.0143	0.700	−0.073	0.142	−0.657	−0.186	−0.131
3rd	0.0078	0.276	0.275	−0.071	0.114	0.792	−0.452
4th	0.0030	0.073	0.928	−0.184	−0.014	−0.281	0.142
5th	0.0027	0.060	0.129	0.612	−0.061	0.383	0.675
6th	0.0004	−0.212	0.175	0.747	0.094	−0.245	−0.545

Table 9. Selected original data from a failed NIR inversion

	VZN	VAZ	SZN	NIR
1	61.3	124.6	28.8	0.165
2	27.6	42.0	35.2	0.287
3	12.4	42.5	34.3	0.298
4	20.2	130.6	32.9	0.216
5	33.7	129.2	32.5	0.210
6	53.0	126.5	32.0	0.195
7	17.0	43.4	37.8	0.190
8	1.3	78.3	37.1	0.181

Table 10. Comparison of true observations and corresponding prior mean estimates

	NIR.ref	EST.ref	EST.var	Dist
1	0.165	0.290	0.247	0.251
2	0.287	0.347	0.168	0.146
3	0.298	0.331	0.176	0.078
4	0.216	0.285	0.236	0.141
5	0.210	0.278	0.248	0.136
6	0.195	0.282	0.252	0.172
7	0.190	0.326	0.190	0.313
8	0.181	0.300	0.221	0.253