Anonymous referee#2

Question 1:

Section 2: The various data sources are not clear. Authors use MODIS surface albedo data for the calibration of their correction method. But it is not clear at all, how the original GLASS surface albedo is estimated. Some references are given, but no satellite or sensor is mentioned. Further clarification is needed here.

Answer:

We are sorry for the unclear description about the original GLASS surface albedo products. Further clarification will be made in our revised paper. Here, we would also like to clarify it as follows:

The GLASS preliminary albedo products are composed of four kinds of global daily 1km albedo datasets which are generated by using two different algorithms (Angular Bin algorithms, AB1 and AB2). Based on Liang's method, the AB algorithms divide solar/view-geometry hemisphere into many grids, and then establish a statistical regression model that links directional reflectance of narrow bands to shortwave albedo in each grid by using POLDER-3 PARASOL BRDF (Bi-directional Reflectance Distribution Function) dataset as training. The AB1 algorithm employs the Aqua or Terra's MODIS land surface reflectance products as inputs while the AB2 algorithm uses the MODIS top-of-atmosphere reflectance data of Aqua or Terra as inputs. The GLASS preliminary albedo products are named as GLASS02A2*x*, where *x*=1 is for AB1+MOD09GA combination, 2 is for AB1+MYD09GA, 3 is for AB2+MOD02 and 4 is for AB2+MYD02. Data layers contained in the GLASS preliminary albedo products are shortwave black-sky albedo (BSA) at local solar noon, shortwave white-sky albedo (WSA) and a quality assessment (QA) flag. The uncertainty of GLASS albedo is encoded as a four-bit word in the QA flag.

Question 2:

Eq. (1): the residual term e_k is missing in the equation and only mentioned in the text.

Answer:

Thank you for pointing out this errant. It has been corrected in our revised manuscript.

Question 3:

p9048,L5: The authors assume that the PDFs of the albedo conditioned by the retrieval and conditioned by the climatological prior are independent. I think that this assumption is not valid, as the retrieved albedo and climatological mean albedo should be highly correlated with each other. The authors need to clarify this point and assess the impact of this assumption on the retrieval method.

Answer:

Thank you for the comment about the assumption of the algorithm. The sentence in the last version is not very clear. We have changed it into "Assume the PDFs $P(\alpha_k|\alpha^*_{k+\Delta k})$ (Δk=-K,…,K) of albedo α_k are independent to each other and independent to $P(\alpha_k|\ \mu_k)$ as well.", which emphasize more on the independence of conditional probability of observation in different days. However, we'd like to argue that the assumption of independence of conditional probability is valid. It is true that "the retrieved albedo and climatological mean albedo should be highly correlated with each other", as the reviewer has pointed out. But our assumption is about the conditional PDFs of albedo, instead of albedo itself. The conditional PDF of albedo is totally different from the PDF of albedo. The conditional PDF of albedo is actually determined by the noise in data and inaccuracy in retrieval method. The noise in data can be usually assumed as white noise. The inaccuracy in retrieval method comes mostly from imperfect atmosphere correction or anisotropic correction, both of which have little correlation from one day to another. So, these conditional PDFs of albedo in different days are independent to each other. It is also a well acknowledged assumption in Bayes theory that the PDF conditioned by observation is independence to the a priori distribution $P(\alpha_k|\ \mu_k)$. So, this assumption is valid and the proposed method is sound in this aspect.

Question 4:

The method proposed by the authors is based on a statistical calibration of the relationship between the albedo of day k and day k+Dk. In addition, they use a climatological mean prior for further constraining the albedo retrieval in a Bayesian framework. In the end, they estimate an albedo value for each day together with its uncertainty. In section 3, the details of the method are outlined. In general I was missing any cross reference to variational parameter estimation methods in the paper. In general, a variational method would be the appropriate method for constraining such a multidimensional statistical problem like the one addressed in the paper. The authors need to clarify in which sense their method is comparable or different to classical variational approaches which are minimizing a cost function.

Answer:

Thank you very much for your suggestions about clarifying the similarity/difference between the traditional parameter optimization methods and our method. References on variational methods will be added in our revised manuscript. Generally, our method has many common features with the variational method. For instance, both of them optimize the parameters by coupling the background and the observational information. The estimation of background information and observational error is very critical since they determine the accuracy of the resultant parameters. Compared to the variational methods, the STF method however doesn't require any complicated optimization methods, and thus is easy to implement.

Question 5:

The general assumption of the manuscript is that there is stable correlation between albedo of day k and day k+Δk. This assumption is valid and the authors show empirically from the data that strong correlations exist. I was however expecting that the authors would be able to estimate some characteristic time scales of the surface albedo. In a variational optimization scheme, one would take the de-correlation of albedo in time into account using e.g. an exponential term like $\exp(-\tau t)$. Is τ assumed to be constant in the manuscript? I guess so, as correlations between k and k+Δk are calculated with same lags. If so, is this really a valid assumption? Temporal de-correlation of albedo should be faster in spring or autumn than during the peak of the vegetation season. The authors are asked to more critically discuss the assumptions made in their manuscript in that sense and to show whether a constant lag is a sufficient approach.

Answer:

It is a good suggestion to compare STF algorithm with the variational optimization algorithm, as the later has been widely used for years. References to the variational optimization has been added into the revised manuscript (see answer to Question 4).These two algorithms have many common features such as the temporal de-correlation of albedo. However, the STF algorithm is proposed as a statistics-based algorithm which relays on statistics rather than physical assumption. As to the question about the lag of temporal de-correlation, the STF algorithm does not assume a constant lag and the empirical expression is different from the $\exp(-\tau t)$. An empirical formula is used to fit the statistical temporal correlation in each day and each pixel. The mathematical description of the empirical formula can be found in the third expression in EQ (6) in the manuscript.

Question 6:

Spatial resolution: The authors aim for a 1km, daily surface albedo product. The prior information they are using is however based on a 5km resolution (p.9050,L1)dataset. Why do they use 5km data, when MODIS is available at higher spatial resolutions? How is the discrepancy in spatial scales considered in the retrieval approach?

Answer:

Thank you for this comment on prior information of albedo. Our former manuscript is not clear in this part. Modifications have been made in the revised manuscript. As we mentioned in the paper, there are a lot of missing data in MODIS albedo product. So the climatology would be unstable if it is calculated on a pixel level. When it is aggregated into 5km, the climatology becomes more reasonable. Besides the GLASS albedo dataset derived from MODIS reflectance data, we also developed the 5km albedo product from 1985-1999by using AVHRR reflectance data which is in 5km resolution. For these reasons, we built a 5km prior information database. Although the database is in 5km resolution, its statistics are derived from 1km resolution MCD43B3 product. So, there is no discrepancy in spatial scales.

Question 7:

The authors provide in Table 4 a statistical comparison of the different albedo products. For this comparison, they provide linear correlation measures for zonal mean values. The found correlation coefficients are very high. Zonal means area first step to compare the different products, but proper approach would show results on a pixel level basis. The authors can easily provide correlation maps, slope/intercept maps and RMSE maps based on the data they have. I guess that this will show a more heterogeneous picture, indicating areas where the different products agree and disagree. I suggest replacing the zonal mean analysis by amore spatially discretized presentation (maps).

Answer:

Thank you for this suggestion. We will add the statistical comparison of the different albedo products. But I doubt it is necessary to map the correlation between different albedo products. Since both MODIS and GLASS preliminary albedo products have a lot of missing data, to calculate the correlation, slope and RMSE maps on a pixel level basis might be unstableand result in very heterogeneous picture. The aim of Table 4 is to prove that there is no bias or other systematic inconsistency between GLASS preliminary albedo products and MCD43B3 product. Data in Table 4 are adequate to this aim.

Question 8:

p9051,L7: The authors motivate the usage of MODIS MCD43B3 product by its*great stability*. In fact my personal experience shows that the MODIS surface albedo product is actually not stable in time and contains a lot of rapid changes in the signal. The signal to noise ratio is rather low in the time series. One can easily see this when looking e.g. on the MODIS subset website for some temporal surface albedo profiles (http://daac.ornl.gov/MODIS/). What might be the impact of noisy input to the stability of the algorithm developed by the authors?

Answer:

Yes, “great stability” is not a clear statement. We have altered the related sentences. It is true that the noise of MODIS albedo may affect the prior statistics which will in turn affect the calculation of filter coefficients in our algorithm. This is one of the reasons that we have to degrade the spatial resolution of climatology database to 5km to reduce the noise effects. Anyway, the MODIS albedo product is one of the most widely used global albedo product it is retrieval with physics-based models. So, it is the best candidate to derive the climatology database.

Question 9:

Merit?: After reading the paper, the reader is left with the question of the scientific merit of the method. The manuscript is rather poor in providing a proper validation of the results and differences to existing products. The authors basically show that their method is capable of filling the gaps, but it is not clear at all from the manuscript where the majority of additional information is coming from. The authors have a framework that allow quantifying the impact of the prior and the actual retrievals on the posterior albedo estimates. I would expect the authors to elaborate more clearly, in which regions/time the information deviates from the prior or not. Further, I would expect a more thorough validation of the final results which goes far beyond the time series shown in Fig.3. To provide a proper justification for the method suggested in the paper, I would expect that the authors show if their method is superior compared to some standard techniques like e.g. Savitzky-Golay filters.

Answer:

Thank you for this comment. Yes, the validation shown in this paper is really poor. But the aim of this manuscript is to introduce the filter algorithm instead of the GLASS albedo product. I think, for a filter algorithm, filling the gaps is an important merit. Other merits for a filter algorithm may include noise reducing, increasing temporal resolution and the reasonability of the gap-fill results. All of these merits can be reflected in the several figures which the manuscript has provided. As to the accuracy of the filtered result, it is determined most by the input of filter algorithm, i.e., the GLASS preliminary products; and it is not in the position for this manuscript to discuss. Actually we intentionally reduce the validation part to minimum.

The reviewer showed concern about “in which regions/time the information deviates from the prior or not”. This is a very good question. We think the currently provided figures can also answer this question. In Figure 3, both the a priori mean (statistics mean) and the filter result have been plotted. The most obvious deviations of filter result from a priori mean occur at snow seasons. The related analysis has been added into the revised manuscript.

It is understandable that the reviewer, as well as most readers, is interested in the accuracy of the GLASS albedo product. Here we will recommend two other papers.

Qu, Y., Q. Liu, S. L. Liang, L. Z. Wang, N. F. Liu & S. H. Liu (2013) Direct-estimation algorithm for mapping daily land-surface broadband albedo from MODIS data. Submitted to IEEE Transactions on Geoscience and Remote Sensing

Liu Q., L.Z. Wang, Y. Qu, N.F. Liu, S.H. Liu, H.R. Tang & S.L. Liang (2012), Preliminary Evaluation of the Long-term GLASS Albedo Product, submitted to International Journal of Digital Earth

The first paper presents validation of the GLASS albedo preliminary products. It will be available to public very soon. The second paper gives an extensive discussion about the overall quality of the GLASS albedo final product, as well as its validation. As it may take time before this paper becomes published, we would like to quote some results of this paper as follows:

In this paper,53 homogeneous FLUXNET sites are selected to validate the accuracy of GLASS albedo product..To minimize the atmosphere and cloud effect, the observations with downward irradiance less than 70% of clear sky irradiance were screened out. For a thorough comparison, the land surface state is categorised into three types: vegetation, snow/ice and bare ground. The statistics of GLASS albedo are summarized in Table1. It can be seen that RMSD (Root Mean Square Difference) is the smallest in vegetation observations, and the largest in snow/ice covered observations. The RMSD for clear-sky ground measurements and "good" quality GLASS data is -0.0005, and the R2 is 0.89. That means the quality flag is a pertinent indicator to the accuracy of GLASS product.

Table 1 Statistics of comparing FLUXNET ground measurements to GLASS albedo product.

Sample criterion	Number Obs.	Bias	RMSD	R2
clear-sky ground measurements & all GLASS data	28881	-2.95e-06	0.0587	0.8033
clear-sky ground measurements & "good" quality GLASS data	16960	-0.0005	0.0455	0.8930
vegetation	14833	-0.0017	0.0303	0.6175
bareground	938	0.0121	0.0532	0.3092
Snow/ice	1189	0.0048	0.1258	0.7231

To check the consistency of GLASS albedo product with MCD43B3 product, black-sky albedo time series from 2003 to 2004 were extracted from 402 BELMANIP sites all around world. The RMSD is 0.031 and coefficient of determination (R2) is 0.922.Statistics in each of the surface states are given in Table 2. We can see that the RMSD for vegetation and bare ground surface is less than 0.02 and the R2 is larger than 0.9, which indicates a good consistency between GLASS and MCD43B3 products in these two surface states. When the surface is probably covered by snow/ice, the discrepancy is larger, with RMSD of 0.0797 and Bias of 0.038. GLASS albedo product is kind of higher than MCD43B3 albedo product for snow/ice covered states. Comparing with Table 1, the bias between GLASS and MCD43B3 products is larger than that between ground truth data and GLASS product. This can be possibly attributed to an under-estimation of albedo in MCD43B3 product when the surface is covered by snow/ice: the algorithm for NASA's MODIS albedo product has dropped the partially snow-covered observations to maintain the accuracy of snow-free albedo at the expense of under-estimation of partial-snow albedo.

Table 2 Statistics of comparing GLASS albedo product to MCD43B3 product.

Sample criterion	Number Obs.	Bias	RMSD	R2
All GLASS data & valid MCD43B3 data	29719	0.0070	0.0310	0.9222
"Good" quality GLASS data & valid MCD43B3 data	18605	0.0073	0.0226	0.9668
Vegetation	14229	0.0068	0.0121	0.9260
Bareground	3319	-0.0004	0.0147	0.9727
Snow/ice	1057	0.0375	0.0796	0.8138

Minor revision:

Thanks very much for the comments about the minor revision. We have thoroughly revised the manuscript, including the following points which have been mentioned by the anonymous referee.