Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-359-AC2, 2018
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Measurement Techniques Discussions
Open Access

Interactive comment on “Bayesian Dark Target Algorithm for MODIS AOD retrieval over land” *by* Antti Lipponen et al.

Antti Lipponen et al.

antti.lipponen@fmi.fi

Received and published: 23 January 2018

We thank Dr. Povey for encouraging and constructive comments. Below we have included the referee’s comments in boldface font and our replies below each of the comments.

P4L21 **The retrieval of $log\ x$ rather than x for positive variables is well documented. You retrieve $log(1+x)$, which I have not encountered before. Could you further discuss this choice, maybe providing references to demonstrate it's use elsewhere? I am concerned that it permits $-1 \leq x$. Did you specifically wish to retain the small but negative τ from the original Dark Target algorithm or are the 'constraints that exclude non-physical solutions' (P5L23) hard limits that prevent this behaviour? If the latter, why not use the more common $log\ x$ formulation? Have you considered how hard limits distort Gaussian uncertainty estimates near those limits?**

In our optimization solution, we constrain the values of $\log(x+1)$ to non-negative values that guarantees us with non-negative AOD retrievals. As we retrieve $\log(\tau+1)$, the prior model is written for the same quantity as well. If we retrieved $\log(\tau)$ instead, it would mean that for really small AOD values we would have really small $\log(\tau)$ values (approaching to $-\infty$ when AOD approaches 0). This in turn would require the use of really high variances (or standard deviations) for the $\log(\tau)$ to allow for higher AOD values to be retrieved. To avoid this problem of small numbers and high variances in the prior models, we decided to retrieve $\log(1+\tau)$ instead of $\log(\tau)$. To make this clear, we added a sentence: "In practice, these constraints are implemented in the optimization algorithm." to the manuscript.

We also agree that the hard limits may distort the Gaussian uncertainty estimates but have not studied it further.

§2 **It's unclear from the text precisely how many pixels are processed at once. Is it an entire granule? Processing 50,000 pixels at once would be an impressive computational task!**

I also recall that you only process pixels for which a DT retrieval was produced (implicitly adopting their cloud flag), but I don't find that mentioned in this text

We retrieve the same dark land, cloud-free pixels as the Dark Target algorithm, not all pixels in a granule. We have clarified this throughout the manuscript. The maximum amount of pixels for a MODIS aerosol granule at 10 km resolution is 27540 (204 by 135) and in most cases only a fraction of these pixels are cloud-free, dark surface pixels. Therefore, in practice the average amount of pixels retrieved simultaneously is about 4000 pixels. In some cases, if the amount of pixels to be retrieved has been larger than ten thousand pixels we have divided the granule into two sub-granules and carried out the retrievals in these sub-granules separately to make the computations faster. We have also made it more clear that our algorithm uses the same preprocessing as the DT by changing the sentence "BDT is a retrieval algorithm based on DT." in Section 2 to "BAR is a retrieval algorithm that uses the same aerosol models and preprocessing of the data, such as cloud-screening, as the DT. Because the same preprocessing is used, the BAR algorithm retrieves the same pixels as the operational DT algorithm."

P6L17 **Though the 50 km correlation length is widely used, you should cite something. doi:10.1175/1520-0469(2003)060<0119:MVOTA>2.0.CO;2 is quite common.**

Thank you for the reference. We have added a citation to this paper.

CC BY

§3.1.3 This method contains a few surprising features. Why use blue sky albedo? Why seasonal averages? Why average the 3 closest values rather than do a bilinear or triangular interpolation? Was the technique overly sensitive to these choices (i.e. were these chosen at random and worked or did it take several attempts to find a stable solution)?

The blue sky albedo (with weight of 0.5) was selected based on initial tests in which it resulted in the best retrieval results. In the tests, we carried out retrievals with white-sky, black-sky, and blue-sky (weight 0.5) albedos for a small amount of granules and compared the results. The differeces between the surface albedos were small but the blue-sky was the best performing one so we selected it to be used in the algorithm. We added the following sentences to the surface reflectance prior model section of the manuscript: "This selection to use the blue sky albedo was done based on a test in which we carried out retrievals with white-sky, black-sky, and blue-sky albedo based prior models. The differences between the different surface albedo types were small but the blue-sky albedo resulted in the best results when compared with the collocated AERONET AOD values."

The nearest-neighbor interpolation was selected as it is a computationally cheap and easy to implement option. As there is a surface reflection prior model with a relatively good spatial resolution, we believe that the use of some other interpolation method would not result in significantly different results.

P8L7 What motivated the addition of coarse mode aerosol to the continental mode? Are the Dark Target team considering removing this step from their own processing?

The coarse mode was added to the continental mode only for convenience as our practical implementation of the algorithm requires two different aerosol models and by

CC BY

doing this choice we can use the same code to process all pixels in all granules. We note that it would be a straightforward task to include the continental mode into the algorithm similarly as in the Dark Target but we think it would not significantly affect the results.

P9L2-4 **I don't understand what you mean by 'marginalize the posterior model'. Marginalize means 'to treat as insignificant' and you use posterior model to describe the cost function, (2). I would expect one to 'minimize the posterior model', but I fail to see why that is relevant to the approximation error approach.**

In the statistical approach, by marginalization we mean that the unknown uncertainty and noise related parameters are integrated out of the posterior probability density. In our case, the posterior probability distribution is a joint conditional probability distribution of AOD, FMF, surface reflectance, and the observation noise and approximation errors given the reflectances observed by the MODIS instrument. Formally the marginalization is:

$$\pi(\tau, \eta, \rho_{\text{surf}} | \rho_{\text{TOA}}) = \int_e \pi(\tau, \eta, \rho_{\text{surf}}, e | \rho_{\text{TOA}}) \mathrm{d}e \tag{1}$$

where τ, η, ρ_{surf}, ρ_{TOA}, and e denote the AOD, FMF, surface reflectances, top-of-atmosphere reflectances observed by the MODIS instrument, and the observation noise and approximation errors, respectively. In the approximation error approach, the integration is carried out approximatively. We have added the following sentence to the manuscript to clarify this: "This means that we integrate the approximation error related variables out of the full posterior probability distribution. This is a typical approach in statistics to treat unknown nuisance parameters." For more information on marginalization of posterior probability distribution and techinques how to carry it out in practice see, for example, doi:10.1615/Int.J.UncertaintyQuantification.v1.i1.10 or

CC BY

doi:10.1007/b138659

P9L31 **'Physical' may be a better word than 'true' here as there arguably is a 'true' FMF as defined by the Dark Target algorithm, but the point is that that value doesn't always mean something in reality.**

We agree, 'true' was changed to 'physical'.

Fig. 1 **Could the urban sites (discussed in §5.4) be displayed in a different colour?**

The figure was edited, the urban sites are now shown in a different color.

Figs. 2&3 **For Angstrom exponent, could you use a colour bar that has grey at the centre so we can distinguish missing data from a value of 1?**

The colormap for the Ångström exponent was changed to distinguish missing data from a value of 1.

App. B **I broadly like this idea, and do something similar myself (though not yet in a published paper), but I'm curious about assuming the MODIS BRDF is accurate. It's a good retrieval but not without substantial uncertainty (of many forms - representational, approximation, etc.). Considering the dominance of the surface in aerosol error budgets, how accurate do you think these estimates of the approximation error are?**

We have not run a more detailed analysis on this. We also agree that the MODIS surface reflectance products may have substantial uncertainties. We assume the MODIS

Printer-friendly version

Discussion paper

surface reflectance as accurate in the approximation error computations but at the same time we use a large amount of data for constructing the approximation error model. As a large amount of data and averaging is used in the computations, it is only necessary to have unbiased (or small bias) surface reflectances in the computations to get a good approximation error model. Therefore, not all single values of the surface reflectances have to be close to correct one. The results show that the retrievals are significantly improved when we introduce the approximation error model. This shows that the models (including surface reflectance) used in the computations are of good quality to be used for this kind of modeling.

- **Several references list a URL twice. Perhaps replace the BibTeX field url with doi?**

URL problem fixed. The BibTeX field url was replaced with doi.

I also include some proofreading recommendations.

We thank for the recommendations. Most of the recommendations were included in the revised manuscript.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-359, 2017.

Printer-friendly version

Discussion paper