Our detailed responses follow with reviewer's comments in bold reminded.

The manuscript of Bernard et al. describes the estimation of aerosol optical thickness (AOT) from the geostationary SEVIRI sensor over land (but also water) surfaces with high temporal resolution suitable for an operational implementation. The authors follow a multi-temporal approach to estimate surface reflectance which is then used together with the measured Top-of-Atmosphere reflectance (corrected for the effects of gases) and a radiative transfer code to invert AOT. The authors compare their SEVIRI derived AOT from three month in spring 2006 with corresponding values from AERONET sites in Europe and with the MODIS level 3 aerosol product and find a generally good agreement. High temporal resolution and spatially homogeneous aerosol information are of interest for different applications, especially when the atmospheric aerosol load exhibits a high diurnal variability like the authors demonstrate. The near real time and operational high resolution AOT product that the authors aim at could therefore be of great interest for different applications and studies. The topic of this manuscript fits well into the scope of AMT and the paper is well organized but needs English proof read. I have no general concerns with this manuscript but a few comments which should be addressed before the paper can be published in AMT.

We are pleased that the topic and motivation for this work was well received by reviewer 1 and are grateful to him for his useful and detailed comments which especially helped us to better discuss uncertainties of the AOT product within the scope of the present paper.

General comments:

The aim of this manuscript should be stated more clearly. Is it the development of the methodology or is it supposed to be a "reference paper" for an operational aerosol product which will be publicly available?

This point was raised by both reviewer and we agree that the title may did reflect correctly the scope of our paper initially. Reviewers are right that the baseline methodology is not new but at the same time we are also proposing a methodology to derive more information on the aerosol model by using the multispectral information available. We have tried to improve the discussion regarding what our current implementation brings to the previously published work.
Regardless, the main purpose of this paper was to document the quality of an operational product that will be made available publicly through the ICARE Data and Services Center (http://www.icare.univ-lille1.fr). Although the methodology is currently being consolidated, the improved product may not be available in a short term so we felt necessary to document the quality of the current version which we think already has some value for specific users over Europe.

With regard to the first, the feasibility of multi-temporal approaches for the AOT retrieval over land from geostationary sensors (e.g. GOES, SEVIRI) has been already demonstrated in previous papers. However, the authors make also use of the 0.8 and 1.6 micron channels in order to determine the best-fit aerosol model for the AOT inversion which is important and (for over land retrievals) novel to my knowledge.

We agree. This is now better identified in the text.

However, this aerosol model selection is not elaborated sufficiently. It should for instance be demonstrated that the aerosol selection leads to superior AOT retrievals than if a (fixed) aerosol model from a predefined or AERONET-derived climatology is used. Further, certain estimated model parameters (e.g. single scattering albedo or Angstrom exponent) could be compared to the ones from AERONET. While water surfaces are usually very dark in the 0.8 and 1.6 micron wavelength region, vegetation is much brighter which might considerably decrease the sensitivity of the sensor to detect an aerosol signal in the NIR. Does that have an impact on your best-model selection?

Regarding the aerosol model derived by the algorithm, no comparison has been made with AERONET data in this study because we initially focused on the AOT product validation. The only preliminary analysis we have done concerns the evaluation of the temporal consistency of the chosen best aerosol model as illustrated in Fig. 7. We would expect that the best model for any particular site should not change or change smoothly overt the course of the day which obviously from Fig. 7 does not seem to be the case.

The point raised by reviewer 1 is relevant and the lack of sensitivity in the NIR over bright surface might be an issue. However there are other potential problems arising from inconsistencies in the spectral dependence of the surface reflectance used as reference.

At this stage of our development, we do not have enough confidence in the spectral coherence of SEVIRI calibration to grant real physical value to the aerosol model retrieved and suspect that diurnal variability of the aerosol model retrieved is at least partly due to inconsistent spectral dependence of the reference surface reflectance. This could be caused calibration inconsistencies, but also to remaining cloud contamination or artifact induced by our temporal fitting of the reference BRDF (this point is currently being improved as discussed later).

We have identified path for improvements of the aerosol model selection which will be implemented in the next version of the algorithm. But for now, we preferred to deliver the best AOT retrieved along with AOT retrieved for all other models (as illustrated by Fig. 7C so that users can get an estimate of the uncertainties associated with model assumption. Also, it allows users to reconstruct a daily evolution of AOT with any model he thinks is most relevant (climatologies, most represented model for a day, average, ….).

With regard to the introduction of an operational aerosol product, I wonder why only data from three months and only over Europe are shown. Second paragraph of the abstract and p.3161, l27 indicate that AOT is derived over the entire SEVIRI disk. There are also other interesting regions with regard to aerosols and I think that the paper can benefit a lot if at least one figure is included illustrating the AOT distribution over the entire SEVIRI disk (e.g. similar to Fig. 8). This figure could also be used to extend the discussion on limitations and potentials of this product.

In fact, the currently available product is derived over the whole disk but suffers from limitations that the authors are trying to correct in an improved methodology. For instance, areas with high aerosols loads for long period currently present a low bias in the retrieved AOT which we are trying to correct through a better estimation of residual aerosol signal in our BDRF estimation. Also, over bright surfaces (desert), a more elaborated decomposition of measured TOA signal should be used instead of what we currently apply. Again we felt necessary to document the product so that it could be made available publicly rapidly and used for what it is worth. Especially over Europe, we believe that the current version provides already meaningful information that can be used for temporal

analysis and contribute to air quality studies for instance.
So we prefer to keep the analysis presented in this paper focused on Europe, knowing that the next version of the algorithm will significantly improve quality elsewhere.

The error discussion is not sufficient and in general quite speculative. Sure, it is known from numerous previous studies that surface reflectance estimates and cloud contamination (especially with the relatively coarse resolution of SEVIRI) are major error sources but no attempts are made by the authors to quantitatively describe and estimate the contribution of these error sources to the total error. With regard to future algorithm improvements, to do so is important. For example, AERONET data can be used for a better error assessment or spatial consistency tests might be applied to investigate the cloud influence. There are also additional sources of errors not or only marginally mentioned in the manuscript which should at least be discussed, e.g. the estimation of the background AOT, estimated aerosol properties or, like Dr. Ham pointed out, calibration issues.

The reviewer comment is well taken and we certainly agree that this is an important step toward building an improved algorithm and we are currently going through this exercise. However, we would like to remind that the current product will be improved in the mid term and we thought that a global estimate of the accuracy was sufficient at this stage.

To answer more specifically though, it is difficult to quantify these various sources of uncertainties separately.
First, evaluating the impact of the residual cloud contamination per say would require a "truth" for cloud masking such as collocated CALIPSO data. This is not easily feasible unfortunately and it is unlikely that we could get statistically significant results over the 3 months considered. We agree with reviewer that could also rely on a more severe selection of AERONET data using temporal analysis to filter out thin clouds. However we felt that this could somehow bias our validation dataset towards "ideal" situations and we wanted to provide a quality assessment of the product as a whole instead.
Secondly, evaluating the impact of the temporal fit actually requires that we assume a more realistic BRDF fitting model. This work is under progress as we are now evaluating the use of kernel-based BRDF models to do our temporal fitting which enables more physical representation of surface reflectance diurnal variation, especially in the hot-spot geometries. However, it was not possible to quantify separately this impact of temporal fitting with the product validated and presented in this study. Again we would like to insist that the purpose of the present analysis is to provide a first evaluation of the overall product quality so that users can rapidly use it for what it is worth. We agree that a more systematic estimation of uncertainty sources would be desirable but this somehow is out of scope of the present analysis.

In addition, I wonder a little that you rely on a "rather simple cloud mask" (P.3156) if cloud contamination is a major error source. So why not using just the best cloud mask available?

Please refer to the answer provided in the specific comments hereafter.

Specific comments:
Please write out all abbreviations (including sensor names) the first time they are mentioned and use the abbreviations afterwards. Please also check the notation of (mathematical) products throughout the manuscript, probably better use the "x" sign.
This has been corrected.

P. 3151, l8: the King et al. (1999) reference is a little old considering all the progress achieved in the field of aerosol remote sensing in the past decade. You could probably also mention a more recent review paper or book besides King et al. (1999).

We are now also referring to publications by Mishchenko et al, 2007 and Kaufman et al, 2002.

Mishchenko, M.~I. and Geogdzhayev, I.~V. and Cairns, B. and Carlson, B.~E. and Chowdhary, J. and Lacis, A.~A. and Liu, L. and Rossow, W.~B. and Travis, L.~D.,"Past, present, and future of global aerosol climatologies derived from satellite observations: A perspective", J. of Quant. Spect. and Rad. Transf., 2007, 106, pp. 325-347, doi = 10.1016/j.jqsrt.2007.01.007

Kaufman, Y.~J. and Tanré, D. and Boucher, O., "A satellite view of aerosols in the climate system", Nature, 2002, 419, pp. 215-223,

P. 3151, l12: The daily sampling increases for geostationary satellites, doesn't it? Or do you want to say that geostationary satellites are only capable to observe a specific part of the entire planet? In addition, the spatial resolution is not necessarily lower than polar orbiting satellites (e.g. OMI).

We meant to say that the geostationary satellites are only capable to observe a specific part of the entire planet.

The previous text : *« From geostationary satellites the surface sampled daily is for obvious reasons reduced and the observations spatial resolution is also lower. However the high temporal resolution ...»*

Now reads : *« From geostationary satellites the surface sampled daily is for obvious reasons limited to the geostationary orbit field of view and the observations spatial resolution tends to be usually lower compared to instruments of the same generation on polar orbits, to a few exceptions such as OMI. However the high temporal resolution ...»*

P.3152, l8 – P. 3153, l1: Think about moving these three paragraphs to section 3 or shorten it otherwise.

We have shortened the three paragraphs which now reads :

"Our algorithm is based on the assumption that the Top Of the Atmosphere (TOA) reflectance in the VIS06 channel increases with the aerosol load {Fraser1985, Kaufman1997}. The TOA reflectance is corrected from the gas and molecule contributions. The thermal IR channels allow the determination of a good cloud mask. Finally, the fixed viewing angle combined with the variable solar angles allow some angular sampling of the surface BRDF and/or aerosol phase function.

The algorithm developed aims to retrieve the aerosol optical thickness in the VIS06 channel in two steps using a method similar to the one developed for GOES-8 by {Knapp2005}. Over land, from a set of 14 days images, a map of estimated surface reflectance is built assuming that the darkest pixel for the period correspond to a clean-sky observation (in fact the clearest). The second step is the retrieval of the AOT for each images using these surface reflectance maps and using a set of 5 aerosol models."

P.3153, l23-l24: Do not most of the gases in the atmosphere consist of molecules? I guess you are referring here to the radiative effects of gases and aerosols, so it would be better and clearer to use here terms like gaseous absorption, Rayleigh scattering, aerosol scattering and aerosol

absorption.

We have rewritten this paragraph to clarify the discussion.
The paragraph :
« In clear-sky conditions, the atmosphere is composed by 3 main elements: gas, molecules and aerosols, which all contribute to the TOA radiance […] Therefore, the atmospheric system we assume is divided in three separated layers: a molecular layer, located above an aerosol layer and beneath a gas absorbing layer.»

Has been replaced with :
« In clear-sky conditions, the TOA radiance derives from the contribution of the absorption and scattering by molecules and the aerosol scattering. Under some assumptions, we can dissociate these three contributions and calculate them separately.
First, for molecules, the absorption is independent from the scattering phenomenon. At the wavelength considered for our retrievals, 0.6µm, the main gas which participate to the absorption is the Ozone.
Regarding the scattering contribution, the molecule and aerosol are mixed in the atmosphere and a coupling effect appears. The radiance produced by such a coupling depends on many parameters: geometry, wavelength, surface pressure, aerosol type and aerosol optical thickness {Santer1999}. Beyond 0.6µm the coupling effect can be neglected {Ramon2001}. In order to simplify the modeling, molecules and aerosols are separated. Thus, the Rayleigh scattering contribution is calculated separately. Therefore, the atmospheric system we assume is divided in three separated layers: a molecular layer responsible for Rayleigh scattering, located above an aerosol layer and beneath a purely absorbing gaseous layer.»

P.3156, l2: In principle, meteorological parameters from model forecast could also be used in a near real time operational aerosol retrieval from SEVIRI. As it is argued in this study that cloud contamination is a major error source in the AOT retrieval why relying on a "rather simple cloud mask" instead of just using the best one available?

This point has been raised by both reviewers and we have improved the manuscript to clarify the rationale for using our own cloud masking scheme.

First, the term "rather simple" refers to the fact that our cloud mask does not have complex dependency to ancillary data and does not require dynamic threshold to be computed online using radiative transfer code as some more "evolved" cloud mask scheme sometimes do. Yet, our cloud mask scheme rely on a variety of spectral threshold which combine quite effectively to produce a quantitative cloud / clear probability index that can easily used to select more or less confident clear pixels. It should be noted that our cloud mask, although simple, has proved to perform quite well and has been used with success for various studies (see for instance Roebeling, R. A., H. M. Deneke, A. J. Feijt, 2008: Validation of Cloud Liquid Water Path Retrievals from SEVIRI Using One Year of CloudNET Observations. J. Appl. Meteor. Climatol., 47, 206–222.)

We used our own cloud mask primarily because we required a cloud masking scheme that we could easily modify to adjust detection level in order to keep all clear sky pixels at the expense of some cloud contamination. Using the more evolved Nowcasting SAF cloud mask for example does not allow for such fine tuning. Again, one need to bear in mind that "cloud masking" can be implemented quite differently depending on target application. Here, because we are looking at aerosols, including extreme events, we choose to allow for some potential cloud contamination which is later removed through temporal analysis, spatial homogeneity consideration and identified

in a quality assurance mask.
Another secondary advantage is that our cloud mask scheme does not rely on any ancillary data which allow for a completely standalone implementation of our aerosol product relying solely on SEVIRI Level 1 data.

We have now clarify these various aspects in the revised manuscript.

P. 3156, l27: Please clarify to what the subscripts 0 are referring to.

T_0 is the Ozone transmittance for a quantity of 344 Dobson Units and an air mass of 1. T_0 is the reference transmittance from which we estimate the transmittance for all the air masses. We are now using T_{03} to remove ambiguity.

P.3158, l23: Better "depending on the latitude"

We agree. Old text was : "..., depending on the Earth's locations,..." and now reads "..., depending on the latitude,..."

P.3159, l21: What is a pseudo-BRDF?
In fact we are not retrieving the full BRDF because for any one pixel viewed by SEVIRI, the viewing geometries are constant and only the sun geometries are changing. So we are only retrieving a sample of the BRDF. A better term instead of "Pseudo-BRDF" could be "A partial BRDF" or better "a BRDF subset". This is the variation of the surface reflectance with the solar angles only.

The correction is also done on page 3163, l.17 and at the legend of the figure 1.

P.3161, l23, 24: Do you mean 2x10^(-5) etc.?

Yes we meant : 2x10^(-3) or 0.001, 1x10^(-4) or 0.0001 and 1x10^(-5) or 0.00001
We added a x sign to remove ambiguity where needed.

P.3162, l21 and P.3163, l20: What happens to isolated pixels? Are they masked out? Please clarify.

Isolated pixels are identified within the quality flag. Text now reads :
p3162 "Isolated pixels have a σ_τ=0 and are identified as such in the Quality Assurance flag so that they can later be rejected."
p3163 "We choose to keep non isolated pixel with a local ... lower than 0.1

P.3163, l5: Is there a reason why you don't interpolate the AERONET AOT (available at different wavelengths) to the SEVIRI wavelength at 630 nm? Doesn't this potentially introduce a small bias?

The comment concerning the wavelength we use to compare AERONET and SEVIRI is interesting.

The error done with this assumption is low compared to the other sources of error (cloud mask, surface reflectance estimation, aerosol background). The bias potentially introduced is negligible compared to other sources like the background aerosol.

P.3164, l7: 4x10^(-2) or 4.1^(-2)?

As above, we indeed meant 4 x 10^(-2).

P.3164, l7 and elsewhere: As far as I understand you don't allow your algorithm to derive negative AOT, do you? While this is physically absolutely correct it would make sense with regard to the statistical analysis to also include slightly negative AOT values (e.g. similar to the MODIS operational algorithm).

Yes, indeed, the algorithm does not allow the retrieval of negative aerosol optical thicknesses. We can understand the logic of the MODIS operational algorithm which uses some statistics of 500 m resolution pixels reflectance over a 10 km x 10 km area and does rely on some assumptions for deriving the surface reflectance. However, our algorithm is based on the primary assumption that we can get an accurate estimate of the surface BRDF on an individual pixel basis so that negative optical thicknesses in our product are expected to be largely caused by cloud shadows that we wish to remove.

P.3164, l17,18: The statement "... comparable to results obtained for the MODIS algorithm ..." is a little strong and I doubt that a meaningful comparison is possible on the basis of comparing three month of SEVIRI AOT over Europe with a decade of worldwide MODIS aerosol retrievals. For example, the study of Levy et al. (2007) reports MODIS/AERONET correlation of 0.9 and a linear regression equation of y=1.01x + 0.03 globally, compared to the correlation of 0.63 and slope of 0.81 in Fig. 4. With regard to the study region of this manuscript, the study of Riffler et al., Atmos. Meas. Tech., 3, 1255-1270, 2010 showed that the MODIS-AERONET correlation is for a majority of (central) European sites better than 0.9 while the slopes vary somewhat between 0.9 and 1.1. Therefore I don't think that "comparable" is the right word here. You might refer to these or additional publications to give a better impression of the SEVIRI performance.

Yes, reviewer 1 is perfectly right on this point. MODIS shows better results compared to AERONET in terms of correlation. But the reviewer is also actually raising a very good point regarding how to compare performances of different retrievals. Going again through the suggested paper, we indeed think it is difficult to draw conclusions on the basis of statistical coefficient that are built from validation dataset selected using different criteria. For instance, we used an averaged AERONET value of +/- 10 min around each SEVIRI observation when Levy et al, (2007 and 2010) (following Ichoku et al, 2002) considered 30 min averages. Also, we used single pixel SEVIRI values (pixel centered on AERONET site) when again previous authors have used area averaged MODIS values over 5x5 pixels. The fact that we get to sample more than 5.000 match-ups during 3 months of data over Europe when Levy et al (2010) for instance ended up with about 85.000 match-ups globally for both Terra and Aqua between start of the mission until September 2008 can also raise question on the differences of representativeness. Also this could result, at least partly, from the somehow stricter criteria used by MODIS for the selection of AERONET validation data (for instance cloud masking of AERONET data over 30 min instead of 10 min for our study).
To conclude, the spatio-temporal matching criteria used for validation in the paper suggested by reviewer 1, are indeed not really comparable to ours, so that results in turns may not be easily compared. To prevent erroneous conclusions, we therefore removed sentences that could be misleading such as the one identified by reviewer 1. Instead we now point out to the fact that validation results obtained for different AOT products should be compared with caution because of

potentially different selection criteria.

P.3164, l19-27: Please quantify the contribution of the different error sources, c.f. General comments.
See discussion in response to General comments.

Section 5: What is the source of this enhanced aerosol loading you describe here?
At that time of July 2006, there were both fires active in Spain and Portugal as well as some dust events coming from North Africa. The event we describe here seems predominantly associated with transport of biomass burning aerosols arising from Iberia peninsula. This is now indicated in the text.

P.3169, l20 and elsewhere: AERONET in capital letters.
This has been corrected.

P.3175, l6: I think this statement is a little misleading. Certainly do the SEVIRI 0.1 resolution AOT maps allow a more detailed picture than the MODIS level 3 product at a resolution of 1x1. But if you want to, you can also construct MODIS 0.1 maps from the level 2 data available at 10 km x 10 km which also are suitable for regional studies. So how do you justify to use MODIS level 3 data instead of level 2 data in this study?

We agree that the statement is misleading and removed it. However, we wanted to make the point that users will usually go for “ease of use” and clearly the level 3 MODIS data are much more straightforward to use than having to rebuild a daily synthesis from level 2, applying all quality assurance filters that are used to created MODIS level 3 from level 2.

P.3175, l27: This can relatively easily be confirmed or rejected by using AERONET AOT values.
Indeed, the assumed aerosol background could be confirm or rejected by using AERONET information but only on particular sites. The goal of using LEO satellites is to benefit from more global information.

Figures:
Fig.4: Please add a colorbar with the corresponding density values.
Figure 4 has been corrected: the color bar now appears.