Interactive comment on “Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR” by C. A. Poulsen et al.

Anonymous Referee #2

Received and published: 1 July 2011

General comments

The authors present an algorithm for the retrieval of cloud properties from ATSR data based on optimal estimation that is used for the compilation of the GRAPE data set. The scheme is described and simulations are performed to assess its accuracy in case of single or multilayer clouds. Finally, an application to real data is shown and discussed.

The paper is well written and should be published in AMT. Furthermore, its companion paper (Sayer et al., 2011) has already been published in ACP. Nevertheless, it must first go through major revisions (see below).

Specific comments
The main deficiencies of the paper consist in

1. the illustration of the retrieval algorithm
2. the description of the RTMs
3. missing literature references about further optimal estimation retrievals
4. missing quantitative conclusions about the accuracy and limitations of the retrieval as extracted from the error simulations.

In the following, I will go through these points.

1. Retrieval Algorithm

Section 3 should contain all the issues regarding the algorithm proposed in this paper.

I start with Eq. (1) describing the cost function $J$:

All terms used in Eq. (1) should be defined, i.e. the subscript $m$ should be explained and also the covariance matrix should be explicitly defined. This will then also explain why $J$ represents a cost function. Furthermore, in Eq. (1) all $x$ should be bold and also all $y$.

Page 2395, line 7–9: Why do you assume / Why can you assume that errors in the measurements (and forward model) are normally distributed with zero mean?

Page 2395, line 17–24: Which values does the cost function $J$ assume? Which range does mean that you have an accurate retrieval? Sayer et al. (2011, Table 2) provides such an information.

Page 2395, line 20–23: Sayer et al. (2011) give more details with respect to $\chi^2$. I think that they should be reproduced here as well (in this or in another form) since they concern the algorithm directly.
Page 2395, line 17–18: How is such a linearisation performed in practice?

Page 2395, Eq. (2): What is the meaning of this equation?

After the main principles have been explained, the authors should give details about covariances, a priori estimates and so on. This means that Section 6 and 7 should be integrated into Section 3. Sections 4 and 5, the description of the cloud model and of the RTMs, are namely not necessary to understand these issues and, most importantly, I think that Sections 3, 6 and 7 belong together. For this reason, I list here my questions regarding Sections 6 and 7:

Page 2402, line 20–21: Is the use of all channels compulsory? Can you apply the retrieval at night using only thermal channels?

Page 2403, line 8: You probably mean the sum of three terms, right? Or is there an additional fourth term that is not mentioned?

Page 2403, line 10–13: Which assumption do you make in order to assume that this covariance matrix is diagonal?

Page 2403, line 14: Please mention the main inadequacies of the cloud model that could play a role here (3D effects, strong aerosol load...).

Page 2403, line 15: Which assumption do you make in order to assume that this covariance matrix is diagonal?

Page 2403, line 14–18: Please explain the origin of the numbers used here and reproduce the derivation of this term avoiding the citation to Watts et al. (1998) which is not a peer-reviewed paper.

Page 2403, line 19–23: Please give a justification for the numbers used here.
Page 2403, line 19–23: Sayer et al. (2011) states that the forward model error is underestimated. Please comment on this.

Page 2404, line 4: What do you mean by *surface temperature-measurements*? As part of the state vector $x$ it represents a result of the retrieval.

Page 2404, line 5, line 21–22: What is the difference between a priori and first guess values? Where are first guess values used?

Page 2404, line 7–9: Is there a particular reason or reference for the choice of these (reasonable) values?

Page 2404, line 16–17: Why do you use such a value of $10^8$? What does it mean *in the absence of useful information*? When do you encounter such a situation?

Page 2404, line 19: Please give a justification for the values of the a priori errors used here.

Page 2404, line 20: Please explain why the a priori covariance is diagonal.

Page 2404, line 5–22: Please discuss (here or elsewhere) how strong the dependence of the retrieval results on the a priori values is. This could be tested in Section 10 by means of the simulated data or shown in Section 12 for that selected example. Are there quantities that show a stronger dependency on a priori values? Is the algorithm capable of resetting a positive a priori cloud fraction (i.e. $f > 0$) to zero? In contrast, would it make sense to apply the retrieval to all pixels (even to those with $f = 0$) and see whether also cloud detection is then refined by the algorithm?

2. Description of the RTMs

Page 2398, line 21–27: I find the list of unknown parameters too long to be presented here. I would rather mention the use of LUTs to account for multiple scattering effects in clouds and present the exhaustive list of parameters at the end of Section 5.1 and
5.2 respectively. Please notice by the way that all quantities mentioned here show an explicit dependency on $\tau$ and $r_{\text{eff}}$ that is not present on pages 2400–2402. It would be interesting as well to know which sample points have been selected for the LUTs, i.e. the grid defined by $\tau$, $r_{\text{eff}}$, $\omega_0$ and $\omega_r$. This information could also be given in Section 4.

**Page 2399, line 1–4:** What does *Radiative is performed in quasi-monotonically* mean? Is this a standard terminology? Please give a reference to this method! What is the inaccuracy when compared to correlated-k methods? Please clarify also that you refer here to the “non-DISORT” part of the RTM used for the FM.

**Page 2399, line 5–6:** How does this work? You need this in order to compute the derivative of the cost function $J$, don’t you?

**Page 2399, Eq. (6):** Please define $L$ (radiance) in general and $L^r_{\lambda}$ in particular. What does the superscript $r$ stand for? Reflected? It is also used in $\omega_r$.

**Page 2399, line 17–20:** This sentence says that

$$R(\bar{\lambda}, \omega_0, \omega_r) = \frac{L^r_{\lambda}(\omega_r)}{E^0_{\lambda}},$$

but in reality $R(\bar{\lambda}, \omega_0, \omega_r)$ is correctly given by Eq. (8). Please adapt the sentence.

**Page 2399, line 20:** For consistency to Eq.(6), you should use $R_{\bar{\lambda}}(\omega_0, \omega_r)$ instead of $R(\bar{\lambda}, \omega_0, \omega_r)$.

**Page 2399, line 18:** $E^0_{\lambda}$ is an irradiance indeed but it cannot represent the irradiance the satellite would measure if .... A satellite can only measure a radiance. $E^0_{\lambda}$ represents the solar irradiance at top of atmosphere on a plane perpendicular to the incoming radiation convolved with the spectral response function of the given sensor channel.
Page 2400, line 4–6: You have already introduced the “Sun-normalised radiance” and the top of atmosphere reflectance. Now you introduce the bidirectional reflectance factor $R(\lambda, \omega_0, \omega_r)$. Comparing to Eq. (8) it seems that this new quantity is simply the spectral analogon of $R(\bar{\lambda}, \omega_0, \omega_r)$. However, if I look at your definition, it turns out to be something different that has nothing to do with a radiance measured in a particular solid angle as it was the case in Eq. (8) but only with radiant fluxes. Please correct this definition, take care of notation, introduce only quantities that are really needed and stick to them through the whole manuscript. By the way, I have never heard of a diffuse surface, while an ideal Lambertian surface is a well-known concept.

Page 2400, line 12: Please discuss the uncertainty due to the neglection of Rayleigh scattering, especially in the short wave channels.

Page 2400, line 13–14: For gas absorption you mention here the use of standard atmospheric profiles (from Anderson et al. (1986)?), while in Section 4 (page 2396, line 6–7) you mentioned The clear-sky atmosphere is defined by temperature and humidity profiles taken from ECMWF analyses (ECMWF, 2008). A fixed mid-latitude ozone profile is assumed (relevant for modelling atmospheric transmission in the visible channels). Does this mean that you use different gas profiles for clear-sky and cloudy-sky? I thought that for clear-sky calculations you used the same model with $R_{CLD} = 0$ and $T_{CLD} = 1$ (see also Eq. (13)). Can you please explain this issue?

Page 2400, line 18: $R_{SFC}$ is called reflectance (as $R$ in Eq. (8)). Do you mean a bidirectional reflection distribution function (BRDF) $R_{SFC}(\omega_0, \omega_r)$? Then it should not be called a Lambertian surface. Do you simply mean a surface albedo? In that case you could forget about the $2\pi$ dependency in Eq. (9).

Page 2400, line 21: Please specify irradiance terms: which terms do you mean? Which irradiance?

Page 2400, line 24: $R(\omega_0, \omega_r)$ is not the spectral bidirectional reflectance factor at
the top of atmosphere because the below cloud absorption is neglected. The correct form of the spectral bidirectional reflectance factor at the top of atmosphere is given in Eq. (12). So what is the meaning of Eq. (9)?

Page 2400–2401, Eq. (9): Many new terms are needed in order to understand this equation. They are all defined afterwards. I think that it would be easier for the reader to know all quantities before Eq. (9) is presented. Please define the quantities prior to Eq. (9).

Page 2401, line 4–6: Terms like spectral directional-hemispherical total transmittance factor or spectral hemispherical-directional total transmittance factor are not immediately clear, so please explain what they mean and write down their definition as integral over $\omega_r$ or $\omega_i$.

Page 2401, Eq. (10): The differentials $d\Omega$ should read $d\omega$. The two integrals (from 0 to $2\pi$ and from 0 to $\pi/2$) refer to the integration of $\Phi$ and $\theta$, but these variables are not explicitly contained in the integrand. Furthermore, two such integrals are meant (i.e. four in total) because the variables are $\Phi_i$, $\theta_i$, $\Phi_r$, $\theta_r$. I find this notation not completely correct, so I would replace these two integrals by two integrals of the type $\int_{2\pi} d\omega$ that should indicate hemispheric integration over $d\omega$. This notation should be then explained in Eq. (7). Finally, I find the choice of $R_{C LD}(2\pi, 2\pi)$ to indicate the quantity described in the right hand side of Eq. (10) not very nice because already in Eq. (8) a similar quantity was introduced. However, in Eq. (8) no integration was contained, while here a twofold integration is meant. Please use a clear, well-defined terminology for all quantities.

Page 2401, line 12: Why $66^\circ$?

Page 2401, Eq. (12): Please tell what $R_*$ is (see also one of my previous remarks about the derivation of the TOA reflectance). In principle, you could show the exact
derivation of this equation instead of Eq. (9).

**Page 2402, line 6:** I thought RTTOV could also provide cloudy radiances. Is this correct? If yes, why don’t you use it?

**Page 2402, line 9:** reflection is a word that is usually reserved for solar radiation. I would rather speak of scattered radiation.

**Page 2402, Eq. (15):** I think that the correct version of this equation reads

\[
L^\uparrow(\omega_r) = \left( L^\uparrow_{bc}(\omega_r) T_{CLD}(2\pi, \omega_r) + \epsilon_{CLD} B(T_{CLD}) + L^\downarrow_{ac}(\omega_r) R_{CLD}(2\pi, \omega_r) \right) \cdot \exp(-\tau_{ac}) + L^\downarrow_{ac}(\omega_r).
\]

Here, however, the same symbol \( T_{CLD} \) is used for transmission and for the temperature of the cloud. Please correct this as well. Furthermore, shouldn’t you use \( \exp(-\tau_{ac}/\cos \theta_r) \) instead of \( \exp(-\tau_{ac}) \)?

**Page 2402, line 15:** What do you mean by effective emissivity? It is usual to term \( f \cdot \epsilon \) the effective emissivity, where \( \epsilon \) is the cloud emissivity and \( f \) the cloud fraction.

Finally, accuracies of the short wave and thermal RTMs with respect to DISORT or other radiative transfer codes should be quantified, if possible. Is it also possible to say how large the RTM error contribution to the overall retrieval error is?

3. Literature References

References about other cloud optimal estimation schemes should be given. These may include but are not limited to Watts et al. (1998); Heidinger and Stephens (2000); Miller et al. (2000); Baran and Havemann (2004); Heidinger (2003); Heidinger and Pavolonis (2005). The present work should also be discussed in view of these algorithms, and differences, similarities, advantages, disadvantages and limitations should be emphasized.

4. Quantitative Conclusions

C976
Sections 10 and 11 deal with the retrieval scheme performance. They are very interesting and many plots are shown. However, only a very few quantitative conclusions are taken from them. For instance, on page 2408, line 4–6, the authors identify cloud types that are difficult to retrieve and talk about thin clouds, clouds with small $r_{\text{eff}}$, and extremely thick clouds. Already these cloud classes should be specified by telling what a thin cloud is and so on. Furthermore, also underestimations and overestimations of cloud parameters as well as indications about the values of the cost function should be quantified in both Sections 10 and 11.

**Further Comments**

Since the validation paper by Sayer et al. (2011) uses the acronym ORAC for the algorithm described here, I would recommend to introduce this acronym in the present paper as well.

**Abstract:** It contains a too long introduction that can be shifted to Section 1 (page 2390, line 1–8), and too few quantitative assertions about the retrieval itself and its performance. It also does not tell anything about the novel aspects of the scheme or about its importance. Finally, the citation of (Sayer et al., 2011) in the last line should also be removed.

**Page 2396, Section 4:** At some point in this section (possibly at the beginning) it should be clarified that the algorithm needs a cloud mask to start with and that cloud thermodynamic phase determination is done separately.

**Page 2396, line 10:** To my knowledge, Cox and Munk (1954a,b) derive BRDF parameters for water reflectance, but you treat surface as a Lambertian albedo. How do you transform BRDF into albedo?

**Page 2396, line 11:** Please insert a citation for the MODIS albedo product (e.g. Schaaf et al. (2002)).
Page 2396, line 14: While you consider surface albedo uncertainties in your OE, you do not consider emissivity uncertainties. Why? What is the effect of this choice on retrieval accuracy?

Page 2396, line 19: Can you quantify how large the impact of this assumption (cloud top=cloud bottom) is for the accuracy of the radiative transfer calculations?

Page 2397, line 7: Ice particle optical properties are not uniquely defined since they strongly depend on shape. Please provide here some detail about the type of ice particles described in Baran and Havemann (2004). Please also give the definition of $r_{\text{eff}}$ used for ice particles.

Page 2397, line 13: Can you give a reference for this size distribution? Which values do you assume for $r_m$?

Page 2397, line 16: This definition of effective radius cannot be correct because you have the fourth moment of the particle size distribution divided by the second moment of the particle size distribution. This has a unit of $\mu m^2$. Usually, according to Hansen and Travis (1974), you use the third moment divided by the second moment of the particle size distribution. Please correct this equation.

Page 2398, line 10: Good convergence means that all channels could be reproduced by the model. On the other hand, the retrieved solution must not be the only one solution of the problem: even good convergence could provide a solution that does not correspond to reality. Please comment on this.

Page 2398, line 10: Why do you think that these cloud classes (with strong vertical variation of $r_{\text{eff}}$ and phase) are among the most difficult ones? The retrieval could find an effective radius, a phase and a cloud top temperature that enable to correctly reproduce the measurements even if the link to the real cloud is difficult to understand.

Page 2403, line 7: $3 \times 3 \text{km} \rightarrow 3 \times 3 \text{km}^2$. By the way: is it $3 \times 3$ or $3 \times 4$ as you state C978.
on page 2404, line 11?

**Page 2405, line 3 + 7 + 10**: Can you give a justification for these numbers (260 K, 23 and 20 µm)? Are they set empirically?

**Page 2405, line 12–13**: Can you please state how unreliable the results are or how often this could take place?

**Page 2405, line 15**: Is this really a solution? Couldn’t there be cases where you have accurate solutions for both cloud phases?

**Page 2405, line 16–17**: A couple of time in the manuscript you mention CPU time. Can you please give some indication about CPU time consumption of the algorithm?

**Page 2406, line 10**: *How well* is maybe not the most appropriate question to ask here. Since your model does not know anything about multilayer clouds you will always get either a liquid water or an ice cloud as a result. This is in any case wrong, so you first have to define what you mean by a “good performance of the algorithm” in this case? Do you expect an optical thickness that is the sum of the optical thicknesses of both layers? And so on... Please specify this in the manuscript.

**Page 2406, line 16–17 + 21–23**: Please explain the meaning of linear error simulations and non-linear error simulations.

**Page 2407, line 12**: Which wind speed do you select for the Cox&Munk parameterisation?

**Page 2407, line 13**: When you fix one of the (output) variables, like cloud fraction here, do you have to implement a new retrieval scheme with one variable less? How does this work? Please comment on this.

**Page 2407, line 4–13**: What about $p_c$? What is the choice you make for this variable?
Page 2407, line 4–13: Do you use DISORT for these simulations? Please specify.

Page 2408, line 6–10: If in your simulation you do not vary ice crystal shape then this cannot be an issue here. However, for an application to real data this is of course an important point. Please clarify.

Page 2408, line 11: Please re-number Section 11 into Section 10.2.

Page 2408, line 19: By simulations you mean here retrievals, don't you?

Page 2409, line 1: How are cloud optical thicknesses varied? According to Section 10.1 or according to the caption of Fig. 3?

Page 2409, line 11: Please specify what you mean by “true” cloud parameters: since in your simulations you always have two parameters (for the liquid water and the ice layer) it is not clear what you mean here.

Page 2411, line 1–3: Multilayer gives high cost functions, but conversely high cost functions could occur for different reasons. This means that high cost functions could not be reliable indicators of multilayer clouds. What about for instance the situation where the initial cloud mask incorrectly identifies a cloud in one pixel that in reality is not there? Please comment on this and make a list of possible situations producing high cost functions (unless it has been already discussed in connection with the cloud model).

Page 2412, line 1: Surface temperature is missing. Please add this quantity as well. It could also be interesting to see the a priori cloud fraction and the a priori surface temperature.

Page 2412, line 5: where the cloud is thin or maybe inexistent? Comparing retrieval results it seems that some pixels flagged as cloudy do not contain cloud (especially over the Mediterranean) and thus produce inaccurate results. Please comment on this.
Page 2412, line 11: The water clouds over Africa as well as the large ice cloud over Central Europe also show some yellow cloud fraction band. Do you think this is realistic? Please comment on this.

Page 2412, line 15: *shadowing*: or reflection by cloud sides or ...

Page 2412, line 19–20: *More well tuned ...*: what does this sentence mean? What is the surface forward model?

Page 2413, Conclusions: I am missing some considerations about the global applicability of the algorithm. Can it be applied over the poles as well as in the Tropics and in mid-latitudes?

Page 2413, line 17: ice crystals: Please remind the reader of the origin of these ice cloud models by adding the reference to Baran and Havemann (2004).

Page 2415: Baran (2005) has not been cited in the text. Please correct this.

Page 2421: I would indicate the smallest $r_{\text{eff}}$ retrieved and not 0 which is a sort of fill value.

Page 2423: Scales for $r_{\text{eff}}$ go from 1 to 27 $\mu$m but the simulations for water clouds run from 3 to 25 $\mu$m. For ice clouds 70 $\mu$m were not simulated neither. Please correct this.

Page 2424: As already remarked, the simulated optical thicknesses mentioned in the caption and those in the text (as well as in the plot itself) differ. Please correct this.

Page 2425: Since the scale of cloud optical depth in panel (a) runs from -2 to +2 I assume that some difference it shown and not the values themselves. Please specify. Furthermore, in this Figure only multilayer cloud results are shown, so I think that the caption should read: *Non-linear retrievals of multilayer cloud performed ...*

Page 2426: The sentence about the “truth” contained in the caption should appear in C981.
the manuscript as well. Furthermore, I do not think this is the best choice you can do. For instance, when retrieving cloud optical thickness I would be happy if the scheme would give me back the sum of the optical thicknesses of the two layers. Please explain why you decided this way.

**Page 2427**: Left panel: It seems that for some cloudy pixels (for which for instance cloud top pressure was derived) the cost function has a white colour (see the large ice cloud over Central Europe). What does this colour mean here? Large cost functions over the Mediterranean seem to indicate the absence of clouds. It would be interesting to see the distribution (histogram) of the cost function and to get the information about how many pixels were retrieved accurately. In addition, the amount of pixels with good convergence would also be helpful to better understand the retrieval. Can you please discuss these issues for this example or more in general in the manuscript?

**Page 2429**: The scale up to 1000 (In Table 2 you say the maximum range is 320!) shows the peaks but reduces the contrast. So it looks like if most clouds had an optical thickness of 10. Please reduce the scale to allow a better evaluation of optical thickness.

**Page 2430**: One has the impression that errors for liquid clouds are larger than for ice clouds. Is it really like this? Why?

**Technical corrections**

Please make sure that only one version of data set, data-set or data set is used in the manuscript.

Please make sure that only one version of on board and on-board is used in the manuscript.

Please make sure that only one version of multilayer and multi-layer is used in the manuscript.
manuscript.

Please make sure that punctuation is present at the end of equations. For instance, Eq. (1) does not contain any period at the end, while Eq. (7) does.

Two acronyms (OE and OEM) have been introduced for almost the same thing and have been used once respectively. Please consider the use of only one of them or the complete avoidance of both of them.

Page 2392, line 15: Please define SLSTR already here and not on page 2393, line 24.

Page 2394, line 6: rear-ir -> near-ir.

Page 2394, line 12: record -> records (?).

Page 2395, line 16: Levenberg-Marquart -> Levenberg-Marquardt.

Page 2397, line 4: Size distributions -> Particle size distributions.

Page 2398, line 1: acknowledges -> acknowledged.

Page 2398, line 4: recognised -> recognise.

Page 2398, line 26: DISTORT -> DISORT.

Page 2399, line 9: field-of-view -> field-of-view FOV.

Page 2399, line 14–15: where \( \omega \) -> where the solid angle \( \omega \).

Page 2400, line 1: Suns -> Sun’s.

Page 2400, line 3: Sun normalised -> Sun-normalised.

Page 2402, line 10: cloud. -> cloud:

Page 2403, line 11: with values on the diagonal equal -> with values equal.
Page 2403, line 14: see Smith (2005) for the origin of these values → (see Smith (2005) for the origin of these values).

Page 2403, line 25: The state-vector used → The state-vector \( x \) in Eq. (1) used.

Page 2403, line 26: optical depth → cloud optical depth.

Page 2404, line 8: are are 15 → are 15.

Page 2404, line 18: unconstrained → unconstrained.

Page 2404, line 18: For cloud fraction, the → For cloud fraction, the.

Page 2405, line 10: the the retrieval → then the retrieval.

Page 2405, line 11: Only one change of phase ... : Please write this sentence separately from this item, on a new line, in line with line 4.

Page 2405, line 14: simple → simply.

Page 2407, line 7: 10„ → 10,

Page 2407, line 16: radii range → radius ranges.

Page 2408, line 5: radii → radius.

Page 2408, line 24–25: Please separate the assumptions about \( p_c \) and \( r_{\text{eff}} \) as you did in Section 10.1.

Page 2409, line 13: noted. → noted:

Page 2409, line 15: optical depth of \( \leq \) → optical depth \( \leq \).

Page 2411, line 10: achive → achieve.

Page 2411, line 11: optical; depth → optical depth.

C984
Page 2412, line 3: made. –> made:


Page 2413, line 23: exected –> expected.

Page 2414, line 16: from ATSR-2 –> from ATSR-2 for the compilation of the GRAPE data set.

Page 2421: and range –> and range for GRAPE.

Page 2422: though multiple scattering between the atmosphere and surface –> through multiple scattering between cloud and surface.

Page 2427: results form –> results from.

References


