Reply to Reviewer #1: *Stephen Wilson*

Adaption of the MODIS aerosol retrieval algorithm by airborne spectral surface reflectance measurements over urban areas: A case study
Author(s): E. Jäkel et al.
MS No.: amt-2015-139

We thank the reviewer for the time and efforts he spent reading our manuscript and providing valuable advises which we truly appreciate. Please find below a discussion of the reviewer's comments (italic). Changes/additions made to the text are underlined and given in quotes.

I think the fundamental question this work poses needs to be discussed more explicitly

- Is it possible to generally improve the retrievals over urban areas without heavy tuning to each and every site? To apply the measurements to Beijing the authors limited their analysis to 6 months of the year and a restricted sza. Was this necessary? What magnitude are the errors introduced if these restrictions are removed?
Thanks for this question. Our original intention was to minimize retrieval uncertainties due to the surface assumption by selecting data with respect to viewing geometry and season. We expected larger variability of the slope parameter for the winter months (different vegetation) and for sensor zenith angles different from nadir direction (BRDF dependence). However, in the new version of the manuscript we used an extended set of MODIS data between 2010 and 2014. The mean difference of the adapted C6 method was improved, whereas for the modified C5 method nearly similar results were found. For this reason we changed Fig. 5 and the discussion of the Beijing-comparison as follows:
"Figure 5 presents the scatter plot and regression lines of the relation between AOD derived from MODIS versus measured by AERONET. Recall that when the retrieval expects a dark surface, a larger observed TOA reflectance is interpreted as enhanced aerosol contribution. Both operational algorithms systematically overestimate the AOD as compared to AERONET, which is consistent with urban surfaces being brighter than expected. However the C6 version is less biased. Since the operational C6 equations include a larger slope parameter a(NDVI0.65/2.1) than does C5, the estimated surface reflectance is larger, leading to reduced retrievals of AOD. Using the measured a(0.47/0.65)-spectral slopes in place of either C5 and C6-assumed slopes will further reduce retrieved AOD. Both modified methods show a reduction of mean differences between AOD(AERONET) and AOD(MODIS) as displayed in Table 3. However, while the mean bias is reduced in either version, modifying the measured slopes also leads to retrieval of negative AOD values, as shown in Fig. 5. Modifying the C6 version, which already had reduced bias compared to C5 operational version, tends to lead to even more retrievals of negative AOD. At the same time, the AOD standard deviations have increased (from 0.12 to 0.17) when using the modified C6 algorithm. One of the reasons might be the different sensitivity of TOA reflectance on changes of AOD over bright and dark surfaces. Over dark surfaces (as assumed in the operational retrieval), changes of AOD result in larger changes of TOA reflectance than over bright urban surfaces (as assumed in the modified retrieval). Consequently, measurement uncertainties have a larger effect on the retrieval uncertainty for an urban surface."

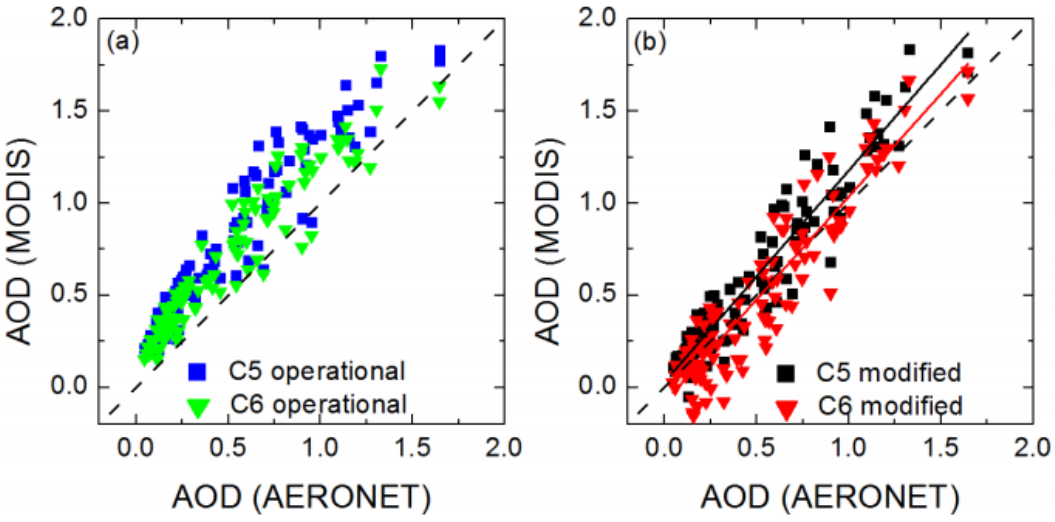

Figure 5. Scatter plot of AOD retrieved from MODIS data using (a) the operational and (b) modified collection 5 and 6 method and the AERONET data for Beijing. The dashed line marks the 1:1 relation.

Fig. 1: New version of Fig. 5 (old manuscript).

In addition we adapted also the Abstract and the Conclusions:
"A significant reduction of the differences between the AOD retrieved by the modified algorithms and AERONET was found, whereby the mean difference decreased from 0.27 ± 0.14 for the operational C5 and 0.19 ± 0.12 for the operational C6 to 0.10 ± 0.15 and 0.02 ± 0.17 by using the modified C5 and C6 retrievals. Since the modified algorithms assume a higher contribution by the surface to the total measured reflectance from MODIS, consequently the overestimation of AOD by the operational methods is reduced."

"While the operational algorithms C5 and C6 have shown a significant positive bias compared to AERONET with a mean difference of 0.27 and 0.19, the modified algorithms revealed a reduced mean difference of 0.10 and -0.02. […] For future studies airborne imaging spectrometer measurements are suggested to characterize the slope parameters for urban areas also for different viewing geometries than the nadir observation. This will improve the corrections of the operational retrievals due to effect of the bidirectional reflectance distribution function (BRDF), since the anisotropy reflection of the surface is better described by the BRDF than assuming a Lambertian surface (Escribano et al., 2014)."

Numbers of the mean differences are now listed in a separate table:

Table 3. Mean value and standard deviation of differences between AOD(MODIS) and AOD(AERONET).

Retrieval Method	Mean Difference ± Standard Deviation
C5 modified	0.10 ± 0.15
C5 operational	0.27 ± 0.14
C6 modified	-0.02 ± 0.17
C6 operational	0.19 ± 0.12

Also, other analyses (cited in this paper) have modified coefficients to retrieve AOD for cities in China – how do these methods compare? Is there any consensus to be observed? (It would appear so.)

Other publications as cited have also shown lower performance for urban surfaces. The analysis performed to improve the AOD retrievals are based on various methods. Bilal et al. (2014) developed a new retrieval using a complex set of additional data (MYD09, MYD04, AERONET). They have shown the relative mean bias of their retrieved AODs and AERONET data for Beijing for the years 2012 and 2013. The first version of the manuscript already included a comparison with our results.

Another publication (de Almeida Castanho et al., 2007) applied modified slope parameter ρsurf(0.66)/ρsurf(2.1) for their AOD retrieval over Mexico City. Oo et al. (2010) adjusted the VIS/SWIR slope parameter for retrievals over New York City and Mexico City. The mean relative bias was not given in these studies. De Almeida Castanho et al. (2007) found a similar correlation coefficient of 0.93 for the relation between Sun photometer measurements and the adapted AOD MODIS retrieval. We did not give this number in the manuscript, since the observed region is Mexico City not Beijing.

It is difficult to compare the results of the publications since different data sets are used (different region, year). But overall we can conclude that adjusting the surface assumptions for urban regions helps to improve the AOD retrievals.

Secondly, you do not have measurements at 2.1um. However, you have values just below 2.0um. Your modelling shows that there is not great variability at these wavelengths. Why have you not looked at (say) 1.95um vs 0.65 to at least test the stability of Equation 2? It may be that the S/N is not good enough?

Fig. 2a shows the scatterplot of the surface reflectances at 0.65 and 2.0 µm as derived from aircraft measurements. We used some mean value of this poor correlation (R=0.62) for the adapted retrievals but it didn't improve at all. Furthermore, since we have more reliable data of the surface albedo we have plotted the mean spectrum between 2.0 and 2.1 µm calculated for measurements along the flight path over Zhongshan (Fig. 2b). It shows that at least for the surface albedo the surface reflection property at 2.0 µm differs from that at 2.1 µm. Therefore we did not apply a new slope assumption for 2.1 versus 0.65 µm for the adapted retrievals.

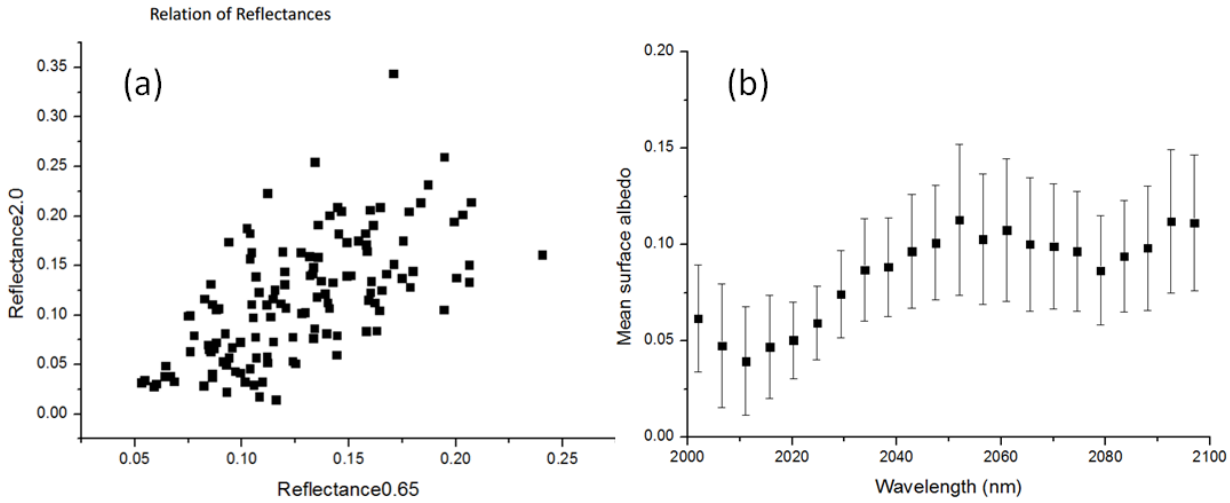

Fig.2: (a) Scatterplot of surface reflectances at 0.65 and 2.9 µm. (b) Wavelength dependence of mean surface albedo.

We discuss the issue briefly in the manuscript:
"According to Eqs. (2) and (3) the surface reflectance at 0.47 µm wavelength in the MODIS AOD algorithm is derived from that at 0.65 µm based on the surface reflectance at 2.1 µm. The airborne measurements of the surface reflectance are not reliable for wavelengths larger than 2.0 µm, due to the low sensitivity of the albedometer. The relation between reflectances at 2.0 µm and 0.65 µm cannot be used as approximation to modify Eq. (2) because of the strong differences between 2.0 and 2.1 µm, and the finding that reflectances at 2.0 µm and 0.65 µm are poorly correlated (R=0.62).
Therefore, our strategy is to modify only the blue/red fit parameter of Eq. (3) to match the airborne observations above Zhongshan."

Thirdly, the calculations indicate that C5 (modified) appears to work better than C6 (modified). Given that C6, amongst other things, corrects coding errors (as outlined in detail in equations 5 – 7), why does it perform worse following correction?
In our first version of the manuscript we found mean differences between AOD(MODIS) and AOD(AERONET) of 0.09±0.18 (adapted C5) and -0.10±0.18 (adapted C6). Changing the slope parameter a(NDVI0.65/2.1) in C6 as described in Eq. (7) results in a larger surface reflectance (valid for most of the cases with 0.25 < NDVI < 0.75) than in C5. This gives lower retrieved AODs derived by the C6 method.
If we use additionally the adapted slope a(0.47/0.65) from aircraft measurements then we get a larger surface reflectance (rho_0.47) resulting in an even lower AOD. In this case an additional higher slope assumption for a(NDVI0.65/2.1) reduces the retrieved AOD even more. That's why we had a lower performance for C6 when filtering the data set for sensor zenith angles < 10° and Summer months.
By using the entire data set without restrictions of sensor zenith angle and season the slope assumption for a(NDVI0.65/2.1) of C6 seems to be a suitable approximation. Better agreement with AERONET data is found when the sensor viewing angle gets larger than 10°, probably caused by BRDF effects. To prove this, further measurements with different viewing geometries would be necessary.
When introducing the operational C5 and C6 methods we added some further information with respect to the main changes made from C5 to C6:
"Changes from C5 to C6 included corrections related to the Rayleigh/aerosol LUTs and modifications of the assumed aerosol type models that are prescribed for season and location (Levy et al., 2013). "

Minor issues

Page 7336 L13 change to "when AERONET data were also available."
Changed as follows:
"The subsequent impact was tested upon two versions of the operational algorithm, Collections 5 and 6 (C5 and C6). AOD retrieval results of the operational and modified algorithms were compared for a specific case study over Zhongshan, to show minor differences between them all. However, the Zhongshan-based spectral surface relationship was applied to a much larger urban sample, specifically to the MODIS data taken over Beijing between 2010–2014.
These results were compared directly to ground-based AERONET (AErosol RObotic NETwork) measurements of AOD."

P 7337 L17 remove comma
Changed.

P7337 L21 (as ocean water) -> (such as ocean water)
Changed.

P7339 L4 by -> using
Changed.

P7339 L21 -> "available at either 250 or 500 m resolution,"
Changed.

P7340 L22 Levy (2007) produced revised estimates that were used in both C5 and C6. (C6 produced in 2013 – 6 years later!)
Changed as follows:
"In particular, Levy et al. (2007) considered more variables for characterizing the wavelength-dependent surface reflectance, including scattering angle and general "greenness" of the surface."

P7340 L25 superscript "m" is not explained.

Changed as follows:
"The effect of geographical and seasonal variations on surface vegetation could be diagnosed through the normalized differenced vegetation index $NDVI_{SWIR}$, combining the measured reflectances ρ^m at 1.2 and 2.1 µm wavelength:…"

P7340 L 26 This sentence confused me. Above (L15) it was declared that atmospheric effects were small at 2.1 micron. Now the atmospheric backscatter ratio, transmissivity and TOA reflectance and the use of Look-Up-Tables are mentioned. A reference to the method would help here. Presumably that would also be the reference for the equations 2 – 4?

C4 assumes that atmosphere is nearly transparent at 2.1 µm, not C5 and C6. To make it more clear we added:
"Levy et al. (2007) relaxed the assumption of atmospheric transparency at 2.1 µm, and reformulated the aerosol retrieval as an inversion of three bands (0.47, 0.65, and 2.1 µm). In addition, detailed atmospheric correction was performed over many AERONET sites, so that the surface reflectance relationship became better characterized.…"
Then we also give a reference for Eqs. 2-4 as suggested by the reviewer:
"The surface reflectance at 2.1 µm wavelength is calculated as a function of measured TOA reflectance, the atmospheric backscatter ratio, the transmissivity, and the TOA reflectance given by the precalculated LUTs depending on AOD and aerosol type. According to Levy et al. (2007) the surface reflectance at 0.65 and 0.47 µm is a function of the surface reflectance at 2.1 µm wavelength, and is derived by:"

P7341 Equations 5 – 7. I could not find this in Levy (2013). I suspect that this is documented in the code as well. If this is the case, state that the summary given is derived from both Levy (2013) and the code.
Since the equations are not explicitly given in Levy et al. (2013) we added the reference for the C5 equations:
"Note, that due to a coding error, the C5 version's $NDVI_{SWIR}$ dependence as given in Eq. (10) from Levy et al. (2007) was reversed. This error was corrected in C6 (Levy et al., 2013)."

P7343 L1 Presumably the quantity quoted is 1 sigma? Please be explicit here and in the associated table (table 2).
We added in the text and in the caption of the table:

"Using Gaussian error propagation, the total uncertainty (one sigma) for the SMART albedometer, ranged from 3 to 14 % and depended on wavelength."
"Relative uncertainties of the different error sources and uncertainty (one sigma) of surface albedo and reflectance after Gaussian error propagation for …"

P7343 L5 delete "as presented in"
Changed.

P7343 L12 -> It took a while to decipher this paragraph, and as written it is still not a constrained problem I believe. I cannot see how the extinction of the well-mixed layer is set. Is it presumed that it has the value at the top of the layer above? Please clarify.

We have rewritten the paragraph. For further explanation (not shown in the manuscript) examples of the extinction profiles are shown in Fig. 3. It should clarify the procedure for the extrapolation.
"Since the LIDAR did not provide information on the vertical aerosol distribution up to 1000 m height, the vertical extinction profile was extrapolated from 1000 m to the surface in correspondence with the vertically integrated extinction coefficient (AOD) of the Sun photometer. The layer between surface and 1000 m height was separated into two sub-layers. Starting from 1000 m altitude the extinction was assumed to increase with decreasing height following the measured slope of the LIDAR extinction profile. The second sub-layer with a constant aerosol extinction coefficient was assumed below to represent a well-mixed layer. "

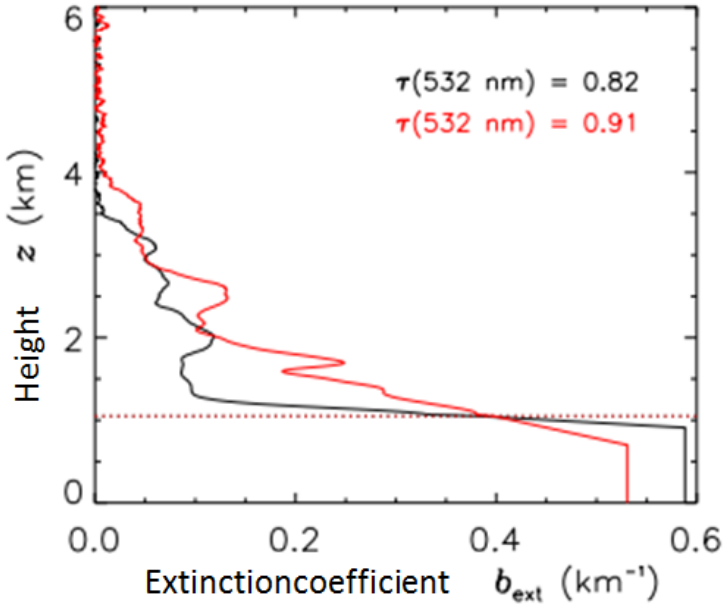

Fig.3: Examples of LIDAR profiles of extinction. Values below dotted lines are extrapolated.

P7344 L5 Remove comma after "Note"
Changed.

P7344 L10 Given that you define the reflectance retrieval, some of the following text would be simpler to follow if you defined the albedo calculation as well.
We added the definition of the surface albedo as follows:
"The spectral surface reflectance was calculated according to the atmospheric correction technique, originally developed for the retrieval of the surface albedo (Wendisch et al., 2004):

$$\alpha_{s,\lambda} = \frac{F_\lambda^{\uparrow}(z_0)}{F_\lambda^{\downarrow}(z_0)}$$

based on airborne measured irradiance."

P7345 L7. The calculations are for a relatively high AOD of 0.9. It would be useful to know how dependent the analysis is on AOD.
We added the following lines:
"In general, lower AODs result in lower uncertainties of the derived surface properties, because the atmospheric contribution to the measured upward radiation decreases while the surface contribution increases. Therefore, uncertainties of the aerosol properties g and ω for lower AODs have lower impact on the retrieval results. Table 2 gives the sensitivities for AOD(0.53 µm)=0.9 representing an upper limit."

We will not show a systematic analysis of the sensitivity, as a function of AOD. But we will give numbers for a lower AOD:
"A similar study was performed for AOD(0.53 µm)=0.12 showing uncertainties between 3.2 – 10.8 % (surface albedo) and 6.6 – 14.4 % (surface reflectance)."

P7345 L 17. Suggest "In wavelength range I (0.4 – 0.6 um) the retrieved surface albedo and surface reflectance is very sensitive to the assumed aerosol properties. As AOD decreases with increasing wavelength, this sensitivity also decreases. At the longest wavelengths uncertainties in the retrieval are dominated by the measurement uncertainty of the SMART detector.
We changed it as follows:
"In wavelength range I (0.4–0.6 µm), the retrieved surface albedo and surface reflectance are sensitive to the assumed optical aerosol. As AOD generally decreases with increasing wavelength, the sensitivity also decreases. At the longest wavelengths (1.8–2.1 µm), the uncertainties are dominated by the measurement uncertainties of the SMART instrument, which has lower sensitivity and poor signal-to-noise-ratio in this spectral range.

P7346 L12 (Symbol theta_0) Do not recall see it being used elsewhere. Delete?
The symbol is removed to minimize the number of symbols.
"The solar zenith angle (SZA) ranged between 44° and 55° during the 1.5 h of the flight. The change of SZA was considered in the derivation of the surface properties."

P7347 L2. Sentence starting with "Therefore" describes what the albedo is dependent upon. However, the issue here is that the albedo as measured is not very dependent on the conditions directly below. I suggest stating this the other way round.
We have rewritten the sentences:
"Since albedo is an integral across all directions, its relative uncertainty is lower due to the directional averaging (or blurring). Consequently, the flight altitude and the aerosol conditions will have the most impact on the variability of the retrieved spatially-averaged surface albedo from airborne measurements (Jäkel et al., 2013). Resolving the small-scale variability of surface albedo requires low-flying aircraft or ground-based measurements."

P7347 L 17 "induces" not sure what this is supposed to mean in this context.
We removed the sentence.

P7347 L 19 "as a function" – add "a"

Changed.

P7347 L 15 "bias between both regression". The meaning is very unclear to me, even following the correction to "regressions"
We changed the sentence as follows:
"Clearly, the measured over-urban surface reflectance relationship for Zhongshan is much different than that assumed by the retrieval algorithm, with the slope coefficient being much larger (0.85 versus 0.49).."

P7348 L8 So what is new here? This seems to imply that, within uncertainty limits, the result would have been replicated by using the value of Levy (2007).
Levy et al. (2007) give a relationship between reflectances at 0.47 and 0.65 nm of 0.766 for urban sites in Summer. First, Levy et al. (2007) did not apply this relationship for the retrieval of AOD for urban sites in their paper. Second, this relationship is based on satellite measurements which may be biased by uncertainties of the atmospheric correction itself. However, as mentioned by the reviewer, the relationship agrees quite well with our observations which can be considered as confirmation of the satellite results.

P7349 L10 "Data and pixel were chosen" – So the "pixel" choice is not selecting data? Please clarify?
We removed the word "pixel" to simplify the sentence:
"The MODIS case was chosen to fit the time and area of the airborne measurements as best as possible in order to minimize differences due to season and viewing geometry."
In the original manuscript the word "data" referred to the Modis data file, whereas the word "pixel" referred to a specific data point within this file corresponding to the latitude/longitude matching with Zhongshan coordinates.

P7350 L8 Change to "Twenty one days were identified that aligned with the AERONET.."
We changed it as follows:
"Since there are such few collocations of AERONET data with MODIS overpass in Zhongshan, it is considered to test the urban surface ratios for use over another city in China. Here the spectral slopes a(0.47/0.65), derived from the airborne measurements over Zhongshan, were applied to MODIS data over Beijing. For Beijing, there are 120 days of concurrent AERONET data, observed between 2010 and 2014 matching in time (± 30 minutes) and space (closest MODIS pixel to AERONET station).

P7350 L 19 "significant" not "significantly"
We changed the entire section. See major comments.

P 7350 L 22 "particular," – also remove "also"
We changed the entire section. See major comments.

P7351 L1 I don't understand this argument. Could you clarify please?
We explained it in another way:
"One of the reasons might be the different sensitivity of TOA reflectance on changes of AOD over bright and dark surfaces. Over dark surfaces (as assumed in the operational retrieval) changes of AOD result in larger changes of TOA reflectance than over bright urban surfaces (as assumed in the modified retrieval). Consequently, measurement uncertainties have a larger effect on the retrieval uncertainty for an urban surface.

P7351 L5 "and"
Changed.

P7351 L6 remove one "correlation"
Changed.

P7353 L6 Reword "First the AOD was retrieved with the operational algorithm for each surface type"
Changed.
.
P7354 L8 "This issue was investigated in two ways."
Changed.

P7355 L4 Not sure that this sentence is sufficiently clear. "show only a low positive bias" implies that the modified algorithms are known to be better, and then states that the difference is apparently not significant. I would suggest rewording with one message
We changed the sentence as follows:
"For the test case over Zhongshan, the AOD values retrieved by the standard algorithms 490 [AOD = 0.84 (C5) and AOD = 0.85 (C6)] agree with those retrieved by the modified algorithms [0.83±0.02 (C5) and 0.82±0.02 (C6)] within the range of measurement uncertainty."

P7355 L14 "alloy"? Not certain what is meant here.
We removed the sentence.

P7356 L14 "certainly enhance" Adding an extra degree of freedom does always allow a reduction in the residual but I am not sure that it is guaranteed that the improvement would be significant. Especially in the general case where the target sites are not well characterized.
We removed the sentence.