Response to Reviewer #1 (OSD-8-C928-2012):

General Comments (Reviewer #1):

C01: While the paper is well written and organized, it lacks scientific rigor. It is highly important that new satellite products be validated with in situ observations. The authors wave their hands at this aspect stating that there is not enough in situ data for a validation and there are limitations to the collocation of an in situ observation to the size of a SCIAMACHY pixel. However, *Bracher et al.* (2009) was able to find in situ data for a proper validation as well as many authors of other satellite phytoplankton functional types retrievals published in the literature.

R01: We agree that we have to show a validation our PhytoDOAS result as commented by the reviewer. Therefore we added further comparisons on global scale for 2005 and regional scale for 2003-2010 to other coccolithophore related satellite products in the manuscript as pointed out below. The success of the PhytoDOAS method to distinguish diatoms and cyanobacteria has already been shown by *Bracher et al.* (2009). Since the spatial resolution of in-situ point measurements and the large SCIAMACHY pixel is too different, collocations will hardly match the same optical average of phytoplankton signatures sampled for both methods. This makes it generally difficult to validate phytoplankton retrievals extracted from SCIAMACHY data with in-situ measurements. However, in the case of coccolithophores this becomes even more difficult, because the different in-situ techniques are not encompassing the whole group of coccolithophores: Via the HPLC method, from marker pigments the biomass of the group of haptophytes are globally assesssed only and the pigments of other haptophyte species (e.g., *Phaeocystis*) are spoiling the measurements of coccolithophores; with flowcytometry only the larger groups of pico- or nanoeukarotic phytoplankton are identified; by microscopic counts cells less than 5 micrometer cannot be identified properly. Hence, there are uncertainties in the determination of in-situ coccolithophore chl-a conc. as a whole, especially on global scale, while the satellite retrievals detect the optical imprints of all coccolithophores regardless of size due to their specific pigment composition. As this uncertainty does not exist in the case of in-situ data of diatoms and cyanobacteria, the validation was relatively straightforward in *Bracher et al.* (2009). Here the chl-a conc. of these groups were directly inferred from a large set of HPLC pigment measurements only and not several techniques had to be combined. Moreover, the diatoms retrieved by the improved PhytoDOAS method are in good agreement with the respective results of *Bracher et al.* (2009), and were therefore not a focus of this manuscript.
In order to verify our PhytoDOAS results for coccolithophores we now added a direct comparison of the monthly mean values over eight years for a region in the North Atlantic of the PIC MODIS satellite product to PhytoDOAS (Fig. 9). The selected region, within the North Atlantic Ocean, was one of the three selected in *Sadeghi* et al. 2012. Within this publication, at three oceanic regions, where coccolithophores blooms are frequently forming, phytoplankton dynamics were studied using PhytoDOAS products in conjunction with other satellite products characterizing phytoplankton and various geophysical parameters. In addition, our global seasonal 2005 coccolithophore PhytoDOAS results were compared to chl-a conc. derived for haptophytes. These data were obtained according to the method developed by *Hirata et al.* (2011) from a synoptic pigment-based relationship of haptophytes chl-a to total chl-a and applying it to the SeaWiFS chl-a product (see Fig. 7 and Fig. 8) The agreement is very good, despite some deviations which can be either due to other haptophytes than coccolithophores dominating in the particular regions or due to that the synoptic relationships are based on empirical relationships inferred from a regionally biased data set (e.g. very few data in the Southern Ocean and tropical Pacific).

NOTE: Motivated by some comments given by the reviewers, we have been convinced that the title

chosen for this MS has not been adequate enough. Since our method improvement so far is focused on the PFT target of coccolithophores, we have modified the title as follows:
Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data.
This was also motivated by the lack of an appropriate in-situ or satellite dinoflagellate product to evaluate the PhytoDOAS dinoflagellate data set. *Hirata et al.* (2011) showed that the currently available global HPLC-based data set on dinoflagellates is not normal distributed and an abundance-based relationship for dinoflagellates chl-a to total chl-a could not been inferred (which can then be used to derive a satellite-based estimate of dinoflagellate distribution).

C02: The comparison of the new PhytoDOAS PFTs with model output (NASA NOBM) and satellite derived products (MODIS PIC) is insufficient. The model output and derived satellite products have there own associated uncertainties. Validation with in situ observations of all the retrieved PFTs needs to be added to this manuscript before it should be accepted for publication.

R02: Regarding the further validation, we already answered and explained this in R01 and addressed it accordingly in the revised manuscript (sec. 2.4). Sure, the other satellite data have their own uncertainties. As discussed in R01, the observed waters of satellite and in-situ point measurements are very different! Therefore, a set of comparisons have been done to evaluate the quality of PhytoDOAS coccolithophores: Particular Inorganic Carbon (PIC), NOBM modeled data and global distribution of haptophytes, based on *Hirata et al*. (2011).
It should be noted that the PhytoDOAS method, compared to other approaches of phytoplankton retrievals, is based on a completely different algorithm, concerning the use of hyperspectral information. Since both satellite products, haptophyte chl-a of *Hirata et al.* (2011) and PIC from MODIS have been validated and their uncertainties have been assessed, the comparisons of PhytoDOAS to these products show clearly our algorithm's functionality. We added now a short discussion on the uncertainties of the PIC data and haptophytes product of *Hirata et al.* (2011) in section 3.2.

C03: There also needs to be a more rigorous statistical treatment as to how the triple-target configuration was determined. It seems that the authors tried various configurations and settled on the one with an appropriate chi-square and lowest residual. This is partially treated in Figure 5. However, this seems incomplete. It seems that a table to more clearly address the quantitative metrics used to make the configuration decision would be a more through treatment.

R03: To incorporate this point, a table has been embedded into the revised version (see Table A1 in appendix), showing the changes in the fit quality by varying the wavelength window and set of PFT targets. The respective explanation has been also given there. In addition, we also added now the testing of the different wavelength windows and target settings by spectral orthogonality survey. These results are discussed now in section 2.4.

C04: The use of various absorption spectra should be treated uniformly. The mix of culture and natural sample spectra is sloppy. There should also be verification with literature that their own spectra match what has been reported in published literature.

R04: Of course, the priority would have been utilizing natural samples for measuring all absorption

spectra. But we have been constrained to obtain an appropriate natural sample for coccolithophores where it was at least dominating the overall phytoplankton biomass by more than 50%. Nevertheless, even though the specific absorption spectrum of coccolithophores used in this study was obtained from an *E. huxleyi* culture sample, its spectral shape is very similar to the specific absorption of natural samples measured by *Siegel et al* (2007) in a coccolithophore bloom off the Namibia Coast (Benguela Upweling). This respective spectrum was also well comparable to the absorption spectra obtained from coccolithophore cultures and the coccolithophore-dominated natural samples in the Kattegat. However, the natural sample of *Siegel et al.* (2007), does not have detailed HPLC analysis to prove the domination via chl-a conc. of coccolithophores among the total phytoplankton biomass and its absorption spectrum. This paragraph was now introduced in the manuscript in section 2.3.
It must be noted that in the future, natural samples of a clear dominance of a certain PFT obtained from the major bio-geographical provinces (according to *Longhurst* 1998) will be used to establish a regional based PhytoDOAS.

Specific Comments (Reviewer #1):

C05: It is not clear why cyanobacteria are not considered in the triple-target PhytoDOAS approach. Please add discussion about this topic.

R05: Cyanobacteria are spectrally more different as compared to other PFTs incorporated in the multi-target fit-mode. Technically, putting cyanobacteria in the simultaneous fit does not help to reach the optimal fit quality and cyanobacteria have to be fitted in a different wavelength window to be retrieved optimal. In addition, the global distribution of cyanobacteria, as shown by instance by *Bracher et al.,* (2009) is also quite different from the other PFTs of interest, which is due to their specific oceanic habitats resulting from their specific biological growth conditions. To address this issue, we have added this explanation in sec. 2.4, where we talk in detail about the different considerations for choosing the right PFT set.

C06: (Introduction) – it needs to be explicitly called out that your use of the term phytoplankton functional types is in the context of taxonomic groups (dominant species). You clarify this on page 2286, line 2 but this needs to be stated in the introduction.

R06: Actually, by phytoplankton functional types we refer to the classification of phytoplankton based on their different biogeochemical functions (according to that *Nair et al.,* (2008)). Wew added the following sentence to the introduction for clarification:
"Even though some taxonomic groups (e.g., diatoms and coccolithophores) at the same time belong to different PFTs, the detection of the most important taxonomic groups is a necessary step towards the understanding of the PFT distribution in the global ocean (because PFT is a concept, while a taxonomic group is an entity)."

C07: (Section 2.1) – I think the description of DOAS can be substantially compressed with citation of the literature. It is not necessary to walk the reader through the DOAS equations if they are published elsewhere.

R07: The initial idea was to have a structure containing the main principal concepts and quantities of DOAS, which are essential for a reader to follow easier the modification exerted in PhytoDOAS.

However, to account for this point, we reduced the description part of DOAS consistently, which made the MS much shorter (about 2 pages).

C08: (Page 2280, line 1) – ".. absorption and scattering of CDOM" – CDOM is dissolved and therefore does not scatter!

R08: That was a sloppy mistake! It has been now corrected in rewriting (shortening) the whole section of the method description (sec. 2.1).

C09: (Section 2.2) – much of this has already been published by *Bracher et al.* (2009). This could be much more concise with citation to the literature.

R09: To incorporate this point, we shortened the description of SCIAMACHY by referring to *Bracher et al.* (2009) and *Bovensmann* et al (1999).

C10: (Page 2284, line 28 to page 2285 line 2) - The sentence starting with "The phytoplankton absorption spectra used in this study. . ." is confusing. Do you mean that *E. huxleyi* was from culture while the dinoflagellates were from a natural sample? Please rewrite to clarify.

R10: The whole sentences have been rewritten as follows:
"*As required by the PhytoDOAS triple-target fit, three phytoplankton absorption spectra were used in this study. The absorption spectrum of coccolithophores was acquired from an E. huxleyi (the dominant species of coccolithophores) culture. A dinoflagellate-dominated natural sample was used to obtain the respective absorption spectrum.*"

C11: (Section 2.3) – Why are you mixing cultures for one species with a natural sample for another? Are there absorption spectra for these species in the literature that could be used instead? If you do use a dinoflagellates natural sample, you need to identify what the predominant species is and include discussion about how variable dinoflagellates absorption spectra can be with various species of dinoflagellates.

R11: This has been already answered in R04.

R12: (Page 2285, line 25) – "right panel" should be "lower panel"

R12: Has been corrected.

C13: (Section 2.4) – It is unclear why cyanobacteria were not considered in the multi-target approach. Please explain!

R13: It has been already answered in R01. To address this issue, we have already added this explanation in sec. 2.4, where we talk about the considerations for choosing the PFT set.

C14: (Page 2286, line 9-10) – "multi-target fit" was previously defined, thus the definition does not need to be repeated here.

R14: This line has been modified, by removing the repeated definition:
"*Investigations proved that multi-target fitting leads to higher fit quality as compared to fitting only one PFT spectrum at the time. This approach results in significantly lower values for the absorption fit factors of each target, compared to the previous approach of the single-target fit.*"

C15: (Page 2287 lines 12-18) – You point out how important comparisons with in situ measurements are. However, I don't find your statements about limited availability of in situ data and collocation of in situ observations sufficient justification not to do this comparison. Validation with in situ observations needs to be included!

R15: This has already been answered in R01. We added part of this explanation into the MS to address this issue (sec. 2.4).

C16: (Page 2288, line 8) – "(as one criterion)" – This leaves the reader wondering what the other criteria are? Please be explicit and add further clarification.

R16: This has been corrected.

C17: (Section 3.1) – You only discuss the retrieval of *E. huxleyi* and dinoflagellates, but leave out diatoms. Please add text to justify your reasons for not considering diatoms in your retrieval discussion or add discussion of diatoms. You use March and October as your comparison months. Please add statements as to why these months were selected.

R17: In R01 it has been already explained why less focus has been made in this study on diatoms. However, at the end of the sec. 3.2 there is a paragraph describing the retrieval of diatoms via the PhytoDOAS *triple-target* mode.
We chose arbitrarily March and October (maybe motivated by that these months cover more or less equally the northern and southern hemisphere).

C18: (Page 2290, lines 4-8) – The sentence starting with "The PhytoDOAS results,"... This has previously been stated and does not need to be repeated.

R18: This was removed.

C19: (Page 2290, lines 8-12) – I find this type of comparison insufficient for validation. Each of these products (NOBM model output and PIC derived product) have their own associated uncertainty. Where the PhytoDOAS does not match well with either the NOBM model or the PIC product, we are left no knowing if PhytoDOAS is performing poorly or if the model or PIC products are erroneous. This is why comparison to in situ observations is imperative!

R19: This has been already discussed in **R01**:

C20: (Page 2291, lines 1-2) – You state, "However, the precise validity test should be done by converting the PIC concentrations into the concentration of living coccolithophore cells. . .". Then this needs to be done or do not include the comparison with the MODIS PIC products.

R20: In rewriting this section, this statement has been removed. Instead, the accuracy of MODIS PIC algorithm has been addressed according to *Balch et al* (2005).

C21: (Page 2291, line 7) – "other comparisons". . . which are? Please describe.

R21: It has been completely modified (see R01, R02, R20).

C22: (Page 2292, lines 5-17) – The ideas in this paragraph need to be proven. The current treatment is insufficient.

R22: The whole paragraph has been rewritten and respective papers have been referenced.

C23: (Page 2292, line 13) – ". . .demanding more investigation. . ." Yes! This investigation should be presented here.

R23: This is also a line with R20 and R22. The whole paragraph has been modified.

C24: (Page 2292, lines 23-24) – ". . .making them not being sufficient for quantitative comparison." Then why do it? So what does any of section 3.3 tell us? You leave the reader with a lot of doubt.

R24: Due to the ambiguity mentioned in this comment, the sentence has been removed. However, the section 3.3 is showing that the PhytoDOAS method is not only functioning for long-term monitoring of coccolithophores, but also is appropriate for detecting short period events, as temporary blooms (with duration of few days). On the other hand, due to high reflectance of coccolithophore-rich surface water surfaces, coccolithophore blooms can be identified clearly as RGB satellite images (not quantitative). These types of satellite images have been used as preliminary evidences that PhytoDOAS is able to detect coccolithophore blooms.

C25: (Page 2295, lines 5-7) – You state, "The global distribution of dinoflagellates retrieved by PhytoDOAS must be compared with an appropriate data set of this taxonomic group." This needs to happen before this manuscript is accepted.

R25: As mentioned in R01, we already changed the title of the manuscript to show more clearly that the focusing target of this retrieval has been the group of coccolithophores and also changes in the text were made accordingly, which state that so far this PhytoDOAS retrieval configuration proved to function mainly for coccolithophores. Therefore, the point mentioned in this comment has been already covered in R01.

C26: (Figure 2 caption) – You state, “The first two spectra were obtained from cultures. . .” The first two listed in the previous sentence are *E. huxleyi* and dinoflagellate. However, in the text you said that the dinoflagellate spectra was taken from a natural sample. Which it is?

R26: It has been corrected.

C27: (Figure 4 caption) – The caption states “scaled to 0.1”. However the legend indicates “scaled by 0.75”. Which is it?

R27: It has been corrected.

C28: (Figure 5 caption) – “(1.07.2005)” Please write out the date. This notation could be confused by some.

R28: It has been corrected.

C29: (Figure 11) – There needs to color-bars associated with the middle and lower panels.

R29: This has been already done.