We would like to thank all the reviewers for their constructive comments which helped improve the manuscript. Our replies to comments are covered below.

General comments:

This manuscript deals with water optical properties which are acquired by remote sensing, and which are then used as input to 1-D lake modeling. This approach is important and needed addition to current efforts of incorporating lakes and reservoirs to weather prediction models. Water clarity is an important factor in defining lake heat budget and thermal stratification and thus is a significant parameter for processes in the air-water interface. With millions lakes of different sizes around the world, comprehensive direct measurements of water clarity are not possible, highlighting the need for indirect estimates of water optical properties. Satellite measurements show great promise in mapping light extinction coefficient, a parameter to define water clarity, and to my knowledge this study is the first one to incorporate lake modeling to water clarity defined from satellite observations. This is an important study with wide interest in scientist from different fields and therefore I find this study appropriate for HESS. However, there are a few major points which prevents me from recommending this manuscript for publication as it is.

The major topic of this manuscript is that satellite-derived light extinction coefficient, Kd, represents well the in situ measurements of Kd, that it can be used as an input to lake modeling, and that it enhances the performance of FLake as compared to the current approach of using constant Kd of 0.2 m-1. These points should be emphasized. Currently, the manuscript seems unbalanced, with much of the focus given to topics not strictly the main theme of this manuscript. Also some restructuring is needed so that the reader does not get distracted from the main focus. In general, the manuscript would benefit from reducing the amount of figures.

Since this is the first approach in combining satellite-derived Kd to lake modeling, it would be of value to describe the strengths and the weaknesses of this approach. How easy and accurate method this is for the modeling community in general; can this be used without in situ measurements of Kd or should there always be e.g. Secchi disk measurements for validation; what are the next steps needed for applying this method in broader context.

Thanks for your comment. Strength: Integrating lake specific Kd values can improve the performance of 1-D lake models. However, field measurements of Kd are not widely available. This study demonstrates that satellite observations are a reliable data source to provide lake models with global estimates of Kd with high spatial and temporal resolutions.

The globally available CC product can be easily used as a source to fill the gaps in Kd in situ observations, and improve the performance of parameterization schemes and, as a result, further improve the NWP and climate models.

Weakness: in situ data of SDD or water clarity are always required for the method development, calibration, and validation.

Next steps: investigate the potential of Sentinel-3 to provide lake modeling community with the water clarity information. Also study the resulted improvement in the performance of NWP and climate models.

These points are mentioned in page 14, lines 5-11.

Specific comments

In regards of restructuring, I will give here an example how the figures (and related discussion) could be rearranged. After the map, I suggest to first show satellite-derived Kd at the site of FLake modeling (current Fig. 8), which is the main input parameter under focus here. For the estimated solar irradiance and incoming long-wave radiation (Figs.2 and 3), the figures add only little to the reported statistics and thus these two figures could be removed. After showing the satellite-derived Kd, it would be logical to show their validation (current Fig. 4). Then the results of models against measurements, i.e. current Figs. 9 and 10. Current Figs. 11, 12, 13 and 14 basically show the same data in different forms. I suggest either to combine Figs. 11 and 12 and show only this, or show only Fig 13. Lastly, current Figs. 5 and 6 could be shown, which would then lead to the discussion of the strengths of satellite-derived Kd and possible future studies (see also the related comment later). This is only a suggestion for restructuring, it could also be done otherwise.

Thanks for the suggestion. The authors, however, would like to keep the current format. This study is based on using satellite-derived shortwave radiation. Longwave radiation is also estimated. Therefore these two parameters need to be evaluated (Figures 2 and 3). Other input of FLake model, the most important one, is water clarity which is also derived from satellite observations. Before any further discussing the potential of combining these observations into the model, first the evaluation was conducted (Figure 4). After that, Figures 5-7 show the extend of spatial and temporal variations of water clarity for the entire water body. Finally, Figure 8 shows how water clarity has changed in the station of interest (NDBC) during the study period. Based on the variations (min, max, average values of Kd at this station), different simulations were designed.

Figure 9 shows how observations and different simulations compare during the study period. The figure demonstrate if there is any specific timing that the difference between simulations and observations is more highlighted; whereas Figure 10 is for statistical evaluation purpose.

Figure 11-14 might have overlaps in the demonstrated information; however, they are presented in this study for different purposes. Figure 11 show LSWT and MWCT; Figure 12: MLD; Figure 13: timing, depth, and temperature of different thermal layers (epilimnion, thermocline) also temperature of MLD; Figure 14: average temperature and depth of each thermal layer.

Page 4, line 9. Air humidity, which is used as FLake input, is taken from land 10 km away from the lake site. Air humidity is important for modelled latent heat fluxes. Could the authors briefly state their opinion how well the measured air humidity represent that over lake, and how or if this affects the modeled results.

As the reviewer mentioned, the humidity is important for modelled latent hear flux and would be different over land and over lake. Since warm air hold more moisture than cold air, the percentage of humidity must change with change in air temperature. We expect that the humidity decreases as temperature increases. Large temperature and humidly differences can lead to a large sensible and latent heat fluxes. Water is a good absorber of the energy but the land absorb much faster the energy from the sun. Water heats up much slowly than land and therefore, the air above land will have higher temperature and therefore less humidity. On the other hand, lack of in-situ data over lakes and the distance of the stations from the shoreline (less than 81km in our case, 10 km was recognized as a mistake in the manuscript) is one of the main limitation of lake studies. In this study all the model forcing comes from the station on land such as air temperature and humidity, therefore in this way the rate of differences between air temperature and humidity kept constant.

Page 4, lines 19-21. Sentences starting with "Available Secchi disk" and "SDD data was" are repetitive and should be merged to one sentence. Also, it could be mentioned here that the SDD data comes from Limnos cruises.

Thanks for your comment. Page 4 line 20 ("Available Secchi disk depth (SDD) field measurements were used to estimate lake water clarity.") has been removed.

Page 6, lines 1-2. Please clarify this sentence. Does this basically mean that only the pixels which were not rejected according to the criteria in Table 1 were used?

Yes.

Page 7, Chapter 3.1.1. A lot of space is dedicated for this, and therefore a justification could be given in the first sentence. E.g. "Validating the satellite-derived Kd with in situ observations is important because. . ." And in the end of the chapter the outcome of the evaluation, e.g. "For these reasons, we deem the satellite-derived Kd correct and thus were confident in using them in the modelling." Also, in Chapter 3.1.1. or later, the authors could discuss whether this kind of validation is always needed with satellite observations, what are the implications, etc.

In the first sentence of this section, line 9, the main focus of that section is mentioned: the reliability of satellite-derived Kd values is highly dependent on comparison of them with independent in situ SDD measurements. So in situ observations are always required for validation of satellite-derived data.

End of this section is closed with the reason and motivation of using satellite-based water clarity measurements, when in situ SDD data are not always describing Kd values (only small values of Kd are described using SDD).

Page 8, Chapter 3.1.2. This chapter seems interesting but out of place. These results are not further elaborated, and therefore I suggest to move them to the end of the Discussion. This way the authors could show what benefits remote sensing of Kd would bring (spatial and temporal variability, which is not achieved well with manual sampling; perhaps good input for 2D and 3D modeling), which would lead to the discussion of possible next studies. This way also the key input parameter, Kd at the NDBC station, would be shown earlier.

Thanks for the suggestion. This section described how Kd value changes spatially and temporally over the full lake; and it ends with the variations of Kd in the location of NDBC station. This is to demonstrate how variant Kd value could be over the lake and over a period of time, demonstrating the shortage of in situ observations to cover these changes temporally and spatially and highlighting the motivation of using remote sensing observations to overcome these concerns. After highlighting this important role that remote sensing observations can play in coupling with lake models, the study continued showing the results of integrating satellite-derived water clarity with FLake. Therefore the authors would prefer to keep the current format of the manuscript.

Page 9, Chapter 3.2.1. If the satellite-derived Kd has been validated sufficiently well and it produces better simulations, what would be needed for the simulations to match the measured LSWT more accurately? This could lead to suggestions for future research.

Thanks for the suggestion. Section 3.2.1 discuss the improvement of modeling using the satellite-derived Kd values. The next section study the sensitivity of FLake to the variations of Kd, and if it is necessary to consider the temporal variations (monthly basis) of Kd in simulation or a constant-lake specific value is sufficient in the modeling for Lake Erie.

Therefore if this comment is suggesting to consider the temporal variations of Kd in simulations, this has been already considered and tested for the range of Kd values in Lake Erie.

Page 9, Paragraph starting ‘Fig. 9 shows the results. . .’. I suggest to first describe the observed behavior in the temperatures and then discuss how the modelled behaviors compare to these.

The temperature changes in three years, 2005-2007, have a normal fluctuation, increasing from spring to summer and decreasing toward winter. Therefore, the authors did not find it necessary to add to the manuscript.

Page 10, lines 17-18. This is quite strong statement and probably not true for all lakes. Lakes are very heterogeneous, be more specific which type of lakes is meant here.

These sentences are for explaining why results of two simulations of Avg and Merged are comparable, while Avg simulation are producing lower MBE. The statement starts with “it is possible”. Therefore it is only a potential reason for such results in Lake Erie, and not a generalized rule for all lakes. However, modification to the sentence has been applied to clearly demonstrate the point.

Page 10, Chapter 3.2.2. This chapter needs the most restructuring. E.g. the paragraph on page 11, lines 25-28, could be removed. The two first sentences are basic limnological knowledge and the last sentence does not really lead the story further. In this chapter, the theme of light penetration and absorption is discussed in many places, e.g. on page 11, lines 8-9, lines 11-12 and lines 29-34, page 12, lines 11-13 and lines 19-20. Remove excess repetition. The last paragraph on page 12 (starting Fig. 14 depicts. . .) repeats what is said earlier and is not the main focus of this study, therefore I suggest to remove that paragraph. The last paragraph of Chapter 3.2.2. discusses about modeled ice cover. This seems a bit out of scope and there really seems to be no ice measurements against which to validate modeling. For this reason, I suggest to either remove this paragraph or significantly shorten it.

Thanks for your comment. Although in situ measurements are not available; however, the sensitivity of FLake to Kd variations to reproduce ice phenology and thickness is investigated and is one of the scopes of this paper.

The repetition has been removed in the new version of manuscript.

Page 11, lines 11-12. The authors seem to mix two concepts here. Darker water color is related to dissolved substances, such as colored dissolved organic matter, not to particulate matter.

Attenuation of light in dark waters is high; and this could be because of the existence of dissolved (absorption) or suspended matters (scattering).

Page 11, lines 13-14. The authors over-simplify the underlying mechanisms for LSWT behavior. The loss of energy to the atmosphere is related to the surface water temperature (and wind), not only in fall but throughout the open-water season. However, the mechanism how mixed layer depth affects the rate of heat loss needs more explaining.

The mixed layer depth (MLD) affects the speed of losing energy to the atmosphere throughout the year. But in these sentences the reason of faster loss of energy in fall is explained. This is because MLD in fall is shallower than in spring, therefor loss of energy is also faster.

Considering MLD to explain the reason is combining the effect of both temperature and wind.

Page 11, lines 18-23. Tie these results from the literature more tightly to the findings in this study, e.g. by writing whether this study supports or opposes previous findings. Also, the sentence on lines 21-23 (starting 'Heiskanen et al. . .') could be removed either from here or from the Summary.

In line 18, the result of our study is that the sensitivity of the model increases from Min to CRCM-12.6 simulation (Kd decreasing from 0.58 to 0.2). The statements after this lines (lines 19-23) are discussing other studies which support the finding of our study. This finding is that FLake is more sensitive to Kd values less than 0.5.

Page 12, paragraph starting with 'Fig. 12 shows'. Here full mixing is described in very atypical way on several occasions, e.g. by 'highest depth of mixing' and 'reaches maximum MLD'. I suggest to describe these occasions either by discussing of overturn, of full mixing or similar.

Lake turnover is the process of lake's water turning over from top to the bottom, which is full mixing. The maximum/highest possible depth of mixing at NDBC station is 12 meter, so when MLD reaches this depth turnover happens. This is the reason for describing turnovers using terms such as 'highest depth of mixing' and 'reaches maximum MLD'.

Page 12, lines 8-9. If this is the reason for earlier overturn in simulations with clearer water, how the authors then explain the results shown in Figs. 11 and 13 where it is evident that there is full mixing in the beginning of September in CRCM-12.6 simulation with temperatures of about 20 deg C? Fig. 14 also shows that the clearer the water, the higher the water temperatures in Oct and Nov. Note that in addition to convection, mixing is related to wind forcing and density gradient in the water column.

CRCM-12.6 has the clearest water compared to other simulations, therefore the water column reaches the same temperature in its layers earlier than other simulations, leading to earlier turnover. Figures 11 and 13 and 14 support this statement.

Page 12, lines 16-17. MLD is not influenced by the thermal structure, but it is part of the thermal structure. I would remove this sentence.

The sentence has been removed.

Page 12, lines 13-14. Fig. 13 is essentially the same data as in Fig. 12 and therefore one cannot be used to confirm the results of the other.

"confirms" has been changed to "also represents".

Page 12, lines 20-22. Deepening of the thermocline is related e.g. to wind forcing and thus it cannot be suggested that thermocline deepening in clear waters is monotonic. Also, it is not clear what is meant with 'stabilize the temperatures'. I suggest to remove this sentence.

Figure 13 shows that in the simulation related to the most clear waters (CRCM-12.6), deepening of thermocline is faster, with a monotonic speed, as opposed to the dark waters.

Convection that transfer heat between layers continues until temperature is fixed (stabilized) in all layers.

Technical corrections

Page 6, line 23. The same result 'was' found for. . .

It has been corrected

Page 8, line 32. 'leads to higher water clarity'. The authors must mean lower water clarity.

Thanks for catching this mistake. It has been corrected

Page 9, line 18. The sentence starting 'The Kd values' and the sentence after that could be merged and rephrased. E.g. 'The monthly-averaged Kd were used to simulate the surface water temperature and produce a merged LSWT (Merged)."

The authors would prefer to keep the sentences separate. This is to more emphasize on the steps taken to produce the merged product. First simulations were run using the monthly averages, and then a merged LSWT is produced.

Page 9, line 20. Comparing LSWT in situ observations (Obs) with. . .

It has been added.

Page 9, line 21. How can the authors compare measured and modelled surface temperatures in April when there seems to be little or no measured LSWT during April, at least according to Fig 9?

Observations for 2006 and 2007 starts from 19 and 18 April, respectively.

Page 10, lines 2-3. Rephrase. Do the authors mean that the annual average of Kd can occasionally be closer to the actual Kd than the monthly-averaged Kd? This same topic is also mentioned in lines 16-17, and at least to me it is unclear how yearly average value (i.e. one single number) can represent the extent of Kd variations (i.e. how big is the range).

Yes. Satellite images are not always available (due to cloud cover or shortage of temporal resolution) to cover the actual variations of Kd in the station on the lake. Having a yearly average has a higher chance of capturing potential variations in Kd value and calculate the average of them; therefore is a better and closer representative of the actual Kd value. The statement has been rephrased.

Page 10, line 10. Please clarify what is specifically meant with 'are as affected'.

It means that no matter which depth we use, the actual depth at station or a tile depth, the large under-prediction is happening for these two simulations of CRCM-12.6 and CRCM-20 (MBE for both is above 1ºC); especially for temperatures above 12 ºC.

Page 10, lines 11-12. 'This can be explained by. . .'. Be more specific in telling how lake depth explains this.

CRCM-12.6 and CRCM-20 only differ in the depth used as input in the simulations. Therefore, if CRCM-20 has the most under-prediction compared to all other simulations (including CRCM-12.6), it is related to the input depth. Clarification has been added to the manuscript.

Page 10, lines 12-13. This should be self-evident if the model is any good, and therefore I suggest to remove this sentence.

The authors would prefer to keep the statement to emphasize on this and other studies results.

Page 11, line 9. '. . . causing thinner mixing depth (Fig. 12)'

It has been corrected.

Page 11, line 35. Change 'when Kd changes. . .' e.g. to 'when maximum (minimum) Kd is used instead of its average value. . .

It has been corrected.

Page 12, line 1. Similar comment as previous. This is a bit misleading wording since it gives the idea that Kd changes naturally, whereas what is meant that different Kd is used as an input.

It has been corrected.

Page 13, line 27. Write open the abbreviation 'LSWT' here. It is not typical abbreviation and not clear for those who only read Summary and conclusions.

It has been corrected.

Page 14, line 3. Change 'has' to 'have'.

It has been corrected.

Comments to figures

Fig. 1. It would be of interest to see the main river inlets and outlets. This way it would be easier to assess how much river inflow possibly affects modeling results.

It has been added to the figure.

Fig. 5. Remove 'Lake Erie boundary' from the legend, it is not needed. Also make the color bar much larger. Same comments for other similar figures.

It has been corrected.

Fig. 8. It would be of interest to see the SDD at this location (or from the nearest location where those exist) together with these CC-derived Kd. These could be marked to the same graph with secondary y-axis.

There is no SDD in situ measurements at NDBC station. According to Fig. 1, the nearest locations with SDD observations are within about 20 km distance from NDBC station. However, water optical properties changes in spatial scale smaller than this distance. Therefore showing SDD values for those station are not a good approximate for SDD at NDBC station.

Fig. 10. It would be interesting to see the performance for each year separately. This could shown by plotting each year with different color. Also, it is more standard to show these kind of scatter plots as box plots (both axes of same length).

The performance of each year separately is shown in Table 3.

Fig. 11. The measured LSWT should be shown. Otherwise, it is impossible to say which simulation performs the best. Use a) and b) for these two graphs. Also in the legend, the Kd values could be shown for each model run.

(a) and b) has been used in the new manuscript. Kd values has been shown in the legend of new manuscript.

However, because this figure is related to sensitivity analysis, there is no need to show the observations. Section 3.2.1, which is more related to the accuracy assessment and improvement of simulation results, has shown observations.

Fig. 15. Model run CRCM-12.6. is not visible. If the resolution can not be increased, describe in the caption where the line is.

Description of the figure was given in the body of manuscript (page 13 lines 33-4), however it has been also added in the caption.