Dear Editor, Dear Reviewers,

Thank you for taking the time again to review our revised manuscript. We are very happy that our efforts have been viewed favourably. We have addressed all the minor comments arising from this second round of reviews below. The original comments are numbered (e.g. R1C1 – Reviewer 1, Comment 1 and R2C1 – Reviewer 2, Comment 1) and shown in black italic text. Our response is shown in blue normal text. We have uploaded two versions of the updated manuscript, one marked up version which shows the changes and one clean version.

Reviewer 1:

R1C1: a very minor point: perhaps it would be more natural to put the comment about the white rectangle in Fig.4 (missing observations) into the Figure caption, rather than in the main text (lines 402-403).

We have done as suggested by the reviewer and moved this comment about the missing observations to the figure caption. See lines 420 – 421.

R1C2: There is a lot to unpack in the 2nd comment from Reviewer 1. Before doing so, I would like to mention that a 2nd paper by Cahill is in the final stages of preparation, which investigates in more detail some of the issues raised in this comment. The title of this 2nd paper is "Bio-optical and bio-physical characteristics of two marine heat wave events in the Arkona Sea in 2018". I would like to carefully delineate between this paper under revision and the work in the 2nd paper.

I have few comments on the validation, in particular on Fig.5, Fig.6 and Table 2: Firstly, why did you decide to do validation only using data from two particular days (29-30/05/2018), is this because of the full satellite coverage on those days? Or because of the bloom on part of the domain?

We chose these days for the validation for a few reasons. First, as the reviewer suggests, i.e. there was full satellite coverage on both of those days. In addition, these days coincide with the May peak OSC-induced heating rates we found in our modelled results which we could link to a bloom event in the model results and a real event which could be seen in the satellite data. The magnitude of our modelled bloom event was lower (c. 8 mg m^{-3}) than that seen in the satellite (c. 30 mg m^{-3}).

How did you exactly determine/define the bloom?

We define a bloom event in our model results following a threshold approach as used by Thomalla et al. (2011) and applied by Lorenz (2019) whereby the initiation of the bloom is understood to be the period of the year which registers a relative increase in chlorophyll concentration, irrelevant of the actual value. The threshold is defined as:

$$CHL_S = CHL_{MEDIAN} * 0.1 * CHL_{MEDIAN}$$

Thus, the bloom start condition is:

$$CHL_t < CHL_S \; \& \; CHL_{t+1} > CHL_S$$

And, the bloom end condition is:

$$CHL_t > CHL_S \; \& \; CHL_{t+1} < CHL_S$$

Lorenz, H., 2019. Marine heatwaves and the phytoplankton spring bloom: a comparative survey of marine heatwaves, spring bloom events and their relationship in the North Sea and Baltic Sea. Master Thesis, University of Potsdam, 86pp.

Thomalla, S. J., Fauchereau, N., Swart, S. and Monteiro, P.M.S., 2011. Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. Biogeosciences, 8, 2849-2866, doi:10.5194/bg-8-2849-2011.

The modelled and observed May bloom event is analysed and discussed in more detail in the 2nd paper.

Wouldn't it be more natural to calculate, for each location in the image, the RMSE (comparing satellite with model) across a longer time-period (matching the model with observations at each location whenever the observations were available) and plot the spatial distribution of RMSE? I know that potentially the different data availability at different locations can skew the RMSE plot, but I would think it would be still more informative than comparison from two selected days (even more, when the days follow each other, so the tracer distributions are due to relatively slow dynamics very similar)?

I understand the reviewer's point but would like to suggest a compromise. We have added statistics (RMSE, bias) to our evaluation and to support statements related to Figure 7 (as per R2C4 – see below). Moreover, we have taken on board the Reviewer's next comment about Figures 5 and 6, and now show just one Figure with the modelled mean values (over two days) for the 4 variables and the related RMSE. We also have changed the colorscale to (hopefully) something more intuitive. Some minor changes to lines 432 to 436 in the marked up version of the manuscript were made to reflect the change to Figure 5. Subsequent figures and related text have been re-numbered.

In the 2nd paper (in prep.) mentioned above, we evaluate our model results over different time periods (May and July 2018) and spatial structures using a combination of satellite data and various dynamic diagnostics (e.g. Rossby number, divergence and coherent structures). We would like to reserve this more detailed analysis for the 2nd paper.

The Table 2 (if you think it is needed) could show values from the RMSE (and maybe some other metrics, such as bias, R) at the specific selected 4 locations. Moreover, the Fig.5 and Fig.6 as they stand is hard to compare due to different ranges in the colorbars. The choice of colorscale is also a little unfortunate as the dark blue and dark green colors at the opposite ends of the colorbar are similar (also I would perhaps expect the colors to change from dark to light, with dark-light on the opposite ends of the colorbar). It seems to me that the chosen colorscale is more suitable for displaying symmetrically spaced values around zero, so the white color runs through the zero value. I leave it to the authors to decide, perhaps I would present a single Figure (instead of two Fig.5-6) plotting spatial distributions of modelled mean values for the 4 variables on the left side and the RMSE on the right side, with perhaps simpler colorbar (like a white-red scale)?

Figure 5 and 6 have been replaced with one Figure showing modelled mean values for the 4 variables and the RMSE. A different colorscale has also been applied. Table 2 has been revised to show mean OLCI and model values over the two days and the bias at the 4 locations.

Reviewer 2:

R2C1: Line 67: after "e.g." a space is required;

Done

R2C2: Line 235: please consider rephrasing as following "where the term" instead of "where the first term", since this is the only term on the right hand side;

Done

R2C3: Line 423: "(Huang et al., 2021)" instead of "(Huang et al., 2021))"

Done

R2C4: Lines 461-464: it is worth supporting the statements by providing statistics;

R^2 and RMSE statistics have been added to support these statements – see lines 468 to 472 in marked up version of the manuscript.

R2C5: Figures 8, 10: please re-plot with the wavelength increasing from 400 to -> 700 from left to -> right.

Done