The authors use BGC-Argo data to compare IOPs (absorption and scattering) and AOPs (Kd and Rrs) estimated using simple water column radiative transfer, in situ measurements of water column constituents, and satellite remote sensing. The focus of this project is the Mediterranean, where the optically complex waters justify comparing different approaches. This paper directly compares different estimates of absorption spectra, and in many cases shows the utility both of more recent estimates of a_w, as well as the benefit of a regional PFT approach for pigment absorption.

While this research is definitely useful, in many places I struggled to understand what exactly the authors were doing. In particular, in Section 2.3 they list 6 different models that they test, comparing measured downwelling irradiance with that derived from different models following Eq. 1-12. Their first model is pure water absorption and scattering, for which they test three different estimates of a_w (Fig. 3). It wasn’t clear to me what downwelling irradiance observations they used to test the different a_w. From the text, it seems like they compared actual profiles of Ed to estimated profiles of Ed assuming clear water, but that doesn’t make any sense. Did they only use BGC-Argo profiles in very clear waters, and if so, how did they define that? Or did they have a “basic” version of the other IOPs that they used in Eq. 1-12, in which case, what was it, and how does it differ from experiment #6? I have a similar question for experiments 2-5; from the text it seems like the authors assume the only constituents are those specifically tested, but that doesn’t make sense unless you have measurement profiles for which that might be true. I’m sure I am missing something here.

Another point of confusion was how the authors use estimates of Chl from BGC-Argo. BGC-Argo floats do not measure Chl, they measure fluorescence from Chl (fChl). During the daytime, near the surface fChl can be reduced even though Chl is still elevated because of non-photochemical quenching (NPQ). There are a number of methods to correct for this, including specifically for BGC-Argo (e.g., Xing et al., 2018). Was any sort of NPQ correction applied to the fChl data? If not, I would be worried that the authors’ “Case I models” are tracking fChl and not Chl.

Finally, I was uncertain about how to interpret Figures 8-10 in light of the text on l. 254-256 and 332. Are the satellite values shown monthly climatological values? If so, I don’t see how these can be compared with Argo considering the time frames and spatial...
locations are different (that is, Argo doesn’t sample uniformly and because of cloud cover neither does the satellite, and they also have different temporal weights based on the different lifetimes of the satellites and Argo floats).

Figures in general: I recommend removing Figures 1 and the top panel of 2 (see comment below), and moving the information in Figures 3-6 into a single Table, after removing experiments that don’t make sense to include (see comments below). What would be more useful to see in a Figure would be some representative profiles of measured Chl, bbp, Ed, and modeled Ed. Instead of Figure 7, what would be more useful to see would be the regionality or where the measurements are close to or far away from the 1:1 line with the model.

Conclusions Section in general: A lot the text in the Conclusions does not logically follow from the results of this paper. For example, I don’t see a clear path between this paper and combining oxygen, nitrate, and pH in numerical models (l. 412-417), and I’m not convinced that this paper in particular demonstrates the need for inclusion of multi-spectral measurements (l. 423-424) or hyperspectral models (425-427). To be clear, I don’t necessarily disagree with any of these statements, I just don’t think these are conclusions or logical next steps that one would arrive at from reading this paper.

Smaller comments:

I don’t think it is useful to show how many BGC-Argo profiles didn’t have the right set of measurements at the right depths to do this analysis. If profiles were discarded because of data quality (rather than data availability) that could be useful to know – e.g., “a condition of less than 30% difference between modelled and computed Ed values was thus added which resulted in 147 profiles less” (l. 77-78), although I don’t quite understand what this means – but right now Figure 1 and the top panel of Figure 2 don’t add anything to the paper.

Sec. 2.2 could use more textual help for the equations – make sure to define variables in the same paragraph where they first appear and provide text to explain the equations. Also, when denoting variables associated with direct downward irradiance sometimes a subscript of “dir” is used (e.g., E_dir) and sometimes “d” (e.g., C_d); please pick one. Similarly for “dif” and “s”. Please also check in this section and throughout to make sure b and b_p mean scattering, and b_b and b_bp mean backscattering, and that variable dependencies are correct (e.g., a_w doesn’t depend on z).

l. 125: Define PFT.

The text around Eq. 13 is confusing; I don’t think this equation is actually helpful. The authors can just say in the text in Sections 2.3.3 and 2.3.4 that profiles either follow Chl, b_bp, or fDOM.

l. 141-142: I am confused by the statement that AOPs “to a certain extent remove the impact of the external environment’s variability”; by definition, AOPs are properties that depend on these external parameters.

Sec. 2.3: I found some of this section backwards – the text in l. 194-195 would be nice to have before the authors explain the text around Eq. 16. Also, considering the authors have chosen to use Mason et al. 2016 instead of Pope and Fry (1997), why show the test with the Pope and Fry-derived values at all? Similarly, in Fig. 5, why show the values without the T/S correction for a_w?
I. 229-230: If the estimated variability in this backscattering spectral power-law slope parameter is from 0-4, using a constant value of 2 seems overly simplistic. Some discussion about error and uncertainty as a result of this would be good to include.

I. 233: What final value of the backscattering ratio was used? What error and uncertainty is the result?

I. 336-337: The authors could be more explicit about what the two Kd estimates actually are here (both are float-derived). There should also be more discussion somewhere – why might IOPs be overestimated?

Need a statement about data availability somewhere (typically in Acknowledgements). Also, should specifically list which BGC-Argo floats were used.

Minor edits (not exhaustive):

I. 157-158: I think it is more correct to say that the absorption spectrum is often modeled (not “follows”) as an exponentially decreasing shape despite its heterogeneous biogeochemical composition.

I. 173: “expressed” should be “parameterized”? Similarly, in I. 225 “obtained” should be “estimated”.

I. 201-202: Organelli et al. (2017c) doesn’t provide spectra for wavelengths shorter than 400 nm; how do you get data at 380 nm? Also, how representative are these of Mediterranean waters? And finally, some explicit response should be given to that paper’s caveat that their results “were not intended for any algorithm development and/or validation” (to quote from their Conclusions).

I. 213-214: How many data points have Chl outside of these bounds? Would it not be preferable to just omit those profiles? If the authors keep them in, some discussion should be present about what error this introduces.

Eq. 21: negative sign missing in the exponent?

Fig. 5: Is it still considered Case I if the CDOM profile follows fDOM and not Chl?

I. 368: It should be noted that satellite-derived Rrs(412) is prone to large uncertainty, especially in optically complex waters; see e.g. Wei et al. (2020) and references therein.

Fig. 7: I don’t think the word BIOPTIMOD was anywhere in the main text?