Response to Interactive comment on "Performance evaluation of ocean color satellite models for deriving accurate chlorophyll estimates in the Gulf of Saint Lawrence

In black our reply to the reviewer #1

Reviewer
The Authors propose that their evaluations can be used to make statements about atmospheric and in-water algorithms independently. Given that no radiometric satellite-to-in situ comparisons are shown, I'm not sure this is truly the case.
In terms of water contributions to radiance, we don't have in situ measurements of Lw or Rrs to compare with AC results, however we can understand the impact of different AC on aph in relative terms by using the same inversion model
Also, we are not validating any sensor, thus our results about using different AC and ocean color inversions are always referred to already validated TOA.

Reviewer
but broader comments about the quality of the AC and the in-water algorithms themselves are unjustified
we agree with the reviewer and we don't talk anymore in terms of 'quality' of the algorithm or model. Our intention was to talk about the most accurate approach

Reviewer
The manuscript would be greatly served by teasing apart the components of the AC and in-water algorithms that succeed and fail in the GSL. Without doing so, the results are neither portable to other regions or to updates to the AC and in-water algorithms (both of which are often improved routinely).
Our intention is to evaluate the performance of each algorithm in our study area. Conversely, we do not try dissecting each technique in order to improve each method and become available to be applied in other environments. This out of the scope of the present contribution.

Reviewer

The Authors show a general misunderstanding of atmospheric correction. The most critical component of AC is the selection of aerosol models and magnitudes. The standard NASA approach still uses Gordon and Wang 1994 (and Kuchinke et al. 2009 uses a modified version of Gordon and Wang 1994).
We are aware about the importance of aerosols for obtaining accurate Lw estimates. We clarified the origin of SA and KU following reviewer's suggestion

Bailey et al. 2010 is not an atmospheric correction method, nor is Stumpf et al. 2003 (see e.g., the text on page 4, lines 1-11 and the first paragraph of Section 2.4). These are simply mechanisms to account for non-zero nLw(NIR)
Now is corrected and we say:

The SA approach is based on the atmospheric correction scheme originally proposed by Gordon and Wang (1994) and modified afterward by Bailey et al. (2010) and Stumpf et al. (2003) to account for non-zero nL_w(NIR) values

Bailey et al. 2010 does not "constrain atmospheric parameters such as aerosol concentration and type" (page 4, line 11).
Bailey et al. (2010) reference was deleted

The Authors repeatedly suggest that NASA has a "default a*ph curve" (see, e.g., page 10, line 3 and Section 3.4.1). This is not the case – there is no default NASA a*ph spectrum.
Sorry about the confusion, we thought that GSM curve for aph* was the SeaDAS default. Now we eliminate NASA and we talk about constant aph* and GSM default that correspond to GSM curve.

Page 5, Line 4: "t_a-w/n_w^2 approximates to 1" is not true. If, for example, t_a-w = 1 and n_w = 1.334, then t_a-w/nw^2 = 0.56

That was a typo. Now is corrected, we referred to ta-w and not to ta-w/nw^2

Section 2.5: The description of GSM is inaccurate.
We improved according to reviewer's suggestions
Also we included the definition of aCDM (CDOM + non-algal particles)

Correct the reference "Levender et al. 2005" to "Lavender et al. 2005".
Done

<u>Other comments</u>
Page 3, Line 3: Reword first sentence. The "concentration" is not a pigment.
done
Page 3, Line 25: The mention of both 443 and 440 nm in the parentheses is confusing.
We only centered band at 443 nm
Page 4, Lines 12-21: Most of this is true, but the paragraph is not easily read. Suggest adding a supporting figure to demonstrate the "spectral overlap between constituents".
No need for a figure, but the sentence was rewritten for easy reading as: "increases, the overlap between optical signatures from different components becomes more severe. Thus, a major overestimation (e.g., mineral-rich waters, Balch et al., 1989; Wozniak and Stramski, 2004) or underestimation (e.g., CDOM-dominated waters, Montes-Hugo et al., 2005) on satellite-derived chl values occur.

Page 5, Lines 22-24: I would argue that Bricaud et al. 1995, 1998, 2004 can be modestly used to estimate a*ph from remotely-derived Chl.
Sorry, but we do not talk about a*aph estimates from chl as used by Bricaud in page 5 line 22-24 .

Page 6, line 27: Change “GLS” to “GSL”.
done
Page 9, line 19: Were sums of pigments used to calculate a total Chl, such as MV + DV + allomers + epimers of Chl?
We used total chlorophyll a (Chl a + epimers and allomers of Chl a + chlorophyllide a) as computed in Roy et al. 2008

Section 2.3: What version of SeaDAS was used?
Version 7
Does “6 h of the satellite overpass”
That means ± 6 h
Were any flags and masks applied?
Yes, we added, **‘Several masks (clouds, glint and land) and flags (brdf_opt = 7, glint_opt = 1, outband_opt = 2, pol_opt = 0, rad_opt = 0) were applied during the processing of each L1 satellite scene’**.

Section 3.1: What statistic is being shown in Fig. 2 and how does that statistic relate to those proposed in Eqs. 6-8?
The error bars correspond to 2 standard errors of satellite data computed within for 3x3 pixels. We clarified that in the legend Fig. 2, Error bars symbolize two standard errors as computed from a 3x3 mask of pixels

Page 15, line 8: It’s been shown many times that spectral matching algorithms such as GSM suffer from fewer retrievals than empirical approaches and QAA. Recommend the Authors comment on this or at least mention how their results fit within a broader community context.
We added text and references supporting these results
We say:
“A consistent pattern during the two surveys was the greater accuracy of KU-QAA with respect to KU-GSM calculations of a_{ph}(443). This was attributed to two main differences between inversion models in terms of a_{ph}(443) calculation and type of AOPs used to carry out the optical inversion…”
“Second, GSM retrievals have a large noise than those obtained from inversion models based on R_{rs} ratios due to a greater influence of atmospheric correction and radiometric calibration in the former model (Siegel et al., 2013)”

Page 16, line 14: I disagree with the statement “not clearly related to variations on time difference”. Fig. 3B appears to show a clear trend for the April 2001 data.

We agree with the reviewer and we complete the sentence

“Although largest biases (i.e., > 0.1 or < -0.1 m^{-1}) were either associated with relatively small (2.2 h) (e.g., sta. 4, April 2001) or large (5.3 h) (e.g., sta. 13, April 2001) temporal offsets (Fig. 3B), data comparisons during April 2001 also suggest a

trend consisting of larger discrepancies between field and satellite estimates as the temporal lag increases"

Page 22, line 18: Recommend adding a discussion of what in situ data was used to parameterize the EC model. Doing so might reveal why it performed well in the GSL.

We don't think is necessary to include that into discussion since it does not provide further insight regarding performance differences.

We checked again the datasource for the EC and we found that samples were derived from west Florida Shelf and Bayboro Harbor. These environments encompass case I and II waters but a narrower range of aph values with respect to the other inversion models. These differences do not explain the better performance of EC over the other models (see Carder et al., 2006).

However, we found that the better performance of EC during April 2001 was related to the greater influence of phytoplankton on IOPs during April 2001.
We wrote:
"The superior performance of EC in April 2001 with respect to May 2000 was mainly attributed to the greater contribution of phytoplankton to IOPs. In fact, further analysis of ac-9 measurements in the GSL showed that a_{ph}(440)/a(440) during May 2000 and April 2001 varied between 9.8 and 34.8%, and 20.3 and 92.9%, respectively. Thus, as waters become less optically complex due to a greater influence of phytoplankton on underwater light fields, accuracy of empirical inversion models improves"

Reviewer #2

Specific comments: Page 9308 -Line 3: I'm curious what kind of result you would obtain if you use aph*(λ) spectrum presented by Bricaud et al.(1995) in addition to the SEADAS default one.
It is a relative question since it would depend on the type of algorithm used for estimating chl in the first place. This will be translated later in more uncertainty as it was highlighted along the text.
We performed some calculations of chl based on EC and in situ aph* and then we couple these results to Bricaud et al. 1995 expression in order to compute aph* These preliminary results show that using a variable aph* parameterization as a function of chl as suggested by Bricaud et al. 1995 increased the median relative bias per cruise up to 11.8%. This error is smaller than that based on constant aph* value per wavelength as computed in our contribution. However, this reduction is part explained by the lack of independency between variables (i.e., chl shows up as independent and dependent variable parameter) used to develop the aph* function of Bricaud et al. (1995) function.

Page 9310 -Line 2: λ0 of QAA is not always 555 nm; for high absorbing waters (a(λ) 0.3 m^-1), Lee et al.(2002) proposes to use rrs(λ0) at 640 nm instead of that at 555 nm.

In QAA version 5 algo (see reference) Lee suggested 550 555 or 560 nm depending on the sensor. This is the version used in our calculations. We agree that 640 nm should be used in very absorbing waters but this has not been the case in the GSL during our surveys where a(440) usually below 0.3 1/m. Also, SeaWiFS has not a 640 nm channel. Despite this, we added to the text this 'potential' situation.

-Line 6: "CDOM" that you call in the manuscript is indeed CDM (colored detrital matter = CDOM + non-algal particles). This term should be modified throughout the text to avoid confusion.
That was modified throughout the text when needed. Some studies are based on acdom but some work with acdm

-Lines 12-13: 6 centered wavelengths (412, 443, 490, 510, 555, and 670 nm) rather than 5 wavelengths.
Ok, done

-Lines 11-22: In the original GSM algorithm, the final aCDM(443) is produced by multiplying a correction factor 0.754188 at the end (Maritorena et al., 2002, Applied Optics, Volume 41, pp. 2705-2714). This factor is no more used (Maritorena, 2010, personal communication). I'm curious what kind of result you would obtain if you remove this factor.
Sorry but I couldn't find that factor 0.754188 in Maritorena et al. 2002. We work in SeaDAS 7.0 with the version of GSM corresponding to Maritorena et al., 2002. It is out of the scope of this contribution to investigate these differences.

Page 9313 -Lines 7-9: Not clear. Do you mean that aph(443) estimates using QAA fell between those using EC and GSM?
It was meant that the number of valid aph443 estimates for all the datasets was EC > QAA > GSM. We improved the text adding intermediate for qaa

-Lines 24-26: What is the definition for low, intermediate, and high aph(443) values?
This is defined in section 2.6

Page 9314 -Line 4: "QAA(0.83) with respect to EC(0.66) and GSM(0.29)". What are these values in the parentheses? This sentence might confuse a reader.

they refer to proportions of samples having a relative bias smaller than 50%
We rewrote

-Line 14: td (time difference, h) should be explained both in the text and the Figure 3 caption.
Done

Page 9315 -Lines 3-29: It's hard to read values in white-black colors of the Figure 4. This can be fixed by using difference colors.
We tried before with other palettes and we found that a grey scale it is a better way to show contrasting values

Page 9319 -Lines 20-21: NOMAD dataset used for SA algorithm includes many case II waters as well.
We agree and eliminate case I waters. We rewrote
"First, SA computes *a*(670) with a model constructed as a function only of phytoplankton concentration as estimated from chlorophyll *a* (Bailey et al., 2010)"

Page 9320 -Line 3: Insert "to" between "respect" and "aerosol"
Done

Page 9320 -Lines 12-14 and 16-18: Any references?
No sorry, this interpretation is based in river discharge rate differences between May 2000 and April 2001

Additional references (not shown in the manucript): Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995), Variations in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization, J. Geophys., Res., 100(C7),13321-13332.

We added this reference along the text