

Interactive comment on “Methods comparison to retrieve the refractive index of small scatterers” by A.-M. Sánchez and J. Piera

A.-M. Sánchez and J. Piera

amsanchez@icm.csic.es

Received and published: 5 April 2016

Dear Reviewer, we are grateful to you for your comments and suggestions, which have helped to improve our contribution. Below you can find our answers to all your comments, addressing the modifications performed in the paper. Find enclosed the last version of the manuscript along with the pdf of the answers.

General comments

Comment: The method 4 – the genetic algorithm that this study is intended to evaluate – is simply a numeric method and it totally different from the “models 1-3”, which in essence are bio-optical models. In a sense, models 1-3 are formulae whereas method 4 is an (advanced) technique seeking solution of a formula. Therefore, there is no com-

C10447

parison between them. Actually, as the authors have already mentioned, the genetic algorithm can be applied to either of those three models in seeking a better solution. In addition, the model-1, as mentioned above, is intended to be used for entire particle populations that are assumed/expected to follow a power-law size distribution, and is fundamentally different from models-2&3, which were developed to apply to a phytoplankton culture (or dominance of one particular phytoplankton species) and require the concurrent measurement of the size distribution. These apple vs. orange comparisons show a poor design of the study.

Answer: The differences between the presented models are known and already stated in the text (see for instance end of Subsection 4.1.1). However, regardless of the scenario, they all perform the same task, which is the estimation of the inner refractive index of the small scatterers. The aim of this study is precisely a fair comparison between the different methods in several theoretical situations which, in some cases, have been adapted to fit in a particular model (note, for instance, that the first example uses a power-law distribution in order to apply the Twardowski Model). With this comparison, it is possible to certify that for the test cases presented in this paper, the Twardowski method is not the best option among all the possibilities (mainly, as you state, because it was not designed for isolated cultures). Even though this was an expected result, only after the numerical experiment the inadequate use of this model was objectively evaluated, considering the relative error as a performance indicator. From our point of view, the results shown in this paper can be useful for other scientists interested in the retrieval of the refractive indices in order to select the most suitable methodology without having to test all of them, at least, when having similar test conditions. To further clarify it in the text, the following modifications (highlighted in black) have been done in the Introduction: “Several inverse models to retrieve the refractive index from optical measurements can be found in the literature. For instance, a single equation based on the Lorenz-Mie theory is used by \citet{Twardowski2001}, to estimate the refractive index of a bulk oceanic distribution. It is indeed a fast method if optical backscattering measurements are feasible. \citet{Stramski1988} presented an extension of a model

C10448

from \citet{Bricaud1986}, designed for isolated phytoplankton cultures (or dominance of one particular phytoplankton species).” And, later, at the end of the introduction, the following paragraph has been added: “It must also be noted that the models are fundamentally different. The model developed by \citet{Twardowski2001} is intended to be used for entire particle populations that are assumed to follow a power-law size distribution while the other models are developed for single phytoplankton cultures (or dominance of one particular phytoplankton species) and require the concurrent measurement of the size distribution. And, besides, such bio-optical models are compared with a numeric method (i.e., the genetic algorithm) in the same conditions. On the other hand, the methodology applied in this contribution allows to obtain an objective comparison of the results of the different methods in those occasions where it is not clear which methodology is most suitable, and therefore, interesting for the ocean optics community.”

Comments: As far as optical modeling is concerned, I'm not sure if the super-accurate estimation of the refractive index offered by the genetic algorithm is meaningful. For one particular wavelength, the genetic algorithm was configured to partition complex refractive index into 2000 random values with real parts between 1.02 and 1.15 and imaginary parts between 0 and 0.02. Each of these complex values is tested to find the best refractive index that reproduce the observed absorption and scattering coefficients. Then this procedure is repeated by generated a new sets of random values following a certain rule (e.g., 50% crossover and 20% mutation). This entire process then moves to the next wavelength. The authored showed that the genetic algorithm can provide a solution with an accuracy of 0.08% for the real part of the index (the error was estimated against $n-1$) and 0.24% for the imaginary part of the index as in the test of spherical particles. Such moot precision can never be verified in an experiment nor can lead to meaningful improvement in optical modeling. Answer: The research presented in this paper comes from the necessity to reproduce the spectra signature of particle absorption and scattering. Using Lorenz-Mie or T-Matrix methods to this end requires retrieving previously the refractive index. Our first attempt in this field was us-

C10449

ing the existing methods of refractive index estimations but, as seen in the paper, they have an associated error. When the results were introduced in the forward simulations (Lorenz-Mie or T-Matrix), the desired spectra was never recovered, which made us to develop some improvements by using, for instance, an accurate optimization method such as the genetic algorithm. In any case, none of the models described in the paper are a realistic representation of real algae where there may be cell walls, chloroplasts, vacuole, nucleus and other internal organelles, each with its own optical properties. The best we can do is a gross approximation (usually using homogeneous spheres) only useful from an optical point of view. The aim of the genetic algorithm is only an attempt to make the approximation a bit closer to the reality. Maybe, using modern techniques (the development of new measurement techniques is obtaining more accurate mappings of refractive-index distributions in live cells and tissues, as seen for instance in Tomographic Phase Microscopy, published in Nature Methods By Choi et al. 2007), the verification could be developed. This has been clarified in the paper by adding the following sentence in the Introduction: “Although new promising techniques such as Tomographic Phase Microscopy (Choi et al., 2007) may provide in the future measurements of the refractive index in live cells, at present current ocean instrumentation do not directly provide it, so it must be estimated somehow (Aas, 1996).” And, in the Discussion: “To conclude, the results presented in Table 2 do not determine which is the best method to estimate the phytoplankton optical properties, since none of them are a realistic representation of real algae where there may be cell walls, chloroplasts, vacuole, nucleus and other internal organelles, each with its own optical properties. However, the assumed particles serve as a first approximation of actual phytoplankton and are useful to extract some preliminary conclusions and to introduce several improvements as an attempt to make the approximations a bit closer to the reality.”

Presentation

Comment: You used relative error in evaluating the performance, but didn't provide a definition. Since all of the relative errors cited in the text are positive values, I'd assume

C10450

you used absolute values. But, this should be defined.

Answer: This has been clarified adding the definition of the relative error at the beginning of Section 4.

Comment: In Page 18739 Lines 3-16 and Figure 5, you compared genetic algorithm with other optimization algorithms, which you did not introduce in the method section. You also mentioned that the BFGS method showed averaged relative error 0.073% for the real part and 0.72% for the imaginary part, which you think are worse than the genetic algorithm. However, this performance measure is actually better than the performance of the genetic algorithm you listed in Table 1 for the real parts of the index.

Answer: As you can find in the text (just above the text you pointed out), the error committed by the genetic algorithm is 0.004% for the real part and 0.24% for the imaginary part, which is better than that committed by the BFGS. In Table 1 there was an error in the real part that has also been corrected. The reason for not introducing BFGS, which is another optimization method as it is the Genetic Algorithm, is because its performance is much worse than the Genetic Algorithm when applied on the examples described in the manuscript. The results were only added in order to state that other optimization algorithms were also tested (a part from BFGS we also tried with Nelder-Mead, Conjugated Gradient, etc.), but still the Genetic Algorithm provided the best solutions. We did not find necessary to make the paper longer with the introduction of a new method that does not present any meaningful improvement. As explained in the text, the Genetic Algorithm presents some disadvantages, being the convergence time the most important one, but its accuracy level is not achievable by any other algorithm. In order to clarify the reason for not introducing in more detail the BFGS algorithm in the paper, the following text has been added in Section 4.1.3: "Other optimization algorithms were also applied to determine if similar results can be obtained with a significant reduction of the computation time. However, since none of them led to any meaningful improvement, they are not introduced here. As an example, Fig. 7b shows the results. . ."

C10451

Comment: I think the test 3 (cylindrical particles) is confusing. First, you used coated spheres to emulate homogeneous cylindrical particles in inversion. Since the cylindrical particles are homogeneous and have only one refractive index, how do you evaluate the results (Fig. 14) of the coated spheres which would give two indices, one for the core and one for the coating. Second, due to the computation constraint, you used equivalent volume spheres to simulate the cylindrical particles in inversion. How can this help you evaluate the performance of the genetic algorithm? It cannot! And it is clearly shown in Fig. 16. Since absorption is proportional to the volume and you used volume equivalent spheres, the inverted imaginary part of the refractive index agree well with the assumed values. However, since scattering depends strongly on the shape of particles, the inverted real part of the refractive index deviate significantly from the assumed values.

Answer: As you state, the two complex refractive indices of the coated sphere cannot be compared with the individual one of the homogeneous cylinder. Instead, we need to use the IOPs that are recovered using the estimated refractive indices in the forward model to analyse if coated spheres are useful to emulate homogeneous cylinders. In this particular case, we used the volume scattering function to this end. To clarify this in the text, the following sentences have been added in Section 4.3.1: "Figure 16a shows the assumed and estimated real part of the inner and outer layers and Figure 16b shows the assumed and estimated imaginary parts. In this case, they cannot be compared with the assumed individual refractive index of the cylindrical particle. Instead, the IOPs obtained from the estimated refractive indices need to be computed using the forward model to analyze if this model is useful to emulate homogeneous cylinders." Regarding the second comment, the real evaluation of the genetic algorithm is done in Section 4.1.3, since it is where the genetic algorithm uses the same shape as the assumed one to estimate the refractive index. In this example, the same could be done if we didn't have the computation constraints when using cylinders. Indeed, simpler examples using cylinders, not reported in the paper, have already been tested with the genetic algorithm and the original refractive index was accurately recovered. But, for

C10452

the paper example, we needed to think in a faster estimation since it was not practical for us to leave a PC computing for several days. Thus, this result cannot be considered as a validation (or invalidation) of the genetic algorithm model, but an alternative technique to find the assumed refractive index. For sure, using better computing resources (as for instance, by means of a computer cluster), this problem disappears and the genetic algorithm can be used with its complete potential. To clarify this in the paper, the following sentence has been added in Section 4.3.2: "For sure, using better computing resources (as for instance, by means of a computer cluster), this problem disappears and the genetic algorithm can be used with its complete potential". And later, in the same Section: "Since absorption is proportional to the volume, the inverted imaginary part of the refractive index agree well with the assumed values (volume equivalent spheres are being used). However, since scattering depends strongly on the shape of particles, the inverted real part of the refractive index deviate from the assumed values".

Comment: While I can understand the text, the writing needs improvement. Also, some figures are difficult to interpret. I will list some specific examples below regarding the writing, figures and others.

Answer: Paper style has been improved using the recommendations described below, the suggestions of Reviewer 2, and others of our own.

Specific comments

Comment: 2π in Eq. (9) is not a normalization factor. It comes naturally from integration w.r.t. the azimuth angle. Sometimes (and often in atmospheric optics), the integration of phase function is normalized to 4π (representing the total solid angle over entire sphere), in this case, the so-called factor is $1/2$.

Answer: Last paragraph of Section 2.1 has been re-edited considering this suggestion.

Comment: The Bernard et al. 2009 reference, which you have cited multiple times and

C10453

is the basis for your model-3, was not in the bibliography list.

Answer: It did not appear due to a problem of compilation in Latex. It has been solved.

Comment: Page 18734, Line 11, attenuation → absorption

Answer: It has been corrected.

Comment: Page 18740, Line 4, ". . . not agree with a perfect power-law distribution since there is minimum size beneath which there are no cells." Any power law function has to stop somewhere in the lower end!

Answer: This sentence tries to state the difference between the PSD of Fig. 4a (performing a classical power-law function), typical in oceanic bulk distributions, and that of Fig. 9, more similar to that of isolated cultures. In order to avoid any confusion, the following text: "Instead of using the PSD of Fig. 4a for this example, the PSD of an isolated culture was simulated with a concentration of 40 particles mm³ ($R_{min} = 0.7 \mu m$, $R_{max} = 12.1 \mu m$). It must be noted that the PSD denotes the external radius (the inner one can be extracted using the VV value). In this case, the function does not agree with a perfect power-law distribution since there is a minimum size beneath which there are no cells. Thus, the PSD of Fig. 7 (using 31 points) better represents the case of a monoculture PSD.", has been simplified to: "In this example, instead of using a PSD describing a power-law function (as in Fig. 4a), the PSD of an isolated culture was simulated with a concentration of 40 particles per mm³ ($R_{min} = 0.7 \mu m$, $R_{max} = 12.1 \mu m$ and using 31 points), as seen in Fig. 9. It must be noted that the PSD denotes the external radius (the inner one can be calculated using the VV value)."

Comment: The way the volume scattering functions were shown in the figures does not help in evaluating the results. Why not draw VSFs at only a few wavelengths using lines instead of the color map.

Answer: VSF figures have been redrawn as suggested. Only three wavelengths (300, 500 and 700 nm) are plotted using intense colors, as seen in the legend, while the rest

C10454

of wavelengths between 300 and 700 nm in steps of 10 nm are plotted in light grey. A clarification has also been added in the last paragraph of Section 4.1: "The volume scattering function is shown in Fig. 5b. Only three wavelengths (300, 500 and 700 nm) are plotted using intense colors, while the rest of wavelengths between 300 and 700 nm in steps of 10 nm are plotted in light grey."

Comment: In section 4.3 Cylindrical-shape particles, you tested coated sphere, but did not mention the size of the core and how did you come up with that size.

Answer: The size of the core can be extracted from the relative chloroplast volume VV . As in previous examples, the assumed value was $VV = 30\%$, since it is a value between that assumed by Bernard et al., 2009, and previous works. This clarification has been added in Section 4.3.1 with: "As in previous examples, the assumed value was $VV = 30\%$ (an averaged value between that assumed by Bernard et al, 2009, and previous works)."

Comment: Cylindrical-shape or spherical-shape should be cylindrical-shaped or spherical-shaped

Answer: It has been corrected.

Comment: In specifying wavelengths (e.g. Page 18743, line 13), longer or shorter are typically used (e.g., longer wavelength), whereas higher or lower are typically used for frequencies (e.g., lower frequency).

Answer: Thank you for this clarification. It has been corrected.

Comment: Both "initial" and "synthetic" refractive indices (as in figure captions) are used to represent the assumed values that have been used to simulate the optical properties. Initial values were also used in running the genetic algorithm. Recommend to use "assumed" to avoid confusion

Answer: The document has been reviewed and the "initial" and "synthetic" adjectives have been replaced by "assumed" when not referring to the initial values of the genetic

C10455

algorithm.

Please also note the supplement to this comment:

<http://www.biogeosciences-discuss.net/12/C10447/2016/bgd-12-C10447-2016-supplement.zip>

Interactive comment on Biogeosciences Discuss., 12, 18723, 2015.

C10456