Ocean Science
Discussions

Interactive comment on "Estimation of phytoplankton pigments from ocean-color satellite observations in the Sénégalo-Mauritanian region by using an advanced neural classifier" *by* Khalil Yala et al.

Anonymous Referee #2

Received and published: 3 November 2019

This paper presents a new neural classifier model for estimating phytoplankton pigments from satellite data, by using an unsupervised learning method instead of the standard supervised learning approach, as the sample size is limited. The method involves blocking the variables as opposed to the unblocked approach in the standard self-organizing map (SOM) approach. I believe the writing needs considerable improvement, as the paper is difficult to read, and the advantage of their 2S-SOM over the standard SOM is not brought out. The English can also be improved.

Specific comments:

There is a lack of comparison with controls for the reader to appreciate the advantage of using this new model. At the minimum, there should be more comparison between the new 2S-SOM model performance scores versus the standard SOM model scores. [The paper would be more interesting if the performance of 2S-SOM is also compared against standard supervised learning models such as multi-layer perceptrons or random forests.]

On line 321, the choice of the elongated 2-dimensional grid of 9x18 is not obvious. Why is a more square grid (e.g. 12x12, 12x13 or 13x13) not used?

The paper is very hard to read as there is a tendency to present many undefined symbols all at once, with the symbols remaining undefined until much later in the paper. For instance, Eq.(5) introduces a large number of symbols and terms all at once. The "block" is not explained in a concrete way until the next section (Sect. 3.3), so I had a misconception on how the data were blocked when reading Sect. 3.2. A much more logical order of presentation is to present the concept of blocking variables first, and try to explain as many of the symbols coming up in Eq.(5) before actually presenting the equation. Also around Eq.(5), there are numerous typos and inconsistent fonts (as listed later in this review).

Minor comments/typos:

Line 22-23: "Thanks to . . . new method. It primarily consists in. . ." is verbose. Simplify to "Our new method consists of . . ."

Line 25: "carried using" should be "carried out using".

Line 69 and throughout the manuscript: Bold fonts are for vectors and matrices (see the journal's manuscript preparation guidelines), but here they are often used for scalars and units. There are many places where the font switches back and forth between bold and Roman and italics (e.g. lines 248-251 and line 272).

Line 151: Need a reference for the OC4V4 algorithm.

Line 162: The last sentence of the paragraph and Table 1 need to be moved to after line 183. The table is currently placed before the terms in it are defined.

Line 174: How can Ra be independent of chl-a if it is divided by rho_wref which is dependent on chl-a?

Line 174: "sensitive the" should be "sensitive to the".

Line 248: W is undefined.

Line 254: should give a specific reference on the kernel and temperature.

Line 276: How were B and Pb chosen?

Line 278: "a" should be alpha.

Line 282: Eta should be beta.

Figure 4: For 2S-SOM, I can see long dash, short dash, space and no space variants.

Line 420: Last sentence of paragraph: I have trouble understanding this sentence.

Fig.13: Top right corner is slightly chopped off.

Line 542: "a" should be alpha.

Fig.16: I don't understand why the black curve tends to lie closer to the blue curve than the red curve is to the blue curve. I would have expected the red curve to lie closer to the blue curve. I might have misunderstood what the curves represent – please give more detailed explanation.

Line 639: Replace "people" with "studies".