MINUTES OF THE GIOP WORKSHOP
HELD DURING OCEAN OPTICS XX
25th September 2010, Anchorage, Alaska, US

Participants
Bryan Franz
Hubert Loisel
Emmanuel Devred
Frederic Melin
Timothy Moore
Constant Mazeran
Zhong Ping Lee
Emmanuel Boss
Samantha Lavender
David Antoine
Stephane Maritorena
Jeremy Werdell
Tim Smyth: videoconference

Cécile Dupouy joins the meeting 9.30

Excused absents
Odile Fanton d'Andon and Antoine Mangin could not attend the meeting and are represented by Constant Mazeran

9h - Meeting begins

Introduction

JW welcomes participants and presents the three items of the meeting:

- Review and feedback on the implementation of GIOP in SeaDAS;
- Metrics identification for evaluating algorithm performance;
- Discussion on how considering input/output uncertainties in the inversion.

BF recalls that the goal of GIOP is to make the Ocean Colour community play and test the algorithms. However NASA has not received much feedback on GIOP, except few questions about what is it.

Also, there has not been a big reactivity of the scientific community in NASA's call about research on uncertainties. SM notices that uncertainties handling is very specific of the inversion, and change with the algorithm. Furthermore EB points out the difficulty of uncertainties propagation, since NASA does not provide uncertainties at Level1.

Bryan Franz's presentation: Review of GIOP implementation in SeaDAS (l2gen)
BF reviews the implementation of GIOP – see OOXX abstract "*A Generalized Framework for Modeling of Inherent Optical Properties in Ocean Remote Sensing Applications"*.

Subjects discussed:

- ***Goal of GIOP and warning to the users***

BF recalls the goal is to have a common framework to evaluate methods and find consensus for a global product. It could also propose options for regional users. SM insists that it is very important to make this “testing approach” extremely clear for users, because some of them who don't know well the project could be lost and make a wrong use of the data.
ED notices that GIOP gives potentially many information that might be difficult to interpret. BF: already some flags for non-physical values....

- ***Accuracy of Lee et al 2002. conversion and its dependence on sun angle***

JW: this has been considered, and does not represent a very big error; it is clearly not the biggest problem in the inversion.

- ***Optimization procedure: choice, issue***

EB emphasizes that the problem of all non-linear inversion methods proposed so far is that they provide always a solution, but there might exist other one’s we do not know. On the contrary, when the problem is linearized, the algorithm finds the unique solution, and if there are more equations than unknowns then it finds the best solution in the least-square sense.
However the advantage of non-linear inversion like Levenberg-Marquardt is that we can weight the cost function. Both options are available in GIOP.

CM asks if the spectral shape of the IOP, notably bbp, could be free parameters. ZL thinks it is impossible, except maybe if a range is given. ED thinks as well than an iterative scheme would not necessarily converge. BF says it could be tried (but if cross correlation then several solutions).

It would be convenient to add the possibility to go through several shapes and find the best one. Currently need to be done one by one.

- ***Output products***

Size fraction is not a product yet, and could be of interest. BF will assess if this is technically possible.

- ***Modeling of the forward model***

EB reminds that it would be important to include salinity in bbw. It is a bias already pointed in the past, which should be addressed.

aph* of Bricaud should maybe use the one published in 2004.

TM thinks there is an issue for splitting aph and adg; there should be an option to partition IOP.

HL suggests testing a power law instead than an exponential for cdm.

EB wonders whether we should include fluorescence in the model, since GIOP proposes to go up to 700 nm. It would be a big issue. If not possible, then the inversion should not consider such bands.

HL wonders about accounting for Raman scattering in the Gordon quadratic formula. It should be intrinsically included because Gordon computed those values from in-situ measurements. In any case, JW says that further developments could not be accomplishable in short-medium term.

HL notices that the f/Q LUT is limited to Case1 waters with Chl as input. Alternative for Case2 waters should be addressed - ED is working on it.

ZhongPing Lee's presentation: Some refinements for global IOP inversion
Main contributions are:

- ***Importance of phase function for IOP computation even in clear waters.***

A bias or "background" in bbp versus Chl may come from the fact that G or g is actually not a monotonic function in bb/a+bb. It depends on phase function shape, different for molecular and particle scattering. The explicit separation of molecules and particle phase function gives a much better fit with respect to Hydrolight simulations. Effect is a 40% reduction in bbp and more for adg443.

- ***Angular dependency of G factor***

Construction of a G LUT with angular-dependent model coefficients for Rrs; to be given to Bryan Franz.

- ***Inversion: QAA vs spectral optimization***

There is a philosophic difference, QAA giving first the total IOP, while optimization gives first individual IOPs. Advantage of QAA is the possibility to retrieve aphi spectral shape, which is a property we want to obtain, whereas optimization assumes a shape before derivation. On the other hand, QAA considers every measurement as signal, while optimization considers the mismatch as noise, uncertainties of the model...
Confidence in the retrieved IOP can come from the agreement between several optimization techniques; no agreement would mean need for further research.

- ***CPU time consideration***

QAA computation time dramatically reduced. Could be important for hyperspectral data?

- ***Future inversion: building an Ensemble of simulations***

EB suggest that if there is no CPU time constraint, then why not building a huge LUT for all combinations of IOP and find the node with best agreement? If several solutions, then it would give directly the uncertainties. That's may be the future, but not possible now.

David Antoine's presentation: some (basic) considerations on our capability to derive bbp from AOPs (R and Kd), in-situ and especially in clear waters – an example using data from BOUSSOLE and Plumes & Blooms
Data shows that relationships do exist between in-situ bbp and Chl, yet scattered. A simple inversion procedure using Kd and R as input (+ LUT of f' and mud from RT computation) allows bbp to be retrieved quite well; some issues however in the range of low bbp < 0.001 m^{-1} (50% of the ocean)
The inverted bbp versus in-situ Chl presents different relationships, which are furthermore different at two bands (443 & 555); this surely means a problem in the spectral dependency of the modeling.
Conclusion is that Chl is not a good predictor for bbp.

A requirement from this study is to get more data in low Chl domain (below 0.3 mg m-3). Half of IOCCG dataset is for bbp > 0.01 m^{-1} (Chl > 5 mg/m3). We need overall error

estimates. The plots should contains error bars for the in-situ data.

EM notices we should be careful with small in-situ bbp measurement because WET Labs discovered problem in the bbp measurement. There might big large absolute difference, important for small values.

Jeremy Werdell's presentation: Analyses and tools for evaluating algorithm performance.

Matchups consists in NOMAD dataset, IOCCG data set and SeaWiFS match-ups. Results are provided on http://oceancolor.gsfc.nasa.gov/cgi/giopval.cgi with the baseline algorithm. JW summarises comparison to in-situ / synthetic data, with the baseline configuration. These results are obviously subject to evolution.

- ***Some recommendations or ideas on the current analysis:***
  - the absolute difference Delta Rrs could be the weight function in the minimization
  - 660 nm is maybe to far for being in the sum of Delta IOP
  - comment on validation plots : at very low value, % difference means nothing. It's not as noisy as it looks like.
  - adding the total "a" would be interesting. More generally, any analysis should be looked at individual component level as well as total.
  - pb in file size to add many wavelengths (ZL : 412 490 560 for total "a" would be great).
  - have a plan to filter the data or have a default validation dataset. What is the dataset considered as satisfying for giving final conclusion on best configuration ?
  - we should understand the different results between f/Q and Gordon G at Level3 : are difference due to low sun angle, or low chlorophyll ? First we should use ZL's separation between water and particles, and check if problem still there.
  - for the sake of visualization and easier SeaDas handling, ED proposes the idea of generating a full Level1 synthetic image (with atmosphere), partitioned into different type of waters ; interest zould be to visualize boxes at once, looking where the algorithm succeeds fails...

- ***Sensitivity analysis***

A sensitivity analyses is presented by JW, which purpose is to find a hierarchy of what should be studied first. Tests include Rrs transmission, Morel f/Q vs Gordong g, LM vs matric inversion, variation on aph, Sdg, eta… Hierarchical summary of sensitivity analysis is:

- Tier1 (very sensitive): LM vs linear matric inversion, aph from Ciotti of GSM, Sdg +/- 33% or from OBPG, Morel f/Q vs Gordon. ZL explains that the huge impact of f/Q vs Gordon in aph for large values surely comes from the quadratic term "g1" of Gordon.
- Tier 2 (sensitive, typically in particular parts of dynamic range): 6 bands vs 5; Sdg+/- 10% or from QAA
- Tier3 not sensitive: transmission, Bricaud ch +/- 10% and 33%, eta +/- 10% and 33%

Hierarchy should be done for each component, including "all".

- ***Validation dataset***

TM remarks that the datasets are very coastal centered: may be not representative for the global ocean. He recommends to build another IOCGG dataset with distinction between Case 1/Case 2.
ZL proposes to separate dataset wrt bathymetry. JW acknowledges that we should stratify the analysis. HL suggests we should use synthetic dataset first (yet it has some assumptions) in order to include other sources of errors only in a second time.
EB says we should analysis the closure of NOMAD IOP with Rrs, in order to filter the data – difficulty is that there are not many matchups with concurrent IOP and AOP.

Note that there is currently no update of NOMAD foreseen.

- ***Inversion and Metrics of the fit***

If we have to decide on a inversion method, EB recommends to use Levenberg-Marquardt as done in GSM because it takes into account uncertainties. However the cost function is highly important (weights between wavelengths): FM notices that global weights (in space and time) may be dangerous, for instance by over-constraining some bands. Ideally it should be pixel dependent.

- ***Validation protocol***

TS wonder how synthesizing the results, which can be better at one wavelength, worse at another ... there is no magic number.
EB states that some variables are meaningless at some wavelengths, and uncertainties within the cost function would have bad impact. ZL says that for instance we mainly see the effect of pure water after 550 nm, and so there is no reason to look after this band for absorption.

→ There is a need to build a list of products + associated wavelengths for validation (where the uncertainties are the smaller, i.e. meaningful). For example bbp could be restricted between 490 and 550 (TBC because 600 nm presents maximal response for coastal water and would be interesting).

→ Should the working group think about creating a validation protocol for IOPs ?

Sam Lavender's presentation: WaterRadiance project

SM gives an overview of the project: literature review, scientific analysis plan, development of water optical properties model, sensitivity analysis. The last task is the development of retrieval algorithms of IOP: GKSS NN and GIOP approach
No boundary between Case1 / Case2. Big part is to improve properties of pure water.

Tim Moore's presentation: role of optical water type classification in the context of GIOP

TM presents a methodology for classification of water type. It shows 60% of IOCCG dataset belongs to 6-7 classes (high scattering water). It could be interesting to add one or two classes, in order to check we are not missing other water types. HL remarks that classification is very dependent on the way the normalisation is handled and on the mathematical method.

There is a possibility for including OWT classes in GIOP and inverse the IOP with locally adapted spectral shapes.

Hubert Loisel's presentation: recent advances for the inversion of the bbp at different wavelength

Recall of the original Loisel & Stramski approach: assessment of the total IOP from basic radiometric measurement, without assumption on spectral shape on IOP. Limitation is the use of R(0-) instead of Rrs. Kd is not measured from remote sensing but estimated from Rrs.

New improvements since the first version are a directly consideration of Rrs instead of R(0-) and a new parameterization between Kd and remote sensing ratios for each wavelength. With the new model, the main problem is retrieval of Kd at different lambda. It is done thanks to a NN approach (Rrs between 412 and 670 nm as input, Kd(490) as output).
This retrieved Kd(490) compares well to other empirical approaches (even better in term of % relative difference). Good and very improved performance at 490 nm for atot and bbp on the IOCCG dataset.

Need to test another dataset. Need to extend learning dataset for very low values of Kd (e.g. Biosope), add algorithm of Lee 2005 in the comparison.

Frederic Melin's presentation: assessment of bio-optical algorithms

Introduction to the BIOMAP dataset: 826 stations in Europe from 2004 to 2009 (see poster as well). Presentation of statistics comparison between QAA and GIOP.
Few maps of aph processed with GIOP seem to be very noisy (more than other products). Does it show sensitivity to the input?

Stephane Maritorena's presentation : uncertainties estimates in input (Rrs) and output ocean color data : a brief review

SM introduces the problematic by recalling the difficulty we are facing in Ocean Colour: not enough matchups for determining uncertainties, which can only give global general accuracy estimate but do not take into account spatial and temporal variability. This is why uncertainties at pixel level are required.
Methods exist in the literature for semi-analytical algorithms: error propagation, covariance matrix in non-linear least-squares techniques, bayesian approaches…
These approaches do not rigorously return the same things:

- Recent implementation with GSM for daily global uncertainties (Maritorena et al. 2010)
- Error propagation through a series of analytical expressions (Lee et al 2010)
- Bayesian approach for error estimation
- Uncertainties in empirically derived variables (T. Moore et al 2009).

- ***Uncertainties in the input data***

Assessing errors in Rrs is not an easy task. Error propagation studies consider TOA radiances, add noise in the NIR bands and propagate. But it requires knowledge of Level1 uncertainties, not easily available – could it be given by sensors inter-comparison?

Including matchups uncertainties, even constant in space and time as in GSM, is necessary and best than using nothing (at least for variation in spectral weight).

- ***Uncertainties in the output data***

A first question in the output uncertainties is to choose, for a given product, at which band we record the error: should it be for one band or for the full spectrum? First option need to define the reference wavelength of each product.

Regarding uncertainties simulation, EB suggests employing a Monte-Carlo method on the input Rrs, and look at the errors in the output IOP. It will maybe depend on Rrs spectra, and will give confidence estimates. An important question is to know whether the error in Rrs constant or relative. FM thinks it is better to consider absolute uncertainties.

An alternative is to consider the GSM approach, for Levenberg-Marquardt inversion intrinsically provides output uncertainties estimates. The main question of this solution is whether the goodness of the fit is a relevant metric wrt uncertainties in the matchups? Results of Maritorena et al 2010 provides encouraging results in that direction.