05/18/23

**Uncertainty considerations for the ZTT inversion algorithm**

Tim Moore and Mike Twardowski

IOP product uncertainties for the ZTT inversion algorithm were characterized through a consolidated, independent data set comprising varied marine environments and optical conditions. The data set was carefully selected and vetted through closure assessments via radiative transfer simulations (using Hydrolight). Independent closure assessment was a criteria for data selected for testing, and requires ancillary data along with concurrent AOP and IOP measurements. As such, some common global data sets such as NOMAD did not meet this requirement, and while used for evaluation, did not factor into our uncertainty estimates. As a side note, the lack of existing hyperspectral data sets meeting these requirements forced evaluation at the multispectral level.

A variety of statistical metrics were calculated to derive IOP product uncertainties. Notably, the mean absolute error (MAE), the root mean square error (RMSE), the bias, and the mean absolute percent error (MAPE) were calculated for individual spectral bands and bulk spectral averages for particle backscatter (bbp), total absorption (at), and the ratio of total backscatter to total absorption. More detailed assessment of absorption product uncertainties were not possible with the vetted evaluation data set due to the lack of absorption component measurements (phytoplankton absorption and colored dissolved and detrital absorption).

These uncertainties were determined for numerous combinations of subcomponent IOP models and other algorithm settings (e.g., the particle backscatter ratio specific to ZTT), in addition to similar configurations for the current GIOP framework. In all, 54 configurations of ZTT were evaluated.

The main constraint was the paucity of high quality independent field data that met requirements for a closure analysis and evaluating all the IOP products across diverse environmental and optical conditions. Ultimately, a small data set (N=51) with multispectral resolution and incomplete IOP measurements covering a range of case 1 and case 2 waters (but not fully representative of the global oceans) was assembled and used for product uncertainty determination. The errors within the field data were difficult to characterize, except at a high level from closure comparisons of predicted AOPs with measured AOPs. Error tracing within the evaluation data set is needed to understand error propagation and significance with model testing. *The paucity of high quality, comprehensive field data sets is a significant constraint in assessing uncertainties.* The limitations apply to multispectral data sets, and even more so to hyperspectral data sets. Hyperspectral backscattering data are not widely available, and widespread capability does not exist beyond a few groups. Part of the constraint could be attributed to the lack of data sharing between groups, more so between non-NASA funded projects. We note that these limitations mentioned apply to the visible spectrum, and even less is known about the UV and NIR regions.

The data sets we eventually used were collected in the last 5 to 10 years (not including NOMAD which extends even further back). Given the resources and time it takes to collect field data, filling the current void with all the requirements needed for complete evaluation of bio-optical algorithms for the hyperspectral domain stretching from the UV to the NIR will likely not be achieved quickly.

This was all drawn from a manuscript we are about to submit for peer review.  Most of what is written is our personal views and conclusions from this manuscript that covers our PACE work.

The overall reference for the manuscript is:

Moore, T.S.,Kolluru, S., Tonizzo, A. and Twardowski, M.S. A bio-optical inversion scheme based on the radiative transfer equation, *prepared for* *submission to Optics Express*, (2023).

We used standard evaluation metrics - Seeger et al 2018 is probably the best source for citing the rationale of metrics used.