**FastMAPOL uncertainty analysis report**

Meng Gao, Kirk Knobelspiesse, Pengwang Zhai, Andy Sayer, Amir Ibrahim, Bryan Franz

1. **Algorithm introduction**

FastMAPOL is a remote sensing algorithm which jointly retrieves aerosol and ocean color properties from multi-angle polarimetric measurements. The algorithm conducts non-linear least square optimization and employs fast neural network forward models to conduct coupled atmospheric and oceanic radiative transfer simulations (Gao et al AMT 2021). The algorithm evolved from the Multi-Angular Polarimetric Ocean coLor (MAPOL) algorithm which was validated through RSP measurements, and further optimized for both AirHARP and HARP2 instruments in terms of accuracy and efficiency. The algorithm is designed to facilitate operationally processing the large volume of multi-angle polarimetric data acquired by PACE.

1. **Data product**

The major data products to be retrieved from FastMAPOL include aerosol microphysical properties (e.g. fine and coarse mode aerosol size, complex refractive index), and optical properties (e.g. spectral aerosol optical depth, single scattering albedo), ocean surface wind speed, and ocean chlorophyll a concentration. Based on the retrieval aerosol properties, atmospheric path radiance will also be calculated and used to derive the water leaving signals. The FastMAPOL algorithm has been used to process the AirHARP measurements from the ACEPOL field campaign (Gao et al, AMT 2021) and synthetic HARP2 measurements (Gao et al, Frontiers, 2021).

1. **Uncertainty estimation components**

Error propagation is a useful tool to evaluate the pixel-wise uncertainty from a retrieval algorithm. However, it requires that the forward model and uncertainty model can sufficiently describe the measurements. Therefore, to ensure the estimated uncertainty and the retrieval product are self-consistent and quality ensured, three keys components are implemented to conduct uncertainty quantification from FastMAPOL: 1) theoretical error propagation which maps the total input uncertainty using forward model Jacobian matrix, 2) an adaptive data screening and cloud masking procedure, which ensure data quality for retrieval and uncertainty evaluation, 3) statistical analysis on the retrieval fitting residuals.

* 1. **Error propagation**

The error propagation method is based upon a Bayesian approach which assumes Gaussian distributions of input uncertainty (including measurements, forward model, and a priori) and output (retrieval) uncertainty (Gao et al, AMT, 2022). These represent the 1 standard deviation (1σ) uncertainty on the retrieved state and are determined by mapping the measurement and forward model uncertainties into retrieval parameter space,

A black text on a white background

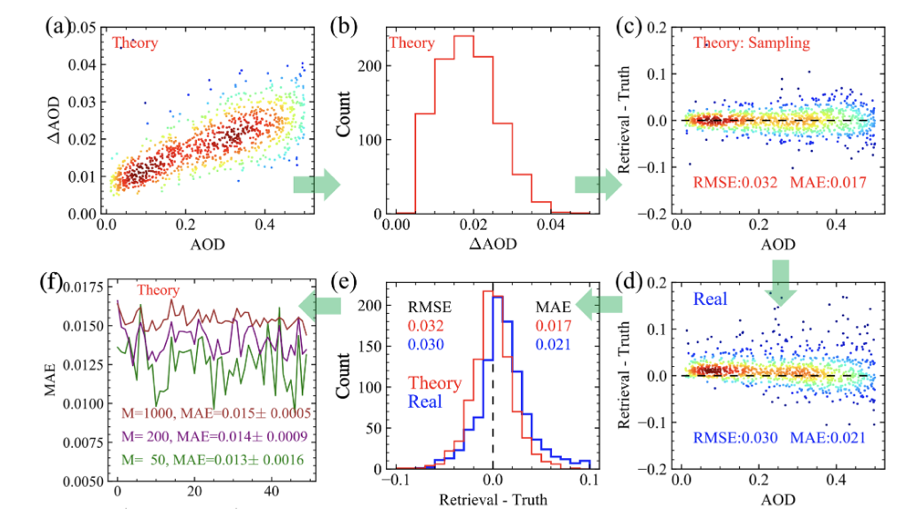
Description automatically generated with low confidence

where *S* is the retrieval uncertainty covariance matrix, is the error covariance matrix which includes contributions from measurement and forward model. *K* is the Jacobian matrix, and *Sa* is the a priori uncertainty covariance matrix. Total input uncertainty is characterized in the error covariance matrix, which includes the instrument uncertainty, as well as contributions from both neural network model, and the uncertainty of the radiative transfer simulation data used for the neural network training (Gao et al, AMT, 2021):

A picture containing font, handwriting, calligraphy, white

Description automatically generated

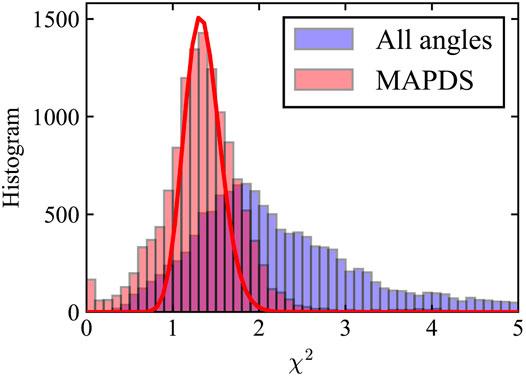
To evaluate the retrieval uncertainty performance from FastMAPOL, the error propagated uncertainties are computed for each pixel, and then compared with the difference between the retrieval results and the truth. To quantify retrieval performance, synthetic data with known truth inputs are used. In real retrievals with field measurements, in-situ measurements can be selected as the truth. The root mean square error (RMSE) and mean absolute error (MAE) are computed based on the difference between retrieval and truth, and also the random numbers sampled from the error propagation uncertainty. Such comparison can be used to indicate the retrieval quality as how well it agrees with the theoretical error propagation. An example of the key procedures is shown in Fig 1:



*Fig 1. Demonstration of the procedures to compare theoretical and real uncertainties. (a) Theoretical uncertainties of AOD retrievals over 1000 synthetic HARP2 measurements; (b) histogram of panel (a); (c) the retrieval error sampled from panel (a); (d) the retrieval error  
derived from the difference between the real retrieval results and truth data; (e) the histogram for the retrieval errors in both panel (c) and (d); (f) the MAE for 50 sets of random theoretical errors considering a total number of cases of 50, 200, and 1000. (Reproduced from Gao et al. AMT, 2022)*

* 1. **Adaptive data screening and multi-angle cloud masking**

When anomalies such as cirrus clouds are presented in the measurements, the aerosol based forward model in FastMAPOL cannot sufficiently describe the measurement, and therefore both the retrieval results and the estimated uncertainty cannot represent the actual retrieval performance. Therefore, an automatic data screening approach is required to ensure the measurement and forward model are self-consistent. Based on synthetic AirHARP and HARP2 data, we find that a total of 20–30 angles across all bands (i.e., five to eight viewing angles per band) are sufficient to achieve good retrieval performance. Furthermore, if not all angles are influenced by the cirrus cloud, the remaining angles can be still used to conduct useful aerosol and ocean color retrieval. Following from this result, we develop an adaptive multi-angle polarimetric data screening (MAPDS) approach to evaluate data quality by comparing measurements with their best-fitted forward model. The data screening method effectively identifies and removes viewing angles affected by thin cirrus clouds and other anomalies, improving retrieval performance. This was tested with AirHARP data, and we found improved agreement with the High Spectral Resolution Lidar-2 (HSRL-2) aerosol data (Gao et al, Frontiers, 2022). The cost function statistics more closely approach to a *χ*2 distribution after removing the cirrus cloud through the multi-angle data screening method (MAPDS) as the example shown in Figure 2 for AirHARP measurement as shown in Fig 2.



***FIGURE 2****. Histogram of the* χ*2 of the FastMAPOL retrievals from AirHARP measurement in ACEPOL field campaign with all angles used and with cirrus cloud angle removed (MAPDS). (Reproduced from Gao et al, Frontiers, 2022)*

* 1. **Fitting residual and correlation analysis**

The retrieval uncertainties estimated by error propagation assume that the error covariance matrix correctly represents the measurement and modeling uncertainty even after the data screening. However, the uncertainty model may perform differently for different scenes and the measurement may also correlate among angles which is often ignored in the retrieval and uncertainty estimation due to lack of knowledge in the data correlation characteristics. To conduct a closure analysis, the fitting residuals which represent the difference between the measurement and the forward model can be compared with the uncertainty model. The disagreement can suggest over fitting, under fitting or an inaccurate uncertainty model. As shown in Figure 3, the standard deviation of the residuals and the actual error in the synthetic data are compared with the uncertainty model for both reflectance and polarization under different strengths of angular measurement correlations. Under stronger correlation, the fitting residuals are found smaller than the measurement uncertainty described, which is due to the impacts of angular correlation and over-fitting of the data. More techniques are provided to account the angular correlation in the retrievals in Gao et al, AMT, 2023.

A picture containing text, screenshot, font, diagram

Description automatically generated

***Figure 3*** *The standard deviation of simulated measurement errors and fitting residual for cases C3 and C4. The model uncertainty of 3 % for reflectance and 0.01 for DoLP are indicated by dashed lines. Results of scenarios C1 and C2 for reflectance are similar, but with estimated DoLP uncertainties closer to the 0.01 line. (reproduced from Gao et al, AMT 2023)*

1. **Validation plan**

The validation of the retrieval product and their uncertainties are through a combination of all above three components. The measurements will go through the data screening and cloud masking process. The resulting measurements may lead to a smaller number of valid angles per pixel and an increased retrieval uncertainty, but the retrieval uncertainty can be better characterized through error propagation. When the validation datasets are available such as from AERONET, the retrieval results will be collocated with the validation data sets, and their difference will be compared with the error propagation uncertainty to quantify the retrieval performance. Furthermore, the fitting residuals will be evaluated to inform the sufficiency of the forward model and uncertainty model.

1. **Reference**

* Gao, M., K. Knobelspiesse, B. A. Franz, P.-W. Zhai, B. Cairns, X. Xu, and J. V. Martins. 2023. "The impact and estimation of uncertainty correlation for multi-angle polarimetric remote sensing of aerosols and ocean color." Atmospheric Measurement Techniques, 16 (8): 2067–2087 [https://doi.org/10.5194/amt-16-2067-2023]
* Gao, M., K. Knobelspiesse, B. Franz, P.-W. Zhai, A. Sayer, A. Ibrahim, B. Cairns, O. Hasekamp, Y. Hu, V. Martins, J. Werdell, and X. Xu. 2022. "Effective uncertainty quantification for multi-angle polarimetric aerosol remote sensing over ocean." Atmos. Meas. Tech., 15 (16): 4859–4879 [10.5194/amt-15-4859-2022]
* Gao, M., K. Knobelspiesse, B. A. Franz, P.-W. Zhai, V. Martins, S. P. Burton, B. Cairns, R. Ferrare, M. A. Fenn, O. Hasekamp, Y. Hu, A. Ibrahim, A. M. Sayer, P. J. Werdell, and X. Xu. 2021. "Adaptive Data Screening for Multi-Angle Polarimetric Aerosol and Ocean Color Remote Sensing Accelerated by Deep Learning." Frontiers in Remote Sensing, 2: [10.3389/frsen.2021.757832]
* Gao, M., B. A. Franz, K. Knobelspiesse, P.-W. Zhai, V. Martins, S. Burton, B. Cairns, R. Ferrare, J. Gales, O. Hasekamp, Y. Hu, A. Ibrahim, B. McBride, A. Puthukkudy, P. J. Werdell, and X. Xu. 2021. "Efficient multi-angle polarimetric inversion of aerosols and oceancolor powered by a deep neural network forward model." Atmospheric Measurement Techniques, 14 (6): 4083–4110 [https://doi.org/10.5194/amt-14-4083-2021]