RC2: 'Review', Anonymous Referee #2, 15 Jun 2022

The Authors discuss a method to evaluate theoretical uncertainties in retrieved aerosol and ocean surface properties from multi-angle polarimetric remote sensing measurements. Through a Monte Carlo sampling approach, they propose a method to compare theoretical uncertainties with observed observed-true parameter differences computed from validation. The method is applied to synthetic as well as to real AirHarp measurements. In addition, the Authors briefly discuss a way to speed up a posteriori uncertainty calculations by analytic differentiation of a neural network-based forward model.

The topic is very interesting and within the scope of AMT, the analyses appear sound and convincing. The only slight criticism that I have is that I found the paper rather long and not always easy to follow. A large amount of detailed analyses are presented, and the task of deciding what are the most important points of the study appears to be mostly left to the reader. I wonder if the information can be synthesised a bit to make the paper easier to read. Furthermore, it seems to me that the aspects of "performance evaluation" and "speed improvement" are a bit intermixed when it comes to the way the sections of the manuscript are organized, which does not help readability. Given that it seems to me that the "performance evaluation" aspect of the manuscript is given a much wider space than the "speed improvement",

We appreciate your time and efforts in reviewing this work. The comments and questions are valuable in improving the clarity of this manuscript. The responses are below (in red) with manuscript revised accordingly (also uploaded).

I wonder whether it would be better to move this latter to an appendix. Apart from this, I think this is an excellent paper. Below are a few other minor comments:

Thank the reviewer for the positive comments. Following the reviewer's suggestion, we have moved section 3 to Appendix with details provided below.

- L113: A more recent reference is

Fougnie, B. et al. (2018), "The multi-viewing multi-channel multi-polarisation imager – Overview of the 3MI polarimetric mission for aerosol and cloud characterization", JQSRT, 23-32, doi: 10.1016/j.jqsrt.2018.07.008

Thank you for suggesting the reference. It has been updated.

- L156-157. Does your forward model covers the entire HARP spectral range?

If so, is it not unrealistic to assume spectrally flat refractive index?

Especially dust aerosols are way more absorbing in the UV than in the VIS/NIR.

For HARP/AirHARP, there are four bands: 440, 550, 670 and 870nm. The refractive index spectral are quite flat within this spectral range as we have been used in our previous study (Gao et al, 2021, AMT). However, UV channels are available for OCI and SPEXone sensors on PACE, which will require a more sophisticated refractive index spectra as the reviewer pointed out.

- L173-175. "The forward calculation of aerosol size etc.".

I suggest rephrasing as... "The forward calculation of aerosol optical depth (AOD)

and single scattering albedo (SSA) from aerosol size and refractive index"

Thank you for the suggestion. We have revised the sentence as

> "The forward calculation of aerosol optical depth (AOD) and single scattering albedo (SSA) from aerosol size and refractive index is also performed using NNs based on simulations using the numerical code based on the Lorenz-Mie theory (Mishchenko et al 2002)"

- L192. "interior method". Do you mean "interior point method"?

Our use of "interior method" is to follow the naming convention in the algorithm paper (Branch et al 1999). The algorithm uses the 'affine scaling' which ensure that the solution is searched strictly within the interior of the feasible region. It seems "interior point method" is often used to refer the algorithm with a barrier function.

We also added more information about the STIR method as follows:

> "The subspace trust-region interior reflective (STIR) algorithm is employed to conduct non-linear least-square minimization of the cost function (Branch et al., 1999). **Its implementation in the Python package SciPy is used in this study (Virtanen et al., 2020).** STIR is based on the Levenberg-Marquardt algorithm combined with an interior method and reflective boundary technique (Branch et al., 1999). **The interior method ensure that the retrieval parameters are searched strictly within the interior of the feasible region as specified in Table 1, while the reflection technique can significantly reduce the number of iterations in the minimization process.**"

- Personally I think section 3 breaks the flow of the paper a bit. While the paper

is mostly about illustrating how theoretical uncertainties compare with observed errors,

here you discuss a technical detail of how to speed up theoretical uncertainty calculations.

Wouldn't it be better to have this as an appendix?

We have further synthesized the manuscript by moving the "speed improvement" section to the appendix A. At the end of Sec 3.1, we added a sentence on the use of automatic differentiation, and leave all the details in Appendix:

> **"Automatic differentiation is used to calculate both the Jacobian matrix and the derivatives defined in Eq. 8 for AOD and SSA, as well as water leaving signals involving atmospheric correction and BRDF correction. More details are discussed in Appendix A."**

Please note that the title is also updated:

> "Effective uncertainty quantification for multi-angle polarimetric aerosol remote sensing over ocean"

- L413. A ratio of 1.5 means a 50% difference. Not sure I would regard this as a "slight underestimate"

Thank you for pointing this out. We have revised the sentence as follows

> "The ratios are mostly in the range 1-1.5, indicating that the theoretical uncertainties work well to represent the real retrieval uncertainties in most cases but are generally underestimates."