Response to RC2

Concerning the introduction

I am missing previous results of (UV-) MFRSR comparisons with other standard AOD measuring instruments. For example in the 2015 the Filter Radiometer comparison in Davos, Switzerland various of this instrument types have participated and the results have been discussed. There are also earlier studies of such comparisons.

In Kazadzis et al., 2018: There were 4 MFRSR instruments in this campaign. The results were summarized in the following paragraph:

"The four MFR instruments showed good agreement for the medians compared to the PFR triad, however, they exhibit larger scatter than the sun-pointing instruments resulting in a lower precision. McArthur et al. (2003) had previously reported that the MFR-derived AOD does not quite meet the accuracy of the sun-pointing instruments under clean atmospheric conditions. MFR_DE showed an AOD overestimation in various instances that gave results that are outside the WMO defined AOD limits (Fig. 2d). This small overestimation of the MFR_DE instrument compared to the PFR Triad could be due to uncertainties introduced while correcting for their angular response, the calibration procedure, or incomplete blocking of the diffuser by the shadow-band. The MFRSRs that are part of the SURFRAD network (MFR_US2 and MFR_US3) give a median AOD at 500-nm that is in very good agreement with the PFR triad and as good or better than some of the other sun-pointing instruments, e.g., CIMEL and POM; these two slightly underestimate the AOD at 865 nm, but are within the WMO defined limits. Again, these two MFRs' medians are comparable to the better sun-pointing instruments, but give larger scatter. These two MFRs are representative of the SURFRAD network that follow network protocols for calibration and alignment and frequent characterizations of the spectral and angular responses (Augustine et al., 2003, Michalsky et al., 2001)."

Would be helpful some of the above aspects to be included in the introduction section or in section 2.4

We have added the following sentences in section 2.4 (lines 214 to 222):

During the recent Fourth Filter Radiometer Comparison held in Davos, Switzerland (between 28 September and 16 October 2015), most AOD values derived from the three AERONET CIMEL sunphotometers are within the ±0.01 range compared with the PFR triad standard (Kazadzis et al. 2018). This includes those determined at 368nm from the extrapolation of AERONET AODs at 340nm and 380nm. The 2015 Davos campaign also included four MFRSR instruments. Overall, the results showed good agreement between the four MFRSRs and the PFR triad standard, though one instrument exhibited a positive bias and low precision compared to the sun-pointing instruments (Kazadzis et al. 2018). However, such errors were likely explained by instrument-specific uncertainties (e.g. angular response correction, responsivity calibration, and shadowband position issues) and do not suggest inherent error in MFRSR AODs (Kazadzis et al. 2018).

Authors are mentioning: “There are no other instruments measuring at 368nm”
The World Meteorological Organization (WMO) instigated the Global Atmosphere Watch (GAW) program in 1989. Based on a recommendation by GAW experts, the World Optical Depth Research Calibration Center (WORCC) was established in 1996 at the PMOD/WRC in Switzerland. WORCC has since been advised by the GAW Scientific Advisory Group for Aerosols. The standard instrument consist of a precision Filter radiometer (PFR measuring at 368, 412, 500 and 862 nm. So actually the WMO reference instruments (triad) is measuring at 368nm.
So the argument of the non existence of instruments measuring at 368nm (thus the choice of the AOD based comparison) is not correct. However, as it is not possible to repeat this study with one of the PFR instruments, probably the only solution could
be the AOD comparison with the collocated cimels as the authors have initiated. Nevertheless, a short comment on the above text could be included in the paper.

Thanks for pointing out the existence of the WMO reference instruments that measure at 368 nm. We have changed the beginning sentence of section 2.4 to the following (lines 199 to 205).

Ideally, to avoid additional uncertainties caused by the interpolation between wavelengths, the calibration factors should be validated via a direct comparison of direct sun signals from the to-be-calibrated UV-MFRSR and a reference instrument measuring at the 368 nm channel (e.g. the standard precision Filter radiometer (PFR) operated by the Physikalisches-Meteorologisches Observatorium Davos, World Optical Depth Research Calibration Center (WORCC)). However, such reference measurements are not available at most UVMRP stations. Therefore, the estimated mean normalized Vo (Vo_norm) values from the Gaussian Process regression and the other two comparison methods (i.e. MA and OPER) are validated indirectly in terms of aerosol optical depth (AOD) against those obtained from the collocated AERONET sites.

Line 224 : probably you should comment also that ozone is also ignored .

We changed the word “ozone” to “O_3” to make it clear ozone is also ignored.

Line 235: I guess that the cloud flagging method is not evaluated here, as comparison with CIMEL data includes only data that CIMEL algorithm considers as cloud free.

Yes, we did not evaluate the cloud flagging method in this study. We mainly relied on AERONET’s cloud flagging for selecting cloud-free UV-MFRSR measurements. We did perform a simple variation check using similar methodologies to those found in many cloud screening algorithms (e.g. Alexandrov et al. 2004) to reduce the potential contamination of aerosol optical depth by broken clouds.

Line 250 : It would be informative to explain why S(lamda) appears in equation 9.

Since F(lamda) is ~4nm the integrated range is ~366-370 nm. There you mention that AOD is the "interpolated ADO spectrum" which I guess you mean the linear (?) interpolated using 340 and 380 AERONET AODs ? Then S(lamda) is used for normalization in this small 4 nm range? Is this so different that the actual interpolated value of AOD at 368 nm? And if S(lamda) is used, why not S(lamda) – Rayleigh optical depth ? Have in mind that spectral function FWHM of the CIMEL is larger than 2nm.

We agree that the S(lambda) in equation 9 has minimal effects on deriving the UVRSR 368 band passed AERONET AOD. For example, for Mauna Loa, Hawaii site, the mean (2.6×10^{-6}) and standard deviation (2.5×10^{-6}) of the difference between the AOD with and without S(lambda) are several magnitudes smaller than the instrument resolution. Therefore, we removed the S(lambda) term in equation 9 for simplicity (lines 266 to 271). The change is so small that it doesn't impact further analysis in the manuscript.

Comparing the AOD directly interpolated to 368 nm with the bandpass AOD (calculated with updated equation 9) for Mauna Loa, Hawaii site, we found that there was a ~0.1% discrepancy. The small discrepancy could be explained by the small difference between the effective wavelength of the instrument (367.91 nm for the MLO UV-MFRSR) and 368 nm. With larger wavelength differences for some instruments, we expect a larger discrepancy. Therefore, we decided to keep using the more accurate bandpass AOD in the manuscript.

The authors chose to evaluate their method by comparing the retrieved (from their Vos data) AODs. Here are my comments on this section:

- The comparison of UVMFR with AERONET would be essential to follow criteria that are defined by WMO –CIMO in order to assess the results in detail.

https://library.wmo.int/pmb_ged/wmo-td_1287.pdf

There (page 8) such conditions and formulas are defined.

For example the U95 criterion where a number (here the lower limit is 95%) of measurements have to be in the range of ±0.005 + 0.010/m

Where 0.005 accounts for instrument related uncertainties and 0.01/m for calibration related uncertainties (calibration uncertainty better than 1%).

Changing the analysis figures with the use of this criterion authors can:

a. better show the agreement and the improvements with their methods by showing the percentage of data within these limits for each case.

b. having in mind that calibration related uncertainties will be inherited in AOD retrievals as a function of air mass, the figures (a) including aod differences vs air mass can point out on Vo related issues.

Still slopes and cor. Coefficients can be reported in the form of a table.

Thank you for pointing out the U95 criterion. We understand that complying with such criterion is critical for achieving traceability of AOD products generated from radiometer measurements.

In this study, the comparison of AOD values derived from UV-MFRSR measurements and AERONET AOD values only serves as an indirect evidence that the calibration of UV-MFRSR is reasonably accurate. However, we do not attempt to argue that the calibration of UVMRP UV-MFRSR is accurate enough to produce AOD values that meet the U95 criterion. Nor do we attempt to argue that UVMRP AOD products are traceable to the AERONET or WMO AOD standard. As mentioned in the manuscript and the citations therein, the stability assumption of the Langley method may not be strictly fulfilled at many UVMRP sites, rendering larger uncertainty of Langley calibration factors at these sites compared with those derived under ideal conditions (such as at Mauna Loa, Hawaii).

This conclusion is supported by the following analysis. Based on the GP uncertainty results (using 1.96 instead of 4.42 standard deviations to mimic 95% CI, data not shown in the manuscript) among the three test sites, only at HI02 (Mauna Loa, Hawaii), the AOD uncertainty (at 95% level) caused by calibration (~0.0078) is lower than the U95 criterion (0.01). The AOD uncertainties (at 95% level) caused by calibration at the other two sites are much larger than 0.01 (i.e. ~0.043 for IL02, and ~0.028 for OK02).

We recognize that the assumption that AERONET AOD represents “ground-truth” is not ideal, however, as indicated by previous work, it is a reasonable assumption for this study. AERONET sunphotometers are routinely calibrated with the uncertainty of AOD around 0.002 to 0.005 in the visible and up to 0.01 in the UV region (Eck et al., 1999; Holben et al., 2001). Additionally, during the recent Fourth Filter

Radiometer Comparison held in Davos, Switzerland (between 28 September and 16 October 2015), most AOD values derived from the three AERONET CIMEL sunphotometers are within the ±0.01 range compared with the PFR triad (Kazadzis et al. 2018). This includes those determined at 368nm from the extrapolation of AERONET AODs at 340nm and 380nm. Therefore, we are confident that the AERONET AOD products used in this study are accurate enough to be a reliable source of AOD values for validation purposes. This assumption is consistent with numerous other field-based evaluations of radiometric AOD accuracy as detailed between lines 213 and 214 in the manuscript. To emphasize the validity of using AERONET as an effective standard, we have revised the following text in the manuscript summarizing the results from the Fourth Filter Radiometer Comparison and Kazadzis et al. (2018) (lines 199 to 222).

Ideally, to avoid additional uncertainties caused by the interpolation between wavelengths, the calibration factors should be validated via a direct comparison of direct sun signals from the to-be-calibrated UV-MFRSR and a reference instrument measuring at the 368 nm channel (e.g. the standard precision Filter radiometer (PFR) operated by the Physikalisches-Meteorologisches Observatorium Davos, World Optical Depth Research Calibration Center (WORCC)). However, such reference measurements are not available at most UVMRP stations. Therefore, the estimated mean normalized Vo (Vo_norm) values from the Gaussian Process regression and the other two comparison methods (i.e. MA and OPER) are validated indirectly in terms of aerosol optical depth (AOD) against those obtained at the collocated AERONET sites. We admit that the uncertainty of UV-MFRSR AODs could exceed the World Meteorological Organization (WMO) U95 criterion (e.g. 95% of the measured data have uncertainty in the range of 0.005 ± 0.01 / airmass, Kazadzis et al. 2018) at many UVMRP sites because the stability assumption of the Langley method may not be strictly fulfilled. Therefore, the AOD comparison in this study can only serve as an indirect evidence to verify whether the calibration of UV-MFRSR is reasonably accurate.

AERONET sunphotometers are routinely calibrated with the uncertainty of AOD around 0.002 to 0.005 in the visible and up to 0.01 in the UV region (Eck et al., 1999; Holben et al., 2001) and are therefore considered a reliable source for AOD intercomparison and radiometer validation [e.g. (Alexandrov et al., 2002, 2008;Augustine et al., 2003;Krotkov et al., 2005a;Krotkov et al., 2005b;Kassianov et al., 2007;Tang et al., 2013;Yin et al., 2015;Zhang et al., 2016)]. During the recent Fourth Filter Radiometer Comparison held in Davos, Switzerland (between 28 September and 16 October 2015), most AOD values derived from the three AERONET CIMEL sunphotometers are within the ±0.01 range compared with the PFR triad standard (Kazadzis et al. 2018). This includes those determined at 368nm from the extrapolation of AERONET AODs at 340nm and 380nm. The 2015 Davos campaign also included four MFRSR instruments. Overall, the results showed good agreement between the four MFRSRs and the PFR triad standard, though one instrument exhibited a positive bias and low precision compared to the sun-pointing instruments (Kazadzis et al. 2018). However, such errors were likely explained by instrument-specific uncertainties (e.g. angular response correction, responsivity calibration, and shadowband position issues) and do not suggest inherent error in MFRSR AODs (Kazadzis et al. 2018).

The AOD retrieval and the differences among two instruments are a consequence not only an uncertainty on the instrument calibrations but also other factors.
Here is a list:
- The calculation of Rayleigh optical depth from both instruments including the pressure measurement. Are the two instruments (UVMFR and CIMEL) use the same formulas ?
- The calculation of Rayleigh and aerosol air mass factors
- The potential differences in the field of view of the instrument
- CIMEL includes NO2 and Ozone optical depths
- The wavelength interpolation from 340 and 380 nm to 368 nm is not by definition linear but aerosol type related.
So in order to assess their results the authors at least have to mention the related uncertainties and the above issues raised by using retrieved AODs from two different instruments with different instrument characteristics and post processing AOD algorithms and procedures, in order to validate the Vos.
In theory a direct comparison of direct sun signals for the UVMFR instrument and a reference instrument measuring at 368nm could be used in order to assess the differences in the Vo calculation, without having the AOD calculation related uncertainties.

We agree that the discrepancy/uncertainties in deriving AOD values from the two instruments' measurements should be highlighted explicitly. The following sentences were added at the end of section 2.4 (lines 273 to 288).

Since AERONET and UV-MFRSR AOD values at 368 nm are derived from measurements involving different instruments and wavelengths, the uncertainties when comparing these AOD values should be noted. Some important sources of uncertainties include:

1) AERONET calibration error – At the time of calibration at MLO, AERONET reference instruments have an uncertainty of ~0.2 to 0.5%, which is equivalent to a 0.002 to 0.005 uncertainty in AERONET AOD (Holben et al. 2001). These calibration factors are likely to shift within the year following calibration, which may result in a total AOD uncertainty of ~0.01 to 0.02 (wavelength dependent, higher in the UV) (Holben et al. 2001).
2) Instrument Field of View (FOV) - AERONET CIMELs have a field-of-view (FOV) of 1.2° while the UV-MFRSR has a larger FOV (e.g. ~6.5°, reported by Kazadzis et al. 2018). AODs obtained from instruments with larger FOVs are associated with greater AOD uncertainty due to larger contributions of scattered light to the direct irradiance measurement (Kim et al. 2005).
3) Instrument maintenance – Periodic soiling and cleaning of the UV-MFRSR diffuser can result in spurious increases and decreases in AOD, respectively. The frequency of on-site maintenance (e.g. cleaning of the UV-MFRSR dome) as well as rainfall events may therefore account for some of the AOD difference (Kim et al. 2005; Kim et al. 2008).
4) Trace gases - As mentioned above, AERONET AOD accounts for NO_2 optical depth (e.g. ~0.002 at OK02) while UV-MFRSR AOD does not.

We do not include the Rayleigh optical depth formula, airmass formula, and interpolation methods in the list.

For the Rayleigh optical depth (RLOD) calculation, both UV-MFRSR and AERONET (version 2, https://aeronet.gsfc.nasa.gov/version2_table.pdf) use the same formula described in Bodhaine et al. 1999. The instantaneous pressure values for UV-MFRSR are obtained from the collocated AERONET measurements. Therefore, RLOD should not introduce additional uncertainty in this study.

For the airmass calculation, both UV-MFRSR and AERONET (version 2, https://aeronet.gsfc.nasa.gov/version2_table.pdf) use the same formula described in Kasten and Young 1989.

For the interpolation between 340 and 380 nm AOD values, we agree that the spectra of aerosol optical depth is aerosol type related and may not be strictly linear (e.g. slightly quadratic). However, we believe that the difference in the interpolated AOD spectrum among different interpolation methods/equations should be negligible because the two wavelengths are so close (e.g. Figure 6 in Krotkov et al. 2005a).