This document contains the authors' responses to the Editor's comments. The Editor's comments are in black, followed by the author's responses in blue.

The new version is significantly improved in relation to the original.

- 5 The authors seem to use the terms OMI aerosol record and OMAERUV interchangeably. For readers unfamiliar with the OMI aerosol products, there may be some confusion, since officially two OMI aerosol records are available. I would suggest the authors to either stick to the 'OMAERUV record' on every instance they refer to it, or include an initial statement stating that the expressions OMAERUV and OMI aerosol record are used Interchangeably throughout in the manuscript.
- 10 Thank you for your correction. We have added the 'It is also noted that there are two official OMI aerosol level 2 products though, the OMI measurements in this paper only refers to the OMAERUV product.' (line 255-257)

A few specific comments on the current version are listed below.

Line 22. Replace 'acceptable' with 'encouraging'

15 We have changed it accordingly. 'The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations.' (line 21-23)

Line 34: Just saying that λ_0 denotes the reference wavelength does not provide useful information. It should be stated that this is the longer wavelength in the pair, and it is used to calculate an assumed wavelengthindependent-scene-reflectivity, needed for calculation of the of the $I_{-\lambda}$ Rayleigh term in Equation (1).

We have added more introduction of the wavelength pair of UVAI. 'λ is the wavelength where the radiance difference between a Rayleigh and a measured scene is calculated, and λ₀ is the longer wavelength where a 25 spectrally constant scene reflectivity is assumed for the calculation of I^{Ray}.' (line 34-36)

Line 42. In addition to the listed aerosol related parameters, the magnitude of the observed UVAI also depends on non-aerosol-related factors such as spectral dependence of land surfaces and ocean color effects, sun-glint effects, and cloud effects. This clarification should be included.

30 This issue was raised in the earlier review to which the authors reply with the statement '... the radiance itself is also affected by surface conditions, clouds, atmospheric gases and aerosols. On the other hand, the UVAI only contains the information of aerosol absorption...'. Their reply reveals a poor understanding of the UVAI concept, since the UVAI is just the measured radiance conveniently packed in a single parameter. It includes any observed spectral dependence other than Rayleigh

35 scattering.

- Thank you for your correction. We have modified our presentation accordingly. 'Although non-aerosol factors exist, such as spectral dependence of the surface, ocean color, sun-glint and cloud contamination, <u>the most</u> dominant are aerosol concentration, aerosol vertical distribution and aerosol optical properties (Wang et al., 2012; Buchard et al., 2017).' (line 44-46)
- 40
 - Line 151. An explanation for the selection criteria UVAI > 1 and CF larger than 0.3 should be added. Is this the MODIS CF? If so, 0.3 represents significant contamination.

Sorry for the confusion. The cloud fraction here is the TROPOMI FRESCO cloud fraction. We may forget to add it from the last version manuscript. We have added its source in the manuscript just after introduction of TROPOMI ALH product (1t the same band, there is TROPOMI FRESCO cloud support product providing).

- 45 TROPOMI ALH product. 'At the same band, there is TROPOMI FRESCO cloud support product providing cloud fraction (CF) for mitigating cloud effects as will be explained later (https://scihub.copernicus.eu last access: 19_Sept_2018) (Apituley et al., 2017).' (line 139-141)
- The explanation for the pixel selection criteria has been added also. 'Before implementing radiative transfer calculations, pre-processing excludes pixels with large solar zenith angle ($\theta_0 \ge \underline{70^\circ}$), weak aerosol absorption (<u>UVAI_{354,388}</u> < 1), insignificant aerosol amount (<u>AOD550</u> < 0.5) or cloud contamination (CF > <u>0</u>.3).' (line 153-155)

Line 183. Replace 'better' with 'closer to AERONET retrieved SSA'. As stated AERONET SSA does not 55 represent the true SSA value. We have changed it accordingly. 'Although our retrieved SSA seems closer to AERONET retrieved SSA than that provided in OMAERUV, one should keep in mind that there is only one record for this event, the meteorological conditions, combustion phases and even the aerosol compositions may change during the 3-hour time difference. (line 186-189)

60

Line 245. Although the same wavelength pair (354, 388) is used in the OMAERUV UVAI definition, the way the observed radiances are calculated have changed [Torres et al., 2018]. The new definition handles the effect of clouds using of Mie Theory (instead of the Lambertian approximation), and reduces significantly the across-track angular dependence. It also accounts for wavelength dependence of surface reflectance as well as sun glint over

- 65 the oceans. For cloud-free conditions, the Mie-UVAI is slightly larger than LER-UVAI (~0.3) for nadir-viewing conditions. The original definition (consistent with that in TROPOMI) is still reported in the OMAERUV as a parameter labeled 'RESIDUE'. Thus, the original definition should be used for algorithm training purposes. This is important because the difference between the LER and Mie definitions increases with viewing sensor solar zenith angle.
- 70 Thank you for your correction. We have re-collected the OMAERUV (as we did not extract the residue from the original files at the first place) and re-collocated it to the AERONET data set using the same criteria.

We have added explanations in the manuscript. 'Note that the UVAI used here is the 'residue' field in the original OMAERUV product, where the simulated radiance $(l_{\lambda}^{Ray} in Eq.(1))$ is calculated by a simple Lambertian 75 approximation that is consistent with TROPOMI UVAI (Torres et al., 2018).' (line 272-274)

The following heatmap gives a quick view of how the UVAI and residue correlates with other parameters. The correlation between the residue and AOD, AAOD and ALH is relatively higher compared with other parameters, but not significantly high (at or less than the level of 0.5) compared with UVAI. As a result, we use a stricter **80** criterion that only keeping OMAERIUN pixels with cloud fraction smaller than 0.1. This enhance the correlation

80 criterion that only keeping OMAERUV pixels with cloud fraction smaller than 0.1. This enhance the correlation between UVAI and AOD, AAOD and ALH (r = 0.66, 0.66 and 0.4, respectively). There is total 5679 samples in the training data set. We have the following changes in the manuscript:

'OMI pixels with θ_0 larger than 70° or cloud fraction larger than 0.1 are excluded.' (line 260-261) 85

'In total 5679 samples are obtained.' (line 266)

'The empirical ratio between a training and testing data set is 70% versus 30%, thus there are 3975 samples in the training data set and 1704 samples in the testing data set,' (line 289-290)

But be noted that in the case applications the input TROPOMI data is still using CF < 0.3 to ensure there is collocated pixels with AERONET sites. In the heatmap, the residue seems has a higher dependence on the solar zenith angle and the surface albedo, too. But we may only focus on the aerosol related features. We have the following changes in the manuscript:

'The UVAI also has a dependence on θ_0 , but in this study we only focus on the aerosol related features.' (line 281-282)

100 'Note that the data includes pixels with CF larger than 0.1 in order to ensure there is satellite measurements collocated with the AERONET sites (though CF is no larger than 0.3).' (line 354-356)

'The input features are selected by the *Spearman's rank* correlation coefficients and a priori knowledge on relationship between UVAI and only aerosol related features.' (line 439-440)

105

90

95

'Other non-aerosol features affecting UVAI should be also taken into consideration.' (line 445-446)

(Deleted:
-(Deleted: 1,
Ì	Deleted: or UVAI354.588 smaller than 0.8, or pixels with extreme high ALH but low UVAI
(Deleted: 4003

Deleted: b

Deleted: The training data set containing the original OMAERUV ALH and the adjusted ALH are referred to as the original and adjusted training data set, respectively. We fit the SVR for AAOD prediction to both training data sets in order to investigate the importance of a reliable ALH input.

Deleted: Pearson



- 120 The new training data set leads to small changes in the SVR model hyper-parameter C, which is 0.08 (used to be 0.11). The accuracy of the updated SVR is slightly lower than that of the previous one (with UVAI in the training data set). The updated SVR tends to predict higher SSA than the previous one, which is as expected, because the Mie-UVAI is overall 0.3 larger than the residue. The higher UVAI (if the ALH and AOD keep the same), the lower predicted SSA as we described in the sensitivity study in Section 3.5. From the case applications, 5 out of 9
- 125 of the predicted SSA records is within +/- 0.03 difference compared with the AERONET records (used to be 66%), and not all predictions are within +/- 0.05 difference (88%).

We also have updated the Figure.6-15 and figures in Appendix B, and results in Table 3 and 4 accordingly.

Line 246. In the description of the OMAERUV record, it should be stated that AOD retrievals have been
 validated using the multi-year AERONET aerosol record [Ahn et al., 2014], and, similarly, the SSA parameter
 have been evaluated with AERONET Almucantar retrievals [Jethva, et al., 2014].

We have added them accordingly. 'Its AOD was validated by the multi-year AERONET record (Ahn et al., 2014) and its SSA was evaluated by AERONET Almucantar retrievals (Jethva et al., 2014).' (line 252-254)

- Line 426 'Acceptable' is a subjective interpretation. Replace 'acceptable (+/- 0.02)' with '+/- 0.02' 135 We have changed it accordingly. 'The accuracy of SVR-predicted SSA is ± 0.02 , with higher tendency to
- overestimate the SSA for relatively absorbing cases.' (line 441-442)

Cited References

- https://doi.org/10.5194/amt-11-2701-2018
 Torres, O., Bhartia, P. K., Jethva, H., and Ahn, C.: Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products, Atmos. Meas. Tech., 11, 2701-2715, 2018.
 - Jethva, H., O. Torres, and C. Ahn (2014), Global assessment of OMI aerosol single-scattering albedo using ground-based AERONET inversion, J. Geophys. Res. Atmos., 119, doi:10.1002/2014JD021672.

10.1002/2013JD020188Ahn, C., O. Torres, and H. Jethva (2014), Assessment of OMI near-UV aerosol optical depth over land, J. Geophys. Res. Atmos., 119, 2457–2473, doi:.

 Jyunting Sun^{1,2}, Pepijn Vecfkind^{1,2}, Swadhin Nanda^{1,2}, Peter van Velhoven³, Pieternel Levelt^{1,2} ¹Department of Satellite Observations, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ¹Department of Geoscience and Remote Sensing (GRS), Civil Engineering and Geosciences, Delft Diversity of Technology, Ddft, 2620 CD, the Netherlands ¹Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ¹Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ¹Correspondence to: Jyunting Sun (ijyunting sum@knmi.nl) Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA tore maintring the UVAI difference between observed ones and that simulations, a significant gap method to retrieve SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectru dependence may cause severe misinterpretations of aerosol absorption neetral data demonstrate the postorial eagend measurements significant gap to extore regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a SSA for some extorem situations. The apatian contrast in SSA retrievely basching cases and fails to prediate SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to oversetimate the SSA for relatively absorbing cases and fails to prediated Fourit. Dis , Not Retal, ferent absorb in gaves of anotal absorption in TWAI records. ¹ Introduction ² The concept of the near-ultraviolet (near-UV) absorbing aerosol al		from ultraviolet satellite observations	Formatted: Font: (Default) Times New Roman
 Department of Satellite Observations, Royal Netherlands Meteorological Institute, Dc Bilt, 3731 GA, the Netherlands Department of Conseince and Remote Seming (GRS), CVII Engineering and Geosciences, Defit University of Technology, Defit, 2628 CD, the Netherlands Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, Dc Bilt, 3731 GA, the Netherlands Correspondence to: Flyunting Sun (ijyunting, sungikami.al) Abstract. The purpose of this study is to demonstrate the role of acrosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet staellite observations for biomass burning acrosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model. With the recently released ALH product 15-5P TROPOMI constraining forward simulations, as significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations as a significant gap priori assumptions on aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements psing the support vector ergression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol absorption. In the second part of this paper, we propose an alternative matcher into externe situations. The spatial contrast in SSA for relatively absorbing acrosol absorption free WVR. Filter 2000 SR and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tuncts to externe situations. The spatial contrast in SSA for model with the ator SVR by severtime situations. The spatial contrast in SSA for model with the orone provectif d he name-ultraviolet (near-UV) absorbing acrosol absorption fre		Jiyunting Sun ^{1,2} , Pepijn Veefkind ^{1,2} , Swadhin Nanda ^{1,2} , Peter van Velthoven ³ , Pieternel Levelt ^{1,2}	
 ¹ Department of Geoscience and Remote Seming (GRS), Civil Engineering and Geosciences. Deff University of Technology, Deff, 2628 CD, the Netherlands ¹ Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ¹ Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ¹ Correspondence to: Jiyunting Sun (jiyunting, sun@kmi.nl) Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model, With the recently released ALH product of S-SP TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer isimulations with spectral dependence may cause severe misintepretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated statellite and ground-based measurements ging the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a forto solution. The spatial contrast in SSA for relatively absorbing earos and absorption so aerosol absorption and exolutions is significantly. ¹ bigher than that of SVR, and the latter better agrees with SSA for multify a clonade at encouse. Moreover, the high resolution TROPOMI UVAI and collocated ALH product swill guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. ¹ Introduction ¹ The concert of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing ae	0	¹ , Department of Satellite Observations, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands	Formatted: Font: (Default) Times New Roman
 Delft University of Technology, Delft, 2628 CD, the Netherlands ¹ Department of Wenther & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands ² Correspondence to: Jiyuuting Sun (jiyuuting.sun@kmi.il) Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model, With the recently released ALH product of 5-5P TROPOMI constraining forward simulations, as gainficant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated stallitie and ground-based measurements jusing the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a 5P transformed in recent years. For all cases, the difference between SVR-retrieved SSA and ALRONET are generally within ±0.05, with over half of samples are within ±0.05. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA for melERA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Marcover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data		² Department of Geoscience and Remote Sensing (GRS), Civil Engineering and Geosciences,	Formatted: Font: (Default) Times New Roman
 ¹ Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands Correspondence to: Jyunting Sun (ijyunting sun@knni.nl) Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieved SSA (0.25) is found between role of x-5P TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated stellite and ground-based measurements gaing the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated stellite and ground-based measurements gaing the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol absorption. The results are encouraging, though at the current base, the model tends to overestimate the SSA for relatively absorbing cases and fails to prefict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly high at not of NCR, and the latter better agrees with SSA for MERA-2 renalysis. In the future, more sophisticated faiture selection procedure and kerned functions should be taken into consideration to improve the SVR mored is fuely for moctor: Aduo,		Delft University of Technology, Delft, 2628 CD, the Netherlands	
 Correspondence to: Jyuming Sun (jiyuming.sun@knmi.nl) Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model. With the recently released ALH product of 5-5P TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a for some serier exists abapened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model lends to varestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA form MERA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. Ilutroduction<		³ Department of Weather & Climate Models, Royal Netherlands Meteorological Institute, De Bilt, 3731 GA, the Netherlands	Formatted: Font: (Default) Times New Roman
Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model. With the recently released ALH product of S-SP TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (025) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosol, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a provint assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encourging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly by higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure ad kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorptio	5	Correspondence to: Jiyunting Sun (jiyunting.sun@knmi.nl)	
scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model. With the recently released ALH product of S-9 TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent earosols, implying that inappropriate assumptions on aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a for priori assumptions on aerosol and record prevents thappened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly by higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a		Abstract. The purpose of this study is to demonstrate the role of aerosol layer height (ALH) in quantifying the single	
retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer model, With the recently released ALH product of S-5P TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA form MERRA-2 reanalysis. In the future, more sophisticated fature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records.		scattering albedo (SSA) from ultraviolet satellite observations for biomass burning aerosols. In the first experiment, we	
model, With the recently released ALH product of S-5P TROPOMI constraining forward simulations, a significant gap in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated stellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR Interduction Formatted: Heading 2, Line spacing: 1.5 lines Deleted: 1 Immoduction! <td></td> <td>retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer</td> <td></td>		retrieve SSA by minimizing the UVAI difference between observed ones and that simulated by a radiative transfer	
 in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a i priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. Ilntroduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detexts elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosp		model. With the recently released ALH product of S-5P TROPOMI constraining forward simulations, a significant gap	
spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements asing the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. I Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference, between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Wittin the out the value of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers wavele		in the retrieved SSA (0.25) is found between radiative transfer simulations with spectral flat aerosols and strong	
cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. 1 Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Utitie: tota of in a Rayleigh atmosphere (Herman et al., 1997): Utitie: tota of in a given wavelength pair (λ and λ_o) (1)		spectral dependent aerosols, implying that inappropriate assumptions on aerosol absorption spectral dependence may	
method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997):		cause severe misinterpretations of aerosol absorption. In the second part of this paper, we propose an alternative	
support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ± 0.05 , with over half of samples are within ± 0.03 . The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. I Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere [Herman et al., 1997): The conc to μh_{μ}^{abs} , $\mu = \mu h_{\mu}^{abs}$, $\mu = \frac{h_{\mu}^{abs}}{abs}$, $\mu = \frac{h_{\mu}^{abs}}{$		method to retrieve SSA based on long-term record of collocated satellite and ground-based measurements using the	Deleted: and ALH
priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ± 0.05 , with over half of samples are within ± 0.03 . The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. Interduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): With the tot of the magnet of t		support vector regression (SVR). This empirical method is free from the uncertainties due to the imperfection of a	
SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and AERONET are generally within ± 0.05 , with over half of samples are within ± 0.03 . The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. 1 Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): With $t = top (t_1, t_2)^{abs}$, $t = top (t_2)^{abs}$		priori assumptions on aerosol micro-physics as that in the first experiment. We present the potential capabilities of the	
AERONET are generally within ±0.05, with over half of samples are within ±0.03. The results are encouraging, though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. 1 Introduction Free concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Uttrate top (the spectral contrast difference (the spectr		SVR by several fire events happened in recent years. For all cases, the difference between SVR-retrieved SSA and	
 though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. <u>1 Introduction</u> The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Witting to a time with ^{abs} at an with ^{Ray}. 		AERONET are generally within ± 0.05 , with over half of samples are within ± 0.03 . The results are encouraging,	
SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. <u>1 Introduction</u> The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): UNITY = tab $(h_{in} e^{h_{in}})$ (1) Field Code Changed		though at the current phase, the model tends to overestimate the SSA for relatively absorbing cases and fails to predict	
 higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. <u>I Introduction</u> Formatted: Heading 2, Line spacing: 1.5 lines Deleted: 1 Introduction¹ Formatted: Formatted: Formatted		SSA for some extreme situations. The spatial contrast in SSA retrieved by radiative transfer simulations is significantly	
sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. I Introduction The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): UNIT of the output of the code Changed Field Code Changed Field Code Changed		higher than that of SVR, and the latter better agrees with SSA from MERRA-2 reanalysis. In the future, more	
 model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. <u>1 Introduction</u> Formatted: Heading 2, Line spacing: 1.5 lines The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): With a tot (l = dl), ^{abs} = l = dl), ^{Ray}. 		sophisticated feature selection procedure and kernel functions should be taken into consideration to improve the SVR	
reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records. <u>1 Introduction</u> Formatted: Heading 2, Line spacing: 1.5 lines The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Deleted: Deleted: Deleted: Uttate tato (la - (la), ^{abs}) (la), ^{Bag}) (1) Field Code Changed		model accuracy. Moreover, the high resolution TROPOMI UVAI and collocated ALH products will guide us to more	
1 Introduction Formatted: Heading 2, Line spacing: 1.5 lines The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): Deleted: Introduction* With a conc (I = - (I), S ^{obs} = I = - (I), S ^{Ray}) (1)		reliable training data set and more powerful algorithms in quantify aerosol absorption from UVAI records.	
The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed <u>radiance</u> in a real atmosphere and a model simulated one <u>in a</u> Rayleigh atmosphere (Herman et al., 1997): UVAL $= 400 \text{ (} h e^{-f_{1}} e^{\frac{h}{2}} e^{\frac{h}{2}} e^{-f_{1}} e^{\frac{h}{2}} e^{\frac{h}{2}} e^{-f_{1}} e^{\frac{h}{2}} e^{\frac{h}{2}} e^{-f_{1}} e^{\frac{h}{2}} e^{\frac{h}{2}} e^{-f_{1}} e^{\frac{h}{2}} e^{-f_{1}} $		<u>1 Introduction</u>	Formatted: Heading 2, Line spacing: 1.5 lines
product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): $\frac{d_{10}}{d_{10}} \frac{d_{10}}{d_{10}} \frac{d_{10}}{d_{10}} \frac{d_{10}}{d_{10}} \frac{Ray}{d_{10}}$ (1) Formatted: Font: 10.5 pt, Not Bold, Font color: Auto, (Chinese (China), (Other) English (US) Deleted: Deleted: for a given wavelength pair (λ and λ_0) Field Code Changed		The concept of the near-ultraviolet (near-UV) absorbing aerosol index (UVAI) initially came along with the ozone	Deleted: 1 Introduction
layers by measuring the spectral contrast difference between a satellite observed radiance in a real atmosphere and a model simulated one in a Rayleigh atmosphere (Herman et al., 1997): $\frac{1}{1}$ (1) (Chinese (China), (Other) English (US) (Deleted: Deleted: Deleted: Deleted: (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)		product of the Nimbus 7/Total Ozone Mapping Spectrometers (TOMS). It detects elevated UV-absorbing aerosol	Formatted: Font: 10.5 pt, Not Bold, Font color: Auto, (Asia
model simulated one in a Rayleigh atmosphere (Herman et al., 1997): $\frac{Deleted: \text{ for a given wavelength pair } (\lambda \text{ and } \lambda_0)}{Field Code Changed}$ (1)		layers by measuring the spectral contrast difference, between a satellite observed radiance in a real atmosphere and a	Chinese (China), (Other) English (US)
(1) Field Code Changed		model simulated one in a Rayleigh atmosphere (Herman et al., 1997):	Deleted: for a given wavelength pair (λ and λ_0)
$OVAI = -100 \left(log_{10} \left(\frac{\pi}{l_{\lambda 0}} \right) - log_{10} \left(\frac{\pi}{l_{\lambda 0}} \right) \right)$		$UVAI = -100 \left(log_{10} \left(\frac{I_{\lambda}}{I_{\lambda 0}} \right)^{obs} - log_{10} \left(\frac{I_{\lambda}}{I_{\lambda 0}} \right)^{Ray} \right) $ (1)	Field Code Changed

, where the superscript obs and Ray denote the radiance from observations and that from simulations, respectively. I_{λ} and $I_{\lambda 0}$ are the radiance at wavelength λ and λ_0 . λ is the wavelength where the radiance difference between a Rayleigh

- 185 and a measured scene is calculated, and λ_0 is the longer wavelength where a spectrally constant scene reflectivity is assumed for the calculation of I_{λ}^{Ray} . Positive UVAI indicates the presence of absorbing aerosols, whereas negative or near zero values imply non-absorbing aerosols or clouds (Herman et al., 1997). The over four-decade UVAI observations (1978 to present) has been widely used for aerosol research. It would be beneficial to derive aerosol absorption properties from the long-term global UVAI records, e.g. the single scattering albedo (SSA), which is the
- 190 ratio of aerosol scattering to aerosol extinction. Aerosols are considered as the largest error source in radiative forcing assessments (IPCC, 2014), and SSA is one of the key parameters to reduce this uncertainty (Haywood and Shine, 1995),

The magnitude of UVAI depends on many factors (Herman et al., 1997; Torres et al., 1998; Hsu et al., 1999; de Graaf et al., 2005). Although non-aerosol factors exist, such as spectral dependence of the surface, ocean color, sun-glint and

- l95 cloud contamination, the most dominant are aerosol concentration, aerosol vertical distribution and aerosol optical properties (Wang et al., 2012; Buchard et al., 2017). To derive SSA from UVAI, the information on other two parameters are necessary. The aerosol concentration is usually provided in terms of the aerosol optical depth (AOD). There are plentiful AOD products with wide spatial-temporal coverage. By contrast, there is much less information on the aerosol vertical distribution. The most well-known aerosol vertical distribution product is provided by the Cloud-
- 200 Aerosol Lidar with Orthogonal Polarization (CALIOP), but the number of measurements is limited because of its narrow tracks (Winker et al., 2009), Passive sensors make efforts on retrieving the aerosol layer height (ALH) from columnar measurements. For example, Chimot et al. (2017) present the feasibility of ALH retrieval using the OMI oxygen band at 447 nm. Tilstra et al. (2018) developed algorithm to derive absorbing aerosol layer height from GOME-2 FRESCO cloud layer height products. Xu et al. (2017; 2019) attend to retrieve ALH from EPIC oxygen

205 absorption bands for dust and carbonaceous layers over both land and ocean surfaces. Recently a new ALH product has been run operationally, based on the measurements at near-Infrared (NIR) oxygen Aband of the TROPOspheric Monitoring Instrument (TROPOMI) on board the Copernicus Sentinel-5 Precursor (S-5P) Sanders et al., 2015). TROPOMI has a wide swath of 2600 km, providing daily global coverage with a spatial resolution of 7×3.5 km² in nadir. The instrument is equipped with both the UV-visible (270-500 nm) and the near-

210 infrared (NIR) (675–775 nm) channels, which can simultaneously provide UVAI and the collocated ALH product (Veefkind et al., 2015)

The purpose of this paper is to demonstrate the role of the ALH in quantifying aerosol absorption from UVAI using newly released TROPOMI level 2 ALH product. At current phase we only focus on the biomass burning aerosols. Two experiments are conducted. First, following previous studies (Colarco et al., 2002; Hu et al., 2007; Jeong and Hsu,

- 215 2008; Sun et al., 2018), we build up lookup tables (LUTs) of simulated UVAI for various aerosol optical properties by radiative transfer models (RTMs). Then SSA is derived by minimizing the difference between pre-calculated UVAI and satellite observed ones. The major uncertainties in the retrieved SSA are caused by assumptions on the wavelength-dependent refractive index and the availability of reliable aerosol vertical distribution information (Sun et al., 2018). Now with the operational TROPOMI ALH constraining forward simulations, it is expected to partly reduce
- 220 the SSA retrieval uncertainty meanwhile quantifying the influence of assumed aerosol properties on the retrieved SSA. Although the availability of ALH in radiative transfer calculations can improve the SSA retrieval, assumptions on aerosol micro-physics remain inevitable. In the second experiment, we therefore propose an empirical method to predict aerosol absorption, based on the long-term records of collocated UVAI, ALH, AOD and absorbing aerosol

Field Code Changed Formatted: Font: 10.5 pt

Formatted: Font: 10.5 pt Field Code Changed

Field Code Changed

Moved (insertion) [1]

Deleted: As the TROPOMI ALH product is not operationally available yet, we focus on the data of one of the largest wildfires that happened in southern California in 2017, i.e. the Thomas Fire incidents/incider ndex/1922). Ignited on 4 December 2017, the fire was expanded quickly northwest by the strong and persistent Santa Ana winds and was fully under control of 12 January 2018. The precise cause of the fire remains unknown, but a prolonged period of heat and absence of precipitation definitely contributed to this devastating fire (https://inciweb.nwcg.gov/incident/5670/). We selected one day (12 December 2017) for our case study. As shown in Fig.1, a brow smoke plume produced by the Thomas Fire was blown away from the continent and transported northwards. The major part of the plum was over the ocean with cloud free conditions, which is favorable for space-borne aerosol observations.

Formatted: Default Paragraph Font, Font: 12 pt, English (UK) Formatted: Default Paragraph Font, Font: 12 pt, English (UK) Deleted:

Field Code Changed

Deleted: The most straightforward approach to derive a relationship between UVAI and quantitative aerosol absorption properties like SSA is through forward radiative transfer simulations Lookup tables (LUTs) of simulated UVAI for various measuring geometries, acrosol properties, atmospheric and surface conditions are constructed by radiative transfer models (RTMs). Then SSA is derived by minimizing the difference between pre-calculated UVAI and satellite observed ones (Colarco et al., 2002; Hu et al., 2007; Jeong and Hsu, 2008; Sun et al., 2018). Hereafter, we refer to this method as the RTM-based retrieval.

Formatted: Font: (Default) Times New Roman, 10.5 pt

- 250 optical depth (AAOD) using machine learning (ML) techniques. ML algorithms learn the underlying behavior of a system from a given training data set. They are particularly useful to address ill-defined inversion problems in the field of geosciences and remote sensing, where theoretical understanding is incomplete but there is a significant amount of observations (Lary et al., 2015). We employ ML techniques in order to avoid explicit assumptions on aerosol microphysics as made in the first experiment. By now, the ALH observations are not abundant and we will use the ALH
- accompanied in the AOD retrieval from the OMAERUV product in the training procedure (Torres et al., 2013). Nevertheless, the recent TROPOMI ALH and other future ALH products make such empirical methods of great potentials. Various ML algorithms have been developed to deal with classification or regression problems. In this paper we choose the support vector regression (SVR), a regression variant form of the support vector machines (SVM) (Drucker et al., 1997). Compared with other algorithms (e.g. the Artificial Neural Network), SVR is less sensitive to
- 260 training data size and can successfully work with limited quantity of data (Mountrakis et al., 2011; Shin et al., 2005). We will present the capability to retrieve SSA from UVAI of using this empirical method with multiple case studies. This paper is organized as follows: the first experiment is outlined in section 2, with description on setting radiative transfer simulations, and the analysis of the uncertainty trigger by the assumption on aerosol absorption spectral dependence; Section 3 starts with introduction of SVR, followed by training data set preparation, SVR model hyper-
- 265 <u>parameter tuning, error analysis</u> and case applications. <u>Finally, the major conclusions and implications for future</u> research are summarized in section 4.

2 Experiment 1: SSA retrieval using radiative transfer simulations

In this section, we present the first experiment that retrieves SSA by radiative transfer calculations as done in previous studies (Colarco et al., 2002; Hu et al., 2007; Jeong and Hsu, 2008; Sun et al., 2018). Forward radiative transfer

- 270 simulations are realized by the KNMI developed radiative transfer model DISAMAR (Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval) (de Haan, 2011). Fig.1 illustrates the model inputs and the procedure. For each pixel, first, aerosol optical properties are computed by Mie theory for various pre-defined aerosol models. Then DISAMAR calculates UVAI using the corresponding satellite information: AOD, ALH, the solar zenith angle (θ_0), the viewing zenith angle (θ_v), the solar azimuth angle (φ_0), the viewing azimuth angle (φ_v), surface
- albedo (A_s) and surface pressure (P_€) of the target pixel. The output of the forward simulations is a LUT of UVAI as a function of the input SSA (determined by the pre-defined aerosol models), which is fit by a second order polynomial function. Finally, by specifying the corresponding satellite observed UVAI, the SSA of the target pixel is estimated from the UVAI-SSA relationship. The retrieved SSA is reported at 500 nm in order to compare with the results of the SVR method. Section 2.1 will introduce the input parameters in radiative transfer simulations, followed by retrieval results in section 2.2.

2.1 Radiative transfer simulation setup

2.1.1 Aerosol models

The aerosol models used for the Mie calculations are a combination of the aerosol models in ESA Aerosol_cci project (Holzer-Popp et al., 2013) and that in the OMAERUV algorithm (Torres et al., 2007; Torres et al., 2013). We assume a fine mode smoke aerosol type and further divide it into 7 subtypes as listed Table 1. We use the particle size distribution of the fine mode strongly absorbing aerosol of ESA Aerosol_cci project. The geometric radius (r_a) is 0.07

 μm (effective radius r_{eff} of 0.14 μm) and the geometric standard deviation (σ_q) is 1.7 (logarithm variance $ln\sigma_q$ of

Field Code Changed

Apart from pre-assumed aerosol micro-physics, the aerosol loading and the aerosol vertical distribution are two key parameters in forward simulations of UVAI. The former is usually provided in terms of the aerosol optical depth (AOD). There are plentiful AOD products providing wide spatial-temporal coverage with various spectral choices. By contrast, only little information on the aerosol vertical distribution is available. The most well-known aerosol profile product is offered by the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), but the number of measurements is limited because of its narrow tracks (Winker et al., 2009). Passive sensors only measure columnar quantities but, in some cases, also provide the aerosol layer height (ALH), a compact form of aerosol profile information indicating where most aerosols are located. Chimot et al. (2017) present the feasibility of ALH retrieval using the oxygen (O₂-O₂) band at 447 nm of the Ozone Monitoring Instrument (OMI), but so far it has not been run operationally vet. ¶

So far it has not been run operationally yet. International (CHR), due so far it has not been run operationally yet. International (CHR), due so far it has been developed for the TROPOSpheric Monitoring Instrument (TROPOMI) on board the Copernicus Sentinel-5 Precursor (S-SP) (Sanders et al., 2015). TROPOMI was launched on 13 October 2017. The instrument is equipped with both the UV-visible (270–500 nm) and the near-infrared (NIR) (675–775 nm) channels (Veefkind et al., 2015), which makes it possible to interpret UVAI using corresponding ALH measurements. Furthermore, TROPOMI has a wide swath of 2600 km, providing daily global coverage with a high spatial resolution of 7×3.5 km² in nadir.⁴ The purpose of this paper is to demonstrate the potential of the TROPOMI ALH product for quantifying aerosol absorption.

Moved up [1]: As the TROPOMI ALH product is not operationally available yet, we focus on the data of one of the largest wildfires that happened in southern California in 2017, i.e. the Thomas Fire

(http://www.fire.ca.gov/current_incidents/incident/details/Index/1922). Ignited on 4 December 2017, the fire was expanded quickly northwest by the strong and persistent Stanta Ana winds and was fully under control on 12 January 2018. The precise cause of the fire remains unknown, but a prolonged period of heat and absence of precipitation definitely contributed to this devasating fire (<u>https://inciweb.nwcg.gov/incident/5670/</u>). We selected one day (12 December 2017) for our case study. As shown in Fig.1, a brown smoke plume produced by the Thomas Fire was blown away from the continent and transported northwards. The major part of the plume was over the ocean with cloud free conditions, which is favorable for space-borne aerosol observations.[§]

Formatted: Default Paragraph Font, Font: 12 pt, English (UK) Formatted: Default Paragraph Font, Font: 12 pt, English (UK)

Deleted: We conduct two experiments to investigate the potential of TROPOMI ALH for quantifying aerosol absorption of the smoke plume generated by the Thomas Fire. First, as concluded in our previous study (Sun et al., 2018), the absence of aerosol vertical distribution information and an improper spectral dependence of aerosol absorption in the near-UV region can be responsible for a large difference between estimated and measured SSA. Now with the TROPOMI ALH as a constraint, we are able to quantitatively determine the influences of assumed wavelength dependent aerosol absorption on retrieved SSA. Similar to our previous study (Sun et[1]

Deleted: introduc

Deleted: es the data sets involved in this stud

Deleted: y

Deleted: the first part of section 3 expresses the setup of the RTMbased method and the first experiment, i.e. a sensitivity study of SSA examining assumptions on the spectral dependence of near-UV aerosol absorption; the second part of section 3 focuses on the secong

Deleted: Data sets

Formatted: Font: 10.5 pt
Formatted: Normal, Line spacing: 1.5 lines
Formatted: Font: 10.5 pt

Formattear Fond: 10.5 p

Formatted: Font: 10.5 pt

Formatted: Heading 2, Line spacing: 1.5 lines

0.53). The real part of the refractive index (n) uses the same value as in the OMAERUV algorithm, which is set to be 1.5 for all subtypes and spectrally flat. We adopt the imaginary part of the refractive index at 388 nm (κ_{388}) of the OMAERUV smoke subtypes (except for BIO-1 whose κ_{388} is 0) in our study and add a subtype with κ_{388} equaling to 405 0.06.

Many studies have shown evidence that absorption by biomass burning aerosols in the near-UV band has a strong spectral dependence (Kirchstetter et al., 2004; Bergstrom et al., 2007; Russell et al., 2010). Accordingly, a constant $20\% \Delta \kappa$ has been applied to all smoke subtypes in the recent OMAERUV algorithm (Jethva and Torres, 2011), where $\Delta \kappa$ is defined as the relative difference between κ_{354} and κ_{388} (i.e. $\Delta \kappa_{\mathbf{k}} = (\kappa_{354} - \kappa_{388})/\kappa_{388}$). In this experiment, we

- 410 will investigate how the retrieved SSA responds to the assumed spectral dependence by considering 9 different $\Delta \kappa$ values from 0% (i.e. 'grey' aerosols) to 40% (very strong spectral dependence). This corresponds to an Absorbing Ångström exponent (α_{abs}) from 1 to 3.4 and from 1.3 to 4.7, depending on aerosol subtype. Note that the $\Delta \kappa$ is only applied between κ_{354} and κ_{388} . As we only investigate the influence due to aerosol absorption spectral dependence in near-UV range in this study, aerosol absorption at wavelengths larger than 388 nm is set equal to that at 388 nm.
- 115 To summarize, the first experiment consists of 9 cases represented by different Δκ. Within each case, there are 7 predefined aerosol subtypes with varying κ_{388} . Thus, 63 forward simulations are performed for each individual pixel.

2.1.2 Inputs from satellite

Fig.1 presents the parameters input for the radiative transfer simulations of UVAI. Satellite measurement geometries $(\theta_0, \theta_n, \varphi_0 \text{ and } \varphi_n)$ and the surface pressure (P_s) accompanied with the TROPOMI UVAI reprocessed product

- 420 (https://scihub.copernicus.eu last access: 8 June 2018) are input for the forward simulations. The TROPOMI UVAI is calculated for two different wavelength pairs. One uses the conventional 340 and 380 nm to continue the heritage of UVAI records from multiple sensors, and the other uses 354 and 388 nm in order to allow comparison with OMI measurements (D.C. Stein Zweers, 2016). In this study we employ the 354 and 388 nm pair. TROPOMI ALH is retrieved at oxygen A-band (759-770 nm), where the strong absorption of oxygen causes the highly
- 425 structed spectrum (https://scihub.copernicus.eu last access: 22 June 2018). This feature is particularly suitable for elevated optically dense aerosol layers (Sanders et al., 2015; Sanders and de Haan, 2016). The ALH is reported in both altitude and pressure. For the forward radiative transfer calculations, the input aerosol profile is parameterized according to the settings in ALH retrieval algorithm: a one-layered box shape profile, with central layer height derived from TROPOMI and an assumed constant pressure thickness of 50 hPa (Sanders and de Haan, 2016). At the same
- 430 band, there is TROPOMI FRESCO cloud support product providing cloud fraction (CF) for mitigating cloud effects as will be explained later (https://scihub.copernicus.eu last access: 19_Sept_2018) (Apituley et al., 2017). The TROPOMI AOD product has not been operational, thus we use AOD from the Level 2 product MYD04 (Collection 6) of Aqua MODIS (http://dx.doi.org/10.5067/MODIS/MYD04 L2.006, last access: 17 July 2019). Aqua has an overpass time similar to S-5P (13:30 local time). The AOD at 550 nm used in the RTM-based method is a
- 435 combination of the Deep Blue Aerosol Optical Depth 550 Land and the Effective Optical Depth Op55um Ocean (Levy et al., 2013).

The surface albedo that used to retrieve TROPOMI UVAI is currently not available in the product. Instead, we use the Aura/OMI Level 3 Lambertian equivalent reflectance (LER) monthly climatology calculated from measurements between 2005 and 2009 (Kleipool et al., 2008) (Kleipool, 2010) (http://dx.doi.org/10.5067/Aura/OMI/DATA3006, last

440 access: 26 September 2018). TROPOMI on S-5P and OMI on Aura have similar overpass times (13:30 local time) and measuring geometries (Levelt and Noordhoek, 2002) (Veefkind et al., 2015).

Field Code Changed

Formatted: Font: 10.5 pt

Formatted: Heading 2, Line spacing: 1.5 lines

 Due to different spatial resolutions, TROPOMI ALH, OMI LER climatology and MODIS AOD are resampled onto the <u>TROPOMI UVAI grid. Before implementing radiative transfer calculations, pre-processing excludes pixels with large</u> solar zenith angle (θ₀.> <u>7</u>0°), weak aerosol absorption (<u>UVAI₃₅₄₃₈₈</u> < 1), insignificant aerosol amount (<u>AOD₅₅₀</u> < 0.5) or cloud contamination (CF > <u>0</u>.3).

2.2 SSA retrieved by radiative transfer simulations

In the first experiment, we focus on one of the largest fire events that happened in southern California in 2017, i.e. the Thomas Fire (http://www.fire.ca.gov/current_incidents/incidentdetails/Index/1922). Fig.A1 in Appendix shows the RGB plume captured by MODIS on 12 December 2017. A brown smoke plume produced by the Thomas Fire was

- 450 blown away from the continent and transported northwards. The major part of the plume was over the ocean and under cloud free condition, which is favorable for space-borne aerosol observations. There are totally 5217 pixels in this case. Fig.2 presents the UVAI, ALH and AOD after pre-processing. The highest UVAI appeared at the south part of the plume, where both aerosol loading and aerosol layering are relatively high (AOD > 2 and ALH is over 2.5 km). Fig.3a displays the mean SSA of all plume pixels retrieved by the RTM-based method as a function of Δκ. The
- 455 retrieved aerosol absorption decreases with Δκ. This finding is in good agreement with Jethva and Torres (2011). 'Gray' aerosols require stronger absorption to reach the same level of UVAI than 'colored' aerosols. This also explains the high SSA standard deviation (filled area) in the cases with little or no spectral dependence in aerosol absorption. The large variability in retrieved SSA (from 0.69±0.13 to 0.94 ±0.03) demonstrates that inappropriate assumptions on the spectral dependence of near-UV aerosol absorption may significantly bias interpretations of smoke aerosol
- 460 absorption and should be carefully handled in forward radiative transfer calculations. The retrieved aerosol absorption is compared with the nearby the version 3 level 1.5 AERONET inversion product (https://aeronet.gsfc.nasa.gov last access: 4 June 2019). Only one site is within 50 km from TROPOMI plume pixels (UCSB, (119.845°W,34.415°N)) with only one record for this case. The SSA at 500 nm at 18:54:47 UTC is 0.98 (sky radiance error 15.8%), which is nearly 3 hours ahead of TROPOMI overpass. There are 15 TROPOMI collocated
- 465 pixels to <u>UCSB</u> with distance within 50 km and time difference within 3 hours. <u>Hereafter we call them AERONET-collocated pixels</u>. As illustrated in Fig.3b, the mean SSA of the collocated pixels also increases with Δκ and eventually levels off at around 0.96. The extremely low SSA and high variation (0.57±0.25) retrieved for 'gray' aerosols prove that the spectral independence assumption is not recommended for smoke aerosols. The differences between the mean SSA of the collocated pixels and the AERONET measurement are shown in Fig.3c.

470 The retrieved SSA starts falling inside the uncertainty range of AERONET (±0.03) (Holben et al., 2006) when Δκ is 25%, where the plume SSA is 0.90±0.05 and the AERONET-collocated SSA is 0.96±0.02 (Table 2). Table 2 also presents the SSA accompanied in AOD retrieval from OMAERUV version 3 product (http://dx.doi.org/10.5067/Aura/OMI/DATA2004 last access: 17 October 2018). OMI pixels are collocated to the AERONET site in the same way as TROPOMI. The SSA of the OMAERUV-AERONET collocated pixels is 0.06

- 475 lower than that of AERONET, which indicates a 20% spectral dependence of aerosol absorption in OMAERUV algorithm may be not sufficient for this case. Although our retrieved SSA seems closer to AERONET retrieved SSA than that provided in OMAERUV, one should keep in mind that there is only one record for this event, the meteorological conditions, combustion phases and even the aerosol compositions may change during the 3-hour time difference.
- 480 Fig.4 presents the spatial distribution of retrieved AAOD and SSA when Δκ is 25%, which shows a strong heterogeneity in the horizontal direction. The plume center is most absorbing, where SSA is even less than 0.70. The

Formatted: Line spacing: 1.5 lines, Tab stops: Not at 8.85

Field Code Changed

Field Code Changed

Field Code Changed

SSA gradually increases when the plume transported northwards. SSA is expected to be low near source flaming regions (Eck et al., 1998; Eck et al., 2003; Eck et al., 2013) while SSA may become higher when aerosols age during transport (Reid et al., 2005; Lewis et al., 2009). The strong spatial variability in SSA is mainly controlled by the

485 heterogeneity of the UVAI (Fig.3a) through the one-to-one numerical relationship. This relationship may differ from one pixel to another as the algorithm focuses on one-pixel retrieval each time. Depending on the combustion phase and meteorological conditions, heterogeneity in aerosol properties is expected for plume of this size. Nevertheless, whether such a large SSA difference 0f 0.38 (maximum SSA – minimum SSA, Table 2) is reasonable needs further investigations (discussed in Section 3.6.3).

490 <u>3 Experiment 2: SSA retrieval using support vector regression</u>

In this section, we propose an empirical method to derive SSA as an alternative of the radiative transfer simulations presented in the first experiment. The motivation is that the assumptions on aerosol micro-physics in forward simulations are inevitable, whereas our knowledge to them is inadequate (particularly the aerosol absorption spectral dependence). An inappropriate assumption may lead to significant bias in retrieved SSA (Fig.3). On the other hand,

495 SVR (and other ML algorithms) is applicable to solve ill-posed inversion problems by learning the underlying behavior of a system from a given data sets without such a priori knowledge on aerosol micro-physics. In this paper, we construct an SVR model with UVAI, AOD and ALH as input features and AAOD as the output, then derive the SSA by the following relationship:

 $SSA = 1 - \frac{AAOD}{C}$

(2) <

The procedure of SVR prediction is presented in Fig.5. We start with a brief introduction of the SVR algorithm,
 followed by input feature selection (section 3.2), training and testing data set preparation (section 3.3), SVR model
 hyper-parameters tuning (section 3.4), error analysis (section 3.5) and case applications (section 3.6).

3.1 Support vector regression

 $y = r_{1}(x) + \delta_{\lambda}$

SVR (Drucker et al., 1997) is the regression variant of SVM, a supervised non-parametric statistical algorithm initially devised by Cortes and Vapnik (1995). SVM algorithm is suitable to solve problems of small training data sets with a
high-dimensional feature space and can provide excellent generalization performance (Durbha et al., 2007; Yao et al., 2008), which has been applied extensively to solve remote sensing problems (Lary et al., 2009; Mountrakis et al., 2011; Noia and Hasekamp, 2018). The basic ideal of SVM in classification problems is finding an optimal hyperplane in a high-dimensional feature space maximizing the margin between the two classes to minimize misclassifications (Durbha et al., 2007). The same principle is applied to regression problems, SVR attends to find an optimal hyperplane

510 that maximizes the margin of tolerance in order to minimize the prediction error. The error within the margin does not contribute to the total loss function, while samples on the margin are called support vectors. For the detailed mathematical formulation of SVR algorithm one can refer to Smola and Scholkopf (2004). Briefly, given the training data with *n* observations {(x₁, y₁), (x₂, y₂), ..., (x_n, y_n)}, assuming the statistical model as the following:

(3)

515 , where x is a multivariate input and y is a scalar output with length $n_{\cdot} \delta$ is the independent zero mean random noise. The input x is first mapped onto a feature space with dimension of m by a non-linear transformation, then a linear model f(x) is constructed based on it:

Moved (insertion) [2]

Moved up [2]: UVAI offline data can be accessed via http://doi.org/10.5270/SSP-0wafvaf). The TROPOMI UVAI is calculated for two different wavelength pairs. One uses the conventional 340 and 380 nm to continue the heritage of UVAI records from multiple sensors, and the other uses 354 and 388 nm in order to allow comparison with OMI measurements (D.C. Stein Zweers, 2016). In this study we retrieve the SSA based on the latter pair. Satellite measurement geometries (solar/viewing zenith angle θ_0/θ_{nv} , solar/viewing azimuth angle ω_0/ψ_{nv} and the sufface pressure (P₃) included in the UVAI product are input for the radiative transfer calculations. The scene albedo ($A_{a,v}$) from the same product is also used in the pre-processing as will be described later. ¶ TROPOMI ALH retrieval is based on the pattern of a highly

Structured spectrum with strong absorption of O₂ in the A-band (759–770 nm), which is particularly suitable for elevated optically dense aerosol layers (Sanders et al., 2015; A.F.J. Sanders and J.F. de Haan, 2016). The TROPOMI ALH data is still in a pre-operational phase of development and only internally available (last access: 22 June 2018). The ALH is reported in both altitude and pressure. The ALH data for this fire event is available for pixels with UVAI values (calculated for the wavelength pair of 340 and 380 nm) larger than 1 to exclude pixels dominated by non-absorbing aerosols. For the forward radiative transfer calculations, the input aerosol profile is parameterized as a one-layered box shape profile, with central layer height derived from TROPOMI and an assumed constant pressure thickness of 50 hPa."

2.2 MODIS satellite data

The corresponding AOD of the plume generated by the Thomas Fire is taken from the Level 2 product MYD04 (Collection 6) from

Deleted: UVAI offline data can be accessed via http://doi.org/10.527/0/SSP-0wafvaf). The TROPOMI UVAI is calculated for two different wavelength pairs. One uses the conventional 340 and 380 nm to continue the heritage of UVAI records from multiple sensors, and the other uses 354 and 388 nm in order to allow comparison with OMI measurements (D.C. Stein Zweers, 2016). In this study we retrieve the SSA based on the latter pair. Statellite measurement geometries (solar/viewing zenith angle $\theta_i/\theta_{i\nu}$, solar/viewing azimuth angle $\varphi_i/\phi_{i\nu}$ and the utransfer calculations. The scene albedo ($A_{i\nu}$) from the same product is also used in the pre-processing as will be described later. ¶

Data sets listed in this section are either used by the RTM-based method or the SVR-based method, or both. The pre-processing and detailed usage of those data are explained in section 3. <u>...[3]</u> Deleted: Methodology

Formetted, Default Davages h

Tornatted. Deladit raragraph ront
Formatted: Default Paragraph Font, English (UK)
Formatted: Font: (Default) Times New Roman
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Moved (insertion) [4]
Formatted: Right, Line spacing: 1.5 lines
Deleted: This section introduces the procedure and technical concerns of the RTM-based method and the SVR-based methodThe
Formatted: Heading 2, Line spacing: 1.5 lines
Formatted: English (US)
Formatted: English (US)
Formatted: (Asian) Chinese (China), (Other) English (US)
Formatted: English (US)
Formatted: English (US)
Field Code Changed
Deleted: major stars of SVR consist of feature selection training

Deleted: major steps of SVR consist of feature selection, training and testing data preparation, hyper-parameters tuning and application. We first collect parameters relevant for the derivation pf]

Formatted: Font: 10.5 pt, Italic

(4)

$$f(x) = \sum_{i=1}^{n} a_{i} b_{i}(x) + b$$
(4)

$$f(x) = \sum_{i=1}^{n} a_{i} b_{i}(x) + b$$
(5)

$$f(x) = b_{i}(x) + b_{i}($$

satellite product containing a long-term UVAI, AOD, SSA and corresponding <u>ALH</u> (Torres et al., 2007; 2013). Its AOD was validated by the multi-year AERONET record (Ahn et al., 2014) and its SSA was evaluated by AERONET Almucantar retrievals (Jethva et al., 2014). The ALH is the best-guessed either from CALIOP climatology or assumed ALH in the retrieval (if the CALIOP climatology is not available) (<u>Torres et al., 2013</u>). As a result, one should keep in

780 mind that the ALH from OMAERUV may suffer from the uncertainties of CALIOP climatology and a priori assumptions, and collocation error between OMI pixels and CALIOP footprint. It is also noted that there are two official OMI aerosol level 2 products though, the OMI measurements in this paper only refers to the OMAERUV product.

We collect the measurements of OMAERUV and AERONET from 2005-01-01 to 2017-12-31, OMI pixels with ρ_0

- 785 larger than 70^o_{ky} or cloud fraction larger than 0.1 are excluded. Then OMI observations are considered as collocated with an AERONET site if their spatial distance is within 50 km and their temporal difference is within 3 hours. To ensure consistency between the different measurement techniques (ground-based and space-borne), we also exclude samples if the SSA difference between OMAERUV and AERONET is larger than 0.03, or the AOD difference between OMAERUV and AERONET is larger than 5%. The AERONET SSA and AAOD are linearly interpolated to
- 790 500 nm as OMAERUV reports them at this wavelength. In total 5679 samples are obtained. Fig.B1 in the Appendix shows the global distribution of the collocated OMAERUV-AERONET samples. Note that these samples are not restricted to biomass burning areas, but may also contain other aerosol types.

3.2.2 Feature selection

The <u>OMAERUV</u>_eAERONET joint data set consists of following parameters: UVAI calculated by 354 and 388 nm wavelength pair, satellite geometries, surface albedo, surface pressure and ALH from <u>OMAERUV</u>_e and SSA, AOD and AAOD from AERONET. Note that the UVAI used here is the 'residue' field in the original OMAERUV product, where the simulated radiance $(I_{\lambda}^{Ray}$ in Eq.(1)) is calculated by a simple Lambertian approximation that is consistent with TROPOMI UVAI (Torres et al., 2018). Fig.6 presents the <u>Spearman's rank</u> correlation coefficients matrix (ρ) of those parameters. It is clear that except for AAOD, SSA is barely associated with other parameters. The correlation between UVAI and SSA is rather low ($\rho = -0.25$). On the other hand, AAOD is highly associated with UVAI ($\rho =$

0.66) as well as AOD (ρ = 0.66) as it carries information on both aerosol absorption and aerosol loading. Therefore, it is preferred to predict AAOD from given UVAI and derive SSA via in Eq. (2) afterwards rather than to directly predict SSA from UVAI. Besides, as mentioned previously, AOD and ALH are the major factors influencing UVAI, which is also reflected by the relatively stronger correlation (ρ = 0.4). Consequently, we construct an SVR model with UVAI, ALH and AOD as the input features, and AAOD as the output. The UVAI also has a dependence on θ₀, but in this

study we only focus on the aerosol related features.

3.3. Preparing training and testing data sets

The SVR model is trained and tested base on the <u>OMAERUV-AERONET joint data set</u> containing <u>8616</u> samples as <u>810</u> described in the last section (consisting of UVAI, ALH from OMAERUV, and AOD, AAOD from AERONET). We further partition it into a training and a testing data set, respectively. The testing data set is used to evaluate the generalization performance of an SVR model trained by training data set, in order to avoid high bias (underfitting) or high variance (overfitting) problems. The empirical ratio between a training and testing data set is 70% versus 30%, thus there are 3975 samples in the training data set and 1704 samples in the testing data set.

Deleted: measurements	
Deleted: as described in section 2	_
Formatted: Font: 10.5 pt	
Formatted: Font: 10.5 pt	
Deleted: or UVAI _{354,388} smaller than 0.8, or pixels with extreme high ALH but low UVAI	
Deleted: an	_
Deleted: pixel	
Formatted: Font: 10.5 pt	
Deleted: ,	
Deleted: 1,	_
Deleted: is collocated	_
Deleted: are also excluded	
Deleted: OMI	_
Deleted: OMI	
Deleted: 4003	
Deleted: Supplement	
Deleted:	
Deleted: OMI	
Deleted: OMI	
Deleted: absolute values,	
Deleted:	
Deleted: decided	
Deleted: and AOD	
Deleted: to	
Deleted: the relationship	
Deleted: ¶ SSA = $1 - \frac{AAOD}{AOD}$	[6]
Moved up [4]: SSA = $1 - \frac{AAOD}{AOD}$	

Deleted: 2.2 Formatted

Deleted: As described in the previous section, the selected features consist of UVAI, ALH, AOD (inputs) and AAOD (output). A good training data set is crucial to ML algorithms, because they learn the underlying behavior of the system from it. One problem we are facing in this study is that the ALH from the OMAERUV product may not have sufficient quality to use. The influence of OMI ALH can be seen by comparing the relationship between 3 input features from TROPOMI-MODIS measurements (Fig.7a) and that from OMI-AERONET measurements (Fig.7b). One may recall that the OMI ALH is climatological whereas the TROPOMI ALH is directly derived from observations. As a result, we propose to create a 'more observational' ALH to replace the OMI ALH in the original OMI-AERONET joint data set. We numerically adjust the OMI ALH, to [7] Formatted: Font: 10.5 pt

Å	Formatted: Font: 10.5 pt
Å	Formatted: Font: 10.5 pt
Á	Formatted: Font: 10.5 pt
λ	Formatted: Font: 10.5 pt
-(Formatted: Font: 10.5 pt
(Deleted: data set
-(Formatted: Font: 10.5 pt
-(Deleted: The rule of thum
-(Deleted: b
1	

Deleted: . The training data set containing the original OMAERUS

900 3.4 SVR hyper-parameters tuning

As described in section 3.1, the generalization performance and model accuracy of the SVR depends on the following hyper-parameters: (1) the width of insensitive zone ε . The cost function does not consider errors in the training data as 905 long as their deviation to the truth is smaller than ε ; (2) the regularization constant C that determines the trade-off between model complexity and the degree to which deviations larger than ε are penalized; (3) choice of the kernel and its parameters. We adopt the methodology from (Cherkassky and Ma, 2004), where SVR parameter C and ε can be

 $C = max(|\overline{y} + 3\sigma_{y}|, |\overline{y} - 3\sigma_{y}|)$

directly determined from the statistics of training data set;

$$=3\sigma \left[\frac{ln\left(n\right) }{n}\right]$$

ε

as:

, where \overline{y} and σ_y are the mean and standard deviation of the output parameter in the training data set, σ is the input 910 noise level (we set it to 0.001) and n is the number of training samples. The determined values for C and ε are in Table 3. We employ the widely used radial basis function (RBF) kernel function to solve the non-linearity in the SVR model. Compared with other kernel functions, RBF is relatively less complex and more efficient. The RBF kernel is defined

$$K(x_i, x) = exp\left(-\frac{\|x_i - x\|^2}{2p^2}\right)$$

(11)

(9)

(10)

, where p is the kernel width parameter that reflect the influencing area of support vectors. This parameter is

p15 determined by hyper-tuning on the testing data set (Durbha et al., 2007) (explained below).

The RMSE of training process may overestimate the accuracy of an SVR model, because the training and predicting process are based on the same data set. Instead, an independent testing data set is used to represent the accuracy of the SVR model. The difference of model accuracy between training and testing process reflects the generalization performance of the SVR model. An ideal SVR model should output a low level RMSE meanwhile the discrepancy

- 920 between training and testing process is also small. If the RMSE of testing process is much larger than that of training process, then the SVR may suffer from overfitting problems. Fig.7 shows the hyper-parameter tuning process. The first row is the RMSE of training process as a function of C and ε . The second row is the RMSE relative difference between testing process and training process. Column indicate different values of p. The cross marker indicates values of the C and ε determined by the Eq. (9) and (10). It is clear that when $p^2=1.67$, the RMSE of training process is relatively
- 925 small, meanwhile the model accuracy difference between training process and testing process is also small. The final <u>value of C, ε and p</u> that will be applied in case studies are <u>listed in Table 3</u>. The corresponding <u>RMSE of AAOD</u> predicted by the training process and testing process are at level of 0.01 (Fig.8a).

3.5 Error analysis

The error sources of SSA retrieval using SVR model depends on the model accuracy as well as the quality of input 930 data. The model accuracy can be represented by the RMSE of the testing process (0.01). As shown in Fig.8a, the SVR model has difficult predicting AAOD larger than 0.05, where most significant biases appear at this range. The uncertainty in AAOD is passed to the SSA by Eq. (2). Fig.8b shows the retrieved SSA in training and testing process. It is noted that the predicted SSA is overall positively biased, particularly in relatively stronger absorption cases (SSA <0.90). The bias is possibly due to that in the feature domain, the UVAI is relatively strongly correlated to others (i.e.

Deleted: To summarize this section, 3 SVR models are applied in this paper: first, an SVR is used to predict ALH in an intermediate step to improve the quality of the original training data set. This is a temporary solution only applied in this paper due to the lack of observational information concerning aerosol vertical distribution. The remaining 2 SVR models are used to predict AAOD with the original OMI ALH and the adjusted ALH, respectively. Fig.5 shows the procedure of the SVRs for ALH (indicated in purple) and AAOD (indicated in green) prediction. Table 2 also summarizes the input features, output parameter and the corresponding data sources of the 3 SVP models discussed in the 3 SVR models discussed in this sub-section. All input features are scaled into the range between 0 and 1 (min-max normalization) before training.

3.2.3 Deleted:

Deleted: training process and

Deleted: For the mathematical formulation of SVR algorithm one can refer to Smola and Scholkopf (2004). Briefly, SVR tries to find the coefficient ω and the bias b of a linear model by m

... [9]

$\frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^n L_{\varepsilon}(\boldsymbol{y}_i, \boldsymbol{g}(\boldsymbol{x}_i)^T \boldsymbol{\omega} + \boldsymbol{b})$

Deleted: SVR can be solved by specifying a kernel function $K(x_i, x_j) = g(x_i)^T g(x_j)$ which is positive defined so that the Mercer's theorem is satisfied (Tuia et al., 2011). It is clear that

Deleted: σ that concerns the influencing area of support vectors

Formatted: Font: 10.5 pt

Deleted: (Eq.(B1) and (B2) in the Supplement).

Deleted: employ Deleted: a

Deleted: applied in this pa

Deleted: . The kernel parameter σ is determined by

Formatted: Font: 10.5 pt

AOD and ALH), which may contain redundant information that adversely impact model performance (Weston et al., 2001; Durbha et al., 2007). More sophisticated feature selection scheme is suggested to reduce the redundancy, e.g. Minimum Redundancy Maximum Relevance (mRMR, Peng et al., 2005). Moreover, the RBF kernel function may not capable enough to solve the non-linearity among the training data sets. The accuracy of SSA predict by testing data set is ± 0.02 , where 82% samples falling the uncertainty range (± 0.03) of the true SSA (AERONET) and their accuracy is

even higher (<u>+0.01</u>).

970

<u>The error the retrieved SSA due to the input features</u> may come from the observational or retrieval uncertainties in each parameter. In our case, the typical UVAI bias requirement is at magnitude of 1 (Lambert et al., 2019). It is reported TROPOMI UVAI suffers from the long-term downward wavelength-dependent trend in irradiance

- 975 (Rozemeijer and Kleipool, 2018). The detected degradation in UVAI_{354,388} is around 0.2 since August 2018 (Lambert et al., 2019). The typical accuracy of TROPOMI ALH is 50 hPa, though in some situations the bias may over this value (e.g. low aerosol loading over bright surface) (Sanders et al., 2016). Depending on the retrieval algorithm the uncertainty of MODIS AOD is ±0.05+15%AOD_{AERONET} (Dark Target algorithm) (Levy et al., 2010) or ±0.03+0.2AOD_{MODIS} (Deep Blue algorithm) (Sayer et al., 2013). The SSA sensitivity to input features is presented in
- 980 Fig.9. We use the mean value of each parameter in the OMAERUV-AERONET data set as reference values (Fig.B2, UVAI = 1.59, ALH = 2.96 km, AOD = 0.39), the corresponding SSA value is 0.94. The positive bias of UVAI always leads to underestimation in SSA, unless the aerosol layer is located at a relatively high altitude or aerosol loading is low. Conversely, the insufficient UVAI causes the overestimation in SSA, except for cases where ALH is low or AOD is high. The sensitivity of SSA to UVAI is weaker when the aerosol layer is close to surface or at a very high altitude.
- 985 <u>The sensitivity of SSA to UVAI</u> always increases with AOD.

3.6 Case applications

Once the hyper-parameters are determined (Section 3.4), the trained SVR model is ready to predict aerosol absorption, The first application is the California fire event in 2017 December (Section 3.6.2), the same as that in the first experiment. To demonstrate the generalization capability of the SVR model, we also apply it to other fire events as

90 long as there are collocated TROPOMI and MODIS measurements and AERONET-retrieved SSA to compare with (Section 3.6.2).

For all applications, the input parameters in the SVR model are TROPOMI UVAI (calculated by 354 and 388 nm wavelength pair), TROPOMI ALH and MODIS AOD, respectively. The MODIS AOD at 550 nm is converted to 500 nm using the Ångström exponent (α) provided by the collocated <u>AERONET site</u>. Note that the data includes pixels

with CF larger than 0.1 in order to ensure there is satellite measurements collocated with the AERONET sites (though CF is no larger than 0.3).

3.6.1 California fire event on 12 December 2017

 Fig.10 presents the retrieved AAOD and corresponding SSA. It is noted that in the center of the plume, where UVAI and AOD are higher while ALH is relatively lower (Fig.2). The SSA should be smaller to compensate the low altitude

 000
 of the aerosol layer according to Fig.9. However, the SVR retrieved SSA is even higher than its surroundings. It is because that at this region, the UVAI and AOD are outside of the distribution of corresponding parameters shown in

Fig.B2. The <u>13-year OMAERUV-AERONET</u> joint <u>data</u> cannot cover some extreme <u>situations</u>. The reason could be the size of the joint data is relatively small as a result of data availability and collocation criteria, or the quality of the joint

Deleted:

The values chosen by the above methods are robust in our case (Fig.B1-B3 in the Supplement, part B), i.e. retaining a relatively low error while preventing overfitting. Table 3 summaries the settings of the SVR models determined by tuning procedure and the evaluation of the algorithm performance. All 3 SVR models present good generalization capabilities as the differences in root mean square error (RMSE) between the training data and the testing data are minor. The accuracy of the SVR model for ALH prediction is 0.26 km. Fig. 7c. shows the relationship of UVAI, AOD and the SVR predicted ALH. The structure is more similar to that in Fig. 7a and |p| between UVAI and ALH increases from 0.30 to 0.61, which is sufficient to mitigate the impact of uncertainties of ALH in the OMAERUV product. Note that this value may be overestimated as the SVR for ALH prediction is only trained by a specific case due to the limited availability of TROPOMI ALH, but it is more reasonable compared with the original UVAI and ALH relationship in the OMI-AERONET AOD and AAOD provides a new training data set for AAOD prediction, i.e. the adjusted training data set. The accuracy of SVR models for AAOD prediction trained by the original and the adjusted training data set are 0.01.[¶]

Deleted: 2.4

Deleted: Data for case application
Deleted: After
Deleted: determining
Deleted: the SVR is applied to predict aerosol absorption in the case of the smoke plume generated by the Thomas Fire event
Deleted: are taken from
Deleted: from near-UV band,
Deleted: from visible band

035 data suffers from observational or retrieval uncertainties. As a result, the SVR model fails to handle the input values outside the range of training data set.

<u>The SSA of the all plume pixels is 0.94 ± 0.01 (including the failed-predicted pixels) and that for the AERONET-</u> <u>collocated pixels (pixels within 50 km from UCSB) is 0.97 ± 0.01 (Table 4).</u> These values may be overestimated while the standard deviation may be underestimated because of the SVR prediction failures of some samples. <u>The SSA</u>

040 <u>difference relative to the AERONET retrieval is only 0.01, which is within the uncertainty range of AERONET</u> (±0.03).

3.6.2 Other case applications

To present the generalization performance of SVR, we apply it to other fire events as long as there is collocated information from TROPOMI, MODIS and AERONET. The same pre-processing as the previous case is applied to exclude pixels with UVAI smaller than 1, AOD smaller than 0.5 or CF larger than 0.3.

Fig.11-13 present California fire events during 9-11 November 2018. The plumes were over ocean but partly contaminated by the underlying clouds (Fig.A2-A4 present the Aqua MODIS RGB images). Fig.14 shows the Canada fire event on 29 May 2019. The case was over land (Fig.A5 present the Aqua MODIS RGB image), which means the brighter surface may cause higher bias in the input AOD and ALH than cases over dark surfaces (Remer 2005; Sanders

- 050 and de Haan, 2016).
 - The retrieved SSA for above events is listed in Table 4. Similar to the California 2017-12-12 case, The SVR fails to retrieve reasonable SSA for pixels if input features outside their corresponding histogram in the OMAERUV-AERONET data (Fig.2B), which may cause overestimations in plume mean SSA. The plume SSA of two California fire events are similar, with values around 0.94-0.95. The retrieved SSA for the Canada fire is relatively higher (0.97)
- 055 We further plot the SSA retrieved by SVR against collocated <u>AERONET records</u> (black crosses in Fig.15)_Including the first case (California fire on 2017-12-12), there are 9 collocated records obtained. The difference between SVR-retrieved SSA and AERONET are alomost within difference of ±0.05, among which over half (5 out of 9) fall in <u>AERONET SSA uncertainty range (±0.03)</u>. We also provide SSA from OMAERUV for these cases (Table 4 and blue circles in Fig.15). Compared with OMAERUV, the SSA retrieved by SVR shows a better consistency with
- 060 AERONET, though one should keep in mind that the accuracy of SVR-retrieved SSA is ±0.02 and the model tends to overestimate the SSA for relatively absorbing cases.

3.6.3 Spatial variability of retrieved SSA

<u>Compared with Fig.</u>4b, the spatial variability of SSA retrieved by SVR is less strong (Fig.10-14), whose difference between maximum and minimum SSA falls in range from 0.09 to 0.10 (Table 4). In the first experiment, SSA is

065 determined by UVAI for each pixel individually. In the SVR model, the spatial variability of the intermediate output AAOD depends on the three input features. Furthermore, SVR predicts SSA for each pixel based on the common relationship between UVAI, AOD and ALH in the training data set.

Heterogeneity in aerosol properties is expected for plume of this size, but to what extend needs further investigations. Here we assess the SSA spatial variability of by an independent data set. We employ the SSA calculated by AOD and scattering AOD from MERRA-2 aerosol reanalysis hourly single-level product

(https://disc.gsfc.nasa.gov/datacollection/M2T1NXAER_5.12.4.htm last access: 16 July 2019). The AOD and aerosol properties of MERRA-2 are proved to be in good agreement with independent measurements (Buchard et al., 2017;

Deleted: The predicted AAOD is further used to derive the SSA according to Eq.(2). The retrieved SSA within 50 km from AERONET site (UCSB) is validated. Information about each data product can be found in the corresponding sub-sections in Section 2.

Formatted: (Asian) Chinese (China)

Deleted: The predicted AAOD is further used to derive the SSA according to Eq.(2). The retrieved SSA within 50 km from AERONET site (UCSB) is validated. Information about each data product can be found in the corresponding sub-sections in Section 2...

Formatted: (Asian) Chinese (China)

Randles et al., 2017). The MERRA-2 AOD and SSA for these cases are shown in Appendix C. The plume can be detected by the high AOD against its surrounding. Although the plume presented by the satellite observations significantly differs from that of model simulations, the SSA spatial difference within the plume is approximately at magnitude of 0.1. From this aspect, the spatial variability of SSA retrieved by the SVR model is in better agreement

085

4 Conclusions

with MERRA-2.

- The long-term record of global UVAI data is a treasure to derive aerosol optical properties such as SSA, which is
 important for aerosol radiative forcing assessments. To quantify aerosol absorption from UVAI, the information of
 AOD and ALH is necessary. There are various AOD products while ALH products are much less accessible. Recently,
 the TROPOMI oxygen A-band ALH product has been run operationally, using which we demonstrate the role of ALH
 in quantifying SSA from satellite retrieved UVAI for biomass burning aerosols.
- In the first experiment, we derive the SSA by forward radiative transfer simulation of UVAI for a fire event in
 OP5 California on 2017-12-12. With the TROPOMI ALH, we are able to quantify the influence of assumed spectral dependence of near-UV aerosol absorption (represented by the relative difference between K₃₅₄ and K₃₈₈) on the retrieved SSA. A significant gap in plume mean SSA (0.25) between 'gray' and strong spectral dependent aerosols (Δκ=0% and 40%, respectively) implies that inappropriate assumptions on spectral dependence may significantly bias the retrieved aerosol absorption. The SSA difference between AERONET and collocated pixels becomes smaller than

100 the uncertainty of AERONET (± 0.03) when $\Delta \kappa = 25\%$. The corresponding plume SSA of 0.90 ± 0.05 and the AERONET-collocated pixels <u>SSA</u> of <u>0.96\pm0.02</u>.

In the second part of this paper, we propose a statistical method based on the long-term records of UVAI, AOD, ALH and AAOD using an SVR algorithm, in order to avoid making assumptions on aerosol absorption spectral dependence over near-UV band, The SVR model is trained by 5679 collocated global observations from OMAERUV and

- 105 AERONET during the period from 2005 to 2017 The SVR-retrieved SSA for the California fire event on 2017-12-12 is 0.97±0.01, which is 0.01 lower than that of AERONET. The SVR algorithm is also applied to other cases. Consider all case applications, the results are encouraging: the SSA discrepancy between retrieval and AERONET for almost all collocated samples are within ±0.05 difference and over half of them fall in the AERONET uncertainty range (±0.03). One should keep in mind the SVR model tends to overestimate the SSA for relatively absorbing cases (e.g.
- 110 SSA<0.90), and sometimes fails to predict reasonable SSA when the input values fall outside the range of the corresponding parameters in the training data set.</p>

In terms of spatial variability, the SSA retrieved by radiative transfer simulations significantly differs from that retrieved by SVR. Spatial heterogeneity in SSA is expected, but to what extent needs further investigations. We employ the SSA provided by MERRA-2 aerosol reanalysis as a reference, whose spatial difference within smoke

115 plume is approximately at magnitude of 0.1. The spatial pattern of SSA retrieved by SVR agrees better with this finding.

In this study, we present the potential to retrieve SSA based on the long-term data records of UVAI, ALH, AOD and AAOD using a statistical method. The motivation is to avoid a priori assumptions on aerosol micro-physics as we made in radiative transfer simulations. At the current phase, the algorithm we choose is SVR as the size of the training

120 data set is relatively small. The input features are selected by the <u>Spearman's rank</u> correlation coefficients and a priori knowledge on relationship between UVAI and only aerosol related features. The model hyper-parameters are analytically determined. The accuracy of SVR-predicted SSA is <u>+0.02</u>, with higher tendency to overestimate the SSA

Deleted: Results and discussions
Deleted: 4.1 Evaluating the retrieved SSA with AERONET
measurements We first analyze the results of the RTM-based method. Fig 8a
displays the mean SSA of all plume pixels retrieved by the RTM-[10]
Formatted: Font: 8 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 8 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Formatted: Font: 10.5 pt
Deleted: used the
Formatted: Font: 10.5 pt
Deleted: 4
Deleted: 'colored'
Formatted: Font: 10.5 pt
Deleted: demonstrates
Deleted:
Deleted: s
Deleted: smoke
Deleted:
Deleted: The TROPOMI ALH product also provides the [11]
Deleted: This is
Deleted: correlation
Deleted: between
Deleted: and
Deleted: , and
Deleted: The new method was realized by SVR, a representative
Deleted: 4003
Deleted: co-located
Deleted: An adjustment on OMI ALH was implemented to [13]
Deleted: . The results of the SVR-based retrieval are more [14]
Formatted: Font: 10.5 pt
Deleted: Pearson

for relatively absorbing cases. The OMAERUV-AERONET data set cannot cover some extreme situations, as a result, the SVR fails to predict reasonable SSA when the input values fall outside the range of the corresponding parameters in the training data set. In the future, more sophisticated feature selection techniques and kernel functions should be

- 255 considered to improve the accuracy the algorithm. Other non-aerosol features affecting UVAI should be also taken into consideration. Moreover, the high-resolution TROPOMI level 2 UVAI and ALH products are expected to significantly increase the size of training data set and improve the quality of the training data set, which will reduce the computational failures of the SVR model and even guide use to more powerful algorithms (e.g. ANN) to retrieve SSA.
- 260 Acknowledgements. This work was performed in the framework of the KNMI Multi-Annual Strategic Research (MSO). The authors thank NASA's GES-DISC and LAADS DAAC for free online access of OMI and MODIS data. The authors thank NASA Goddard Space Flight Center AERONET Project for providing the data from the AERONET. The authors thank Swadhin Nanda for providing TROPOMI aerosol layer height product.

References

265 Ahn, C., Torres, O., and Jethva, H. (2014). Assessment of OMI near-UV aerosol optical depth over land. Journal of Geophysical Research: Atmospheres, 119(5), 2457-2473.

Alex J.Smola and Scholkopf, B.: A tutorial on support vector regression, Stat. Comput., 199–222, doi:10.1023/B:STCO.0000035301.49549.88, 2004.

Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., Loyola, D. and Wang, P.: Level 2 Product User Manual KNMI level 2 support 270 products, 2017.

Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D. and Yu, H.: The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies, J. Clim., 30(17), 6851– 6872, doi:10.1175/JCLI-D-16-0613.1, 2017.

Cherkassky, V. and Ma, Y.: Practical selection of SVM parameters and noise estimation for SVM regression, Neural Networks, 17(1), 113–126, doi:10.1016/S0893-6080(03)00169-2, 2004.

Chimot, J., Pepijn Veefkind, J., Vlemmix, T., De Haan, J. F., Amiridis, V., Proestakis, E., Marinou, E. and Levelt, P. F.: An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nmO2O2spectral band using a neural network approach, Atmos. Meas. Tech., 10(3), 783–809, doi:10.5194/amt-10-783-2017, 2017.

Colarco, P. R., Toon, O. B., Torres, O. and Rasch, P. J.: Determining the UV imaginary index of refraction of Saharan dust particles from
 Total Ozone Mapping Spectrometer data using a three-dimensional model of dust transport, J. Geophys. Res., 107(July 1997), 1–19, 2002.
 Cortes, C. and Vapnik, V.: Support-Vector Networks, IEEE Expert, 7(5), 63–72, doi:10.1109/64.163674, 1995.
 D.C. Stein Zweers: TROPOMI ATBD of the UV aerosol index., 2016.

De Graaf, M. and Stammes, P.: SCIAMACHY Absorbing Aerosol Index – calibration issues and global results from 2002–2004, Atmos. Chem. Phys, 5, 2385–2394, doi:10.5194/acpd-5-3367-2005, 2005.

285 de Haan, J. F.: DISAMAR Determining Instrument Specifications and Analyzing Methods for Atmospheric Retrieval., 2011. Drucker, H., J.C.Burges, C., Linda, K., Alex.Smola and Vladimir, V.: Support vector regression Machines, Adv. Neural Inf. Process. Syst., 155–161, doi:10.1145/2768566.2768568, 1997.

Durbha, S. S., King, R. L. and Younan, N. H.: Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer, Remote Sens. Environ., doi:10.1016/j.rse.2006.09.031, 2007.

290 Guyon, I. and Elisseeff, A.: An Introduction to Variable and Feature Selection, 3(03), 1157–1182, doi:10.1016/j.aca.2011.07.027, 2003. Haywood, J. M. and Shine, K. P.: The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget, Geophys. Res. Lett., 22(5), 603–606, doi:10.1029/95GL00075, 1995.

Herman, J. R. and Celarier, E. a.: Earth surface reflectivity climatology at 340–380 nm from TOMS data, J. Geophys. Res., 102(D23), 28003, doi:10.1029/97JD02074, 1997.

- Herman, J. R., Bhartia, P. K., Torres, O., Hsu, C., Seftor, C. and Celarier, E.: Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data, J. Geophys. Res., 102(D14), 16911, doi:10.1029/96JD03680, 1997.
 Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I. and Smirnov, A.: AERONET A federated instrument network and data archive for aerosol characterization, Remote
- Holben, B. N., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., Dubovik, O. and Lille, U. S. T. De: AERONET's Version 2.0 quality assurance criteria, Int. Soc. Opt. Photonics, 6408(64080Q), 2006.
 Hsu, N. C. and Herman, @bullet J R: Comparisons of the TOMS aerosol index with Sun-photometer aerosol optical thickness: Results and applications, J. Geophys. Res., 104(27), 6269–6279, doi:10.1029/1998JD200086, 1999.
 Hu, R. M., Martin, R. V. and Fairlie, T. D.: Global retrieval of columnar aerosol single scattering albedo from space-based observations, J.

Sens. Environ., 66(1), 1-16, doi:10.1016/S0034-4257(98)00031-5, 1998.

- 305 Geophys. Res. Atmos., doi:10.1029/2005JD006832, 2007. IPCC: Climate Change 2014 Synthesis Report, Geneva. [online] Available from: http://www.ipcc.ch/pdf/assessmentreport/ar5/syr/SYR_AR5_FINAL_full.pdf, 2014. Jeong, M. J. and Hsu, N. C.: Retrievals of aerosol single-scattering albedo and effective aerosol layer height for biomass-burning smoke: Synergy derived from "A-Train" sensors, Geophys. Res. Lett., 35(24), 1–6, doi:10.1029/2008GL036279, 2008.
- 310 Jethva, H. and Torres, O.: Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from Ozone Monitoring Instrument, 10541–10551, doi:10.5194/acp-11-10541-2011, 2011. Jethva, H., Torres, O. and Ahn, C.: Global assessment of OMI aerosol single-scattering albedo using ground-based AERONET inversion, J. Geophys. Res., 119(14), 9020–9040, doi:10.1002/2014JD021672, 2014.
- 315 Kleipool, Q.: OMI Surface Reflectance Climatology README, 2010. Kleipool, Q. L., Dobber, M. R., de Haan, J. F. and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res. Atmos., 113(18), 1–22, doi:10.1029/2008JD010290, 2008. Lary, D. J., Alavi, A. H., Gandomi, A. H. and Walker, A. L.: Machine learning in geosciences and remote sensing, Geosci. Front., (7), 3– 10, 2015.
- 320 Levelt, P. F. and Noordhoek, R.: OMI algorithm theoretical basis document, J (August), 1–50, 2002. Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F. and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6(11), 2989–3034, doi:10.5194/amt-6-2989-2013, 2013. Mountrakis, G., Im, J. and Ogole, C.: Support vector machines in remote sensing: A review, ISPRS J. Photogramm. Remote Sens., 66(3), 247–259, doi:10.1016/j.isprsjprs.2010.11.001, 2011.
- 325 Noia, A. Di and Hasekamp, O. P.: Neural Networks and Support Vector Machines and Their Application to Aerosol and Cloud Remote Sensing : A Review., 2018.

Peng, H., Long, F. and Ding, C.: Feature Selection Based on Mutual Information (mRMR), Ieee Trans. Pattern Anal. Mach. Intell., 27(8), 1226–1238, doi:10.1007/978-3-319-03200-9_4, 2005.

Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products - #03: July 2018 - May 2019. Lambert, J.-

- 330 C., A. Keppens, D. Hubert, B. Langerock, K.-U. Eichmann, Q. Kleipool, M. Sneep, T. Verhoelst, T. Wagner, M. Weber, C. Ahn, A. Argyrouli, D. Balis, K.L. Chan, S. Compernolle, I. De Smedt, H. Eskes, A.M. Fjæraa, K. Garane, J.F. Gleason, F. Goutail, J. Granville, P. Hedelt, K.-P. Heue, G. Jaross, ML. Koukouli, J. Landgraf, R. Lutz, S. Niemejer, A. Pazmiño, G. Pinardi, J.-P. Pommereau, A. Richter, N. Rozemeijer, M.K. Sha, D. Stein Zweers, N. Theys, G. Tilstra, O. Torres, P. Valks, C. Vigouroux, and P. Wang. S5P MPC Routine Operations Consolidated Validation Report series, Issue #03, Version 03.0.1, 125 pp., June 2019.
- 335 Randles, C. A., da Silva, A. M., Buchard, V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y. and Flynn, C. J.: The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation, J. Clim., 30(17), 6823–6850, doi:10.1175/JCLI-D-16-0609.1, 2017. Remer, L. a.: The MODIS Aerosol Algorithm, Products, and Validation, J. Atmos. Sci., 62, 947–973, 2005. Rozemeijer, N. C. and Kleipool. Q.: SSP Mission Performance Centre Level 1b Readme., 1–16, 2018.
- 340 Sanders, A. F. J., Haan, J. F. De, Sneep, M., Apituley, A., Stammes, P., Vieitez, M. O. and Tilstra, L. G.: Evaluation of the operational Aerosol Layer Height retrieval algorithm for Sentinel-5 Precursor : application to O 2 A band observations from GOME-2A, , 4947–4977, doi:10.5194/amt-8-4947-2015, 2015.

Sanders A.F.J. and de Haan J.F.: TROPOMI ATBD of the Aerosol Layer Height product., 2016. Shin, K. S., Lee, T. S. and Kim, H. J.: An application of support vector machines in bankruptcy prediction model, Expert Syst. Appl., 345 28(1), 127-135, doi:10.1016/j.eswa.2004.08.009, 2005. Smola, A. J. and Scholkopf, B.: A tutorial on support vector regression, Stat. Comput., 199-222, doi:10.1023/B:STCO.0000035301.49549.88, 2004. Sun, J., Veefkind, J. P., Velthoven, P. Van and Levelt, P. F.: Quantifying the single scattering albedo for the January 2017 Chile wildfires from simulations of the OMI absorbing aerosol index, Atmos. Meas. Tech. Discuss., 2(February), 1-24, 2018. 350 T. Holzer-Popp, Leeuw, G. de, Griesfeller, J. and Martynenko, D.: Aerosol retrieval experiments in the ESA Aerosol cci project, Atmos. Meas. Tech., 6(8), 1919-1957, doi:10.5194/amt-6-1919-2013, 2013. Tilstra, L. G., Wang, P. and Stammes, P.: ALGORITHM THEORETICAL GOME-2 Absorbing Aerosol Height, 2018. Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z. and Gleason, J.: Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis, J. Geophys. Res. Atmos., 103(D14), 17099-17110, doi:10.1029/98JD00900, 1998. 355 Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P. and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations : An overview, J. Geophys. Res., 112(D24S47), 1-14, doi:10.1029/2007JD008809.2007 Torres, O., Ahn, C., Chen, Z. and Space, G.: Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations. . 3257-3270. doi:10.5194/amt-6-3257-2013. 2013. 360 Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P. and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations : An overview, J. Geophys. Res., 112(D24S47), 1-14, doi:10.1029/2007JD008809, 2007 Torres, O., Ahn, C., Chen, Z. and Space, G.: Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations, , 3257-3270, doi:10.5194/amt-6-3257-2013, 2013. 365 Torres, O., Bhartia, P. K., Jethva, H. and Ahn, C.: Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products. 2(2), 2701-2715, 2018. Tuia, D., Verrelst, J., Alonso, L., Perez-Cruz, F. and Camps-Valls, G.: Multioutput support vector regression for remote sensing biophysical parameter estimation, IEEE Geosci. Remote Sens. Lett., 8(4), 804-808, doi:10.1109/LGRS.2011.2109934, 2011. Veefkind, J. P., Aben, I., Mcmullan, K., Förster, H., Vries, J. De, Otter, G., Claas, J., Eskes, H. J., Haan, J. F. De, Kleipool, Q., Weele, M. 370 Van, Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H. and Levelt, P. F.: Remote Sensing of Environment TROPOMI on the ESA Sentinel-5 Precursor : A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120(2012), 70-83, doi:10.1016/i.rse.2011.09.027.2015. Wang, P., Stammes, P., Van Der A, R., Pinardi, G, and Van Roozendael, M.: FRESCO+: An improved O2A-band cloud retrieval 375 algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8(21), 6565–6576, doi:10.5194/acp-8-6565-2008, 2008. Wang, P., Tuinder, O. N. E., Tilstra, L. G., De Graaf, M. and Stammes, P.: Interpretation of FRESCO cloud retrievals in case of absorbing aerosol events, Atmos. Chem. Phys., 12(19), 9057-9077, doi:10.5194/acp-12-9057-2012, 2012. Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T. and Vapnik, V.: Feature selection for SVMs, Adv. Neural Inf. Process. Syst., 668-674 [online] Available from: http://papers.nips.cc/paper/1850-feature-selection-for-svms.pdf, 2001. 380 Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H. and Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, J. Atmos. Ocean. Technol., 26(11), 2310-2323, doi:10.1175/2009JTECHA1281.1, 2009. Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y., Marshak, A., Reid, J. and Miller, S.: Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: First results from EPIC/DSCOVR at Lagrange-1 point, Geophys. Res. Lett., 44(14), 7544-7554, doi:10.1002/2017GL073939, 2017. 385 Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Reid, J. S., Miller, S. D., Martins, J. V. and Remer, L. A.: Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: Hourly results from EPIC/DSCOVR satellite in deep space, Atmos, Meas, Tech, Discuss., 1-31, doi:10.5194/amt-2018-414, 2019. Yao, X., Tham, L. G. and Dai, F. C.: Landslide susceptibility mapping based on Support Vector Machine: A case study on natural slopes of Hong Kong, China, Geomorphology, doi:10.1016/j.geomorph.2008.02.011, 2008. 390





Figure 2: Satellite data from California fire event on 2017-12-12: (a) TROPOMI UVAI calculated by reflectance at 354 and 388 nm; (b) TROPOMI ALH (unit: km); (c) MODIS AOD at 550 nm.



<u>Figure 3: SSA retrieved by</u> radiative transfer simulations as a function of $\Delta \kappa (\Delta \kappa = (\kappa_{354} - \kappa_{388})/\kappa_{368})$; (a) SSA mean and standard deviation (filled region) of plume pixels; (b) SSA mean and standard deviation (filled region) of the 15_AERONET-collocated pixels; (c) absolute difference between the mean SSA of the 15_collocated pixels and the AERONET retrieval.



Deleted:

Moved (insertion) [3]

[17]







Figure 7: The performance of SVR model as a function to hyper-parameters (C, ε and p). The cross marker represents the values of C and ε according to Cherkassky and Ma (2004). p^2 equaling 1.67 is sufficient to obtain a relatively high accuracy, meanwhile prevents overfitting on the training data set.

505



525 Figure 8: The accuracy of the trained SVR model: (a) the predicted AAOD at 500 nm against true AAOD at 500 nm. The dashed line is the 1:1 line and the solid line is the linear fitting for the testing data set; (b) the predicted SSA at 500 nm against true SSA at 500 nm. The grey and red color indicates samples in training and testing data set, respectively. The values inside parenthesis is the statistics for samples fall in AERONET uncertainty of 0.03.



530 Figure 9: The sensitivity of the SVR-retrieved SSA: (a) the response of predicted SSA at 500 nm as a function of changes in UVAI and ALH; (b) the response of predicted SSA at 500 nm as a function of changes in UVAI and AOD.



Figure 10: SVR retrievals for California fire event on 2017-12-12: (a) retrieved AAOD at 500 nm; (b) retrieved SSA at 500 nm.



535

Figure11: SVR retrievals for California fire event on 2018-11-09: (a) TROPOMI UVAI calculated by reflectance at 354 and 388 nm; (b) TROPOMI ALH; (c) MODIS AOD at 550 nm; (d) retrieved AAOD at 500 nm; (e) retrieved SSA at 500 nm.

Deleted:



540 Figure 12: SVR retrievals for California fire event on 2018-11-10: (a) TROPOMI UVAI calculated by reflectance at 354 and 388 nm; (b) TROPOMI ALH; (c) MODIS AOD at 550 nm; (d) retrieved AAOD at 500 nm; (e) retrieved SSA at 500 nm.



Figure 13: SVR retrievals for California fire event on 2018-11-11: (a) TROPOMI UVAI calculated by reflectance at 354 and 388 545 nm; (b) TROPOMI ALH; (c) MODIS AOD at 550 nm; (d) retrieved AAOD at 500 nm; (e) retrieved SSA at 500 nm.



Figure14: SVR retrievals for Canada fire event on 2019-05-29: (a) TROPOMI UVAI calculated by reflectance at 354 and 388 nm; (b) TROPOMI ALH; (c) MODIS AOD at 550 nm; (d) retrieved AAOD at 500 nm; (e) retrieved SSA at 500 nm.



Figure15: SVR-retrieved SSA (black cross) and OMAERUV-retrieved SSA (blue circle) against AERONET SSA at 500 nm for all 5 cases in this study.

555







Moved up [3]: Figure 9: SSA retrieved by the: (a) RTM-based method; (b) SVRbased method with OMAERUV ALH; (c) SVR-based method with adjusted ALH.¶

Geometri	Effective	Geometry	Variance	Refractive	Spectral	Refractive index	Refractive index
(r_a)	(r_{eff})	deviation	(inte _{eff})	part (n)	dependence (Δk)	354 nm	other wavelengths
(y)		(σ_g)		1 ()		(K354)	(≥388 nm)
0.07 µm	0.14 µm	1.7	0.53	1.5	0%, 5%, 10%,	$(1 + \Delta \kappa) \times \kappa_{388}$	0.005
					15%, 20%, 25%,		0.010
					30%, 35% and		0.020
					40%		0.030
							0.040
							0.048
							0.060

Table 2 Retrieved SSA by the radiative transfer simulations for the California fire on 2017-12-12.

Retrieval methods	Number of	Retrieved SSA	$SSA_{max}-SSA_{min} \\$	Retrieved SSA	AERONET	OMAERU
	plume	(plume pixels)		(collocated-pixels)	SSA	SSA
	pixels					
RTM with	5217	0. <u>90</u> ±0.05	0.38	0.95 ± 0.02	0.98	0.92 <u>+</u> 0.01

Table 3 Values for hyper-tuning decided regularization constant C, the width of the insensitive zone ε and the BRF kernel parameter p^2 .

	SVR hyp	er-parameters		-
Parameters	С	ε	p^2	-
Values	0.09	0.0001	1.67	

585

Table 4 SVR-retrieved SSA. If there is no standard deviation followed, then it indicates there is only one record.

Case	Num. of	Retrieved	SSA _{max} -	Collocated	<u>SSA</u>	AERONET	OMAERUV
	Plume	SSA (plume	$\mathrm{SSA}_{\mathrm{min}}$	AERONET	(collocated-	SSA	SSA
	pixels	pixels)			pixels)		
California 2017-12-12	5217	0.94 <u>+</u> 0.01	0.09	UCSB	0.97 <u>+</u> 0.01	0.98	0.92 <u>+</u> 0.01
California 2018-11-09	1944	<u>0.9</u> 4 <u>±0.01</u>	0.10	Santa_Monica_Colg	<u>0.93±0.0</u> 1	<u>0.</u> 89 <u>±0.0</u> 6	<u>0.</u> 89 <u>±0.0</u> 6
California 2018-11-10	2184	<u>0.9</u> 4 <u>±0.0</u> 2	0.10	CalTech	<u>0.9</u> 6 <u>±0.01</u>	<u>0.</u> 89 <u>±0.0</u> 7	-
				Fresno_2	<u>0.9</u> 3 <u>±0.0</u> 2	<u>0.</u> 91 <u>±0.0</u> 1	-
				Modesto	<u>0.9</u> 4 <u>±0.01</u>	<u>0.</u> 92 <u>±0.0</u> 1	<u>0.</u> 96 <u>±0.0</u> 1
				USC_SEAPRISM_2	<u>0.9</u> 3 <u>±0.00</u>	<u>0.</u> 90	-
California 2018-11-11	2815	<u>0.9</u> 5 <u>±0.0</u> 2	0.09	Modesto	<u>0.9</u> 8 <u>±0.00</u>	<u>0.</u> 96 <u>+0.0</u> 1	<u>0.95±0.00</u>
Canada 2019-05-29	8013	<u>0.9</u> 7 <u>±0.0</u> 2	0.10	Fort_McKay	<u>0.9</u> 7 <u>±0.02</u>	<u>0.</u> 95 <u>+0.0</u> 0	0.93
				Fort_McMurray	<u>0.9</u> 8±0.01	0.93	1.00

Formatted: Font: (Default) Times New Roman Formatted: Left, Line spacing: 1.5 lines Deleted: validation

Deleted: ¶

199

-1	Deleted: 896	
4	Deleted: 45	
۲	Deleted: 57	
Ì	Deleted: 17	
()	Deleted: 02	
Ù	Deleted: 57	
Ì	Deleted: 17	
Ì	Deleted: SVR with adjusted ALH	[19]
Y	Deleted: SVR model parameters	
Y	Deleted: SVR model	
Ì	Deleted: ALH prediction	[20]
Ì,	Deleted: (original training data set)	
()	Deleted: 010	
Q	Deleted: 2	
Ņ	Deleted: 1	
Ì	Deleted: AAOD prediction (adjusted training data set)	[21]
4	Deleted: 56	
Y	Deleted: 49	
Y	Deleted: 57	

1

Appendix

Part A: Case information



Figure A1: Smoke plume captured by Aqua MODIS for California fire event on 2017-12-12 (source:https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.



Figure A2: Smoke plume captured by Aqua MODIS for California fire event on 2018-11-09 (source: https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.



Figure A3: Smoke plume captured by Aqua MODIS for California fire event on 2018-11-10 (source: https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.



Figure A4: Smoke plume captured by Aqua MODIS for California fire event on 2018-11-11 (source: https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.



Figure A5: Smoke plume captured by Aqua MODIS in for Canada fire event on 2019-05-29 (source: https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.



Figure B2: Statistics of the OMAERUV-AERONET joint data set: (a) OMAERUV UVAI calculated from reflectance at 354 and 388 nm; (b) OMAERUV ALH; (c) AERONET AOD at 500 nm; (d) AERONET AAOD at 500 nm; (e) AERONET SSA at 500 nm.

Part C: MERRA-2 aerosol reanalysis.



Figure C1: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for California fire event on 2017-12-12: (a) AOD at 500 nm; (b) SSA at 500 nm.



675

Figure C2: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for California fire event on 2018-11-09: (a) AOD at 500 nm; (b) SSA at 500 nm.



2.0

Figure C3: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for California fire event on 2018-11-10:

1.5

0.80

0.85

135<u>°W 130°W 125°W 120°W</u>

(b) SSA₅₀₀

0.90

0.95

1.00

135°W 130°W 125°W 120°W

(a) AOD₅₀₀

1.0

40°N

35°N

30°N

25°N

680 (a) AOD at 500 nm; (b) SSA at 500 nm.

0.0

0.5

Figure C4: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for California fire event on 2018-11-11: (a) AOD at 500 nm; (b) SSA at 500 nm.



Figure C5: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for Canada fire event on 2019-05-29: (a) AOD at 500 nm; (b) SSA at 500 nm.

Page 6: [1] Deleted	Sun ji	7/15/19 4:15:00 PM	
▼			
Page 6: [2] Deleted	Sun ji	7/15/19 6:48:00 PM	
Page 9: [3] Deleted	Sun ji	7/18/19 3:48:00 PM	
Page 9: [4] Deleted	Sun ji	7/18/19 4:16:00 PM	
Page 9: [5] Deleted	Sun ji	7/18/19 4:16:00 PM	
▼			
Page 11: [6] Deleted	Sun ji	7/20/19 12:11:00 PM	
Page 11: [7] Deleted	Sun ji	7/20/19 12:20:00 PM	
x			
Page 11: [8] Deleted	Sun ji	7/20/19 12:32:00 PM	
▼			
Page 12: [9] Deleted	Sun ji	7/18/19 4:17:00 PM	
Page 15: [10] Deleted	Sun ji	7/20/19 2:04:00 PM	
<u>×</u>			
Page 15: [11] Deleted	Sun ji	7/21/19 10:39:00 PM	
▼		_	
Page 15: [12] Deleted	Sun ji	7/21/19 10:41:00 PM	
▼			
Page 15: [13] Deleted	Sun ji	7/21/19 10:41:00 PM	
v			
Page 15: [14] Deleted	Sun ji	7/21/19 10:48:00 PM	
X			
Page 21: [15] Deleted	Sun ji	7/20/19 11:51:00 AM	
Page 21: [16] Deleted	Sun ji	7/18/19 5:36:00 PM	
Page 21: [17] Deleted	Sun ji	7/18/19 7:02:00 PM	
Page 21: [18] Deleted	Sun ji	7/20/19 11:52:00 AM	
▼	-		
Page 27: [19] Deleted	Sun ji	7/22/19 8:23:00 AM	
Page 27: [20] Deleted	Sun ji	7/20/19 1:18:00 PM	
Values	0.09 0.0001 1.67		
Page 27: [21] Deleted	Sun ji	7/20/19 1:18:00 PM	