Responses to Reviewer 1

This paper shows the dust and non-dust aerosol components over the Middle East that are available from two reanalysis products, MERRA-2 and CAMS, and WRF-Chem model simulations. It first compares the 10-m wind speed among the products, then compares AOD and size distributions with remove-sensing data (such as MODIS MAIAC and AERONET) and PM2.5/PM10 concentrations with ground-based measurements. With the results from WRF-Chem simulations and reanalysis products, the aerosol composition of PM2.5 and PM10 is presented and days of PM above the reulatory standard are estimated.

I found this paper is interesting in the sense that an evaluation of model and reanalysis products is specifically performed for the Middle East region with, albeit limited, remote sensing and ground-based measurements, and the seasonal and annual levels of PM are presented. I however have quite some comments regarding the presentation and understanding of the products used in the paper, and recommend substantial revision before accepting for publication on ACP. I believe that the revision is not difficult to deal with although it could be extensive.

We thank the reviewer for the valuable comments. Despite poor air quality, the Middle East has very sparse air quality observations. So it is essential to thoroughly test the modeling tools. This is the first attempt to reconcile observation, models, and reanalysis products in this region.

All references supporting our response are placed at the end of the text. The reviewer's questions are in black. Our answers are in blue.

For the presentation style in general:
- Abstract should contain only one paragraph with acronyms spelled out (e.g., ME).
- Introduction section is too long
- should be more concise and more relevant to the point of the study. It is not a literature review.
- Conclusion section is also too long and unfocused. It has 14 paragraphs! It should be consolidated with key points summarized and highlighted, not list everything you have done.

We agree and have revised the abstract, introduction, and conclusions sections accordingly.

Aerosol composition: there are no data to evaluate the models. The surface measurement data are for PM, not chemical species. Besides, the models do not include nitrate and ammonium, and it seems they don't have the chemical mechanisms for producing secondary organic aerosols. Therefore, the chemical composition from the model omits some important components. The problem should be acknowledged at least. Is there any reference for the aerosol composition in the region?

In-situ air quality observations in the Middle east are scarce. It is one of the known problems for air quality research in this area. The things are simplified a bit by the fact that in the ME dust dominates aerosol pollution. E.g., Calipso records dust in 95% of profiles (Osipov et al., 2015). The effect of nitrates, ammonia, and organics on AOD and PMs is insignificant in comparison with dust therefore the employed chemical scheme **(***GOCART-RACM***)** is adequate.

To support this conclusion, we have conducted a laboratory analysis of the chemical composition of soil and dust deposition samples that show a little presence of organics and ammonium (Prakash et al. 2016; Engelbrecht et al., 2017). According to (Engelbrecht et al.,

2017) in 2015 the annual average weight percentages of soluble ions of ammonium (NH4) and sulfate (SO4) in deposition samples taken at four sites at the KAUST campus are 0.05% and 2.513%, respectively. It means that available ammonium may neutralize at maximum 5% of sulfate mass. The actual contribution of ammonium sulfate should be lower, as some ammonium may also be bound as ammonium nitrate, ammonium phosphate, or ammonium chloride. We have added this explanation to the revised text (see Sec 4.1, last paragraph).

Reanalysis products: It should be pointed out that the reanalysis products from MERRA-2 and CAMS-OA are the reanalysis of AOD, not the mass concentrations of individual aerosol species. The mass of individual aerosol specie is adjusted mostly proportionally according to the differences between the AOD from native model simulation and after the assimilation of satellite data. Also, in general, a better understanding of the reanalysis products and other products is needed.

We are aware of the reanalysis machinery and mentioned in the original text on p. 4, line 88: "*They improve the aerosol total column loadings through the assimilation of observed AOD but are not capable of assimilating the aerosol vertical structure and chemical composition.*" The representatives of both MERRA-2 and CAMS development teams are co-authors on this paper.

Comparisons with data: The comparison with AERONET AOD is not an independent evaluation of WRF-Chem and MERRA-2, because the WRF-Chem is "tuned" to match AERONET AOD and MERRA-2 assimilates AERONET AOD. This evaluation should be properly addressed.

This is partially correct. All satellite retrievals use AERONET observations for calibration. MERRA-2 assimilates AERONET AOD, but CAMS-OA does not. In WRF-Chem, we tuned to the annual average AOD to fit AERONET observations. We did not tune the temporal correlations between the model and AERONET data, just the mean bias. In this sense, the correlation coefficient, which is about (0.62-0.85), between WRF-Chem and AERONET AOD provides an independent evaluation of the WTF-Chem performance (see Table 4). We clarified this issue in the text (Sec. 5.2.2, last paragraph).

Also, for the AOD comparisons with both AERONET and satellite data, it is not clear if the comparisons were done under the same spatial and temporal conditions (e.g., models are sampled under clear-sky only condition or all-sky, if model and data are temporally matched).

The model, reanalysis, and observations are temporarily matched. It was mentioned in the original text (p. 17, line 364): "*Because AERONET conducts observations only during the daylight time, we interpolated WRF-Chem, MERRA-2, CAMS-OA AODs to the AERONET measurements times and then conducted time averaging to make simulated and observed AODs consistent.*"

To me, a major conclusion is that the PM2.5 concentrations over the Middle East (at least at the places the study was examined) almost never below the WHO standard because of the dominance of dust in PM2.5 which cannot or very hard to mitigate. Even if all anthropogenic emissions are shut down, the air quality in the Middle East will not improve. What is the implication for that? How to improve the air quality in the Middle East under such circumstance? This problem should be discussed.

WHO provides the guidelines, not air quality standards, that are subject to the national regulations. Yes, one of the important implications of this study is that anthropogenic pollutants in the cities are coming on top of the high aerosol background maintained by natural dust aerosols. This puts stricter requirements on anthropogenic pollution control. The effect of

natural pollution could be alleviated by using specific architectural planning, increasing in-city vegetation cover, and providing air quality forecasts to alarm the population about hazardous air quality. Our work is well in line with these ideas. The text is extended to include this discussion in page 32, in the conclusion section.

Specific comments:
Page 1, line 4: Spell out “ME”.

Fixed

Page 2, line 26: “mass budget” – it should be “emission budget”.

Fixed

Page 2, line 55-56: AVHRR was not designed to measure column aerosol properties. It was designed to observe clouds, surface temperature, and vegetation but later was expanded to retrieve aerosols over the ocean.

This sentence was removed due to the reorganization of the introduction section.

Page 3, line 57: Change “CALIOPE” to “CALIOP”.

Fixed

Page 4, line 92: “we improve the latest. . .emission. . .” does not sound appropriate. You just use the new SO2 emission data set. What is the spatial resolution of the new SO2 emission data set from Liu et al. 2018? Does it match the WRF-Chem spatial resolution? Do you have to do “downscaling” or interpolation?

It is not precisely correct. We added ship emissions to OMI-HTAP and implemented (and improved in comparison what was there) this dataset in WRF-Chem. We modified the text to make it sound more appropriate (Sec. 4.1, 2nd paragraph). We also added a reference to our recently published paper, where this emission dataset has been used. The dataset is built initially on a 0.10^0x 0.10^0 grid, and we conservatively interpolated the emissions on WRF-Chem 10 km x 10 km grid.

Page 6, line 123: Which wavelength is your chosen reference wavelength? What the wavelengths pair you used to calculate the Angstrom Exponent? Or did you use the Angstrom Exponent provided by AERONET?

We now mention (Sec. 2.1, after formula 1) in the text that we use the Angstrom exponent from AERONET that is provided for the 440-675 nm waveband.

Page 7, line 167-168: “MERRA-2 assimilates AOD at 550 nm from the AVHRR over the oceans”: This was done before the MODIS observations. MERRA-2 assimilates the MODIS AOD over the oceans since 2000.

This is correct. As explained in Table 2 from (Randles et al., 2017), since 2000 MERRA-2 assimilates MODIS and MISR data (over land and ocean) on the Terra satellite which has an equatorial overpass at 10:30 am UTC, while AVHRR has mostly orbited with the afternoon equatorial crossing time. Therefore MERRA-2 continued using AVHRR data over the ocean until 2002 when the Aqua satellite was launched. Since Aqua has an orbit with the equator

overpass at 2:30 pm, AVHRR data was no longer needed for coverage. This information was added to the revised paper (see Sec. 3.1).

Page 7, line 169: “specially processed MODIS observations. . .”: What product is that? Any references for such “non-standard” product?

Randles et al. (2017) in section 3 (subsection d) gives details on the aerosol observing system used in MERRA-2 for assimilation, including bias correction (see Sec. 3.1).

Page 7, line 184: “CAMS-OA assimilates MODIS observations”: Be more specific on what MODIS product(s) it assimilates.

CAMS-OA assimilates MODIS AQUA and TERRA AODs. It uses observations from Collection 5 since 20090901, and Deep Blue since 20150902. The clarification is added to the revised text (Sec. 3.2, last 2 sentences).

Page 8, line 200: “wavelengths larger than 450 km”???

We nudge only long waves. The text is corrected to read: “We only nudge waves with wavelengths longer than 450 km.” (Sec 4, 2nd paragraph).

Page 9, section 4.1: SO2 is oxidized to form sulfate aerosol. It is described that gas phase SO2+OH reaction is done with the RACM, but it is not clear how the heterogeneous reactions are treated. Such description should be added.

Oxidation of SO2 into sulfate is calculated within the GOCART aerosol module. Calculation of OH and other chemical reactions is done within RACM. There are no heterogeneous reactions in the RACM chemical mechanism, only gas-phase chemistry. The reference to (Stockwell et al., 1997) is in the original text. We clarified this point (Sec 4.1, 1st paragraph).

Page 9, line 234: “the first bin appears to be very poorly populated”: Why? The small particles should be transported by the winds more easily than the larger particles. Explain.

This discussion is related to sea salt. The sea salt droplets are relatively large and there is little mass accumulated in the first bin, therefore it is relatively unimportant. The text was corrected in a few places to clarify this issue.

Page 10, line 260-261: What was the error that you are correcting? Simply saying it was corrected because it was incorrectly calculated does not help the readers/users.

We have corrected three essential drawbacks in the code. These corrections have been tested and implemented in the official WRF-Chem release v4.1.3 (released on Nov 25, 2019). In the text, we provided a brief description of each of them (see Sec. 4.2). We also submitted a paper to GMD, where the effect of those errors has been quantified. *Firstly, we show that the diagnostic output of PM2.5 surface concentration was underestimated by 7% and PM10 was overestimated by 5%. Secondly, we demonstrate that the contribution of sub-micron dust particles was underestimated in the calculation of optical properties and thus, AOD was consequently underestimated by 25-30%. Thirdly, we show that an inconsistency in the process of gravitational settling led to the overestimation of the dust column loadings by 4-6%, PM10 surface concentrations by 2-4%, and the rate of dust gravitational settling by 5-10%.*

Page 10, section 4.2: I don't think the bug-fix needs to be described in a devoted section. It can be summarized in a few sentences in the model description.

Sorry, we prefer to present this information in a separate section.

Page 12, Figure 3 caption: change "for" to "from".

Fixed

Page 11-13, section 5.1: Why not compare soil moisture and precipitation, since you mentioned on page 11 that dust emission and deposition are sensitive to the soil moisture and precipitation.

The ME, and especially the Arabian Peninsula, where primary dust sources are concentrated, are arid regions. Winds are the primary driver of dust generation there. The precipitation is sporadic, and soil moisture is always low. A comparison of soil moisture and precipitation could be essential in other regions of the world.

Page 14, line 321: What is "the lower atmospheric layer"? i.e., what is the altitude range the dust is emitted into? Or is it emitted into the lowest atmospheric layer? Please clarify

In the model, dust is emitted in the lowest model layer, but here we discuss the physical process in the real atmosphere. Saltation injects dust particles at about 0.1 m height (Martin and Kok, 2017). Dust is mixed up by turbulence in the near-surface atmospheric layer. It is a well-known process, and we do not mean giving here extra details.

Page 14, line 324-325: "But because. . ." this sentence has been said in the WRF-Chem description section. It does not belong here anyway.

We agree. This sentence is removed.

Page 14, line 328-329: "WRF-Chem underestimated. . ." What is the evidence for that? Is there any reference or from your own simulation describing that problem? This contradicts the findings by Kok et al. that global models overestimate the fine mode aerosols but underestimate the coarse mode aerosols.

Please see Figure 4 and explanations therein. Kok (2011) found that the models overestimate the emission of a fine dust mode, and Adebiyi and Kok. (2020) suggested that the models underestimate the mass of the coarse (with radius r > 2.5 um) dust mode in the atmosphere almost four times, because of too fast removal processes. The argument is not entirely valid for the dust source regions like the Middle East where deposition, which Adebiyi and Kok. (2020) blame for too-quick removal of coarse dust from the atmosphere, does not have enough time to do this. Adebiyi and Kok. (2020) also analyzed dust size distribution in the near-surface layer where in-situ measurements are available.

Here we compare the column integrated dust volume size distribution from the model with the column integrated aerosol volume size distribution from AERONET and find that WRF-Chem underestimates the volume of fine particles with 0.1 um < r < 1 um and overestimate the volume of particles with 1 um < r < 2 um. We have to increase emissions in the first bin and decrease emissions in the second bin to correct this deficiency. The text is expanded to clarify this issue (Sec. 5.2.1).

Strictly speaking, our new s_p settings (see the modified text in Sec. 5.2.1, last sentence of the 2nd paragraph) are in line with (Adebiyi and Kok., 2020), as in comparison with the default s_p set we decreased the dust mass influx fraction into two finest bins 1 and 2 (0.1 um < r < 1.8 um) from 0.3175 to 0.25 (see explanation to the reviewer's comment to Page 14, line 331, below), slightly increase the mass flux fraction from 0.2275 to 0.25 for the intermediate bin 3 (1.8 um < r < 3 um), and increased dust mass flux fraction into two coarsest bins 4 and 5 (3 um < r < 10 um) from 0.455 to 0.5. We have added the new Appendix 3 to the paper to explain these points.

Page 14, line 330: How do you know that the total emitted dust mass is overestimated, since there is absolutely no measurements of dust emission?

We here do not mean to compare the simulated dust emissions with the absolute value of real dust emissions that are not measured. The measured physical quantity is AOD that in the model is controlled by emissions. If the model overestimates AOD, this is associated with overestimating dust emission (assuming we do not touch the dust removal processes). If the model excessively emits large particles, it generates higher dust mass flux than if it would generate emitting more fine particles, because finer particles produce a larger AOD per unit mass. We changed the wording to clarify this issue. (Sec. 5.2.1, 2nd paragraph)

Page 14, line 331, adjusted sp fraction: The first size bin represents clay and the rest four bins represents silt. The 0.1, 0.25, 0.25, 0.25, 0.25 fractions is based on the assumption that 10% of clay will be emitted but 100% of silt is subject to be emitted to the atmosphere based on the early work in the 1990s from Tegan. Even though these numbers are arbitrary, but the sum of adjusted silt fractions (0.15, 0.17, 0.38, 0.1) is only 0.8. Please explain why you do not account for the rest of 0.2 fraction in the silt group.

The GOCART dust emission formula (2) calculates dust mass flux into the atmosphere within five dust bins. In this formula the factor C controls the total mass flux, and the s_p coefficients split the total mass flux into five different size fractions. Following this logic, we have to assume that the sum of s_p equals 1 as we stated in the text. In the revised paper, we reiterated the sensitivity of the results to the choice of s_p and slightly readjusted the s_p values. Now we use the set of s_p =(0.15; 0.1; 0.25; 0.4; 0.1). It means that 15% of the total dust mass flux is coming as clay and 85% as silt.

In the original formulation the sum of s_p equals to 1.1. It is not crucially important, as the total flux is multiplied by the factor C that is tuned to fit the observed optical depth. So we can normalize the original s_p coefficients by dividing them to 1.1 and multiplying constant C to 1.1. It will not change any results in (2) but gives the s_p set of (0.09, 0.2275, 0.2275, 0.2275, 0.2275) that is normalized to 1 consistently with our approach. We have added the new Appendix 3 to the paper to discuss these points.

Page 16-18, Section 5.2.2: As I mentioned at the beginning, the comparison with AERONET AOD is not an independent evaluation of WRF-Chem and MERRA-2, because the WRF-Chem is "tuned" to match AERONET AOD and MERRA-2 assimilates AERONET AOD.

It is not exactly correct, at least for the model. The model is tuned to match the average value of AOD at AERONET sites using a spatially uniform time-independent factor C, which controls total dust emission. So we tuned the time-averaged AOD bias concerning available AERONET observations, not the correlation coefficient. Therefore, the high correlation of (0.62-0.85) between simulated and AERONET AOD is an independent proof of the model performance. To clarify this issue, we have expanded the text (Sec. 5.2.2, last paragraph).

Page 17, Figure 6: There are several very large spikes of AOD from the WRF-Chem simulations in Mesaira and Sede Boker in 2016. What causes these spikes?

Thanks for catching these spurious AOD spikes. We have analyzed the meteorological fields from our run for July 2016. We found that on the 27th of July 2016, a high-pressure system in the Eastern Mediterranean moving south-eastward formed high-pressure gradients reaching 3 hPa/100km. This system forced a strong gradient wind with speed exceeding 15 m/s and associated dust generation. MERRA-2 and CAMS, as well as synoptic charts based on *in-situ* observations, suggest that WRF-chem overestimates the sea level pressure gradient (see Figure A below). The preliminary analysis indicates that the boundary conditions calculated using MERRA-2 fields generated the spurious meteorological system. We re-calculated the entire July of 2016 with the boundary conditions from ERA-Interim reanalysis (see Figure A, top right panel). In the new run, the sea level pressure looks similar to observations, and spurious AOD spikes disappeared (see Figure B). We have incorporated the new July-2016 results in our analysis and corrected the figures and tables in the paper accordingly.

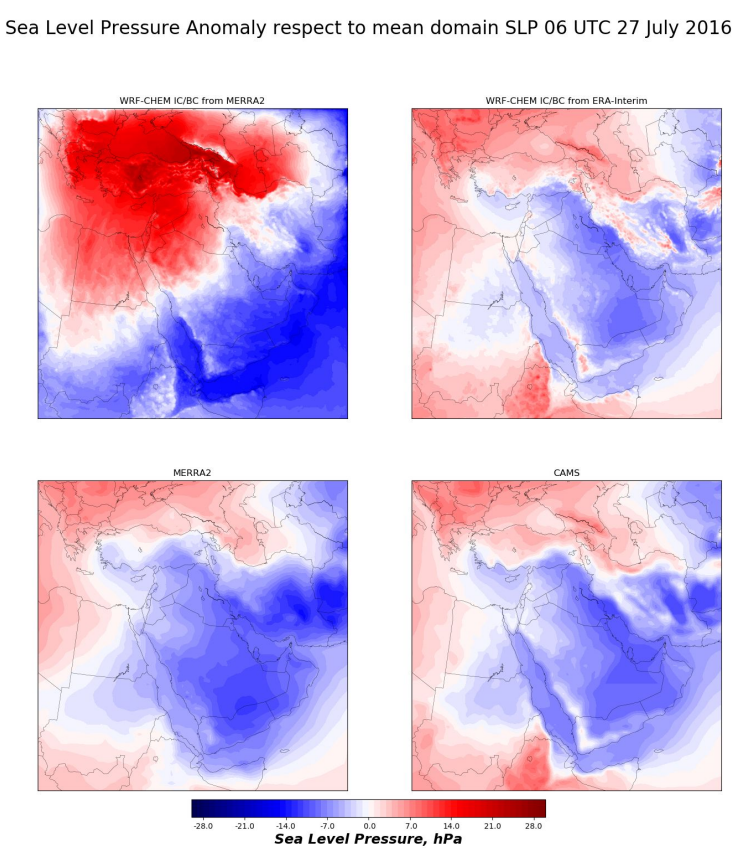


Figure A. *Sea Level Pressure anomaly from MERRA-2, CAMS-OA, and two WRF-CHEM runs with MERRA-2 and ERA-Interim boundary conditions.*

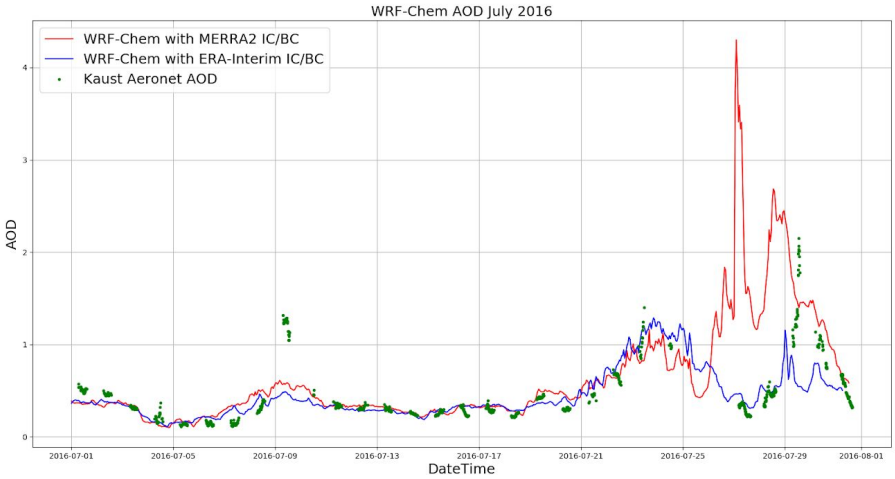


Figure B. Simulated and Observed AERONET AODs at the KAUST site for July 2016.

Page 19, Figure 7: How are the models sampled when compared to satellite data? Are they temporally matched (i.e., model results are concurrent with the satellite data, or model results are averaged for the clear sky only during the season)?

In the revised paper, we sampled WRF-Chem, MERRA-2, CAMS-OA during a day-light time (6 am-2 pm UTC or 9 am-5 pm local time) (Sec. 5.2.2, 1st paragraph). The results are not visibly different from our previous estimates when we applied 24-hour sampling.

Page 20, line 388-389: ". . .in good agreement with. . .": What is your criteria for "good agreement"? In general, such subjective statement should be avoided. Instead, you could say something more quantitative, such as "with xx%" or "correlation coefficient within xx-yy".

We agree and have corrected the sentence to account for the reviewer's comment. Now it reads, as *"Based on the comparison of WRF-Chem AOD with the AOD from MODIS and AERONET observations, we conclude that spatial and temporal WRF-Chem's AOD distribution is in good agreement with the available satellite and ground-based observations, i.e. annual mean correlation coefficient R exceeds 0.6 (see Tab. 5) and correlation with AERONET is 0.43-0.85 (see Tab. 4)."* (Sec. 5.2.3, last sentence).

Page 20, line 406: ". . .good agreement. . ." again! See my comments above.

This sentence was corrected. See the previous comment.

Page 21, line 420: sulfate ion: So for PM2.5 you only consider the mass of sulfate ion, not neutralized sulfate that exists in the atmosphere, such as ammonium sulfate? The mass of ammonium sulfate is 37% more than just sulfate ion.

We understand that ammonium sulfate has a bigger mass than sulfate ion. But in our region of interest, there is little ammonia to neutralize a significant amount of sulfate. Therefore, we assume that most of the ME anthropogenic aerosol is sulfate.

We repeat here our response to the major concern:
To support this conclusion, we have conducted laboratory analysis of the chemical composition of soil and dust deposition samples that show a little presence of organics and ammonium (Prakash et al. 2016; Engelbrecht et al., 2017). According to (Engelbrecht et al., 2017) in 2015 the annual average weight percentages of soluble ions of ammonium (NH_4) and sulfate (SO_4) in deposition samples taken at four sites at the KAUST campus are 0.05% and 2.513%, respectively. It means that available ammonium may neutralize at maximum 5% of sulfate

mass. The actual contribution of ammonium sulfate should be lower, as some ammonium may also be bound as ammonium nitrate, ammonium phosphate, or ammonium chloride.

Page 21, line 422: again, what is the reason that “the first sea salt bin is poorly populated”?

This is related to sea salt size distribution. The sea salt is a relatively coarse aerosol with very poor fine fraction therefore the fine model bin is poorly populated. We have modified the analysis and accounted for the contribution of the first sea salt bin in PM for MERRA-2 in the revised paper. The results did not visibly change.

Page 21, line 426-429, PM calculations: It should be noted that all models do not include nitrate and ammonium when calculating the PM mass. Associated error/uncertainty should be estimated.

Yes, for consistency, we show the contribution of only SO_4 for all models.

Page 24, 3rd line from the bottom: “As we have shown, WRF-Chem provides reliable estimates. . .”: What is the criteria for "reliable"? From Fig. 8, WRF-Chem underestimates PM2.5 at Jeddah and Riyadh by a factor of 2 and overestimates PM2.5 at Dammam. Its performance for total PM2.5 is inferior to CAMS. In addition, its chemical composition of PM have not been evaluated at all.

In this study, we evaluate the performance of the WRF-Chem and the best available assimilation products over the Arabian Peninsula using observed PM concentrations. This region has a poorer observation coverage in comparison with Europe or the US. Therefore model estimates are valuable to plan further analysis and mitigation measures.

The situation in the Middle East is simplified by the dominance of dust in the PM. WRF-Chem does a good job in comparison with MODIS and AERONET AOD observations, as well as predicts well the distribution of SO_2, which is the only sulfate precursor, see (Ukhov et al., 2020).

The calculations of surface aerosol concentration within a city is challenging for the 10x10-km^2 resolution model in comparison with the point observations. E.g., we do not account for in-city dust generation, although there could be a significant amount of resuspended dust. So the larger discrepancies in PM concentrations within the city are expected.

Page 27, Figure 10: The labels and legends on this figure are way too small to be legible.

Fixed

Page 28, line 516-517: Is this a “drift of sulfate”? What is the emission patter of SO2?

As it is stated in the paper, Figure 10f shows SO_4 concentration. The OMI-HTAP SO_2 combined emissions are presented in (Ukhov et al., 2020).

Page 28, line 517-521: I don’t understand what the relevancy is to refer the sulfate concentration over the US.

We added this sentence for comparison of sulfate concentrations over the US and the ME.

Page 28, line 523, MERRA-2 underestimates the SO2 emission: Do you know if indeed sulfate is too low or SO2 emission is too low in MERRA-2? Several issues here to challenge such statement. First, sulfate mass in MERRA-2 is not necessarily corresponding to SO2 emission because the aerosol masses (including sulfate) are adjusted after the AOD simulation, which has nothing to do with SO2 emission. Second, van Donkelaar's work "retrieved" PM2.5 based on the satellite AOD and the GEOS-Chem model such that the sulfate (and other aerosols) concentration is adjusted based on the adjustment of model AOD to satellite total AOD. As a result, the sulfate from van Donkelaar's work is not necessarily representative of the "true" sulfate concentrations.

We clarified the text and added a reference on our recently published paper (Ukhov et. al., 2020), where we compare different SO_2 emission dataset including EDGAR-4.2 used in MERRA-2. Ukhov et. al. (2020) shown that EDGAR-4.2 underestimates SO_2 emissions over the Arabian Peninsula in comparison with the new OMI-HTAP SO_2 emission dataset.

Page 28, line 538-539: Again, I don't understand what the relevancy of the US-EPA standard being applied here. The Saudi Arabia's standard should be used. And in line 539, now you use the WHO guidelines as reference. This is confusing.

We can not avoid the comparison of air quality in the Middle East with air quality in the US and Europe. For this purpose, we specifically discussed all PM air quality limits in Table 1, and apply them when appropriate. We specifically discuss WHO guidelines, European, US, and Saudi Arabian air pollution limits to comprehensively evaluate PM pollution in the ME, and quantify its sources.

Page 30, line 554 and 556-557: I would not emphasize "for the first time" to elevate the significance of the paper. Simply state what you've done and found is more appropriate.

We agree, the wording "for the first time" is removed from the text.

Page 30, line 564: "The air pollution in the major Middle Eastern cities is evaluated" sounds overstatement. The evaluation is rather limited to only three cities and only with PM2.5 and PM10, not all major cities and not all pollutants.

This sentence is replaced by "*We evaluated the AOD and PM air pollution over the Arabian Peninsula and in the ME major cities.*" (Sec. 6, 4th sentence)

Page 30, line 576: " improve calculation of sulfate aerosol": there is no approve that sulfate simulation is improved because there is no data to evaluate it.

The reviewer technically is correct. Strictly speaking, we did calculations with the improved SO_2 emissions (see Ukhov et al., 2020) that affected sulfate concentrations. The text is revised to clarify this point.

Page 31, line 582-583: CAMS-OA deficiency has been corrected: Then why don't you use the latest version that is available in 2019? What is the point to evaluate the results from an obsolete model version?

The paper evaluates the operational CAMS product. So for any given time, only the forecast and analyses of the current operational version is available. Further, rerunning the CAMS system (with data assimilation at the full resolution) is quite expensive. So it can not be easily redone. The re-analysis has a frozen model version for the whole period. So we always

evaluate the best product at the time. E.g., CAMS-OA had an important upgrade of the horizontal resolution of the operational system from T255 (80km) to T511 (40 km) on 21.6.2016. The CAMS-OA product is still in use and is distributed by ECMWF, so an independent evaluation of the existing product is useful. The evaluation period of 2015-2016 does not cover the time when the latest changes in CAMS-OA were introduced, so the evaluation of the newest version can not be done in the current study.

Page 31, line 589: “quite well” – again! Please avoid using such subjective statement.

We agree. The text is corrected.

References used in the response:

Adebiyi AA, Kok JF. Climate models miss most of the coarse dust in the atmosphere. Science Advances. 2020 Apr 1;6(15):eaaz9507.

Engelbrecht J, Stenchikov GL, Prakash PJ, Lersch T, Anisimov A, Shevchenko I. Physical and chemical properties of deposited airborne particulates over the Arabian Red Sea coastal plain.

Kok, J. F. (2011). A scaling theory for the size distribution of emitted dust aerosols suggests climate models underestimate the size of the global dust cycle. Proceedings of the National Academy of Sciences, 108(3), 1016-1021.

Martin, R. L., & Kok, J. F. (2017). Wind-invariant saltation heights imply linear scaling of aeolian saltation flux with shear stress. *Science advances*, *3*(6), e1602569.

Osipov S, Stenchikov GL, Brindley H, Banks J. Diurnal cycle of the dust instantaneous direct radiative forcing over the Arabian Peninsula.

Prakash PJ, Stenchikov GL, Tao W, Yapici T, Warsama BH, Engelbrecht J. Arabian Red Sea coastal soils as potential mineral dust sources.

Randles CA, Da Silva AM, Buchard V, Colarco PR, Darmenov A, Govindaraju R, Smirnov A, Holben B, Ferrare R, Hair J, Shinozuka Y. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. Journal of Climate. 2017 Sep;30(17):6823-50.

Stockwell WR, Kirchner F, Kuhn M, Seefeld S. A new mechanism for regional atmospheric chemistry modeling. Journal of Geophysical Research: Atmospheres. 1997 Nov 27;102(D22):25847-79.

Ukhov A, Mostamandi S, Krotkov N, Flemming J, da Silva A, Li C, Fioletov V, McLinden C, Anisimov A, Alshehri Y, Stenchikov G. Study of SO2 pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem simulations. Journal of Geophysical Research: Atmospheres. 2020 Mar 6:e2019JD031993.

Responses to Reviewer 2

General comments

The manuscript by Ukhov et. al., presents a detailed comparison of WRF-Chem, MERRA-2 and CAMS aerosol data with respect to the two MODIS data products, AERONET and ground-based network of PM measurements. The paper is well written and extensive sets of data are considered for comparison which represents aerosol optical depth, PM10, PM2.5 and their spatial and temporal patterns. In addition to the comparison, the composition of aerosol among dust, sulfate, sea salt and other constituents have been discussed. The impact on aerosol air pollution has also been investigated. This study could have been completed by also including some comparison for vertical profiles of aerosol extinction or various components of aerosol (e.g. dust, sulfate), with the measurements (if available) or at least among the model and assimilated products. I have a few major and several minor concerns with the manuscript, which upon being addressed, I recommend publication in ACP.

We thank the reviewer for the valuable comments. We agree that comparing vertical aerosol distribution in the models and in observation is very useful. But it is not a small side issue. To address it, we have established a micropulse lidar (MPL) site at the KAUST campus and observed the aerosol vertical profile since 2014. There is no way currently to separate in observations the vertical profile of dust from that of sulfate and other aerosols. We are not sure we have enough space in this paper to address the important issue of the aerosol vertical profile properly; therefore, we refrain from doing this here. This work is mostly the subject of another paper submitted to ACP recently.

All references used to support our responses to the reviewer's comments are presented at the end of the text. The reviewer's questions are in black. Our answers are in blue.

Major issues:

1. The manuscript primarily focuses on the various aerosol product and WRF- Chem code has been modified to calculate these parameters. In section 4.2, authors mention that the code modification will be published in a forthcoming publication. Since the data produced for use in this publication is simulated with modified model code, yet not peer-reviewed, I can only recommend publication after the technical publication.

The model modifications have been reviewed, tested, and implemented in the v4.1.3 official version of the WRF-Chem code (released on Nov 25, 2019). The forthcoming paper in GMD is about the quantitative evaluation of the effects of those changes on simulations. The GMD paper is independent of the current study, and, we believe, should not delay the publication of the present manuscript.

2. For sections 2.2, 5.2.2 and 5.2.3 authors use the MODIS combined deep blue (DB) and dark target (DT) product. It a level 3 gridded product at a much coarse spatial resolution of 1◦ × 1◦. DB has poorer performance over water, while DT has limitation over land. In my opinion, authors should use separate DB and DT, level 2 gridded products, which are available at much finer (10km × 10km) resolution (comparable to WRF-Chem and MAIAC). Moreover, level 2 product also allows the possibility of applying a quality assurance criterion, which has shown improvement in the comparison previously (for e.g. Liu, N., et al. (2019)).

Following the reviewer's recommendation, we calculated the 10x10 km^2 MODIS DB&DT level 2 AOD product. The AOD fields with 10x10 km^2 and 100x100 km^2 resolution are shown below. We see that qualitatively the AOD structure is similar at both resolutions, but 10x10 km^2 fields have much more fine details. In the revised manuscript, we now use the 10x10 km^2 resolution MODIS DB&DT product, and all tables in the revised paper are corrected accordingly.

Figure A. 10 x 10 km² MODIS AOD

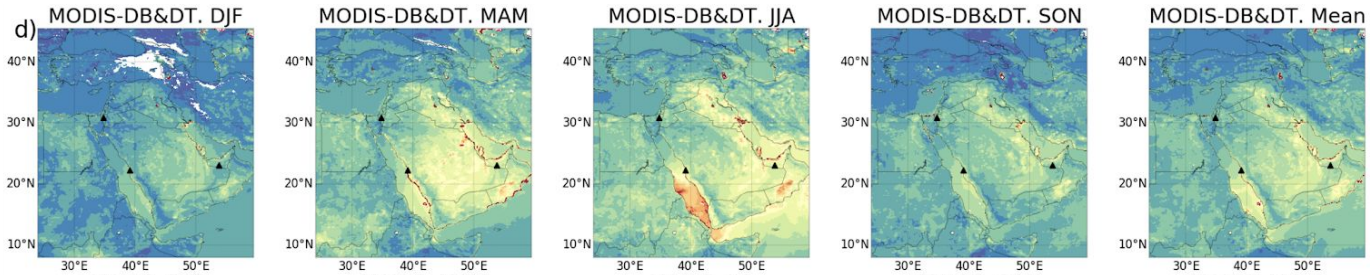
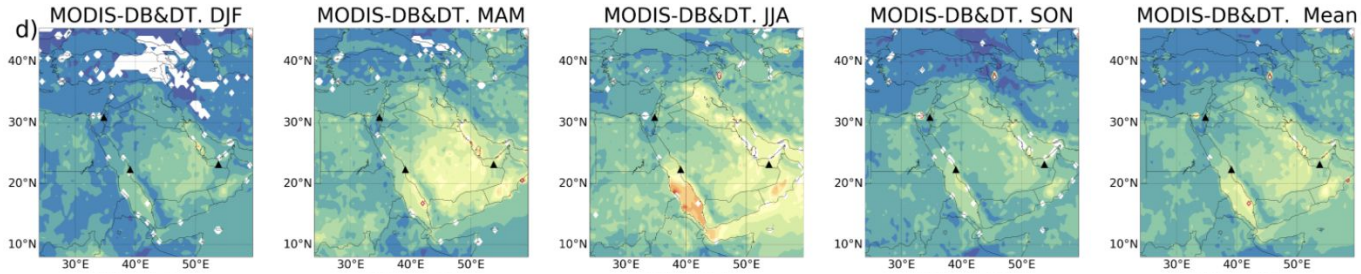


Figure B. 100 x 100 km² MODIS AOD



2. For comparison with MODIS data products, the model data should be sampled around satellite overpass or at most averaged ± 1h around satellite overpass. Further, in order to avoid sampling bias, only those days should be considered for calculating seasonal means when both measurements (AERONET/Satellite) and model (or assimilation) data are available.

Because we analyze the seasonal mean AOD fields over the entire ME, we have to use multiple overpasses to compile a map for the whole domain. Therefore AOD's from WRF-Chem, MERRA-2, CAMS-OA are sampled at the day-light time (6 am-2 pm UTC or 9 am-5 pm local time). We added this explanation in the text (Sec. 5.2.3, 1st paragraph).

There is no missing data in the model and the reanalysis outputs, and only a little in the MODIS products, therefore all available observations and model outputs were used for calculating seasonal means. The undefined pixels detected in observations are synchronously excluded from the statistical analysis in all datasets.

4. The introduction is very long in general and can be curtailed by only keeping the content most important to the study. Some (not all) suggestions: - Lines 55-60- MISR, AVHRR and CALIOPE are not relevant to this study. Some restructuring is also needed. For example, the description of the work to be presented in the text fits better towards the end of the introduction. In line 75, authors mention about evaluation to be presented in the subsequent section but this is followed by further literature review. Line 93 again starts with the highlight of work to be presented in this work.

We revised the text to remove the redundancies

5. The conclusion needs to be curtailed. Redundancy in the conclusion can be reduced. Some examples (not all): Lines 558-559 and lines 571-543; Lines 560- 561 and 574-575. Numbers should be provided in conclusion rather than only qualitatively stating “overestimate” / “contribute” etc.

We revised the conclusion and made it more concise.

Detailed Comments:

1. Abstract: Please use abbreviations only after providing their full form at the first use (e.g. ME). Abstract does not do justice to the manuscript. Some more key finding should be added.

Added full form for ME (Middle East). More key findings were added.

2. Line 6: WRF-CHEM code was modified but this is not described in detail in this manuscript. Authors wish to publish it as separate manuscript and hence this does not fit to abstract.

We agree and have removed this sentence from the abstract.

3. Line 15: rich – reach

Fixed

4. Lines 15-16: Contribution of both organic matter and black carbon are negligible. Is is important to mention this comparison?

The PM speciation is vital to plan air quality mitigation measures. There are few observations available, so model results that provide spatially resolved information are valuable for understanding the effect of different types of aerosols on air quality in the ME.

5. Line 35: Essential – Important/crucial

Fixed

6. Line 46: PM10 and PM2.5 are defined with respect to “aerodynamic diameter”.

Yes

7. Lines 61-63: Is it justified to compare the 21 days’ mean with air quality regulation standards for 1 year. Please note that some of the measured mean PM10 concentrations are smaller than 24 hours’ limit.

Due to the modification of the introduction, this part was removed from the text.

8. Lines 79-85 : What are the conclusions of these comparisons?

The text was rephrased to: “*These data assimilation products adequately reproduce AOD and PM concentrations at different regions of the world (Provençal et al., 2017; Buchard et al., 2017; Cesnulyte et al., 2014; Cuevas et al., 2014).*”

9. Lines 89-92: Given that mineral dust contributes 75-95% of the PM, how much discrepancy is caused by outdated emission inventories in MERRA-2 and CAMS-OA?

The anthropogenic emissions certainly make an essential contribution in the air pollution in the cities and this information is important for air quality control in urban centers. To clarify this issue we have added the following sentence (p.4 line 85): “*E.g.,SO2 emissions used in MERRA-2 and CAMS-OA differ by 45-50% in some ME regions (Ukhov et al., 2020)”.* In the 2nd paragraph of the Sec. 4.1 we also mentioned that 14 previously unaccounted SO_2 point sources located in the ME were included in the new OMI-HTAP dataset.

10. Line 113: What are CIMEL and PREDE?

To clarify the text the sentence is updated to read "*AERONET comprises more than 1000 observation sites equipped with CIMEL sunphotometers and PREDE skyradiometers manufactured in France by CIMEL and in Japan by PREDE.*"

11. Lines 118-120: Authors should also provide a statistical comparison for the case when only cloud screened and quality-assured data are used in the results and discussion.

In our analysis now we use "Utilizes AERONET AOD, which is pre- and post-field calibration applied, automatically cloud cleared and manually inspected (Level 2.0 AOD)." The text (Sec 2.1, 1st paragraph) is clarified to read: *"We utilized level 2.0 (cloud screened and quality assured) AERONET AOD data."*

12. Line 119, 122: Angstrom – Ångström

Fixed

13. Line 139: MAIAC also provides AOD at 470nm.

Added

14. Section 2.2: Please mention the Quality assurance filter criteria if applied!

We did not use a quality assurance filter.

15. Line 153? quarterly refers to what?

"Quarterly" refers to the calibration audit. Sentence has been rephrased to read (Sec 2.3. Last sentence) "*...audit is conducted quarterly by Ricardo-AEA Ltd...*".

16. Line 164 (DMS)

Fixed

17. Line 173: This line is not clear to me.

CAMS-OA is the operational analysis, not reanalysis. The model, its horizontal resolution, and assimilation routine are improving on the way, so we always use the best available product. E.g., the important upgrade of the horizontal resolution of the operational system from T255 (80km) to T511 (40 km) was accomplished on 21.6.2016.

18. Line 179 and later in the text: dustbins – dust-bins.

Fixed

19. Line 2019: OH is hydroxyl radical and not "Hydroxide radical".

Fixed

20. Line 219: I had difficulties understating the treatment of PM, BC and OC emissions. Black carbon, organic carbon and dust, these are already included in PM. So if the emission of both PM and its constituents are specified separately, this would end up in doubling of certain constituents of PM.

Reviewer meant line 229 not 219. We agree that it sounds confusing, because we followed the emission categories used in the WRF-Chem. It meant that the "PM" emissions comprise the additional aerosol biogenic and fossil components. Now the text reads as (Sec 4.1, 2nd paragraph): "*All other constituents (other PM from biogenic and fossil components, black and organic carbon, etc.), ...*"

21. Equation 3: Use of S in both LHS and RHS are confusing. I would suggest using Smod or S' or something different.

We agree, this is confusing. We defined S' as a modified topographic source function.

22. Lines 250-253: How is the value of C=0.5 achieved? The tuning of C with respect to measured AOD should be discussed in more detail.

WRF-Chem is tuned to reduce the seasonal mean AOD biases with respect to AERONET observations. The value of *C*=0.5 obtained in the course of multiple WRF-Chem runs with different values of *C* gives the best AOD fit. Three references with detailed description of the tuning procedure were provided in the original text (page 10 line 248): (Kalenderski et al., 2013; Jish Prakash et al., 2015; Anisimov et al., 2017).

23. Section 4.2: How are the diagnostic output of PM are different from those calculated in section 5.3?

As we mention in the original text, if we would use the default WRF-Chem v3.7.1 code we would overestimate the PM10 and underestimate the PM2.5 surface concentrations. For typical Middle East conditions, *diagnostic output of PM2.5 surface concentration could be underestimated by 7% and PM10 surface concentrations could be overestimated by 5%. (See Sec. 4.2)*

24. Section 5.1 Lines 271-275 fit better for methods/domain description.

Sorry, we prefer to concentrate on the ME climate description in section 5.1.

25. Figure 2: What is the physical significance of the topographic source function? Do the high values represent higher dust emission potential?

The topographic source function defines a spatial pattern of emission. The factor *C* - controls the total amount of emissions. The topographic source function has been built under an assumption that low-land areas accumulate fine-scale material (Ginoux et al., 2001). The areas with the higher values of source function generate higher dust emission flux, see eq. (2).

26. Line 283: Missing ")".

Fixed

27. Lines 305-309: Higher R and lower RMSD for V are not specific only for summers.

We agree. The text has been corrected accordingly.

28. Table 3: How are the statistics for Autumn and Spring

The dry subtropics have essentially two seasons, warm Winter, and very warm Summer. The intermediate seasons are not so essential. We prefer not to spend much time on their discussion.

29. Table 3,4 and 5: Slope/Bias should also be provided in addition to the R and RMSD. These quantities provide an idea about overestimation/underestimation/trend.

Table 4 shows bias for the AOD time series. We added the scatter plots for the AOD time series in Fig. 6. The bias has been added for the spatial distributions of AOD in Table 5. Figure 3 shows that the seasonal mean wind field in WRF-Chem and both reanalyses do not have systematic differences. We believe it is not needed to add bias for the wind in Table 3.

30. Line 315: Aerosol content is also characterized by other quantities apart from AOD.

We talk here about the satellite observed quantities. Of course, one needs aerosol size distribution or mass extinction coefficient to convert AOD to mass loading.

31. Lines 327-330: It would be nice to see the underestimation/overestimation with default sp fraction and its magnitude as a figure (at least in appendix).

In the revised paper, we reiterated the sensitivity of the dust size distribution to the choice of s_p and slightly readjusted the s_p values. Now we use the set of s_p=(0.15; 0.1; 0.25; 0.4; 0.1). Below we compare the size distributions obtained in the simulations with this updated and the default s_p=(0.1; 0.25; 0.25;0.25;0.25) values for summer of 2015. Using the updated s_p values improves the size distribution fit (see Figure below). We have added the new Appendix 3 to the paper to discuss these points.

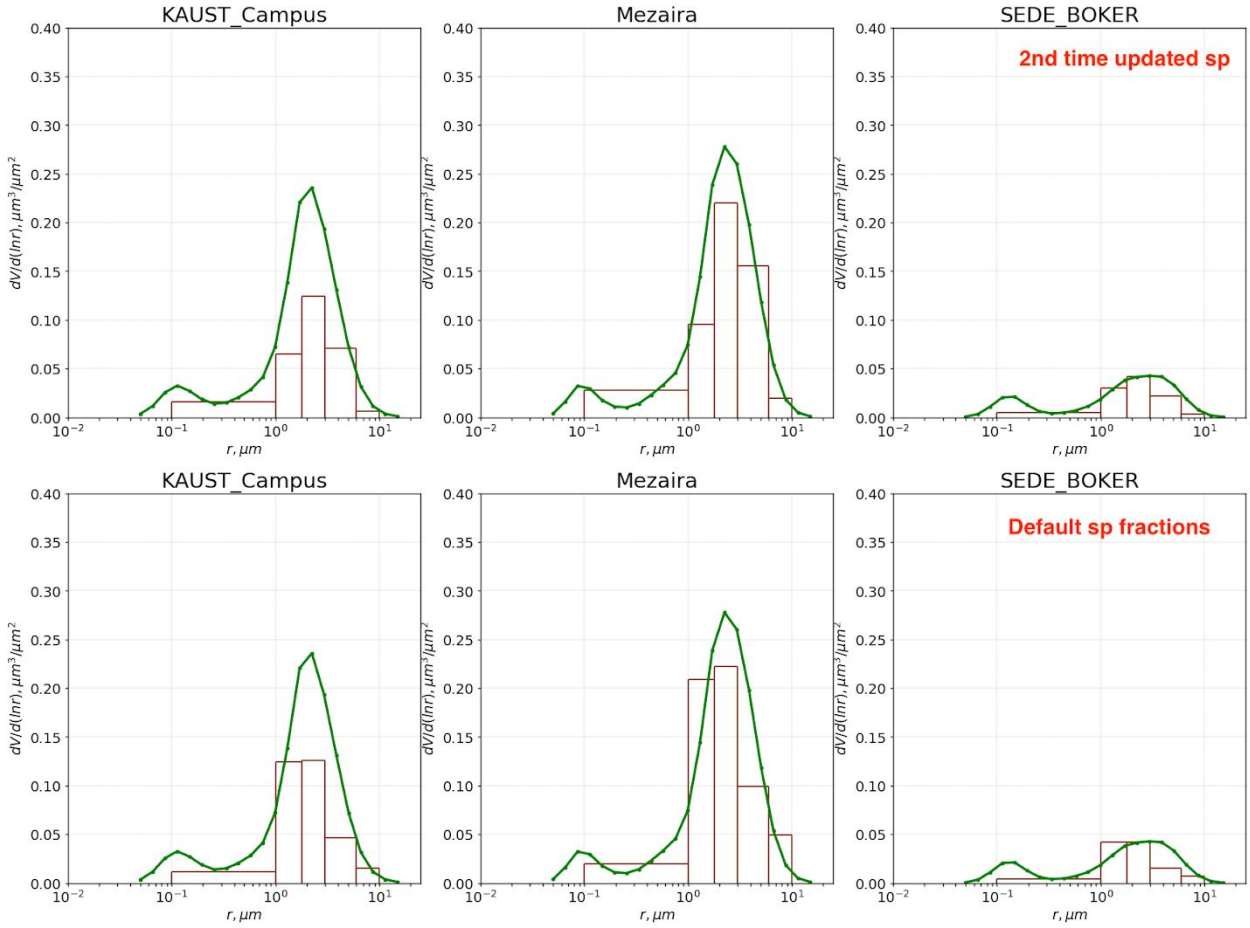


Figure C. Volume size distribution at KAUST AERONET site averaged for JJA of 2015 from WRF-Chem simulation with default sp={0.1,0.25,0.25,0.25,0.25} fractions (bottom) and updated sp={0.15; 0.1; 0.25; 0.4; 0.1} (top).

32. Line 345: This line should only be kept if the evaluation of updated CAMS-OA is shown in the manuscript.

Figure 4 compares the volume size distributions from WRF-Chem, MERRA-2, and CAMS-OA with the AERONET retrievals. The paper evaluates the operational CAMS product, CAMS-OA. So, for any given time, only the forecast and analyses of the current operational version are available. Further, rerunning the CAMS system (with data assimilation at the full resolution) is quite expensive. So it can not be easily redone. The product is still in use and is distributed by ECMWF, so an independent evaluation of the existing product is useful. The evaluation period of 2015-2016 does not cover the time when the latest changes in CAMS-OA were introduced, so the comparison can not be made in the current study.

33. Figure 6: Please mention that panel A corresponds to 2015 and B correspond to 2016.

The caption has been changed to address this issue.

34. Line 366: At a given location, up to 4 measurements are possible on several days due to overlap of two orbits each for TERRA and AQUA.

Thanks. The text is corrected.

35. Section 5.2: I was surprised to see that MAIAC underestimates AOD with respect to AERONET. The evaluation of MAIAC by Lyapustin et. al., shows overestimation at all the three AERONET sites shows in this study. Authors should address, why even for a similar dataset, an underestimation is observed in this study by MAIAC. Authors could also refer to the finding of Liu et. al., 2019, where they have found that applying a QA filter significantly reduces the Deep blue (over land) AOD from MODIS over China. There are other evaluation studies (e.g. Liu et al., 2019, Mhawish et. al., 2019), which have found MAIAC to be more accurate than Deep blue and Dark Target. Authors should address, why for their domain this is not the case.

According to (A. Lyapustin personal communication, April 2020), MAIAC underestimates AERONET in the ME (at *KAUST_Campus* and *Mezaira* sites). So, our results are consistent with this. We do not apply a QA filter in our calculations.

36. Figure 7, I wonder how there are NAN values at around 40 ◦N 40 ◦E in MODIS DB&DT products in the annual mean but there are no NAN values in MAIAC annual mean. If the seasonal NAN values are removed by annual mean, this should hold valid for both the MODIS data products. I would recommend the authors to recheck the calculation of spatial means. Please also indicate the location of three AERONET site in Figure 7. This would help the reader to follow the discussion.

There are some undefined pixels in the MAIAC product that we referred to as NANs. This confusing terminology has been corrected in the revision. In Figure 7 in the original manuscript, we interpolated MODIS and MAIAC AODs to the MERRA-2 grid. That caused some discrepancies, e.g., led to an artificial increase in undefined areas. Now model outputs and satellite products are plotted in its original resolution (see Figure 7 in the manuscript and figure below). We have recalculated all statistical characteristics in Table 4 using MAIAC on its original grid. Table 4 shows that MAIAC now compares better with AERONET than MODIS DB&DT in terms of bias and correlation coefficient. Locations of 3 AERONET sites are now shown on the plots, as requested by the reviewer. We also fixed an error in the calculation of spatial means.

Figure D. MODIS 100x100km^2 and MAIAC (interpolated on MERRA-2 grid) and error in calculation of seasonal means.

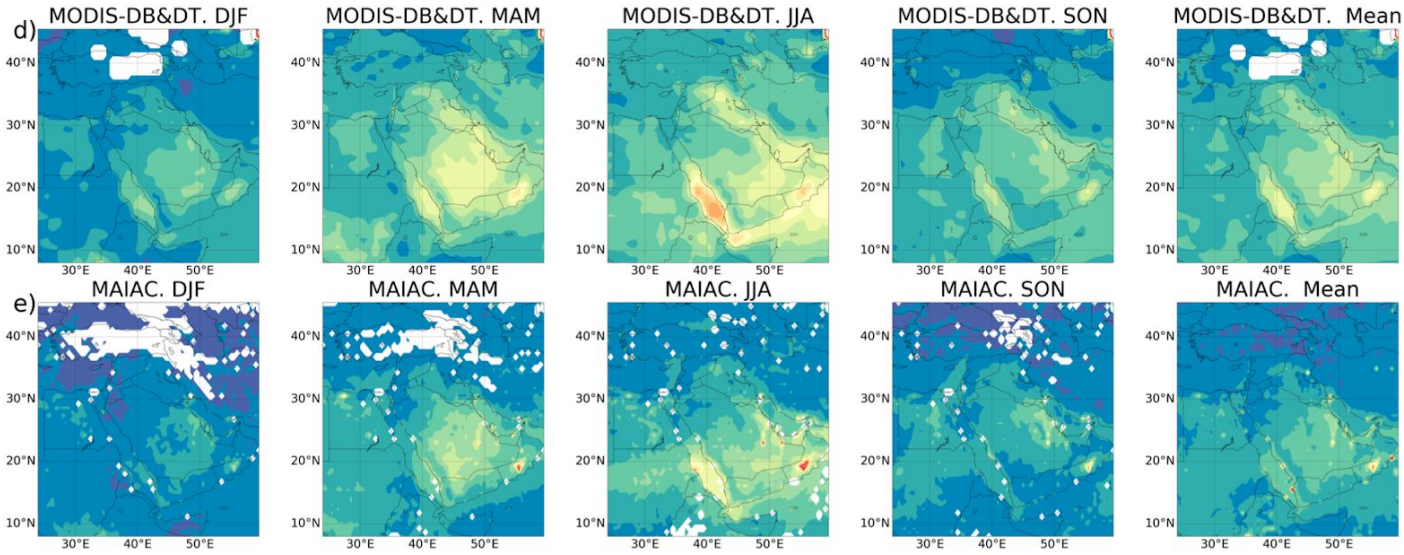
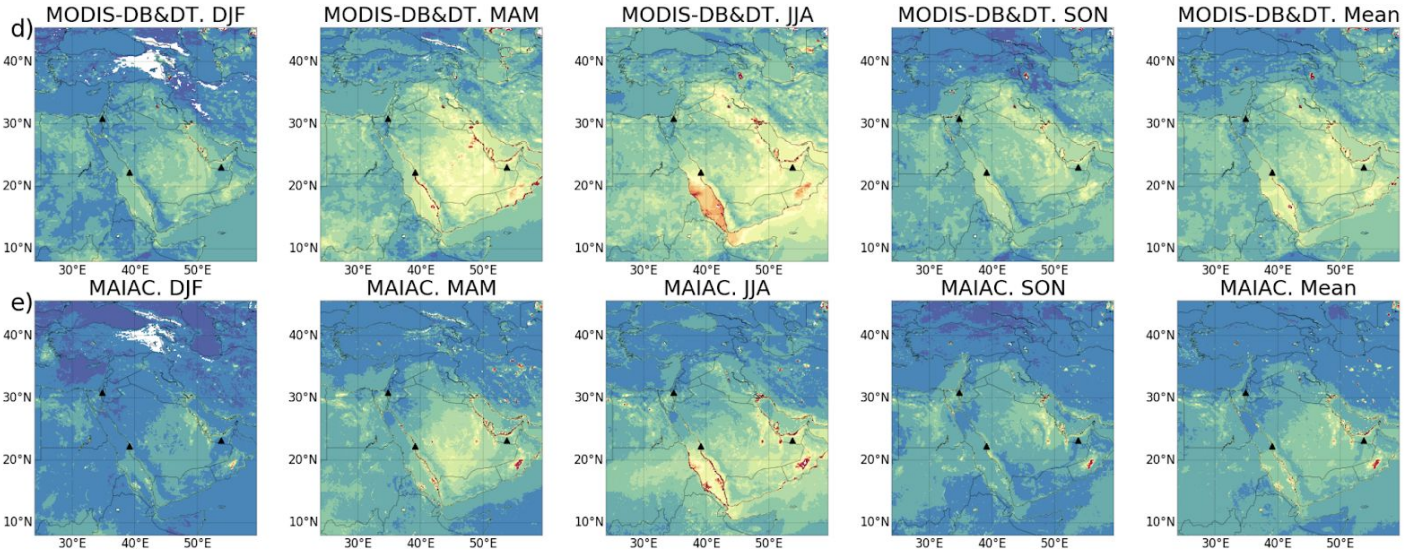


Figure E. MODIS 10x10km^2 and MAIAC (on its original resolution). Fixed error in calculation of seasonal means.



37. Section 5.3: Please provide references from where the formulas for calculation of PM2.5 and PM10 are adapted. What is the rationale behind the choice of the coefficients used in equation 4 and 5?

These formulas are taken from the WRF-Chem source code and Copernicus knowledge base (https://confluence.ecmwf.int/display/CUSF/PM10+and+PM25+global+products). The coefficients in those formulas account for the contributions of dust and sea salt bins to PM2.5 and PM10. Dust and sea salt have different bin sizes in WRF-Chem and CAMS-OA; therefore, those coefficients are different for WF-Chem and CAMS-OA. Both WRF-Chem and MERRA-2 use the GOCART aerosol module with the same bin sizes; therefore, the coefficients for WRF-Chem and MERRA-2 are the same.

38. Sections 5.3, 5.3.3, 5.3.4 and 5.4: Air Quality and Air pollution are very broad terms which also include trace gases in addition to the aerosol. Hence, the subtitles of these sections should be made more specific.

The reviewer is generally correct and we changed the titles 5.3, 5.3.3, 5.3.4 (there is now 5.4) to "*PM Air Pollution*", "*Spatial patterns of PM air-pollution*", "*PM air-pollution in the ME major cities*", correspondingly.

We have to mention here that according to the US EPA the air quality index is defined by the leading pollutant, which in the ME, almost exclusively is PM2.5.

39. Line 443: How does the calculated concentration of 298 µg/m3 compare against the measurements?

It is not possible to compare because there are no observations of total dust concentration, only PM2.5 and PM10 are available from MODON observations. Daily average PM10 surface concentration on 8 July 2016 registered by Jeddah AQMS is 184 µg/m3.

40. Figure 8 and 9: Please provide the uncertainty marks in the histogram which represent the variability over the mean.

Uncertainty marks are shown now in both figures. PM2.5 and PM10 measurement error is +/- 5%. This information was added to the text (Sec. 2.3).

41. Lines 496-503 Authors evaluate the PM2.5/PM10 ratio to evaluate the dominance of coarse/fine particles. A more quantitative evaluation would be PM10-PM2.5, which provides a more exclusive number for larger particles.

Both PM10-PM2.5 and PM2.5/PM10 are informative. The PM2.5/PM10 ratio is widely accepted in air pollution literature, e.g., see Gehrig et al., 2003; Parkhurst et al., 1999; Querol et al., 2001. Therefore we prefer to use this ratio in this study.

42. What are the major non-sulfate constituents in total PM2.5 non-dust aerosol?

In coastal areas it is sea salt and organic matter, over inland only organic matter. BC has a very little effect (see Table 6).

43. Lines 587-588: In addition to the AOD retrieval uncertainty, there are several other differences e.g. Spatial resolution, Quality assurance filter which contribute the observed difference.

The discussion is expanded to add more detailed explanation.

44. Line 600: Please use the same convention for the naming of seasons. “Fall” season is nowhere discussed in the text and appears for the first time in the conclusion.

Changed to autumn.

45. Line 609: Air quality should be replaced with PM air quality.

According to the US EPA the air quality is evaluated based on the concentration of the most significant leading pollutant, which are PM2.5 and PM10 in the Middle East, so PM air quality and air quality terms are almost equivalent in the ME.

References:

[1]. Liu, N., et al. (2019), Evaluation and comparison of multiangle implementation of the atmospheric correction algorithm, Dark Target, and Deep Blue aerosol products over China, Atmos. Chem. Phys., 19(12), 8243-8268, doi:10.5194/acp-19-8243-2019.

[2]. Mhawish, A., et al. (2019), Comparison and evaluation of MODIS Multiangle Implementation of Atmospheric Correction (MAIAC) aerosol product over South Asia, Remote Sensing of Environment, 224, 12-28, doi:10.1016/j.rse.2019.01.033.

[3]. Lyapustin, A., et al. (2018), MODIS Collection 6 MAIAC algorithm, Atmos. Meas. Tech., 11(10), 5741-5765, doi:10.5194/amt-11-5741-2018.

References used in the response:

Anisimov A, Axisa D, Kucera PA, Mostamandi S, Stenchikov G. Observations and cloud-resolving modeling of Haboob dust storms over the Arabian Peninsula. Journal of Geophysical Research: Atmospheres. 2018 Nov 16;123(21):12-47.

Gehrig R, Buchmann B. Characterising seasonal variations and spatial distribution of ambient PM10 and PM2. 5 concentrations based on long-term Swiss monitoring data. Atmospheric Environment. 2003 Jun 1;37(19):2571-80.

Ginoux P, Chin M, Tegen I, Prospero JM, Holben B, Dubovik O, Lin SJ. Sources and distributions of dust aerosols simulated with the GOCART model. Journal of Geophysical Research: Atmospheres. 2001 Sep 16;106(D17):20255-73.

Kalenderski S, Stenchikov G, Zhao C. Modeling a typical winter-time dust event over the Arabian Peninsula and the Red Sea. Atmospheric Chemistry & Physics Discussions. 2012 Oct 1;12(10).

Querol X, Alastuey A, Rodriguez S, Plana F, Ruiz CR, Cots N, Massagué G, Puig O. PM10 and PM2. 5 source apportionment in the Barcelona Metropolitan area, Catalonia, Spain. Atmospheric Environment. 2001 Dec 1;35(36):6407-19.

Prakash PJ, Stenchikov G, Kalenderski S, Osipov S, Bangalath H. The impact of dust storms on the Arabian Peninsula and the Red Sea. Atmospheric Chemistry & Physics Discussions. 2014 Dec 2;14(13).

Parkhurst WJ, Tanner RL, Weatherford FP, Valente RJ, Meagher JF. Historic PM2. 5/PM10 concentrations in the southeastern United States—Potential implications of the revised particulate matter standard. Journal of the Air & Waste Management Association. 1999 Sep 1;49(9):1060-7.

Ukhov A, Mostamandi S, Krotkov N, Flemming J, da Silva A, Li C, Fioletov V, McLinden C, Anisimov A, Alshehri Y, Stenchikov G. Study of SO2 pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem simulations. Journal of Geophysical Research: Atmospheres. 2020 Mar 6:e2019JD031993.

Assessment of natural and anthropogenic aerosol air pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem model simulations

Alexander Ukhov[1], Suleiman Mostamandi[1], Arlindo da Silva[2], Johannes Flemming[3], Yasser Alshehri[1], Illia Shevchenko[1], and Georgiy Stenchikov[1]

[1]King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]European Centre for Medium-Range Weather Forecasts, Reading, UK

Correspondence: Georgiy Stenchikov (georgiy.stenchikov@kaust.edu.sa)

Abstract. Modern-Era Retrospective analysis for Research and Applications v.2 (MERRA-2), Copernicus Atmosphere Monitoring Service Operational Analysis (CAMS-OA)~~data assimilation products~~, and a high-resolution regional Weather Research and Forecasting model ~~(10 resolution)~~ coupled with Chemistry (WRF-Chem) were used to evaluate natural and anthropogenic ~~aerosol~~ Particulate Matter (PM) air pollution in the ~~ME~~ Middle East (ME) during 2015-2016. ~~Satellite and ground-based AOD~~Two Moderate Resolution Imaging Spectrometer (MODIS) retrievals: combined product Deep Blue and Deep Target (MODIS-DB&DT), Multi-Angle Implementation of Atmospheric Correction (MAIAC), and Aerosol Robotic Network (AERONET) aerosol optical depth (AOD) observations, as well as *in situ* ~~Particulate Matter (PM)~~ PM measurements for 2016, were used for validation of the WRF-Chem output and both assimilation products.

MERRA-2 and CAMS-OA assimilate AOD observations. WRF-Chem ~~code was modified to correct the calculation of dust gravitational settling and aerosol optical properties. The~~ is a free-running model, but dust emission in WRF-Chem is ~~calibrated to fit Aerosol Optical Depth (AOD)~~ tuned to fit AOD and aerosol volume size distributions obtained from ~~Aerosol Robotic Network (AERONET) observations~~AERONET. MERRA-2 was used to construct WRF-Chem initial and boundary conditions both for meteorology and chemical/aerosol species. ~~SO_2~~ SO_2 emissions in WRF-Chem are based on the novel ~~NASA SO_2 emission datasetthat reveals unaccounted sources over the ME.~~ OMI-HTAP SO_2 emission dataset.

The correlation with the AERONET AOD is highest for MERRA-2 (0.72-0.91), MAIAC (0.63-0.96), and CAMS-OA (0.65-0.87), followed by MODIS-DB&DT (0.56-0.84) and WRF-Chem (0.43-0.85). However, CAMS-OA has a relatively high positive mean bias with respect to AERONET AOD. The spatial distributions of seasonally averaged AODs from WRF-Chem, assimilation products, and MAIAC are well correlated with MODIS DB&DT AOD product. MAIAC has the highest correlation (R=0.8) followed by MERRA-2 (R=0.66), CAMS-OA (R=0.65), and WRF-Chem (R=0.61). WRF-Chem, MERRA-2, and MAIAC underestimate, and CAMS-OA overestimates MODIS-DB&DT AOD.

The simulated and observed PM concentrations might differ of a factor of two, because of it is more challenging to the model and the assimilation products to reproduce PM concentration measured within the city. Although aerosol fields in WRF-Chem and assimilation products are ~~quite~~ entirely consistent, WRF-Chem, due to its higher spatial resolution and better ~~SO_2~~ SO_2

emissions, is preferable for analysis of regional air-quality over the ME. The WRF-Chem's PM background concentrations exceed the World Health Organization (WHO) guidelines over the entire ME. ~~The~~ Mineral dust makes the major contributor to PM (≈75–95%) ~~is mineral dust. In the ME urban centers and near oil recovery fields~~compared to other aerosol types. Near and down the wind from the SO_2 emission sources, non-dust aerosols (primarily sulfate) contribute up to ~~26~~30% into PM2.5. The contribution of sea salt into PM ~~can rich up to~~ in coastal regions can reach 5%. The ~~contribution~~ contributions of organic matter~~into PM prevails over black carbon.~~ , black and organic carbon into PM over the Middle East are insignificant. In the major cities over the Arabian peninsula, the 90th percentile of PM_{10} and $PM_{2.5}$ daily mean surface concentrations exceed the corresponding Kingdom Saudi Arabia air-quality limits. The contribution of the non-dust component to $PM_{2.5}$ is $< 25\%$, which limits the emission control effect on air quality. The mitigation of the dust effect on air quality requires the development of environment-based approaches like growing tree belts around the cities and enhancing in-city vegetation cover. The presented in this study WRF-Chem configuration could be a prototype of a future air quality forecast system that warn the population against air pollution hazards.

1 Introduction

PM is a complex mixture of sea salt, sulfate, black carbon, organic matter, and mineral dust, suspended in the air. The dramatic increase in the level of air pollution in developing countries over the last decades is forced by rapid economic and population growth, burning of fossil fuels, construction, and agricultural activities (Janssens-Maenhout et al., 2015). However, the primary cause of air pollution in the ME is mineral dust, and it is on the rise (Klingmüller et al., 2016). Along with Asia and Africa, the ME significantly contributes to global dust emissions, which are in the range of 1000-2000 $\mathrm{Tg/year}$ (Zender et al., 2004). According to Prospero et al. (2002), the Middle East and North Africa (MENA) regions account for about half of global dust emissions. By integrating surface emissions in MERRA-2 reanalysis we found that, the total global dust emission averaged over the 2015-2016 period is about 1600 $\mathrm{Tg/year}$, right in the middle of the Zender et al. (2004) estimate. The dust emission from our simulation domain (see Fig. 1) that covers the ME and nearby areas is about 500 $\mathrm{Tg/year}$, contributing ≈30% to the global dust ~~mass~~ emission budget. Also, frequent inflows of pollutants from Europe, Africa, and India, worsen the air quality over the Arabian Peninsula (Jish Prakash et al., 2015; Kalenderski et al., 2013; Notaro et al., 2013; Reid et al., 2008; Mohalfi et al., 1998; Kalenderski and Stenchikov, 2016; Parajuli et al., 2019). Because of the large amount of dust, the ME is one of the most polluted areas in the world. Located in the center of the northern subtropical dust belt, the Arabian Desert is the third-largest (after the Sahara and the East Asian deserts) region of dust generation, where dust plays a significant role in controlling regional climate (Cahill et al., 2017; Banks et al., 2017; Jish Prakash et al., 2016; Farahat, 2016; Kalenderski and Stenchikov, 2016; Munir et al., 2013; Alghamdi et al., 2015; Lihavainen et al., 2016; Anisimov et al., 2017; Osipov and Stenchikov, 2018).

In addition to natural dust aerosols, the ME receives high concentrations of anthropogenic PM ~~(Karagulian et al., 2015; Al-Taani et al., 20~~(Karagulian et al., 2015; Al-Taani et al., 2019; Alharbi et al., 2015; Khodeir et al., 2012). The most ~~essential~~ important anthropogenic aerosol in ME is sulfate with ~~SO_2~~ SO_2 as a precursor, the contributions of other types of aerosols in PM, sea salt, organic matter, and black carbon are of lesser importance in the ME (Randles et al., 2017). ~~SO_2~~ ME emits about 10% of the

total global anthropogenic SO_2 (Klimont et al., 2013). SO_2 produced in the course of power generation, water desalination, and oil recovery operations (Al-Jahdali and Bisher, 2008) is converted photochemically into sulfate aerosol, which contributes to PM and has significant adverse effects on human health (Lelieveld et al., 2015). ~~Sulfate aerosol concentration strongly depends upon the strength of the SO_2 sources, and is subject to diurnal variability due to the photochemical reactions, and thus exhibits substantial temporal and spatial heterogeneity. Along with air-pollution, aerosols (natural and anthropogenic) alter the Earth's radiative balance and generally cool the climate playing an essential role in regional and global climate changes (Carlson and Benjamin, 1980; Miller and Tegen, 1998; Bangalath and Stenchikov, 2016; Charlson et al., 1992; Chuang et al., 1997; Myhr . This aerosol direct climate effect is associated with the total content of aerosols in the atmospheric column~~Ukhov et al. (2020) simulated SO_2 transport and distribution over the Middle east using the high-resolution WRF-Chem model and demonstrated high surface concentrations of SO_2 along the west and east coasts of Arabian Peninsula.

The impact of aerosols on air-quality is characterized by near-surface concentrations of PM, which comprise both PM_{10} and $PM_{2.5}$ (particles with diameters less than 10 μm and 2.5 μm, correspondingly). Extended exposure to PM may cause cardiovascular and respiratory disease, lung cancer, and cause premature mortality on a global scale (Lelieveld et al., 2015). According to the WHO, outdoor air pollution caused 4.2 million premature deaths worldwide in 2016 (WHO, 2018). To protect human health and the environment WHO (WHO, 2006), and the National Agencies, e.g., the United States Environmental Protection Agency (US-EPA) (USEPA, 2010), European Commission (EC) (EUEA, 2008), and Kingdom Saudi Arabia Presidency of Meteorology and Environment (KSA-PME) (PME, 2012) issued the air quality regulations for PM that are presented in Table 1. The WHO regulations are the strictest, while KSA-PME regulations are the softest.

Table 1. Air quality regulations for $PM_{2.5}$ and PM_{10} prescribed by WHO, US-EPA, EC, and KSA-PME, $\mu g/m^3$.

	Aver. period	WHO	US-EPA	EC	KSA-PME
$PM_{2.5}$	24 hours	25	35[1]	-	35
	1 year	10	15[2]	25	15
PM_{10}	24 hours	50	150[4]	50[3]	340
	1 year	20	-	40	80

[1] 98th percentile, averaged over 3 years
[2] annual mean, averaged over 3 years
[3] 35 permitted exceedances per year
[4] not to be exceeded more than once per year on average over 3 years

~~Satellite observations provide information on global aerosol abundance and spatial-temporal variability. This enables a better understanding of the sources of aerosols and associated physical processes. Passive satellite aerosol sensors like Moderate Resolution Imaging Spectrometer (MODIS) , Multiangle Imaging SpectroRadiometer (MISR), Advanced Very High-Resolution Radiometer (AVHRR) are designed to measure column integrated aerosol properties. Active instruments such as Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOPE) can, in principle, provide information on the vertical structure of aerosols, although the current generation of spaceborne lidars do not possess the necessary vertical or spectral resolution to properly characterize near surface aerosol properties. The~~ Global satellite observations of aerosol optical depth (AOD) inform about

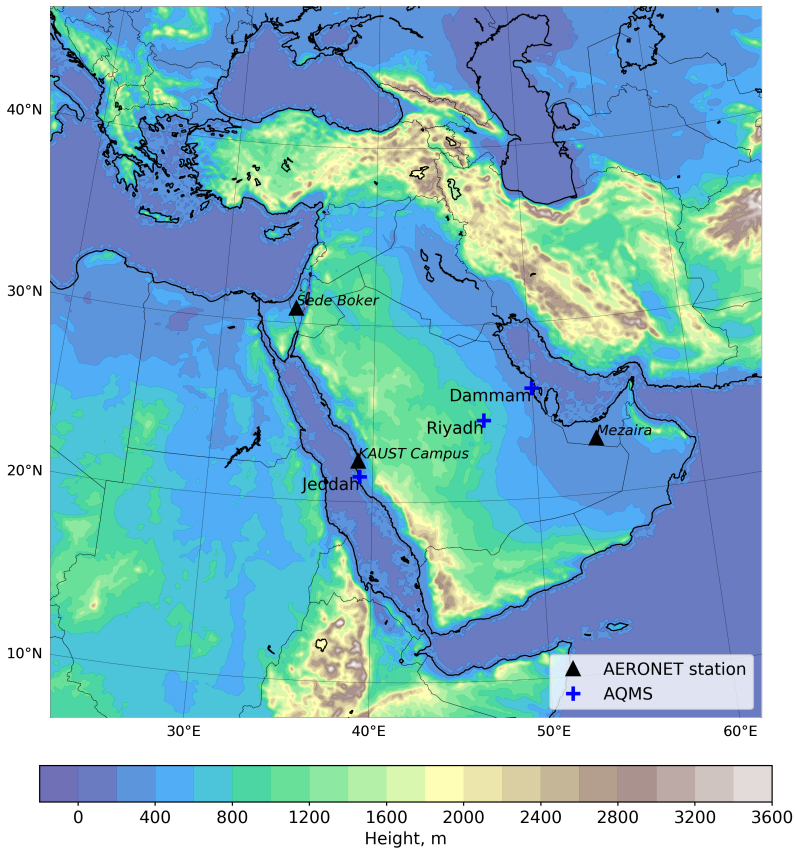


Figure 1. Simulation domain with marked locations of the AQMS and AERONET sites.

vertically-integrated aerosol loading in an entire atmospheric column. But the near-surface PM concentration can not be observed from the space~~but is measured *in situ*~~. These measurements could be conducted only *in situ* in a limited number of locations. ~~For example, in Tawabini et al. (2017) the measurements of PM_{10} concentration conducted for 21 days at Dhahran, Khobar, and Dammam (cities on the west coast of Saudi Arabia) in October-December of 2015 have shown average PM_{10} concentrations of 177, 380, and 126 respectively, which are higher than WHO, US-EPA, and even KSA-PME limits. The $PM_{2.5}$ observations in the United Arab Emirates conducted for the period 1980-2016 showed that annual mean $PM_{2.5}$ concentrations varied from 49 to 77 with an overall average of 61.25 exceeding the US-EPA standards (Al-Taani et al., 2019) , and over the last 14 years $PM_{2.5}$ concentrations showed a positive trend. In Karagulian et al. (2019), the WRF-Chem v3.8.1 was used to simulate a dust storm over the UAE on 2 April 2015. The simulated PM_{10} concentration peaked at 1500 . During another severe dust storm that occurred on 18 - 22 March 2012, the AOD reached 4.5 at *KAUST Campus* AERONET station (Jish Prakash et al., 2015). This dust storm covered a huge area, including Iraq, Iran, Kuwait, Syria, Jordan, Israel, Lebanon, UAE, Qatar, Bahrain, Saudi Arabia, Oman, Yemen, Sudan, Egypt, Afghanistan, and Pakistan. Dust source regions along the western coast of the Arabian Peninsula were also activated. The dust emission rate calculated in WRF-Chem~~

~~exceeded 500 (Jish Prakash et al., 2015). Nayebare et al. (2016); Khodeir et al. (2012); Al-Jeelani (2009); Munir et al. (2013) conducted chemical characterization of airborne PM and the general state of air pollution in Saudi Arabia's cities.~~

~~Simulation domain with marked locations of the AQMS and AERONET sites.~~

~~Along with~~ Along with instrumental observations, modern data assimilation products provide valuable information about AOD and near-surface PM concentration even in areas where satellite sensors are unreliable due to factors such as the high reflectivity of land surfaces (Shi et al., 2011). ~~Here we evaluate two~~ Assimilation products improve the aerosol total column loadings through the assimilation of observed AOD but are not capable of assimilating the aerosol vertical structure and chemical composition. There are two well known data assimilation products that assimilate atmospheric constituents: MERRA-2 (Randles et al., 2017; Buchard et al., 2017) from National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and CAMS-OA ~~(Flemming et al., 2015; Inness et al., 2015)~~ (Inness et al., 2019; Flemming et al., 2015; Inness et al., from European Centre for Medium-range Weather Forecast (ECMWF). ~~The accuracy of these~~ These data assimilation products ~~in terms of~~ adequately reproduce AOD and PM concentrations ~~was evaluated~~ at different regions of the world (Provençal et al., 2017; Buchard et al., 2017; Cesnulyte et al., 2014; Cuevas et al., 2014). E.g., Provençal et al. (2017) tested PM surface concentrations from the MERRA Aerosol Reanalysis (predecessor of MERRA-2) against observations over Europe. Buchard et al. (2017) evaluated MERRA-2 surface $PM_{2.5}$ on the global scale and over the continental United States. Excessive validation of the Monitoring Atmospheric Composition and Climate (MACC) reanalysis (predecessor of CAMS) has been conducted ~~in Cesnulyte et al. (2014), where~~ by Cesnulyte et al. (2014), who compared the model AOD ~~is compared~~ with the AERONET observations. ~~In Cuevas et al. (2014) ,~~ Cuevas et al. (2014) evaluated atmospheric mineral dust from the MACC reanalysis has been evaluated over the MENA region for 2007–2008 using satellite and ground-based observations. MERRA-2 and CAMS-OA are global and have a relatively low spatial resolution (in comparison with the regional models), which diminishes their ability to resolve fine-scale regional spatial features. ~~They improve the aerosol total column loadings through the assimilation of observed AOD but are not capable of assimilating the aerosol vertical structure and chemical composition.~~ Like any other model, MERRA-2 and CAMS-OA use emission inventories of anthropogenic pollutants that may be outdated and incomplete, especially in the rapidly developing parts of the world, like the ME region (McLinden et al., 2016). ~~Here we improve the latest inventories of anthropogenic emissions in WRF-Chem using the novel SO_2 emissions data set (Liu et al., 2018)~~E.g., SO_2 emissions used in MERRA-2 and CAMS-OA differ by 45-50% in some ME regions (Ukhov et al., 2020).

~~Thus in~~ In this study, we ~~test~~ evaluate aerosol outputs from MERRA-2, CAMS-OA, and WRF-Chem over the ME, against satellite, ground-based AOD observations, and *in situ* $PM_{2.5}$ and PM_{10} measurements ~~, and evaluate air-quality~~ during 2015-2016 period, and assess air pollution over the ME focusing on the following science questions:

1. How accurately do WRF-Chem, MERRA-2, and CAMS-OA capture the abundance of dust aerosol, its volume size, and spatial distributions over the ME, in comparison with AERONET and satellite observations?

2. How accurately do WRF-Chem, MERRA-2, and CAMS-OA capture PM surface concentrations compared with *in situ* measurements?

3. What are the contributions of dust, sea salt, sulfate, black carbon, and organic matter in PM surface concentrations?

4. What is the overall impact of PM pollution on air quality over the ME region and in the ~~most significant ME urban centers~~ME major cities?

The paper is organized as follows: Section 2 describes the observational datasets used in this study. Section 3 ~~briefly describes~~ characterizes data assimilation products. In Section 4, the WRF-Chem model setup is described. In Section 5, ~~a comparison of~~ the capabilities of WRF-Chem, MERRA-2, and CAMS-OA to simulate dust aerosol abundance over the ME ~~is presented. PM pollution maps and PM levels in major urban centers of the ME~~ are compared; the PM spatial distributions and PM air pollution in the ME major cities obtained from the WRF-Chem ~~simulation~~ simulations are also discussed. Conclusions are presented in Section 6.

2 Observational datasets

To evaluate the data assimilation products and WRF-Chem output, we use ~~MODIS~~ Moderate Resolution Imaging Spectrometer (MODIS) AOD retrievals, ground-based Aerosol Robotic Network (AERONET) AOD observations, and aerosol volume size distribution retrievals, as well as *in situ* measurements of PM surface concentrations.

2.1 AERONET

AERONET comprises more than 1000 ~~CIMEL and PREDE robotic sunphotometers (made in France and Japan, respectively) which~~ observation sites equipped with CIMEL sunphotometers and PREDE skyradiometers manufactured in France by CIMEL and in Japan by PREDE. They measure direct sun and sky radiances at eight wavelengths (340, 380, 440, 500, 670, 870, 940, and 1020 nm) every 15 minutes during daylight time (Holben et al., 1998). In 2012 we established the *KAUST Campus* site, which is currently the only permanently operational AERONET site in Saudi Arabia. For this study we have chosen three AERONET sites (*KAUST Campus*, *Mezaira*, and *Sede Boker*, see Fig. 1) that routinely collected data in 2015-2016 and are located within our domain. We ~~primarily~~ utilized level 2.0 (cloud screened and quality assured) AERONET AOD data~~but used level 1.5 (cloud screened) data when level 2.0 data were not available~~. To facilitate comparison with the model output the 550 nm AOD is calculated using ~~Angstrom exponent according to~~ the following relation:

$$\frac{\tau_\lambda}{\tau_{\lambda_0}} = \left(\frac{\lambda}{\lambda_0}\right)^{-\alpha} \tag{1}$$

where α is the ~~Angstrom exponent~~ Ångström exponent for the 440-675 nm wavelength range provided by AERONET, τ_λ is the optical thickness at wavelength λ, and τ_{λ_0} is the optical thickness at the reference wavelength λ_0. From here forward, we ~~will~~ presume that AOD is given or calculated at 550 nm.

In addition to direct observations of AOD, the AERONET retrieval algorithm provides column integrated Aerosol Volume Size Distribution (AVSD) $dV/d\ln r$ ($\mu m^3/\mu m^2$) on 22 logarithmically equidistant discrete points in the range of radii between 0.05 and 15 μm (Dubovik and King, 2000). We use these retrievals to evaluate the AVSDs ~~produced~~ calculated by WRF-Chem, CAMS-OA, and MERRA-2.

2.2 MODIS

MODIS instruments ~~onboard the NASA Terra and Aqua~~ on-board the NASA TERRA and AQUA satellites provide aerosol properties over both land and ocean with near-daily global coverage. The ~~high surface albedo over the desert surfaces complicates the AOD retrievals (Shi et al., 2011). The~~ standard MODIS AOD aerosol product combines two retrieval algorithms: 1) the MODIS dark-target (DT) algorithm (Kaufman et al., 1997) is used over the ocean and dark areas with sufficient vegetation, 2) the Deep Blue (DB) algorithm is used over bright desert surfaces of the Sahara and the ME. ~~The uncertainties of AOD obtained with the DB algorithm are ≈25–30% (Hsu et al., 2006).~~ From this combined product (MODIS-DB&DT v6.1) we use AOD at 550 nm level ~~3~~ 2 data from the daily dataset at ~~1°×1°~~ 10 km spatial resolution, downloaded from ~~(Acker and Leptoukh, 2007)~~ https://ladsweb.modaps.eosdis.nasa.gov/about/purpose (Levy et al., 2015).

Recently, a new MODIS AOD product became available that was obtained using the Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm (Lyapustin et al., 2018). This algorithm uses time series analysis and image processing to derive the surface bidirectional reflectance function at fine spatial resolution. MAIAC uses empirically tuned, spatially varying, aerosol properties derived from the AERONET climatology, and provides AOD at 470 and 550 nm with 1 km spatial resolution over land globally. We include the new MAIAC product (version 6, level 2) in the comparison between simulated and retrieved AODs.

2.3 Surface *in situ* PM observations

To test the model-produced PM concentrations, we use observations conducted by the air quality monitoring stations (AQMS) that measure surface concentrations of $PM_{2.5}$ and PM_{10} in Riyadh, Jeddah, and Dammam (megacities of Saudi Arabia), see Fig. 1. Observations are available starting from 2016. The measurements were conducted by the Saudi Authority for Industrial Cities and Technology Zones (MODON). MODON uses MP101M analyzer to continuously detect $PM_{2.5}$ and PM_{10} concentrations by measuring the absorption of low-energy β-radiation that is proportional to the mass of aerosol particles independently of their physicochemical nature (measurement Method ISO 10473). $PM_{2.5}$ and PM_{10} measurement error is $\pm$ 5%. The system satisfies the European Standards EN 12341 and US EPA (40CFR part 53) for PM_{10} and EN 14907 for $PM_{2.5}$ continuous monitoring. The PM measurements are conducted every 15 minutes, and collected data are transmitted in real-time to servers at MODON for processing and storage. To provide confidence in the operational status of the each AQMS, a comprehensive physical audit is conducted quarterly by Ricardo-AEA Ltd, (https://www.ctc-n.org/network/network-members/ricardo-aea-ltd)~~quarterly~~.

3 Data assimilation products

MERRA-2 and CAMS-OA assimilate satellite ~~and ground-based~~ observations to provide aerosol abundance and air-quality data globally. MERRA-2 also assimilates AERONET AODs. In contrast, WRF-Chem is a free-running model and does not assimilate observations. ~~Here, we specifically evaluate these products against observations over the ME, and compare them with the WRF-Chem output.~~

3.1 MERRA-2

MERRA-2 (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2) provides meteorological and atmospheric composition fields on $0.625^\circ \times 0.5^\circ$ latitude-longitude grid and 72 terrain-following hybrid $\sigma - p$ model layers (Randles et al., 2017; Buchard et al., 2017). The pressure at the model top equals 0.01 hPa. MERRA-2 uses the Goddard Earth Observing System, version 5 (GEOS-5) atmospheric model (Rienecker et al., 2008), which is interactively coupled to the Goddard Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) model ~~(Chin et al., 2002)~~ (Chin et al., 2002; **?**) (i.e., it takes into account the effects of aerosols on radiation and model dynamics). This model simulates dust and sea salt in five size bins (see Tab. 2), ~~SO_2~~SO_2, sulfate, organic and black carbon (hydrophobic and hydrophilic), ~~O_3, CO~~O_3, CO, dimethyl sulfide ~~DMS~~(DMS), and methane sulfonic acid (~~MSA~~MSA). The dust density is 2600 $\mathrm{kg/m^3}$ for all sizes. Dust and sea salt emissions are calculated in the model, depending on the near-surface wind. The dust source function is taken from Ginoux et al. (2001). For anthropogenic emissions, MERRA-2 employs the EDGAR-4.2 (Janssens-Maenhout et al., 2013) emission inventory available on a $0.1^\circ \times 0.1^\circ$ grid. MERRA-2 ~~assimilates~~ assimilated AOD at 550 nm from the ~~AVHRR~~ Advanced Very High-Resolution Radiometer (AVHRR) (Heidinger et al., 2014) over the oceans ~~, AOD retrievals from the MISR over bright surfaces (Kahn et al., 2005), as well as~~ until 2002. Since 2000 MERRA-2 started assimilating MODIS and Multiangle Imaging SpectroRadiometer (MISR) (Kahn et al., 2005) data over land and ocean. Both instruments are on the TERRA satellite which has an equatorial overpass at 10:30 am UTC, while AVHRR has mostly orbited with the afternoon equatorial crossing time. Therefore MERRA-2 continued using AVHRR data over the ocean until 2002 when the AQUA satellite was launched. Since AQUA has an orbit with the equator overpass at 2:30 pm UTC, AVHRR data was no longer needed for coverage. We have to mention that MERRA-2 assimilates specially processed MODIS observations~~(but~~ , not the standard MODIS-DB&DT aerosol product~~) and AERONET to constrain the atmospheric aerosols~~. It also assimilates AERONET AODs (Randles et al., 2017).

3.2 CAMS-OA

CAMS-OA (https://atmosphere.copernicus.eu/) has been conducted in almost real-time since July 2012. The CAMS-OA product has a resolution of $0.8^\circ \times 0.8^\circ$ before 21 June 2016, and $0.4^\circ \times 0.4^\circ$ after that, with 60 vertical levels. It employs the ECMWF aerosol data assimilation system developed within the Integrated Forecast System (IFS) (Morcrette et al., 2009; Benedetti et al., 2009). The extended version of the Carbon Bond chemical mechanism 5 (CB05) (Yarwood et al., 2005) is implemented in the IFS (Flemming et al., 2015). CB05 describes tropospheric chemistry with 54 species and 126 reactions. The chemistry scheme is coupled with the aerosol module.

CAMS-OA simulates five aerosol species: dust, sea salt, sulfate, organic carbon, and black carbon. To ~~simulate~~ calculate dust and sea salt, it uses three ~~dustbins~~ dust-bins (see Tab. 2). The dust density is 2600 $\mathrm{kg/m^3}$ for all bins. Emissions of mineral dust and sea salt depend on simulated near-surface wind speed. Dust emission is parameterized following Marticorena and Bergametti (1995) with the source function adopted from Ginoux et al. (2001). ~~SO_2~~ SO_2 oxidation into sulfate aerosol is parameterized using a prescribed latitude-dependent e-folding timescale ranging from 3 days at the equator to 8 days at the poles. The anthropogenic emissions for the chemical species are taken from the MACCity inventory (Granier et al., 2011),

Table 2. Radii ranges (μm) of dust and sea salt bins used in GOCART model (WRF-Chem, MERRA-2) and in CAMS-OA.

	Bin				
	1	2	3	4	5
CAMS-OA dust	0.03-0.55	0.55-0.9	0.9-20.0	-	-
CAMS-OA sea salt	0.03-0.5	0.5-5.0	5.0-20.0	-	-
GOCART dust	0.1-1.0	1.0-1.8	1.8-3.0	3.0-6.0	6.0-10.0
GOCART sea salt	0.03-0.1	0.1-0.5	0.5-1.5	1.5-5.0	5.0-10.0

which is available on a $0.5^\circ \times 0.5^\circ$ grid and covers the period 1960–2010. CAMS-OA assimilates MODIS ~~observations.~~ AQUA and TERRA AODs. It uses observations from Collection 5 since 2009, and Deep Blue since 2015.

4 WRF-Chem

To calculate fine-resolution PM and sulfate fields, we use the Weather Research and Forecasting (WRF) model (Skamarock et al., 2005) coupled with chemistry (WRF-Chem v3.7.1) (Grell et al., 2005). The WRF-Chem is used for prediction and simulation of weather, air quality, and dust storms, accounting for the aerosol effect on radiation. WRF-Chem can be configured with one of the few gas-phase chemical mechanisms, photolysis, and aerosols parameterization models. WRF-Chem has been widely used for air quality simulations in different parts of the globe: East Asia (Wang et al., 2010), ~~US~~ North America (Kim et al., 2006; Chuang et al., 2011), Europe (Forkel et al., 2012; Ritter et al., 2013), South America (Archer-Nicholls et al., 2015) and Middle East (Parajuli et al., 2019).

To reduce the clock-time of our two-year calculations, we simulated each month of the 2015-2016 period separately. Each simulation starts from the last week of the previous month. This time is considered a spin-up and is excluded from post-processing. The simulation domain, shown in Fig. 1, is centered at 28°N, 42°E, and a 10×10 km^2 horizontal grid (450×450 grid nodes) is employed. The vertical grid comprises 50 vertical levels with enhanced resolution closer to the ground comprising 11 model levels within the near-surface 1-km layer. The model top boundary is set at 50 hPa.

To improve the representation of the meteorological fields, we apply spectral nudging (Miguez-Macho et al., 2004) above the planetary boundary layer (PBL) (>5.0 km) to horizontal wind components (U and V) toward the MERRA-2 wind field. The nudging coefficient for U and V is set to be 0.0001 s^{-1}. We ~~nudge modes with wavelengths larger~~ only nudge waves with wavelengths longer than 450 km. This allows us to keep the large-scale motions close to reanalysis, and leave the resolved small-scale, high-frequency features unaffected.

The aerosol/chemistry initial ~~conditions~~ and boundary conditions (IC&BC) are calculated using MERRA-2 output ~~by means of~~ using the newly developed *Merra2BC* interpolation utility ~~(see Appendix ??)~~(Ukhov and Stenchikov, 2020). To be consistent with aerosol/chemistry IC&BC, we also define the meteorological IC&BC using MERRA-2 output (see Appendix ??A1).

The following set of physical parameterizations was used in WRF-Chem runs. The Unified Noah land surface model (*sf_surface_physics*=2) and the Revised MM5 Monin-Obukhov scheme (*sf_sfclay_physics*=1) are chosen to represent land sur-

face processes and surface layer physics. The Yonsei University scheme is chosen for PBL parameterization (*bl_pbl_physics*=1). The WRF single moment microphysics scheme (*mp_physics*=4) is used for the treatment of cloud microphysics. The New Grell scheme (*cu_physics*=5) is used for cumulus parameterization. The Rapid Radiative Transfer Model (RRTMG) for both short-wave (*ra_sw_physics*=4) and long-wave (*ra_lw_physics*=4) radiation is used for radiative transfer calculations. Only the aerosol direct radiative effect is accounted for. More details on the physical parameterizations used can be found at http://www2.mmm.ucar.edu/wrf/users/phys_references.html.

4.1 Gas-phase chemistry and aerosols

To calculate the atmospheric chemistry within WRF-Chem, we employ the Regional Atmospheric Chemistry Mechanism (RACM, *chem_opt*=301) (Stockwell et al., 1997) containing 77 species and 237 reactions, which include 23 photolysis reactions~~. It~~ , but no heterogeneous chemistry. The RACM chemical module is embedded into WRF-Chem using the Kinetic PreProcessor (KPP) (Damian et al., 2002). The role of KPP is to integrate the system of stiff nonlinear ordinary differential equations, which represents the specified set of chemical reactions. The photolysis rates are calculated on-line according to Madronich (1987) (*phot_opt*=1). Similar to MERRA-2, the GOCART chemistry module is used to calculate ~~SO_2~~ SO_2 to sulfate oxidation ~~(Chin et al., 2002) by the hydroxide~~ (Chin et al., 2002; **?**) by the hydroxyl radical OH whose abundance is interactively simulated by RACM.

~~Here we~~ We use the novel OMI-HTAP ~~SO_2~~ SO_2 emission dataset (Liu et al., 2018) based on the combination of distributed ~~SO_2~~ SO_2 emissions from residential and transportation sectors, taken from the HTAP-2.2 inventory (Janssens-Maenhout et al., 2015) with the catalogue of the strong (>30 $\mathrm{kt/year}$) ~~SO_2~~ SO_2 point emissions (Fioletov et al., 2016) built using satellite observations by Ozone Monitoring Instrument (OMI) (Levelt et al., 2006; Li et al., 2013). The catalogue contains more than 500 point sources of industrial origin, some of which are not present in the widely used EDGAR-4.2 and HTAP-2.2 emission datasets. For example, 14 previously unaccounted ~~SO_2~~ SO_2 point emissions located in the ME (mostly in the Arabian Gulf) were detected, most of them are related to oil and gas industry. OMI-HTAP divides ~~SO_2~~ SO_2 emissions into surface and elevated ones. We distribute the surface ~~SO_2~~ SO_2 emissions with a constant mixing ratio in the 0-1000 m layer, and elevated emissions in 120-1000 m layer. All other constituents (other PM from biogenic and fossil components, black and organic carbon, etc.), including ~~SO_2~~ SO_2 shipping emissions, are taken from the HTAP-2.2 inventory and are treated as surface emissions. OMI-HTAP SO_2 emissions are provided on $0.1^\circ \times 0.1^\circ$ grid (Liu et al., 2018). We conservatively interpolated them on the WRF-Chem 10×10 km^2 grid. See Ukhov et al. (2020) for details.

To calculate aerosols we employ the GOCART (Chin et al., 2002) aerosol model (*chem_opt=301*). It is the same microphysical model as that used in MERRA-2 (see Sec. 3.1). Dust and sea salt size distributions in WRF-Chem are approximated by the same five dust and sea-salt size bins as those in MERRA-2 (Tab. 2). However, only the last four "salt" bins in Tab. 2 are used in WRF-Chem, as the first bin appears to be very poorly populated. Dust density is assumed to be 2500 $\mathrm{kg/m}^3$ for the first ~~dustbin~~ dust-bin and 2650 $\mathrm{kg/m}^3$ for 2-5 ~~dustbins~~dust-bins. Emission of sea salt is calculated according to Gong (2003).

Dust emission from the surface is calculated using the GOCART emission scheme (Ginoux et al., 2001) (*dust_opt=1*). Dust emission mass flux, F_p (μg m^{-2} s^{-1}) in each ~~dustbin~~ dust-bin p=1,2,...,5 is defined by the relation:

$$F_p = \begin{cases} CSs_p u_{10m}^2 (u_{10m} - u_t), & \text{if } u_{10m} > u_t \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where, C has the dimension of (μg s^2 m^{-5}) and is a spatially uniform factor which controls the magnitude of dust emission flux; S is the spatially ~~nonuniform~~ varying topographic source function (Ginoux et al., 2001) that characterizes the spatial distribution of dust emissions; u_{10m} is the horizontal wind speed at 10 m height; u_t is the threshold velocity, which depends on particle size and surface wetness; s_p is a fraction of dust mass emitted into ~~dustbin~~ dust-bin p, $\sum s_p = 1$.

To avoid natural dust emission in urban areas, we use the built-in WRF-Chem the U.S. Geological Survey (USGS) 24-category land-use data set (Anderson, 1976). We modify the source function S using the following expression:

$$S' = (1.0 - URBAN\_MASK) \cdot S \quad (3)$$

where S' is the modified topographic source function, $URBAN\_MASK$ is the USGS "*Urban and Built-up Land*" mask field. It has the sense of a fraction of urban area in a grid-cell and ranges from 0 to 1. Grid cells with $URBAN\_MASK$=1 do not produce natural dust emissions. We do not account for anthropogenic dust emissions within cities, and therefore potentially underestimate urban dust pollution.

As in our previous studies (Kalenderski et al., 2013; Jish Prakash et al., 2015; Anisimov et al., 2017), we tune dust emissions to fit the ~~daily average~~ AOD from the AERONET stations located within the domain. For this purpose, the ~~parameter~~ factor C from Eq. (2) has been adjusted to achieve the best agreement between simulated and observed AOD at *KAUST Campus, Mezaira*, and *Sede Boker* AERONET sites, see Fig. 1. Assuming that factor C does not depend on time and geographical coordinates, we can only tune the annual average AOD bias. Both simulations and observations represent the total AOD with contributions from all types of aerosols. Because dust dominates all other aerosols in the ME, we choose to tune only the dust emissions.

Obtained during test runs, C value of 0.5 is kept constant in all subsequent production runs. We also tune s_p from Eq. (2) to better reproduce the AVSDs provided by AERONET inversion algorithm. This tuning and the comparisons of AOD and AVSDs from the assimilation products and WRF-Chem simulations are discussed in detail below.

In situ air quality observations in the Middle east are scarce. It is one of the known problems for air quality research in this area. The things are simplified a bit by the fact that in the ME dust dominates aerosol pollution. E.g., Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) (Vaughan et al., 2004) records dust in 95% of profiles (Osipov et al., 2015). The effect of nitrates, ammonia, and organics on AOD and PMs is insignificant; therefore, the employed aerosol-chemical scheme (GOCART-RACM) is adequate for the ME conditions. To support this conclusion, we have conducted laboratory analysis of the chemical composition of soil and dust deposition samples that show a little presence of organics and ammonium

(Jish Prakash et al., 2016; Engelbrecht et al., 2017). According to Engelbrecht et al. (2017), in 2015, the annual average weight percentages of soluble ions of ammonium (NH_4) and sulfate in deposition samples taken at four sites at the *KAUST Campus* are 0.05% and 2.513%, respectively. It means that available ammonium may neutralize at a maximum of 5% of sulfate mass. The actual contribution of ammonium sulfate should be lower, as some ammonium may also be bound as ammonium nitrate, ammonium phosphate, or ammonium chloride.

4.2 WRF-Chem code modification

We have corrected the source code of the WRF-Chem v3.7.1 with GOCART aerosol module in several places. These corrections were implemented in the WRF-Chem v4.1.3 official release~~and will be described~~. We evaluate how they change the results in the forthcoming technical publication. Here we only briefly discuss the introduced changes and their effects.

Firstly, the diagnostic output of PM concentrations was corrected, because contributions of the individual dust and sea salt bins were incorrectly calculated. Therefore, $PM_{2.5}$ surface concentrations were erroneously underestimated by 7% while PM_{10} - were overestimated ~~.~~

by 5%. Secondly, we found that the contribution of fine dust particles with radii $<0.46\ \mu m$ was omitted in the calculation of AOD, AOD was consequently underestimated by 25-30%. This led to an overestimation of the dust emission flux because we force the simulated AOD to match the AERONET observations.

Thirdly, we fixed the dust and sea salt gravitational settling subroutine, since initially, the calculations of mass fluxes of settling particles did not account for changes in air density. Due to this error, the total mass of dust and sea salt aerosols increased, violating mass conservation.

5 Results

5.1 Regional climate and circulation

The ME is one of the hottest and driest regions on the Earth. Summer in the ME is long and hot with little precipitation. Precipitation mainly occurs in the south-west of the Arabian Peninsula. Winter is mild, with rainfall being mostly associated with cold fronts and cyclones propagating from the Eastern Mediterranean (Climate.com, 2018). Emission and transport of dust are driven by winds. Emission and deposition of dust are also sensitive to soil moisture and precipitation (Furman, 2003; Shao, 2008; Yu et al., 2015). However, because the ME is an arid region, the soil moisture and precipitation effects are insignificant.

Figure 2 shows contours of sea level pressure, topographic source function S (Ginoux et al., 2001), and seasonally 2015-2016 averaged wind speed barbs at 10 m height over the ME during winter (DJF) and summer (JJA) from WRF-Chem simulations. Over northeast Africa in winter (see Fig. 2a), the strong pressure gradient between the Red Sea trough and the stationary high-pressure system over Egypt predominantly generates moderate north-easterly winds (up to 10 m/s). Therefore in winter, dust storms occur more frequently in the west of the Arabian Peninsula. Over the Central and Eastern Arabian Peninsula and

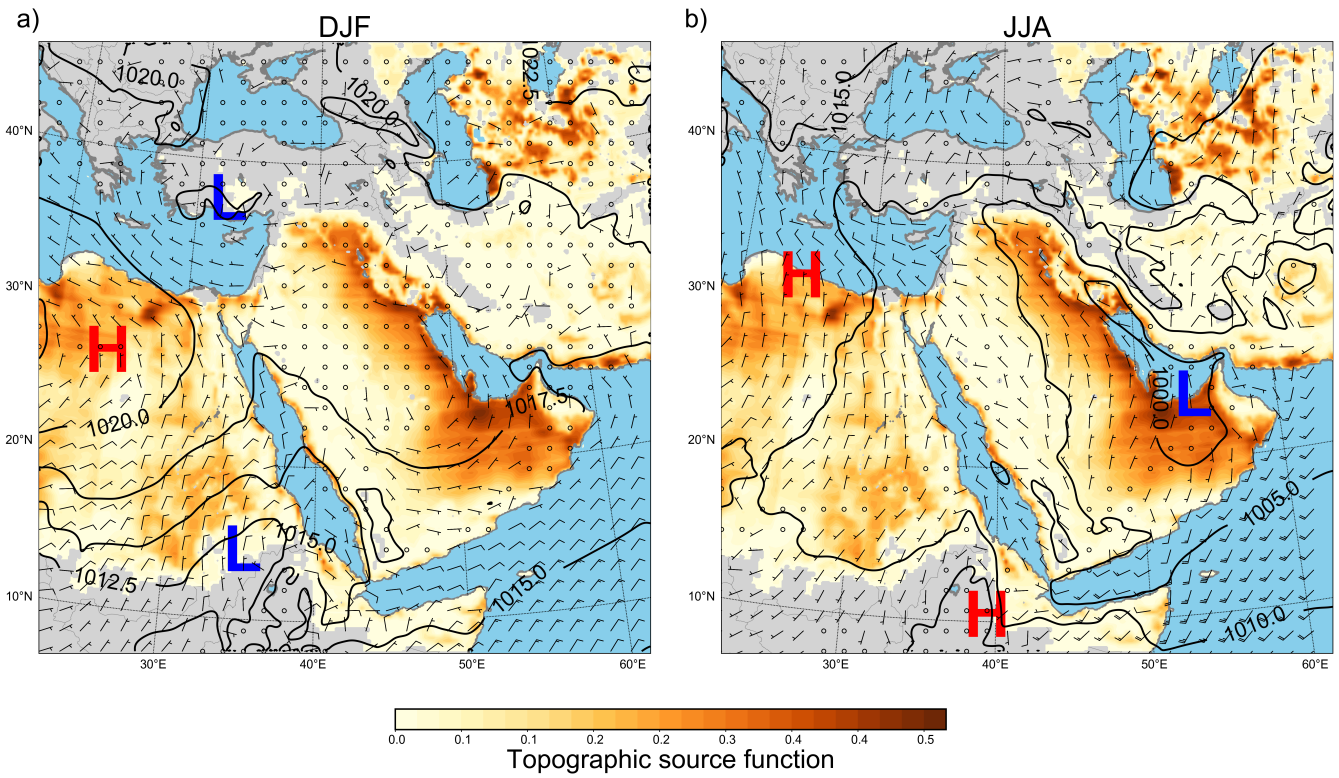


Figure 2. Seasonally averaged for 2015-2016 wind barbs (m/s) at 10 m, sea level pressure (contours), and erodibility function (shading) (Ginoux et al., 2001). a) Winter (DJF), b) Summer (JJA).

the eastern part of the ME, winds are relatively weak and do not have a clear direction. However, cold fronts generated by Mediterranean cyclones can cause dust storms and dust transport to central regions of the Arabian Peninsula.

In summer (see Fig. 2b) the high-pressure system over the eastern Mediterranean and low-pressure system over the Arabian Gulf promote moderate north-northwesterly winds known as *Shamal* (Yu et al., 2016; Hamidi et al., 2013), which dominate over the central part of the Arabian peninsula. *Shamal* is the primary driver of dust storm events over this area (Yu et al., 2016; Shao, 2001; Middleton, 1986; Goudie and Middleton, 2006; Notaro et al., 2015). *Shamal* brings dust to the Arabian Gulf, north, and central part of Saudi Arabia, from the Tigris-Euphrates basin of Syria and Iraq (Anisimov et al., 2018).

Figure 3 shows wind speed seasonally averaged for 2015-2016 at 10 m from MERRA-2, CAMS-OA, and WRF-Chem during winter (DJF) and summer (JJA). WRF-Chem spatial distributions of wind speed agree well with MERRA-2 and CAMS-OA, but due to the higher spatial resolution, WRF-Chem better resolves the fine-scale spatial structures of the 10 m ~~wind~~ -wind field over complex terrain. All panels have similar seasonal variations of wind speed. In winter, maximum winds are stronger over the south-east of the domain. In the Central and northern parts of the domain winds are weak. In summer, wind speed increases in the northern and central parts of the ME. *Somali Jet* produces strong (10-15 m/s) winds in the Arabian Sea along the coasts of Somalia and Oman.

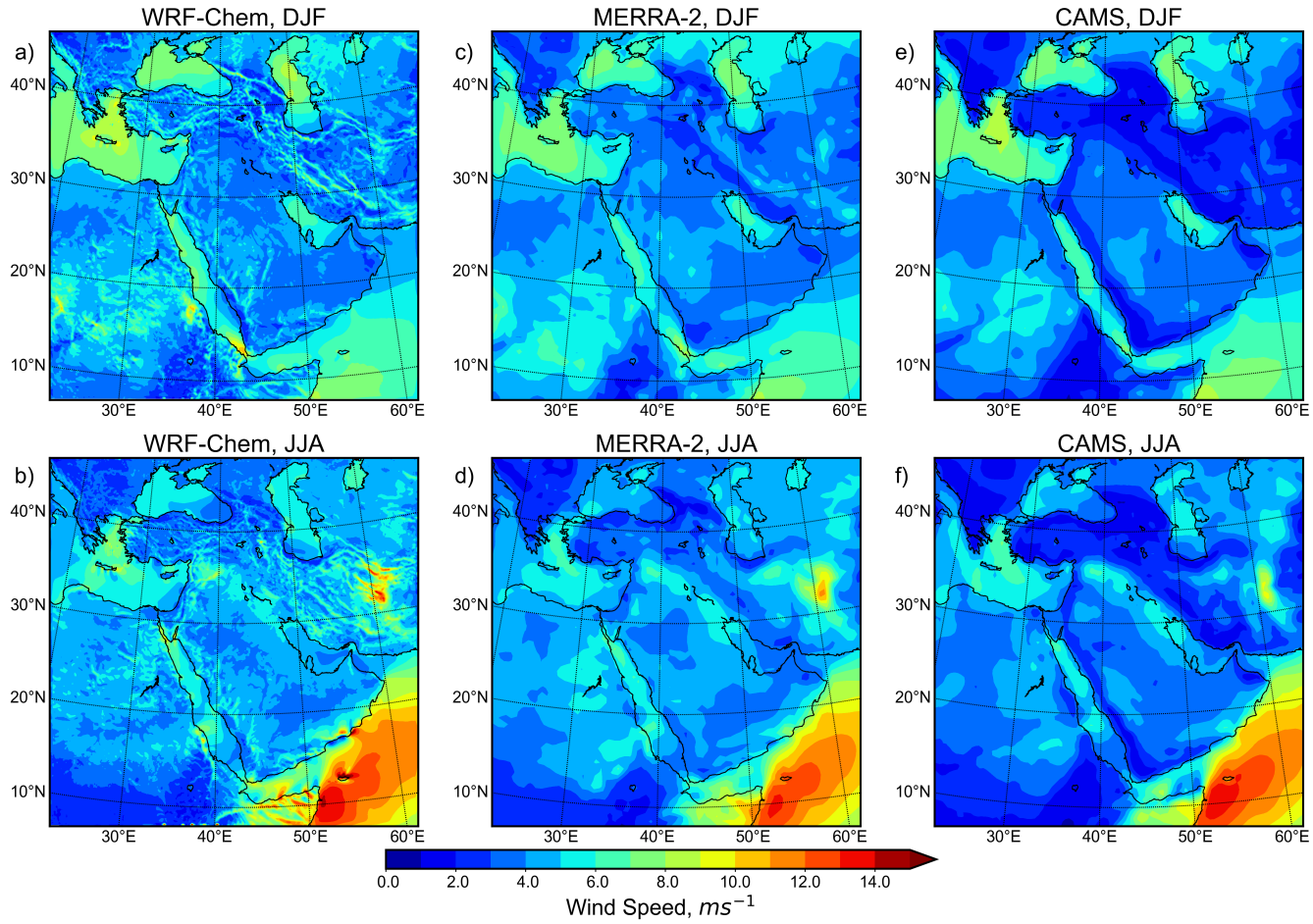


Figure 3. Seasonally averaged 2015-2016 wind speed at 10 m ~~for~~ from WRF-Chem, MERRA-2, and CAMS-OA during winter (DJF), summer (JJA).

To conduct the statistical analysis, we interpolated the seasonally averaged 2015-2016 zonal and meridional wind components (U and V) at 10 m from WRF-Chem, and CAMS-OA on MERRA-2 grid and calculated Pearson correlation coefficient (R), and root mean square differences ($RMSD$) between each pair, see Tab. 3, correspondingly. $RMSD$ is calculated using the same formula as the root mean square error ($RMSE$). The procedure of calculation of these parameters is given in Appendix A2. Pearson correlation coefficients provided in Tab. 3 are close to 1. The highest correlation is achieved between MERRA-2 and CAMS-OA. WRF-Chem's correlation coefficient with respect to MERRA-2 is smaller but exceeds that of the WRF-Chem - CAMS-OA pair. The WRF-Chem and MERRA-2 wind fields are close partly because WRF-Chem boundary conditions are built using MERRA-2 reanalysis, and the large-scale winds are nudged (see Sec. 4) to the ones from MERRA-2 over the PBL.

The $RMSDs$ (see Tab. 3) are lower in winter than in summer. All $RMSDs$ are in the range of 0.45-0.85 $\mathrm{m/s}$. The lowest $RMSDs$ are between MERRA-2 and CAMS-OA. Notably, ~~in all three cases,~~ the correlation coefficients for the meridional component V are higher, and the $RMSDs$ are lower when compared with the zonal wind component U. This is because ~~,~~ ~~in summer,~~ the northern winds are stable since they are maintained by the large-scale processes. In contrast, the zonal wind component, which is affected by small-scale processes like sea-breezes, is variable. The results of the statistical analysis in Tab. 3 and the clear similarity of the spatial patterns (among all products) of the averaged 10 m wind fields presented in Fig. 3, suggest that WRF-Chem captures the magnitude and spatial distribution of the 10 m wind. Thus, we conclude that WRF-Chem with the selected set of physical parameterizations satisfactorily simulates both the large- and meso-scale atmospheric processes in the ME.

Table 3. Pearson correlation coefficient R and root mean square difference $RMSD$ $(\mathrm{m/s})$ for the seasonally averaged 2015-2016 wind components U and V at 10 m.

Season	WRF-Chem wrt CAMS-OA		WRF-Chem wrt MERRA-2		CAMS-OA wrt MERRA-2	
	R	$RMSD$	R	$RMSD$	R	$RMSD$
	$U \mid V$	$U \mid V$	$U \mid V$	$U \mid V$	$U \mid V$	$U \mid V$
Winter (DJF)	0.918 | 0.954	0.716 | 0.593	0.954 | 0.963	0.572 | 0.537	0.954 | 0.974	0.558 | 0.449
Summer (JJA)	0.929 | 0.981	0.853 | 0.704	0.938 | 0.982	0.833 | 0.669	0.965 | 0.986	0.636 | 0.593
Annual mean	0.924 | 0.968	0.785 | 0.649	0.946 | 0.973	0.703 | 0.603	0.960 | 0.980	0.597 | 0.521

* wrt - with respect to

5.2 AOD

In this section, we evaluate the ability of WRF-Chem, CAMS-OA, and MERRA-2 to reproduce the aerosol content in the atmosphere accurately. This content is characterized by AOD. In the ME, mineral dust contribution to the total AOD is dominant ($\approx$87%) (Kalenderski and Stenchikov, 2016; Osipov et al., 2015). The treatment of optically active dust within the model is therefore vitally important. AOD is calculated based on aerosol concentrations and aerosol optical properties, which depend upon aerosol size distribution. We, therefore, evaluate how well WRF-Chem and assimilation products reproduce aerosol volume size distribution.

5.2.1 Aerosol volume size distributions

Dust particles are emitted into the lower atmospheric layer with some predominant size distribution ~~(Kok, 2011)~~(Martin and Kok, 2017; Kol . Emitted dust is processed by the atmosphere to produce the atmospheric dust size distribution that is retrieved by the AERONET inversion algorithm (Dubovik and King, 2000) and reported as column integrated AVSD. Strictly speaking, AERONET AVSD incorporates contributions from all types of aerosols. But the size distribution of emitted dust has the strongest effect on

column integrated AVSD, because dust dominates all other aerosols in the ME. Therefore, we have to tune the dust emission parameters in the first place.

Eq. (2) assumes that ~~dust particles of different sizes have different~~ emission mass fluxes ~~that~~ into five dust size bins are controlled by the s_p fractions. In WRF-Chem the default values of s_p fractions for the five ~~dustbins~~ dust-bins (see Tab. 2) are $\{0.1, 0.25, 0.25, 0.25, 0.25\}$. We found that with these default s_p fractions, WRF-Chem underestimated the volume of fine dust particles in ~~comparison~~ first bin 0.1 μm< r <1 μm compared with AERONET AVSD, whereas the ~~coarse mode~~ volume of the second bin 1 μm< r <1.8 μm was overestimated. In combination with ~~the tuning of~~ fitting the observed AOD by tuning of factor C~~parameter~~, this led to an ~~overestimation~~ the increase of the total emitted dust mass, since fine particles are optically more efficient per unit mass than coarse particles. To achieve a better agreement between the simulated and AERONET AVSDs we adjusted fractions s_p to be {~~0.2,~~ 0.15, ~~0.17, 0.38,~~ 0.1, 0.25, 0.4, 0.1}. A similar approach was implemented in Khan et al. (2015) using the MADE/SORGAM chemistry/aerosol scheme. ~~Updated~~ This s_p ~~values were kept in further~~ modification is in line with (Adebiyi and Kok, 2020) as it effectively decreased emission of dust particles with radii r<2.5 μm and increased emission of coarse particles with radii r>2.5 μm (see Appendix A3). We use the updated s_p values in all our WRF-Chem simulations.

Figure 4 shows seasonally averaged 2015-2016 volume size distributions obtained from MERRA-2, CAMS-OA, AERONET and WRF-Chem with updated s_p fractions. The comparison is conducted for the *KAUST Campus, Mezaira, Sede Boker* AERONET sites (see Fig. 1), since only these sites have information on AVSDs during the 2015-2016 period. ~~Because dust bins are coarse, especially in the sub-micron range, model and assimilation products struggle to correctly reproduce the fine mode of the AERONET AVSD (see~~ The effect of s_p modification could be seen in Fig. 4 ~~). The volume size distributions from the model and assimilation products demonstrate pronounced seasonal variability with the increased amount of dust in the atmosphere during spring and summer. Since the *KAUST Campus* and *Mezaira* sites are located in the vicinity of the strong dust sources, the coarse mode at these sites is more pronounced than at the *Sede Boker* site, which is farther from the strong dust emission sources~~by comparing AVSDs from WRF-Chem with updated set of s_p and MERRA-2 that uses the default s_p set. A direct comparison of AVSDs from the WRF-Chem runs with the updated and default s_p sets is shown in Appendix A3.

~~Summer (JJA) averaged 2015-2016 AVSD () at *KAUST Campus* AERONET site obtained from the AERONET inversion algorithm and from WRF-Chem.~~

Both MERRA-2 and WRF-Chem use the GOCART aerosol scheme with the same five ~~dustbins~~dust-bins, and they approximate the shape of the AERONET AVSD relatively well. CAMS-OA uses only three ~~dustbins~~ dust-bins (see Tab. 2) and fails to reproduce the AERONET AVSD even qualitatively. It overestimates the volume of particles with radii of 0.55-0.9 μm and underestimates the volume of particles with radii of 0.9-20 μm. With the latest system upgrade in 2019, this weakness of CAMS-OA has been corrected by introducing of a new dust scheme (Nabat et al., 2012).

The volume size distributions from the model and assimilation products demonstrate pronounced seasonal variability with the increased amount of dust in the atmosphere during spring and summer. Since the *KAUST Campus* and *Mezaira* sites are located in the vicinity of the strong dust sources, the coarse mode at these sites is more pronounced than at the *Sede Boker* site, which is farther from the strong dust emission sources.

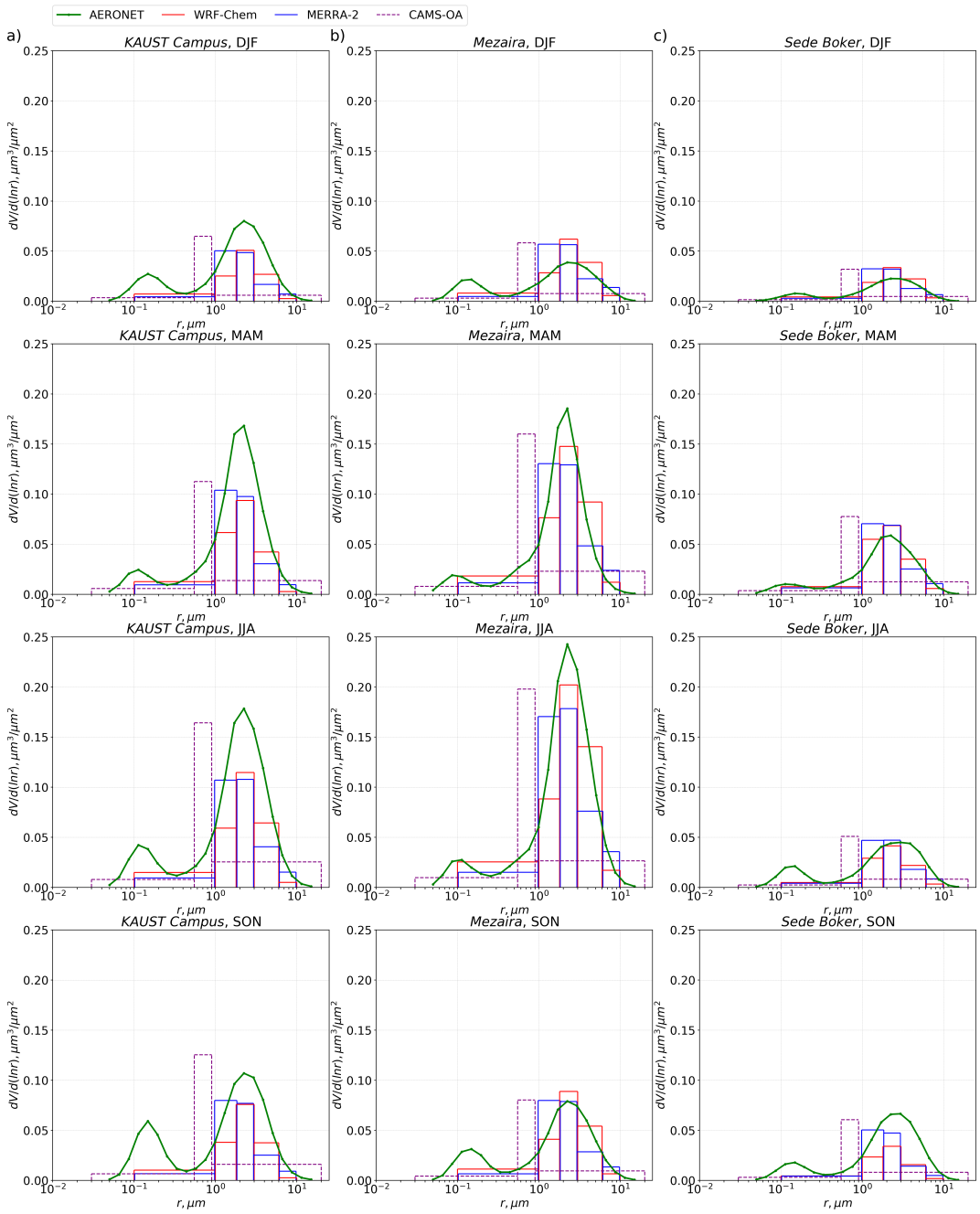


Figure 4. Seasonally averaged 2015-2016 AVSDs ($\mu m^3/\mu m^2$) obtained from MERRA-2, CAMS-OA, WRF-Chem, and from the AERONET inversion algorithm at a) *KAUST Campus*, b) *Mezaira* and c) *Sede Boker* AERONET sites. Winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

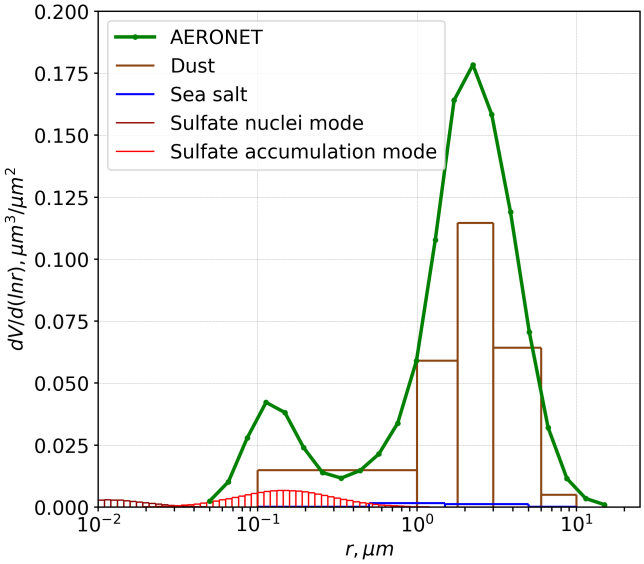


Figure 5. Summer (JJA) averaged 2015-2016 AVSD ($\mu m^3/\mu m^2$) at *KAUST Campus* AERONET site obtained from the AERONET inversion algorithm and from WRF-Chem.

The fine mode in the AERONET AVSD is more pronounced at the *KAUST Campus* in comparison with the other AERONET sites due to its proximity to strong ~~SO_2~~ SO_2 sources located along the west coast of Saudi Arabia (Ukhov et al., 2020). This proximity leads to a higher contribution of fine sulfate particles to the fine mode. The smaller volume of fine particles in the WRF-Chem and MERRA-2 simulated ~~dust~~ AVSD (see Fig. 4) is in part because the ~~sulfate contribution is not shown~~simulated AVSDs show only dust omitting the contributions of sulfate and sea salt. Sea salt particles/droplets are relatively large and mostly contribute to the coarse mode.

Figure 5 shows the contributions of dust, sea salt, and sulfate aerosols into the AVSD at the *KAUST Campus* AERONET site in WRF-Chem simulation averaged for two summer seasons (JJA) of 2015-2016. In WRF-Chem, sulfate aerosol is computed using a bulk approach. ~~However, for~~ For calculating of aerosol optical properties, it is assumed that sulfate aerosol ~~is described by~~ comprises two log-normal modes: nuclei and accumulation. According to WRF-Chem source code, the nuclei mode median radii μ_{nuc}=0.005 μm and geometric width σ_{nuc}=1.7, the accumulation mode median radii μ_{acc}=0.035 μm and geometric width σ_{acc}=2.0. The nuclei mode comprises 25% of the sulfate aerosol mass, and accumulation mode - 75%. It is assumed that sulfate aerosol density is 1800 kg/m^3 and sea salt density is 2200 kg/m^3. Figure 5 demonstrates that the contribution of the sulfate nuclei mode in the aerosol volume is almost negligible, while the sulfate accumulation mode adds in the volume of aerosol particles with radii <1 μm. The contribution of the sea salt aerosol into AVSD in WRF-Chem simulations is very little.

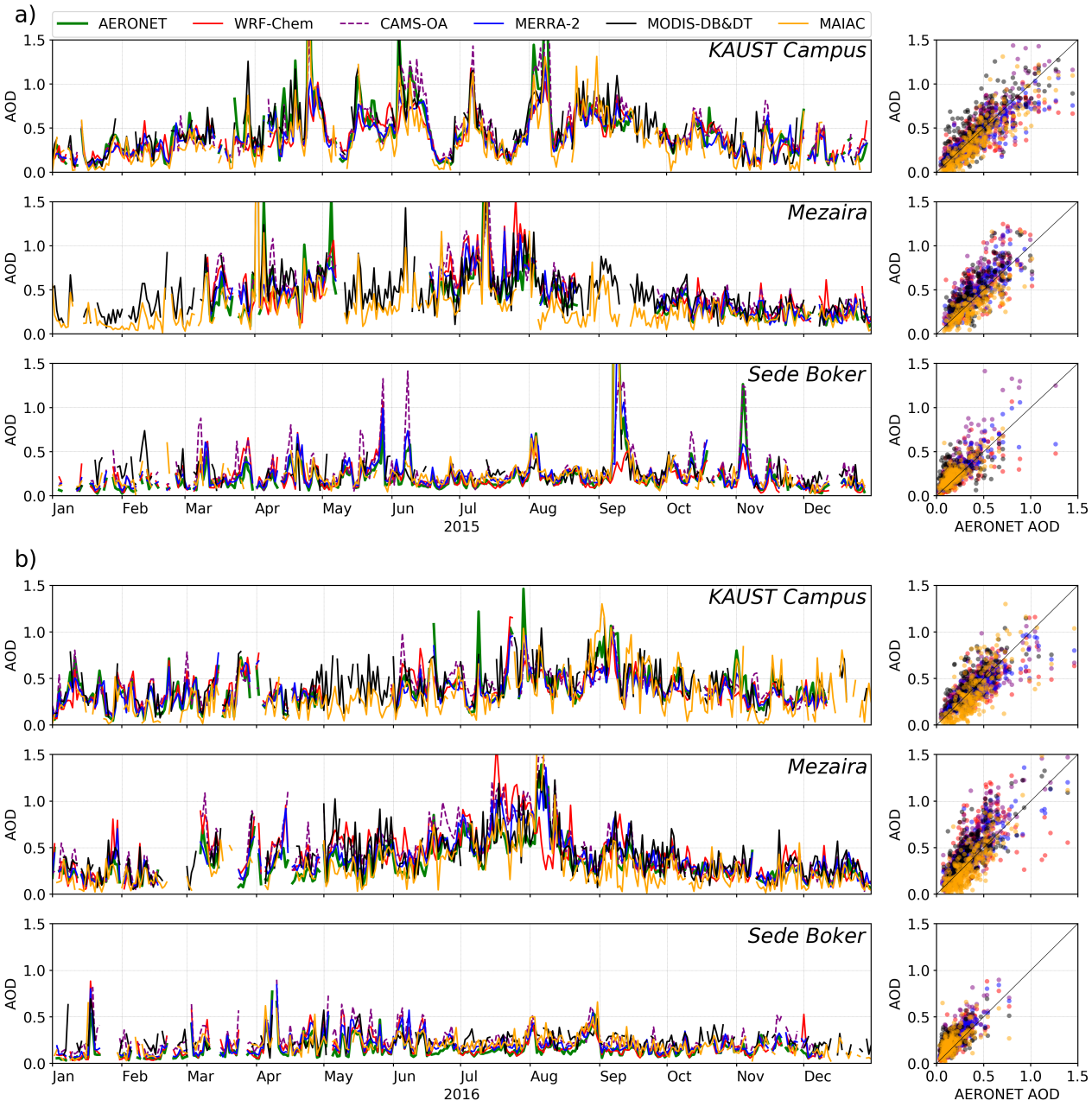


Figure 6. Daily averaged AOD ~~2015-2016~~ at three AERONET sites (*KAUST Campus, Mezaira, Sede Boker*) and corresponding scatter plots computed for WRF-Chem, AERONET, MERRA-2, CAMS-OA, MODIS-DB&DT, and MAIAC~~.~~: a) 2015, b) 2016.

Table 4. Pearson correlation coefficient *R* and mean ~~*BIAS*~~ bias calculated for daily averaged AOD time-series from WRF-Chem, CAMS-OA, MERRA-2, MODIS-DB&DT, and MAIAC with respect to AERONET AOD observations.

	WRF-Chem		CAMS-OA		MERRA-2		MODIS-DB&DT		MAIAC	
	~~*BIAS*~~*bias*	*R*	~~*BIAS*~~*bias*	*R*	~~*BIAS*~~*bias*	*R*	~~*BIAS*~~*bias*	*R*	~~*BIAS*~~*bias*	*R*
	2015									
KAUST Campus	~~-0.05~~ -0.04	~~0.73~~ 0.74	0.01	0.86	-0.05	0.85	~~0.02~~ 0.06	~~0.83~~ 0.81	~~-0.07~~ -0.08	~~0.88~~ 0
Mezaira	~~0.05~~ 0.07	0.73	0.11	~~0.80~~ 0.81	~~0.03~~ 0.04	0.83	~~0.10~~ 0.07	~~0.78~~ 0.79	-0.07	0.8
Sede Boker	~~-0.03~~ -0.01	~~0.44~~ 0.43	0.07	0.65	0.02	0.72	~~0.02~~ 0.06	~~0.92~~ 0.84	0.04	0.9
	2016									
KAUST Campus	~~0.00~~ -0.01	~~0.60~~ 0.75	0.01	0.76	-0.03	0.88	~~-0.05~~ 0.06	~~0.77~~ 0.73	~~-0.07~~ -0.05	~~0.64~~ 0
Mezaira	0.09	~~0.52~~ 0.62	~~0.11~~ 0.12	~~0.88~~ 0.87	0.06	~~0.87~~ 0.85	0.08	~~0.70~~ 0.77	-0.04	0.8
Sede Boker	~~0.02~~ 0.03	~~0.66~~ 0.85	~~0.10~~ 0.09	~~0.84~~ 0.83	0.04	0.91	~~0.03~~ 0.08	~~0.47~~ 0.56	0.05	0.6

5.2.2 Comparison with AERONET AOD

The comparison of the daily averaged AOD ~~from~~ time series and corresponding scatter plots calculated using WRF-Chem, MERRA-2, CAMS-OA, MODIS-DB&DT, and MAIAC data with AERONET AOD observations conducted at *KAUST Campus*, *Mezaira* and *Sede Boker* during 2015-2016 period is presented in Fig. 6. Because AERONET conducts observations only during the daylight time, we interpolated WRF-Chem, MERRA-2, CAMS-OA AODs to the AERONET measurements times and then conducted time averaging to make simulated and observed AODs consistent. AODs from MODIS-DB&DT and MAIAC are provided as a daily average. ~~However, MODIS conducts~~ Although MODIS routinely provides observations only twice a day during daylight time, up to four observations might be collected on some days due to overlap of the TERRA and AQUA orbits at some locations.

The scatter plots show that the model and assimilation products ~~reproduce~~ are capable of reproducing the magnitude and temporal evolution of the observed AERONET AOD at all sites~~quite well~~. During both years, *KAUST Campus* and *Mezaira* sites show higher AOD in summer and lower AOD in winter.

To quantify the capability of the WRF-Chem, MERRA-2, and CAMS-OA models, and the MODIS-DB&DT and MAIAC products to reproduce the AERONET AOD, we calculate Pearson correlation coefficient R and mean ~~*BIAS*~~ bias (see Appendix A2) with respect to the AERONET AOD observations for the 2015-2016 period, see Tab. 4. The correlation coefficients are the highest for MERRA-2 and MAIAC. MAIAC shows better correlation than MERRA-2 during 2015 (0.88-0.96), but MERRA-2 is better correlated with AERONET (~~0.88~~0.85-0.91) than MAIAC in 2016. CAMS-OA, despite it does not assimilate AERONET, shows better correlations (0.65~~-0.88~~-0.87) than MODIS-DB&DT (~~0.47-0.92~~0.56-0.84). However, CAMS-OA overestimates AOD, particularly during acute dust events, and has a relatively high positive mean ~~*BIAS*~~bias. The R coefficient for the WRF-Chem AOD is (~~0.44-0.73~~0.43-0.85). MERRA-2 and WRF-Chem have the lowest ~~*BIAS*~~ mean bias in comparison with the other models and products. ~~Both~~ MODIS-DB&DT ~~and MAIAC have a slightly larger *BIAS* than WRF-Chem.~~

~~MODIS-DB&DT *BIAS* is positive except *KAUST Campus* in~~ mean bias is positive in 2015 and 2016, while MAIAC ~~*BIAS*~~ mean bias is negative for *KAUST Campus* and *Mezaira* and positive for *Sede Boker* during both years.

We have to mention here that the satellite retrievals and MERRA-2 use AERONET observations for calibration. WRF-Chem is tuned to reduce the annual mean bias with respect to AERONET observations. CAMS-OA does not assimilate AERONET AODs. In WRF-Chem, we did not tune the temporal correlation between the model and AERONET AOD. In this sense, the correlation coefficient between WRF-Chem and AERONET AOD provides an independent evaluation of the model performance (see Tab. 4). It is expected that the temporal correlation for the assimilation products and satellite retrievals will be higher than for the free-running WRF-Chem.

5.2.3 Comparison of spatial AOD distributions

We also examine how well MERRA-2, CAMS-OA, MAIAC, and WRF-Chem reproduce spatial patterns and seasonal variability of the AOD in comparison with the conventional MODIS-DB&DT retrievals. The seasonally and annually-averaged 2015-2016 AOD fields from WRF-Chem, CAMS-OA, MERRA-2, and the two MODIS retrievals DB&DT and MAIAC are presented in Fig. 7. ~~For the~~ The seasonally averaged AOD's from WRF-Chem, MERRA-2, CAMS-OA are shown at their original spatial resolution and were calculated using only daytime (6 am-2 pm UTC or 9 am-5 pm local time) output. The AODs were sampled under all-sky conditions, which in the ME does not make much of a difference, as cloud fraction is low. For the statistical comparison, we ~~interpolate~~ interpolated AOD fields (preserving the area average AODs) on the MERRA-2 grid and ~~calculate~~ calculated the Pearson correlation coefficient R ~~and~~ root mean square error $RMSE$ ~~(see Tab. 5)~~ and mean bias with respect to MODIS-DB&DT AOD~~.~~

~~The spatial and seasonal variability of the AOD from WRF-Chem, CAMS-OA, MERRA-2, and MAIAC is in good agreement with the MODIS-DB&DT AOD. The WRF-Chem and assimilation products exhibit similar AOD patterns. MAIAC underestimates AOD over the whole domain during all seasons in comparison with MODIS-DB&DT, which is consistent with the MAIAC and MODIS-DB&DT AOD comparison with AERONET AOD (see Fig. 6). CAMS-OA overestimates MODIS-DB&DT AOD in spring and summer but has the best agreement with MODIS-DB&DT in autumn and winter. In summer, when dust emission is at its maximum, WRF-Chem shows higher AOD over the south-eastern part of the Arabian Peninsula in comparison with MODIS-DB&DT AOD, but MERRA-2 underestimates MODIS-DB&DT AOD in summer. In winter and autumn, WRF-Chem and MERRA-2 demonstrate similar patterns and AOD levels, but both of them slightly underestimate MODIS-DB&DT AOD. The extremely high MODIS-DB&DT AOD over the southern part of the Red Sea is real but is overestimated in comparison with the findings of Brindley et al. (2015); Osipov and Stenchikov (2018).~~

see Tab. 5. When conducting statistical analysis, the grid-cells with ~~NAN-values~~ undefined pixels in MODIS-DB&DT and MAIAC retrievals were excluded~~from the analysis both in observations and the model outputs.~~.

The statistical scores provided in Tab. 5 show that the annual mean AOD from ~~MERRA-2 and CAMS-OA have the highest spatial correlations (0.720 and 0.696)and one of the lowest~~ MAIAC has the highest correlation (R=0.796), but also the highest *RMSE*~~(0.088 and 0.093)~~ =0.123, and the biggest bias=-0.095 with respect to MODIS-DB&DT AOD~~, followed by WRF-Chem (~~. MERRA-2 annual mean AOD has R=~~0.653, 0.663 with respect to MODIS-DB&DT AOD, CAMS-OA - ~~*RMSE*~~R=~~0.090)and~~

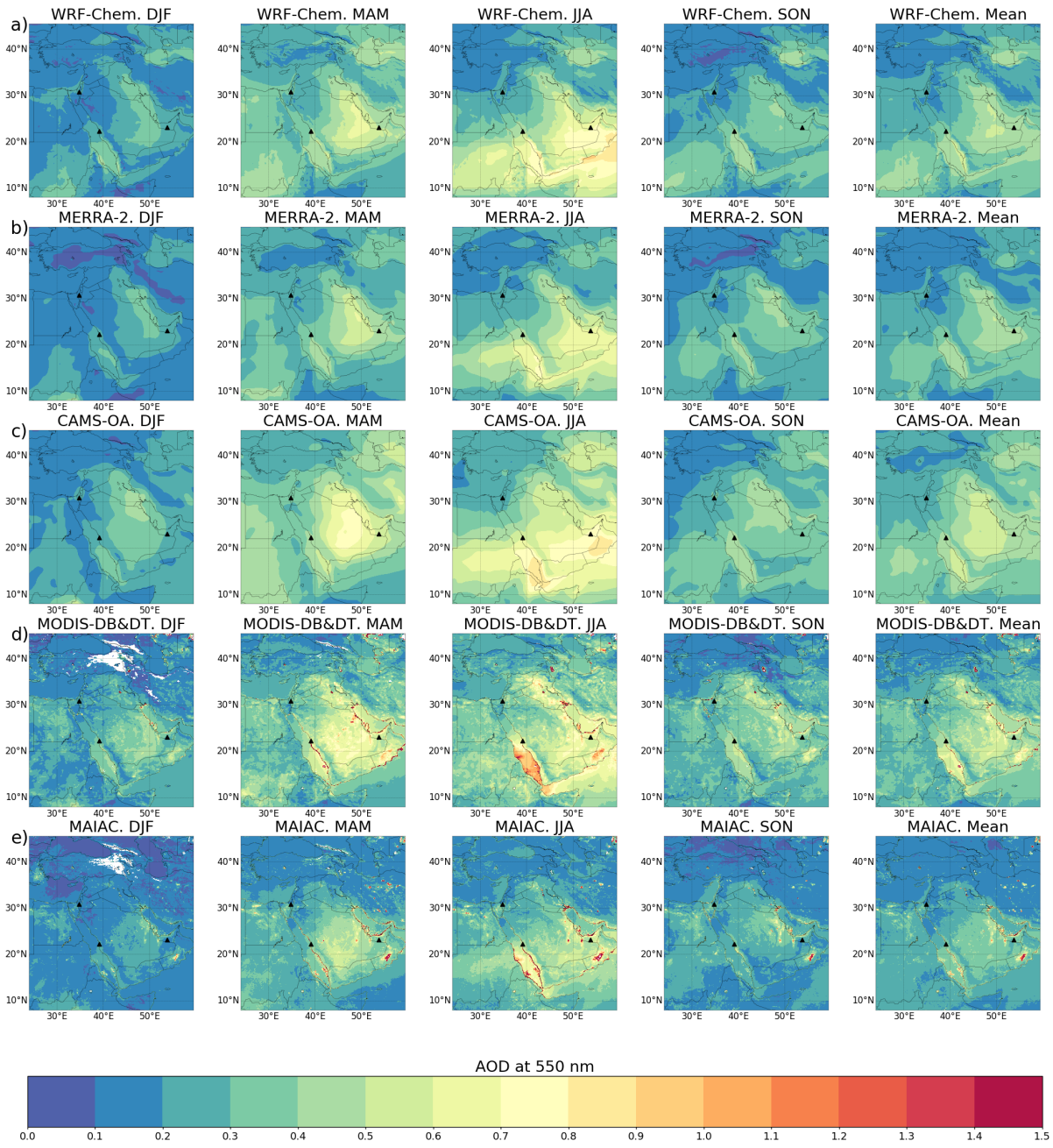


Figure 7. Seasonally averaged 2015-2016 AOD. Right column is annual mean AOD. Rows: a) WRF-Chem, b) MERRA-2, c) CAMS-OA, d) MODIS-DB&DT, and e) MAIAC. Winter (DJF), spring (MAM), summer (JJA), and autumn (SON). White dots are ~~NAN values~~undefined pixels. Black triangles denote locations of *KAUST Campus*, *Mezaira*, and *Sede Boker* AERONET stations.

~~MAIAC AOD~~ 0.650), and WRF-Chem - *R*=0.609 with *RMSE's*=0.116 for all of them. WRF-Chem, ~~which has the lowest correlation (0.608) and highest *RMSE* (0.135). Notably, the difference in terms of *R* and *RMSE* between the two retrieval algorithms~~ MERRA-2, and CAMS-OA demonstrate similar AOD patterns, but WRF-Chem and MERRA-2 underestimate, and CAMS-OA overestimates MODIS-DB&DT ~~and MAIAC is bigger than the difference between WRF-Chem~~ AOD during all seasons with the annual mean bias=-0.009, -0.042, 0.039, correspondingly. MAIAC underestimates AOD in comparison with MODIS-DB&DT, which is consistent with the MAIAC and MODIS-DB&DT ~~.~~ AOD comparison with AERONET AOD (see Tab. 4 and Fig. 6).

Table 5. Pearson correlation coefficient (R)~~and~~ , root mean square error ($RMSE$) and mean bias calculated for seasonally and annually averaged 2015-2016 AOD geographic distributions from CAMS-OA, MAIAC, MERRA-2, and WRF-Chem with respect to MODIS-DB&DT AOD.

	CAMS-OA			MAIAC			MERRA-2		
	R	*RMSE*	*bias*	*R*	*RMSE*	*bias*	*R*	*RMSE*	*bi*
Winter (DJF)	~~0.625~~ 0.599	~~0.074~~ 0.084	~~0.459~~ 0.019	~~0.109~~ 0.794	~~0.577~~ 0.092	~~0.078~~ -0.072	~~0.464~~ 0.569	~~0.082~~0.090	-0.0
Spring (MAM)	~~0.726~~ 0.700	~~0.112~~ 0.129	~~0.526~~ 0.052	~~0.180~~ 0.802	~~0.742~~ 0.142	~~0.102~~ -0.107	~~0.670~~ 0.717	~~0.103~~0.127	-0.0
Summer (JJA)	~~0.737~~ 0.702	~~0.142~~ 0.152	~~0.601~~ 0.069	~~0.190~~ 0.782	~~0.792~~ 0.160	~~0.112~~ -0.117	~~0.747~~ 0.742	~~0.128~~ 0.133	-0.0
Autumn (SON)	~~0.620~~ 0.559	~~0.085~~ 0.111	~~0.455~~ 0.027	~~0.140~~ 0.717	~~0.647~~ 0.111	~~0.087~~ -0.084	~~0.539~~ 0.595	~~0.098~~ 0.108	-0.0
Annual mean	~~0.696~~ 0.650	~~0.093~~ 0.116	~~0.608~~ 0.039	~~0.135~~ 0.796	~~0.720~~ 0.123	~~0.088~~ -0.095	~~0.653~~ 0.663	~~0.090~~0.116	-0.0

Based on the comparison of WRF-Chem AOD with the AOD from ~~AERONET and MODIS~~ MODIS and AERONET observations, we conclude that spatial and temporal WRF-Chem's AOD distribution is in good agreement with the available ~~ground-based and satellite observations.~~ satellite and ground-based observations, i.e. annual mean correlation *R* exceeds 0.6 (see Tab. 5) and correlation with AERONET is 0.43-0.85 (see Tab. 4).

5.3 ~~Air-quality~~PM air pollution

To test the ability of the data assimilation products and models to characterize ~~air-quality~~ PM air pollution in the ME, we compare surface daily mean $PM_{2.5}$ and PM_{10} concentrations from WRF-Chem, MERRA-2, and CAMS-OA, with daily averaged measurements conducted by the three AQMS, see Fig. 8 and 9. The AQMS are installed in Jeddah, Riyadh, and Dammam (Fig. 1), the Saudi Arabian mega-cities. PM measurements conducted by MODON (see Sec. 2.3) are ~~only~~ available starting from 2016. The modeled $PM_{2.5}$ and PM_{10} concentrations were sampled from the model fields at the exact AQMS locations. The following formulas were used to calculate $PM_{2.5}$ and PM_{10} surface concentrations using WRF-Chem and MERRA-2 output:

$$\begin{aligned} PM_{2.5} &= DUST_1 + DUST_2 * 0.38 + SEAS_1 + SEAS_2 + SEAS_3 * 0.83 \\ &\quad + sulfate + (OC_1 + OC_2) * OC_{mfac} + BC_1 + BC_2 \\ PM_{10} &= DUST_1 + DUST_2 + DUST_3 + DUST_4 * 0.74 + SEAS_1 + SEAS_2 + SEAS_3 + SEAS_4 \\ &\quad + sulfate + (OC_1 + OC_2) * OC_{mfac} + BC_1 + BC_2 \end{aligned} \quad (4)$$

where $DUST_{1,2,3,4}$, ~~$SEAS_{2,3,4}$~~ $SEAS_{1,2,3,4}$, $OC_{1,2}$, $BC_{1,2}$, and $sulfate$ are respectively the concentrations of the dust, and sea-salt in the first four bins, ~~sea-salt in the three bins,~~ organic and black carbon (hydrophobic and hydrophilic) and sulfate ion (SO_4^{2-}). As was mentioned in Sec. 4.1, $SEAS_1$ is not present in the WRF-Chem output. So for WRF-Chem we assume $SEAS_1 = 0$. The factor OC_{mfac} = 1.8 accounts for the conversion of organic carbon into organic matter. ~~WRF-Chem simulates only the last four sea salt bins from the GOCART model (see Tab. 2). Because the first sea salt bin is poorly populated, we omit it from calculations of PM for MERRA-2.~~

CAMS-OA $PM_{2.5}$ and PM_{10} were calculated using the following relations (https://confluence.ecmwf.int/display/CUSF/PM10+and+PM25+global+products):

$$\begin{aligned} PM_{2.5} &= DD_1 + DD_2 + SS_1/4.3 + 0.5 * SS_2/4.3 + 0.7 * (OM_1 + OM_2 + sulfate) + BC_1 + BC_2 \\ PM_{10} &= DD_1 + DD_2 + DD_3 * 0.4 + SS_1/4.3 + SS_2/4.3 + OM_1 + OM_2 + sulfate + BC_1 + BC_2 \end{aligned} \quad (5)$$

where $DD_{1,2,3}$, $SS_{1,2}$, $sulfate$, $BC_{1,2}$, $OM_{1,2}$ are surface concentration of dust in three bins, sea salt in two bins, sulfate, black carbon, and organic matter (hydrophobic and hydrophilic). The size ranges of dust and sea salt bins from CAMS-OA are presented in Tab. 2.

The histograms at the right-side panels in Fig. 8 and 9 show the annual mean PM concentrations from WRF-Chem, MERRA-2, and CAMS-OA split into the dust and non-dust components. The dashed and dash-dotted horizontal lines correspond to KSA-PME limits and WHO air quality guidelines for daily (on the left-side panels) and annual mean (on the right-side panels) PM concentrations. We also calculated the separate contributions of sulfate, sea salt, organic matter, and black carbon into the non-dust $PM_{2.5}$ and PM_{10}, see Tab. 6 and 7, respectively. ~~The dashed and dash-dotted horizontal lines correspond to KSA-PME limits and WHO air quality guidelines for daily (on the left-side panels) and annual mean (on the right-side panels) PM concentrations.~~

The sporadic peaks in the observations which are not captured by the model and assimilation products are due to unaccounted factors, such as nearby traffic, construction works, and local anthropogenic or natural emissions, which are not present in the emission inventories, or due to meteorological fluctuations that are not resolved in the models. Talking about extreme dust pollution cases, we analyzed dust surface concentrations using WRF-Chem output during the dust storm, which took place in the Jeddah region on 8th July in 2016. The calculated surface concentrations in all ~~dustbins~~ dust-bins $DUST_{1,2,3,4,5}$ at the peak of the storm were {55,58,63,111,11} $\mu g/m^3$, respectively. The sum of all ~~dustbins~~ dust-bins yields the total dust concentration of 298 $\mu g/m^3$.

5.3.1 $PM_{2.5}$

Fig. 8 shows that the daily averaged $PM_{2.5}$ concentrations observed by MODON AQMS at all locations never drop below the WHO limit of 25 $\mu g/m^3$. During the severe dust events, this limit is exceeded in 2016 10-15 times. The less restrictive KSA-PME limit of 35 $\mu g/m^3$ is exceeded 7-11 times during the dust outbreaks. Annually averaged MODON measurements

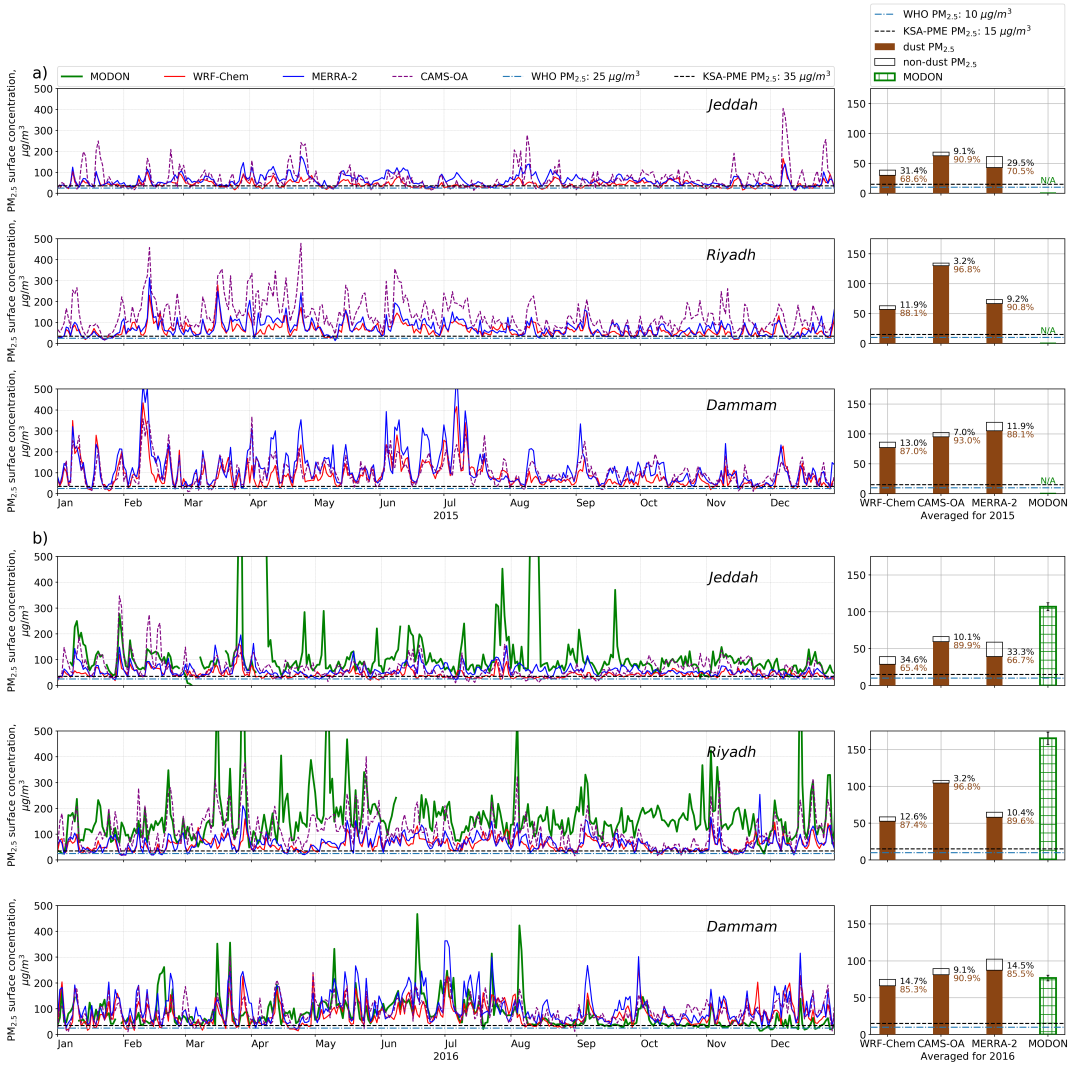


Figure 8. Left: WRF-Chem daily averaged $PM_{2.5}$ surface concentrations ($\mu g/m^3$) with MODON observations, MERRA-2, CAMS-OA at Jeddah, Riyadh, Dammam. The dash-dotted line corresponds to the 25 $\mu g/m^3$ WHO daily average guideline. Right: stacked bars show the decomposition of the $PM_{2.5}$ annual mean surface concentrations into dust and non-dust components. The dash-dotted line corresponds to the 10 $\mu g/m^3$ WHO annual guideline. Numbers on the right hand side of WRF-Chem, CAMS-OA, and MERRA-2 bars show the contribution (%) of the dust and non-dust into the total $PM_{2.5}$ concentration. a) 2015, b) 2016.

are 8-18 times higher than the 10 $\mu g/m^3$ WHO limit and 5-12 times higher than the 15 $\mu g/m^3$ KSA-PME limit for annual mean $PM_{2.5}$ concentrations.

Both data assimilation products and WRF-Chem underestimate ≈3 times annual mean $PM_{2.5}$ concentrations in Jeddah and Riyadh and slightly overestimate, though WRF-Chem slightly underestimates, $PM_{2.5}$ in Dammam in comparison with

observed concentrations during 2016. The CAMS-OA annual mean surface $PM_{2.5}$ concentrations in Jeddah and Riyadh are higher than those from WRF-Chem and MERRA-2, providing the best fit for MODON observations, at least on an annual mean (during 2016) basis.

Annual mean $PM_{2.5}$ concentrations from WRF-Chem and MERRA-2 exceed the WHO limit of 10 $\mu g/m^3$ ~~5 and 11~~ ≈4-7 and ≈6-10 times, respectively, in all locations. The KSA-PME limit of 15 $\mu g/m^3$ for annual average $PM_{2.5}$ concentrations is exceeded ~~3-8~~ ≈2.5-4.5 and ≈4-6.5 times, respectively, for WRF-Chem and MERRA-2.

In Jeddah and Dammam, WRF-Chem and MERRA-2 show similar relative contributions of non-dust components to $PM_{2.5}$ (~~25-33~~30-34% in Jeddah and ~~10-14~~12-14% in Dammam), but in MERRA-2 sea salt is a major contributor into non-dust $PM_{2.5}$, while in WRF-Chem it is sulfate, see Tab. 6. This difference between WRF-Chem and MERRA-2 is mainly because MERRA-2 generates more sea salt, but also because MERRA-2 underestimates ~~SO_2~~ SO_2 emissions located in the Arabian Gulf and along the west coast of Saudi Arabia ~~(Janssens-Maenhout et al., 2013)~~(Ukhov et al., 2020), and hence underestimates sulfate concentrations, as discussed in Sec. 4.1. In Riyadh, the contribution of the non-dust component to $PM_{2.5}$ is ≈~~9-11~~9-12% for both MERRA-2 and WRF-Chem. In CAMS-OA, the contribution of non-dust particulates to $PM_{2.5}$ in Jeddah and Dammam is ≈~~7-11~~7-10%, and the contribution of sea salt is little. According to Tab. 6, in all considered cities, the contribution of black carbon (BC) to $PM_{2.5}$ is not significant ~~in WRF-Chem, CAMS-OA, and MERRA-2~~for all models. In MERRA-2, the contribution of organic matter (OM) to $PM_{2.5}$ is more substantial (but still minor) in comparison with WRF-Chem and CAMS-OA. In general, among all models contribution of dust to $PM_{2.5}$ in Jeddah is 65-90%, while in Riyadh and Dammam this contribution is 85-95%, see Tab. 6.

5.3.2 PM_{10}

Daily averaged MODON measurements almost continuously exceed the WHO guideline of 50 $\mu g/m^3$ at all locations~~.~~, see Fig. 9. In Riyadh and Dammam, PM_{10} concentration is higher than in Jeddah, where the KSA-PME limit of 340 $\mu g/m^3$ for daily averaged PM_{10} is exceeded in 2016 about a dozen times. In Dammam, this limit is more frequently exceeded, especially during the summer period. During acute dust events in Dammam, daily averaged PM_{10} concentrations can exceed the WHO guideline limit by more than 10-20 times. Annually averaged MODON measurements are ~~8-11~~ 7-11 times higher than the 20 $\mu g/m^3$ WHO guideline, and in 2-3 times higher than the 80 $\mu g/m^3$ KSA-PME limits for annual mean PM_{10} concentrations.

~~WRF-Chem simulations,~~ In contrast with MERRA-2 ~~,~~ and CAMS-OA~~data assimilation products compare well~~ , WRF-Chem compares better with PM_{10} observations by MODON in ~~Jeddah and Riyadh. The~~ all locations. MERRA-2 ~~and WRF-Chem~~ overestimates ≈1.2-1.8 times and CAMS-OA underestimates ≈1.5-2 times annual mean PM_{10} ~~concentrations time-series are quite close to each other. Both WRF-Chem and MERRA-2 overestimate observations in Dammam. WRF-Chem reproduces annually-averaged MODON PM_{10} observations quite well, especially in Jeddah and Riyadh.~~

~~CAMS-OA, in general, underestimates PM_{10} concentrations in comparison with the observations.~~ MODON observations in all locations. This is in agreement with Cuevas et al. (2014), who stated that MACC (the predecessor of CAMS-OA) underestimates PM_{10} daily and monthly means all year long, and with our findings in Sec. 5.2.1, where we have shown that CAMS-OA underestimates the volume of particles with radii 0.9-20 μm. Annual mean PM_{10} concentrations from WRF-Chem

Table 6. Contributions (%) of dust and non-dust components into $PM_{2.5}$ for Jeddah, Riyadh, and Dammam during 2015-2016.

	Jeddah			Riyadh			Dammam		
	WRF-Chem	CAMS-OA	MERRA-2	WRF-Chem	CAMS-OA	MERRA-2	WRF-Chem	CAMS-OA	MERRA-2
	2015								
dust	~~74.8~~ 68.6	90.9	70.6	~~91.1~~ 88.1	96.8	90.8	~~90.2~~ 87.0	93.0	88.1
sulf	~~15.6~~ 19.9	5.1	6.1	~~6.8~~ 9.0	2.1	5.0	~~7.5~~ 10.0	3.9	3.6
BC	~~1.7~~ 2.1	0.7	0.6	0.1	0.2	0.3	0.1	0.7	0.3
OM	~~3.3~~ 4.0	3.1	5.1	~~1.4~~ 1.8	0.8	2.7	~~1.2~~ 1.5	2.3	3.1
salt	~~4.5~~ 5.5	0.1	17.6	~~0.7~~ 0.8	0.1	1.3	~~1.0~~ 1.3	0.1	4.9
	2016								
dust	~~71.9~~ 65.4	89.9	66.8	~~90.6~~ 87.4	96.8	89.6	~~89.1~~ 85.3	90.9	85.5
sulf	~~18.9~~ 23.2	5.7	6.8	~~7.0~~ 9.5	1.9	5.7	~~8.4~~ 11.3	4.7	4.3
BC	~~1.7~~ 2.1	0.8	0.7	~~0.1~~ 0.2	0.3	0.3	~~0.1~~ 0.2	0.9	0.4
OM	~~3.2~~ 4.0	3.4	5.4	~~1.6~~ 2.1	0.9	2.9	~~1.3~~ 1.8	3.4	4.1
salt	~~4.4~~ 5.4	0.1	20.4	~~0.7~~ 0.9	0.1	1.4	~~1.1~~ 1.5	0.1	5.7

* for WRF-Chem and MERRA-2: $dust = DUST_1 + DUST_2 * 0.38$, $BC = BC_1 + BC_2$, $sulf = sulfate$, $OM = (OC_1 + OC_2) * OC_{mfac}$, $salt = SS_1 + SS_2 + SS_3 * 0.83$

**for CAMS: $dust = DD_1 + DD_2$, $sulf = 0.7 * sulfate$, $BC = BC_1 + BC_2$, $OM = 0.7 * (OM_1 + OM_2)$, $salt = SS_1/4.3 + 0.5 * SS_2/4.3$

Abbreviations of the aerosols' names correspond to those given in Sec. 5.3.

and MERRA-2 exceed the WHO limit of 20 $\mu g/m^3$ ≈6-15 and ≈10-20 times, respectively, in all locations. The KSA-PME limit of 80 $\mu g/m^3$ for annual average PM_{10} concentrations is exceeded ≈1.5-4 and ≈1.5-5 times, respectively, for WRF-Chem and MERRA-2.

According to Tab. 7 MERRA-2 shows the highest contribution of the sea salt into PM_{10} in the coastal cities of Jeddah ~~and Dammam~~(≈21-25%) and Dammam (≈6-7%). MERRA-2 demonstrates the lowest (≈1-2%) contribution of sulfate to PM_{10}, while WRF-Chem and CAMS-OA produce similar sulfate contribution to PM_{10} in Jeddah ((≈~~6~~7%)) and Riyadh ((≈~~2~~2.4%)). MERRA-2 also shows the lowest contribution (≈0.1-0.2%) of black carbon (*BC*) to PM_{10} in all considered cities. CAMS-OA organic matter (*OM*) contribution to PM_{10} is prevailing ~~(up to 3-4 times)~~ in 2-8 times over the WRF-Chem and MERRA-2 contributions. CAMS-OA demonstrates the lowest (0.1-0.2%) contribution of sea salt to PM_{10}. Contribution of dust to PM_{10} in Jeddah is 70-90%, while in Riyadh and Dammam this contribution is 90-96%. Minimal contribution (≈3.5-4%) of non-dust components to PM_{10} is observed among all models in Riyadh.

5.3.3 Spatial patterns of ~~air-pollution~~PM air pollution

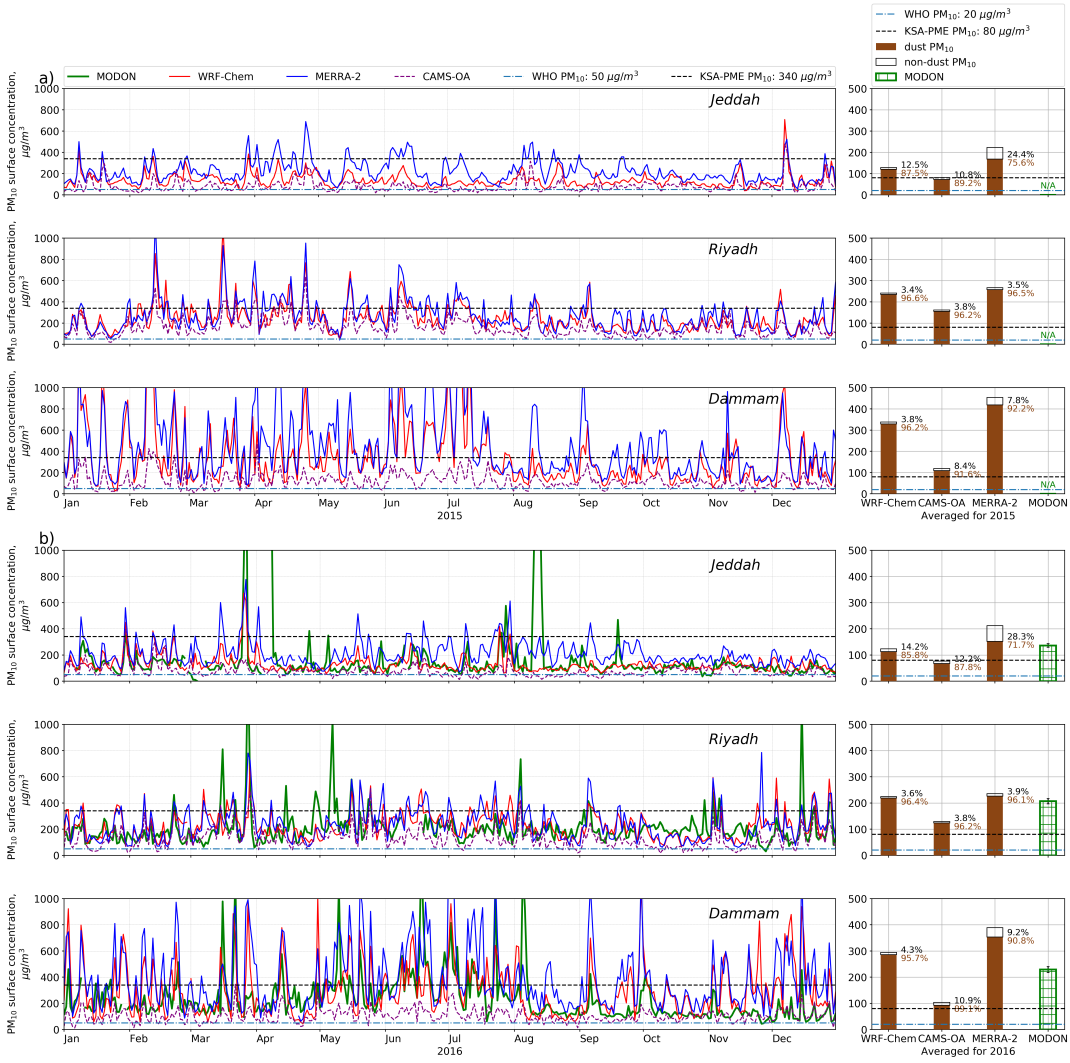


Figure 9. Left: WRF-Chem daily averaged PM_{10} surface concentrations ($\mu g/m^3$) with the MODON observations and MERRA-2 and CAMS-OA, at Jeddah, Riyadh, and Dammam. The dash-dotted line corresponds to the 50 $\mu g/m^3$ WHO daily - guideline. Right: stacked bars show the decomposition of the PM_{10} annual mean surface concentrations into dust and non-dust components. The dash-dotted line corresponds to the 20 $\mu g/m^3$ WHO annual guideline. Numbers on the right-hand side of WRF-Chem, CAMS-OA, and MERRA-2 bars show the contribution (%) of the dust and non-dust particulates to the total PM_{10} concentration. a) 2015, b) 2016 year.

590 ~~As we have shown, WRF-Chem provides reliable estimates of aerosol pollution in major Saudi Arabia's cities, Riyadh, Jeddah, and Dammam. Further, we will~~ In this section we use WRF-Chem output averaged for 2015-2016 to discuss the spatial patterns of aerosol pollution and partial contributions from natural and anthropogenic aerosols into PM over the ME.

Table 7. Contributions (%) of dust and non-dust components into PM_{10} for Jeddah, Riyadh, and Dammam during 2015-2016.

	Jeddah			Riyadh			Dammam		
	WRF-Chem	CAMS-OA	MERRA-2	WRF-Chem	CAMS-OA	MERRA-2	WRF-Chem	CAMS-OA	MERRA-2
	2015								
dust	~~86.6~~ 87.5	89.2	75.6	~~96.2~~ 96.6	96.2	96.5	~~95.9~~ 96.2	91.6	92.2
sulf	~~6.0~~ 6.3	6.2	1.7	~~2.3~~ 2.4	2.6	1.4	~~2.5~~ 2.6	4.8	0.9
BC	~~1.9~~ 0.7	0.6	0.2	~~0.5~~ 0.0	0.2	0.1	~~0.4~~ 0.0	0.6	0.1
OM	1.3	3.8	1.4	0.5	1.0	0.7	0.4	2.8	0.8
salt	4.2	0.2	21.2	0.5	0.1	1.4	0.8	0.2	6.0
	2016								
dust	~~85.1~~ 85.8	87.8	71.7	~~96.0~~ 96.4	96.2	96.1	~~95.4~~ 95.7	89.1	90.8
sulf	~~7.5~~ 7.8	7.0	1.9	~~2.4~~ 2.5	2.3	1.6	~~2.8~~ 2.9	5.7	1.1
BC	~~1.9~~ 0.7	0.7	0.2	~~0.6~~ 0.0	0.2	0.1	~~0.5~~ 0.0	0.8	0.1
OM	1.3	4.2	1.5	0.5	1.1	0.8	~~0.4~~ 0.5	4.2	1.1
salt	~~4.3~~ 4.4	0.2	24.8	0.5	0.1	1.5	0.9	0.2	6.8

* for WRF-Chem and MERRA-2: $dust = DUST_1 + DUST_2 + DUST_3 + DUST_4 * 0.74$, $sulf = sulfate$, $BC = BC_1 + BC_2$, $OM = (OC_1 + OC_2) * OC_{mfac}$, $salt = SS_1 + SS_2 + SS_3 + SS_4$

**for CAMS: $dust = DD_1 + DD_2 + DD_3 * 0.4$, $sulf = sulfate$, $BC = BC_1 + BC_2$, $OM = OM_1 + OM_2$, $salt = SS_1/4.3 + SS_2/4.3$

Abbreviations of the aerosols' names correspond to those given in Sec. 5.3.

Figures 10 a, b, c show the spatial distributions of the $PM_{2.5}$ and PM_{10} surface concentrations and the $PM_{2.5}/PM_{10}$ ratio. The left limits of the color bars for $PM_{2.5}$ and PM_{10} are set to be equal to the corresponding WHO annual guideline ~~limits~~concentrations. Over the whole domain, the annual mean surface concentrations of $PM_{2.5}$ and PM_{10} exceed WHO guidelines of 10 and 20 $\mu g/m^3$, correspondingly. The regions of high surface concentrations coincide with the main dust sources, which span from Northern Iraq to Oman, include Sudan, Egypt, ~~Algeria~~Libya, and Turkmenistan. PM concentrations in these regions exceed even the less restrictive KSA-PME air quality limit for annual mean $PM_{2.5}$ and PM_{10} concentrations by more than 5 times.

In the entire domain, the max, min, and mean values of the $PM_{2.5}/PM_{10}$ ratio (see Fig. 10c) are ~~0.89, 0.24, and 0.38~~ 0.73, 0.20, and 0.31 respectively. As expected, the lower $PM_{2.5}/PM_{10}$ ratios (0.2-0.3) are observed over the dust source regions (i.e., along the eastern Arabian peninsula, Iraq, and northern Africa), where both coarse and fine particles are generated. However, large particles can not be transported as far from source regions as small particles, due to the shorter lifetime of large particles compared with small particles with respect to deposition processes. The higher values (0.4-0.6) of the $PM_{2.5}/PM_{10}$ ratio are observed over south-eastern Europe, Turkey, Ethiopia, and western parts of the Arabian Peninsula that are farther from the main dust sources.

Figure 10d shows the sum of surface concentrations of ~~black carbon and organic matter~~ organic matter and black carbon $(OC_1 + OC_2) * OC_{mfac} + BC_1 + BC_2)$. Their max, min, and mean concentration values are 31.8, 0.2, and 1.3 $\mu g/m^3$ respectively. Their contribution to aerosol pollution over the Arabian Peninsula in WRF-Chem simulations is ~~fairly~~ insignificant. Figure 10e shows the surface concentration of sea salt calculated as a sum of concentrations in each bin $SEAS_2 + SEAS_3 + SEAS_4$. Over the seas and coastal areas, the average concentration of sea salt is 3-12 $\mu g/m^3$. In summer, strong winds in *Somali jet* (see Fig. 3b intensify sea salt emission over the Arabian Sea, creating high surface concentrations of sea salt (27-42 $\mu g/m^3$) along the coasts of Somalia and Oman. Due to prevailing northern winds, the transport of sea salt from the Mediterranean Sea to Egypt and Libya is observed.

The relatively high sulfate surface concentration (see Fig. 10f) is observed in the vicinity of the strong ~~SO_2~~ SO_2 sources located along the west and east coast of Saudi Arabia and over the Arabian Gulf, as well as downwind from those sources, see (Ukhov et al., 2020) for details. Figure 10f also denotes the locations of the AERONET stations, as in Fig. 1. The sulfate concentration at the *KAUST Campus* site is higher than at the *Mezaira* and *Sede Boker* AERONET sites (see Sec. 5.2.1) so it experiences more pronounced contribution of sulfate particulates into the fine mode of the AVSD (see Fig. 4a and Fig. 5).

Due to the prevailing northern winds along the Red Sea, sulfate aerosols originating from ~~SO_2~~ SO_2 emissions along the Red Sea coast spread along the west coast of the Arabian Peninsula towards Yemen. The drift of sulfate from the Arabian Gulf into the interior of the Eastern part of the Arabian Peninsula is also seen. The sulfate annual mean background surface concentration over the US for the period 2003–2012, obtained in Buchard et al. (2016), is 2-3 $\mu g/m^3$, similar to the background concentrations we see in the ME. But in the downwind or in the vicinity of strong ~~SO_2~~ SO_2 point emissions, sulfate concentrations are 3-4 times higher. Similar sulfate surface concentrations for the period 2000-2016 over the US were obtained in van Donkelaar et al. (2019), where the concentrations reach $\approx$10 $\mu g/m^3$ over the eastern part of the US during summer. In Al-Taani et al. (2019) the average 1980-2016 sulfate concentration computed for UAE is 2.5-3 times lower. This difference is caused by the fact that ~~in Al-Taani et al. (2019) , sulfate fields were taken~~ Al-Taani et al. (2019) took the sulfate fields from MERRA-2 reanalysis, which underestimates the ~~SO_2 emissions~~ SO_2 emissions as shown in (Ukhov et al., 2020).

Contributions of dust to $PM_{2.5}$ and PM_{10} calculated as ratios of dust $PM_{2.5}$ to total $PM_{2.5}$, and dust PM_{10} to total PM_{10}, are shown in Fig. 10g and 10h, respectively. Due to relatively low dust surface concentrations over the eastern Mediterranean, Turkey, and south-eastern Europe, the contribution of dust to $PM_{2.5}$ and PM_{10} is 20-50% and 50-80%, respectively. In other areas that are closer to the dust source regions, the contribution of dust to PM is above ~~90~~80%.

Figure 10i shows the ratio between the concentration of sulfate aerosol with respect to the total concentration of $PM_{2.5}$ non-dust aerosols. The max, min, and mean values of this ratio are ~~0.76~~0.84, 0.07, and ~~0.45~~0.52, respectively. This ratio is low over the seas where sea salt is prevalent but consistently exceeds 0.6 over land. Sulfate, therefore, is the primary anthropogenic pollutant over land. In the ~~Arabian Coastal areas,~~ central and southern parts of Saudi Arabia, and over ~~south~~ Iran, sulfate contributes ~~60-80~~60-90% to the total $PM_{2.5}$ non-dust aerosols concentration. Over the other parts of the Arabian Peninsula, the northern part of Sudan, Libya, and Egypt, sulfate contributes approximately 40-60% to total $PM_{2.5}$ non-dust aerosols concentration.

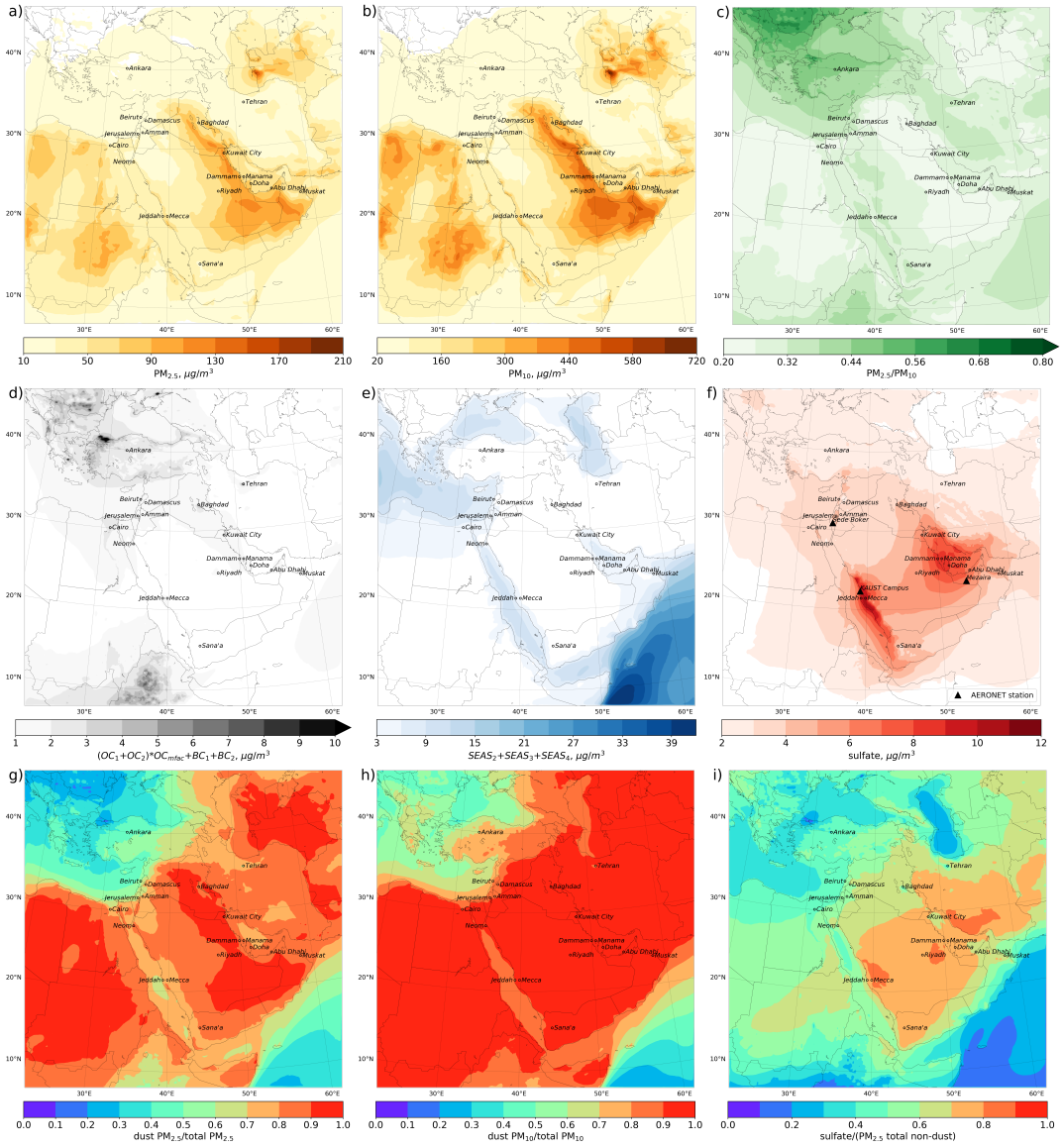


Figure 10. WRF-Chem's average 2015-2016 surface concentrations, ($\mu g/m^3$): a) $PM_{2.5}$, b) PM_{10}; c) ratio $PM_{2.5}/PM_{10}$; d) black carbon and organic matter ($(OC_1 + OC_2) * OC_{mfac} + BC_1 + BC_2$), ($\mu g/m^3$), e) sea salt ($SEAS_2 + SEAS_3 + SEAS_4$), ($\mu g/m^3$), f) sulfate, ($\mu g/m^3$) and locations of AERONET stations, g) ratio dust $PM_{2.5}$/(total $PM_{2.5}$), h) ratio dust PM_{10}/(total PM_{10}), i) ratio sulfate/($PM_{2.5}$ total non-dust). Abbreviations of the aerosols' names correspond to those given in Sec. 5.3.

5.3.4 ~~Air-pollution~~ PM air pollution in ~~urban centers~~the ME major cities

To evaluate the air-quality in the ME's major cities, we calculate for their locations the average for 2015-2016 daily $PM_{2.5}$ and PM_{10} surface concentrations, their 90th percentiles, and we also calculate the contribution of the dust and non-dust components

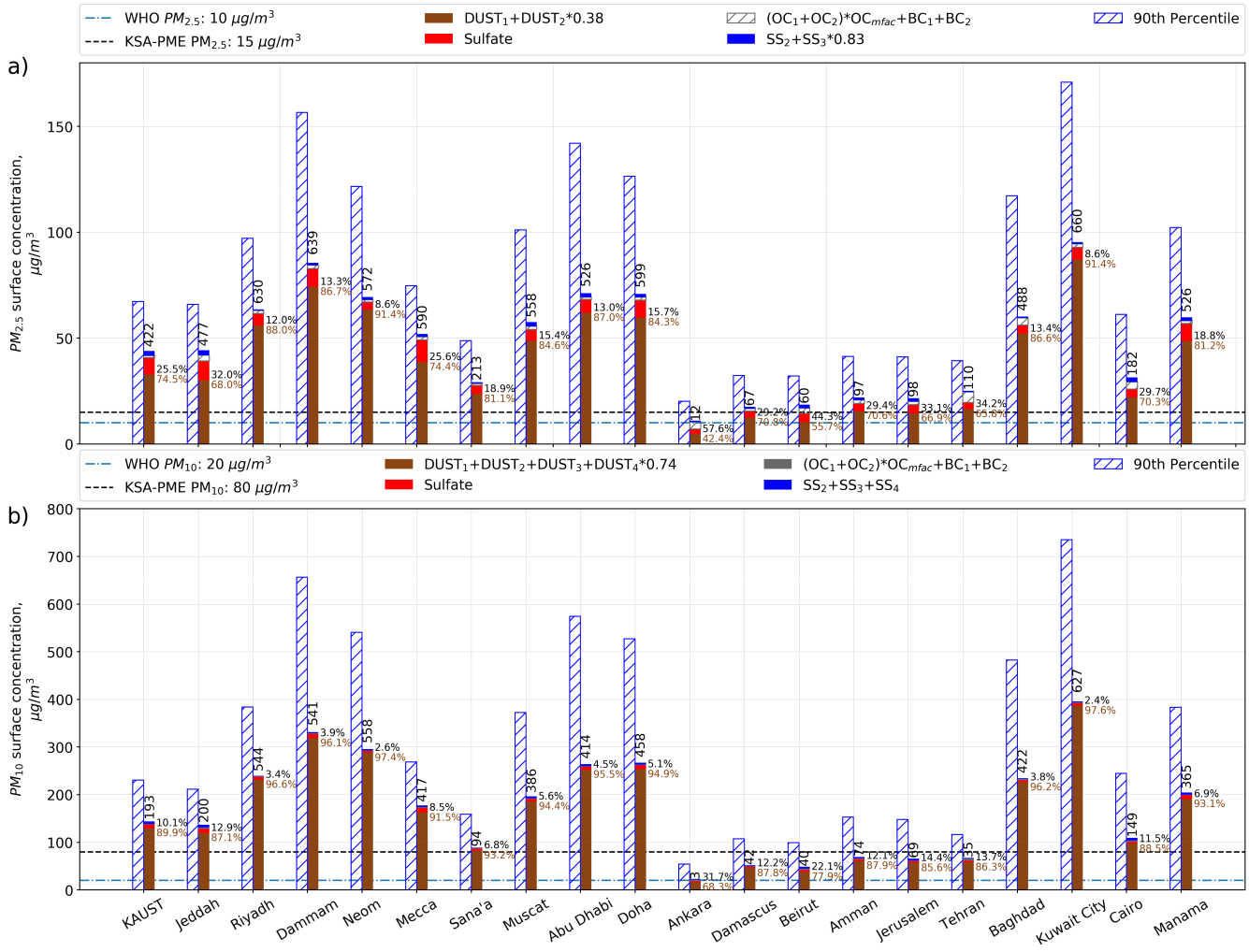


Figure 11. Annual mean 2015-2016 PM surface concentrations ($\mu g/m^3$) calculated for the ME major cities and PM decomposition into dust and non-dust (sulfate, sea salt, black carbon and organic matter) components (stacked bars). Abbreviations of the aerosols' names correspond to those given in Sec. 5.3. Hatched bars denote 90th percentiles ($\mu g/m^3$) calculated using daily mean PM concentrations. WHO guidelines and KSA-PME air-quality limits for annual averaged PM are shown by dash-dotted and dashed lines. Numbers over the stacked bars correspond to the number of days during 2015-2016, when daily averaged PM surface concentration exceeded US-EPA air-quality limit: a) $PM_{2.5}$. Daily averaged US-EPA air-quality limit is 35 $\mu g/m^3$. Annual WHO guideline and KSA-PME limit are 10 and 15 $\mu g/m^3$, respectively; b) PM_{10}. Daily averaged US-EPA air-quality limit is 150 $\mu g/m^3$. Annual WHO and KSA-PME limits are 20 and 80 $\mu g/m^3$, respectively.

into PM (see Fig. 11). We calculate the number of days during the 2015-2016 period when the daily $PM_{2.5}$ and PM_{10} surface concentrations exceed the US-EPA air-quality limit of 35 $\mu g/m^3$ and 150 $\mu g/m^3$ respectively.

Figure 11 shows that the annually-averaged PM_{10} ~~and $PM_{2.5}$~~ $_{2.5}$ and PM_{10} exceed the WHO air-quality guidelines ~~2-20 and 2-12~~ 2-9 and 3-20 times, respectively in all major cities of the ME, except Ankara. The KSA-PME air-quality limit for annual mean PM_{10} ~~$_{2.5}$~~ is exceeded by up to ~~5~~ 6 times, and by up to ~~8~~ 5 times for $PM_{2.5}$ ~~$_{10}$~~. Due to the lack of strong dust sources nearby, air-quality conditions in the cities in the eastern Mediterranean are more favorable when compared with those in the Arabian Peninsula. In these cities, the ~~air-pollution~~ air pollution shifts from natural to anthropogenic, as the contribution of non-dust aerosols to $PM_{2.5}$ increases up to ~~40~~30-45%, in contrast with the cities located in the Arabian Peninsula, where this contribution is up to ~~6-26~~8-25%. Sulfate aerosol is the major contributor to non-dust $PM_{2.5}$.

The cities at the eastern coast of the Arabian Peninsula ~~and Baghdad~~ have the highest 90th percentiles of daily mean PM concentrations. For example, in Dammam, Abu Dhabi, Doha, and Kuwait City, ~~and Baghdad,~~ the 90th percentiles of daily mean surface concentration of PM_{10} and $PM_{2.5}$ ~~exceed 500-750 and 150-230~~ are in the range of 400-740 and 130-180 $\mu g/m^3$ respectively. This is above the KSA-PME air-quality limits for daily mean PM_{10} and $PM_{2.5}$.

In the Eastern Mediterranean cities, within the 2015-2016 period, the US-EPA air-quality daily mean limits are exceeded ~~5-75~~ 40-75 days for PM_{10} and ~~7-208~~ 60-100 days for $PM_{2.5}$. In the cities of the Arabian Peninsula, Iraq, and Iran, the US-EPA PM daily mean limits are exceeded ~~95-626~~ 94-627 days for PM_{10} and ~~230-684~~ 213-640 days for $PM_{2.5}$ during the same period.

6 Conclusions

This study ~~presents, for the first time,~~ assesses the impact of aerosols on air pollution in the Middle East. It presents an evaluation of ~~the advanced components of~~ two aerosol data assimilation products, MERRA-2 and CAMS-OA, as well as high-resolution WRF-Chem simulations over the Middle East. ~~For this purpose, we collected the unique $PM_{2.5}$ and PM $_{10}$ measurements taken in the biggest Saudi Arabian cities, retrievals of aerosol size distribution, and advanced satellite products. For the first time, we compared the~~ In the scope of this study, we conducted high-resolution WRF-Chem simulations for the 2015-2016 period. We evaluated the AOD and PM air pollution over the Arabian Peninsula and in the ME major cities. We also tested the new MODIS AOD retrieval, MAIAC, with the conventional MODIS-DB&DT and AERONET AOD over the ~~dust source regions of the~~ Middle East. ~~In the scope of this study, we conducted advanced two-year high-resolution WRF-Chem simulations.~~

The WRF-Chem ~~code was corrected to better~~ v3.7.1 code was fixed to describe the aerosol effects ~~, and a new capability of using~~ correctly, and MERRA-2 ~~output for~~ has been used for constructing boundary and initial conditions ~~has been developed~~for meteorological and chemistry/aerosol variables. To improve the calculation of sulfate aerosols, the most accurate ~~SO_2~~ SO_2 emission dataset retrieved from OMI observations using the ~~“~~"top-down" approach was implemented in WRF-Chem ~~. The contribution of natural dust, sea salt, and anthropogenic aerosols into the PM was estimated. We found that the three-bin approximation in CAMS-OA is not enough to correctly represent the aerosol size distribution, and MERRA-2 overestimates sea salt and underestimates sulfate concentrations. The air pollution in the major Middle Eastern cities is evaluated.~~

We evaluate the PM air-pollution over the ME during the 2015-2016 period using the regional calculations. We also tuned in WRF-Chem model v3.7.1, MERRA-2 and CAMS-OA assimilation products, and satellite and ground-based AOD observations, as well as *in situ* $PM_{2.5}$ and PM_{10} surface concentration measurements available for 2016.

The regional WRF-Chem model has an advantage of higher spatial resolution (10×10) versus global MERRA-2 ($0.625^\circ\times0.5^\circ$) and CAMS-OA ($0.8^\circ\times0.8^\circ$ before 21 June 2016, and $0.4^\circ\times0.4^\circ$ thereafter). We have modified the WRF-Chem code to correct the discovered model deficiencies. These bug fixes have been implemented in the official WRF-Chem v4.1.3 release. The dust emission and its size distribution are tuned to fit the observed AOD and AVSD. To consistently account for the trans-boundary pollutant transport, MERRA-2 reanalysis was used to construct initial and boundary conditions for WRF-Chem both for meteorological fields and chemistry species. For this, we developed the new pre-processing utility named *Merra2BC* available at (). We equipped WRF-Chem with the novel OMI-HTAP SO_2 the dust emission data-set (Liu et al., 2018) based on the combination of SO_2 emissions taken from the HTAP-2.2 inventory with the catalogue of strong SO_2 point emission sources (Fioletov et al., 2016). This allows us to improve calculation of sulfate aerosol and its contribution to PM. size distribution to match the retrieved AERONET AVSDs.

The model and assimilation products struggle to fit the retrieved AERONET AVSD, failing to correctly reproduce the fine mode in the sub-micron range. MERRA-2 and WRF-Chem use the five-bin GOCART aerosol model with the same five dustbins and demonstrate a better agreement with the AERONET AVSD observations than CAMS-OA that uses only three dustbinsa three dust-bins microphysical model. CAMS-OA overestimates the volume of fine dust particles with radii of 0.55-0.9 μm and underestimates the volume of coarse dust particles with radii of 0.9-20 μm in comparison with AERONET 's AVSD. This CAMS-OA deficiency has been corrected with the latest system upgrade in 2019 by the introduction of a new desert dust scheme (Nabat et al., 2012). AOD and aerosol concentrations are linked through aerosol size distributionand optical propertiesaerosol volume size distribution.

We use in this study AOD observed by three AERONET stations and AOD from two MODIS retrieval products: MODIS-DB&DT and MAIAC. The AOD observations reflect the effect of the mixture of all aerosols, but over the ME land areas dust is dominant. The MODIS-DB&DT and MAIAC retrievals are not identical and the differences between them, with the most accurate AERONET AOD, advise the level of intrinsic AOD retrieval uncertainty. At all At all considered AERONET sites, WRF-Chem, CAMS-OA, MERRA-2, MODIS-DB&DT, and MAIAC reproduce quite well can reproduce the magnitude and temporal evolution of the AERONET AOD time series during the whole considered period. period. MAIAC and MERRA-2 and MAIAC have the highest correlation with respect to AERONET AOD. CAMS-OA shows better correlation than MODIS-DB&DT, although CAMS-OA overestimates AERONET AOD, especially during the severe dust events and exhibits the a relatively high positive *BIAS*bias. The MODIS-DB&DT and MAIAC retrieval products have similar absolute values of the mean *BIAS*bias, which is slightly larger than that of MERRA-2 and WRF-Chem. The MODIS-DB&DT AOD is biased positivelyexcept at *KAUST Campus* in 2016, but the MAIAC AOD is biased negatively except for *Sede Boker* for both years.

The model fields The spatial AOD fields from WRF-Chem and assimilation products exhibit similar AOD seasonal patterns. CAMS-OA overestimates MODIS-DB&DT AOD during spring and summer but has the best agreement with MODIS-DB&DT during the autumn and winter. During summer, patterns, but WRF-Chemshows higher AOD over the south-eastern part of the

~~Arabian Peninsula in comparison with MODIS-DB&DT AOD. In general, MERRA-2 underestimates the MODIS-DB&DT AOD during summer. During winter and fall, WRF-Chem~~ , MAIAC, and MERRA-2 ~~AODs are close, but both of them slightly underestimatethe~~ underestimate, and CAMS-OA overestimates MODIS-DB&DT AOD. ~~The AOD fields from~~ MAIAC has the highest correlation, followed by MERRA-2~~and~~ , CAMS-OA~~have the highest spatial correlationand lowest *RMSE* with respect to MODIS-DB&DT AOD. MAIAC underestimates AOD over the whole domain during all seasons in comparison with MODIS-DB&DT.~~ , and WRF-Chem.

~~MODON AQMS observations of daily mean PM $_{2.5}$ concentrations at all locations never drop below the WHO limit of 25 . The less restrictive~~ The capability of WRF-Chem, MERRA-2, and CAMS-OA in reproducing PM air pollution over the ME was tested against *in situ* measurements. Annual mean PM concentrations from WRF-Chem and MERRA-2 exceed the corresponding WHO limits up to 20 times. The KSA-PME limit ~~of 35 is exceeded by 7-11~~ for annual average concentrations is exceeded up to 6.5 times. The ~~daily mean PM$_{10}$ MODON measurements exceed the WHO guideline of 50 at all locations. PM$_{10}$ concentration is higher in Riyadh and especially in Dammamin comparison with Jeddah. In Dammam the KSA-PME limit for daily averaged PM$_{10}$ of 340 is more frequently exceeded than in Jeddah~~ CAMS-OA annual mean PM$_{2.5}$ fits the MODON AQMS observations better than other products. CAMS-OA, ~~especially during the summer period. Annually averaged MODON measurements are 8-11 times higher than the WHO guideline of 20 and 2-3 times higher than the KSA-PME limit of 80 .~~

~~The capability of WRF-Chem,~~ MERRA-2, and WRF-Chem underestimate ≈3 times annual mean PM$_{2.5}$ observed concentrations during 2016 in Jeddah and Riyadh. CAMS-OA, ~~and CAMS-OA in reproducing the ME air-quality is tested against AQMS measurements.~~ MERRA-2 overestimate, and WRF-Chem underestimates observed annual mean PM$_{2.5}$ in Dammam. In Jeddah ~~and Riyadh~~ and Dammam, WRF-Chem and MERRA-2 show similar relative contributions of non-dust components to PM$_{2.5}$. Still, in MERRA-2, sea salt is a major contributor in non-dust aerosol concentration, while in WRF-Chem~~and~~ , it is a sulfate. This difference is both because MERRA-2 ~~are able to reproduce the PM~~generates more sea salt and underestimates SO_2 emissions, hence underestimating sulfate concentrations. In Jeddah, Dammam, and Riyadh, the contribution of black carbon to PM$_{2.5}$ is insignificant for all models. WRF-Chem results compare better with PM$_{10}$ ~~measurements quite well, but both of them overestimate observations in Dammam.~~ MODON observations in all locations than ones from MERRA-2 and CAMS-OA. MERRA-2 overestimates ≈1.2-1.8 times. CAMS-OA underestimates ~~the~~ ≈1.5-2 times annual mean PM$_{10}$ ~~surface concentrations in the ME region~~ MODON observations in all locations primarily due to its deficiency in the dust size distribution. ~~PM$_{10}$ concentrations from CAMS-OA are 1.5-3.0 times smaller in comparison with those from~~ MERRA-2 ~~and~~ shows the highest contribution of the sea salt and the lowest contribution of black carbon and sulfate to PM$_{10}$ in all locations. CAMS-OA organic matter contribution to PM$_{10}$ is prevailing over the WRF-Chem ~~. In Riyadh and Jeddah WRF-Chem, and MERRA-2 underestimate PM$_{2.5}$ in comparison with MODON observations during 2016 about twice. However, both WRF-Chem~~ and MERRA-2 ~~PM$_{2.5}$ concentrations are in relatively good agreement with MODON observations in Dammam~~contributions. CAMS-OA ~~PM$_{2.5}$ concentrations are higher than in WRF-Chem and MERRA-2 capturing PM$_{2.5}$ observations better than other products in Jeddah and~~ demonstrates the lowest contribution of sea salt to PM$_{10}$. Minimal contribution of non-dust components to PM$_{10}$ is observed among all models in Riyadh.

~~We use WRF-Chem output to conduct the~~ The PM composition analysis ~~. We found that~~ shows that in WRF-Chem, the annual average $PM_{2.5}/PM_{10}$ ratio over the ME is ~~0.38~~0.3. It decreases to ~~0.2-0.3~~ 0.25 over the major dust source regions, i.e.~~along~~ , in the eastern Arabian peninsula, Iraq, and northern Africa. In most parts of the ME, dust is the major contributor to PM~~, but in the eastern Mediterranean and Turkey contribution of the dust component to $PM_{2.5}$ and PM_{10} decreases to 20-50% and 50-80%, respectively.~~

~~Sulfate aerosol contributes~~ . The sulfate aerosol contribution to $PM_{2.5}$ is essential in the areas where strong ~~SO_2~~ SO_2 sources are present, i.e., in the west and east coasts of Saudi Arabia and over the Arabian Gulf. In these areas sulfate surface concentration reaches 8-11 $\mu g/m^3$, while the "clean" background level is 2-4 $\mu g/m^3$. High sulfate content along the west coast of Saudi Arabia is consistent with the increased volume of the fine mode in the *KAUST Campus* AERONET AVSD in comparison with AVSDs from other sites. ~~Sulfate~~ In WRF-Chem, sulfate is the major non-dust pollutant over the ME. Sulfate aerosols contribute ~~60-80~~ 60-90 % to the total $PM_{2.5}$ non-dust aerosols over the ~~western and eastern Arabian coasts, over the~~ central and southern parts of Saudi Arabia ~~, and over the southern~~ and Iran. Over the other parts of the Arabian Peninsula, northern Sudan, Libya, and Egypt, sulfate contributes approximately 40-60 % to the total $PM_{2.5}$ non-dust aerosol concentration. In Jeddah and Dammam, WRF-Chem and MERRA-2 show similar relative contributions of the non-dust component to $PM_{2.5}$ (30-34% in Jeddah and 12-14% in Dammam). In MERRA-2, in contrast with WRF-Chem, sea salt is a major non-dust contributor to $PM_{2.5}$. In CAMS-OA contribution of the non-dust particulates to $PM_{2.5}$ in Jeddah and Dammam is ≈7-10% and the contribution of sea salt is little. In Riyadh, the contribution of the non-dust component to $PM_{2.5}$ is ≈9-12% for both MERRA-2 and WRF-Chem.

The analysis of the annually averaged $PM_{2.5}$ and PM_{10} surface concentrations in the ME major cities ~~conducted using WRF-Chem output~~ shows a very high ~~level of PM pollution~~ PM pollution level. In Dammam, Abu Dhabi, Doha, and Kuwait City, ~~and Baghdad,~~ the 90th percentile of PM_{10} and $PM_{2.5}$ annual mean surface concentrations exceed ~~500-750 and 150-230~~ 400-740 and 130-180 $\mu g/m^3$ respectively, which is above the KSA-PME air-quality limit. In the eastern Mediterranean, dust concentration drops, and ~~the contribution of~~ non-dust aerosols' contribution to $PM_{2.5}$ increases up to ~~25-40~~30-45%. In the cities located in the Arabian Peninsula contribution of the non-dust component to $PM_{2.5}$ is ~~6-26~~8-25%, which limits the effect of the emission control on air-quality. In the eastern Mediterranean cities during the 2015-2016 period, the daily mean surface PM concentrations exceed the US-EPA air quality daily mean limit ~~5-75~~ 40-75 days for PM_{10} and ~~7-208~~ 60-100 days for $PM_{2.5}$. In the ME cities over the Arabian peninsula, Iraq, and Iran, the US-EPA air-quality daily mean limit is exceeded ~~95-626~~ 94-627 days for PM_{10} and ~~230-684~~ 213-640 days for $PM_{2.5}$.

~~In Jeddah and Dammam, WRF-Chem and MERRA-2 show similar contributions of the non-dust component to $PM_{2.5}$ (25-33% in Jeddah and 10-14% in Dammam). In MERRA-2, however, sea salt is a major non-dust contributor to $PM_{2.5}$, while in WRF-Chem it is sulfate. In CAMS-OA contribution of the non-dust particulates to $PM_{2.5}$ in Jeddah and Dammam is ≈7-11%. In Riyadh, the contribution of the non-dust component to $PM_{2.5}$ is ≈9-11% for both MERRA-2 and WRF-Chem.~~

~~In Jeddah, Riyadh, and Dammam, the contribution of black carbon to $PM_{2.5}$ and PM_{10} is not significant in WRF-Chem and both assimilation products. In MERRA-2, the contribution of organic matter to $PM_{2.5}$ is more substantial in comparison with WRF-Chem and CAMS-OA. However, in CAMS-OA, PM_{10} has more organic matter than in WRF-Chem and MERRA-2.~~

~~We also observe the relative increase of organic matter in $PM_{2.5}$ (except WRF-Chem in some cases) and PM_{10} in Jeddah and Dammam in comparison with Riyadh.~~

Thus, in this study, we found that MERRA-2 and CAMS-OA assimilation products, as well as WRF-Chem output despite some intrinsic uncertainties, could be ~~successfully~~ used for evaluation the ~~air-quality over the Arabian Peninsula~~PM air pollution over the ME. All products show the dominant contribution of mineral dust into PM. However, in the Arabian coastal areas where ~~SO_2~~ SO_2 emissions are high, both contributions of sulfate and sea salt could be significant. The broad effect of natural aerosols on air quality in the ME puts stricter requirements on anthropogenic pollution control. The impact of dust could be alleviated by employing specific to desert areas architectural solutions, increasing in-city vegetation cover, and providing air-quality forecasts to alarm the population on hazardous air quality. The developed WRF-Chem modeling framework can be used to simulate other pollutants like ~~NO_2 and O_3~~ NO_x and O_3. The results of the current research could serve as the basis for an improved air-quality forecast system that interactively calculates high-resolution radiative, dynamical, atmospheric chemistry and aerosol processes, driven by natural and anthropogenic emissions. This system will be especially valuable for the prediction of extreme pollution events. It will also improve understanding of the impact of anthropogenic and natural pollution on air quality and human health in the ME region.

Code availability.

1. *Merra2BC* interpolation utility is available at http://github.com/saneku/Merra2BC

Data availability.

1. MERRA-2 ~~is~~ data are available at https://disc.gsfc.nasa.gov/daac-bin/FTPSubset2.pl
2. CAMS-OA ~~is~~ data are available at http://apps.ecmwf.int/datasets/data/cams-nrealtime
3. MODIS-DB&DT AOD level 2 data are available at https://ladsweb.modaps.eosdis.nasa.gov/about/purpose
4. AERONET data are available at https://aeronet.gsfc.nasa.gov/
5. MAIAC data are available at https://lpdaac.usgs.gov/products/mcd19a2v006/
6. HTAP-2.2 emission inventory is available at http://edgar.jrc.ec.europa.eu/htap_v2/index.php?SECURE=123
7. OMI-HTAP SO_2 emission dataset is available at https://avdc.gsfc.nasa.gov/pub/data/project/OMI_HTAP_emis/

Appendix A

A1 ~~Merra2BC interpolation utility~~

~~*Merra2BC* interpolator (available at) creates time-varying chemical boundary conditions based on MERRA-2 reanalysis for a WRF-Chem simulation by interpolating chemical species mixing ratios defined on the MERRA-2 grid to the WRF-Chem grid~~

for initial conditions and boundary conditions. In the case of initial conditions, interpolated values are written to each node of the WRF-Chem grid. In the case of boundary conditions, only boundary nodes are affected.

Merra2BC utility is written in Python. The utility requires additional modules which need to be installed in the Python environment: NetCDF4 interface (to work with netCDF files) and SciPy's interpolation package.

The full MERRA-2 reanalysis dataset including aerosol fields is publicly available online (see "*Code and data availability*" section). Depending on the requirements, one or both of the following aerosol and gaseous collections need to be downloaded: $inst3\_3d\_aer\_Nv$ - gaseous and aerosol mixing ratios () and $inst3\_3d\_chm\_Nv$ - Carbon monoxide and Ozone mixing ratios (). In addition to downloaded mixing ratios, pressure thickness DELP and surface pressure PS fields also need to be downloaded. Spatial coverage of the MERRA-2 files should include the area of the WRF-Chem simulation domain. The time span of the downloaded files should match with the start and duration of the WRF-Chem simulation. For more information regarding MERRA-2 files specification please refer to Bosilovich et al. (2016).

A0.1 Mapping chemical species between MERRA-2 and WRF-Chem

The *Merra2BC* input file $config.py$ contains multiplication factors to convert MERRA-2 mixing ratios of gases given in to . Aerosols are converted from to . When using the GOCART aerosol module in the WRF-Chem simulation, all MERRA-2 aerosols and gases are matched with those from WRF-Chem. To convert MERRA-2 aerosol mixing ratios given in into , multiply by a factor of 10^9. In the case of gases, multiply MERRA-2 mixing ratios by a ratio of molar masses M_{air}/M_{gas} multiplied by 10^6 to convert to , where M_{gas} and M_{air} are the corresponding molar masses. If another aerosol module is chosen in WRF-Chem, then different multiplication factors should be used.

A0.1 Typical workflow

Below are the steps describing how to work with the *Merra2BC* utility:

1. Run $real.exe$, which will produce the initial $wrfinput\_d01$ and boundary conditions $wrfbdy\_d01$ files required by WRF-Chem simulation
2. Download required MERRA-2 collection files
3. Download the *Merra2BC* code from
4. Edit $config.py$ file which contains:
5. Mapping of chemical species and aerosols between MERRA-2 and WRF-Chem
6. Paths to $wrfinput\_d01$, $wrfbdy\_d01$, $met\_em...*$ files;
7. Path to the downloaded MERRA-2 collection files

~~Program *real.exe* sets default boundary and initial conditions for some chemical species. *Merra2BC* adds interpolated values to the existing ones and it may cause incorrect concentration values. To avoid this, run script "*zero_fileds.py*", which will zero the required fields Run script "*main.py*", which will perform the interpolation; as a result, files *wrfinput_d01*, *wrfbdy_d01* will be updated by the values interpolated from MERRA-2 Modify WRF-Chem *namelist.input* file at section *&chem:* set *have_bcs_chem = .true.* to activate updated boundary conditions from MERRA-2 and, if it is needed, *chem_in_opt = 1* to activate updated initial conditions; Run *wrf.exe* program.~~

~~For the usage of the *Merra2BC* interpolator the following python modules need to be installed:~~

- ~~netcdf4:~~
- ~~scipy:~~

A0.1 ~~Meteorological Boundary and Initial Conditions~~

A1 Meteorological Boundary and Initial Conditions

To be consistent with ~~BC&IC~~IC&BC for chemical species and aerosols, we ~~utilized the same~~ developed a procedure to build meteorological ~~BC&IC~~IC&BC from MERRA-2 reanalysis for all required by WRF-Chem meteorological parameters. In particular, the following 3D parameters were processed: pressure (Pa), geopotential height (m), temperature (K), meridional and zonal wind components ($\mathrm{m/s}$), relative humidity (%); 2D parameters: surface pressure (Pa), sea level pressure (Pa), meridional and zonal wind components at $10\mathrm{m}$ ($\mathrm{m/s}$), temperature at $2\mathrm{m}$ (K), relative humidity at $2\mathrm{m}$ (%), skin temperature (K), ice mask (0/1), terrain height (m), land/sea mask (1/0), soil temperature at 0-10 (cm), 10-40 (cm), 40-100 (cm) and 100-200 (cm); soil moisture at 0-10 (cm), 10-40 (cm), 40-100 (cm) and 100-200 (cm); snow depth (m); snow water equivalent ($\mathrm{kg/m^2}$).

A2 Statistics

We calculated the following statistical parameters to quantify the level of agreement between estimations and observations:

Pearson correlation coefficient (R):

$$R = \frac{\sum_{i=1}^{N}\left(F_i - \bar{F}\right)\left(O_i - \bar{O}\right)}{\sqrt{\sum_{i=1}^{N}\left(F_i - \bar{F}\right)^2 \sum_{i=1}^{N}\left(O_i - \bar{O}\right)^2}}. \tag{A1}$$

Root mean square error ($RMSE$):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(F_i - O_i\right)^2} \tag{A2}$$

Mean ~~bias ($BIAS$):~~ *bias*:

$$BIAS\,bias = \frac{1}{N}\sum_{i=1}^{N}\left(F_i - O_i\right) \tag{A3}$$

where F_i is the estimated value, O_i is the observed value, $\bar{F} = \frac{1}{N}\sum_{i=1}^{N} F_i$ and $\bar{O} = \frac{1}{N}\sum_{i=1}^{N} O_i$ their averages and N is the number of data.

A3 Comparison of AERONET and WRF-Chem volume size distributions

The GOCART dust emission formula (2) calculates dust mass flux into the atmosphere within five dust-bins. In this formula the factor C controls the total dust emission mass flux, and the s_p fractions split this flux into five different dust-bins. We assume that $\sum s_p$=1. To match the observed AERONET AVSD we changed the default s_p={0.1, 0.25, 0.25, 0.25, 0.25} to {0.15, 0.1, 0.25, 0.4, 0.1}. It means that 15% of the total dust mass flux is coming as clay and 85% as silt.

In the original formulation the fractions s_p are not normalized and $\sum s_p$=1.1. It is not essential, as the total flux is multiplied by the factor C that is tuned to fit the observed AOD. So we can normalize the original s_p fractions by dividing them to 1.1 and multiplying factor C to 1.1. It will not change any results in eq. (2) but gives the s_p set of {0.09, 0.2275, 0.2275, 0.2275, 0.2275} that is normalized to 1 consistently with our approach. Figure A1 compares the AVSDs calculated with the updated and default s_p fractions for Summer (JJA) of 2015.

Author contributions. A. Ukhov wrote the manuscript and took part in planning and performing the calculations. S. Mostamandi performed the calculations, constructed meteorological IC&BC based on MERRA-2 reanalysis, prepared MAIAC AOD fields, wrote the section on meteorological conditions, and took part in the discussions. G. Stenchikov planned the calculations, led the discussion, and reviewed and improved the manuscript. I. Shevchenko maintained the *KAUST Campus* AERONET station. Y. Alshehri collected, filtered, and validated PM observational data and wrote the section on the PM measurement procedures. J. Flemming and A. Dasilva participated in the discussion, helped to formulate the research program, and reviewed the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. In this work we used AERONET data from the *KAUST Campus* site that was established and maintained by our group with the help of the NASA Goddard Space Flight Center AERONET team. We thank Brent Holben and Alexander Smirnov for the monitoring and regular calibrations of our instruments. We also used data from the *Sede Boker* and *Mezaira* sites and would like to thank their principal investigators Arnon Karnieli and Brent Holben.

We would like to thank Dr. Fei Liu for providing the OMI-HTAP dataset.

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST). For computer time, this research used the resources of the Supercomputing Laboratory at KAUST. The authors would like to thank the Saudi Authority for Industrial Cities and Technology Zones (MODON) for sharing their air quality observational data.

We are thankful to Linda and Mark Everett for proofreading this manuscript.

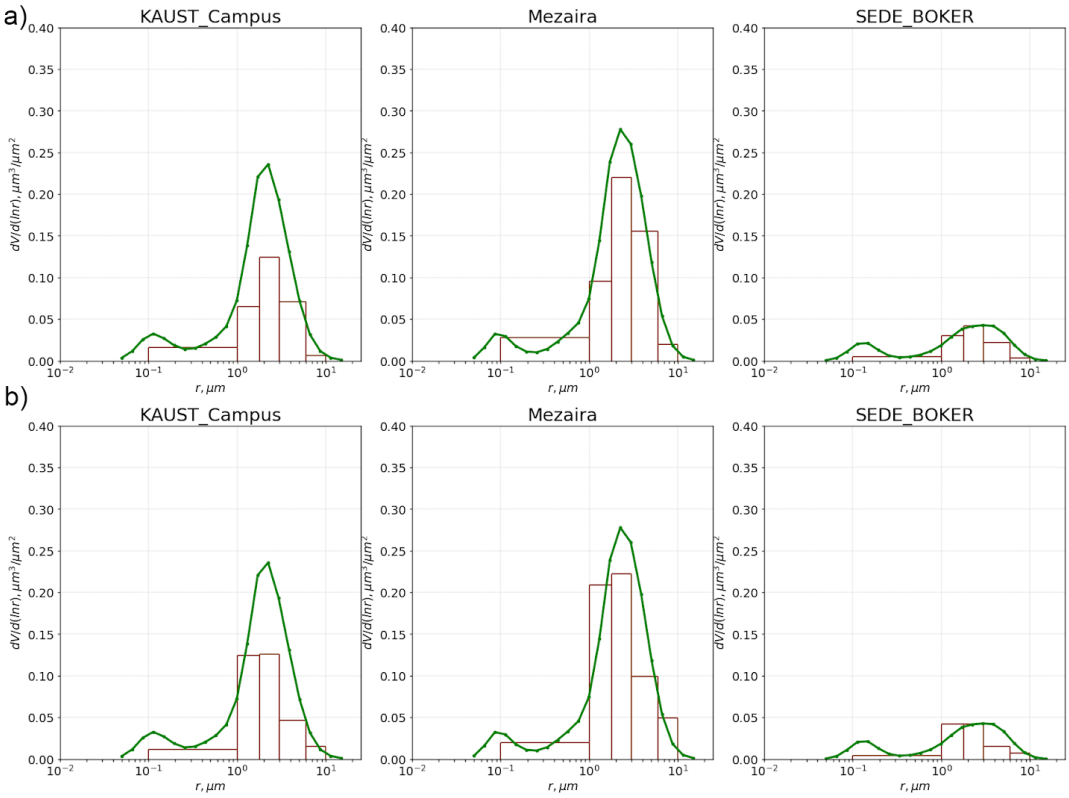


Figure A1. Volume size distributions at *KAUST Campus*, *Mezaira* and *Sede Boker* AERONET sites averaged for JJA of 2015 from WRF-Chem (bars) and from AERONET (solid line): a) updated s_p={0.15, 0.1, 0.25, 0.4, 0.1}, b) default s_p={0.1, 0.25, 0.25, 0.25, 0.25} fractions.

References

Acker, J. G. and Leptoukh, G.: Online analysis enhances use of NASA earth science data, Eos, Transactions American Geophysical Union, 88, 14–17, 2007.

Adebiyi, A. A. and Kok, J. F.: Climate models miss most of the coarse dust in the atmosphere, Science Advances, 6, eaaz9507, 2020.

Al-Jahdali, M. and Bisher, A. B.: Sulfur dioxide (SO2) accumulation in soil and plant's leaves around an oil refinery: A case study from Saudi Arabia, American journal of environmental sciences, 4, 84–88, 2008.

Al-Jeelani, H. A.: Air quality assessment at Al-Taneem area in the Holy Makkah City, Saudi Arabia, Environmental monitoring and assessment, 156, 211, 2009.

Al-Taani, A. A., Nazzal, Y., Howari, F. M., and Yousef, A.: Long-term trends in ambient fine particulate matter from 1980 to 2016 in United Arab Emirates, Environmental monitoring and assessment, 191, 143, 2019.

Alghamdi, M. A., Almazroui, M., Shamy, M., Redal, M. A., Alkhalaf, A. K., Hussein, M. A., and Khoder, M. I.: Characterization and elemental composition of atmospheric aerosol loads during springtime dust storm in western Saudi Arabia, Aerosol Air Qual. Res, 15, 440–453, 2015.

Alharbi, B., Shareef, M. M., and Husain, T.: Study of chemical characteristics of particulate matter concentrations in Riyadh, Saudi Arabia, Atmospheric Pollution Research, 6, 88–98, 2015.

Anderson, J. R.: A land use and land cover classification system for use with remote sensor data, vol. 964, US Government Printing Office, 1976.

Anisimov, A., Tao, W., Stenchikov, G., Kalenderski, S., Prakash, P. J., Yang, Z.-L., and Shi, M.: Quantifying local-scale dust emission from the Arabian Red Sea coastal plain, Atmos. Chem. Phys, 17, 993–1015, 2017.

Anisimov, A., Axisa, D., Kucera, P. A., Mostamandi, S., and Stenchikov, G.: Observations and Cloud-Resolving Modeling of Haboob Dust Storms Over the Arabian Peninsula, Journal of Geophysical Research: Atmospheres, 123, 12–147, 2018.

Archer-Nicholls, S., Lowe, D., Darbyshire, E., Morgan, W., Bela, M., Pereira, G., Trembath, J., Kaiser, J., Longo, K., Freitas, S., et al.: Characterising Brazilian biomass burning emissions using WRF-Chem with MOSAIC sectional aerosol, Geoscientific Model Development, 8, 549–577, 2015.

Bangalath, H. K. and Stenchikov, G.: Sensitivity of the Middle East–North African Tropical Rainbelt to Dust Shortwave Absorption: A High-Resolution AGCM Experiment, Journal of Climate, 29, 7103–7126, 2016.

Banks, J. R., Brindley, H. E., Stenchikov, G., and Schepanski, K.: Satellite retrievals of dust aerosol over the Red Sea and the Persian Gulf (2005–2015), Atmospheric Chemistry and Physics, 17, 3987–4003, 2017.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J., et al.: Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation, Journal of Geophysical Research: Atmospheres, 114, 2009.

Bosilovich, M., Lucchesi, R., and Suarez, M.: MERRA-2: File specification GMAO Office Note No. 9 (Version 1.1), 2016.

Brindley, H., Osipov, S., Bantges, R., Smirnov, A., Banks, J., Levy, R., Jish Prakash, P., and Stenchikov, G.: An assessment of the quality of aerosol retrievals over the Red Sea and evaluation of the climatological cloud-free dust direct radiative effect in the region, Journal of Geophysical Research: Atmospheres, 120, 10–862, 2015.

Buchard, V., da Silva, A., Randles, C., Colarco, P., Ferrare, R., Hair, J., Hostetler, C., Tackett, J., and Winker, D.: Evaluation of the surface PM2. 5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States, Atmospheric Environment, 125, 100–111, 2016.

Buchard, V., Randles, C., Da Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A., Ziemba, L., et al.: The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies, Journal of Climate, 30, 6851–6872, 2017.

Cahill, B., Toumi, R., Stenchikov, G., Osipov, S., and Brindley, H.: Evaluation of thermal and dynamic impacts of summer dust aerosols on the Red Sea, Journal of Geophysical Research: Oceans, 122, 1325–1346, 2017.

Carlson, T. N. and Benjamin, S. G.: Radiative heating rates for Saharan dust, Journal of the Atmospheric Sciences, 37, 193–213, 1980.

Cesnulyte, V., Lindfors, A., Pitkänen, M., Lehtinen, K., Morcrette, J.-J., and Arola, A.: Comparing ECMWF AOD with AERONET observations at visible and UV wavelengths, Atmospheric Chemistry and Physics, 14, 593–608, 2014.

Charlson, R. J., Schwartz, S., Hales, J., Cess, R. D., Coakley, J. J., Hansen, J., and Hofmann, D.: Climate forcing by anthropogenic aerosols, Science, 255, 423–430, 1992.

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A., Higurashi, A., and Nakajima, T.: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, Journal of the atmospheric sciences, 59, 461–483, 2002.

Chuang, C. C., Penner, J. E., Taylor, K. E., Grossman, A. S., and Walton, J. J.: An assessment of the radiative effects of anthropogenic sulfate, Journal of Geophysical Research: Atmospheres, 102, 3761–3778, 1997.

Chuang, M.-T., Zhang, Y., and Kang, D.: Application of WRF/Chem-MADRID for real-time air quality forecasting over the Southeastern United States, Atmospheric Environment, 45, 6241–6250, 2011.

Climate.com: Climate of Middle East, Climate.com, http://climateof.com/middleeast/index.asp, 2018.

Cuevas, E., Camino, C., Benedetti, A., Basart, S., Terradellas, E., Baldasano, J., Morcrette, J.-J., Marticorena, B., Goloub, P., Mortier, A., et al.: The MACC-II 2007-2008 reanalysis: atmospheric dust evaluation and characterization over Northern Africa and Middle East., Atmospheric Chemistry & Physics Discussions, 14, 2014.

Damian, V., Sandu, A., Damian, M., Potra, F., and Carmichael, G. R.: The kinetic preprocessor KPP-a software environment for solving chemical kinetics, Computers & Chemical Engineering, 26, 1567–1579, 2002.

Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, Journal of Geophysical Research: Atmospheres, 105, 20 673–20 696, 2000.

Engelbrecht, J. P., Stenchikov, G., Prakash, P. J., Lersch, T., Anisimov, A., and Shevchenko, I.: Physical and chemical properties of deposited airborne particulates over the Arabian Red Sea coastal plain, Atmospheric Chemistry and Physics, 17, 11 467–11 490, 2017.

EUEA: Air Quality Standards, European Environment Agency, http://ec.europa.eu/environment/air/quality/standards.htm, 2008.

Farahat, A.: Air pollution in the Arabian Peninsula (Saudi Arabia, the United Arab Emirates, Kuwait, Qatar, Bahrain, and Oman): causes, effects, and aerosol categorization, Arabian Journal of Geosciences, 9, 196, 2016.

Fioletov, V. E., McLinden, C. A., Krotkov, N., Li, C., Joiner, J., Theys, N., Carn, S., and Moran, M. D.: A global catalogue of large SO 2 sources and emissions derived from the Ozone Monitoring Instrument., Atmospheric Chemistry & Physics, 16, 2016.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R., Gaudel, A., Inness, A., et al.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF., Geoscientific model development, 8, 2015.

Forkel, R., Werhahn, J., Hansen, A. B., McKeen, S., Peckham, S., Grell, G., and Suppan, P.: Effect of aerosol-radiation feedback on regional air quality–A case study with WRF/Chem, Atmospheric environment, 53, 202–211, 2012.

Furman, H. K. H.: Dust storms in the Middle East: sources of origin and their temporal characteristics, Indoor and Built Environment, 12, 419–426, 2003.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions of dust aerosols simulated with the GOCART model, Journal of Geophysical Research: Atmospheres, 106, 20 255–20 273, 2001.

Gong, S.: A parameterization of sea-salt aerosol source function for sub-and super-micron particles, Global biogeochemical cycles, 17, 2003.

Goudie, A. S. and Middleton, N. J.: Desert dust in the global system, Springer Science & Business Media, 2006.

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., van Der Gon, H. D., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S., Klimont, Z., et al.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period, Climatic Change, 109, 163, 2011.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, 2005.

Hamidi, M., Kavianpour, M. R., and Shao, Y.: Synoptic analysis of dust storms in the Middle East, Asia-Pacific Journal of Atmospheric Sciences, 49, 279–286, 2013.

Heidinger, A. K., Foster, M. J., Walther, A., and Zhao, X.: The pathfinder atmospheres–extended AVHRR climate dataset, Bulletin of the American Meteorological Society, 95, 909–922, 2014.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., et al.: AERONET—A federated instrument network and data archive for aerosol characterization, Remote sensing of environment, 66, 1–16, 1998.

Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R.: Deep blue retrievals of Asian aerosol properties during ACE-Asia, IEEE Transactions on Geoscience and Remote Sensing, 44, 3180–3195, 2006.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., et al.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, Atmospheric chemistry and physics, 15, 5275–5303, 2015.

Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J., Engelen, R., Eskes, H., Flemming, J., et al.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, 2019.

Janssens-Maenhout, G., Pagliari, V., Guizzardi, D., and Muntean, M.: Global emission inventories in the Emission Database for Global Atmospheric Research (EDGAR)–Manual (I), Gridding: EDGAR emissions distribution on global gridmaps, Publications Office of the European Union, Luxembourg, 2013.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., et al.: HTAP_v2: a mosaic of regional and global emission gridmaps for 2008 and 2010 to study hemispheric transport of air pollution., Atmospheric Chemistry & Physics Discussions, 15, 2015.

Jish Prakash, P., Stenchikov, G. L., Kalenderski, S., Osipov, S., and Bangalath, H. K.: The impact of dust storms on the Arabian Peninsula and the Red Sea, Atmospheric Chemistry and Physics, 2015.

Jish Prakash, P., Stenchikov, G., Tao, W., Yapici, T., Warsama, B., and Engelbrecht, J. P.: Arabian Red Sea coastal soils as potential mineral dust sources, Atmospheric Chemistry and Physics, 16, 11 991–12 004, 2016.

Kahn, R. A., Gaitley, B. J., Martonchik, J. V., Diner, D. J., Crean, K. A., and Holben, B.: Multiangle Imaging Spectroradiometer (MISR) global aerosol optical depth validation based on 2 years of coincident Aerosol Robotic Network (AERONET) observations, Journal of Geophysical Research: Atmospheres, 110, 2005.

Kalenderski, S. and Stenchikov, G.: High-resolution regional modeling of summertime transport and impact of African dust over the Red Sea and Arabian Peninsula, Journal of Geophysical Research: Atmospheres, 121, 6435–6458, 2016.

Kalenderski, S., Stenchikov, G. L., and Zhao, C.: Modeling a typical winter-time dust event over the Arabian Peninsula and the Red Sea, Atmospheric Chemistry and Physics, 2013.

Karagulian, F., Belis, C. A., Dora, C. F. C., Prüss-Ustün, A. M., Bonjour, S., Adair-Rohani, H., and Amann, M.: Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level, Atmospheric environment, 120, 475–483, 2015.

Karagulian, F., Temimi, M., Ghebreyesus, D., Weston, M., Kondapalli, N. K., Valappil, V. K., Aldababesh, A., Lyapustin, A., Chaouch, N., Al Hammadi, F., et al.: Analysis of a severe dust storm and its impact on air quality conditions using WRF-Chem modeling, satellite imagery, and ground observations, Air Quality, Atmosphere & Health, pp. 1–18, 2019.

Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E., Chu, A., and Holben, B.: Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer, Journal of Geophysical Research: Atmospheres, 102, 17 051–17 067, 1997.

Khan, B., Stenchikov, G., Weinzierl, B., Kalenderski, S., and Osipov, S.: Dust plume formation in the free troposphere and aerosol size distribution during the Saharan Mineral Dust Experiment in North Africa, Tellus B: Chemical and Physical Meteorology, 67, 27 170, 2015.

Khodeir, M., Shamy, M., Alghamdi, M., Zhong, M., Sun, H., Costa, M., Chen, L.-C., and Maciejczyk, P.: Source apportionment and elemental composition of PM2. 5 and PM10 in Jeddah City, Saudi Arabia, Atmospheric pollution research, 3, 331–340, 2012.

Kim, S.-W., Heckel, A., McKeen, S., Frost, G., Hsie, E.-Y., Trainer, M., Richter, A., Burrows, J., Peckham, S., and Grell, G.: Satellite-observed US power plant NOx emission reductions and their impact on air quality, Geophysical Research Letters, 33, 2006.

Klimont, Z., Smith, S. J., and Cofala, J.: The last decade of global anthropogenic sulfur dioxide: 2000–2011 emissions, Environmental Research Letters, 8, 014 003, 2013.

Klingmüller, K., Pozzer, A., Metzger, S., Stenchikov, G. L., and Lelieveld, J.: Aerosol optical depth trend over the Middle East, Atmospheric Chemistry and Physics, 16, 5063–5073, 2016.

Kok, J. F.: Does the size distribution of mineral dust aerosols depend on the wind speed at emission?, Atmospheric Chemistry and Physics, 11, 10 149–10 156, 2011.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367, 2015.

Levelt, P. F., van den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on geoscience and remote sensing, 44, 1093–1101, 2006.

Levy, R., Hsu, C., et al.: MODIS Atmosphere L2 Aerosol Product. NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA, 2015.

Li, C., Joiner, J., Krotkov, N. A., and Bhartia, P. K.: A fast and sensitive new satellite SO2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument, Geophysical Research Letters, 40, 6314–6318, 2013.

Lihavainen, H., Alghamdi, M., Hyvärinen, A.-P., Hussein, T., Aaltonen, V., Abdelmaksoud, A., Al-Jeelani, H., Almazroui, M., Almehmadi, F., Al Zawad, F., et al.: Aerosols physical properties at Hada Al Sham, western Saudi Arabia, Atmospheric Environment, 135, 109–117, 2016.

Liu, F., Choi, S., Li, C., Fioletov, V. E., McLinden, C. A., Joiner, J., Krotkov, N. A., Bian, H., Janssens-Maenhout, G., Darmenov, A. S., et al.: A new global anthropogenic SO 2 emission inventory for the last decade: a mosaic of satellite-derived and bottom-up emissions, Atmospheric Chemistry and Physics, 18, 16 571–16 586, 2018.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC algorithm., Atmospheric Measurement Techniques, 11, 2018.

Madronich, S.: Photodissociation in the atmosphere: 1. Actinic flux and the effects of ground reflections and clouds, Journal of Geophysical Research: Atmospheres, 92, 9740–9752, 1987.

Marticorena, B. and Bergametti, G.: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, Journal of Geophysical Research: Atmospheres, 100, 16 415–16 430, 1995.

Martin, R. L. and Kok, J. F.: Wind-invariant saltation heights imply linear scaling of aeolian saltation flux with shear stress, Science advances, 3, e1602 569, 2017.

McLinden, C. A., Fioletov, V., Shephard, M. W., Krotkov, N., Li, C., Martin, R. V., Moran, M. D., and Joiner, J.: Space-based detection of missing sulfur dioxide sources of global air pollution, Nature Geoscience, 9, 496–500, 2016.

Middleton, N.: A geography of dust storms in South-west Asia, Journal of Climatology, 6, 183–196, 1986.

Miguez-Macho, G., Stenchikov, G. L., and Robock, A.: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations, Journal of Geophysical Research: Atmospheres, 109, 2004.

Miller, R. and Tegen, I.: Climate response to soil dust aerosols, Journal of climate, 11, 3247–3267, 1998.

Mohalfi, S., Bedi, H., Krishnamurti, T., and Cocke, S. D.: Impact of shortwave radiative effects of dust aerosols on the summer season heat low over Saudi Arabia, Monthly weather review, 126, 3153–3168, 1998.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J., Razinger, M., et al.: Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: Forward modeling, Journal of Geophysical Research: Atmospheres, 114, 2009.

Munir, S., Habeebullah, T. M., Seroji, A. R., Morsy, E. A., Mohammed, A. M., Saud, W. A., Abdou, A. E., and Awad, A. H.: Modeling particulate matter concentrations in Makkah, applying a statistical modeling approach, Aerosol Air Qual. Res, 13, 901–910, 2013.

Myhre, G., Shindell, D., Bréon, F., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J., Lee, D., Mendoza, B., et al.: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, K., Tignor, M., Allen, SK, Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, PM, Cambridge University Press Cambridge, United Kingdom and New York, NY, USA, 2013.

Nabat, P., Solmon, F., Mallet, M., Kok, J., and Somot, S.: Dust emission size distribution impact on aerosol budget and radiative forcing over the Mediterranean region: a regional climate model approach., Atmospheric Chemistry & Physics Discussions, 12, 2012.

Nayebare, S. R., Aburizaiza, O. S., Khwaja, H. A., Siddique, A., Hussain, M. M., Zeb, J., Khatib, F., Carpenter, D. O., and Blake, D. R.: Chemical characterization and source apportionment of PM2. 5 in Rabigh, Saudi Arabia, Aerosol and Air Quality Research, 16, 3114–3129, 2016.

Notaro, M., Alkolibi, F., Fadda, E., and Bakhrjy, F.: Trajectory analysis of Saudi Arabian dust storms, Journal of Geophysical Research: Atmospheres, 118, 6028–6043, 2013.

Notaro, M., Yu, Y., and Kalashnikova, O. V.: Regime shift in Arabian dust activity, triggered by persistent Fertile Crescent drought, Journal of Geophysical Research: Atmospheres, 120, 2015.

Osipov, S. and Stenchikov, G.: Simulating the regional impact of dust on the Middle East climate and the Red Sea, Journal of Geophysical Research: Oceans, 123, 1032–1047, 2018.

Osipov, S., Stenchikov, G., Brindley, H., and Banks, J.: Diurnal cycle of the dust instantaneous direct radiative forcing over the Arabian Peninsula, Atmospheric Chemistry and Physics, 15, 9537–9553, 2015.

Parajuli, S. P., Stenchikov, G. L., Ukhov, A., and Kim, H.: Dust emission modeling using a new high-resolution dust source function in WRF-Chem with implications for air quality, Journal of Geophysical Research: Atmospheres, 2019.

PME: Ambient Air Quality Standard, The Presidency of Meteorology and Environment, 2012.

Prospero, J. M., Ginoux, P., Torres, O., Nicholson, S. E., and Gill, T. E.: Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product, Reviews of geophysics, 40, 2–1, 2002.

Provençal, S., Buchard, V., da Silva, A. M., Leduc, R., and Barrette, N.: Evaluation of PM surface concentrations simulated by Version 1 of NASA's MERRA Aerosol Reanalysis over Europe, Atmospheric pollution research, 8, 374–382, 2017.

Ramanathan, V., Chung, C., Kim, D., Bettge, T., Buja, L., Kiehl, J., Washington, W., Fu, Q., Sikka, D., and Wild, M.: Atmospheric brown clouds: Impacts on South Asian climate and hydrological cycle, Proceedings of the National Academy of Sciences, 102, 5326–5333, 2005.

Randles, C., da Silva, A. M., Buchard, V., Colarco, P., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., et al.: The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation, Journal of climate, 30, 6823–6850, 2017.

Reid, J. S., Piketh, S. J., Walker, A. L., Burger, R. P., Ross, K. E., Westphal, D. L., Bruintjes, R. T., Holben, B. N., Hsu, C., Jensen, T. L., et al.: An overview of UAE2 flight operations: Observations of summertime atmospheric thermodynamic and aerosol profiles of the southern Arabian Gulf, Journal of Geophysical Research: Atmospheres, 113, 2008.

Rienecker, M. M., Suarez, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H., Gu, W., Sienkiewicz, M., Koster, R., Gelaro, R., et al.: The GEOS-5 Data Assimilation System: Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0, Tech. rep., NASA Goddard Space Flight Center, Greenbelt, Maryland, 2008.

Ritter, M., Müller, M. D., Tsai, M.-Y., and Parlow, E.: Air pollution modeling over very complex terrain: an evaluation of WRF-Chem over Switzerland for two 1-year periods, Atmospheric research, 132, 209–222, 2013.

Shao, Y.: A model for mineral dust emission, Journal of Geophysical Research: Atmospheres, 106, 20 239–20 254, 2001.

Shao, Y.: Physics and modelling of wind erosion, vol. 37, Springer Science & Business Media, 2008.

Shi, Y., Zhang, J., Reid, J., Holben, B., Hyer, E., and Curtis, C.: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation, Atmospheric Chemistry and Physics, 11, 557–565, 2011.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 2, Tech. rep., National Center For Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div, 2005.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, Journal of Geophysical Research: Atmospheres, 102, 25 847–25 879, 1997.

Tawabini, B. S., Lawal, T. T., Shaibani, A., and Farahat, A. M.: Morphological and Chemical Properties of Particulate Matter in the Dammam Metropolitan Region: Dhahran, Khobar, and Dammam, Saudi Arabia, Advances in Meteorology, 2017, 2017.

Ukhov, A. and Stenchikov, G.: Merra2BC. Interpolation utility for boundary and initial conditions used in WRF-Chem, https://doi.org/10.5281/zenodo.3695911, 2020.

Ukhov, A., Mostamandi, S., Krotkov, N., Flemming, J., da Silva, A., Li, C., Fioletov, V., McLinden, C., Anisimov, A., Alshehri, Y., et al.: Study of SO2 pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem simulations, Journal of Geophysical Research: Atmospheres, p. e2019JD031993, 2020.

USEPA: National Ambient Air Quality Standards, USEPA, https://www.epa.gov/criteria-air-pollutants/naaqs-table, 2010.

van Donkelaar, A., Martin, R. V., Li, C., and Burnett, R. T.: Regional Estimates of Chemical Composition of Fine Particulate Matter using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors, Environmental science & technology, 2019.

Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., Hu, Y., and Hostetler, C. A.: Fully automated analysis of space-based lidar data: An overview of the CALIPSO retrieval algorithms and data products, in: Laser radar techniques for atmospheric sensing, vol. 5575, pp. 16–30, International Society for Optics and Photonics, 2004.

Wang, X., Liang, X.-Z., Jiang, W., Tao, Z., Wang, J. X., Liu, H., Han, Z., Liu, S., Zhang, Y., Grell, G. A., et al.: WRF-Chem simulation of East Asian air quality: Sensitivity to temporal and vertical emissions distributions, Atmospheric Environment, 44, 660–669, 2010.

WHO: Air quality guidelines: global update 2005, WHO, 2006.

WHO: Ambient air quality and health, WHO, http://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health, 2018.

Yarwood, G., Rao, S., Yocke, M., and Whitten, G.: Updates to the carbon bond chemical mechanism: CB05, Final report to the US EPA, RT-0400675, 8, 2005.

Yu, Y., Notaro, M., Liu, Z., Wang, F., Alkolibi, F., Fadda, E., and Bakhrjy, F.: Climatic controls on the interannual to decadal variability in Saudi Arabian dust activity: Toward the development of a seasonal dust prediction model, Journal of Geophysical Research: Atmospheres, 120, 1739–1758, 2015.

Yu, Y., Notaro, M., Kalashnikova, O. V., and Garay, M. J.: Climatology of summer Shamal wind in the Middle East, Journal of Geophysical Research: Atmospheres, 121, 289–305, 2016.

Zender, C. S., Miller, R., and Tegen, I.: Quantifying mineral dust mass budgets: Terminology, constraints, and current estimates, Eos, Transactions American Geophysical Union, 85, 509–512, 2004.