reviewer comments *response to comment*

Note: The results of the study have been entirely reprocessed with the last pyaerocom version and completed observation and model data, leading to possible changes in the observation and model trends. The values have been changed accordingly in the manuscript. An error has been detected in the script generating Table 4 which conducted to false values in the parenthesis starting from the $PM_{2.5}$ row. This has been corrected.

Reviewer #1

The paper by Mortier et al. reports the results of an interesting study on regional trends in aerosols, focussing on properties which are sufficiently observed by global and regional observational networks. The way the results are presented is appealing, in particular figures 5 and 6. The summary is nicely presented as a series of bullets. The impact of restricted sampling at the monitoring sites on the trends is well explained. I am in favour of publication, but have a number of minor comments that should be addressed before publication.

The authors would like to thank the reviewer for his detailed comments, which contributed to make the manuscript clearer and more intelligible. Below are the answers to the minor comments made by the reviewer.

Minor comments:

- Abstract: Please include a list of the models included in the study.
- A list of the models groups used in this study has been added in the abstract:
    - ***Former version***: The set of observed regional trends has then been used for the evaluation of the climate models and their skills in reproducing the aerosol trends
    - ***Updated version***: The set of observed regional trends has then been used for the evaluation of 11 models (CAMS-reanalysis, 6 AeroCom Phase 3 models, and 4 CMIP6 models) and their skills in reproducing the aerosol trends

- The introduction is too short to my opinion. There is very little information on previous aerosol model-model and model-observations intercomparisons. Activities like AeroCom, CMIP6, should be introduced.
- The introduction has been expanded with references to studies on model-model / model-obs comparisons. The mentioned activities are also introduced in the revised introduction. A reference has also been added regarding the large global AOD increase between 1950 and 1990.

- l89: "Samples are collected every third day". Is this only for the speciated analysis or also for total PM2.5 and PM10?
- Samples are collected every 3rd day for both speciated and PM2.5 and PM10 filters. They all follow the same schedule, and some speciated analyses are actually done on the same filters.

- l102: "The data have been screened" By the authors, by Aas et al. or by the instrument teams? Please give some more details: when is a site regionally representative? Is there a link with the model grid box size? Also for PM the representativity would be good to discuss (in sec 2.1.2).
- The screening has been done by Aas et al., which included instrument teams from the different countries. The words ‘regionally representative’ are referring to the fact

that urban sites were not included in this study, as strongly affected by local emissions. This sentence has been removed as it might be confusing with the representativity study presented in this paper.

Regarding the PM (sec 2.1.2), the following text has been added:
"The stations are mostly located in remote and rural locations. This ensures a good representativity of the measurements as some chemical species contributing to PM observations (i.e., Organic carbon) can present different seasonality and spatial variability."

- l117: Since Gliss et al. is in preparation, it would be good to provide more details on the selection/correction procedure.
- Gliss et al., is now publically available in ACPD. The reference has been updated in the manuscript.

- Sec 2.2.1. It is a bit strange to call the CAMS reanalysis a "climate model" (line 120). Maybe it is good to explicitly refer to "climate models, aerosol models and aerosol reanalyses" in the paper or at least clearly explain the types of models/reanalyses.
- Agreed and changed accordingly.

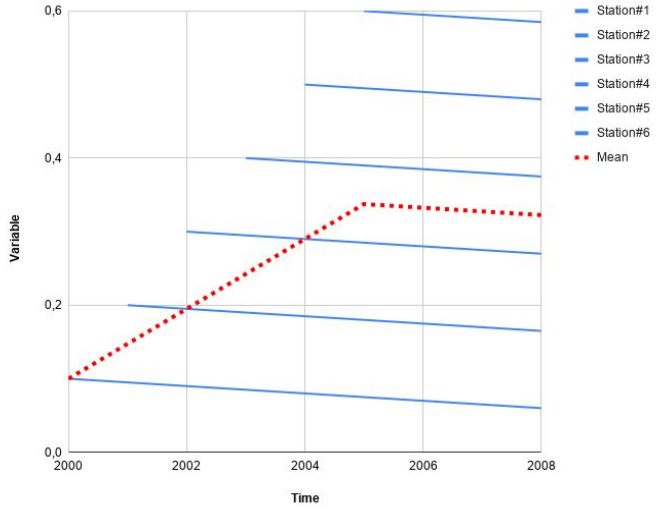
- l129: Please place reference to Inness et al between brackets.
- Done.

- l147: "ESGF nodes". Please explain the acronym and provide a web link or reference.
- An explanation of the acronym has been added as well as a link to the web interface. (Earth System Grid Federation: https://esgf.llnl.gov/)

- l154: ".. into one average time series." This may be removed from the sentence to avoid repetition.
- Done.

- l155: "not as easy to define when combining the trends for individual sites together." Why not? Is there a difference between the trend of a sum and the sum of individual trends?
- We acknowledge that the sentence was not clear enough as the main point was intended to be about the definition of an uncertainty from several individual trends.
  We have clarified and have rewritten the part:
  - ***Former version***: A first advantage of this method is that a single trend can be computed in a given region, with an associated significance and uncertainty, which is not as easy to define when combining the trends for individual sites together.
  - ***Updated version***: One advantage of this method is that a single trend can be computed in a given region, with an associated significance and uncertainty. It is difficult, apart from a diversity analysis, to define such an uncertainty when combining individual trends

The question raised about the difference between the trend of a sum and a sum of individual trends is relevant though. Differences can actually happen due to the fact that the observations might start at different times in areas associated with contrasted absolute levels.



This is an extreme example illustrating that the sum of individual negative trends (solid lines) could be associated with an overall increasing trends (dashed line) if the representativity of the observation network changes dramatically in time.

- l156: "with our aggregation method, even a station that has not provided a sufficient amount of data for computing a trend at its location can still contribute to the computation of a regional time series." This sounds a bit dangerous. Including incomplete time series that have an offset with respect to the mean will introduce a spurious trend. Please provide more discussion on how much this may impact the trends derived?

  The comment is directly connected to the comment below on l180.

- l180: "minimum of 300 valid daily measurements". This is basically one year of data. Why are the authors not more selective, e.g. allowing only stations with measurements for 50% of the time? Why this choice?

  The authors acknowledge that taking into account incomplete time series could introduce erroneous trends. This aspect is covered by the representativity study section (3.3) which compares trends computed from incomplete and complete time series. However, the derivation of the trends from the observations is mostly used, in this work, for making a model trends validation dataset, when colocating observation and model data.

  The minimum of 300 valid daily measurements has been chosen in order to remove temporary AERONET stations (DRAGON campaigns). This is an arbitrary choice that the authors assume to provide the best compromise, considering the different

regions, and different variables, between availability and robustness. For comparisons, AOD trends were computed with the 50% time criteria. This resulted in no available trends in Australia, due to the lack of remaining observations. However, the AOD trends computed with the 300 daily measurements criteria shows a non significant decrease of 1.5%/yr which is assumed to be representative of the trend over the whole period and region, as shown by the representativity study. This illustrates the fact that in some cases, the partial information provided by the observations remains valuable.

- l160: "Seven regions". Please provide the corner locations of the regions, for instance in a small table. Is there overlap between Asia and NAfrica? Are some stations used in multiple regions?
- A table with the regions coordinates has been added in supplementary.
- In its former version, an overlap was observed between NAFRICA and ASIA, which caused the Aeronet Solar_Village station to be considered in both regions. We have now removed the overlap in the updated version.

- l182: "three valid points". Please specify more explicitly. Is a valid point a daily mean observation for one station? Or something else?
    - ***Former version*:** A minimum of three valid points (daily or monthly depending on the available resolution) is required per month to be present in the overall regional time series
    - ***Updated version*:** A minimum of three valid stations is required to be present in the overall regional time series to produce a valid point. In other words, if the available time resolution is daily, at least three stations need to provide valid data for a certain day in order to produce a valid regional mean for that day.

- l201: "Mann-Kendall test ... Theil-Sen". Please provide a reference.
  The following references have been added to the manuscript:
    - Mann-Kendall: Hamed, Khaled H., and A. Ramachandra Rao. "A modified Mann-Kendall trend test for autocorrelated data." Journal of hydrology 204.1-4 (1998): 182-196.
    - Theil-Sen: Sen, Pranab Kumar (1968), "Estimates of the regression coefficient based on Kendall's tau", Journal of the American Statistical Association, 63 (324): 1379–1389, doi:10.2307/2285891, JSTOR 2285891, MR 0258201.

- l208: "residuals". Please provide a definition of the residuals. Is this computed based on the yearly-mean, regional-mean values?
  We have clarified this in the manuscript: Indeed, the residuals are computed based on the difference between the linear trend and the yearly mean values of the regional time series.

- l228: "model subset of data". What does this mean? Make very clear that this study is based on model data only.
    - ***Former version:*** In order to evaluate [...] two sensitivity studies, focusing on the time sampling and the space sampling, have been conducted using model subsets of data.
    - ***Updated version:*** In order to evaluate [...] two sensitivity studies, focusing on the time sampling and the space sampling, have been conducted using NorESM2 model data subsets.

- l230-236: I read the description of the datasets a few times and still I am not sure I understand what is done. Please define the "Ref" and "Exp" very explicitly, maybe even using formula's.

- A sentence has been added before the defintion of the 'ref' and 'exp' datasets.
  The definitions have also been reformulated:
    - ***Former version:***
        - Time representativity study
            - Ref_{time}: Collocation in space and time
            - Exp_{time}: Collocation in space using complete time-series
        - Space representativity study
            - Ref_{space}: Collocation in space using complete time-series (=Exp_{time})
            - Exp_{space}: All grid-points in region using full time-series
    - ***Updated version:***
      The reference dataset corresponds to the model data co-located to the available observations while the experiment dataset uses all model points.
        - Time representativity study
            - Ref_{time}: Model data collocated in space and time with available observations
            - Exp_{time}: Model data collocated in space with available observations using the complete model time-series
        - Space representativity study
            - Ref_{space}: Model data collocated in space with available observations using using the complete model time-series (=Exp_{time})
            - Exp_{space}: All of the model grid-points in the region using the complete model time-series

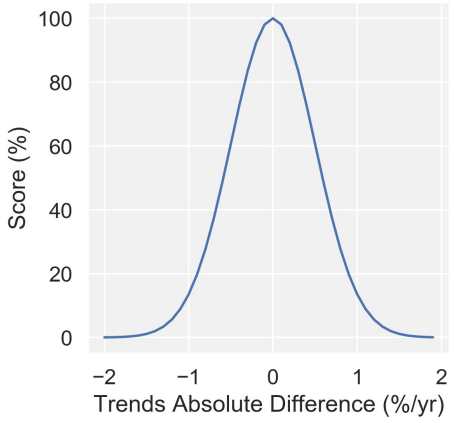
- l238: sigma = 0.5. Why 0.5? It seems sigma has a unit %/year?!
  As mentioned l240, the choice of Sigma=0.5%/yr is an arbitrary choice which results in a representativity score of 50% when the trends difference (between a reference and an experiment subsets) is of 0.5%/yr. The unit was indeed forgotten, and has been added in the manuscript.

- Eq.2: I'm struggling with the representativity and the normal distribution. A relative trend is expressed as %/year. Therefore this has a dimension. But a normal

distribution takes a dimensionless quantity as argument? Therefore Eq.2 does not make sense.

The authors agree that this part of the manuscript was misleading and have modified the manuscript in order to clarify the representativity strategy.

The normal distribution has been used in order to map the trends difference to a score, expressed in percent, as shown by the following figure.



The equation 2 was misleading and has been removed since it was not further used in the manuscript.

- ***Former version***: The difference between the relative trends are computed for each parameter and region. Those differences are then converted into a score (\unit{\%}) by using a normal distribution f described by a mean μ = 0 and a standard deviation of σ = 0.5. The choice of these parameters leads to a representativity score of 100% when there is no difference in the trends of a reference and an 240 experiment dataset, while a difference of 0.5\unit{\%/yr} would indicate a representativity score of 50%.
- ***Updated version***: The difference between the relative trends are computed for each parameter and region. In order to summarize the representativity, those differences are then converted into a score (\unit{\%}) by using a mapping function which has been defined based on a normal distribution. The choice of the parameters describing this function leads to a representativity score of 100\% when there is no difference in the trends computed for a reference and an experiment dataset, while a difference of 0.5\unit{\%/yr} obtained with these two datasets would indicate a representativity score of 50\%.

- Fig. 4: Why is the number of points a coloured region, and not a simple line? Fig. 4: Which model is used for this?
- There was no particular reason for the fact that the points are displayed were colored regions and not simple lines. The Figure has been re-processed with simple lines in order to avoid any confusion.
- NorESM2 was used for the computation of the representativity study since all the variables were available with this model. A sentence in the text and the Figure caption has been added.

- l278: Is there an explanation why PM2.5 is a larger fraction of PM10 in Europe compared to North America?
  The following text has been added in the manuscript:
  “This difference in the relative proportion of fine particles against coarse particles in Europe and North America may be due in part to our definition of regions. Putaud et al. (2010) presented a phenomenology of PM data in Europe showing coarse aerosol tended to be highest in southern Europe which in our study is part of the North Africa region. The discrepancy in the relative proportions of coarse and fine aerosol in Europe and North America may be exacerbated by both a decrease in North America of the fine particles concentration due to pollution mitigation strategies coupled with the growth of the coarse mass due to increasing contributions of natural and agricultural sources, particularly in the western half of the U.S. (Hand et al., 2019a).”

- l305: "Collaud Coen". Provide reference.
- Done

- l310: "smoother". What does this mean?
- The word ‘smoother’ has been replaced by ‘more homogeneous’

- l314: "somewhat higher". Do you mean "somewhat more negative" or "less negative"?
- The authors meant ‘more negative’. This has been replaced in the manuscript.

- l321: The difference with Collaud Coen deserves more discussion. Is this trend significant? Is the difference understood?
  The following text has been inserted in the manuscript:

  This probably illustrates the difference of methodology which consists of computing the mean of station trends in one case, and the trend of a regional time series in the other case, especially when only few measurements are available. However, as shown by the representativity study (\ref{fig:obs_trends}), the non-significant increase of +0.0\%/yr found, in this study, with the observations is similar to the trend derived over the whole region and using complete time series of the NorESM2 model data.

- l339: second increase: should this be "decrease"?
- Yes indeed, this error has been fixed in the manuscript

- l394: "could be caused by increased wet scavenging". How does this match with the SO4 negative trends?
  Agreed. The hypothesis has been removed from the manuscript.

- l464: The large trend in Arctic? Do the authors have any idea how to explain this?
- The following text has been added to the manuscript:
  “In addition, one finds large and significantly increasing trends in the high Arctic that could be explained by a change in the air mass circulation pattern, or by the increase of open sea, which might contribute to a higher production of sea salt aerosols.”

- l542: "brightening Streets et al. (2006); Norris and Wild (2007)." Please place references between brackets.
- Done

- References: Please provide the DOIs for all the cited papers.
- Done

- Olivié, D. et al.: in preparation. Please remove or provide full author list and title. The same remark holds for a few other preprints/in preparation papers.
- These papers were removed from the manuscript

Reviewer #2

Received and published: 21 April 2020

The authors derived trends of total aerosol optical depth, small particle optical depth, large particle optical depth etc. from ground based observations and models. The authors analyzed these trends separated by regions and from 2000 through 2014. The authors show that a limited spatial coverage of ground based observations leads to the AOD trend derived from them not representing the trend over most of the regions except for Europe where ground based observations are most densely populated. The authors also compared observed trends with trends derived from models. In addition, using one of CMIP6 models, the authors show regional trends as well as global trends.

I have **two major issues** on the current version and **one suggestion**. Once the paper emphasizes sampling issues in ground based observations in an application of validating global models, the topic discussed in the paper is relevant within the scope of ACP. The authors did a sensitivity study to test how well the trend derived from ground-based observations represents the trend for the entire region. The result shows that only the AOD trend derived over Europe and Australia represents the entire region (i.e. f factor discussed in Section 3.3 is less than 0.5 so that the true trend falls within a 60% confidence interval). The result of this sensitivity study is only presented as thick black borderlines in Figure 5. In addition, the result of the sensitivity study is not treated the uncertainty in Figure 6. **Because the sampling uncertainty is a part of the uncertainty in observed trends, the error bar attached to the observation need to include this sampling error. When the sampling error is included as the uncertainty, the error bars of the observed trends are much larger. I suggest including the sampling uncertainty in the error bar. Then significant modeled trends consistent with observations are those within the error bar.**

The authors appreciate the two major comments and the suggestion of the reviewer. We believe that the intended objectives of the paper were not precisely enough described in the manuscript. We added some transitions to reinforce the connections in between the different sections and clarified the objectives of the paper in a revised introduction.

The aim of the representativity study is to assess whether the single use of ground-based observations can be utilized to derive representative trends over regions during the considered time period. The result of this study shows that most of them do not actually permit the derivation of such accurate trends due to partial coverage in time and space. However, those observed computed trends can still be used for the evaluation of the model trends, when co-locating the dataset in time and space with available observations. Figure 6 describes how well the models can reproduce the observed trends, whether these trends are representative for the whole region/time period or not. For this reason, the authors decided not to include the representativity study as an uncertainty in the Figure 6 since the models are co-located with the observations and are computed with the same amount of data.

The second point is related my comment above. **The connection between the first paragraph of the Section 4.3.1 and second paragraph is weak**. The first paragraph seem to conclude that regional trends derived from limited number of ground based observations can lead a misleading trend. Then why do the authors need to discuss global trends where ground based observations even represent less? Could you elaborate more the reason for discussing the global trend without showing any observations to compare (given the point the authors made in the first paragraph)? One cannot even estimate the uncertainty in the global trend other than perhaps discussing spreads among the models. But the spread is not the uncertainty in the modeled trends. Moreover, Section 4.3 focuses on mostly one model (NorESM2). Furthermore, the authors mention briefly that the ADO trend agrees with the trend derived from MODIS but the trends derived in this study are from 2000 to 2014 while the study by Zhang and Reid was published in 2010, i.e. their period is shorter than the period used in this study. Therefore, I do not think that their result cannot compare with the trend derived from 2000 to 2014 data.

The end of the first paragraph of the Section 4.3.1 indeed relates the lack of observations for describing accurate regional trends for most of the parameters considered in this study. The assessment of the global trends is performed in this section, without the use of observations since, as indicated in the first paragraph, the partial coverage of the observations in space and time do not permit derivation of such global trends. The single model used in this section (NorESM, as being the only model for which all of the nine parameters were available for this study) provides data at the global scale and for each timestamp of the study period. All of the model data (grid boxes and timestamps) are used to derive the global trends presented in this section. While only one model is used in this section, Figure 6 shows that NorESM2 presents, for most of the parameters/regions, similar trends to other models. This suggests that the use of this single model would probably not deliver a wrong picture of the aerosol global trends.
While the authors agree to the fact that no uncertainty can be associated with the derived global trends due to the lack of observations, the authors also provide global trends of AOD for all of the models used in this study (l448). The spread of the global trends (which is indeed not similar to an uncertainty) indicates that 90% of the models reveal increasing global AOD over the study period. In addition, while the study period is not the same, the comparison of the global AOD trend with MODIS is (+0.003/decade against +0.0028/decade) also tends to confirm this global slight increase, that the authors do not expect to change dramatically within 4 years.

The authors have reinforced the connection between the first and the second paragraphs of the Section 4.3.1:
l441: ‘At the opposite of observations, models provide data at a global scale and along the whole study period. The completeness of those datasets offers the opportunity to derive global aerosol trends. ’

Given my two comments above, my suggestion is to significantly shorten Section 4.3 and focus on analysis of the representativeness of ground based observations. The results of Section 3.3 are only briefly presented in Figure 5 and are not discussed in detail. The

number of ground sites was dramatically changed during the period analyzed in this study (2000 to 2014) as shown in Figure 1. The authors seem to have done the analysis of the impact so why not discus in detail?

According to our previous answers, the authors would like to preserve the structure of the document by not emphasizing the representativity study as the main part of this paper. This study intended to bring the attention on the potential artificial trends produced by the lack of data. We believe that a more detailed analysis of this representativity issues could be the subject of a separate dedicated paper.

Some minor comments Section 3.3. The description of the method needs to be given more. For example:

- Line 230 to 236, the authors say “collocation”. But I did not understand what was collocated with what till I read the caption of Figure 4. Figure 4 only shows two regions. Perhaps include a table showing “f “factors for all regions?
  Reviewer#1 also mentioned the lack of clarity in these definitions. The manuscript has been reworked as follows in order to make the text more intelligible:
  - ***Former version:***
    - Time representativity study
      - Ref_{time}: Collocation in space and time
      - Exp_{time}: Collocation in space using complete time-series
    - Space representativity study
      - Ref_{space}: Collocation in space using complete time-series (=Exp_{time})
      - Exp_{space}: All grid-points in region using full time-series
  - ***Updated version:***
    The reference dataset corresponds to the model data co-located to the available observations while the experiment dataset uses all model points.
    - Time representativity study
      - Ref_{time}: Model data collocated in space and time with available observations
      - Exp_{time}: Model data collocated in space with available observations using the complete model time-series
    - Space representativity study
      - Ref_{space}: Model data collocated in space with available observations using using the complete model time-series (=Exp_{time})
      - Exp_{space}: All of the model grid-points in the region using the complete model time-series

- Also it is not clear how the number of points shown in the top plots of Figure 4 is related to the number of observations.
  The Figure 4 caption has been completed as follows:
  - ***Former version:*** Ref_{time} corresponds to the model output collocated in space and time to the available observations. Exp_{time}/Ref_{space} corresponds to the model output collocated in space to the stations providing

measurements, using the complete time series from 2000 to 2014. Exp_{space} corresponds to the model output in the whole geographic region (see \ref{fig:map_obs}) without any collocation to the observations.
    - ***Updated version:*** The blue color (Ref_{time}) corresponds to the model output collocated in space and time with the available observations. The upper graphs show an overall increase in the number of available observations (more stations) combined with a seasonal cycle (less AOD available in wintertime). The orange color (Exp_{time}/Ref_{space}) corresponds to the model output collocated in space to the stations providing measurements, using the complete time series from 2000 to 2014. The green color (Exp_{space}) corresponds to the model output in the whole geographic region (see \ref{fig:map_obs}), using all of the grid boxes without any collocation to the observations.

- Line 362. “sign” instead of “direction”?
  Agreed:
    - ***Former version****:* the models show trends in the same direction as the observations [...]
    - ***Updated version:*** the models show trends with the same sign as the observed trends [...]
- Line 400 to 402. The statement might be true, but it is also possible that AE is less sensitive to the change in a relative sense.
  The authors agree with that remark and updated the manuscript as follows:
    - ***Former version*****:** the trends are usually smaller than for AOD in the respective regions, meaning that the amount of the particles is more subject to variations than the size (type) of these particles.
    - ***Updated version*****:** the trends are usually smaller than for AOD in the respective regions. This can mean that the amount of the particles is more subject to variations than the size (type) of these particles but could also illustrate that AE is less sensitive to the change in a relative sense.

- Conclusions stated in the conclusion section need to be more specific. For example, please state the regions instead of saying “some observations”
  Agreed. Some specifications have been added to the manuscript:
  i.e., “Significant decreases are found in Europe, North America, South America, North Africa and Asia”

Evaluation of climate model aerosol trends with ground-based observations over the last two decades - an AeroCom and CMIP6 analysis

Augustin Mortier[1], Jonas Gliß[1], Michael Schulz[1], Wenche Aas[2], Elisabeth Andrews[3], Huisheng Bian[4], Mian Chin[5], Paul Ginoux[6], Jenny Hand[7], Brent Holben[5], Zhang Hua[8], Zak Kipling[9], Alf Kirkevåg[1], Paolo Laj[10], Thibault Lurton[11], Gunnar Myhre[12], David Neubauer[13], Dirk Olivié[1], Knut von Salzen[14], Ragnhild Skeie[12], Toshihiko Takemura[15], and Simone Tilmes[16]

[1]Norwegian Meteorological Institute, Oslo, Norway
[2]NILU, Norwegian Institute for Air Research, Kjeller, Norway
[3]Cooperative Institute for Research in Environmental Sciences,University of Colorado, Boulder, Colorado, USA
[4]Maryland Univ. Baltimore County (UMBC), Baltimore, MD, USA
[5]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[6]NOAA, Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[7]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA
[8]Laboratory for Climate Studies, National Climate Center, China Meteorological Administration, Beijing, China
[9]European Centre for Medium-Range Weather Forecasts, Reading, UK
[10]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, Institute for Geosciences and Environmental Research, Grenoble, France
[11]Met Office Hadley Centre, Exeter, UK
[12]CICERO Center for International Climate and Environmental Research, Oslo, Norway
[13]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[14]Environment Canada, Montréal, Canada
[15]Research Institute for Applied Mechanics, Kyushu University, 6-1 Kasuga-koen, Kasuga, Fukuoka, Japan
[16]National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA

Correspondence: Augustin Mortier (augustinm@met.no)

Abstract. This study presents a multi-parameter analysis of aerosol trends over the last two decades at regional and global scales. Regional time series have been computed for a set of nine optical, chemical composition and mass aerosol properties by using the observations ~~of~~ from several ground-based networks. From these regional time series the aerosol trends have been derived for different regions of the world. Most of the properties related to aerosol loading exhibit negative trends, both at the surface and in the total atmospheric column. Significant decreases of aerosol optical depth (AOD) are found in Europe, North America, South America~~and North Africa~~, North Africa and Asia, ranging from ~~-1.3~~-1.2%/yr to -3.1%/yr. An error and representativity analysis of the ~~incomplete~~ spatially and temporally limited observational data has been performed using model data subsets in order to investigate how likely the observed trends represent the actual trends happening in the regions over the full study period from 2000 to 2014. This analysis reveals that significant uncertainty is associated with some of the regional trends due to time and space sampling deficiencies. The set of observed regional trends has then been used for the evaluation of ~~the climate models~~ 10 models (6 AeroCom Phase III models, and 4 CMIP6 models) and the CAMS-reanalysis dataset, and their skills in reproducing the aerosol trends. Model performance is found to vary depending on the parameters and the regions

of the world. The models tend to capture trends in AOD, column ~~Angstrom~~ Ångström exponent, sulfate and particulate matter well (except in North Africa), but show larger discrepancies for coarse mode AOD. The rather good agreement of the trends, across different aerosol parameters between models and observations, when co-locating them in time and space, implies that global model trends, including those in poorly monitored regions, are likely correct. The models can help to provide a global picture of the aerosol trends by filling the gaps in regions not covered by observations. The calculation of aerosol trends at a global scale reveals a different picture from ~~the one~~ that depicted by solely relying on ground based observations. Using a model with complete diagnostics (NorESM2) we find a global increase of AOD of about 0.2%/yr between 2000 and 2014, primarily caused by an increase ~~of~~ in the loads of organic aerosol, sulfate and black carbon.

1 Introduction

As one of the key gears involved in the climate mechanism (Pöschl, 2005), and as a predominant component of air quality that affects human health (Burnett et al., 2014), aerosols have been increasingly subject to observation over the last two decades, both from ground and space-based platforms (Holben et al., 2001; Kaufman et al., 2002). Aerosols are also recognized to have an important role for the fertilization of the Amazon forest (Yu et al., 2015), and in other socioeconomic fields such as the solar energy production ~~(Li et al., 2017; Labordena et al., 2018)~~(Li et al., 2017; Labordena et al., 2018; Sweerts et al., 2019).

Through their direct, semi-direct and indirect effects (Rap et al., 2013; Johnson et al., 2004; Lohmann and Feichter, 2005), aerosol particles are crucial for the estimation of the radiative forcing. Currently, the overall estimate of aerosol radiative forcing is associated with high uncertainties (Haywood and Boucher, 2000; Stocker, 2014). Some of the reasons for these uncertainties reside in the heterogeneity of atmospheric particles, both in terms of their microphysical and optical properties, as well as the high variability of these aerosols in space and time. ~~The different~~ Different regions of the world exhibit contrasting aerosol properties (Holben et al., 2001), which can vary depending on the seasons, from year to year, and possibly exhibit inter-annual trends (Streets et al., 2009). In addition to natural emissions such as sea salt and dust, anthropogenic sources of aerosol add another layer of complexity. The Second Industrial Revolution, which relied on the use of fossil fuel energy, has had a significant impact on the aerosol load on a global scale, and on the local air quality, resulting in severe pollution episodes, such as the famous 1952 smog event in London ~~, 1952~~ (Bell et al., 2004) that caused the death of thousands of people within a few days. Starting in the 1970s in the US, and in the 1990s in Europe, mitigation measures were ~~established~~ implemented to limit the emission of particles and other pollutants (Bryner, 1995; Turnock et al., 2016) resulting in significant improvements in terms of air quality and particle concentration levels (Likens et al., 2001). In recent decades there has been a shift of anthropogenic emissions from Europe and North America to the developing nations, which are now facing, in varying degrees, the air quality issues that affected Europe and North America 40 years ago (Streets et al., 2008; Ramachandran et al., 2012).

In order to provide realistic radiative forcing estimates and projections, it is important for the atmospheric models to be able to capture the long-term aerosol trends caused by both natural and anthropogenic variations. ~~With a consistent multi-parameter analysis, this~~

Assessing and improving the modelling of aerosols in global earth system models is the main objective of the AeroCom-project. Specific experiments are conducted within this initiative with a focus on individual aerosol species, such as dust (Huneeus et al., 2011) or organic aerosols (Tsigaridis et al., 2014), while dedicated control experiments aim to enable an assessment of the global aerosol modelling. Both Kinne et al. (2006) and, more recently, Gliß et al. (2020) present evaluations of global aerosol optical properties simulations by focusing on AeroCom control experiment data for a specific year.

This study presents an overview of ~~aerosol trends~~ the aerosol trends for multiple aerosol parameters (optical and chemical) over the last two decades using ground based observation network data as a reference for the evaluation of the models' skills in reproducing ~~the aerosols~~ those trends.

To serve that purpose, this study addresses the following ~~three~~ two questions:

- What are the observed aerosol trends over the last two decades in the different regions of the world? (Section 4.1)
- Can the climate models reproduce these observed trends? (Section 4.2)

Then, having developed an understanding of the models' skills in reproducing the observed aerosol trends, the last section of this study aims to answer the following question:

- What are the global aerosol trends derived from the model data? (Section 4.3)

The CAMS reanalysis dataset and output from six AeroCom models and four CMIP6 models (both model groups performed historical experiments) are evaluated in this work. CMIP6 (Coupled Model Intercomparison Project, Phase 6) is an intercomparison project organised by the WCRP (World Climate Research Program). Participating models will contribute to the assessment of the climate change in the upcoming 2024 IPCC (Intergovernmental Panel on Climate Change) report.

Figure 1 presents the time series of ~~global modeled AOD~~ modeled global AOD (aerosol optical depth) between 1850 and 2014. All of the climate models appear to exhibit a large increase in AOD, especially between 1950 and 1990 (Tegen et al., 2000) , followed by more stable conditions up to the present. While the models show some diversity in absolute values, the trends (focus of this paper) seem to be consistent ~~,~~ among the different models ~~,~~ at a global scale. ~~The aerosol optical measurements, which started to develop~~ Long-term monitoring of many optical and chemical parameters were initiated in the late 1990's ~~, allow investigation of the~~ (e.g., Holben et al., 2001; Laj et al., 2020), providing a high quality dataset for the investigation of aerosol trends over the last two decades ~~, and~~ (e.g., Collaud Coen et al., 2020; Hand et al., 2019a; Aas et al., 2019). These observational datasets also offer an opportunity to validate the modeled trends in this period. Since 2014 is the last year available from the CMIP6 historical runs, we focus this study on the aerosol trends in the 2000-2014 period.

2 Datasets

A set of nine column and in situ surface aerosol datasets are used in this study. The observation networks and the models providing output for these parameters are reported in Table 1.

2.1 Observations

For each of the parameters used in this study, data of the highest quality level provided by the different observation networks were used. Mountain sites, corresponding to an elevation above 1000 m, were excluded, mainly because global models have problems ~~to simulate~~ simulating the aerosol distribution in ~~steep~~ complex terrain (Kinne et al., 2013).

2.1.1 Columnar aerosol optical properties

The AErosol RObotic NETwork (AERONET) is a network established by NASA (National Aeronautics and Space Administration), and expanded by national and international collaborations. AERONET operates aerosol ground-based measurements in the different regions of the world (Holben et al., 2001). The observation of the columnar aerosol properties is performed by standardized and calibrated solar-powered CIMEL Electronique sunphotometers. These instruments measure the solar radiation reaching the surface of the Earth at different wavelengths and for different optical geometries. A new version of the sunphotometer (CE318-T) is also able to perform night-time measurements using the moon as light-source (Barreto et al., 2016). The direct measurements (aiming at the light-source) allow for the derivation of the aerosol optical depth (AOD), and the ~~Angstrom~~ Ångström exponent (AE) which are related to the amount and size of the particles, respectively. The spectral information can be further utilized to derive the AOD for the fine and the coarse particles, split by diameter less than or greater than ~~1 μm~~ 1.2 µm (O'Neill et al., 2003). Three different data quality levels are available depending on the application of cloud filtering and correction for instruments calibration derivations (Smirnov et al., 2000, 2004). The level 2.0 version 3 daily data, which provides automatic instrument anomaly quality controls (Giles et al., 2019), are used in this study for four different parameters: AOD (calculated at 550 nm), AE (calculated using 440 nm and 870 nm channels), ~~AOD<1μm~~ $\mathrm{AOD_f}$ (or fine AOD), and ~~AOD>1μm~~ $\mathrm{AOD_c}$ (or coarse AOD) corresponding to the AOD of the particles whose diameter is less than and greater than ~~1 μm~~1.2 µm, respectively.

2.1.2 Particulate matter concentrations

The particulate matter (PM) measurements are from EMEP (covering Europe), and IMPROVE (for North America). The PM data have been made available either via the EBAS database infrastructure (http://ebas.nilu.no), or in the original IMPROVE data to be found in the VIEWS database (http://views.cira.colostate.edu/). Both $\mathrm{PM_{10}}$ and $\mathrm{PM_{2.5}}$ (with unit $\mu\mathrm{g\,m^{-3}}$) are used in this study. Note that the $\mathrm{PM_{10}}$ size fraction of particles below 10 ~~μm~~ µm encompasses the $\mathrm{PM_{2.5}}$ aerosol mass below 2.5 ~~μm~~µm.

The first PM measurements in EMEP started in 1996 and the number of sites increased steadily in the following decade (Tørseth et al., 2012). Most of the sites use the gravimetric method for both size fractions, though some used automated

monitors, i.e. TEOM FDMS or b-attenuation. The EMEP monitoring complies with the European standards, i.e EN12341:2014 for the gravimetric methods and EN16450:2017 for the automatic methods.

The IMPROVE network has been operating since 1988 at predominantly remote and rural sites across the United States. This ensures a good representativity of the measurements as some chemical species contributing to PM observations (i.e., Organic carbon) can exhibit different seasonality and spatial variability. IMPROVE uses four separate modules to collect samples for speciated $PM_{2.5}$ analysis and gravimetric $PM_{2.5}$ and PM_{10} bulk mass measurements. Samples are collected every third day for 24 h and reported at local conditions. $PM_{2.5}$ and PM_{10} mass concentrations are determined from Teflon filters from two separate modules sampling with $PM_{2.5}$ and PM_{10} inlets, respectively. The gravimetric mass measurements are not performed at controlled relative humidity and temperature, and a laboratory relocation in 2011 resulted in unstable weighing conditions. Therefore, gravimetric mass measurements from 2011-2018 were subject to potentially high relative humidity conditions and likely contain particle bound water on the filters that could bias trends (Hand et al., 2019b).

2.1.3 ~~air~~ Sulphate aerosol concentration

The sulphate aerosol concentration (SO_4) dataset is a subset of the global data presented in Aas et al. (2019) and is based on measurements obtained in different regional networks as described in Table 1. The sulfate aerosol concentrations are obtained from analysis of aerosol filters. In the EMEP, CASTNET, CAPMON and EANET networks, these are either sampled with a PM_{10} inlet or a total aerosol inlet, with no specific size cut off effective, using a filterpack sampler. In the IMPROVE network, sulfate measurements are done using a filterpack sampler with a $PM_{2.5}$ inlet. The filters are typically analysed by ion chromatography after water extraction of the aerosol filter.

The data have been screened to be ~~regionally representative and~~ of satisfactory quality. Urban sites are not included, nor are sites where the surroundings have changed considerably in the period in question. In Aas et al. (2019) the data were averaged to monthly means. When the data have lower sampling frequency than daily, samples are weighted prior to averaging in accordance with how many days were sampled in a given month.

2.1.4 Scattering and absorption coefficients

~~Due to the scarcity of stations available for long-term trend analysis (only 28), the presence of regionally non-representative stations (e.g., stations located near roads, in cities), difficult to capture by global models, can have large effects on the computation of the regional average time series. The urban stations have therefore been removed from this analysis~~For the surface in-situ PM measurements, the scattering and absorption coefficients measurements were accessed through EBAS database infrastructure. The level 2 data (quality controlled, hourly averaged, reported at standard temperature and pressure (STP) conditions) were used~~for two parameters~~. Detailed information on the quality assurance and quality control procedures for GAW aerosol in-situ data are available in (Laj et al., 2020). The difference in measurement conditions (i.e., observations being made at STP versus models simulating at ambient conditions) was not expected to impact the calculated trends so no adjustment was made to account for this.

Scattering and absorption coefficients are measured by different instruments:

- Scattering coefficients (σ_{sp}, in Mm^{-1}), were measured by integrating nephelometers. For better consistency in the model comparisons (model data for these parameters are reported for RH=0%), only the measurement data obtained when the relative humidity in the instrument was lower than 40% were utilized (Pandolfi et al., 2018).

- Absorption coefficients (σ_{ap}, in Mm^{-1}), were obtained from filter-based absorption photometers.

Due to the scarcity of stations available for long-term trend analysis (only 28), the presence of regionally non-representative stations (e.g., stations located near roads, in cities), difficult to capture by global models, can have large effects on the computation of the regional average time series. The urban stations have therefore been removed from this analysis.

Altogether the same data selection procedures (exclusion of stations, removal of outliers) and corrections (conversion to coefficients at 550 nm wavelength) were applied as in ~~(Gliß et al., 2020),~~ Gliß et al. (2020), which describes the AeroCom evaluation ~~analysis~~ of the Control 2019 experiment, analysing AeroCom simulations of the year 2010 in detail.

2.2 Models

A set of ~~11 climate models is~~ 10 climate and aerosol models and a aerosol reanalysis dataset are used in this study. Their main characteristics are reported in Table 2. These models can be separated into three main groups.

2.2.1 CAMS-Reanalysis

The CAMS reanalysis, which is the successor to the MACC reanalysis (Monitoring Atmospheric Composition and Climate), is the latest global reanalysis dataset of atmospheric composition produced by the Copernicus Atmosphere Monitoring Service (Inness et al., 2019). It is produced using 4DVar data assimilation in the CY42R1 model cycle of the ECMWF (European Centre for Medium-Range Weather Forecasts) Integrated Forecast System (IFS), with 60 hybrid sigma/pressure vertical levels. The model used in the CAMS reanalysis includes several updates to the aerosol and chemistry modules on top of the standard CY42R1 release. The IFS model assimilates several satellite products, from aerosols (AOD) to greenhouse gases (CO2, CH4) ~~Inness et al. (2019)~~(Inness et al., 2019), where most relevant for aerosol trends are data from both MODIS sensors and AATSR/ATSR2. Daily data, from the ECMWF data archive (MARS), were used in this study. The CAMS reanalysis data set covers the period January 2003 to near real time. The three first years of this study period (2000-2002) are missing for this model.

2.2.2 AeroCom phase III

~~The~~ Initiated in 2000, the AeroCom-project (https://aerocom.met.no) is an open international initiative of scientists interested in the advancement of the understanding of the global aerosol and its impact on climate (Schulz et al., 2006). Different model experiments have been conducted during the third phase of this project, ~~initiated~~ started in 2015, in order to investigate specific topics (~~eg~~ e.g., dust, volcanic aerosols, aerosol absorption, hygroscopicity, etc.) but also the general modelling of the aerosols (control experiments). The model versions ~~are also as close as possible linked to those GCM~~ and parametrizations

used in AeroCom are closely linked to the versions used for CMIP6 and~~for instance AerChemMIP~~, for instance, AerChemMIP (Aerosol Chemistry Model Intercomparison Project) climate experiments.

In this study, we use the model outputs from the historical AeroCom experiment, whose main aim is to understand the regional trends in aerosol distribution from 1850 to 2015 and to quantify the aerosol forcing with a main emphasis on the direct aerosol effect. The models were also run in various configurations, providing different degrees of constraints on the evolving meteorological conditions, such as using monthly fixed sea-surface temperature (SST), historically evolving SSTs, and basic meteorology fields, e.g., wind for a given year.

2.2.3 CMIP6

The upcoming 2024 IPCC sixth assessment report (AR6) will feature new state-of-the-art CMIP6 ~~(Couple Model Intercomparison Project, Phase 6)~~ models with model runs in higher resolution and with new physical processes. An overview of the experimental design and organisation can be found in Eyring et al. (2016). In this study, we use a preliminary extract of the data of four CMIP6 models from the historical experiment, as available on ESGF nodes (Earth System Grid Federation: https://esgf.llnl.gov), which provided output from 1850 to 2014. 2014 was selected as the last year of the study period of the analysis presented here.

3 Methods

3.1 Regional time series

Due to the nature of the processes involved in the emission and the deposition of aerosols, one can expect different trends in different regions of the world. Instead of investigating the trends obtained at each individual observation station in a given region, we resort here to the analysis of average regional time series as computed by assembling all measurements at stations in each region~~into one average time series. A first~~. One advantage of this method is that a single trend can be computed in a given region, with an associated significance and uncertainty~~, which is not as easy to define when combining the trends for individual sites together~~. It is difficult, apart from a diversity analysis, to define such an uncertainty when combining individual trends. Also, with our aggregation method, even a station that has not provided a sufficient amount of data for computing a trend at its location can still contribute to the computation of a regional time series. The computation of such aggregated regional time series makes most sense in regions exhibiting similar seasonal patterns.

3.1.1 Regions definition and observations coverage

Seven regions are considered in this study. The definition of these regions has been done in a pragmatic way to limit the number of geographic areas investigated, but altogether also provides a global coverage when considering the ensemble of all regions. The Americas and Africa have been separated ~~in a~~ into northern and southern ~~section~~sections. In order to assemble the sites most affected by Saharan dust, the North Africa region has been extended ~~in the North~~ to the north beyond the continent itself.

Stations located in the south of Spain, Cyprus and Greece contribute to the regional time series in the region we are calling North Africa. The regions coordinates can be found in the supplementary info.

As seen in Figure 2, the regions do not have a similar coverage in terms of observations. North America and Europe have the highest concentrations of instruments monitoring aerosol trends.

- AERONET is the most important network in terms of number of instruments. More than 1000 observation points, with more or less long time series, are found across the globe. The highest density of instruments is in Europe and in the central part of North America (US). The lowest densities are found in southern Africa and Australia.
- Particulate Matter: ~~212~~ 227 instruments are used in this study and are spread mostly over Europe and North America.
- SO_4: Altogether 346 instruments have been operating, mostly in North America and Europe. A few stations are also located in Asia and North Africa.
- σ_{sp} and σ_{ap}: Combined for both parameters ~~circa~~ approximately 50 stations are spread over North America, Europe, North Africa and Asia. Due to time coverage issues (2005 is the first year where in situ optical data are available in the European time series), the data from 2000 up to the year 2018 were used to compute the regional time series of these two parameters.

3.1.2 Time series aggregation requirements

The regional time series are computed by combining, for each month, the valid data of all the stations in the corresponding region. In order to construct consistent and robust regional time series, some additional criteria are required to be met to provide a valid point (a station with valid measurements) going into the regional time series. Stations ~~having operated very shortly~~ with very short time series (e.g AERONET DRAGON campaign stations) are eliminated by requiring a minimum of 300 valid daily measurements in the whole period from 2000 to 2014, which reduces, as an illustration, the number of AERONET stations from ~~1010~~ 1015 to 437. A minimum of three valid ~~points (daily or monthly depending on the available resolution) is required per month~~ stations is required to be present in the overall regional time series to produce a valid point. In other words, if the available time resolution is daily, at least three stations need to provide valid data for a certain day in order to produce a valid regional mean for that day. The list of the station names contributing to the computation of the regional time series can be found in the supplementary info.

When all criteria are fulfilled for a given month in the regional time series, the median ~~,~~ and the first and third quartiles are computed from all valid data points available. The quartiles provide an indication of the intra-regional variability. An example of regional time-series ~~is~~ are shown in Figure 3 for AOD.

3.2 Trends calculation

3.2.1 Yearly, regional time series

For all of the parameters, the trends are computed based on the yearly averages of the regional time series. Using the yearly averages eliminates any issues caused by the seasonal cycles (observed for most of the aerosol parameters used in this study) during the calculation of the trend slope. In order to ensure the statistical robustness of these yearly averages, the time averaging is performed step-by-step with specific time constraints. By starting at the finest time resolution available in the data, monthly, seasonal and then yearly averages are computed when the following criteria are fulfilled:

- at least 5 days per month are available (when daily observations are available).
- at least 1 month with data per season is present ~~with data~~ (seasons defined as JFM, AMJ, JAS, OND).
- all 4 seasons are available for a given year.

These temporal constraints offer a reasonable compromise between the availability and robustness of the yearly statistics.

3.2.2 Trends computation

We use the same methodology as described by Aas et al. (2019) to derive the trends of the regional time series. The significance of the trends is tested with the Mann-Kendall test (Hamed and Rao, 1998). The related p-value is used to determine if the trend is significant or not within a confidence interval of 95%. The slope is calculated with the Theil-Sen estimator which is less sensitive to outliers than standard least-squares methods (Sen, 1968). At least least 7 valid yearly regional averages (50% of time coverage) are required in the regional time series for the computation of a slope.

An uncertainty is provided for each trend by combining the error of the slope calculation itself to the error of the residuals:

$$Uncertainty = \sqrt{\left(\frac{\Delta m}{y(2000)}\right)^2 + \left(\frac{m \cdot \Delta r}{y(2000)^2}\right)^2} \quad (1)$$

where Δm is the Theil-Sen estimator 95% confidence interval, $y(2000)$ is the value of the regression line at the year 2000, m is the value of the Theil-Sen slope and Δr is the averaged error on the residuals computed based on the difference between the linear trend and the yearly mean values of the regional time series.

The trend is provided as a relative trend (%/yr) with respect to the first year of the time period (2000).

3.3 Representativity of the trends

The number of available points used to compute the regional time series is not constant in time. For a given observation station, the number of points available might vary in time due to the nature of the measurements. For instance, classic ~~sun photometers~~ sunphotometers only measure in the daytime and in cloud free conditions. Due to seasonal daylight and cloud condition

variations, clear seasonal cycles are observed in the number of observations of AOD. The density of the different observation networks can also change with time. The early development of the different observation networks usually coincided with an increase in the number of observation stations. More recently, primarily for funding reasons, some networks have reduced the number of stations. This variation in the number of available measurements raises the question of time representativity for the computation of the trends.

Associated with this time representativity issue comes the space representativity issue. The data coverage is uneven across the different regions. Moreover, within a single region, the observation stations might be located in contrasting environments. Stations located in environments that are more urban, or rural, or mostly affected by natural particles, might have trends differing from the trend associated with the whole region.

Some studies have focused on the representativity of the observation stations by investigating the biases of different optical properties (Wang et al., 2018; Schutgens et al., 2017; Schutgens, 2019). The analysis here is dedicated to ~~characterise~~ characterising the representativity of the observation networks specifically for the purpose of computing the trends. These two perspectives on representativity might give different results, since a station associated with a bias ~~,~~ could still have a representative tendency in time. In order to evaluate the effect of the partial space and time sampling of the observations for the evaluation of the trends, two sensitivity studies, focusing on the time sampling and the space sampling, have been conducted using ~~model subsets of data~~ NorESM2 model data subsets. For each of these studies, the trends are computed for one reference (Ref) and one experiment (Exp) dataset, and compared with each other. The reference dataset corresponds to the model data co-located to the available observations while the experiment dataset uses all model points.

- Time representativity study
  - Ref_{time}: ~~Collocation~~ Model data collocated in space and time with available observations
  - Exp_{time}: ~~Collocation in space using complete~~ Model data collocated in space with available observations using the complete model time-series
- Space representativity study
  - Ref_{space}: ~~Collocation in space using complete~~ Model data collocated in space with available observations using the complete model time-series (=Exp_{time})
  - Exp_{space}: All of the model grid-points in ~~region using full~~ the region using the complete model time-series

The difference between the relative trends are computed for each parameter and region. ~~Those~~ In order to summarize the representativity, those differences are then converted into a score (%) by using a ~~normal distribution f described by a mean $\mu = 0$ and a standard deviation of $\sigma = 0.5$~~mapping function which has been defined based on a normal distribution. The choice of ~~these parameters~~ the parameters describing this function leads to a representativity score of 100% when there is no difference in the trends ~~of~~ computed for a reference and an experiment dataset, while a difference of 0.5~~%/yr~~ %/yr obtained with these two datasets would indicate a representativity score of 50%.

~~For a given parameter p and a region r, the representativity $Rep(p,r)$ is calculated as following~~

$$Rep_{space,time}(p,r) = f\left(\left|\tilde{t}_{Exp_{space,time}(p,r)} - \tilde{t}_{Ref_{space,time}(p,r)}\right|\right)$$

~~where $\tilde{t}$ is the relative trend of the corresponding dataset.~~

Finally, the total score is computed as the mean of the time and the space representativities.

An example of the calculation is presented in Figure 4 for AOD in Europe and North America. In both regions, the Ref_{time} dataset, corresponding to the available observations, reveals strong seasonal cycles when considering the number of points used to compute the regional time-series. These cycles are observed with most of the ~~sun photometer~~ sunphotometer datasets since the instruments only operate during daytime and cloud free conditions, and the amount of daylight and clouds varies with the season. Together with this seasonal cycle, one observes an increase in the number of points with time, which reflects the increasing number of stations over these two regions.

The trends in Europe show similar values for the time study, which means that the trend is not greatly affected by the variation of the available measurements in time. The difference is larger when considering all the grid-boxes of the domain, but the overall difference of the two studies corresponds to a representativity of ~~69~~76%. In North America, the difference in the three trends is larger, ~~outstanding is~~ particularly for the space study trend. This means that the trend obtained in the whole region is significantly different from the trend obtained when considering only the grid points where observation stations are located. It should be mentioned that the ocean grid-points are not filtered out when computing the trends over the whole domain. For this reason, the regions containing a greater proportion of ocean grid-points, where the trends are most likely to differ from those observed over land, will tend to have a lower spatial representativity, such as North America.

This representativity study illustrates that the partial coverage in time and space of the observations leads, in some cases, to artificial trends. The representativity scores are discussed for each parameter in the following section together with the trend estimate results.

4 Results

4.1 Trends in observations

This ~~sections~~ section presents the trends in the observations computed for the different parameters and over the predefined regions. In order to compare the trends observed for the set of nine aerosol parameters in a consistent manner, we focus on the relative trends, with the reference set to the year 2000, as the first year of the study period. The means for the year 2000, reported in Table 3, reveal a large inter-regional variability.

The AOD is more than three times higher in Asia (AOD=~~0.35~~0.37) than in North America and Australia (AOD=0.10). Intermediate AOD values are found in Europe and South Africa, while the second highest load is found in North Africa (AOD=0.26). In most regions, the AOD is largely dominated by its fine ~~fraction (AOD<1 μm~~mode fraction (AOD_f), but this is not the case

in North Africa (or Australia), where the persistent presence of desert dust makes the coarse mode (~~AOD>1μm~~$\mathrm{AOD_c}$) contribution to the total AOD similar in size to the fine mode contribution. This predominance of coarse particles is reflected in the AE values which exhibit lower values in North Africa (AE=~~0.72~~0.70) and Australia (AE=~~0.97~~1.00).

The PM observations are primarily available from Europe and North America. $\mathrm{PM_{10}}$ observations are also available in the North Africa region as defined in this analysis, but these stations are located in the northern part of the region, i.e., in southern Europe, which is less affected by the dust sources than the AERONET stations, which cover the whole region including the surrounding deserts. Both $\mathrm{PM_{10}}$ and $\mathrm{PM_{2.5}}$ are larger in Europe than in North America, with different relative proportions. In Europe, $\mathrm{PM_{2.5}}$ represent ~~75~~76% of the $\mathrm{PM_{10}}$, as compared to ~~on~~ only 57% in North America. This difference in the relative proportion of fine particles against coarse particles in Europe and North America may be due in part to our definition of regions. Putaud et al. (2010) presented a phenomenology of PM data in Europe showing coarse aerosol tended to be highest in southern Europe which in our study is part of the North Africa region. The discrepancy in the relative proportions of coarse and fine aerosol in Europe and North America may be exacerbated by both a decrease in North America of the fine particles concentration due to pollution mitigation strategies coupled with the growth of the coarse mass due to increasing contributions of natural and agricultural sources, particularly in the western half of the U.S. (Hand et al., 2019a).

$\mathrm{SO_4}$ means (surface mass concentrations) for the year 2000 ~~ranges~~ range between 1.45 and 2.98 $\mu g\,m^{-3}$ with the lowest value occurring in North America and the highest value in North Africa (sites in southern Europe). Similar means are found in Europe and Asia, around 2 $\mu g\,m^{-3}$, though one should bear in mind that there are relatively few sites in Asia and they are not located in the most polluted areas in China and India (Aas et al., 2019).

Analogous to the surface $\mathrm{PM_{10}}$ measurements, σ_{sp} is higher in Europe (~~33~~ 34 Mm^{-1}) than in North America (~~25~~ 23 Mm^{-1}). The same feature is found for σ_{ap} which also has higher values in Europe than North America.

The relative trends for the 2000-2014 period are shown in Figure 5. The heatmap is dominated by blue color, which indicates mostly negative trends, especially when considering the ~~extensive parameters~~parameters related to aerosol burden (i.e., the extensive parameters). Usually, the lowest p-values (<0.05) are associated with the lowest uncertainties not shown in the same figure though. The largest circles (highest significance of trend) are more confidently associated with a ~~rather certain~~ decrease/increase of the aerosol property in the time period 2000-2014 since the value of the trend is greater than the uncertainty. The uncertainties are presented in Figure 6. Among the 38 computed trends, 22 are associated with a representativity score higher than 50% and 24 are significant at a 95% confidence level.

- In Europe, both columnar and surface parameters reveal statistically significant decreases~~, with~~. With the exception of σ_{ap}, for which the ~~observed decrease is not significant. For this last parameter, the~~ associated uncertainty of the trend exceeds the trend itself~~. This large uncertainty is linked to the low data coverage in the earliest years. For the other parameters, the uncertainties are~~, the trends computed for other parameters are associated with uncertainties lower than the trends values. A decrease in AOD (-2.8%/yr) is found for both fine and coarse mode particles. This is consistent with the negative trends found at some individual stations in this region (Glantz et al., 2019). The fine mode is decreasing more than the coarse mode, which is consistent with the decrease observed for AE. The same shift in aerosol size is found at the surface since $\mathrm{PM_{2.5}}$ has decreased by a factor of two relative to $\mathrm{PM_{10}}$. These trends could result from the

mitigation measures ~~aiming~~ aimed at reduced anthropogenic aerosol emissions. This is more directly observed in the decrease of SO_4 (-1.5%/yr). We find a somewhat lower trend than what was reported in Aas et al. (2019) (-2.67%/yr), but that could be explained by the differences in the methodology (trends computed from the regional time series, in this study, against a statistical average of the trends computed at the individual stations) and/or the definition of the region. The stations in the Mediterranean Basin, where a larger decrease is found (-4.3%/yr), are attributed to the North African region in this study.

The representativity study reveals that the observed trends are actually representative for the whole period and region for all of the parameters~~, except for σ_{sp} and σ_{ap} due to the lack of observations in the earliest period~~. A good agreement is found with the trends obtained at individual stations ~~and reported by ?, which reports on~~ reported by Collaud Coen et al. (2020), who found decreases of -2.92%/yr for σ_{sp} and -4.2%/yr for σ_{ap}, as compared to -2.5%/yr and -2.0%/yr in this study.

– In North America, similar trends are found for the columnar properties as were found for Europe. AOD is decreasing at a rate of 1.3%/yr, a 55% percent smaller trend than observed in Europe, but the North America reference value in 2000 is 40% lower than the reference value in Europe. One can note that the representativity scores are higher for AE than for AOD, while these two parameters have the same amount of data. This means that the trends ~~are probably smoother,~~ in AE are probably more homogeneous in space and time, ~~when comparing AE with AOD,~~ which makes the same amount of available observations more representative in the case of AE. The decreases observed for both $PM_{2.5}$ (~~-2.1~~-2.0%/yr) and PM_{10} ~~-1.6~~-1.2%/yr are significant and in the same range of values ~~than~~ as the trends found in Europe. However, the actual trends for PM_{10} and $PM_{2.5}$ are probably somewhat ~~higher~~ more negative than found here. The possible bias is caused by increased relative humidity during weighing, resulting in more particle bound water and thus higher mass, after the relocation of the laboratory in 2011. Hand et al. (2019b) reported that the decrease in $PM_{2.5}$ from 2005 through 2016 was -2.6%/yr, while it was -3.9%/yr for the reconstructed fine mass correcting for the possible bias in the measurements. SO_4 decreases by about 3%/yr, which is twice as large as the decrease observed in Europe, where the reference value is however larger than in North America. The sulfate trend is similar to the trend reported by Aas et al. (2019) in this region (-3.15%/yr). The regional time series ~~are~~ extend farther back in time for σ_{sp} and σ_{ap} in North America than in Europe. However, no significant trends are found for these data sets. ~~? finds~~ This is in contrast to Collaud Coen et al. (2020) which found a large decrease for σ_{sp} (~~-2.57~~-2.66%/yr)~~which is not found in this study , when using regional~~. Our study used regionally averaged time series to calculate the trend rather than regionally averaged trends as was done by ~~?~~Collaud Coen et al. (2020). This probably illustrates the difference of methodology which consists of computing the mean of station trends in one case, and the trend of a regional time series in the other case, especially when only few measurements are available. However, as shown by the representativity study (5), the non-significant increase of +0.0%/yr found, in this study, with the observations is similar to the trend derived over the whole region and using complete time series of the NorESM2 model data. Similar values are found in this study and by ~~?~~ Collaud Coen et al. (2020) for σ_{ap} (-1.85%/yr) ~~despite the fact~~ although the trend is, here, not significant. The IMPROVE

network also measures filter absorption using a Hybrid Integrating Plate and Sphere (HIPS) system (White et al., 2016). These data are not included in this study, but White et al. (2016) reports a significant decrease (-2.7%/y) in the light absorption coefficients from 2005 to 2015.

- All of the columnar properties show significant decreasing trends in South America~~. All of the trends are significant~~, except for ~~AOD>1μm~~AOD_c. As shown in the regional time series in Figure 3, the observed decrease in AOD coincides with a global diminution of the intensity of the seasonal peaks happening around September and resulting from the Amazonian forest fires (Aragão et al., 2018). These peaks are highly variable from year to year and could greatly affect the trend when considering another time period. With a rate of -2.0%/yr, the largest decrease of AE is found in this region. While no significant trend is found for ~~AOD>1μm~~AOD_c, the tendency towards increasing coarse particles is probably due to the production of local dust as a result of the increasing deforestation (Werth and Avissar, 2002; Betts et al., 2008).

- In North Africa, while significant decreases are found for all AOD parameters, an increase of AE (+~~1.1~~1.2%/yr) is observed, which indicates an increase in the proportion of fine particles with time. This is consistent when considering the AOD of the fine and coarse modes, which reveal a larger decrease for ~~AOD>1μm~~AOD_c. Chin et al. (2014) also found a decrease in dust in the Sahara/Sahel in the time period 1980-2009 due to reduced 10m-wind speed, possibly caused by an increase in sea surface temperature (SST) in the North Atlantic.

- AE is also increasing in Asia as a combination of a (not significant) increase in ~~AOD<1μm~~ AOD_f and a significant ~~increase in AOD>1μm~~decrease in AOD_c. The increase in AE is likely tied to increases in anthropogenic emissions which are associated with fine mode aerosol. This result is consistent with the trend reported by Yoon et al. (2012) at some individual stations. At the same time, we observe an increase of SO_4 of 3.8%/yr, which is consistent with the trend reported in Aas et al. (2019). This increase is associated with a large uncertainty ($\pm$4%/yr) due to a drop in the already small number of stations available in the region, especially between 2010 and 2012. Indeed, with a maximum of 12 stations, a few stations missing can greatly affect the computation of the regional time series. This is reflected by the representativity study which reveals a score lower than 40% for this parameter.

- No significant trends ~~could be~~ are found in Australia, although the representativity is greater than 50% for ~~AOD, AOD<1μm and AE~~AOD_f.

This multi-parameter trend analysis reveals a decrease in most of the parameters relating to aerosol burden~~(extensive parameters)~~, both in the total column and at the surface level. In Asia, the trends in ~~AOD<1μm~~AOD_f, AE and SO_4 suggest an increase in the proportion of the finer particles. While differences might be expected when comparing regional trends with trends computed at individual stations, the trends are usually consistent with those previously reported in the literature. de Meij et al. (2012) focused on regional AOD trends in the 2000-2009 period; despite the differences in the study periods and the methodologies involved, ~~consistent trends can be found in~~ trends consistent with those found in this study are found for most of the regions~~with the trends obtained in this study~~.

4.2 Evaluation of the models trends against observations

In order to evaluate the trends from the models, the regional time series have been computed with the model output collocated in space and time to the available observations at the station level. The model trends are computed in a similar manner to the trends for the observation datasets. ~~However, for the few models providing output every 5 years (in addition to 2014), the minimum required number of points has been reduced from 7 to 4, so the trends can be computed using the years 2000, 2005, 2010 and 2014.~~ The results, shown in Figure 6, reveal ~~different performances~~ (a) the differing abilities of the various models ~~, for the reproduction~~ to reproduce of the observed trends, ~~depending~~ and (b) the model performance depends on the parameters and the regions.

- AOD: the models show trends ~~in the same direction as the observations~~ with the same sign as the observed trends over all the regions except in Asia, where the associated uncertainties are ~~, however,~~ usually larger than the trend values. Some differences among the three model groups are observed when investigating the different regions:
  - EUROPE: all the groups underestimate the observed decrease in AOD. With an average decrease of -1.0%/yr, the CMIP6 models exhibit the largest underestimation, while the best performance is obtained with CAMS-Rean (~~-2.1~~-2.0%/yr). The AP3 models' trends range from -1.3%/yr to ~~-2.0~~-2.3%/yr.
  - NAMERICA: in contrast to the results for EUROPE, on average, all of the models overestimate the observed AOD decrease in NAMERICA even though two models of the AP3 group simulate lower trends than are found for observations. The consistency in the trends is very high within the CMIP6 group over this region.
  - SAMERICA: ~~CAMS Rean~~ CAMS-Rean slightly overestimates the observed AOD decrease while ~~all the models of~~ almost all of the models in the two other groups underestimate this decrease. A few of the models simulate positive trends, but these are associated with large uncertainties.
  - NAFRICA: all the models capture the observed decreasing AOD tendency. With a trend of -3.0%/yr, CAMS-Rean is the closest to the observed trend (~~-2.7~~-2.8%/yr). AP3 and CMIP6 multi-model trend averages are ~~-2.0%/yr and -2.2~~both equal o -2.1%/yr~~, respectively~~.
  - ASIA: A large inter-model variability is found in this region where the uncertainty is also ~~significant~~important. The means of the trends of each group ~~range from -0.2%/yr to +0.2~~are close to 0%/yr.
- ~~AOD<1μm~~AOD_f: usually, the same patterns are found as for AOD. The models that underestimated the AOD underestimate ~~AOD<1μm~~ AOD_f and vice versa. For ~~AOD<1μm~~ AOD_f and the following parameters, only NorESM2 provides data for the CMIP6 group.
  - ~~in~~ EUROPE: the underestimation of the decrease in AOD_f captured by the models is larger than the underestimation of AOD.

- ASIA: ~~an increase, associated with large uncertainties is found in both models of the AP3 group (+1.3%/yr) and observations (+0.8%/yr)~~as for AOD, the trends are associated to large uncertainties and show a large inter-model variability.

- ~~AOD>1μm~~AOD_c: the performance of the models is not as good as for ~~AOD<1μm~~AOD_f. This is also observed when evaluating the models for a single year (Gliß et al., 2020). The inter model variability is also higher since some models simulate ~~AOD>1μm~~ AOD_c trends in opposite directions in some regions.

  - EUROPE: while the observations exhibit a significant decrease, CAMS-Rean and all of the AP3 models exhibit increasing values for ~~AOD>1μm~~AOD_c. NorESM2 from CMIP6 ~~simulate~~ simulates a decrease consistent with the observations.

  - SAMERICA: All of the models simulate large increases, from +~~4.3~~1.2%/yr up to +~~14.6~~8.5%/yr which are not visible in the observations (~~-0.1~~+0.1%/yr).

  - NAFRICA: the models reproduce the observed decrease of ~~3.3~~3.1%/yr to some extent (from ~~-0.7~~-0.6%/yr to ~~-2.5~~-2.7%/yr). The fact that some models with fixed SST (e.g ECHAM-HAM) reproduce this decrease does not support the hypothesis of the SST changes ~~. The decrease in dust could be caused by increased wet scavenging of dust after coating with anthropogenic sulfate aerosols. The production of high levels of readily soluble materials on the dust surface makes dust aerosols effective cloud condensation nuclei~~ impacting dust emissions. (Fan et al., 2004; Bauer and Koch, 2005; Bauer et al., 2007; Neubauer et al., 2019).

  - ASIA: CAMS-Rean captures ~~the same~~ a similar negative trend as computed with the observations dataset. ~~Like for AOD<1μm~~As with AOD_f, no certain trend can be identified in this region with the NorESM2 CMIP6 model.

- AE: the trends are usually smaller than for AOD in the respective regions~~, meaning~~ . This can mean that the amount of the particles is more subject to variations than the size (type) of these particles but could also illustrate that AE is less sensitive to the change in a relative sense. This feature is visible with both observations and models.

  - EUROPE and NAMERICA: one model ~~of~~ in the AP3 group (ECHAM-HAM) simulates a significant positive trend in AE while negative tendencies are found in the observation and with the other models.

  - SAMERICA: all of the models simulate negative AE trends, most of them significant, in agreement with the observations. CAMS-Rean and the AP3 models tend to underestimate the decrease, while ~~the~~ NorESM2 CMIP6 model tends to overestimate it.

  - NAFRICA: CAMS-Rean does an excellent job of reproducing the observed AE increase (+1.3%/yr versus +~~1.1~~1.2%/yr). The significant trends of the AP3 models range from ~~-0.5~~-0.4%/yr to +2.0%/yr. The increase of AE supports the theory of enhanced scavenging of dust by anthropogenic aerosols.

  - ASIA: the AP3 models and the NorESM2 CMIP6 model exhibit significant positive trends, which is also the case for the observations. CAMS-Rean does not capture any significant trend in this region.

- $PM_{2.5}$: Almost all the models simulate significant decreases over Europe and North America, in good agreement with the observations. The CMIP6 model performs however better in North America, while it underestimates the extent of the decrease in Europe. Further analysis reveals that, despite the fact that it does a good job reproducing the $PM_{2.5}$ trend in North America, CAMS-Rean exhibits a large positive bias in ~~North America~~ this region when considering the absolute values (+100%). In North Africa, both CAMS-Rean and AP3 models capture the significant decrease seen in the observations.

- PM_{10}: In North Africa, only CAMS-Rean reproduces the observed significant decrease. Positive trends are found for all the models ~~of~~ in the AP3 (except GEOS) and CMIP6 groups. As for $PM_{2.5}$, ~~NorESM~~ NorESM2 has better performance in North America. CAMS-Rean produces a trend twice as high as the observed trends both over Europe and North America.

- SO_4: The AP3 and CMIP6 models perform quite well for the SO_4 surface concentration. The magnitude of the model trends is however higher than the observed trends in all the regions except North Africa.

- σ_{sp} and σ_{ap}: as mentioned in the previous section, the observations trends have been computed for these two parameters using data until 2018. The two models providing output for these parameters are NorESM2 and ~~SPRINTARS~~OsloCTM3. NorESM2 provides data until 2014, so the NorESM2 trends correspond to the period [2000-2014], while ~~SPRINTARS~~ OsloCTM3 provides data until ~~2018 and thus covers the whole observation period~~ 2017 and the respective trends correspond to [~~2000-2018~~2000-2017].

  - EUROPE: a significant decrease is found in the observations for both σ_{sp} and σ_{ap} but this is not captured by the models ~~where positive trends are found, although~~ for which the calculated trends are associated with large uncertainties.

  - NAMERICA: A significant decrease of -1.3%/yr is found with both NorESM2 and OsloCTM3 for σ_{sp} which is not seen in the observations. For σ_{ap}, ~~NorESM2 captures~~ the two models capture a similar trend as derived from the observations ~~, while SPRINTARS does not~~(-1.5%/yr).

This model trends evaluation reveals some ~~key-points~~key points. First, CAMS-Rean, which assimilates AOD, performs the best for capturing the trends of this parameter. Second, a large inter-model variability is generally found over Asia, where the observed trends are also the most uncertain. Considering the total column, the models usually perform rather well for AOD, ~~AOD<1μm~~AOD_f, and AE, but show lower skill for ~~AOD>1μm~~AOD_c. At ground level, the models perform well for both SO_4 concentration and PM. The trends in σ_{sp} and σ_{ap} computed from regional time series are associated with large uncertainties due to the limited number of stations. This is exacerbated by the fact that data was only available from two models for these parameters.

4.3 Trends in models

4.3.1 Global trends

As discussed previously, the regional trends found are probably not always representative of the trends in the extended regions and over the whole study period. The reasons are the partial spatial and temporal coverage of the ground based observations. Moreover, the observation stations are obviously located on land. This does not allow for a depiction of a global aerosol trends and is unfortunate as sea salt particles are among the most predominant aerosols on Earth (Schulz et al., 2004).

Unlike observations, models provide data at a global scale and for the entire study period. The completeness of these model datasets offers the opportunity to derive global aerosol trends. In order to provide an assessment of the aerosol trends at a global scale, we present, in this section, the trends computed with ~~the~~ NorESM2 ~~data~~ (CMIP6 group)~~using all grid boxes~~, which provides data for all of the nine parameters considered in this study. The calculation of the global trend is made by averaging the absolute trends computed at each grid-point of the model and using all timestamps in the study period. In order to provide a relative trend, this absolute trend is normalized to the global average of the considered parameter for the year 2000. The global trends are reported for the nine aerosol parameters in Table 4. The global maps, shown in Figure 7, enable investigation of the spatial variability of these trends.

While the observed trends of the three AOD parameters show a decrease in most of the regions of the world, the global AOD trend is actually positive (+0.2%/yr). This global increase is also found with other models. Averages of the models from the CAMS-Rean and the AP3 groups simulate global trends of about +0.2%/yr and +0.3%/yr respectively. Within the CMIP6 group, IPSL and CESM2 also exhibit positive trends (+0.7%/yr and +0.3%/yr), consistent with NorESM2, while CanESM simulates a negative trend (-0.8%/yr). The relative increase of 0.2%/yr found with NorESM2 corresponds to an absolute rate of +0.0028/decade, which is in excellent agreement with the global trend (over the oceans) of +0.003/decade reported by Zhang and Reid (2010) using MODIS data. The increase of AOD is observed to be larger for the fine fraction, with an increase of about +0.6%/yr, as compared to +0.1%/yr for ~~AOD>1μm~~$\mathrm{AOD_c}$. As seen in Figure 7, similar geographical patterns are found for the three AODs: increase in South-Africa and East-Asia and decrease in Europe and in the US. The increasing AOD observed in Canada is dominated by an increase of ~~AOD<1μm~~ $\mathrm{AOD_f}$ in this region. The ~~important~~ prominent increase of AOD in Indonesia seems to be linked to a large increase of ~~AOD>1μm. Over the~~ $\mathrm{AOD_c}$. The tropical Pacific Ocean, ~~one region~~ off the west coast of South America, has significant positive modelled trends in both AOD and ~~AOD<1μm~~$\mathrm{AOD_f}$. Almost no significant trend is found south of 60°S.

The model also simulates an increase for AE on a global scale, with a rate of +0.3%/yr. This suggests a shift towards smaller particles. The largest increases are found over Canada, Greenland, Siberia and the Pacific Ocean. There are some distinct outliers around 60°S. In the Atlantic, we find a decrease of AE off the east coast of the US, which is consistent with the decrease of ~~AOD<1μm~~ $\mathrm{AOD_f}$ in the same region.

The trends in both $\mathrm{PM_{2.5}}$ and $\mathrm{PM_{10}}$ exhibit similar geographical features as are observed for AOD. In addition, one finds large and significantly increasing trends in the high Arctic that could be explained by a change in the air mass circulation pattern, or by the increase of open sea, which might contribute to a higher production of sea salt aerosols (Willis et al., 2018; Abbatt et al., 20

. The global averages show that $PM_{2.5}$ is increasing faster than PM_{10} (+0.2%/yr vs. +0.1%/yr), which is consistent with the increasing AE, suggesting a relatively higher fraction of fine particles with time.

The surface SO_4 concentration trends map reveals two large contrasting regions. Significant decreases are found over North America and Europe, while significant increases are found over southern and eastern Asia and southern to central parts of Africa. This illustrates the shift of polluting activities from the developed countries to the developing countries during the last two decades. With an overall increase of +0.4%/yr, the global trend is positive.

The σ_{sp} trends are very similar to those observed for both $PM_{2.5}$ and PM_{10}. The same geographical patterns are found, as well as the global average trend which amounts to an increase of 0.2 %/yr over the study period.

σ_{ap} reveals increasing tendencies over most of the grid-boxes of the model, except in Europe, the eastern part of the US, and Australia. This explains why ~~the~~ a large positive global trend is obtained for this parameter, with an average of +1.5%/yr. Further analysis shows a good spatial correlation with the ~~BC~~ black carbon OD (Optical Depth) that exhibits a strong global positive trend of +2.3%/yr, as discussed below.

Table 4 also contains the trends computed for the different aerosol parameters when combining only the grid-points where an observation station is located, whether measurements are available or not. Significant differences in 'global' trends can be found when observations are not provided over some regions. This is most obvious for SO_4 for which the observation stations are located mostly in Europe and North America and exhibit decreasing values, while only a few stations are located in the regions associated with increasing values. In this case, the computation of the trends by considering only observation station grid-boxes leads to a global decrease of ~~-1.1~~-3.9%/yr while consideration of all of the grid-boxes of the model leads to a global increase of +0.4%/yr.

4.3.2 Contribution of main aerosol species to the AOD trends

The averaged global trend computed by NorESM2 indicates an increase of AOD in the 2000-2014 period with a rate of about 0.2%/yr. The trends in AE, ~~AOD<1μm and AOD>1μm indicate that the fine~~ AOD_f and AOD_c indicate that fine mode particles are primarily responsible for this increase in the atmospheric column.

In this section, we investigate the trends of the major aerosol species simulated by NorESM2. For that purpose, the absolute trends of the individual contribution of these species to the AOD were computed, as well as the trends in the loads and the emissions. The trends of OD and loads are shown in Figure 8. In this version, NorESM2 simulates a large proportion of sea salt. This is the result of a model tuning used for reaching climate equilibrium. While the model attributes too much OD to SS, the trends should not be affected by this tuning.

The relative increase of AOD of +0.2%/yr corresponds to an absolute increase in AOD of +3.1e^{-4}/yr. This positive trend is dominated by an increase in the species specific ODs of the organic aerosols (OA), SO_4 and black carbon (BC), which are responsible for an increase of the OD of about +2.0e^{-4}/yr, +0.7e^{-4}/yr and +0.4e^{-4}/yr, respectively. The relative OD trends give a different ranking since the highest increase is found for BC (+2.5%/yr), followed by OA (+0.5 %/yr). On average, the trends for dust and sea salt OD are slightly negative (-0.1e^{-4}/yr). Note - these species trends include any associated water which can change as function of relative humidity.

The trends in OD do not necessarily represent the trends in the aerosol loads ~~, since the~~ which do not include associated water. The different species have different global mass extinction coefficients (~~from this study~~ calculated in this study as OD/load, dust: 1.8 m^2g^{-1}, SS: 4.3 m^2g^{-1}, OA: 5.6 m^2g^{-1}, SO_4: 5.3 m^2g^{-1}, BC: 7.6 m^2g^{-1}). For sea salt, opposite trends are ~~even~~ observed for the sea salt OD (positive trend) and the sea salt load (negative trend). The analysis of the global maps (not shown in this study) reveals that the largest increases of the sea salt load happen in Indonesia and near the North Pole and result in a relatively larger increase of OD in these areas. ~~This effect relates~~ These localized increases in sea salt OD drive the global sea salt OD trend and are due, at least in part, to the higher relative humidity at these latitudes which makes the sea salt, which is very hygroscopic, more efficient at light extinction.

5 Conclusions

The main findings of this multi parameter trends analysis ~~can be listed as follows~~are listed below:

- The observations exhibit mostly negative trends regarding the extensive parameters in the different regions of the world. Significant decreases are found in Europe, North America, South America, North Africa and Asia. In Asia, AE is increasing in time and is consistent with increases in ~~AOD<1μm~~ AOD_f and SO_4, which reflects the regional increase of the anthropogenic aerosols in that region in the overall study period from 2000 to 2014.

- Some observation networks allow for the derivation of representative trends over the whole study period. In other cases, the ~~partial~~ limited temporal and spatial coverage of the observations can induce artificial and/or highly uncertain trends when using regional time series. ~~60% of the 37 trend values computed in this study are significant at the 90% level~~Among the 38 computed trends with observation data, 22 are considered as representative of the actual trends occurring in the whole region and study period.

- The models tend to capture observed AOD, AE, SO_4 and PM trends but show larger discrepancies regarding ~~AOD>1μm~~AOD_c. The smaller amount of data available for establishing σ_{sp} and σ_{ap} trends makes the validation of the modeled trends more uncertain.

- The ~~global trends computed using model data give a different picture than the trends obtained when using only ground-based observations. The~~ rather good agreement of the trends, across different aerosol parameters between models and observations, when co-locating them in time and space, implies that global model trends, including those in poorly monitored regions, are likely correct.

- The global trends computed with ~~the model data show mostly positive trends for all~~ NorESM2 (CMIP6 group) model data give a different picture than the trends obtained when using only ground-based observations. Global positive trends are found for all of the parameters related to aerosol loading. The trends in AOD are dominated by the increase of the fine particles both in the column and at the surface. This tendency toward finer particles is consistent with the positive trend in AE. This increase appears to be dominated by ~~the organic aerosols~~organic aerosol, for which the emissions have

increased in the study period, and by ~~the~~ SO_4 ~~whose emissions~~ aerosol whose sources were shifted from Europe and North America to Africa and East Asia where a global positive SO_4 trend is found.

Some elements were not considered in this study which could be investigated in order to complete the aerosol trends picture:

- Some regions are associated with strong seasonal cycles. In South America, the regional time series shows high peaks in AOD, associated with forest fires in the late summer, whose intensity greatly varies from year to year. In Africa, a strong seasonal contrast is also found due to the transport of desert dust at altitude in the summer months (Mortier et al., 2016; Ogunjobi et al., 2008). The computation of the seasonal trends would allow characterization of the tendencies in such extreme or synoptic aerosol events.

- This study shows that the trends computed from the ground-based observations networks are not representative of the global aerosol trends due to the inhomogeneities in data spatial coverage. The satellites providing a global Earth observation could be utilized for the evaluation of the model trends in the regions lacking observations and over the oceans (Hsu et al., 2012; Zhang and Reid, 2010).

- The trends in the meteorological parameters could be investigated in parallel with the aerosol trends because they affect the ~~aerosols~~ aerosol life cycle and their optical properties (Che et al., 2019). Hypothetical trends in wind velocity could produce trends in the loads of sea salt and dust and, as seen in the last section, trends in OD could also be enhanced by relative humidity changes. Changes in temperature could impact the magnitude of the biogenic emissions. Indeed, increasing temperatures, associated with changes in land use and high atmospheric ~~CO2~~ CO_2 concentrations have been shown to lead to an increase of the BVOC emissions (Peñuelas and Staudt, 2010). Finally, trends in precipitation that are responsible for aerosol wet scavenging would directly ~~produce~~ impact trends in aerosol loads.

- Several studies have linked the trends in anthropogenic aerosols to radiative forcing variations while investigating sources of global dimming and brightening ~~Streets et al. (2006); Norris and Wild (2007)~~(Streets et al., 2006; Norris and Wild, 2007) . It could be of interest to evaluate how much the modeled trends deviations, as compared to the observations, are affecting the calculation of the radiative forcing, in the different regions of the world, and at a global scale.

- While the mountain sites were excluded from this study, it could be of interest to investigate the trends at higher altitude (which may be related to changes in long range transport) by including the in situ and remote sensing stations higher than 1000 m (Jungfraujoch, Mauna Loa Observatory, ~~..~~etc.).

- ~~While assembling the dataset for this analysis, it appeared that more observations (σ_{ap} in the US) could be utilized. Due to time limitations, these data could not be integrated in the study, but could be considered in the future to enrich both databases.In addition, more models and diagnostics from the AeroCom and CMIP6 ensemble should be added into the analysis when data become available to eventually confirm the regional and global trends for all parameters~~Similarly, it may also be of interest to look at trends in smaller regions (e.g., split North America into several sub-regions which are more internally consistent in terms of climate and environment than the large NAMERICA region defined here, or consider southern Europe as its own region rather than combining it with the North Africa region as was done here).

Code availability. The observation and model data were read and collocated with the pyaerocom python library (https://github.com/metno/pyaerocom, version 0.10.0).

Author contributions. A. M. has coordinated the study, has been responsible for the statistical calculation and analysis and wrote the paper, J. G. is the main developer of the pyaerocom library, M. S. has provided feedback on the methods and the manuscript, W. A, E. A, J. H, and P. L have provided in situ data, contributed to the observation dataset section writing and provided feedback on the manuscript, B. H. is the PI of AERONET, H. B., M. C., P. G., Z. H., Z. K., A. K., T. L., G. M., D. N., D. O., K. S., T. T., and S. T. have provided model output data and feedback on the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. Data providers from all the regional and global networks are greatly acknowledged for sharing and submitting their data to be used. DN acknowledges funding from the European Union's Horizon 2020 research and innovation programme project FORCeS under grant agreement No 821205. The ECHAM-HAMMOZ model is developed by a consortium composed of ETH Zurich, Max Planck Institut für Meteorologie, Forschungszentrum Jülich, University of Oxford, the Finnish Meteorological Institute and the Leibniz Institute for Tropospheric Research, and managed by the Center for Climate Systems Modeling (C2SM) at ETH Zurich.

The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. All simulations were carried out on the Cheyenne high-performance computing platform https://www2.cisl.ucar.edu/user-support/acknowledging-ncarcisl, and are available to the community via the Earth System Grid.

References

Aas, W., Mortier, A., Bowersox, V., Cherian, R., Faluvegi, G., Fagerli, H., Hand, J., Klimont, Z., Galy-Lacaux, C., Lehmann, C. M., et al.: Global and regional trends of atmospheric sulfur, Scientific reports, 9, 953, https://doi.org/10.1038/s41598-018-37304-0, 2019.

Abbatt, J. P. D., Leaitch, W. R., Aliabadi, A. A., Bertram, A. K., Blanchet, J.-P., Boivin-Rioux, A., Bozem, H., Burkart, J., Chang, R. Y. W., Charette, J., Chaubey, J. P., Christensen, R. J., Cirisan, A., Collins, D. B., Croft, B., Dionne, J., Evans, G. J., Fletcher, C. G., Galí, M., Ghahremaninezhad, R., Girard, E., Gong, W., Gosselin, M., Gourdal, M., Hanna, S. J., Hayashida, H., Herber, A. B., Hesaraki, S., Hoor, P., Huang, L., Hussherr, R., Irish, V. E., Keita, S. A., Kodros, J. K., Köllner, F., Kolonjari, F., Kunkel, D., Ladino, L. A., Law, K., Levasseur, M., Libois, Q., Liggio, J., Lizotte, M., Macdonald, K. M., Mahmood, R., Martin, R. V., Mason, R. H., Miller, L. A., Moravek, A., Mortenson, E., Mungall, E. L., Murphy, J. G., Namazi, M., Norman, A.-L., O'Neill, N. T., Pierce, J. R., Russell, L. M., Schneider, J., Schulz, H., Sharma, S., Si, M., Staebler, R. M., Steiner, N. S., Thomas, J. L., von Salzen, K., Wentzell, J. J. B., Willis, M. D., Wentworth, G. R., Xu, J.-W., and Yakobi-Hancock, J. D.: Overview paper: New insights into aerosol and climate in the Arctic, Atmospheric Chemistry and Physics, 19, 2527–2560, https://doi.org/10.5194/acp-19-2527-2019, 2019.

Aragão, L. E., Anderson, L. O., Fonseca, M. G., Rosan, T. M., Vedovato, L. B., Wagner, F. H., Silva, C. V., Junior, C. H. S., Arai, E., Aguiar, A. P., et al.: 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions, Nature communications, 9, 536, https://doi.org/10.1038/s41467-017-02771-y, 2018.

Barreto, Á., Cuevas Agulló, E., Granados-Muñoz, M. J., Alados-Arboledas, L., Romero Campos, P. M., Gröbner, J., Kouremeti, N., Almansa Rodríguez, A. F., Stone, T., Toledano, C., et al.: The new sun-sky-lunar Cimel CE318-T multiband photometer-a comprehensive performance evaluation, https://doi.org/10.5194/amt-9-631-2016, 2016.

Bauer, S. and Koch, D.: Impact of heterogeneous sulfate formation at mineral dust surfaces on aerosol loads and radiative forcing in the Goddard Institute for Space Studies general circulation model, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2005JD005870, 2005.

Bauer, S., Mishchenko, M., Lacis, A., Zhang, S., Perlwitz, J., and Metzger, S.: Do sulfate and nitrate coatings on mineral dust have important effects on radiative properties and climate modeling?, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2005JD006977, 2007.

Bell, M. L., Davis, D. L., and Fletcher, T.: A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution., Environmental health perspectives, 112, 6–8, https://doi.org/doi.org/10.1289/ehp.6539, 2004.

Betts, R., Sanderson, M., and Woodward, S.: Effects of large-scale Amazon forest degradation on climate and air quality through fluxes of carbon dioxide, water, energy, mineral dust and isoprene, Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 1873–1880, https://doi.org/10.1098/rstb.2007.0027, 2008.

Bian, H., Chin, M., Hauglustaine, D. A., Schulz, M., Myhre, G., Bauer, S. E., Lund, M. T., Karydis, V. A., Kucsera, T. L., Pan, X., et al.: Investigation of global particulate nitrate from the AeroCom phase III experiment, Atmospheric Chemistry and Physics, 17, 12 911, https://doi.org/10.5194/acp-17-12911-2017, 2017.

Bryner, G. C.: Blue skies, green politics: The clean air act of 1990, 1995.

Burnett, R. T., Pope III, C. A., Ezzati, M., Olives, C., Lim, S. S., Mehta, S., Shin, H. H., Singh, G., Hubbell, B., Brauer, M., et al.: An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure, Environmental health perspectives, 122, 397–403, https://doi.org/10.1289/ehp.1307049, 2014.

Che, H., Gui, K., Xia, X., Wang, Y., Holben, B. N., Goloub, P., Cuevas-Agulló, E., Wang, H., Zheng, Y., Zhao, H., et al.: Large contribution of meteorological factors to inter-decadal changes in regional aerosol optical depth, Atmospheric Chemistry and Physics, 19, 10 497–10 523, https://doi.org/10.5194/acp-19-10497-2019, 2019.

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, https://doi.org/10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2, 2002.

Chin, M., Diehl, T., Tan, Q., Prospero, J., Kahn, R., Remer, L., Yu, H., Sayer, A., Bian, H., Geogdzhayev, I., et al.: Multi-decadal aerosol variations from 1980 to 2009: a perspective from observations and a global model, https://doi.org/10.5194/acp-14-3657-2014, 2014.

Colarco, P., da Silva, A., Chin, M., and Diehl, T.: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD012820, 2010.

Collaud Coen, M., Andrews, E., Alastuey, A., Arsov, T. P., Backman, J., Brem, B. T., Bukowiecki, N., Couret, C., Eleftheriadis, K., Flentje, H., Fiebig, M., Gysel-Beer, M., Hand, J. L., Hoffer, A., Hooda, R., Hueglin, C., Joubert, W., Keywood, M., Kim, J. E., Kim, S.-W., Labuschagne, C., Lin, N.-H., Lin, Y., Lund Myhre, C., Luoma, K., Lyamani, H., Marinoni, A., Mayol-Bracero, O. L., Mihalopoulos, N., Pandolfi, M., Prats, N., Prenni, A. J., Putaud, J.-P., Ries, L., Reisen, F., Sellegri, K., Sharma, S., Sheridan, P., Sherman, J. P., Sun, J., Titos, G., Torres, E., Tuch, T., Weller, R., Wiedensohler, A., Zieger, P., and Laj, P.: Multidecadal trend analysis of aerosol radiative properties at a global scale, Atmospheric Chemistry and Physics Discussions, 2020, 1–54, https://doi.org/10.5194/acp-2019-1174, 2020.

Danabasoglu, G., Lamarque, J.-F., B. J., Bailey, D. A., DuVivier, A. K., and et al.: The Community Earth System Model version 2 (CESM2), Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2019MS001916, submitted.

de Meij, A., Pozzer, A., and Lelieveld, J.: Trend analysis in aerosol optical depths and pollutant emission estimates between 2000 and 2009, Atmospheric Environment, 51, 75 – 85, https://doi.org/10.1016/j.atmosenv.2012.01.059, 2012.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development (Online), 9, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fan, S.-M., Horowitz, L. W., Levy, H., and Moxim, W. J.: Impact of air pollution on wet deposition of mineral dust aerosols, Geophysical research letters, 31, https://doi.org/10.1029/2003GL018501, 2004.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., et al.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmospheric Measurement Techniques, 12, 169–209, https://doi.org/10.5194/amt-12-169-2019, 2019.

Glantz, P., Freud, E., Johansson, C., Noone, K. J., and Tesche, M.: Trends in MODIS and AERONET derived aerosol optical thickness over Northern Europe, Tellus B: Chemical and Physical Meteorology, 71, 1–20, https://doi.org/10.1080/16000889.2018.1554414, 2019.

Gliß, J., Mortier, A., Schulz, M., Andrews, E., Balkanski, Y., Bauer, S. E., Benedictow, A. M. K., Bian, H., Checa-Garcia, R., Chin, M., Ginoux, P., Griesfeller, J. J., Heckel, A., Kipling, Z., Kirkevåg, A., Kokkola, H., Laj, P., Le Sager, P., Lund, M. T., Lund Myhre, C., Matsui, H., Myhre, G., Neubauer, D., van Noije, T., North, P., Olivié, D. J. L., Sogacheva, L., Takemura, T., Tsigaridis, K., and Tsyro, S. G.: Multi-model evaluation of aerosol optical properties in the AeroCom phase III Control experiment, using ground and space based columnar observations from AERONET, MODIS, AATSR and a merged satellite product as well as surface in-situ observations from GAW sites, Atmospheric Chemistry and Physics Discussions, 2020, 1–62, https://doi.org/10.5194/acp-2019-1214, 2020.

Hamed, K. H. and Rao, A. R.: A modified Mann-Kendall trend test for autocorrelated data, Journal of hydrology, 204, 182–196, https://doi.org/10.1016/S0022-1694(97)00125-X, 1998.

Hand, J., Gill, T., and Schichtel, B.: Urban and rural coarse aerosol mass across the United States: Spatial and seasonal variability and long-term trends, Atmospheric Environment, 218, 117 025, https://doi.org/10.1016/j.atmosenv.2019.117025, 2019a.

Hand, J., Prenni, A., Schichtel, B., Malm, W., and Chow, J.: Trends in remote PM2.5 residual mass across the United States: Implications for aerosol mass reconstruction in the IMPROVE network, Atmospheric Environment, 203, 141 – 152, https://doi.org/https://doi.org/10.1016/j.atmosenv.2019.01.049, 2019b.

Haywood, J. and Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review, Reviews of geophysics, 38, 513–543, https://doi.org/10.1029/1999RG000078, 2000.

Holben, B. N., Tanre, D., Smirnov, A., Eck, T., Slutsker, I., Abuhassan, N., Newcomb, W., Schafer, J., Chatenet, B., Lavenu, F., et al.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, Journal of Geophysical Research: Atmospheres, 106, 12 067–12 097, https://doi.org/10.1029/2001JD900014, 2001.

Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S.-C., and Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, Atmospheric Chemistry and Physics, 12, 8037–8053, https://doi.org/10.5194/acp-12-8037-2012, 2012.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J.-J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T., and Zender, C. S.: Global dust model intercomparison in AeroCom phase I, Atmospheric Chemistry and Physics, 11, 7781–7816, https://doi.org/10.5194/acp-11-7781-2011, 2011.

Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., et al.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Johnson, B., Shine, K., and Forster, P.: The semi-direct aerosol effect: Impact of absorbing aerosols on marine stratocumulus, Quarterly Journal of the Royal Meteorological Society, 130, 1407–1422, https://doi.org/10.1256/qj.03.61, 2004.

Kaufman, Y. J., Tanré, D., and Boucher, O.: A satellite view of aerosols in the climate system, Nature, 419, 215, https://doi.org/10.1038/nature01091, 2002.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment – optical properties in aerosol component modules of global models, Atmospheric Chemistry and Physics, 6, 1815–1834, https://doi.org/10.5194/acp-6-1815-2006, 2006.

Kinne, S., O'Donnel, D., Stier, P., Kloster, S., Zhang, K., Schmidt, H., Rast, S., Giorgetta, M., Eck, T. F., and Stevens, B.: MAC-v1: A new global aerosol climatology for climate studies, Journal of Advances in Modeling Earth Systems, 5, 704–740, https://doi.org/10.1029/2002JD002975, 2013.

Kirkevåg, A., Grini, A., Olivié, D., Seland, O., Alterskjær, K., Hummel, M., Karset, I. H., Lewinschal, A., Liu, X., Makkonen, R., et al.: A production-tagged aerosol module for Earth system models, OsloAero5. 3-extensions and updates for CAM5. 3-Oslo, Geoscientific Model Development, https://doi.org/10.5194/gmd-11-3945-2018, 2018.

Labordena, M., Neubauer, D., Folini, D., Patt, A., and Lilliestam, J.: Blue skies over China: The effect of pollution-control on solar power generation and revenues, PloS one, 13, e0207 028, https://doi.org/10.1371/journal.pone.0207028, 2018.

Laj, P., Bigi, A., Rose, C., Andrews, E., Lund Myhre, C., Collaud Coen, M., Wiedensohler, A., Schultz, M., Ogren, J. A., Fiebig, M., Gliß, J., Mortier, A., Pandolfi, M., Petäjä, T., Kim, S.-W., Aas, W., Putaud, J.-P., Mayol-Bracero, O., Keywood, M., Labrador, L., Aalto, P., Ahlberg, E., Alados Arboledas, L., Alastuey, A., Andrade, M., Artíñano, B., Ausmeel, S., Arsov, T., Asmi, E., Backman, J., Baltensperger, U., Bastian, S., Bath, O., Beukes, J. P., Brem, B. T., Bukowiecki, N., Conil, S., Couret, C., Day, D., Dayantolis, W., Degorska, A., Dos Santos, S. M., Eleftheriadis, K., Fetfatzis, P., Favez, O., Flentje, H., Gini, M. I., Gregorič, A., Gysel-Beer, M., Hallar, G. A., Hand, J., Hoffer, A., Hueglin, C., Hooda, R. K., Hyvärinen, A., Kalapov, I., Kalivitis, N., Kasper-Giebl, A., Kim, J. E., Kouvarakis, G., Kranjc, I., Krejci, R., Kulmala, M., Labuschagne, C., Lee, H.-J., Lihavainen, H., Lin, N.-H., Löschau, G., Luoma, K., Marinoni, A., Meinhardt, F., Merkel, M., Metzger, J.-M., Mihalopoulos, N., Nguyen, N. A., Ondracek, J., Peréz, N., Perrone, M. R., Petit, J.-E., Picard, D., Pichon, J.-M., Pont, V., Prats, N., Prenni, A., Reisen, F., Romano, S., Sellegri, K., Sharma, S., Schauer, G., Sheridan, P., Sherman, J. P., Schütze, M., Schwerin, A., Sohmer, R., Sorribas, M., Steinbacher, M., Sun, J., Titos, G., Tokzko, B., Tuch, T., Tulet, P., Tunved, P., Vakkari, V., Velarde, F., Velasquez, P., Villani, P., Vratolis, S., Wang, S.-H., Weinhold, K., Weller, R., Yela, M., Yus-Diez, J., Zdimal, V., Zieger, P., and Zikova, N.: A global analysis of climate-relevant aerosol properties retrieved from the network of GAW near-surface observatories, Atmospheric Measurement Techniques Discussions, 2020, 1–70, https://doi.org/10.5194/amt-2019-499, 2020.

Li, X., Wagner, F., Peng, W., Yang, J., and Mauzerall, D. L.: Reduction of solar photovoltaic resources due to air pollution in China, Proceedings of the National Academy of Sciences, 114, 11 867–11 872, https://doi.org/10.1073/pnas.1711462114, 2017.

Likens, G. E., Butler, T. J., and Buso, D. C.: Long-and short-term changes in sulfate deposition: effects of the 1990 Clean Air Act Amendments, Biogeochemistry, 52, 1–11, https://doi.org/10.1023/A:1026563400336, 2001.

Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, Atmospheric Chemistry and Physics, 5, 715–737, https://doi.org/10.5194/acp-5-715-2005, 2005.

Lund, M. T., Myhre, G., Haslerud, A. S., Skeie, R. B., Griesfeller, J., Platt, S. M., Kumar, R., Myhre, C. L., and Schulz, M.: Concentrations and radiative forcing of anthropogenic aerosols from 1750 to 2014 simulated with the Oslo CTM3 and CEDS emission inventory, https://doi.org/10.5194/gmd-11-4909-2018, 2018.

Lurton, T. et al.: Implementation of the CMIP6 forcing data in the IPSL-CM6A-LR model, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2019MS001940, 2019.

Mortier, A., Goloub, P., Derimian, Y., Tanré, D., Podvin, T., Blarel, L., Deroo, C., Marticorena, B., Diallo, A., and Ndiaye, T.: Climatology of aerosol properties and clear-sky shortwave radiative effects using Lidar and Sun photometer observations in the Dakar site, Journal of Geophysical Research: Atmospheres, 121, 6489–6510, https://doi.org/10.1002/2015JD024588, 2016.

Myhre, G., Berglen, T. F., Johnsrud, M., Hoyle, C., Berntsen, T. K., Christopher, S., Fahey, D., Isaksen, I. S., Jones, T., Kahn, R., et al.: Modelled radiative forcing of the direct aerosol effect with multi-observation evaluation, Atmospheric Chemistry and Physics, 9, 1365–1392, https://doi.org/10.5194/acp-9-1365-2009, 2009.

Neubauer, D., Ferrachat, S., Drian, S.-L., Stier, P., Partridge, D. G., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., Lohmann, U., et al.: The global aerosol-climate model ECHAM6. 3-HAM2. 3–Part 2: Cloud evaluation, aerosol radiative forcing and climate sensitivity, Geoscientific Model Development Discussions, https://doi.org/10.3929/ethz-b-000361964, 2019.

Norris, J. R. and Wild, M.: Trends in aerosol radiative effects over Europe inferred from observed cloud cover, solar "dimming," and solar "brightening", Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD007794, 2007.

Ogunjobi, K., Ajayi, V., Balogun, I., Omotosho, J., and He, Z.: The synoptic and optical characteristics of the harmattan dust spells over Nigeria, Theoretical and Applied Climatology, 93, 91–105, https://doi.org/10.1007/s00704-007-0332-2, 2008.

O'Neill, N., Eck, T., Smirnov, A., Holben, B., and Thulasiraman, S.: Spectral discrimination of coarse and fine mode optical depth, Journal of Geophysical Research: Atmospheres, 108, 2003.

Pandolfi, M., Alados-Arboledas, L., Alastuey, A., Andrade, M., Angelov, C., Artiñano, B., Backman, J., Baltensperger, U., Bonasoni, P., Bukowiecki, N., et al.: A European aerosol phenomenology–6: scattering properties of atmospheric aerosol particles from 28 ACTRIS sites, Atmospheric Chemistry and Physics, 18, 7877–7911, https://doi.org/10.5194/acp-18-7877-2018, 2018.

Peñuelas, J. and Staudt, M.: BVOCs and global change, Trends in plant science, 15, 133–144, https://doi.org/10.1016/j.tplants.2009.12.005, 2010.

Pöschl, U.: Atmospheric aerosols: composition, transformation, climate and health effects, Angewandte Chemie International Edition, 44, 7520–7540, https://doi.org/10.1002/anie.200501122, 2005.

Putaud, J.-P., Dingenen, R. V., Alastuey, A., Bauer, H., Birmili, W., Cyrys, J., Flentje, H., Fuzzi, S., Gehrig, R., Hansson, H., Harrison, R., Herrmann, H., Hitzenberger, R., Hüglin, C., Jones, A., Kasper-Giebl, A., Kiss, G., Kousa, A., Kuhlbusch, T., Löschau, G., Maenhaut, W., Molnar, A., Moreno, T., Pekkanen, J., Perrino, C., Pitz, M., Puxbaum, H., Querol, X., Rodriguez, S., Salma, I., Schwarz, J., Smolik, J., Schneider, J., Spindler, G., ten Brink, H., Tursic, J., Viana, M., Wiedensohler, A., and Raes, F.: A European aerosol phenomenology – 3: Physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe, Atmospheric Environment, 44, 1308 – 1320, https://doi.org/10.1016/j.atmosenv.2009.12.011, 2010.

Ramachandran, S., Kedia, S., and Srivastava, R.: Aerosol optical depth trends over different regions of India, Atmospheric Environment, 49, 338–347, https://doi.org/10.1016/j.atmosenv.2011.11.017, 2012.

Rap, A., Scott, C. E., Spracklen, D. V., Bellouin, N., Forster, P. M., Carslaw, K. S., Schmidt, A., and Mann, G.: Natural aerosol direct and indirect radiative effects, Geophysical Research Letters, 40, 3297–3301, https://doi.org/10.1002/grl.50441, 2013.

Schulz, M., de Leeuw, G., and Balkanski, Y.: Sea-salt aerosol source functions and emissions, in: Emissions of Atmospheric Trace Compounds, pp. 333–359, Springer, https://doi.org/10.1007/978-1-4020-2167-1_9, 2004.

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Dentener, F., Guibert, S., et al.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations, Atmospheric Chemistry and Physics, 6, 5225–5246, https://doi.org/10.5194/acp-6-5225-2006, 2006.

Schutgens, N., Tsyro, S., Gryspeerdt, E., Goto, D., Weigum, N., Schulz, M., and Stier, P.: On the spatio-temporal representativeness of observations, Atmospheric Chemistry and Physics Discussions, https://doi.org/10.5194/acp-2017-149, 2017.

Schutgens, N. A.: Site representativity of AERONET and GAW remotely sensed AOT and AAOT observations, Atmos. Chem. Phys. Discuss., https://doi. org/10.5194/acp-2019-767, in review, https://doi.org/10.5194/acp-2019-767, 2019.

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M.: The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations, Geoscientific Model Development Discussions, 2020, 1–68, https://doi.org/10.5194/gmd-2019-378, 2020.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, Journal of the American statistical association, 63, 1379–1389, https://doi.org/10.1080/01621459.1968.10480934, 1968.

Smirnov, A., Holben, B., Eck, T., Dubovik, O., and Slutsker, I.: Cloud-screening and quality control algorithms for the AERONET database, Remote sensing of environment, 73, 337–349, https://doi.org/10.1016/S0034-4257(00)00109-7, 2000.

Smirnov, A., Holben, B., Lyapustin, A., Slutsker, I., and Eck, T.: AERONET processing algorithms refinement, in: AERONET Workshop, El Arenosillo, Spain, pp. 10–14, 2004.

Stocker, T.: Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.

Streets, D. G., Wu, Y., and Chin, M.: Two-decadal aerosol trends as a likely explanation of the global dimming/brightening transition, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026471, 2006.

Streets, D. G., Yu, C., Wu, Y., Chin, M., Zhao, Z., Hayasaka, T., and Shi, G.: Aerosol trends over China, 1980–2000, Atmospheric Research, 88, 174–182, https://doi.org/10.1016/j.atmosres.2007.10.016, 2008.

Streets, D. G., Yan, F., Chin, M., Diehl, T., Mahowald, N., Schultz, M., Wild, M., Wu, Y., and Yu, C.: Anthropogenic and natural contributions to regional trends in aerosol optical depth, 1980–2006, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2008JD011624, 2009.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Sweerts, B., Pfenninger, S., Yang, S., Folini, D., van der Zwaan, B., and Wild, M.: Estimation of losses in solar energy production from air pollution in China since 1960 using surface radiation data, Nature Energy, 4, 657–663, https://doi.org/10.1038/s41560-019-0412-4, 2019.

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., and Nakajima, T.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, Journal of Geophysical Research: Atmospheres, 105, 17 853–17 873, https://doi.org/10.1029/2000JD900265, 2000.

Takemura, T., Nakajima, T., Dubovik, O., Holben, B. N., and Kinne, S.: Single-scattering albedo and radiative forcing of various aerosol species with a global three-dimensional model, Journal of Climate, 15, 333–352, https://doi.org/10.1175/1520-0442(2002)015<0333:SSAARF>2.0.CO;2, 2002.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2004JD005029, 2005.

Tegen, I., Koch, D., Lacis, A. A., and Sato, M.: Trends in tropospheric aerosol loads and corresponding impact on direct radiative forcing between 1950 and 1990: A model study, Journal of Geophysical Research: Atmospheres, 105, 26 971–26 989, https://doi.org/10.1029/2000JD900280, 2000.

Tegen, I., Neubauer, D., Ferrachat, S., Drian, S.-L., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., et al.: The global aerosol-climate model ECHAM6. 3-HAM2. 3-Part 1: Aerosol evaluation, Geoscientific Model Development, 12, 1643–1677, https://doi.org/10.5194/gmd-12-1643-2019, 2019.

Tilmes, S., Hodzic, A., Emmons, L. K., Mills, M. J., and Gettelman, A.and Kinnison, D. E. P. M. L. J.-F. V. F. S. M. J. P. C. J. J. L. X.: Climate forcing and trends of organic aerosols in the Community Earth System Model (CESM2), Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2019MS001827, 2019.

Tørseth, K., Aas, W., Breivik, K., Fjæraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg, S., and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972ndash;2009, Atmospheric Chemistry and Physics, 12, 5447–5481, https://doi.org/10.5194/acp-12-5447-2012, 2012.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., Artaxo, P., Bahadur, R., Balkanski, Y., Bauer, S. E., Bellouin, N., Benedetti, A., Bergman, T., Berntsen, T. K., Beukes, J. P., Bian, H., Carslaw, K. S., Chin, M., Curci, G., Diehl, T., Easter, R. C., Ghan, S. J., Gong, S. L., Hodzic, A., Hoyle, C. R., Iversen, T., Jathar, S., Jimenez, J. L., Kaiser, J. W., Kirkevåg, A., Koch, D., Kokkola, H., Lee, Y. H., Lin, G., Liu, X., Luo, G., Ma, X., Mann, G. W., Mihalopoulos, N., Morcrette, J.-J., Müller, J.-F., Myhre, G., Myriokefalitakis, S., Ng, N. L., O'Donnell, D., Penner, J. E., Pozzoli, L., Pringle, K. J., Russell, L. M., Schulz, M., Sciare, J., Seland, Ø., Shindell, D. T., Sillman, S., Skeie, R. B., Spracklen, D., Stavrakou, T., Steenrod, S. D., Takemura, T., Tiitta, P., Tilmes, S., Tost, H., van Noije, T., van Zyl, P. G., von Salzen, K., Yu, F., Wang, Z., Wang, Z., Zaveri, R. A., Zhang, H., Zhang, K., Zhang, Q., and Zhang, X.: The AeroCom evaluation and intercomparison of organic aerosol in global models, Atmospheric Chemistry and Physics, 14, 10 845–10 895, https://doi.org/10.5194/acp-14-10845-2014, 2014.

Turnock, S., Butt, E., Richardson, T., Mann, G., Reddington, C., Forster, P., Haywood, J., Crippa, M., Janssens-Maenhout, G., Johnson, C., et al.: The impact of European legislative and technology measures to reduce air pollutants on air quality, human health and climate, Environmental Research Letters, 11, 024 010, https://doi.org/10.1088/1748-9326/11/2/024010, 2016.

Wang, R., Andrews, E., Balkanski, Y., Boucher, O., Myhre, G., Samset, B. H., Schulz, M., Schuster, G. L., Valari, M., and Tao, S.: Spatial Representativeness Error in the Ground-Level Observation Networks for Black Carbon Radiation Absorption, Geophysical Research Letters, 45, 2106–2114, https://doi.org/10.1002/2017GL076817, 2018.

Wang, Z., Zhang, H., and Lu, P.: Improvement of cloud microphysics in the aerosol-climate model BCC_AGCM2. 0.1 _CUACE/Aero, evaluation against observations, and updated aerosol indirect effect, Journal of Geophysical Research: Atmospheres, 119, 8400–8417, https://doi.org/10.1002/2014JD021886, 2014.

Werth, D. and Avissar, R.: The local and global effects of Amazon deforestation, Journal of Geophysical Research: Atmospheres, 107, LBA–55, https://doi.org/10.1029/2001JD000717, 2002.

White, W. H., Trzepla, K., Hyslop, N. P., and Schichtel, B. A.: A critical review of filter transmittance measurements for aerosol light absorption, and de novo calibration for a decade of monitoring on PTFE membranes, Aerosol Science and Technology, 50, 984–1002, https://doi.org/10.1080/02786826.2016.1211615, 2016.

Willis, M. D., Leaitch, W. R., and Abbatt, J. P.: Processes Controlling the Composition and Abundance of Arctic Aerosol, Reviews of Geophysics, 56, 621–671, https://doi.org/10.1029/2018RG000602, 2018.

Yoon, J., von Hoyningen-Huene, W., Kokhanovsky, A., Vountas, M., and Burrows, J.: Trend analysis of aerosol optical thickness and Ångström exponent derived from the global AERONET spectral observations, Atmos. Meas. Tech, 5, 1271–1299, https://doi.org/10.5194/amt-5-1271-2012, 2012.

Yu, H., Chin, M., Yuan, T., Bian, H., Remer, L. A., Prospero, J. M., Omar, A., Winker, D., Yang, Y., Zhang, Y., et al.: The fertilizing role of African dust in the Amazon rainforest: A first multiyear assessment based on data from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations, Geophysical Research Letters, 42, 1984–1991, https://doi.org/10.1002/2015GL063040, 2015.

Zhang, H., Wang, Z., Wang, Z., Liu, Q., Gong, S., Zhang, X., Shen, Z., Lu, P., Wei, X., Che, H., et al.: Simulation of direct radiative forcing of aerosols and their effects on East Asian climate using an interactive AGCM-aerosol coupled system, Climate Dynamics, 38, 1675–1693, https://doi.org/10.1007/s00382-011-1131-0, 2012.

Zhang, H., Jing, X., and Li, J.: Application and evaluation of a new radiation code under McICA scheme in BCC_AGCM2. 0.1, Geoscientific Model Development, 7, 737, https://doi.org/10.5194/gmd-7-737-2014, 2014.

Zhang, J. and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products, Atmospheric Chemistry and Physics, 10, 10 949–10 963, https://doi.org/10.5194/acp-10-10949-2010, 2010.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K., Huo, H., Kannari, A., Klimont, Z., Park, I., Reddy, S., Fu, J., et al.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmospheric Chemistry and Physics, 9, 5131–5153, https://doi.org/10.5194/acp-9-5131-2009, 2009.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L., Dunne, J., et al.: The GFDL global atmosphere and land model AM4. 0/LM4. 0: 2. Model description, sensitivity studies, and tuning strategies, Journal of Advances in Modeling Earth Systems, 10, 735–769, https://doi.org/10.1002/2017MS001208, 2018a.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L., Dunne, J., et al.: The GFDL global atmosphere and land model AM4. 0/LM4. 0: 1. Simulation characteristics with prescribed SSTs, Journal of Advances in Modeling Earth Systems, 10, 691–734, https://doi.org/10.1002/2017MS001209, 2018b.

Table 1. List of observations and model datasets used in this study (see text for explanation).

Parameter	Type	Observation networks	Models
AOD	Column	AERONET[1]	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; GEOS; OsloCTM3; GFDL-AM4; BCC-CUACE; CanESM5; CESM2; IPSL-CM6A
~~AOD<1μm~~ AOD_f	Column	AERONET	NorESM2; SPRINTARS; ECHAM-HAM; GEOS; OsloCTM3; GFDL-AM4
~~AOD>1μm~~ AOD_c	Column	AERONET	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; OsloCTM3; GFDL-AM4 ~~; BCC-CUACE~~
AE	Column	AERONET	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; GEOS; OsloCTM3; GFDL-AM4
$PM_{2.5}$	Surface	EMEP[2]; IMPROVE[3]	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; GEOS
PM_{10}	Surface	EMEP; IMPROVE	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; GEOS
SO_4	Surface	EMEP; IMPROVE; CASTNET[4]; CAPMoN[5]; EANET[6]	ECMWF-Rean; NorESM2; SPRINTARS; ECHAM-HAM; GEOS; OsloCTM3; BCC-CUACE
σ_{sp}	Surface	GAW-WDCA[7] (incl. IMPROVE; NOAA-FAN[8]; ACTRIS[9]; EMEP)	NorESM2; ~~SPRINTARS~~ OsloCTM3
σ_{ap}	Surface	GAW-WDCA (incl. NOAA-FAN; ACTRIS; EMEP)	NorESM2; ~~SPRINTARS~~ OsloCTM3

[1]Aerosol Robotic Network [2]The European Monitoring and Evaluation Program [3]Interagency Monitoring of Protected Visual Environments [4]Clean Air Status and Trends Network [5]The Canadian Air and Precipitation Monitoring Network [6]Acid Deposition Network in East Asia [7]Global Atmosphere Watch - World Data Centre for Aerosol [8]National Oceanic and Atmospheric Administration Federated Aerosol Network [9]Aerosol, Clouds, and Trace Gases Research Infrastructure

Table 2. Information on models used in this study (CAMS-Rean: CAMS reanalysis, AP3: AeroCom phase ~~3~~III; CMIP6: historical experiments from CMIP6).

Model	Group	Natural interactive emissions	Anthropogenic emissions	Meteorology	Resolution (degree)	References
ECMWF-Rean	CAMS-Rean	D, SS	MACCity	RA	0.7x0.7	Inness et al. (2019); Zhang et al. (2009)
SPRINTARS	AP3	D, SS, DMS, Oce VOC	C: SO_2, BC, OC	N	0.56x0.56	Takemura et al. (2000, 2002, 2005)
ECHAM-HAM	AP3	D, SS, DMS	C: SO_2, BC, OC	fSST	1.875x1.875	Tegen et al. (2019); Neubauer et al. (2019)
GEOS	AP3	D, SS, DMS, Oce VOC	O: SO_2, SO_4, BC, OC, NH_3	N	1.00x1.00	Bian et al. (2017); Chin et al. (2002); Colarco et al. (2010)
OsloCTM3	AP3	D, SS	C:	S	2.25x2.25	Lund et al. (2018); Myhre et al. (2009)
GFDL-AM4	AP3	D, SS, DMS, Oce&Veg OC,	C: SO_2, SO_4, BC, OC	fSST&N	1.x1.25	Zhao et al. (2018a, b)
BCC-CUACE	AP3	D, SS, DMS	C:SO_2, BC, OC	F	2.8x2.8	Zhang et al. (2012, 2014); Wang et al. (2014)
NorESM2	CMIP6	D, SS, DMS, MSA, BVOC	C: SO_2, SO_4, OC, BC	F	1.89x2.50	~~??Kirkevåg et al. (2018)~~ Seland et al. (2020); Kirkevåg et al. (2018)
CanESM5	CMIP6	D, SS, DMS	C: SO_2, OC, BC	F	2.77x2.81	Swart et al. (2019)
CESM2	CMIP6	D, SS, DMS_{clim}	C: SO_2, OC, BC	F	0.94x1.25	Danabasoglu et al. (submitted); Tilmes et al. (2019)
IPSL-CM6A	CMIP6	D, SS, DMS_{clim}	*C:SO_2, BC, OC, NH_3	fSST	2.50×1.27	Lurton et al. (2019)

Anthropogenic emissions (C=CMIP6-CEDS, O=other, *C=CMIP6 modified); Interactive natural emissions (D=dust, SS=sea salt, O=biogenic organic, V=volcanic, Oce=oceanic, Veg=vegetation, DMS=dimethyl sulfide, DMS_{clim}=dimethyl sulfide from climatology, VOC=volatile organic compounds, MSA=methane sulfonate; Meteorology (S=prescribed varying meteorology CTM, N=GCM nudged to analysed meteorology, fSST=fixed SST/SIC monthly fields GCM not nudged, F= free running coupled GCM, RA=combined reanalysis of meteorology and composition)

Table 3. Observational mean values for the year 2000, the reference year used for computing relative trends. Each value is extracted as the intercept of the linear trend computed in the 2000-2014 period, except for σ_{sp} and σ_{ap}, where the trends have been computed over 2000-2018. Because the required minimum number of yearly averages was set to seven, no trend could be computed in the southern African region.

	EUROPE	NAMERICA	SAMERICA	NAFRICA	ASIA	AUSTRALIA
AOD	0.17	0.10	0.15	0.26	~~0.35~~ 0.37	0.10
~~AOD<1μm~~ AOD_f	~~0.14~~ 0.15	0.08	~~0.12~~ 0.11	0.11	~~0.18~~ 0.22	~~0.05~~ 0.04
~~AOD>1μm~~ AOD_c	0.03	0.02	0.03	0.10	~~0.11~~ 0.09	0.03
AE	~~1.44~~ 1.43	~~1.46~~ 1.48	~~1.30~~ 1.26	~~0.72~~ 0.70	~~1.06~~ 1.16	~~0.97~~ 1.00
$PM_{2.5}$ ($\mu g\,m^{-3}$)	~~12.8~~ 12.3	~~7.3~~ 6.9	-	~~-~~9.0	-	-
PM_{10} ($\mu g\,m^{-3}$)	16.8	~~12.8~~ 12.4	-	~~19.6~~ 19.7	-	-
SO_4 ($\mu g\,m^{-3}$)	2.01	1.45	-	2.98	1.97	-
σ_{sp} (Mm^{-1})	~~33.2~~ 34.4	~~25.0~~ 23.4	-	-	-	-
σ_{ap} (Mm^{-1})	~~9.7~~ 6.3	~~2.7~~ 2.6	-	-	-	-

Table 4. Global means and trends of aerosol parameters using NorESM2 model data. The value in parenthesis is obtained by aggregating only grid-points where observation stations are located while using the complete model time series. The relative trends are calculated by averaging the absolute trends within the considered grid-points and normalizing it to the global mean for the year 2000.

	$Mean_{2000}$	Trend (%/yr)
AOD	(0.16) 0.14	(+0.1) +0.2
~~AOD<1μm~~ AOD_f	(0.09) 0.05	(+0.4) +0.6
~~AOD>1μm~~ AOD_c	(0.06) 0.09	(-0.2) +0.1
AE	(0.78) 0.43	(+0.2) +0.3
$PM_{2.5}$ ($\mu g\,m^{-3}$)	(~~12.4~~6.3) 9.1	(~~+0.2~~-1.1) +0.2
PM_{10} ($\mu g\,m^{-3}$)	(~~19.3~~11.7) 18.7	(~~+0.1~~-0.6) +0.1
SO_4 ($\mu g\,m^{-3}$)	(~~2.33~~1.86) 0.64	(~~-1.1~~-3.9) +0.4
σ_{sp} (Mm^{-1})	(~~28.0~~13.9) 21.2	(~~+0.3~~-0.3) +0.2
σ_{ap} (Mm^{-1})	(~~3.1~~2.0) 0.9	(~~+1.8~~1.4) +1.5

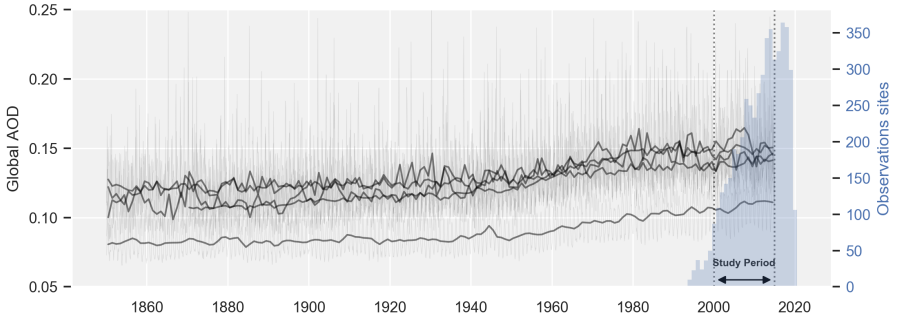


Figure 1. Global AOD computed from model historical runs (OsloCTM3, GFDL-AM4, CanESM5, CESM2, IPSL-CM6A, ECHAM-HAM) at monthly (gray lines) and yearly resolutions (black lines), overlayed with the number of active observation sites in the ~~sun photometer network~~ AERONET sunphotometer network.

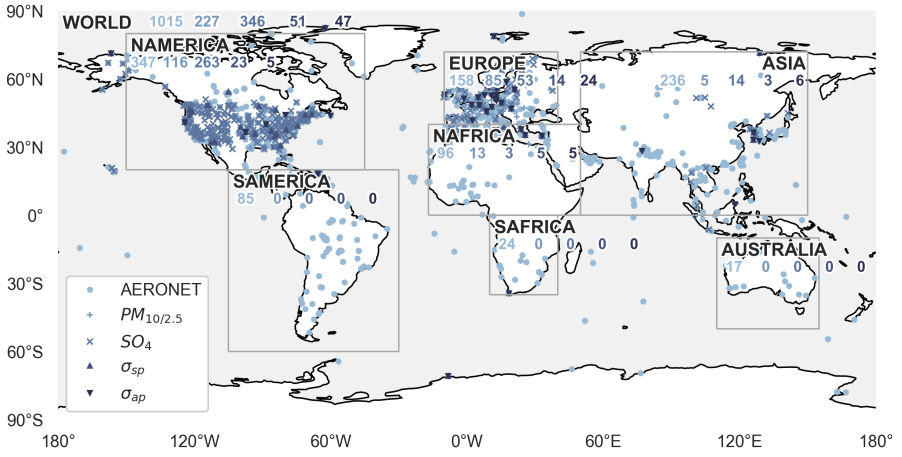


Figure 2. Distribution of the observations within the different regions considered in this study. The numbers reported within each region correspond to the maximum number of stations given for the observation networks corresponding to the five observation types found in the legend.

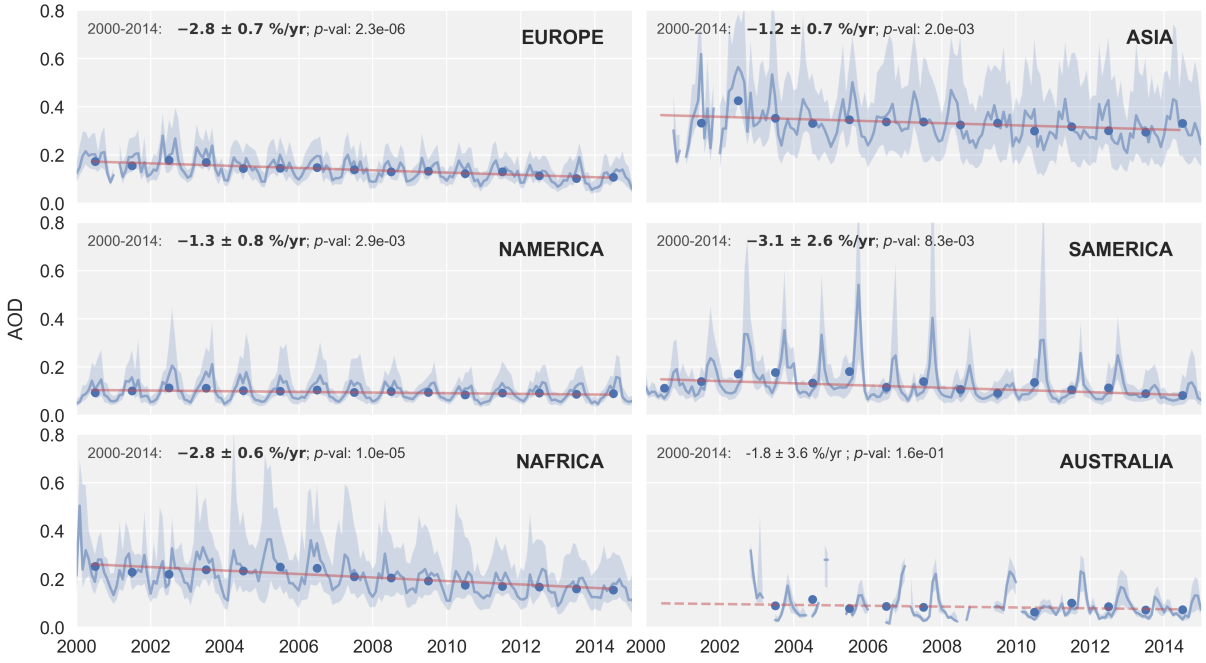


Figure 3. Regional time series of AOD. The dark blue line corresponds to the median and the light blue envelope is bound by the first and third quartiles of all valid points at the corresponding month, respectively. The blue dots correspond to the yearly averages which are used to compute the linear trend. The latter is displayed as a continuous line when the trend is significant and as a dashed line when it is not. Trend values, an error estimate and significance value are given in ~~the figure~~each pane.

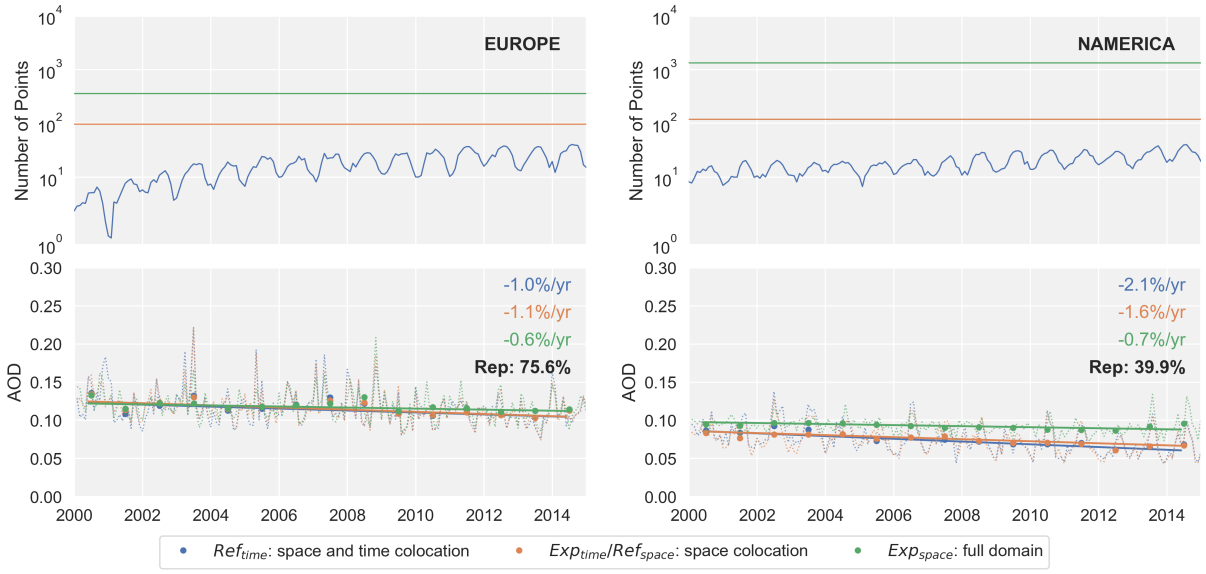


Figure 4. Three regional AOD time series and ~~respectived~~ respective trends, constructed from model data (NorESM2) for the investigation of ~~representativeness~~ the representativity of the observational data. The upper figures correspond to the number of points used to compute the regional time series for the three different datasets. The lower figures show the time series, the trends, and the resulting representativity value (black, bold). The blue color (Ref_{time}) corresponds to the model output collocated in space and time ~~to~~ with the available observations. The upper graphs show an overall increase in the number of available observations (more stations) combined with a seasonal cycle (less AOD available in wintertime). The orange color (Exp_{time}/Ref_{space}) corresponds to the model output collocated in space to the stations providing measurements, using the complete time series from 2000 to 2014. The green color (Exp_{space}) corresponds to the model output in the whole geographic region (see 2), using all of the grid boxes without any collocation to the observations.

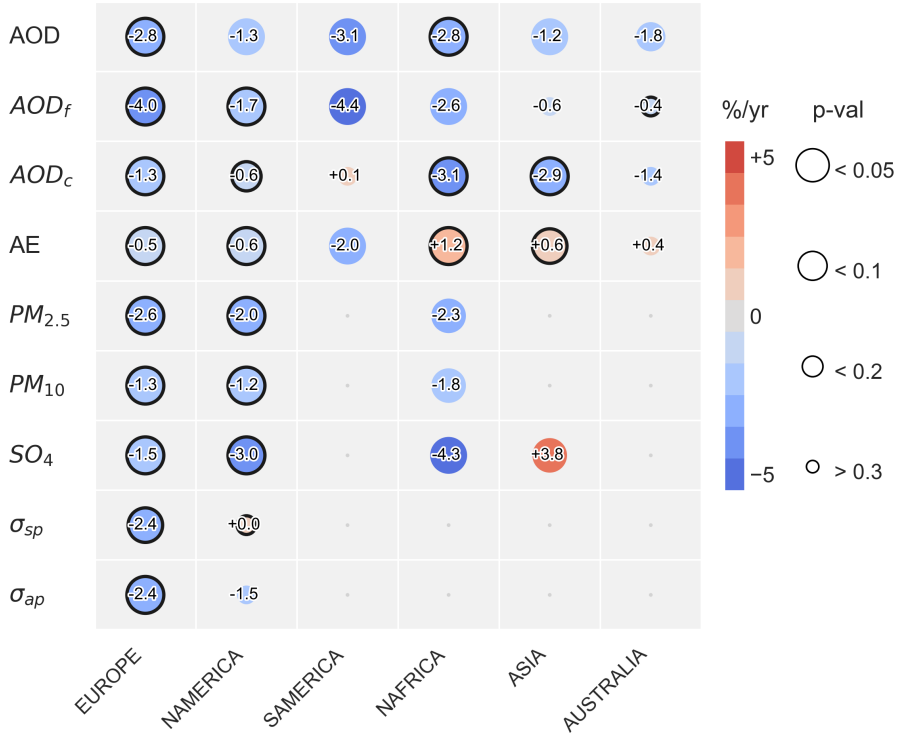


Figure 5. Regional trends of the aerosol properties computed with the observation datasets. The color of the circles corresponds to the slope, while the radius indicates the p-value. The largest circles represent the trends significant with a confidence level of 95%. The circles bordered with a black line indicate the trends associated with a representativity greater than 50%.

Figure 6. Regional trends of the aerosol properties computed with observations and models collocated in space and time to the observations. The error bars correspond to the uncertainty of the trend as calculated using both the uncertainty ~~on~~ in the Theil-Sen slope and the residuals. The bold font indicates that the trends are significant ~~with an expectancy~~ at a confidence level of 95% (p-val<0.05).

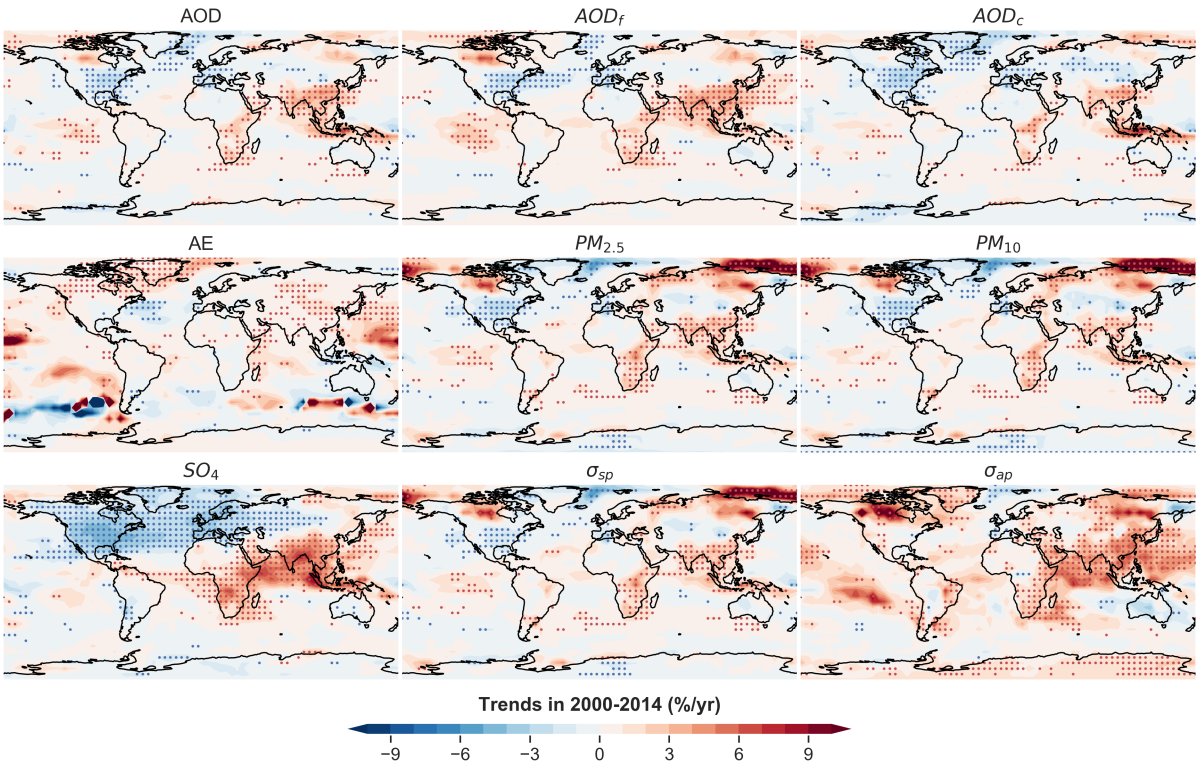


Figure 7. Global trends of aerosol properties using NorESM2 data regridded at a 5x5 degrees resolution. The blue and red dots dots indicate ~~respectively~~ significant negative and positive trends, respectively.

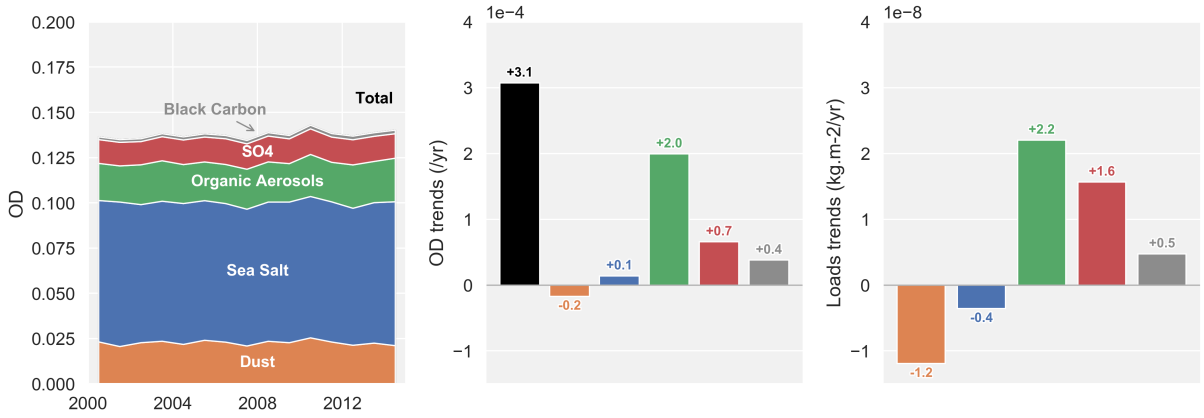


Figure 8. Absolute trends in OD and emissions of the main aerosol species computed with NorESM2. The y-axis of the trends in OD and the emissions is given according to the power of 10 indicated at the top left corner of each of the subplots.