Author's Response to Review Comments on the ACPD manuscript **Aerosol optical properties over Europe: an evaluation of the AQMEII Phase 3 simulations against satellite observations (acp-2017-1119)**

by Laura Palacios-Peña

A: First of all, we would like to thank the anonymous referees for their very valuable comments in the interactive comment on "Aerosol optical properties over Europe: an evaluation of the AQMEII Phase 3 simulations against satellite observations (acp-2017-1119)" by Laura Palacios-Peña et al. We strongly appreciate the referees' comments and following these, the manuscript has been revised in order to correct errors and to introduce the suggestions for improving the quality of the paper. Please see below our point-by-point replies:

Anonymous Referee #2

The authors present the evaluation of the aerosol optical properties simulated in the frame of the Air Quality Model Evaluation International Initiative. As a reference, MODIS AOD is used.
I have several concerns about the paper, which are specified below.
The manuscript needs a major revision before it can be considered for publication in ACP.

General comments

The language should be thoroughly checked.
A: The language has been checked by a native proofreader.

The authors say that MODIS "combined" product is not validated (page 8, line 3). If this is true, I would be very careful to evaluate the model results with the product, which is not validated. However, this is not true. Several validation papers for MODIS C6 are published, among them:
http://www.mdpi.com/2072-4292/10/3/475
https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022453
https://www.atmos-meas-tech.net/6/2989/2013/amt-6-2989-2013.pdf
http://home.iitk.ac.in/~snt/pdf/Mhawish_RSE_2017.pdf
Please, include MODIS validation results and possible bias of the model results comparison, related to bias in MODIS AOD into the discussion.
A: As the referee suggested, the results of the MODIS validation have been included in the manuscript, indicating the information about the error in: section 2.2 Observational Data.

"The used data from MODIS were Level 2 of the Atmospheric Aerosol Product (both MOD04_L2 and MYD04_L2) from the collection 6, with a resolution of 10 km which are estimated by two different algorithms, Dark Target (DT) and Deep Blue (DB). The used variables were: (1) a "combined" variable of the DT and DB algorithms which provide information about AOD at 550 nm for both ocean and land; and; (2) AE between 550 and 860 nm over the ocean estimated by the DT algorithm. There are several evaluations of this "combined" AOD products of MODIS C6 against AERONET sites around of the world (Sayer et al., 2014; Mhawish et al., 2017; Bilal et al., 2018). All of these established that a high percentage of retrievals are within the estimated error (EE) of the DT and DB algorithms, which is (±0.05 + 15%) (Levy et al., 2013). Moreover, in Sayer et al. (2014) and Bilal et al. (2018) the performance of combined retrievals outperformed DT or DB retrievals in terms of correlation (around 10%), meanwhile relative mean bias values were similar at a global scale. The preliminary estimated error (EE) for the used AE products was 0.45 in the pixels with an AOD > 0.2 (Levy et al., 2013)."

Moreover, the comparison of model results has been revised in order to include in the discussion the possible model's bias related to MODIS' bias.

MODIS provides very good but not perfect product, and validation results show weaknesses of the AOD product over certain areas. Thus, I strongly recommend to include AOD validation with AERONET (https://aeronet.gsfc.nasa.gov/) and MAN (https://aeronet.gsfc.nasa.gov/new_web/maritime_aerosol_network.html). Otherwise, the results of your analysis are biased by the MODIS quality.
A: As the referee #2 and the referees for the first revision of this work suggested, we have included the AOD and AE validation with AERONET and MAN. The description of the used product has been included in the Observational Data section. Moreover, the Results sections has been re-written taking into account these observations and their results.

Page 8, Line 21-26: *"From AERONET, AOD at 675 nm and AE between 440 and 870 nm retrievals of level 2.0 from the available European stations during the entire 2010 year were used. For this network data, the total uncertainty for the AOD data under cloud-free conditions is established as < ± 0.01 for λ > 440 nm and < ± 0.02 for shorter wavelengths (Holben et al., 1998). The same variables were used from MAN, whose estimates uncertainty of AOD in each channel is, as well as AERONET, < ± 0.02 because MAN is affiliated with the AERONET calibration and data processing standards and procedures (Smirnov et al., 2009)."*

MODIS coverage is surprisingly low in your analysis. How many (in %) pixels you discarded with your mask? Please, show it on the map. Please, repeat the analysis for all reported MODIS AOD pixels. As far as I know, MODIS deliver AOD data which passed the quality control.
A: The number of occurrences in the AOD (Figure S1) and AE (Figure S2) retrievals are showed in the Supplementary Material. We decided to apply "our mask" in order to show temporal means calculated by a number of observations large enough to ensure confident and representative results.

MODIS abroad instruments (Terra and Aqua) in Sun-synchronous orbits around the Earth does not provide data over the entire studied domain for each time step. This, together with the fact that we are using the highest quality of MODIS data, sometimes leads to a small number of occurrences in each pixel. As our results are seasonal means, we cannot calculate the seasonal mean with a really small number of occurrences. So, our mask is established to avoid this fact and in order to build a seasonal mean with a number of observations large enough to make our results robust.

To build our mask we first estimated the maximum of occurrences in the AOD or AE retrievals over the entire domain and period (JFM, AMJ, JAS and OND), which is the maximum number of observations that we can have in our domain. Then, we established the criteria that the pixels shown as final results have to present at least a number of occurrences higher than the 10% of this maximum. So, we show our validation results over the pixels which accomplish with this criterion, but we never discarded any observations over these pixels.

It would not be acceptable to show mean results where means are calculated from a really small number of occurrences, which could lead to a misunderstanding of the validation.

For the new AERONET validation, we followed the same criterion. As AERONET stations are located in a fixed place and we are using hourly means for our evaluation, we establish the maximum of possible occurrences as the maximum of hours in which AERONET stations can retrieve measurements (light hours) during the target periods (JFM, AMJ, JAS and OND). Then, we did not include those stations in which the number of hourly observations is lower than the 10% of the maximum occurrences.

This criterion has not been implemented for the MAN evaluation because this database displays punctual observation with a specific coordinate and time.

We have been re-written Page 8, Lines 27 to 29 in order to improve the explanation of "our mask".

"All the observations used in this work are not provided in a temporal regularly way. This means that the number of occurrences in each of the pixel for satellite data or in each station for AERONET data are not the same. As the results in this work are represented as seasonal means and in order to show robust means estimated with a reasonable number of occurrences, a mask showing only those pixels (stations) where the satellite (station) occurrences were higher that the 10% of the maximum possible occurrences was implemented. The maximum possible occurrences for satellite data were selected as the maximum of occurrences over the studied season (JFM, AMJ, JAS or OND) and the domain. Figures S1 and S2, in the appendix, show the number of observations used when the mask was implemented. For AERONET, the maximum possible occurrences were established as the maximum of solar-light hours in each station during each season."

Specific comments

Abstract.

- Please, add numbers (abs, or %) for under/overestimation results Done
- I disagree with the conclusion that "spatial and temporal variability of this variable is well-represented by all the models". Is that conclusion done from the comparison with MODIS? As I see from figure S3, almost all models have similar tendencies to underestimate, compared to MODIS, AOD over Siberia and underestimate AOD it the southern part of AOI.

A: We meant that special and temporal variability of AOD was better represented than AE. We have been re-written this sentence:
"Despite this behaviour, the spatial and temporal variability of AOD was better represented by all the models than AE variability, which was strongly underestimated in all the simulations."

Page 4, Line 20. Please, correct to "ATSR".Done
Page 7, Line 14. Please, specify "x" in MxD. Done
Page 8, Line 22. What is EE for AE MODIS product? Please, include the discussion to explain your choice for AE between -0.5 and 4.
A: The EE for AE MODIS product is indicated in lines 5-6, page 8: "*The preliminary estimated error (EE) for the used AE products was 0.45 in the pixels with an AOD > 0.2 (Levy et al., 2013).*"
The following discussion about the choice for AE between -0.5 and 4 has been included as:
"Moreover, and according to the EE for the AE products of MODIS, we set the AE values range between -0.5 and 4.0. It is widely known that AE values spread from 0 to 4 and even sometimes, when really coarse particles are presented, they can reach negative values. Then, we choose AE values between -0.5 as lower limit in order to cover possible negative values in a close smoothing value to the EE for the AE products of MODIS".

Page 8, lines 4-5. Please, discuss briefly the results here.
A: “The MAE was calculated as in Equation 4 and provides an estimation of the magnitude of the error independently of over-or underestimation.”

Page 8, Line 27. Please, explain the mask in the other words. The current explanation is not clear. Was that mask applied for AOD? Or AE? Do you mean, that if for certain location max AOD was 1, you discarded all the cases when AOD<0.1 for that location? Why? For AE, AOD limit of 0.1 is acceptable.
A: As we explained above, we have been re-written Page 8, Lines 27 to 29 in order to improve the explanation of the mask.

Page 8, Line 29. What is the measure of confidence here?
A: The mask is not a proper measure of the confidence of our results. The mask ensures that we are calculating the mean results with a large enough number of occurrences.

Technical comments
Please change the color scale for MODIS AOD to max. 1. With the current color scale, the AOD variation below 0.5 is hardly visible. I advice to include also the red color to the color panel.
A: We followed the advice of the reviewer and the maximum of the color scale has been changed. Moreover, we reversed the color scale in order to improve the representation of the AOD variation. We did not include the red color to this panel because this panel contain the green color and people with deuteranopia, protanopia, tritanopia of similar diseases could misunderstanding the color panel. Some examples in this link: https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html

Anonymous Referee #3
General comments
This is an interesting work that attempts to evaluate modeled AOD and AE over Europe using six model simulations performed under the AQMEII3 framework. The AQMEII initiative has provided a great opportunity for air quality model evaluation and model intercomparison across two continents that allows the community to assess the accuracy of the modeling systems, the drivers of their differences and make suggestions for future model improvements.

The main objection regarding the scientific methodology is the use of only one MODIS product to conduct model evaluation and intercomparison. The authors should take advantage of all possible AOD/AE observations in the model domain over Europe, to enhance their understanding of model behavior.
A: As the referee #3 and the referees for the first revision of this work suggested, we have included the AOD and AE validation with AERONET and MAN.

The paper needs grammatical editing to improve the language and some restructuring to improve the flow. There are a lot of clarifications needed for the methodology and discussion of the results. I am in favor of publishing this paper with Atmospheric Chemistry and Physics with Major Revisions. The specific comments that follow will help improve the discussion of the methodology and significance of the findings so that the overall quality of the manuscript is enhanced.
A: We have revised the entire text and tried to do our best to improve the language and the structure of the text. As above commented, the language had been checked by a native proofreader.

Specific comments
Abstract:
1. Line 13: “this variable” refers to AOD? Please be specific Done

2. General comment: It would be more beneficial to have a quantitative description of the conclusions. If AE is predicted more erroneously than AOD, add some quantitative measures to that statement in terms of errors/biases etc. Descriptive characterizations like "more serious errors" do not add any substantial information.

A: We have corrected this type of expression and added a quantitative description as the referee suggested.

Introduction:

1. Lines 4-5 and elsewhere: please remove the phrase "so-called". The aerosol-cloud- radiation interactions are widely known; the definition of the acronym is enough. Done
2. General comment: The introduction includes a lot of information on past and recent work similar to this study. What is not clear is how this study is different from others. What is the new contribution made by this study to the scientific knowledge of modeled AOD and AE? A high-level brief statement would be appropriate in the introduction.

A: Following the referee suggestion, a brief statement has been included in the last part of the introduction:

"Curci et al. (2017) used AQMEII Phase 3 simulations to evaluate the black carbon absorption against AERONET but no works have evaluated against satellite the seasonal representation of optical properties over Europe by the regional models involved in AQMEII Phase 3. This represents an added value because three main reasons: 1) all the regional models evaluated here were run using the same boundaries and initial conditions which permit investigate the importance of different processes and feedbacks in each models; 2) the use of two different emissions data set permits evaluated the influence of these in the aerosol optical properties representation; and 3) the use of online coupled chemistry-meteorology/climate models (as were some of the used here) will permit to investigate the influence of the ARI and ACI. As above-mentioned, aerosol optical properties influence ARI and ACI, a good representation of them is, thus, a key issue to reduce the uncertainty of aerosol effects on the Earth's climate system. For this reason, our main study aim was to evaluate the representation of two main aerosol optical properties, AOD and AE, using the models of the AQMEII Phase 3 initiative over Europe. The evaluation was made by using the remote-sensing observations from the MODIS sensor and from AERONET and MAN (Maritime Aerosol Network). Section 2 provides a brief description of the observational and models data, and the evaluation methodology. Section 3 presents the evaluation results. Finally, Section 4 summarises the main conclusions reached."

Methodology:

1. Page 6, IT1 simulation: ". . .Meteorological inputs were generated using WRF-Chem version 3.4.1. Anthropogenic emission were MACC and biogenic were computed by MEGAN.WRF-Chem was adopted to predict GOCART (Goddard Chemistry Aerosol Radiation and Transport) dust emissions (Ginoux et al., 2001) along with meteorology." What is the role of WRF-Chem in the WRF-CAMx simulations? It is not clear how those two modeling frameworks are combined. It doesn't make sense to use WRF-Chem for meteorology while WRF is already the modeling system used combined with CAMx.

A: There was an error and the referee was right, WRF is used comvined with CAMx and not WRF-Chem. This has been corrected.

2. Table 1 should include one more column that describes the origin of AOD/AE calculations: prognostic (i.e. during runtime) or diagnostic (i.e post-processing).

A: Done

3. Description of dust sources for each simulation is very important since the domain covers North Africa and the signal from the satellite is much stronger there. Prescribed dust emissions (as total PM), online dust sources, etc. This information can be added to Table 1 under Aerosol Mechanism.

A: Done

4. It's a surprise that the authors did not make use of AERONET data for such comprehensive AOD and AE model evaluation. This is strongly recommended for the revised version.

A: As all the referees suggested the evaluation against AERONET and MAN network has been included in the manuscript

5. In addition, the merged AOD product has been used in a number of recent publications with one of the most important being Sayer et al. (2014) (https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD022453) which is a point of reference for the usage of these products.

A: This reference, among others, has been listed as results of the merged AOD product, as well as the error of this product.

6. Observational data: The authors state that "The selection of this observational data was due to results found by Palacios-Peña et al. (2017a)". More detailed explanation is needed for the selection of observational data than the reference to a previous publication. The readers should not have to read other papers to understand the basics of the approach.

A: A more detailed explanation has been done.

"The selection of this observational data was due to results found by Palacios-Peña et al. (2018) where they evaluated the uncertainty in the satellite representation by comparing MODIS, OMI (Ozone Monitoring Instrument) and SeaWIFS (Sea-viewing Wide Field-of-view Sensor) AOD retrievals against AERONET. They found that MODIS presented the best agreement with the AERONET observations compared to other satellite AOD observations during two studies with high aerosol load took place in 2010 all over Europe."

7. Observational data (page 8, line 26): The description of the threshold (mask) is not clear. Is it applied to AOD and AE? Why specifically 10% of the maximum? Does that mean that the threshold varies with location, depending on the maximum observed satellite retrieval? Why?

A: The use of the threshold or mask has been explained above. Moreover, we did our best to a better description of the mask in the manuscript.

"All the observations used in this work are not provided in a temporal regularly way. This means that the number of occurrences in each of the pixel for satellite data or in each station for AERONET data are not the same. As the results in this work are represented as seasonal means and in order to show robust means estimated with a reasonable number of occurrences, a mask showing only those pixels (stations) where the satellite (station) occurrences were higher that the 10% of the maximum possible occurrences was implemented. The maximum possible occurrences for satellite data were selected as the maximum of occurrences over the studied season (JFM, AMJ, JAS or OND) and the domain. Figures S1 and S2, in the appendix, show the number of observations used when the mask was implemented. For AERONET, the maximum possible occurrences were established as the maximum of solar-light hours in each station during each season."

8. Figure S1 shows no satellite observations over land for JFM and limited data for OND, assuming that grey colored areas have zero observations. The picture is worse for AE where no observations are seen over land. With these data points

available for evaluation, how can the authors estimate seasonal statistical metrics? This is a clear case where AERONET data and/or additional satellite AOD/AE products should be used to cover for these gaps (MODIS, MISR, or whatever is available for that specific period),

A: We have included AERONET in our evaluation.

Results

1. Page 9, line 24: The phrase "The seasonal break down is presented besides named." is confusing. Please rephrase.

A: The phrase has been rephrased as: *"The seasonal means are presented in the named column".*

2. The results section needs some restructuring to allow a nice reading flow. Paragraphs should contain more than one sentence and it will be beneficial to keep the discussion on one topic/figure/statistic in one paragraph. For example, devote one paragraph describing spatial patterns of MODIS AOD, one paragraph for each season and so on.

A: Following the referee comment, we have restructured the results section.

3. The entire AOD evaluation section does not include an interpretation of the under- or over- estimation seen for each model simulation. Also, why is there a difference between model simulations? This is much more important and interesting than just presenting the statistical findings.

A: Following the referee's comments (here and below), a major effort has been done to take into account the referee's suggestion and the Section 3 (Results) and 4 (Summary and Conclusions) has been restructured.

4. Section 3.2, page 18, line 6: "In this section, the simulations run with the available data were less than they were for AOD." What is the meaning of this sentence? Wasn't there one annual simulation performed by each modeling system? This is confusing; please explain.

A: This phrase has been removed. Instead this has been added at the beginning of the section 3.2: *"AE models simulations are less than for AOD because some models did not provide AOD at different wavelengths and then, it was no possible to estimate AE following the methodology stablished above."*

5. Section 3.2 (Variability): page 22, line 4: why is ES1 showing a remarkable AOD representation? The PDF for ES1 shows that the simulation underrepresented low AOD and overestimated high AOD for the majority of the seasons (except JAS maybe).

A: This ES1 behaviour in the AOD representation is doing in the section Summary and Conclusions.

"ES1, which used WRF-Chem, presented a high AOD overestimation due to the dust outbreaks. This marked overestimation took place because of a bug in the used dust scheme, which lacks the gravitational settling."

…….

"For the AOD representation, all the simulations presented similar PDF to the observed values. …. ES1 presented higher probabilities for high AOD values than those observed due to the above-explained lack of dust gravitational settling."

6. Figures S5 to S9 in the supplement are never mentioned in the text. In addition, Figure S9 (annual PM2.5 emissions in Europe and North America) does not seem to align with the contents of this manuscript. Please specify the role of all additional figures in the supplementary material.

A: For the sake of brevity and in order to show only relevant figures for the work these figures have been remove from the Supplementary Material. Moreover, the figures displayed the total values for the entire year as well as the text regarding they has been removed because they do not provide added valuable information.

Summary and Conclusions

1. Page 22, line 29: What is the meaning of front observations in the following phrase “an evaluation of the simulations of the front observations was needed”? Please revise accordingly.

A: This has been rephrased as: *“an evaluation of the simulations against observations was needed.”*

2. Page 23, line 7: Please rephrase the following sentence by replacing “misunderstanding” with a more appropriate characterization (“due to a misunderstanding of the simulation of the aerosol vertical. . .”).

A: The phrase has been rewritten: *“As established in Palacios-Peña et al. (2018), this underestimation may be due to a misinterpretation of the simulation of the aerosol vertical and may, therefore, be due to the AOD representation given the understated injection height of the total biomass burning emissions found for the MACC emissions by Soares et al. (2015).”*

3. A lot of the text in section 4 could be included in the respective discussion of the results (see comment #3 in the Results section above). I suggest that the authors discuss here the main key results drawn by their analysis in a clear and concise way.

A: A major effort has been done to take into account the referee’s suggestion and the Section 3 (Results) and 4 (Summary and Conclusions) has been restructured.

Aerosol optical properties over Europe: an evaluation of the AQMEII Phase 3 simulations against satellite observations

Laura Palacios-Peña[1], Pedro Jiménez-Guerrero[1], Rocío Baró[1], Alessandra Balzarini[2], Roberto Bianconi[3], Gabriele Curci[4,5], Tony Christian Landi[6], Guido Pirovano[2], Marje Prank[7,8], Angelo Riccio[9], Paolo Tuccella[4,5], and Stefano Galmarini[10]

[1]Physics of the Earth, Department of Physics, Regional Campus of International Excellence (Campus Mare Nostrum), University of Murcia (UMU-MAR), 30100 Murcia, Spain
[2]Ricerca sul Sistema Energetico (RSE SpA), Milano, Italy
[3]Enviroware srl, Concorezzo, MB, Italy
[4]CETEMPS, University of L'Aquila, Italy
[5]Dept. Physical and Chemical Sciences, University of L'Aquila, Italy
[6]CNR - Institute for Atmospheric Sciences and Climate, Bologna, Italy
[7]Finnish Meteorological Institute, Atmospheric Composition Research Unit, Helsinki, Finland
[8]Cornell University, Atmospheric and Earth Sciences, Ithaca, NY, USA
[9]University Parthenope of Naples, Dept. of Science and Technology, Napoli, Italy
[10]European Commission, Joint Research Centre (JRC), Directorate for Energy, Transport and Climate, Air and Climate Unit, Ispra (VA), Italy

Correspondence to: Pedro Jiménez-Guerrero (pedro.jimenezguerrero@um.es)

Abstract. The main uncertainties in estimates of changes in the Earth's energy budget are related to the role of atmospheric aerosols. These changes are caused mainly by aerosol-radiation (ARI) and aerosol-cloud interactions (ACI), which heavily depend on aerosol properties. From the 1980s, many international modelling initiatives have studied atmospheric aerosols and their climate effects. Phase 3 of the Air Quality Model Evaluation International Initiative (AQMEII) focuses on evaluating and intercomparing regional and linked global/regional modelling systems by collaborating with the Task Force on the Hemispheric Transport of Air Pollution Phase 2 (HTAP2) initiative. Within this framework, the main aim of this work was to evaluate the representation of aerosol optical depth (AOD) and the Ångström exponent (AE) by the AQMEII Phase 3 simulations over Europe. The evaluation was made using satellite data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on board the Terra and Aqua platforms. The results indicated that the skills of AQMEII simulations in the AOD representation (mean absolute errors ranged from 0.05 to .30) produced fewer errors than in the AE (mean absolute errors ranged from 0.30 to 1) . Regardless of the models and emissions used, models were skilful at representing the low and medium AOD values observed (below 0.5). However, high values (close to 1.0) were underestimated for biomass burning episodes, and were overestimated for desert dust contributions, related mainly to emission and boundary conditions. Despite this behaviour, the spatial and temporal variability of ~~this variable was well-represented~~ AOD was better represented by all the models ~~.~~ ~~Generally, the AE evaluation showed more serious errors than the AOD evaluation. Moreover, the observed variabilityof this parameter~~ than AE variability, which was strongly underestimated in all the simulations.

1 Introduction

The Fifth Assessment Report (AR5) of the Intergovernmental Panel of Climate Change (IPCC) ascribes the ~~gravest~~ large uncertainty to estimate ~~and interpret~~ changes to the Earth's energy budget to aerosol and clouds. Atmospheric aerosols produce these changes mainly by two different ways: influencing the Earth's radiation, the ~~so-called~~ aerosol-radiation interactions (ARI); and modifying clouds and precipitation, the ~~so-called~~ aerosol-clouds interactions (ACI), which also increase uncertainty due to clouds (Boucher et al., 2013).

ARI and ACI depend on the optical properties of atmospheric aerosols along with their atmospheric distribution and hygroscopicity, and their ability to act as cloud condensation nuclei (CCN) and ice nuclei (IN). All these properties are highly variable on space and time scales due to aerosol particles' short-lived, non-uniform emissions, and the dependence of sinks on meteorology (Boucher et al., 2013; Randall et al., 2007). Thus the determination of atmospheric aerosol properties, by a complex interplay between their sources, atmospheric transformation processes and their removal from the atmosphere (Boucher et al., 2013) plays a part in the ~~gravest~~ large uncertainty of aerosol effects on the Earth's climate.

It was in the 1980s when the atmospheric science community began to pay increasing attention to the atmospheric aerosol subject (Fuzzi et al., 2015). Since then, major efforts have been made to acquire better knowledge of atmospheric aerosol properties and their interactions with the Earth's climate to reduce the above-mentioned ~~grave~~ large uncertainty. Many regional field measurement campaigns have taken place; e.g., the Integrated Campaign for Aerosols, Gases and Radiation Budget (Moorthy et al., 2008, ICARB); the Megacity Impact on Regional and Global Environments field experiment (Paredes-Miranda et al., 2009, MILAGRO); the Integrated Project on Aerosol Cloud Climate and Air Quality interactions (Kulmala et al., 2011, EUCAARI); Aerosol, Radiation, and Cloud Processes affecting Arctic Climate (Warneke et al., 2010, ARCPAC); among many others (Boucher et al., 2013). Moreover, global long-term aerosol measurements are taken by surface networks, such as Global Atmosphere Watch (Ogren, 2011, GAW), Aerosol Robotic Network (Holben et al., 1998, AERONET), the European Monitoring and Evaluation Programme (Tørseth et al., 2012, EMEP) or by satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer (Remer et al., 2005, MODIS) or the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (Winker et al., 2003, CALIPSO) among many other base measurements, base networks and instruments on board satellites (Boucher et al., 2013).

Measurements provide incomplete sampling, but can be combined with information from global and regional aerosol models. There are a number of international initiatives that study, among other climate issues, atmospheric aerosols and their climate effects. Some examples are the Aerosol Comparisons between Observations and Models project, now in their Phase II (Schulz et al., 2009, AEROCOM-II), the Coupled Model Intercomparison Project, now in Phase 6 (Eyring et al., 2016, CMIP6), the Chemistry-Climate Model Initiative (Eyring et al., 2013, CCMI) or the Aerosol and Chemistry Model Intercomparison Project (Collins et al., 2016, AerChemMIP). Among these initiatives and many others, the primary purpose of the Air Quality Model Evaluation International Initiative (Rao et al., 2011, AQMEII) is to coordinate international efforts in scientific research on regional air quality model evaluations across the modelling communities of North America and Europe.

AQMEII Phase 1 (Galmarini et al., 2012) focused on developing general model-to-model and model-to-observation evaluation methodologies, Phase 2 (Galmarini et al., 2015) on simulating aerosol/climate feedbacks with online coupled modelling systems. As part of this Phase 2, some studies have evaluated aerosol properties and their effects on the climate system. Balzarini et al. (2015) analysed online model sensitivity to the chemical mechanisms of WRF-Chem in reproducing aerosol properties, and found that although different chemical mechanisms give different Aerosol Optical Depths (AOD), the latter was underestimated. Forkel et al. (2015) found pronounced feedback effects, such as a reduction in seasonal mean solar radiation of 20 Wm^{-3} and temperature of 0.25° in the summer of 2010 when ARI were taken into account. High aerosol concentrations resulted in a 10-30% decreased precipitation and low concentrations in very low cloud droplet numbers (5-100 droplets cm^{-1}) and a 50-70% lower cloud liquid water, which led to an increase in downward solar radiation of almost 50% when ACI were taken into account. Makar et al. (2015) evaluated the effect on chemistry due to feedback between aerosol and meteorology. In this study, ACI were usually found to have a stronger effect on the predictions of ozone, particulate matter and other species, and also on the atmospheric transport and chemistry of large emitting sources such as plumes from forest fires and large cities. A work similar is that of Wang et al. (2015), in which a multi-model assessment of major column abundances of gases, radiation, aerosol and cloud variables was made using available satellite data. The evaluation results showed an excellent agreement between all the simulations and satellite-derived radiation variables, as well as precipitable water vapour. Other aerosol-/cloud-related variables, such as AOD, cloud optical thickness, cloud liquid water path, CCN and cloud droplet number concentration were moderately to largely underestimated by most simulations due to the underestimations of aerosol loadings (Baró et al., 2018). They also indicated ~~grave~~ large uncertainties associated with the current model treatments of ACI.

Moreover, and through the AQMEII Phase 2, the working group 2 of the COST Action ES1004 EuMetChem (European framework for online integrated air quality and meteorology modelling, http://www.eumetchem.info/) investigated the importance of different processes and feedbacks in online coupled chemistry-meteorology/climate models for air quality simulations and weather predictions. As part of this initiative, an important aerosol load episode, the Russian wildfires in 2010, was investigated. Some results indicated that the inclusion of ARI led to a drop of between 10 and 100 Wm^{-2} of the average downward shortwave radiation on the ground and an almost 1° in the mean temperature (Forkel et al., 2016; Toll et al., 2015a). During the same episode, Baró et al. (2017) found a reduction in the 10-metre wind speed of 0.2 ms^{-1} (10%) since presence of biomass burning implied a reduction in shortwave downwelling radiation on the surface which, in turn, led to a reduction in the 2-metre temperature, and thus a reduction in the turbulence flux, and developed a stabler planetary boundary layer. Kong et al. (2015) and Palacios-Peña et al. (2017) evaluated the observation of the inclusion effects of ARI and ACI for this wildfires episode and a desert dust outbreak. The results showed that a minor, but significant, improvement was observed when ARI and ACI were taken into account.

In AQMEII Phase 3, to which this work is a contribution, the AQMEII initiative focused on evaluating and intercomparing regional and coupled global/regional modelling systems by collaborating with the Task Force on Hemispheric Transport of Air Pollution, Phase 2 (Dentener et al., 2015, HTAP2). Simulation performance followed the strategy adopted in the first two AQMEII Phases, as described in Galmarini et al. (2012, 2015, 2017).

On the other hand, several previous studies had evaluated modelled aerosol optical properties against satellite data from a global view. In Ghan et al. (2001), simulated AOD were within a factor of 2 of AVHRR (Advanced Very High Resolution Radiometer) products and the behaviour of the Ångström Exponent (AE), estimated from POLDER (POLarization and Directionality of the Earth's Reflectances) and SeaWiFS (Sea-Viewing Wide Field-of-View Sensor), was similar to that simulated. Otherwise, both the simulated AOD in Chin et al. (2002) and Reddy et al. (2005) were reproduced with most of the notable features in TOMS (Total Ozone Mapping Spectrometer), AVHRR and MODIS. Moreover, Ginoux et al. (2006) revealed sensitivity to humidity when evaluated against satellite data. Kinne et al. (2003) compared aerosol modules from seven models with MODIS and TOMS, and found large discrepancies over tropical and Southern Hemisphere oceans due to sea salt treatment. Kinne et al. (2006) also discovered a lower simulated AOD among 20 different modules from the AEROCOM Project (0.11 to 0.14) than the satellite AOD composite of MODIS, MISR (Operational Microwave Integrated Retrieval System), AVHRR, TOMS, and POLDER retrievals (0.15).

More recent studies are Colarco et al. (2010), who assessed simulated AOD *versus* MODIS and MIRS, and found similar seasonal and regional variability and magnitude over downwinds of the Saharan dust plume, a high bias in sulphate-dominated regions of North America and Europe, and a better agreement over ocean when the sea salt burden was reduced by a factor of 2. Furthermore, Zhang et al. (2012) reported a relative difference in AE of 13.8% with a negative (positive) bias over high latitude regions (oceans), but correlated well for AOD in comparison with MODIS. Finally, Liu et al. (2012) evaluated long-term simulations compared with the satellite composite derived by Kinne et al. (2006) and identified a low bias for AOD, but a good representation of the observed geographical and temporal variations of aerosol optical properties.

Similar studies to ours, which made a seasonal comparison over Europe, are: Jeuken et al. (2001), in which the comparison of AOD calculations with ~~ARST-2~~ ATSR-2 (Along Track Scanning Radiometer 2) on board the European ERS-2 satellite showed an average difference of 0.17-0.19, but a good representation of the observed patterns; Solmon et al. (2006), in which the simulated AOD presented a general underestimation (more pronounced over the Mediterranean Basin), but within the range of AERONET and MIRS over northern Europe, and followed spatial patterns of MODIS and TOMS over both Europe and Africa.

~~As the above-mentioned aerosol optical properties influence ARI and ACI, a good representation of them is, thus, a key issue to reduce the uncertainty of aerosol effects on the Earth's climate system.~~ Curci et al. (2017) used AQMEII Phase 3 simulations to evaluate the black carbon absorption against AERONET but no works have evaluated against satellite the seasonal representation of optical properties over Europe by the regional models involved in AQMEII Phase 3. This represents an added value because three main reasons: 1) all the regional models evaluated here were run using the same boundaries and initial conditions which permit investigate the importance of different processes and feedbacks in each models; 2) the use of two different emissions data set permits evaluated the influence of these in the aerosol optical properties representation; and 3) the use of online coupled chemistry-meteorology/climate models (as were some of the used here) will permit to investigate the influence of the ARI and ACI. As above-mentioned, aerosol optical properties influence ARI and ACI, a good representation of them is, thus, a key issue to reduce the uncertainty of aerosol effects on the Earth's climate system. For this reason, our main study aim was to evaluate the representation of two main aerosol optical properties, AOD and AE, using the models of the AQMEII Phase 3 initiative over Europe. The evaluation was made by using the remote-sensing observations from the MODIS sensor and from AERONET and MAN (Maritime Aerosol Network). Section 2 provides a brief description of the observational and models data, and the evaluation methodology. Section 3 presents the evaluation results. Finally, Section 4 summarises the main conclusions reached.

2 Methodology

In this work, we focused on evaluating aerosol optical properties representation, AOD and AE, over Europe throughout the year 2010. The evaluation was made using remote sensing data from the MODIS sensors on board the Terra and Aqua satellites and AERONET and MAN ground-based networks.

2.1 Model simulations

The evaluated simulation data were from the regional chemical-meteorology simulations made over Europe within Phase 3 of the AQMEII framework.

Two different emission data were used. On the one hand, the so-called HTAP_v2.2 (referred to from this point onwards as HTAP emissions). These data were harmonised by the Joint Research Centre's (JRC) Emission Data Base for Global Research (EDGAR) team in collaboration with regional emission experts from different agencies from the United States, Europe and Asia. HTAP emissions covered the years 2008 and 2010, with yearly and monthly time resolutions, and a global geo-coverage with a spatial resolution of $0.1°$. The chemical species were $\mathrm{SO_2}$,$\mathrm{NO_x}$, NMVOC, $\mathrm{CH_4}$, CO, $\mathrm{NH_3}$, $\mathrm{PM_{10}}$, $\mathrm{PM_{2.5}}$, BC and OC at the sector-specific level, and there were seven emission sectors (air, ships, energy, industry, transport, residential and agriculture) (Janssens-Maenhout et al., 2015; Galmarini et al., 2017). There were also the so-called MACC emissions (Pouliot et al., 2015), which have been previously used for AQMEII Phase 2 (Galmarini et al., 2015). The data set is a follow-on to the widely used TNO-MACC database (Pouliot et al., 2012), with a base resolution of $\sim 7km$. The provided species were: $\mathrm{CH_4}$, CO, $\mathrm{NO_x}$, $\mathrm{SO_x}$, NMVOC, $\mathrm{NH_3}$, $\mathrm{PM_{coarse}}$, $\mathrm{PM_{2.5}}$. A separate PM bulk composition profile file was composed, based on information per source sector per country. The different represented chemical components were EC, OC, $\mathrm{SO_4^{-2}}$, sodium and other mineral components. Fire emissions were included but volcanic and dimethyl sulphide emission (DMS) were not considered (Galmarini et al., 2017).

The study period was 2010 and the study domain was Europe. A detailed description of the simulations can be found in Solazzo et al. (2017). However, a brief description of them which focuses on aerosol treatment is provided below, and this information is summarised in Table 1.

The FI1 simulations were made at the Finnish Meteorological Institute (FMI), and the only difference between both FI1 simulations was the type of emissions used (HTAP or MACC). The System for Integrated modeLling of Atmospheric coMposition (SILAM), version 5.4. (Sofiev et al., 2015), was employed with the meteorological input extracted from the European Centre for Medium-Range Weather Forecasts (ECMWF). Sea salt emissions were included, as in Sofiev et al. (2011) (but not from boundaries), as were biogenic volatile organic compounds (VOC) emissions as in Poupkou et al. (2010) and wild-land fire emissions as in Soares et al. (2015). Wind-blown dust was included only from lateral boundary conditions. Gas phase chemistry was simulated with Carbon-Bond Mechanism-IV (CBM-IV), and with updated reaction rates according to IUPAC recommendations (http://iupac.pole-ether.fr) and JPL (http://jpldataeval.jpl.nasa.gov). Secondary inorganic aerosol (SIA) formation was computed with the updated DMAT scheme (Sofiev, 2000) and secondary organic aerosol (SOA) formation with the Volatility Basis Set (Donahue et al., 2006, VBS). AOD in SILAM was computed assuming external mixture of spherical particles, taking into account their hygroscopic growth. The optical properties used in the Mie computations originate from the OPAC dataset (Hess et al., 1998).

The ES1 simulation was run by the Regional Atmospheric Modelling Group at the University of Murcia (UMU, Spain). They used the Weather Research Forecasting model online coupled with Chemistry (Grell et al., 2005, WRF-Chem), version 3.6.1. Meteorological inputs were driven by ECMWF analysis fields. The aerosol module based on the Modal Aerosol Dynamics Model for Europe (Ackermann et al., 1998, MADE), in which secondary organic aerosols (SOA) were incorporated by the Secondary Organic Aerosol Model (Schell et al., 2001, SORGAM), was used. The employed gas phase chemistry model was the Regional Acid Deposition Model, version 2 (Stockwell et al., 1990, RADM2), with 57 chemical species and 158 gas phase reactions, of which 21 are photolytic. Anthropogenic emissions were MACC emissions. Biogenic VOC emissions were computed by applying the Model of Emissions of Gases and Aerosols from the Nature (MEGAN) emissions model (Guenther, 2006), version 2.04. The MADE/SORGAM sea salt (Gong, 2003) and dust (Shaw et al., 2008) emissions were used.

The IT1 simulation was made at the Ricerca sul Sistema Energetico (RSE, Italy) using the Weather Research Forecasting (WRF) model coupled with the Comprehensive Air Quality Model with Extensions (CAMx), version 6.10. Meteorological inputs were generated using ~~WRF-Chem~~ WRF version 3.4.1. Anthropogenic emission were MACC and biogenic were computed by MEGAN. WRF-Chem was adopted to predict GOCART (Goddard Chemistry Aerosol Radiation and Transport) dust emissions (Ginoux et al., 2001) along with meteorology. Sea salt emissions were computed using published algorithms (de Leeuw et al., 2000; Gong, 2003). The WRF-CAMx pre-processor (ENVIRON, 2014, version 4.2) was used to create the CAMx ready input files by collapsing the 33 vertical layers used by WRF to 14 layers in CAMx, but by maintaining the layers up to 230 m above ground level identical. Aerosol optical properties were estimated by means of the Aerosol Optical DEpth Module (Landi, 2013, AODEM) post-processing tool that was coupled to CAMx regional model. AODEM calculated the optical properties (e.g. AOD, extinction and scattering coefficients, and particle number concentrations) at different wavelengths and size bins starting from the aerosol mass concentration predicted by CAMx. In this work, the Mie theory was applied by dividing the size range (40 nm to 10 μm) into 10 bins, and calculating the hygroscopic growth of each aerosol species in each bin with the Hanel formula. Moreover, particles were assumed to be internally mixed.

The IT2 simulations were run at the University of L'Aquila (Italy) using WRF-Chem (Grell et al., 2005), version 3.6. The modified MADE/VBS aerosol scheme (Tuccella et al., 2015) was included in this version. This scheme is based on MADE to treat inorganic aerosols, along with the VBS approach (Ahmadov et al., 2012). MADE/VBS allows a better representation of the SOA mass. The Regional Atmospheric Chemistry Mechanism - Earth System Research Laboratory (RACM-ESRL) gas phase chemical mechanism (Kim et al., 2009) was used. Anthropogenic emissions were MACC emissions which had been adapted to the chemical mechanism used following the method of Tuccella et al. (2012). As in the IT1 and ES1 simulations, biogenic emissions were calculated online by the MEGAN model (Guenther, 2006). Finally, the meteorological analyses used to initialise WRF were provided by the ECMWF with a horizontal resolution of 0.5° every 6 h. IT2_M-ARI was run with ARI, while large-scale clouds were solved by a simple module. IT2_M-ARI+ACI took into account ARI and ACI, while aqueous chemistry was solved in convective clouds. As well as ES1, IT2 simulations used the MADE/SORGAM sea salt and dust emissions.

WRF-Chem simulations, ES1 and IT2, determined aerosol optical properties according to wavelength following Fast et al. (2006), Barnard et al. (2010) and Chapman et al. (2009). The composite aerosol optical properties were determined by the Mie theory, summation over all the size bins and wet particles diameters. An overall refractive index for a given size bin, as determined by an volume averaging of complex indexes of refraction associated with each chemical constituent of the aerosol, was used. The inclusion of ACI and ARI in WRF-Chem is described in Chapman et al. (2009).

Table 1. Model simulations

Model Code	**Insti-tution**	**Meteorolo-gical model**	**Dispersion model**	**Emissions**	**Aerosol mech. ~~mechanism~~(dust sources)**	**AOD/AE estimation**	**Gas Phase mechanism**	**Resolution (XY, Z)***
FI1_HTAP	FMI	ECMWF	SILAM v.5.4.	HTAP	DMAT -VBS ~~VBS~~ (boundaries)??	??	CBM-IV	0.25°, 12 uneven level below 13km (1^{st} to $\sim 30m$)
FI1_MACC				MACC				
ES1_MACC	UMU	WRF	~~WRF-~~WRF-Chem v3.6.1	MACC	~~MADE-~~ MADE-Sorgam (online+boundaries)	prognostic /diagnostic	RADM2	23km, 33 levels ~~u~~ up to 50hPa (1^{st} to $\sim 21m$)
IT1_MACC	RSE	~~v.3.4~~WRF v.3.4	~~v6~~CAMx v6.10	MACC	Coarse -Fine (online+boundaries)	diasnostic	CB05	23km, 14 levels ~~u~~ up to 8km (1^{st} to $\sim 25m$)
IT2_M-ARI	UAq	WRF	WRF-Chem v3.6	MACC	(ARI) MADE/VBS	prognostic	RACM-ESRL (Aq. conv. clouds)	23km, 33 level up to 50hPa 12 below 1km (1^{st} to $\sim 12m$)
IT2_M-ARI+ACI					(ARI+ACI) (online+boundaries)	/diagnostic		

FMI (Finnish Meteorological Institute, Finland), UMU (University of Murcia, Spain), RSE (Ricerca sul Sistema Energetico, Italy), UAq (University of LÁquila, Italy)
*XY: Horizontal resolution; Z: Vertical resolution.

A multimodel ensemble (henceforth referred to as ENSEMBLE) of the available simulations was also evaluated. The results presented herein did not intend to represent an ensemble of opportunity, but were merely calculated as the mean of all the participating simulations. As part of the AQMEII Phase 3 initiative, the available variables of aerosol optical properties were AOD at 470, 550 and 670 nm.

2.2 Observational Data

The used observational data came from the twin MODIS (Moderate Resolution Imaging Spectroradiometer) sensors. These instruments, aboard the Terra (MOD04_L2) and Aqua (MYD04_L2) satellites, provide information about aerosol optical properties around the world. Moreover, and in order to a reliable and complete analysis, we used ground-base observation from all the available station of AERONET (Aerosol Robotic Network,https://aeronet.gsfc.nasa.gov/) and the available data from MAN (Maritime Aerosol Network, https://aeronet.gsfc.nasa.gov/new_web/maritime_aerosol_network.html) as as a component of AERONET.

The used data from MODIS were Level 2 of the Atmospheric Aerosol Product (MxD04both MOD04_L2 and MYD04_L2) from the collection 6 (C6), with a resolution of 10 km which are estimated by two different algorithms, Dark Target (DT) and Deep Blue (DB). The used variables were: (1) a "combined" variable of the DT and DB algorithms which provide information about AOD at 550 nm for both ocean and land; and; (2) AE between 550 and 860 nm over the ocean estimated by the DT algorithm. Although There are several evaluations of this "combined" AOD product has not yet been validated, it offers a wider coverageproducts of MODIS C6 against AERONET sites around of the world (Sayer et al., 2014; Mhawish et al., 2017; Bilal et al., 2018) . All of these established that a high percentage of retrievals are within the estimated error (EE) of the DT and DB algorithms, which is ($\pm 0.05 + 15\%$) (Levy et al., 2013). Moreover, in Sayer et al. (2014) and Bilal et al. (2018) the performance of combined retrievals outperformed DT or DB retrievals in term of correlation (around 10%), meanwhile showed relative mean bias values similar at a global scale. The preliminary estimated error (EE) for the used AE products was 0.45 in the pixels with an AOD $>$ 0.2 (Levy et al., 2013). The selection of this observational data was due to results found by ?. Palacios-Peña et al. (2018) where they evaluated the uncertainty in the satellite representation by comparing MODIS, OMI (Ozone Monitoring Instrument) and SeaWIFS (Sea-viewing Wide Field-of-view Sensor) AOD retrievals against AERONET. They found that MODIS presented the best agreement with the AERONET observations compared to other satellite AOD observations during two studies with high aerosol load took place in 2010 all over Europe.

As Terra and Aqua are in Sun-synchronous orbits around the Earth, MODIS does not provide data over the entire studied domain for each time step. According to Levy et al. (2013), who have established that there is no significant difference between MODIS/AERONET comparability for Terra and Aqua data, we combined the hourly data from both satellites in order to obtain a whole year of data with a wider coverage for each time step than by using the Terra and Aqua data separately.

From AERONET, AOD at 675 nm and AE between 440 and 870 nm retrievals of level 2.0 from the available European stations during the entire 2010 year were used. For this network data, the total uncertainty for the AOD data under cloud-free conditions is established as $< \pm 0.01$ for $\lambda > 440$ nm and $< \pm 0.02$ for shorter wavelengths (Holben et al., 1998). The same variables were used from MAN, whose estimates uncertainty of AOD in each channel is, as well as AERONET, $< \pm 0.02$ because MAN is affiliated with the AERONET calibration and data processing standards and procedures (Smirnov et al., 2009) .

2.3 Evaluation Method

Simulations (Table 1) and satellite data had a different spatial resolution. Henceforth and beforehand, all the data were preprocessed and bilinearly interpolated to a common working grid with a resolution 0.25°.

As mentioned above, our objective was to evaluate the representation of the main aerosol optical properties: AOD and AE. MODIS provides AOD at 550 nm and AE between 550 and 860 nm, but from simulations, the available variables were AODs at different wavelengths. Thus in order to evaluate AE from simulations, this variable had to be estimated through the Ångström empirical expression (Ångström, 1929, eq. 1), where λ is the wavelength and β is Ångström's turbidity coefficient.

$$AOD = \beta\lambda^{-AE} \tag{1}$$

By rationing equation 1 at two different wavelengths and taking algorithms, AE can be computed from the spectral AOD values (Eck et al., 1999, eq. 2). Hence it is possible to estimate AE between two known wavelengths, and to also use this AE to estimate AOD at other different wavelengths. However, as established Ignatov et al. (1998), retrievals of AE under AOD conditions lower than 0.1 are highly uncertain. For this reason, we chose the criteria to estimate AE over areas with AOD > 0.1. Moreover, and according to the EE for the AE products of MODIS, we set the AE values range between -0.5 and 4.0. It is widely known that AE values spread from 0 to 4 and even sometimes, when really coarse particles are presented, they can reach negative values. Then, we choose AE values between -0.5 as lower limit in order to cover possible negative values in a close smoothing value to the EE for the AE products of MODIS.

$$AE = -\frac{ln(\frac{AOD_{\lambda_2}}{AOD_{\lambda_1}})}{ln(\frac{\lambda_2}{\lambda_1})} \tag{2}$$

Once all the data had the same resolution, and following Equation 2, the simulated AOD and AE were at the MODIS wavelengths. Then the hourly data were evaluated using classical statistics such as: the mean of the individual model-prediction error or bias (e_i); i.e., the mean bias error (MBE); the mean of the absolute error (MAE); and the coefficient of determination (r), according to Willmott et al. (1985), Weil et al. (1992) and Willmott and Matsuura (2005). Before the statistical analysis

All the observations used in this work are not provided in a temporal regularly way. This means that the number of occurrences in each of the pixel for satellite data or in each station for AERONET data are not the same. As the results in this work are represented as seasonal means and in order to show robust means estimated with a reasonable number of occurrences, a mask that showed the areas those pixels (stations) where the satellite observations were higher than (station) occurrences were higher that the 10% of their maximum the maximum possible occurrences, was implemented. This mask was implemented in order to carry out a statistical analysis with a number of observations which show more confident resultsThe maximum possible occurrences for satellite data were selected as the maximum of occurrences over the studied season (JFM, AMJ, JAS or OND) and the domain. Figures S1 and S2, in the appendix, show the number of observations used when the mask was implemented. For AERONET, the maximum possible occurrences was established as the maximum of solar-light hours in each station during each season.

The MBE was estimated as in Equation 3, where i represents each time step, P is the simulation and O is the observational value. MBE provides an idea about the behaviour of the models, and indicates whether the model over- or underestimates the variable values measured by the satellite sensor.

$$MBE = n^{-1}\sum_{i=1}^{n} e_i = \overline{P_i - O_i} \tag{3}$$

The MAE was calculated as in Equation 4 and provides an estimation of the magnitude of the error independently of over- or underestimation.

$$MAE = \langle n^{-1}\sum_{i=1}^{n} |e_i| \rangle = \overline{|P_i - O_i|} \tag{4}$$

The temporal determination coefficient was estimated as in Equation 5 and was used as a measure of the strength of the linear relationship between two variables, in our case, the satellite and simulations values.

$$r^2 = \langle \frac{n^{-1}\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sigma_O \sigma_P} \rangle^2 \tag{5}$$

Finally, the Kernel probability density functions (PDF) with a broadband of 0.05 were used to evaluate the skills of the simulations to reproduce the variability (temporal and spatial) of the studied variables (AOD and AE).

3 Results

The next section evaluates the skills of the different AQMEII Phase 3 simulations with respect to the representation AOD and AE representation. The first section shows the model evaluation for the AOD representation and the second for AE. The numerical result of each case ~~is~~ for MODIS (M) and AERONET(A) are indicated by the numbers represented in each figure. Finally, the skills of the simulations to reproduce the variability of the studied variables (AOD and AE) are analysed using the PDF of each variable.

3.1 Model evaluation of the AOD representation

AOD is defined as the integrated extinction coefficient over a vertical atmospheric column and indicates to what degree aerosols avoid light transmission. AOD is not a direct function of the atmospheric load of particles, but can provide us an approximate idea of both atmospheric load of particles and the interaction of these particles with radiation.

First, temporal means of AOD at 550 nm values from a combination of the two MODIS satellites and of AOD at 675 nm from AERONET station. MAN data are displayed as diamond linked by a colored line, each color represented the track of a boat and they show instantaneous observations. All of these observations were analysed in the first row of Figures 1, 2 and

3. The temporal ~~mean for the whole year is show in figures in the supplementary material (SM). The seasonal break down is presented besides named~~ seasonal means and corresponding MAM data are presented in the named column. JFM represents the temporal mean for January, February and March (from now on winter); AMJ for April, May and June (spring); JAS for June, August and September (summer); and finally, OND for October, November and December (autumn).

~~The temporal mean of AOD (Figure S3 at SM) for the whole year 2010 presented the highest values (between 0.5 and 1) over Russia and the surrounding areas, and over the south of the domain and the Mediterranean Region.~~

When seasonal figures were analysed (Figure 1), the highest values (around 1) were found over the southern part of the domain for all seasons due to frequent Saharan desert dust outbreaks, which affects the Mediterranean Region. Moreover, these desert dust outbreaks were more frequent and strong for spring and summer~~because the mean AOD values were higher.~~

~~In spring, the highest temporal mean AOD values, between 0.5 and 1,~~ ; reached the southern part of the domain ~~. These mean values can be higher than 1 over specific areas.~~

~~In~~ with AOD values above 0.4. But in summer, the ~~temporal mean AOD values over the southern part of the domain fell between 0.5 and 0.8, but the~~ higher mean AOD values (above ~~1.5~~1 for MODIS values) were found over a large area in Russia and its surrounding areas ~~. During this period,~~ due to a heat wave and wildfires occurred over this area~~, which explained this fact. For these two reasons, this season presented the highest mean AOD values when both space and time were considered.~~

~~However for~~ . However in autumn and winter, high mean AOD values were also found over the southern part of the domain, but were lower than ~~0.5~~0.4. The lower mean AOD value when considering space and time was found in autumn. It is noteworthy that AOD satellite values agree on those values displayed by the available AERONET station and MAN data. The gap over the northern part of the terrestrial domain in winter and autumn is explained because ice, snow and clouds were avoided for the MODIS sensor, and aerosol properties were not retrieved over the areas where they were present (http://darktarget.gsfc.nasa.gov/)~~, which explains the gap over the northern part of the terrestrial domain in winter and autumn. The gaps~~ . The gap in the rest of the season are explained by the implemented mask, which is explained in the Observational Data Section.

Moreover, as the number of solar light hours is lower in the north during winter and autumn, this affected also to the number of AERONET station with available data which explain the lack of AERONET data because our criterion of number of occurrences equal or higher 10% of the maximum of solar hours was not passed for some AERONET station. Throughout the seasons, high AOD values were obtained over the south-eastern part of the domain; Syria, Iraq, Kuwait and the Persian Gulf.

According to MBE, all the simulations spatially presented a similar behaviour ~~throughout the year (SM) and~~ in different seasons (Figure 1). The main feature was an overestimation of AOD over the southern part of the domain, the main area affected by desert dust outbreaks; and an underestimation (to a greater or lesser extent) over the Russian area, affected by the wildfire emissions in summer. One of the most remarkable issue in this evaluation is that MODIS and AERONET displayed similar results when the simulations are evaluated in front of each one. Sometimes, the numerical result of each case, represented as the temporal and spatial mean of the results, were lower for AERONET, but this could be because AERONET station covered a total space lower that MODIS and they are not located over the main affected areas by over-underestimations. MAN results could be not as similar as because they were displaying instanstaneus results and not temporal mean as did MODIS and AERONET.

Throughout the yearAs a general behaviour during all the seasons, FI1_HTAP and FI1_MACC, which use the ECMWF model for meteorology and SILAM for chemistry, showed a slight overestimation of the AOD values. Thus, these simulations gave high AOD values than the other models and this could be because SILAM is known to have slower dry particle deposition. This could explain that, although the band quiet crudely represents size distribution, AOD is also very sensitive.These values were overestimated over the southern part of the domain (northern part of the Saharan Desert), with values around 0.1. These values were spatially consistent with the higher MAE values (see second column of FigureS3 at SM). Over north-western areas and Russia (due to the wildfires that occurred in summer), these simulations underestimated AOD (negative MBE values around -0.05). Figure 2). One main issue is that no clear differences were found when HTAP and MACC emissions were used.

Idem Figure 1 for the MAE results of AOD at 550nm satellite values *vs.* simulations.

The other simulations; ES1_MACC, IT1_MACC, IT2_M-ARI and IT2_M-ARI+ACI; used the WRF meteorological model. When a different chemistry model was employed, minor differences in the error of simulations were found between those made using the CAMx chemistry model (IT1_MACC) and the WRF-Chem (IT2 simulations). These differences were of a similar order of magnitude to that of the differences between the IT2 simulations by including ARI and ACI. However, the ES1_MACC simulation which, like the IT2 simulations, used WRF-Chem, presented remarkable differences by displaying a strong overestimation of AOD over the southern areas of the domain. This marked overestimation took place because of a bug in the used dust scheme, which lacks the gravitational settling. Although the IT2 simulations used the same dust scheme and model version, the dust flux was modified for these simulations to estimate accurate dust concentrations. Thus ES1_MACC showed the higher MBE and MAE values throughout the year . Despite this overestimation, a slight underestimation was seen over Russian areas (negative MBE values around -0.05).

and for both, MODIS and AERONET (see labels in Figures 1 and 2). IT1_MACC presented a general weak overestimation of AOD over the whole domain. This simulation over Russia underestimated the AOD values (negative MBE values around -0.1), which was due mainly to the wildfires that occurred in summer. The IT2 simulations displayed a different behaviour. These simulations presented a general weak underestimation over the whole domain(MBE of -0.03 and -0.05 for ARI and ARI+ACI, respectively), , except over the southern part of the domain (areas affected by the Saharan dust outbreaks), where AOD was overestimated with low values. The underestimation of AOD was stronger over the northern part of the domain (particularly over Russian areas) and reached mean values of around -0.2. The IT1_MACC and the IT2 simulations presented the lowest absolute error values .

(see labels in Figures 1 and 2). The ENSEMBLE notably overestimated the AOD values over the southern part of the domain, with very high values over the northern part of the Saharan Desert, which is consistent with the higher MAE values obtainedfor the whole year of ENSEMBLE. The underestimation was milder than the overestimation, and pointed it out over the Russian areas affected by wildfires.

. One notable point in the underestimation was presented by all the simulations over the south-eastern part of the domain (represented as a blue spot), centred over Iraq. The ES1_MACC simulation did not show this underestimation because of its high AOD values, but presented lower overestimation values (close to 0) over this area than over its surroundings. This small spot was also notable when MAE (Figure 2) was analysed. This can be explained by the fact that the emissions inventories used herein only covered European areas (see the Figure S7 in the SM), thus the emissions over that area were not considered. Moreover, all the simulations through the seasons overestimated the AOD over the southern part of the domain. This was related mainly to the high dust concentrations according to the boundary conditions because Solazzo et al. (2017) found that the error in primary species as dust was strongly affected by the emission and boundary conditions in the AQMEII Phase 3 simulations.

Idem Figure 1 for the determination coefficient of AOD at 550nm satellite values *vs.* simulations.

In winter (the first column in Figures 1 and 2), all the simulations presented a weak underestimation over the Atlantic Ocean, except for IT1_MACC, which presented a weak overestimation in the northern part (MBE mean of 0.03MODIS mean of 0.01). The above-mentioned blue spot was clearly defined over a small south-easterly area and was stronger during this season, even for the ES1_MACC simulation with negative MBE values. For the IT1_MACC, IT2_M-ARI and IT2_M-ARI+ACI simulations, the higher MAE values were consistent over the last mentioned area. The FI1 simulations presented an overestimation over North Africa. This area was larger with a stronger overestimation (mean MBE MBE MODIS mean of 0.23 and AERONET mean of 0.07) for the ES1_MACC simulations for the same reason explained above. All year long, these simulations showed higher MAE values (a mean of 0.28). The ENSEMBLE presented an intermediate behaviour, with milder MBE and MAE values (0.02 and 0.12, respectively for MODIS; and 0 and 0.06 for AERONET): an overestimation of the AOD values over North Africa, a very weak underestimation over the Atlantic Ocean and the blue spot centred over Iraq and Syria.

In spring (the second column in Figures 1 and 2), the underestimation of AOD was similar to that in winter, but with steeper values. All the simulations presented an overestimation (with different degrees) over the southern part of the domain, the Balkan Peninsula and southern Russia. This overestimation was larger and stronger for the ES1_MACC simulation (MBE MODIS mean of 0.21 and AERONET mean of 0.13) and once again presented higher MAE values (0.29, MODIS; 0.19, AERONET). All the simulations, except for IT1_MACC, presented a weak underestimation over the Atlantic Ocean. The IT simulations gave fewer errors than the rest. As in winter, a small south-easterly area (the blue spot) appeared, but was consistent with the maximum MAE values for the IT simulations.

The underestimation of AOD due to the wildfire emissions over Russia and the surrounding areas was one of the most important issues in summer (the third column in Figures 1 and 2). This underestimation was larger and stronger for the IT2 simulations, and was smaller and weaker for the FI1 simulations. Moreover, the aforementioned small area in the south-eastern part of the domain presented higher underestimation values over a larger area than during the other seasons, and reached as far as the Persian Gulf. The overestimation behaviour was the reverse of the underestimation; the FI1 simulations presented higher values and the IT2 ones gave lower values. While the overestimation was stronger and affected a large area than during another season, this time the higher overestimation values were found over the north-west areas of Africa and the Iberian Peninsula. As during the other seasons, the ES1_MACC simulation showed the stronger and larger overestimation. During this season, with higher AOD values, all the simulations presented the highest error values. The ENSEMBLE represented the most relevant behaviour of MBE and MAE, which means that the important evaluation results were found over those areas where other simulations presented a notable issue (mainly the south-western part of the domain, Russia and the surrounding areas, and the "blue spot"). As established in Palacios-Peña et al. (2018), this underestimation may be due to a misinterpretation of the simulation of the aerosol vertical and may, therefore, be due to the AOD representation given the understated injection height of the total biomass burning emissions found for the MACC emissions by Soares et al. (2015). A different hypothesis ascribes this underestimation to underestimated emissions. Toll et al. (2015b) found that while the daytime plumes from large fires were indeed lifted higher, the night time emissions and emissions from small fires were injected closer to the ground, making the average smoke transport distance even smaller than for the fixed emission height. Also Soares et al. (2015) point out, referring to Wooster et al. (2005), that MODIS is not sensitive enough to register the fire radiative power of small or smoldering fires, and thus large fraction of those is missed in the emission data, including also strongly emitting peat fires. The 2010 Russian fires included some huge fires, but also numerous small ones over large areas, and a large fraction of those was probably missed by MODIS.

In autumn most of the domain presented error values that came close to 0 for all the simulations. Thus autumn was the season when the lowest error values were found. All the simulations showed an overestimation that came close to the south boundary and an underestimation over Tunisia and Algeria. Both the overestimation and underestimation were weaker for the IT simulations than for the FI1 ones. ES1_MACC was the only simulation that displayed a different behaviour during this season with a high overestimation over almost all the domain (0.25, MODIS; 0.10, AERONET).

Regarding the coefficient of determination, only the areas where the results were significant at 90% are shown. All the all the simulations presented values above 0.5 over most of the domain . Throughout 2010 (Figure S3 at SM) the values of this statistics were lower than when the seasonal breakdown was analysed (Figure 3). ENSEMBLE presented the highest values and over a largest area (in detail over Russian and the surrounding areas, and over the south-western part of the domain), while the ES1_MACC simulation presented the smallest area and, thus, the lowest mean value.

when are compared against MODIS. In winter the highest determination values (close to 1.0) were found over the north-eastern part of the African continent. In spring, these high values were found over central and eastern parts of Europe and North Africa. In summer, the highest values were mainly over Russian and the surrounding areas and a part of the Atlantic Ocean in the south-western part of the domain. Finally in autumn, the coefficient of determination values were lower than for the other seasons, and were mainly over the Mediterranean sea and the Atlantic Ocean.

3.2 Model evaluation of the AE representation

AE is a parameter that indicates the relationship between the size of the particles suspended in the atmosphere and the wavelength of the incident light, and, although there is not a direct correspondence between aerosol size and AE, provides an idea of the size of particles. Low AE values are related to coarse particles, such as desert dust or sea salt, and high values are associated with fine particles, such as anthropogenic source particles or biomass burning. The AE values are between 0 (or even slightly negative in coarse mode aerosol episodes) and 4 (Boucher, 2015). AE models simulations are less than for AOD because some models did not provide AOD at different wavelengths and then, it was no possible to estimate AE following the methodology stablished above.

Temporal mean of AE between 550 and 860 nm satellite values, which are only estimated over the sea , is and the temporal mean of AE between 440 and 860 nm from the available AERONET station and MAN data, are showed in the first row in the AE figures (Figures 4, 5 and 6and S4 at SM).

Idem Figure 4 for the MAE results of AE between 550 and 860nm satellite values *vs.* simulations.

Throughout the whole year (Figure S4 at SM)). Generally, through the different seasons, low AE values (below 0.5) were found offshore, where sea salt particles (coarse) predominated. Over the Mediterranean coast near the Saharan desert, low values were found due to the frequency of desert dust outbreaks. High values (between 1.5 and 2.0) were obtained over coasts and inland in central Europe. These values were due mainly to anthropogenic emissions, such as traffic road, which presents fine particles. Moreover, these values lowered from the coast inland to offshore. This pattern also showed when seasons were separately evaluated. Two small areas over the north of the Black and Caspian sea, showed values around 2.5.

In winter (JFM, the first column in the first row in the AE figures) and autumn (OND, fourth column), the lowest values were found over the Atlantic Ocean and the Mediterranean Sea. In the same wayas throughout the year, high AE values (around 1.5) were shown over coasts and inland in central Europe. In autumn, a small area over the north of the Caspian sea with values of 2.5 was found.

In spring, as represented in the second column (AMJ) in the first row in the AE figures, the AE values presented a narrow range between 1.0 and 1.5 over most of the domain. Some exceptions were values that came close to 0.5 near the African continent, and values close to 2.0 over coasts in central and in north Europe. It is noteworthy that low AE values (close to 0.5) were uniformly distributed in spring over the southern part of the domain, while in summer (JAS, the third column in the first row in the AE figures) the lowest AE values (lower than 0.5) were found mainly over the southern Atlantic Ocean. Values between 2.0 and 2.5 were estimated over north-eastern coasts and over central and north Europe and the north of the Black and Caspian sea, which were lower than in summer.

In this section, the simulations run with the available data were less than they were for AODAs well as happened with AOD, AERONET station and MAN data show values similar to those displayed by MODIS. As AERONET stations are loceted over the continent the temporal and spatial mean of the results in higher values due to a higher influence of anthropogenic emissions. The errors of the simulations made for the whole year are shown in Figure S4 at SM.

Throughout the yearOn a broad view through the different seasons, FI1_HTAP (the ECMWF meteorological model and the SILAM chemistry model) underestimated the AE over most of the domain(MBE of -0.60). This underestimation was higher over areas near European coasts and inland, where the satellite observations showed values of around 1.5, which were lower over the south-western part of the domain, where the satellite observations gave AE values that came close to 0.5. This simulation also presented the highest MAE values.

This model estimated larger-sized particles than that retrieved by observations. As aforementioned, SILAM crudely represents size distribution, which impacted the AE representation because it may have be centered on particles with a larger diameter. In spite of the results obtained for the AOD representation evaluation (due to the lack of dust gravitational settling), ES1_MACC is the simulation which presented the lowest presented low error values (MBE and MAE) over the whole yearthrough the different seasons.This could be explained by the high dust concentration over southern areas, resulting in low AE values and thus compensating the tendency for producing high $PM_{2.5}/PM_{10}$ ratios (Solazzo et al., 2012; Balzarini, 2013; Solazzo et al., 2014) . A very low overestimation was found over areas close to Africa, and a more notable underestimation was found over areas near the European coast (MBE of -0.1).

and inland. The IT1_MACC simulation generally overestimated the AE values over the Atlantic Ocean and the Mediterranean Sea (the areas with AE values that came close to 0.5). Over the areas near the coast of central and northern Europe, where the satellite observations gave values around 1.5, this simulation presented a weaker underestimation than in the other simulations.

The IT2_M-ARI+ACI simulation showed an overestimation over the Atlantic and Mediterranean coast near North Africa(MBE of 0.14), and a weak underestimation over the coasts of the North and Baltic Seas . This simulation along with and inland over the AERONET available station. IT_MACC (WRF coupled CAMx) and both WRF-Chem simulations (ES1_MACC , which used the and IT2_M-ARI+ACI) underestimated high AE values and overestimated low AE values and thus, they underpredicted the variability of this variable as was also found in Palacios-Peña et al. (2017, 2018). On the other hand, Solazzo et al. (2012); Balzarini (2013); Solazzo et al. (2014) found a severely underestimate for PM_{10} concentrations over Europe for WRF-CAMx and WRF-Chem model, gave the lowest valuesof MAE.

models, which could explain the overestimation of low AE values. Moreover, they also found an underestimation of $PM_{2.5}$ concentrations which could also explain the underestimation of high AE values since simulated particles underestimate the variability of the size. Finally, ENSEMBLE presented a notable underestimation of the AE values over the European Coast (including the Mediterranean Sea) and inland, probably due to the strong underestimation provided by the FI1_HTAP simulation. Very low overestimation values were obtained over the Atlantic Ocean and near African coasts in the south Mediterranean Sea. Moreover, ENSEMBLE and the other simulations presented a strong underestimation over the two small areas with AE values of around 2.5. AE values were underestimated by all the studies simulations when these were compared with AERONET station and MAN. One difference was "Polarstern Fall" boat of MAM during autumn which showed values close to 0.

In order to gain better knowledge of model skills, a seasonal study was conducted. Winter is represented in the first column in Figures 4 and 5. The FI1_HTAP simulation generally showed an underestimation of the AE values (-0.30 MODIS and -0.46 AERONET), which was stronger over areas near the European coast. The ES1_MACC simulations presented the lowest error values(0.14 MODIS and -0.32 AERONET) . This simulation displayed a weak overestimation of the AE values over the Atlantic Ocean and the Mediterranean Sea, and an underestimation over small areas close to the European coast and over AERONET stations. Both the IT1_MACC and IT2_M-ARI+ACI simulations gave a general overestimation over most of the domain. IT1_MACC showed a very weak underestimation close to the European coast, but this simulation had the highest MAE values due to the strong overestimation. However, IT2_M-ARI+ACI displayed really low bias (temporal and spatial AERONET MBE of 0) when is compared with available AERONET stations. ENSEMBLE gave high overestimation values over the North Atlantic and the Mediterranean Sea, and weak underestimation values close to the European coast and inland.

The second column of Figures 4 and 5 shows the results obtained in spring. For this season, FI1_HTAP underestimated the AE values over most of the domain and presented the highest error values (MBE of -0.62 and MAE of 0.64). against MODIS). This underestimation was larger when this simulations is compared with AERONET station (MBE of -0.99 and MAE of 0.99).

The ES1_MACC simulation displayed a behaviour somewhere between the other simulations, with a weak overestimation over the North Africa coast and a more notable underestimation over the northern part of the domain. Notwithstanding, the ES1_MACC simulations presented the lowest absolute error ~~values~~value when was compared with MODIS. For AERONET comparison, the lowest values was found for T2_M-ARI+ACI, but his is probably due to this simulation did not shown northerly stations. The IT1_MACC simulation overestimated the AE values over the Atlantic Ocean and the southern part of the domain, and the underestimation was found in areas over the European coast. The IT2_M-ARI+ACI simulation overestimated the AE values over the Moroccan Atlantic coast and the south of the Mediterranean Sea, but small areas of underestimation were found over the Azores Islands and the northern coast of France. Moreover, of the AERONET station over Europe overestimated these values. Finally, ENSEMBLE produced a general underestimation over most of the domain, for both MODIS and AERONET. The overestimation was produced mainly over an area that lies north of the British Isles, where satellite values came close to 0.5.

All the simulations made in summer (the third column in Figures 4 and 5) displayed similar skills as in spring. Generally speaking, FI1_HTAP underestimated the AE values and presented the ~~greatest~~ highest errors. During this season, ES1_MACC showed a larger area of underestimation and a smaller one of overestimation, but with similar error values as in spring. The overestimation of the IT1_MACC simulations was weaker, but the underestimation was stronger and over a large area over the North and Baltic Seas. The IT2_M-ARI+ACI simulation also produced an overestimation over most of the domain, but it was weaker than that presented in spring. Notwithstanding, this simulation presented a small area of underestimation over the Baltic sea. However, ENSEMBLE displayed a general underestimation that lowered from ~~the coast~~ inland to offshore.

In autumn (the fourth column in Figures 4 and 5), the behaviour of simulations was similar to that shown in winter. FI1_HTAP produced a general, but weaker underestimation than in spring and summer. During this season, ES1_MACC produced a general overestimation over the Atlantic ocean and the Mediterranean sea and an underestimation over the inland AERONET station, but once again it gave the lower error values. The IT1_MACC and IT2_M-ARI+ACI simulations overestimated AE values over most of the domain with similar MBE and MAE values but underestimated AE values in east AERONET stations. Finally, ENSEMBLE produced a weak overestimation over the Atlantic Ocean and the Mediterranean Sea, and a weak underestimation over the Black, Caspian and Red seas. ~~ENSEMBLE was the simulation with the lowest MAE error~~The values of all AERONET stations also were underestimated.

Figure 6 shows the results of the determination coefficient~~, which were significant at 90%. These results were worse throughout the year (column third in Figure S4 at SM) than when the different seasons were evaluated.~~ . FI1_HTAP and IT1_MACC showed relatively high values (around 0.5) over the Mediterranean Sea, but over this area, all the other simulations presented values above 0.25. ~~Even though the determination values were higher during the seasonal study, it~~ However, determination coefficient values were really low when simulations were compared with AERONET station and MAM data. It was very difficult to find a clear coefficient of determination pattern. During each season FI1_HTAP was used to present the highest determination values and ENSEMBLE the lowest ones.

3.3 Variability

A good approach to evaluate the spatial and temporal variability of a variable is the Probability Density Function (PDF). This represents the density of counts for each value of the variable. In order to study how the AQMEII Phase 3 simulations represented the variability of AOD and AE, the PDF of both variables for each studied season are shown in Figure 7. In that Figure, first left column corresponds to the ~~AOD~~ PDF of AOD at 550nm, second column of AOD at 675nm, third column of AE between 550 ad 860 nm and fourth of AE between 440 and ~~right to the AE PDF; first~~ 870nm. First row corresponds to winter (JFM), second to spring (AMJ), third to summer (JAS), and bottom row to autumn (OND). Observed ~~satellite values~~ values (MODIS in first and third column and AERONET in second and fourth) was represented by a black line; the ENSEMBLE by a red line; FI1 simulations by green dashed lines; ES1 by a yellow dashed line; IT1 by a cyan dashed line; IT2_M-ARI by a blue dashes line and finally IT2_M-ARI+ACI by a blue dotted line. Due to the small number of MAM occurrences this data are not shown in this section. PDF for MODIS and AERONET were evaluated separately because they were not represented the same variable and over the same space and time. However, they represented a similar behaviour regarding the comparison of the variability of the simulations against observations.

The PDFs of AOD for the data that corresponded to winter (JFM), spring (AMJ) and autumn (OND) presented a similar behaviour, both MODIS and AERONET. The observed ~~satellite~~ values showed a high probability for low values (between 0 and 0.5). The PDF of the IT1_MACC values for these seasons was the most similar one to ~~the observed MODIS~~ both observed values. For these three seasons, this was the simulation with a lower absolute error when the temporal standard deviation from the simulations was evaluated against observations, as we can see in the SM.

~~Simulations FI1 and IT2 displayed analogous PDFs with the highest probability for the lower AOD simulated valuesthan those observed~~ During autumn, AERONET data and their respective PDF from simulation are narrower than those from MODIS, so the AOD values, both observed and modelled, were lower over AERONET station. In winter (JFM) and autumn (OND), ~~the IT2 simulations presented higher probabilities~~ simulations FI1 and IT2 displayed analogous PDFs with the highest probability for the lower AOD ~~values than the FI1 simulations~~simulated values than those observed. However in spring (AMJ), these four simulations gave almost equal PDFs.

The ES1_MACC simulation showed a remarkable representation of AOD in all seasons when was compared with AERONET. The PDFs for this simulation estimated higher probabilities for the high AOD values than the other simulations and the observed values. For this reason, the probability of low AOD values was lower than for the rest. This behaviour was not observed when this simulation was compared against AERONET values, when the ES1_MACC PDF was similar to the rest. ENSEMBLE displayed in JFM, AMJ and OND a high probability for the lower AOD simulated values than those observed, but with ENSEMBLE, the probability for the higher AOD values was higher than for those observed.

The PDFs of the AOD representation were different in summer (JAS). For this season, both IT2 were the simulations that displayed the nearest behaviour to the observed PDF values. As seen in the SM, these simulations displayed the lowest MAE compared with the observed standard deviation. All the simulations and ENSEMBLE in this season presented a higher probability for the high AOD values than those observed.

The ~~second~~ third and fourth column of Figure 7 represents the PDFs for the ~~AE values~~MODIS and AERONET AE values, respectively. As for AOD, winter and autumn presented similar PDFs of MODIS observed values. The observed MODIS AE values showed a high probability for the low AE values, around 0.5, and a low probability for the high AE values. For AERONET, winter and autunm PDFs displayed high probability values for low AE values but these showed their highest probability for higher values than those observed in MODIS. For spring and summer, the PDFs for the MODIS observed values were tray-shaped, with a high probability for the AE values between 0.5 and 1.5. As well as that in autumn, these PDFs displayed their highest probabilities for AE values around 1.5. AERONET stations in spring and summer showed a probability which increased from AE values of 0 to high density for values from 0 to 2 where the probability decreased.

The behaviour of the other simulations and ENSEMBLE was similar in all seasons and showed a medium behaviour somewhere in the middle of the rest of the simulations. The FI1_HTAP simulation displayed a high probability for very low AE values, between 0 and 0.6 for MODIS and 0 and 0.4 for AERONET. IT1_MACC gave similar PDFs for all the seasons, and a high probability was found for the AE values between 1 and 1.5. FI1_HTAP and IT1_MACC were the simulations with the narrowest PDFs, which indicates that these simulations produced a strong underestimation of the observed variability of AE (as also indicated in the evaluation of the temporal standard deviation shown in the SM).

The PDFs for ES1_MACC, IT2_M-ARI+ACI and ENSEMBLE were wider than for the other ~~two~~ simulations. ES1_MACC ~~showed a high~~ and IT2_M-ACI+ACI showed a higher probability for the AE values ~~, between~~ from 0 ~~and 1.5.~~ to 2 for MODIS and 0 to 0.7 for AERONET. But IT2_M-ARI+ACI ~~for the AE values was between 0.5 and 1.5~~used to show a high probability for a slightly higher AE values than ES1_MACC. ENSEMBLE showed a high probability that ranged from 0 to 1.5 for MODIS and 0 and around 0.7 for AERONET. Notwithstanding, ~~these~~ all the simulated PDFs were narrower than the PDF for the observed values, thus all the simulations underestimated the representation of the AE values. This is observed in the SM, where the estimation of the MBE of the standard deviation gave negative results for all the seasons and simulations. It should be pointed out the low variability of the simulations over inland, that means over AERONET station, meanwhile during all the season AERONET displayed a PDF between 0 and 2 AE values, the evaluated simulations displayed PDF between o and 0.6 AE values which indicate that simulations displayed really low AE values inland.

4 Summary and Conclusions

Although AQMEII Phase 3 focuses on evaluating and intercomparing regional and linked global/regional modelling systems, an evaluation of the simulations ~~of the front~~ against observations was needed. Solazzo et al. (2017) analysed the performance of models for different meteorological variables and chemical species. In order to perform a more detailed analysis of the models' performance, this work focused on evaluating the aerosol optical properties representation by means of AQMEII Phase 3 simulations using satellite sensors. The evaluation of these variables is important because they strongly influence ARI and ACI and, thus, influence the atmospheric aerosol effect on the climate system.

As the Mediterranean Region is frequently affected by Saharan desert dust outbreaks, and an area over Russia and the surrounding during summer affected by wildfires, presented the highest AOD values for 2010. It shoul be pointed out the similarity between AOD values displayed by MODIS and AERONET when the simulations are evaluated in front of each one. When the representation of this variable was evaluated~~from the MODIS observation~~, generally all the simulations presented similar spatial patterns and gave a good representation of the low and medium AOD values. The lowest AOD underestimations for all the simulations were found over the Atlantic Ocean and, thus, sea salt emissions could be underestimated. However, a major underestimation occurred during the wildfires episode over Russia in summer. As established in ~~?~~Palacios-Peña et al. (2018) , this underestimation may be due to a ~~misunderstanding~~ misinterpretation of the simulation of the aerosol vertical and may, therefore, be due to the AOD representation given the understated injection height of the total biomass burning emissions found for the MACC emissions by Soares et al. (2015). A different hypothesis ascribes this underestimation to underestimated emissions. Toll et al. (2015b) found that while the daytime plumes from large fires were indeed lifted higher, the night time emissions and emissions from small fires were injected closer to the ground, making the average smoke transport distance even smaller than for the fixed emission height. Also Soares et al. (2015) point out, referring to Wooster et al. (2005), that MODIS is not sensitive enough to register the fire radiative power of small or smoldering fires, and thus large fraction of those is missed in the emission data, including also strongly emitting peat fires. The 2010 Russian fires included some huge fires, but also numerous small ones over large areas, and a large fraction of those was probably missed by MODIS.

Moreover a high underestimation was produced for all the simulations, irrespectively of the used meteorological and chemical model, over a small area in the south-eastern part of the domain (the above-called "blue spot"). This can be explained by the fact that the emissions inventories used herein only covered European areas (see the Emission Map in the SM), thus the emissions over that area were not considered.

The AOD over the southern part of the domain was overestimated and related mainly to the high dust concentrations according to the boundary conditions. In line with this, Solazzo et al. (2017) found that the error in primary species as dust was strongly affected by the emission and boundary conditions in the AQMEII Phase 3 simulations.

On the whole, the FI1 simulations, which used the ECMWF files as SILAM model input, gave high AOD values because SILAM is known to have slower dry particle deposition than other models. This could explain that, although the band quiet crudely represents size distribution, AOD is also very sensitive. IT1, which used the WRF meteorological conditions as CAMx model input, displayed quite a reasonable AOD representation skill. ES1, which used WRF-Chem, presented a high AOD overestimation due to the dust outbreaks. This marked overestimation took place because of a bug in the used dust scheme, which lacks the gravitational settling. Although the IT2 simulations used the same dust scheme and model version, the dust flux was modified for these simulations to estimate accurate dust concentrations. The IT2 simulations presented similar patterns to simulations FI1 and IT1, but obtained lower AOD values. No important differences were observed for the AOD representation among the IT2 simulations when the ARI and ACI and aqueous chemistry in convective clouds were solved (IT2_M-ARI+ACI) *versus* the ARI and large-scale clouds only being solved by a simple module (IT2_M-ARI).

It is important to highlight that for all the simulations and seasons, the highest determination values were obtained over the areas with medium AOD values (observed values between 0.5 and 1.0), which were approximately the areas with the lowest error values. Thus the temporal representation of the medium AOD values by all the simulations was acceptable. The use of an ensemble as the means for all the participant simulations improved this statistical figure.

The AE satellite values were obtained only over sea. High AE values, which indicate fine particles, were found near central European coasts which were higher in summer, probably due to the wildfires that occurred thenand inland, which represent low particles, probably influenced by the anthropogenic emissions. Low AE values, which indicate coarse particles, were observed over the southern part of the domain, close to the Saharan desert and over the Atlantic Ocean. It was also noteworthy that the AE values over the Atlantic Ocean were generally much higher in spring and summer than in autumn and winter. This means that the aerosol particles over ocean areas and near the coast in warm months were apparently finer than in colder months. This might be related to two different hypotheses: weaker winds in warm months or hygroscopic growth, which could be greater in cold months generally because of higher relative humidity (RH).

AE modelling skills were lower than for AOD (larger errors). The simulation run with the SILAM model and driven by ECMWF meteorological inputs (FI1_HTAP) largely underestimated AE over most of the domain. Hence, this model estimated larger-sized particles than that retrieved by satellite observations. As aforementioned, SILAM crudely represents size distribution, which impacted the AE representation because it may have be centered on particles with a larger diameter. The simulations using WRF coupled CAMx model (IT_MACC) and both WRF-Chem simulations (ES1_MACC and IT2_M-ARI+ACI) underestimated high AE values and overestimated low AE values. Thus, they underpredicted the variability of this variable. These results are similar to those established in Palacios-Peña et al. (2017); ?Palacios-Peña et al. (2017, 2018). On the other hand, Solazzo et al. (2012); Balzarini (2013); Solazzo et al. (2014) found a severely underestimate for PM_{10} concentrations over Europe for WRF-CAMx and WRF-Chem models, which could explain the overestimation of low AE values. These authors also found an underestimation of $PM_{2.5}$ concentrations which could also explain the underestimation of high AE values since simulated particles underestimate the variability of the size. An interesting fact is shown for ES1_MACC. Despite the lack of dust gravitational settling, it presented the lowest error values for AE. This could be explained by the high dust concentration over southern areas, resulting in low AE values and thus compensating the tendency for producing high $PM_{2.5}$/PM_{10} ratios. One remarkable issue is that AE values were highly underestimated by all the studies simulations when these were compared with AERONET station.

It was not possible to find any clear spatial pattern for the coefficient of determination of the AE representation. One striking fact in this case was that using the mean of all the simulations as ENSEMBLE did not improve this statistical figure. In fact the worse determination results were found for ENSEMBLE. The FI1_HTAP simulation showed the highest determination values. Thus despite the high underestimation of the AE values, it displayed a good skill in the temporal AE representation.

As mentioned above, a good approach to evaluate the spatial and temporal variability of a variable is PDF. A wide PDF indicates wide variability for the studied variable, and a narrow and low variability. For the AOD representation, all the simulations presented similar PDF to the observed values. The behaviour of all the simulations was similar in winter, spring and autumn; simulations FI1 and IT2 presented higher probabilities for lower AOD values than those observed; ES1 presented higher probabilities for high AOD values than those observed due to the above-explained lack of dust gravitational settling. Finally, IT1 presented the most skilful PDF, except during summer when was presented by IT2_M-ARI+ACI . Given the probability of obtaining AOD values around 0.5 being higher, the IT2 simulations presented the best skills in summer. One general conclusion was reached from the PDF of the AE values. For this variable, all the simulations in all the studied seasons underestimated temporal and spatial variability and this behaviour was stronger inland.

In conclusion, the skills of all the simulations in the AOD representation produced lower errors than in the AE representation. For AOD, low and medium values were well-represented, but high values presented larger errors. High values due to dust were overestimated because of an overestimation in the boundary conditions. The high AOD values due to biomass burning were underestimated, which should be ascribed to an understated injection height of the total biomass burning emissions or directly to underestimated emissions. Other high AOD values were underestimated because the emissions which produced these high values were not considered. The errors in the AOD representation evidenced the strong influence of emissions and boundary conditions in the estimation of aerosol optical properties. Generally speaking, the models' skills to represent the variability of AOD were acceptable. For AE, the SILAM simulation underestimated the observed values and the WRF coupled CAMx simulation and the WRF-Chem simulations were those with the best skills in the representation of this variable. But for all the simulations, the variability of this variable was underestimated.

Following these results, further studies are needed to improve the representation of aerosol optical properties, along with other properties such as atmospheric distribution, hygroscopicity, or the ability to act as cloud condensation nuclei (CCN) and ice nuclei (IN). The matter noted in the representation of aerosol properties can help to gain a better representation of ARI and ACI and aerosol effects on meteorology and climate, and could reduce the grave uncertainty in the estimations of changes in the Earth's radiation budget due to aerosols and clouds.

5 Data availability

The outputs from the simulations can be obtained by emailing to *rbianconi@enviroware.com*. MODIS data are publicly available on the MODIS Atmosphere website (https://modis-atmos.gsfc.nasa.gov/MOD04_L2/acquiring.html).

Acknowledgements. This work was conducted under the support of the AQMEII/HTAP Phase III initiative. The authors acknowledge Project REPAIR-CGL2014-59677-R of the Spanish Ministry of Economy and Competitiveness and the FEDER European programme for support to conduct this research. This work has been also possible thanks to fellowship 19677/EE/14, funded by the "Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia", and the Programme "Jiménez de la Espada de Movilidad, Cooperación e Internacionalización", within the II PCTIRM 2011-2014 framework. L. Palacios-Peña acknowledges the FPU scholarship (ref. FPU14/05505) of the Spanish Ministry of Education, Culture and Sport. G. Curci and P. Tuccella thank the EuroMediterranean Center for Climate reserach (CMCC) for providing the computational resources. We also thank the researchers and their staff who have been involved in the MODIS datasets (NASA).

References

Ackermann, I. J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F. S., and Shankar, U.: Modal aerosol dynamics model for Europe: development and first applications, Atmospheric Environment, 32, 2981 – 2999, doi:10.1016/S1352-2310(98)00006-5, 1998.

Ahmadov, R., McKeen, S. A., Robinson, A. L., Bahreini, R., Middlebrook, A. M., de Gouw, J. A., Meagher, J., Hsie, E.-Y., Edgerton, E., Shaw, S., and Trainer, M.: A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2011JD016831, d06301, 2012.

Ångström, A.: On the atmospheric transmission of sun radiation and on dust in the air, Geografiska Annaler, 11, 156–166, 1929.

Balzarini, A.: Implementing the WRF-Chem modeling system to investigate the interactions between air quality and meteorology, Ph.D. thesis, University of Milano-Bicocca, 2013.

Balzarini, A., Pirovano, G., Honzak, L., Ẑabkar, R., Curci, G., Forkel, R., Hirtl, M., José, R. S., Tuccella, P., and Grell, G.: WRF-Chem model sensitivity to chemical mechanisms choice in reconstructing aerosol optical properties, Atmospheric Environment, 115, 604 – 619, doi:10.1016/j.atmosenv.2014.12.033, 2015.

Barnard, J. C., Fast, J. D., Paredes-Miranda, G., Arnott, W. P., and Laskin, A.: Technical Note: Evaluation of the WRF-Chem "Aerosol Chemical to Aerosol Optical Properties" Module using data from the MILAGRO campaign, Atmospheric Chemistry and Physics, 10, 7325–7340, doi:10.5194/acp-10-7325-2010, 2010.

Baró, R., Lorente-Plazas, R., Montávez, J. P., and Jiménez-Guerrero, P.: Biomass burning aerosol impact on surface winds during the 2010 Russian heat wave, Geophysical Research Letters, doi:10.1002/2016GL071484, 2016GL071484, 2017.

Baró, R., Jiménez-Guerrero, P., Stengel, M., Brunner, D., Curci, G., Forkel, R., Neal, L., Palacios-Peña, L., Savage, N., Schaap, M., Tuccella, P., Denier van der Gon, H., and Galmarini, S.: Comparison of regional meteorology-chemistry models with satellite cloud products over Europe, Atmospheric Chemistry and Physics Discussions, 2018, 1–31, doi:10.5194/acp-2018-114, https://www.atmos-chem-phys-discuss.net/acp-2018-114/, 2018.

Bilal, M., Nazeer, M., Qiu, Z., Ding, X., and Wei, J.: Global Validation of MODIS C6 and C6. 1 Merged Aerosol Products over Diverse Vegetated Surfaces, Remote Sensing, 10, 475, 2018.

Boucher, O.: Atmospheric Aerosols: Properties and Climate Impacts, Springer, 2015.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and aerosols, in: Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, pp. 571–657, Cambridge University Press, 2013.

Chapman, E. G., Gustafson Jr., W. I., Easter, R. C., Barnard, J. C., Ghan, S. J., Pekour, M. S., and Fast, J. D.: Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of elevated point sources, Atmospheric Chemistry and Physics, 9, 945–964, doi:10.5194/acp-9-945-2009, 2009.

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

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, doi:10.1029/2009JD012820, d14207, 2010.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6, Geoscientific Model Development Discussions, 2016, 1–28, doi:10.5194/gmd-2016-139, 2016.

Curci, G., Bieser, J., Im, U., Christensen, J. H., Baró, R., na, L. P.-P., Jiménez-Guerrero, P., Prank, M., Colette, A., García-Vivanco, M., Tucella, P., Manders, A., Toros, H., Sokhi, R., Hogrefe, C., Galmarini, S., Solazzo, E., and Bianconi, R.: Black carbon absorption of solar radiation: combining external and internal mixing assumptions, to be submitted to Atmospheric Chemistry and Physics Discussions, https://www.atmos-chem-phys.net/special_issue390.html, 2017.

de Leeuw, G., Neele, F. P., Hill, M., Smith, M. H., and Vignati, E.: Production of sea spray aerosol in the surf zone, Journal of Geophysical Research: Atmospheres, 105, 29 397–29 409, doi:10.1029/2000JD900549, 2000.

Dentener, F., Galmarini, S., Hogrefe, C., Carmichael, G., Law, K., and Denby, B., eds.: Global and regional assessment of intercontinental transport of air pollution: results from HTAP, AQMEII and MICS, Special Issue on Atmospheric Chemistry and Physics, https://www.atmos-chem-phys.net/special_issue390.html, 2015.

Donahue, N. M., Robinson, a. L., Stanier, C. O., and Pandis, S. N.: Coupled partitioning, dilution, and chemical aging of semivolatile organics, Environmental Science and Technology, 40, 2635–2643, doi:10.1021/es052297c, 2006.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research: Atmospheres, 104, 31 333–31 349, doi:10.1029/1999JD900923, 1999.

ENVIRON, O.: User's Guide to the Comprehensive Air Quality Model with Extensions Version 5.40, ENVIRON International Corporation, Novato, CA., www.camx.com, 2014.

Eyring, V.and Lamarque, J.-F., Hess, P., Arfeuille, F., Bowman, K., Chipperfield, M. P., Duncan, B., Fiore, A., Gettelman, A., Giorgetta, M. A., Granier, C., Hegglin, M., Kinnison, D., Kunze, M., Langematz, U., Luo, B., Martin, R., Matthes, K., Newman, P. A., Peter, T., Robock, A., Ryerson, T., Saiz-Lopez, A., Salawitch, R., Schultz, M., Shepherd, T. G., Shindell, D., Stähelin, J., Tegtmeier, S., Thomason, L., Tilmes, S., Vernier, J.-P., Waugh, D. W., and Young, P. J.: Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) Community Simulations in Support of Upcoming Ozone and Climate Assessments, SPARC Newsletter, pp. 48–66, 2013.

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, 9, 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006721, d21305, 2006.

Forkel, R., Balzarini, A., Baró, R., Bianconi, R., Curci, G., Jiménez-Guerrero, P., Hirtl, M., Honzak, L., Lorenz, C., Im, U., Pérez, J. L., Pirovano, G., José, R. S., Tuccella, P., Werhahn, J., and Žabkar, R.: Analysis of the WRF-Chem contributions to AQMEII phase2 with respect to aerosol radiative feedbacks on meteorology and pollutant distributions, Atmospheric Environment, 115, 630 – 645, doi:10.1016/j.atmosenv.2014.10.056, 2015.

Forkel, R., Brunner, D., Baklanov, A., Balzarini, A., Hirtl, M., Honzak, L., Jiménez-Guerrero, P., Jorba, O., Pérez, J., San José, R., et al.: A Multi-model Case Study on Aerosol Feedbacks in Online Coupled Chemistry-Meteorology Models Within the COST Action ES1004 EuMetChem, in: Air Pollution Modeling and its Application XXIV, pp. 23–28, Springer, 2016.

Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., Denier van der Gon, H., Facchini, M. C., Fowler, D., Koren, I., Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M., and Gilardoni, S.: Particulate matter, air quality and climate: lessons learned and future needs, Atmospheric Chemistry and Physics, 15, 8217–8299, doi:10.5194/acp-15-8217-2015, 2015.

Galmarini, S., Rao, S. T., and Steyn, D. G.: Preface to the AQMEII p1 Special issue, Atmospheric Enviroment, 53, 1–3, 2012.

Galmarini, S., Hogrefe, C., Brunner, D., Makar, P., and Baklanov, A.: Preface, Atmospheric Environment, pp. 340–344, 2015.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmospheric Chemistry and Physics, 17, 1543–1555, doi:10.5194/acp-17-1543-2017, 2017.

Ghan, S., Laulainen, N., Easter, R., Wagener, R., Nemesure, S., Chapman, E., Zhang, Y., and Leung, R.: Evaluation of aerosol direct radiative forcing in MIRAGE, Journal of Geophysical Research: Atmospheres, 106, 5295–5316, doi:10.1029/2000JD900502, 2001.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions of dust aerosols simulated with the GOCART model, Journal of Geophysical Research: Atmospheres, 106, 20 255–20 273, doi:10.1029/2000JD000053, 2001.

Ginoux, P., Horowitz, L. W., Ramaswamy, V., Geogdzhayev, I. V., Holben, B. N., Stenchikov, G., and Tie, X.: Evaluation of aerosol distribution and optical depth in the Geophysical Fluid Dynamics Laboratory coupled model CM2.1 for present climate, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006707, d22210, 2006.

Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, Global Biogeochemical Cycles, 17, doi:10.1029/2003GB002079, 1097, 2003.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957 – 6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.

Guenther, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmospheric Chemistry and Physics, 6, 2006.

Hess, M., Koepke, P., and Schult, I.: Optical Properties of Aerosols and Clouds: The Software Package OPAC, Bulletin of the American Meteorological Society, 79, 831–844, 1998.

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

Ignatov, A., Stowe, L., and Singh, R.: Sensitivity study of the Ångström exponent derived from AVHRR over the oceans, Advances in Space Research, 21, 439–442, 1998.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmospheric Chemistry and Physics, 15, 11 411–11 432, doi:10.5194/acp-15-11411-2015, 2015.

Jeuken, A., Veefkind, J. P., Dentener, F., Metzger, S., and Gonzalez, C. R.: Simulation of the aerosol optical depth over Europe for August 1997 and a comparison with observations, Journal of Geophysical Research: Atmospheres, 106, 28 295–28 311, doi:10.1029/2001JD900063, 2001.

Kim, S.-W., Heckel, A., Frost, G. J., Richter, A., Gleason, J., Burrows, J. P., McKeen, S., Hsie, E.-Y., Granier, C., and Trainer, M.: NO_2 columns in the western United States observed from space and simulated by a regional chemistry model and their implications for NO_x emissions, Journal of Geophysical Research: Atmospheres, 114, doi:10.1029/2008JD011343, d11301, 2009.

Kinne, S., Lohmann, U., Feichter, J., Schulz, M., Timmreck, C., Ghan, S., Easter, R., Chin, M., Ginoux, P., Takemura, T., Tegen, I., Koch, D., Herzog, M., Penner, J., Pitari, G., Holben, B., Eck, T., Smirnov, A., Dubovik, O., Slutsker, I., Tanre, D., Torres, O., Mishchenko, M., Geogdzhayev, I., Chu, D. A., and Kaufman, Y.: Monthly averages of aerosol properties: A global comparison among models, satellite data, and AERONET ground data, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2001JD001253, 4634, 2003.

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, doi:10.5194/acp-6-1815-2006, 2006.

Kong, X., Forkel, R., Sokhi, R. S., Suppan, P., Baklanov, A., Gauss, M., Brunner, D., Baró, R., Balzarini, A., Chemel, C., Curci, G., Jiménez-Guerrero, P., Hirtl, M., Honzak, L., Im, U., Pérez, J. L., Pirovano, G., José, R. S., Schlünzen, K. H., Tsegas, G., Tuccella, P., Werhahn, J., Ẑabkar, R., and Galmarini, S.: Analysis of meteorology–chemistry interactions during air pollution episodes using online coupled models within AQMEII phase-2, Atmospheric Environment, 115, 527 – 540, doi:10.1016/j.atmosenv.2014.09.020, 2015.

Kulmala, M., Asmi, A., Lappalainen, H. K., Baltensperger, U., Brenguier, J.-L., Facchini, M. C., Hansson, H.-C., Hov, Ø., O'Dowd, C. D., Pöschl, U., Wiedensohler, A., Boers, R., Boucher, O., de Leeuw, G., Denier van der Gon, H. A. C., Feichter, J., Krejci, R., Laj, P., Lihavainen, H., Lohmann, U., McFiggans, G., Mentel, T., Pilinis, C., Riipinen, I., Schulz, M., Stohl, A., Swietlicki, E., Vignati, E., Alves, C., Amann, M., Ammann, M., Arabas, S., Artaxo, P., Baars, H., Beddows, D. C. S., Bergström, R., Beukes, J. P., Bilde, M., Burkhart, J. F., Canonaco, F., Clegg, S. L., Coe, H., Crumeyrolle, S., D'Anna, B., Decesari, S., Gilardoni, S., Fischer, M., Fjaeraa, A. M., Fountoukis, C., George, C., Gomes, L., Halloran, P., Hamburger, T., Harrison, R. M., Herrmann, H., Hoffmann, T., Hoose, C., Hu, M., Hyvärinen, A., Hõrrak, U., Iinuma, Y., Iversen, T., Josipovic, M., Kanakidou, M., Kiendler-Scharr, A., Kirkevåg, A., Kiss, G., Klimont, Z., Kolmonen, P., Komppula, M., Kristjánsson, J.-E., Laakso, L., Laaksonen, A., Labonnote, L., Lanz, V. A., Lehtinen, K. E. J., Rizzo, L. V., Makkonen, R., Manninen, H. E., McMeeking, G., Merikanto, J., Minikin, A., Mirme, S., Morgan, W. T., Nemitz, E., O'Donnell, D., Panwar, T. S., Pawlowska, H., Petzold, A., Pienaar, J. J., Pio, C., Plass-Duelmer, C., Prévôt, A. S. H., Pryor, S., Reddington, C. L., Roberts, G., Rosenfeld, D., Schwarz, J., Seland, Ø., Sellegri, K., Shen, X. J., Shiraiwa, M., Siebert, H., Sierau, B., Simpson, D., Sun, J. Y., Topping, D., Tunved, P., Vaattovaara, P., Vakkari, V., Veefkind, J. P., Visschedijk, A., Vuollekoski, H., Vuolo, R., Wehner, B., Wildt, J., Woodward, S., Worsnop, D. R., van Zadelhoff, G.-J., Zardini, A. A., Zhang, K., van Zyl, P. G., Kerminen, V.-M., S Carslaw, K., and Pandis, S. N.: General overview: European Integrated project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI); integrating aerosol research from nano to global scales, Atmospheric Chemistry and Physics, 11, 13 061–13 143, doi:10.5194/acp-11-13061-2011, 2011.

Landi, T. C.: AODEM, ISBN 10: 3659318027 / ISBN 13: 9783659318023, LAP Lambert Academic Publishing, 2013.

Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., and Hsu, N.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech, 6, 2989–3034, 2013.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and

Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geoscientific Model Development, 5, 709–739, doi:10.5194/gmd-5-709-2012, 2012.

Makar, P., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Ẑabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M., Pabla, B., Pérez, J., Pirovano, G., José, R. S., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on chemistry, Atmospheric Environment, 115, 499 – 526, doi:10.1016/j.atmosenv.2014.10.021, 2015.

Mhawish, A., Banerjee, T., Broday, D. M., Misra, A., and Tripathi, S. N.: Evaluation of MODIS Collection 6 aerosol retrieval algorithms over Indo-Gangetic Plain: Implications of aerosols types and mass loading, Remote Sensing of Environment, 201, 297–313, 2017.

Moorthy, K. K., Satheesh, S. K., Babu, S. S., and Dutt, C. B. S.: Integrated Campaign for Aerosols, gases and Radiation Budget (ICARB): An overview, Journal of Earth System Science, 117, 243–262, doi:10.1007/s12040-008-0029-7, 2008.

Ogren, J.: WMO/GAW Standard Operating Procedures for In-Situ Measurements of Aerosol Mass Concentration, Light Scattering and Light Absorption, WMO/GAW, Tech. rep., World Meteorological Organization Report, 2011.

Palacios-Peña, L., Baró, R., Guerrero-Rascado, J. L., Alados-Arboledas, L., Brunner, D., and Jiménez-Guerrero, P.: Evaluating the representation of aerosol optical properties using an online coupled model over the Iberian Peninsula, Atmospheric Chemistry and Physics, 17, 277–296, doi:10.5194/acp-17-277-2017, 2017.

Palacios-Peña, L., Baró, R., Baklanov, A., Balzarini, A., Brunner, D., Forkel, R., Hirtl, M., Honzak, L., López-Romero, J. M., Montávez, J. P., Pérez, J. L., Pirovano, G., San José, R., Schröder, W., Werhahn, J., Wolke, R., Žabkar, R., and Jiménez-Guerrero, P.: An assessment of aerosol optical properties from remote-sensing observations and regional chemistry–climate coupled models over Europe, Atmospheric Chemistry and Physics, 18, 5021–5043, doi:10.5194/acp-18-5021-2018, https://www.atmos-chem-phys.net/18/5021/2018/, 2018.

Paredes-Miranda, G., Arnott, W. P., Jimenez, J. L., Aiken, A. C., Gaffney, J. S., and Marley, N. A.: Primary and secondary contributions to aerosol light scattering and absorption in Mexico City during the MILAGRO 2006 campaign, Atmospheric Chemistry and Physics, 9, 3721–3730, doi:10.5194/acp-9-3721-2009, 2009.

Pouliot, G., Pierce, T., van der Gon, H. D., Schaap, M., Moran, M., and Nopmongcol, U.: Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, Atmospheric Environment, 53, 4–14, 2012.

Pouliot, G., van der Gon, H. A. D., Kuenen, J., Zhang, J., Moran, M. D., and Makar, P. A.: Analysis of the emission inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project, Atmospheric Environment, 115, 345–360, 2015.

Poupkou, A., Giannaros, T., Markakis, K., Kioutsioukis, I., Curci, G., Melas, D., and Zerefos, C.: A model for European biogenic volatile organic compound emissions: software development and first validation, Environmental Modelling & Software, 25, 1845–1856, 2010.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and Taylor, K. E.: Climate models and their evaluation, in: Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR), pp. 589–662, Cambridge University Press, 2007.

Rao, S. T., Galmarini, S., and Puckett, K.: Air Quality Model Evaluation International Initiative (AQMEII): Advancing the State of the Science in Regional Photochemical Modeling and Its Applications, Bulletin of the American Meteorological Society, 92, 23–30, doi:10.1175/2010BAMS3069.1, 2011.

Reddy, M. S., Boucher, O., Bellouin, N., Schulz, M., Balkanski, Y., Dufresne, J.-L., and Pham, M.: Estimates of global multicomponent aerosol optical depth and direct radiative perturbation in the Laboratoire de Météorologie Dynamique general circulation model, Journal of Geophysical Research: Atmospheres, 110, doi:10.1029/2004JD004757, d10S16, 2005.

Remer, L. A., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R., Kleidman, R., Eck, T. F., E, V., and Holben, B. N.: The MODIS aerosol algorithm, products, and validation, Journal of the atmospheric sciences, 62, 947–973, 2005.

Sayer, A., Munchak, L., Hsu, N., Levy, R., Bettenhausen, C., and Jeong, M.-J.: MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and ?merged? data sets, and usage recommendations, Journal of Geophysical Research: Atmospheres, 119, 2014.

Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., and Ebel, A.: Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, Journal of Geophysical Research: Atmospheres, 106, 28 275–28 293, doi:10.1029/2001JD000384, 2001.

Schulz, M., Chin, M., and Kinne, S.: The aerosol model comparison project, AeroCom, phase II: Clearing up diversity, IGAC Newsletter, 2009.

Shaw, W. J., Allwine, K. J., Fritz, B. G., Rutz, F. C., Rishel, J. P., and Chapman, E. G.: An evaluation of the wind erosion module in DUSTRAN, Atmospheric Environment, 42, 1907 – 1921, doi:https://doi.org/10.1016/j.atmosenv.2007.11.022, 2008.

Smirnov, A., Holben, B. N., Slutsker, I., Giles, D. M., McClain, C. R., Eck, T. F., Sakerin, S. M., Macke, A., Croot, P., Zibordi, G., Quinn, P. K., Sciare, J., Kinne, S., Harvey, M., Smyth, T. J., Piketh, S., Zielinski, T., Proshutinsky, A., Goes, J. I., Nelson, N. B., Larouche, P., Radionov, V. F., Goloub, P., Krishna Moorthy, K., Matarrese, R., Robertson, E. J., and Jourdin, F.: Maritime Aerosol Network as a component of Aerosol Robotic Network, Journal of Geophysical Research: Atmospheres, 114, doi:10.1029/2008JD011257, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011257, 2009.

Soares, J., Sofiev, M., and Hakkarainen, J.: Uncertainties of wild-land fires emission in AQMEII phase 2 case study, Atmospheric Environment, 115, 361–370, 2015.

Sofiev, M.: A model for the evaluation of long-term airborne pollution transport at regional and continental scales, Atmospheric Environment, 34, 2481–2493, 2000.

Sofiev, M., Soares, J., Prank, M., de Leeuw, G., and Kukkonen, J.: A regional-to-global model of emission and transport of sea salt particles in the atmosphere, Journal of Geophysical Research: Atmospheres, 116, 2011.

Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., and Genikhovich, E.: Construction of the SILAM Eulerian atmospheric dispersion model based on the advection algorithm of Michael Galperin, Geoscientific Model Development, 8, 3497–3522, doi:10.5194/gmd-8-3497-2015, 2015.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Appel, K. W., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Miranda, A. I., Nopmongcol, U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operational model evaluation for particulate matter in Europe and North America in the context of AQMEII, Atmospheric Environment, 53, 75 – 92, doi:10.1016/j.atmosenv.2012.02.045, 2012.

Solazzo, E., Galmarini, S., Bianconi, R., and Rao, S. T.: Model evaluation for surface concentration of particulate matter in Europe and North America in the context of AQMEII, in: Air Pollution Modeling and its Application XXII, pp. 375–379, Springer, 2014.

Solazzo, E., Bianconi, R., Hogrefe, C., Curci, G., Tuccella, P., Alyuz, U., Balzarini, A., Baró, R., Bellasio, R., Bieser, J., Brandt, J., Christensen, J. H., Colette, A., Francis, X., Fraser, A., Vivanco, M. G., Jiménez-Guerrero, P., Im, U., Manders, A., Nopmongcol, U., Kitwiroon, N., Pirovano, G., Pozzoli, L., Prank, M., Sokhi, R. S., Unal, A., Yarwood, G., and Galmarini, S.: Evaluation and error apportionment of an ensemble of atmospheric chemistry transport modeling systems: multivariable temporal and spatial breakdown, Atmospheric Chemistry and Physics, 17, 3001–3054, doi:10.5194/acp-17-3001-2017, 2017.

Solmon, F., Giorgi, F., and Liousse, C.: Aerosol modelling for regional climate studies: application to anthropogenic particles and evaluation over a European/African domain, Tellus B, 58, 51–72, 2006.

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X.: The second generation regional acid deposition model chemical mechanism for regional air quality modeling, Journal of Geophysical Research: Atmospheres, 95, 16 343–16 367, doi:10.1029/JD095iD10p16343, 1990.

Toll, V., Reis, K., Ots, R., Kaasik, M., Männik, A., Prank, M., and Sofiev, M.: SILAM and MACC reanalysis aerosol data used for simulating the aerosol direct radiative effect with the NWP model HARMONIE for summer 2010 wildfire case in Russia, Atmospheric Environment, 121, 75–85, doi:10.1016/j.atmosenv.2015.06.007, 2015a.

Toll, V., Reis, K., Ots, R., Kaasik, M., Männik, A., Prank, M., and Sofiev, M.: SILAM and MACC reanalysis aerosol data used for simulating the aerosol direct radiative effect with the NWP model HARMONIE for summer 2010 wildfire case in Russia, Atmospheric Environment, 121, 75–85, 2015b.

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

Tuccella, P., Curci, G., Visconti, G., Bessagnet, B., Menut, L., and Park, R. J.: Modeling of gas and aerosol with WRF/Chem over Europe: Evaluation and sensitivity study, Journal of Geophysical Research: Atmospheres, 117, 2012.

Tuccella, P., Curci, G., Grell, G. A., Visconti, G., Crumeyrolle, S., Schwarzenboeck, A., and Mensah, A. A.: A new chemistry option in WRF-Chem v. 3.4 for the simulation of direct and indirect aerosol effects using VBS: evaluation against IMPACT-EUCAARI data, Geoscientific Model Development, 8, 2749–2776, doi:10.5194/gmd-8-2749-2015, 2015.

Wang, K., Yahya, K., Zhang, Y., Hogrefe, C., Pouliot, G., Knote, C., Hodzic, A., José, R. S., Pérez, J. L., Jiménez-Guerrero, P., Baró, R., Makar, P., and Bennartz, R.: A multi-model assessment for the 2006 and 2010 simulations under the Air Quality Model Evaluation International Initiative (AQMEII) Phase 2 over North America: Part II. Evaluation of column variable predictions using satellite data, Atmospheric Environment, 115, 587 – 603, doi:10.1016/j.atmosenv.2014.07.044, 2015.

Warneke, C., Froyd, K. D., Brioude, J., Bahreini, R., Brock, C. A., Cozic, J., de Gouw, J. A., Fahey, D. W., Ferrare, R., Holloway, J. S., Middlebrook, A. M., Miller, L., Montzka, S., Schwarz, J. P., Sodemann, H., Spackman, J. R., and Stohl, A.: An important contribution to springtime Arctic aerosol from biomass burning in Russia, Geophysical Research Letters, 37, doi:10.1029/2009GL041816, l01801, 2010.

Weil, J., Sykes, R., and Venkatram, A.: Evaluating air-quality models: review and outlook, Journal of Applied Meteorology, 31, 1121–1145, 1992.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate research, 30, 79, 2005.

Willmott, C. J., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., O'Donnell, J., and Rowe, C. M.: Statistics for the evaluation and comparison of models, Journal of Geophysical Research: Oceans, 90, 8995–9005, doi:10.1029/JC090iC05p08995, 1985.

Winker, D. M., Pelon, J. R., and McCormick, M. P.: The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds, in: Third International Asia-Pacific Environmental Remote Sensing Remote Sensing of the Atmosphere, Ocean, Environment, and Space, pp. 1–11, International Society for Optics and Photonics, 2003.

Wooster, M. J., Roberts, G., Perry, G., and Kaufman, Y.: Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release, Journal of Geophysical Research: Atmospheres, 110, 2005.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J.: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, Atmospheric Chemistry and Physics, 12, 8911–8949, doi:10.5194/acp-12-8911-2012, 2012.

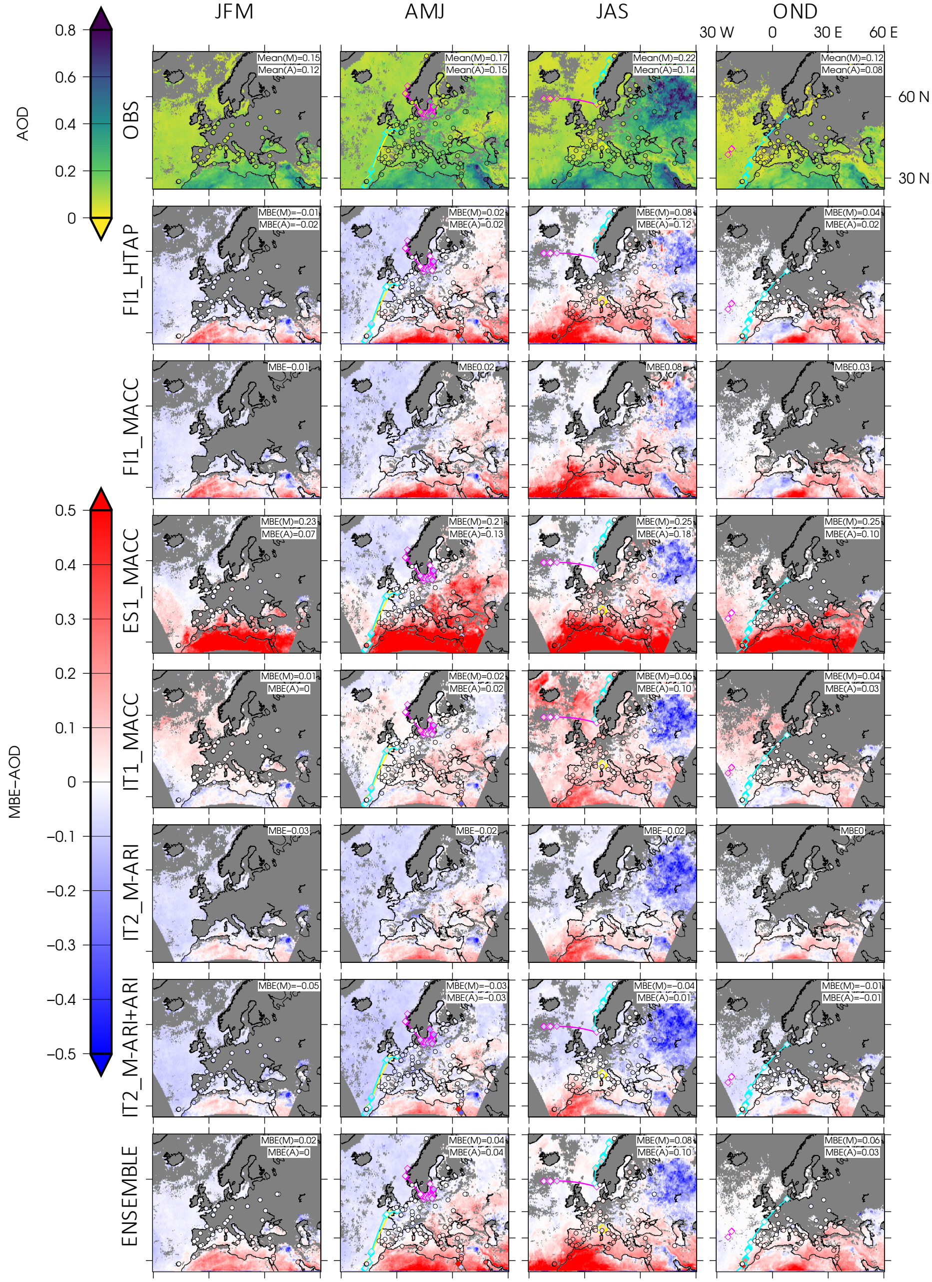

Figure 1. The MBE results of AOD at 550nm satellite and AOD at 675 nm AERONET (points) values *vs.* simulations. Columns from left to right, temporal mean of: winter (JFM), spring (AMJ), summer (JAS) and autumn (OND). First row: satellite values; and from second row to the bottom, the MBE values of: FI1_HTAP, FI1_MACC, ES1_MACC, IT1_MACC, IT2_M-ARI, IT2_M-ARI+ACI and ENSEMBLE. MAM value are represented by colored lines. In JFM: Ak Fedorov (yellow), Oceania (magenta), Polarstern (cyan) and Zim Iberia(chocolate). In JAS: Alliance (yellow), Ak Ioffe (magenta) and Oceania (cyan). And in OND: Ak Fedorov (yellow), James Cook (magenta) and Polarstern (cyan).

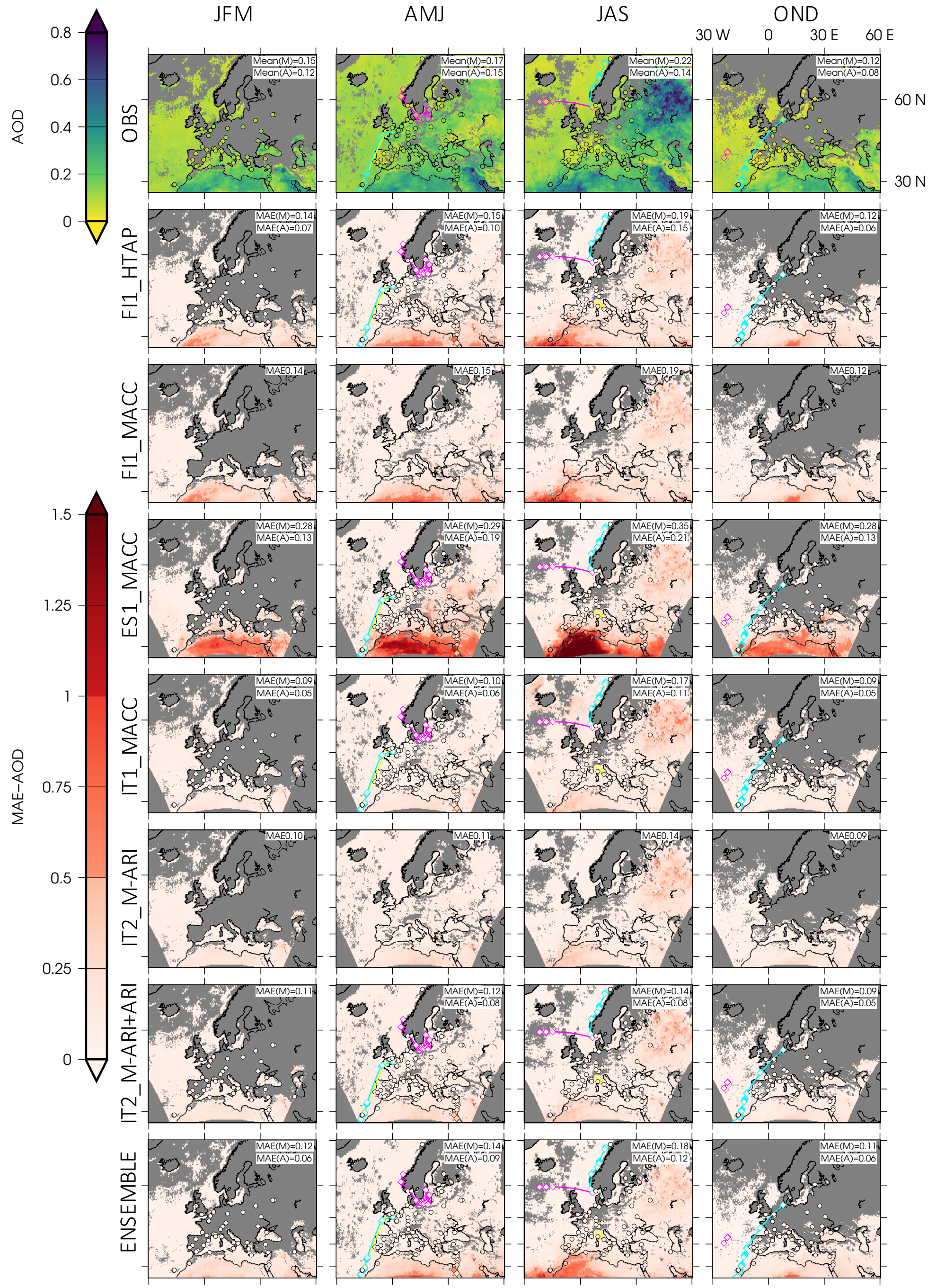

Figure 2. Idem Figure 1 for the MAE results of AOD at 550nm satellite values *vs.* simulations.

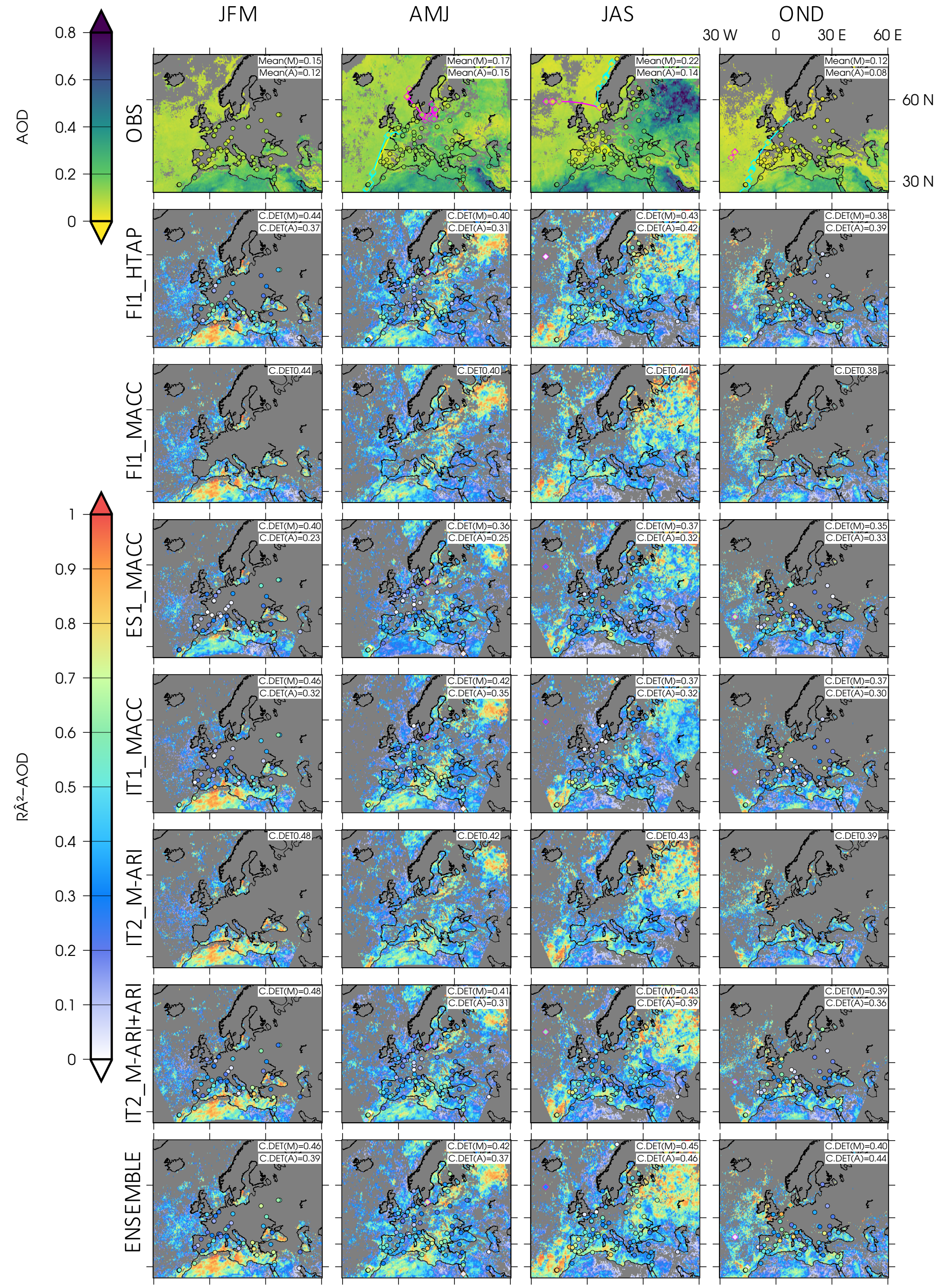

Figure 3. Idem Figure 1 for the determination coefficient of AOD at 550nm satellite values *vs.* simulations.

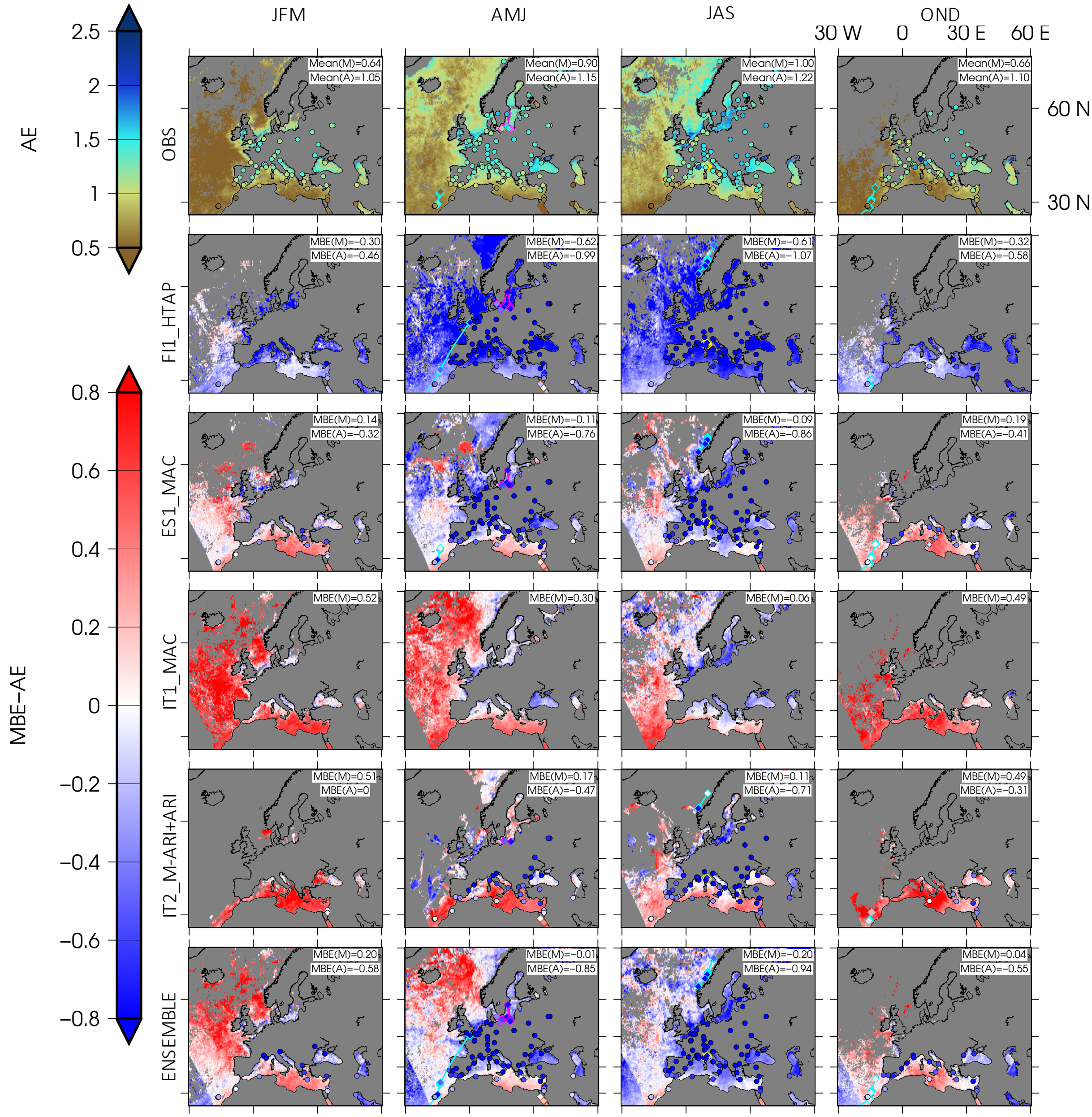

Figure 4. MBE results of AE between 550 and 860nm satellite and AE between 440 and 870 nm AERONET (points) values *vs.* simulations. Columns from left to right, temporal mean of: winter (JFM), spring (AMJ), summer (JAS) and autumn (OND). First row: satellite values; and from the second row to the bottom, the MBE values of: FI1_HTAP, ES1_MACC, IT1_MACC, IT2_M-ARI+ACI and ENSEMBLE. MAM value are represented by colored lines. In JFM: Oceania (magenta), Polarstern (cyan) and Zim Iberia(chocolate). In JAS: Alliance (yellow) and Oceania (cyan). and In OND: Polartstern (cyan).

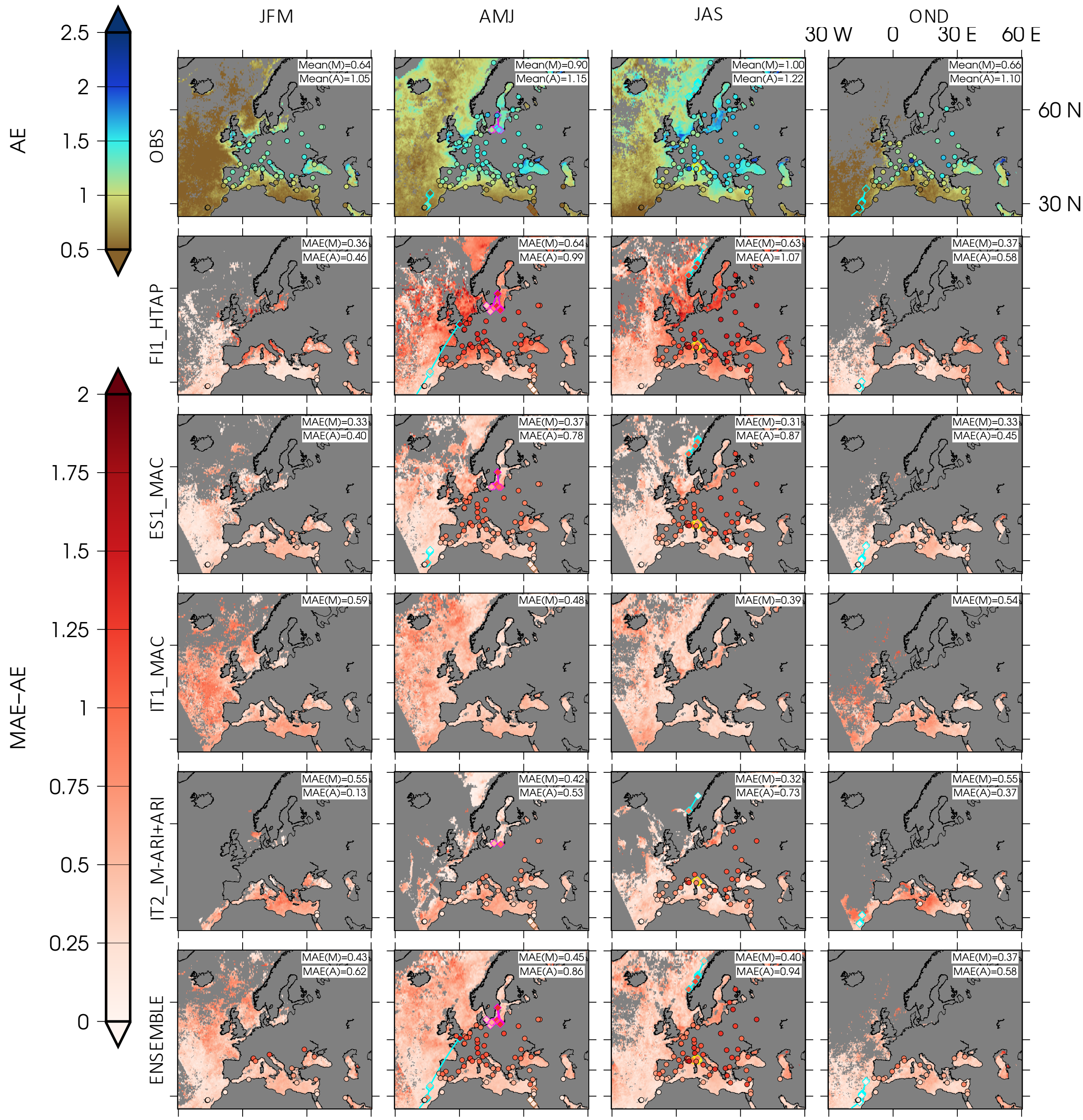

Figure 5. Idem Figure 4 for the MAE results of AE between 550 and 860nm satellite values *vs.* simulations.

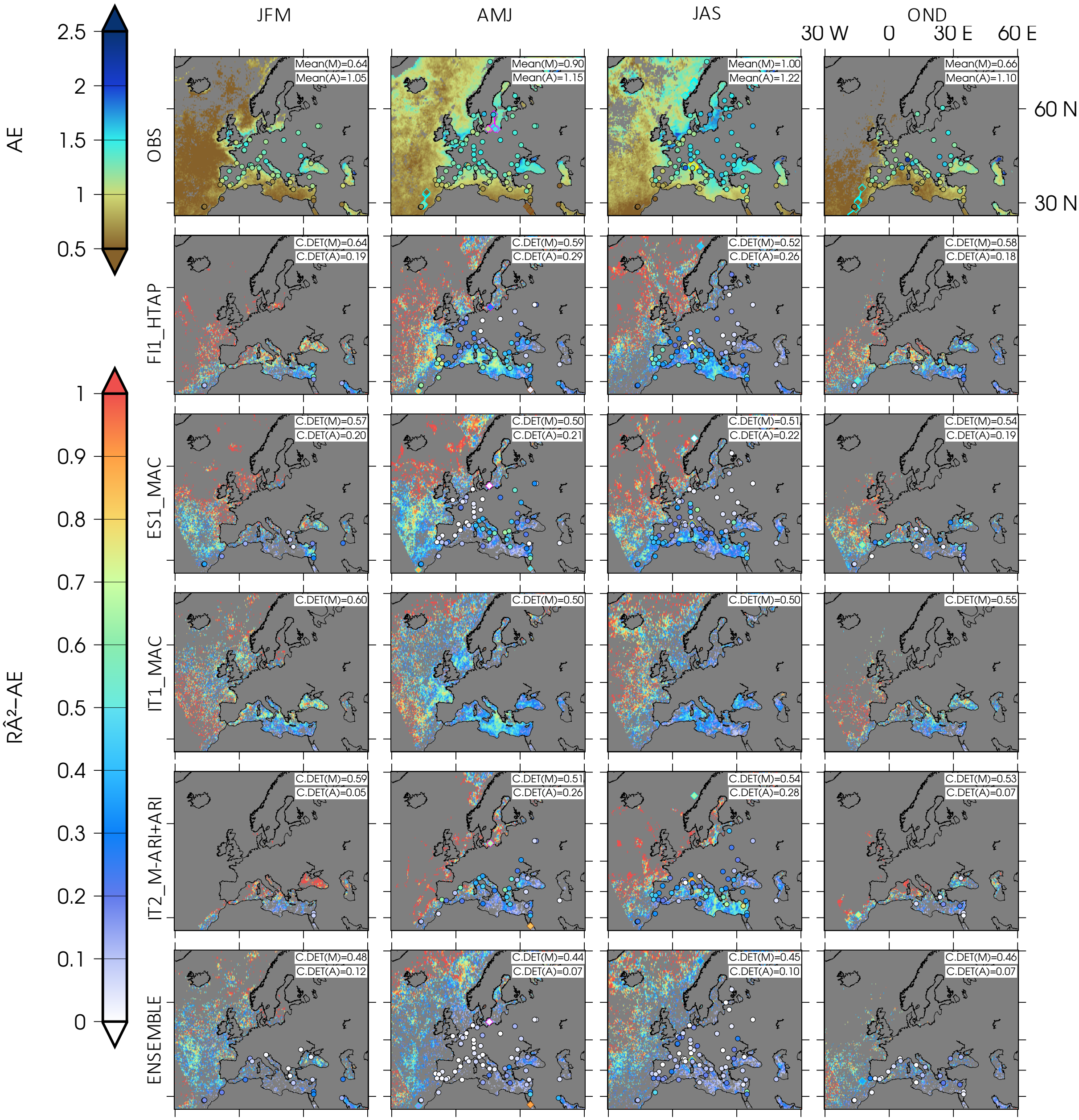

Figure 6. Idem Figure 4 for the determination coefficient of AE between 550 and 860nm satellite values *vs.* simulations.

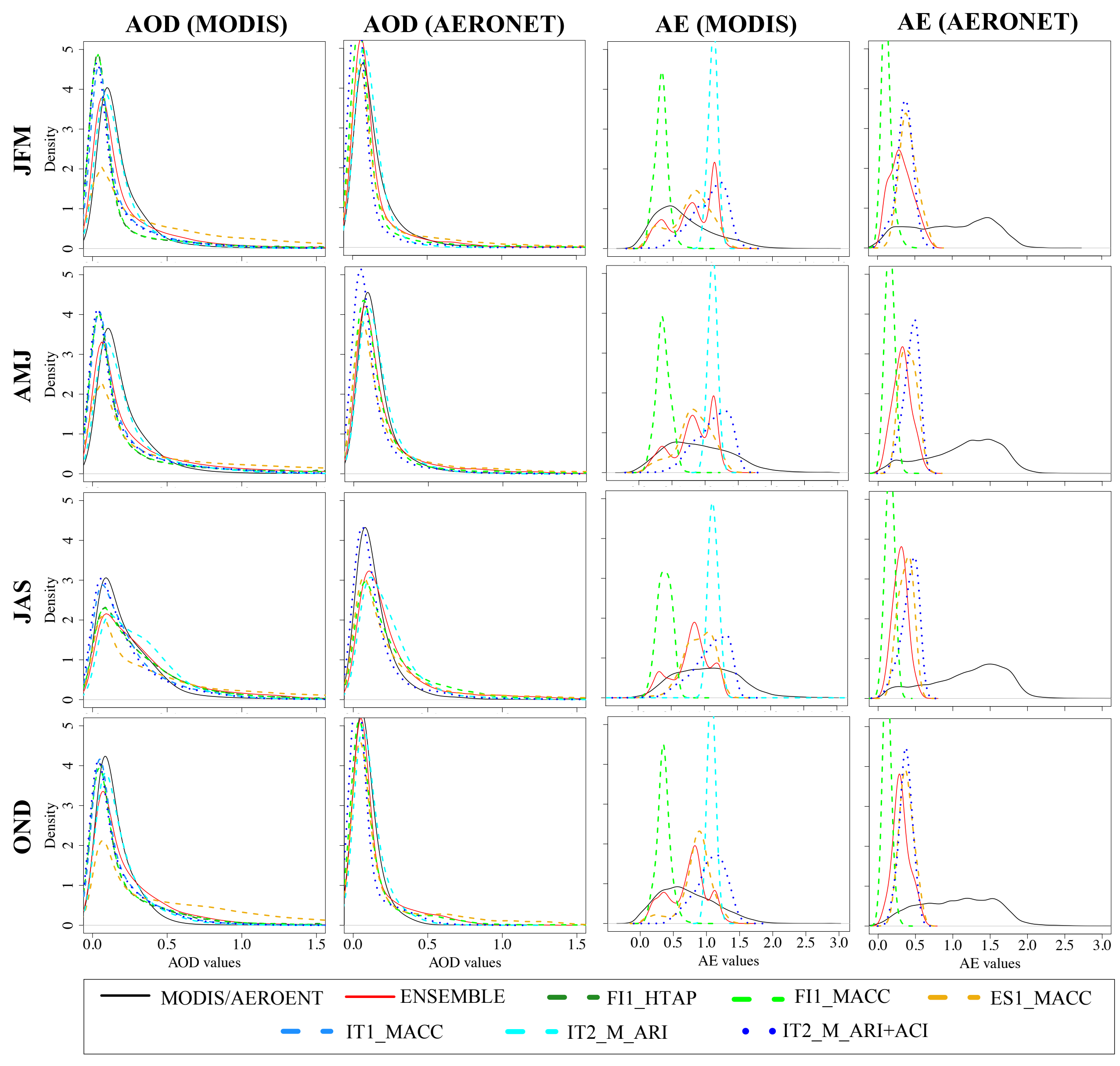

Figure 7. PDF for AOD (~~left column~~first, MODIS and second, AERONET columns) and AE (third, MODIS and fourth right, AERONET) values. From the top to the bottom: JFM, AMJ, JAS, OND.