An assessment of aerosol optical properties from remote sensing observations and regional chemistry-climate coupled models over Europe

Laura Palacios-Peña[1], Rocío Baró[1,2], Alexander Baklanov[3], Alessandra Balzarini[4], Dominik Brunner[5], Renate Forkel[6], Marcus Hirtl[2], Luka Honzak[7], José María López-Romero[1], Juan Luis Pérez[8], Guido Pirovano[4], Roberto San José[8], Wolfram Schröder[9], Johannes Werhahn[6], Ralf Wolke[9], Rahela Zabkar[10], and Pedro Jiménez-Guerrero[1]

[1]Department of Physics, Regional Campus of International Excellence Campus Mare Nostrum, University of Murcia, Murcia, Spain
[2]Section Chemical Weather Forecasts, Division Data/Methods/Modeling, ZAMG - Zentralanstalt für Meteorologie und Geodynamik, Austria
[3]World Meteorological Organization, Geneve, Switzerland
[4]Ricerca sul Sistema Energetico (RSE), Italy
[5]Laboratory for Air Pollution/Environmental Technology, Empa, Swiss Federal Laboratories for Materials Science and Technology, Switzerland
[6]Karlsruher Institut für Technologie (KIT), Institut für Meteorologie und Klimaforschung, Atmosphärische Umweltforschung (IMK-IFU), Germany
[7]BO-MO d.o.o, Slovenia
[8]Environmental Software and Modelling Group, Computer Science School - Technical University of Madrid, Spain
[9]Leibniz Institute for Tropospheric Research, Permoserstr, Germany
[10]Slovenian Environment Agency, Slovenia

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

Abstract. Atmospheric aerosols modify the radiative budget of the Earth due to their optical, microphysical and chemical properties, and are considered one of the most uncertain climate forcing agents. In order to characterise the uncertainties associated with satellite and modelling approaches to represent aerosol optical properties, mainly aerosol optical depth (AOD) and the Ångström exponent (AE), their representation by different remote sensing sensors and regional online coupled chemistry-climate models over Europe is evaluated. This work also characterises whether the inclusion of aerosol-radiation (ARI) or/and aerosol-cloud interactions (ACI) helps improve the skills of modelling outputs.

Two case studies were selected within the EuMetChem COST Action ES1004 framework, when important aerosol episodes in 2010 all over Europe took place: a Russian wildfire episode and a Saharan desert dust outbreak that covered most of the Mediterranean Sea. The model data came from different regional air quality-climate simulations performed by working group 2 of EuMetChem, which differed according to whether ARI or ACI was included or not. The remote sensing data came from three different sensors: MODIS, OMI and SeaWIFS. The evaluation used classical statistical metrics to first compare satellite data versus the ground-based instrument network (AERONET) and to then model versus the observational data (both satellite and ground-based).

Regarding the uncertainty in the satellite representation of AOD, MODIS presented the best agreement with the AERONET observations compared to other satellite AOD observations. The differences found between remote sensing sensors highlighted the uncertainty in the observations, which have to be taken into account when evaluating models. When modelling results were considered, a common trend for underestimating high AOD levels was observed. For the AE, models tended to underestimate its variability, except when considering a sectional approach in the aerosol representation. The modelling results showed better skills when ARI+ACI interactions were included; hence this improvement in the representation of AOD (above 30% in the model error) and AE (between 20 and 75%) is important to provide a better description of aerosol-radiation-cloud interactions in regional climate models.

1 Introduction

The uncertainty of atmospheric aerosol effects on the Earth radiative budget is much greater than for any other climate-forcing agent (Boucher et al., 2013). In this sense, aerosol properties, with emphasis placed on optical properties, are vastly variable on space and time scales due to short-lived aerosol particles and non-uniform emissions (Ramanathan et al., 2001; Kaufman et al., 2002; Randall et al., 2007). According to the International Panel on Climate Change (IPCC) Fifth Assessment Report (AR5), aerosol effects, which depend on their properties, can be classified into Aerosol-Radiation and Aerosol-Cloud Interactions (respectively ARI and ACI) (Boucher et al., 2013).

In order to reduce, or to at least characterise, this uncertainty, there are two main approaches to study aerosol optical properties and their influences on the climate system: (1) observations and (2) climate models. Aerosol properties cannot be constrained by measurements alone and using models can improve knowledge and our understanding of physical, chemical and optical aerosol properties. However, models are based on the previous observational study of these properties in order to define and implement the behaviour of aerosol particles in modelling systems.

When aerosol optical properties are evaluated, remote sensing is one of the most widely used techniques in the observational approach. The main advantages of remote sensing are: (1) it does not perturb the observed sample (aerosol particles in this case), and are sensitive to different properties, particularly to aerosol optical properties on which this study focuses; (2) they can provide point, column or profile data. For these reasons, several studies using this approach were carried out to improve knowledge of aerosol properties (e.g., Dubovik et al., 2000; Tanré et al., 2001; Dubovik et al., 2002; Levy et al., 2007; Garland et al., 2008) and to establish their radiative effects (Haywood et al., 2001; Chou et al., 2002; Bellouin et al., 2005; Huang et al., 2006, among many others).

These studies used data that were measured chiefly from two main platforms: from the Earth's surface, the so-called ground-based measurement; and from space by using satellites. Moreover, different instruments, such as sun photometers, spectro-radiometers or Lidar were used. Major efforts have been made to create networks of ground-based measurements of aerosol optical properties around the world; e.g., as the Aerosol Robotic Network (Holben et al., 1998, AERONET), the European Aerosol Research Lidar Network to Establish an Aerosol Climatology (Pappalardo et al., 2014, EARLINET) or the Latin American Lidar Network (Antuña-Marrero et al., 2016, LALINET). There are instruments aboard satellites that provide infor-

mation about aerosol optical properties with a wide spatial coverage; e.g., the Multi-angle Imaging SpectroRadiometer (Diner et al., 1998, MISR), the Moderate Resolution Imaging Spectroradiometer (Remer et al., 2005, MODIS), the Spinning Enhanced Visible and Infrared Imager (Aminou et al., 1997, 1999, SEVIRI), the Ozone Monitoring Instrument (Schoeberl et al., 2006, OMI) or the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (Winker et al., 2003, CALIPSO).

5 At this point, it is important to highlight that several studies (Fuzzi et al., 2015; Van Donkelaar et al., 2015) have discussed the differences between aerosol properties measured by ground-based stations, with a limited spatial coverage, but with more reliable information, and by satellite, with broader spatial coverage, but not with such reliable information as ground-based observations. These aspects are important, together with the wide variability of aerosol optical properties on space and time scales (Ramanathan et al., 2001; Kaufman et al., 2002; Randall et al., 2007), the huge number of different instruments aboard
10 satellites and the algorithms used, as they produce considerable uncertainty in the measured data.

  By considering the modelling approach to reduce or characterise the uncertainty of aerosol effects as part of the COST Action ES1004 EuMetChem (European framework for online integrated air quality and meteorology modelling; see http://www.eumetchem.info/), a list of interactions were identified as the most relevant coupling processes for regional air quality and weather predictions. These interactions were chosen because experts considered them to be the most important, but were
15 poorly represented at the same time in the current online coupled models (Kong et al., 2015). This type of model offers the possibility to account for the feedback mechanisms between simulated aerosol concentrations and meteorological variables, and thus permits the simulation of ARI and ACI, which form part of the above-mentioned list of interactions.

  The efforts made by EuMetChem action focused on two case studies, which were chosen from the previous experience of phase 2 of the Air Quality Model Evaluation International Initiative (AQMEII Galmarini et al., 2015), when important aerosol
20 episodes took place in 2010 all over Europe with potential aerosol effects on meteorology. These cases consist in the Russian wildfires and heat wave episode (in July and August 2010) and a Sahara desert dust outbreak with enhanced cloud and rain over the Mediterranean Sea (in October 2010). These case studies were chosen given the evidence of the particularly significant interactions between meteorology and chemistry during such strong air pollution episodes (Konovalov et al., 2011; Chen et al., 2014; Wong et al., 2012).

25 Several studies have highlighted the relevance of including aerosol radiative feedbacks during such episodes to improve meteorological forecasts (among many others Pérez et al., 2006; Bangert et al., 2012; Makar et al., 2015a, b). For example, the impacts of high aerosol loading during the desert dust events have been described by Shao et al. (2011) to demonstrate interactions with climate and ecosystem, impacts on the Earth radiative budget and a drop in photolysis rates.

  Furthermore, Konovalov et al. (2011) and Chubarova et al. (2012) have found that the very high aerosol concentration due
30 to the fire emissions during the specific Russian wildfires episode significantly changed the regional conditions of the climate system by changes in atmospheric gas composition, as well as the optical and radiative characteristics of aerosols. In Péré et al. (2014), was found a reduction of solar radiation on the ground up to 80-150 Wm^{-2} in diurnal averages, with a major reduction in the near-surface air temperature between 0.2 and 2.6 K on a regional scale over most of eastern Europe, a reduction in the atmospheric boundary layer height from 13% to 65% and the vertical wind speed from 5% to 80% when the ARI effects during

the Russian wildfire episode were included. Similar results have been reported by Baró et al. (2017), which reveal the effects of ARI on surface winds.

Moreover, Forkel et al. (2016) analysed the different chemistry and physics options of the same online coupled simulations used in this work to evaluate the effects of ARI on radiation and temperature during both episodes studied herein. During the fires episode, the inclusion of this effect led to a reduction between 10 and 100 Wm^{-2} in the average downward short-wave radiation at the ground level and a drop in the mean 2 m-temperature of almost 1 K over the area where the fires took place. During the dust outbreak episode, the ARI effect resulted in lower mean 2 m-temperature (-0.25 K over the Mediterranean Region).

However, none of the aforementioned studies have evaluated the representation of aerosol optical properties and the effects of ARI+ACI on these properties. As the previous revision reveals, a strong influence of the ARI+ACI exists on meteorological variables. Several studies untangled the influence of atmospheric aerosols on the atmospheric boundary layer (Roy and Sharp, 2013), atmospheric stability and winds (Péré et al., 2014; Baró et al., 2017) with a consequent effect on aerosol concentrations (Zhang et al., 2015). The winds variation affects emissions of wind-dependent particles over land, such as desert dust, or ocean, sea salt (Boucher et al., 2013; Prijith et al., 2014; Li et al., 2015). Moreover, the transport of pollutants (e.g Yang et al., 2017) and the aerosol vertical distribution (Mishra et al., 2015) could be altered. The modification of all of these processes influences aerosol optical properties (e.g. Huang et al., 2010). On the other hand, ARI+ACI effects on relative humidity (Baró et al., 2017) modify the size of the particles due to the hygroscopic growth affecting the particle extinction (Curci et al., 2015).

Due to the uncertainty of aerosol effects, the above-mentioned coupling processes are treated differently in online coupled chemistry-meteorology models. This fact highlights the need to study the response of this type of models to the same aerosol emissions. Although the scope of this work is not to study ensemble forecasting, an ensemble mean was performed by using a set of simulations. The reason was because the ensemble mean is likely to provide the most skilful simulations compared to individual ones (Baklanov et al., 2014), like the results found by Fernández et al. (2009); Knutti et al. (2010); Kjellström et al. (2011). These results can be explained by the paradigm of different models being considered independent samples from a given distribution that is centred on the truth (Annan and Hargreaves, 2010). Hence the ensemble mean improves compared to the performance of individual models, and this ensemble could be expected to converge to the truth as more models are added to the ensemble.

Hence the main objectives can be summarised as: (1) assessing the representation of aerosol optical properties by a set of online coupled chemistry-climate simulations; (2) determining whether the inclusion of aerosol radiative feedbacks in this type of models improves the modelling outputs of aerosol optical properties over Europe. In order to achieve these objectives and to characterise the high uncertainty of the measured data, the first step was to determine the "best" satellite data according to AERONET to assess simulations. Afterwards, the representation of the aerosol optical properties by regional online coupled models was evaluated by comparing the ensemble of the models against this "best" satellite estimation.

2 Data and Methods

Aerosol optical properties considered for the evaluation in this work were aerosol optical depth (AOD) and Ångström exponent (AE) at different wavelengths. These optical properties were assessed during two different episodes with a high aerosol load over Europe in 2010, as established in COST Action ES1004 EuMetChem. They were identified from the previous experience of phase 2 of the AQMEII modelling inter-comparison exercise (Galmarini et al., 2015), and consists in: the Russian wildfires and heat wave episode (from 25 July to 15 August 2010) and a Saharan desert dust outbreak with enhanced cloud and rain (from 2 to 15 October 2015).

2.1 Model simulations

The used simulations were run by working group 2 (WG2) of COST Action ES1004 EuMetChem (http://www.eumetchem. info/). WG2 of EuMetChem investigated the importance of different processes and feedback in online coupled chemistry-meteorology models for air quality simulations and weather forecasts.

The used simulations are summarised in Table 1. All these simulations, according to the model, chemical/physical options and feedbacks used, were denoted as an experiment. Two different online coupled models were used: COSMO-MUSCAT (Wolke et al., 2012) and WRF-Chem, version 3.4.1 (Grell et al., 2005) with different configurations (CS1, CS2, ES1 and ES3).

Simulations covered two different episodes with a high aerosol load over Europe in 2010. The Russian wildfires episode lasted 22 days and the Sahara desert dust outbreak lasted 14 days. Both episodes were simulated following the common strategy for AQMEII phase 2, a sequence of 2-day time slices in which the meteorology is reinitialised every 2 days and the chemical state is adopted from the final step of the previous time slice (Galmarini et al., 2012, 2015, 2017). A spin-up of 5 days was used for the chemistry.

The target domains covered Europe. However for an easy clear analysis of aerosol effects, the assessment was done in a window of the larger domain (see Figure 1). The majority of the simulations were carried out with a grid width of approximately 23 km. Exceptions were CS2 (9.9 km) and DE3 ($0.125°$, around 14 km), both with higher resolutions for the Russian wildfires episode. The outputs of all the simulations were interpolated to a common lat-lon regular grid at a resolution of $0.1°$. The analysis grid for the Russian wildfires/heat wave case (Figure 1, blue box) covered between $40°$ and $60°$ north and $20°$ and $60°$ east, with a grid size of 50,000 cells; that of the Saharan desert dust outbreak (Figure 1, green box) covers $25°$ and $55°$ north and $-10°$ and $30°$ east, with a grid size of 120,000 cells.

The simulations were run for three different feedback configurations: a base case called No Radiative Feedbacks (NRF), which does not take into account the radiative feedbacks of atmospheric aerosols in meteorology; a case that only bears in mind aerosol-radiation interactions (ARI); and another case that considers aerosol-radiation and aerosol-clouds interactions (ARI+ACI).

The meteorological initial and boundary conditions came from 3-hourly data with $0.25°$ of resolution from the European Centre for Medium-Range Weather Forecasts (ECMWF) operational archive. The chemistry boundary conditions for the main trace gases and particulate matter concentrations are available from the ECMWF IFS-MOZART model run in the MACC-II

Table 1. Model simulations

	CS1	CS2	DE3	ES1	ES3
Lead Institution	KIT/IMK-IFU*	KIT/IMK-IFU*	TROPOS Leipzig	UMU-MAR	UPM-ESMG
Case	Russian wildfires & Dust event	Russian wildfires	Russian wildfires & Dust event	Russian wildfires & Dust event	Russian wildfires & Dust event
Runs**	NRF, ARI+ACI	NRF, ARI, ARI+ACI	NRF, ARI	NRF, ARI, ARI+ACI	NRF, ARI, ARI+ACI
Spatial Resolution	$23km$	$9.9km$	$0.125°$ (Russian wildfires); $0.25°$ (Dust event)	$23km$	$23km$
Model	WRF-Chem v.3.4.1	WRF-Chem v.3.4.1.	COSMO-MUSCAT	WRF-Chem v.3.4.1.	WRF-Chem v.3.4.1.
Chemical option	RADM2 modified (Stockwell et al., 1990)	RADM2 modified (Stockwell et al., 1990)	RACM-MIM2 (Karl et al., 2006; Tilgner et al., 2008)	RADM2 (Stockwell et al., 1990)	CBMZ (Zaveri and Peters, 1999)
Aerosol option	MADE-SORGAM (Ackermann et al., 1998; Schell et al., 2001)	MADE-SORGAM (Ackermann et al., 1998; Schell et al., 2001)	Simpson et al. (2003)	MADE-SORGAM (Ackermann et al., 1998; Schell et al., 2001)	MOSAIC (Zaveri et al., 2008)
Microphysic option	Morrison et al. (2009)	Lin et al. (1983)	Kessler-type bulk (Doms et al., 2011)	Lin et al. (1983)	Morrison et al. (2009)
Wet deposition	Simple	Easter et al. (2004)		Easter et al. (2004)	Simple
Aqueous chemistry	-	Fahey and Pandis (2001)		Fahey and Pandis (2001)	-

*joint effort with ZAMG, RSE and UPM-ESMG

**NRF: No Radiative Feedbacks, ARI: Aerosol-Radiation Interactions, ARI+ACI: Aerosol-Radiation and Aerosol-Clouds Interactions

project (Inness et al., 2013, Monitoring Atmospheric Composition and Climate) with a grid width of $1.125°$ and a 3-hourly temporal resolution.

Anthropogenic emissions details can be found in Im et al. (2015a, b). They came from the TNO (Netherlands Organization for Applied Scientific Research) MACC emissions inventory (Pouliot et al., 2012; Kuenen et al., 2014; Pouliot et al., 2015, http: 5 //www.gmes-atmosphere.eu/) with a spatial resolution of $\sim$7 km. Annual emissions of methane (CH_4), carbon monoxide (CO), ammonia (NH_3), total non-methane volatile organic compounds (NMVOC), nitrogen oxides (NO_x), particulate matter (PM_{10} & $PM_{2.5}$) and sulphur dioxide (SO_2) were made available by ten activity sectors. As part of the AQMEII and EuMetChem
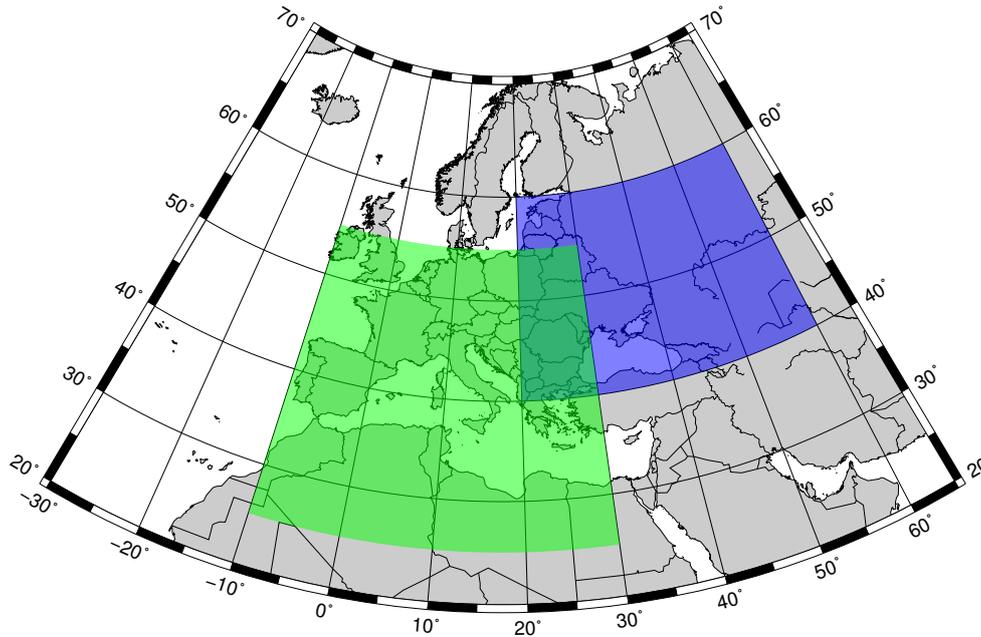
Figure 1. Target (grey) and analysis domains (blue for the Russian wildfires/heat wave case, green for the Saharan desert dust outbreak).

initiatives, temporal profiles (diurnal, day-of-week, seasonal) were provided from Schaap et al. (2005) and vertical distributions were also made available.

The biomass burning emission data of the total PM emissions with a spatial resolution of $0.1°$ were estimated from the project IS4FIRES (Sofiev et al., 2009, Integrated monitoring and modelling system for wild-land fires). Other biomass burning
5 emission species were estimated as detailed in Im et al. (2015b). No heat release due to fires was taken into account.

In the COSMO-MUSCAT simulations, biogenic emissions were treated with the model described in Guenther (1993). Dust emission and transport were computed on the basis of meteorological and hydrological fields from COSMO (Heinold et al., 2007). In the WRF-Chem simulations, MEGAN (Guenther, 2006, Model of Emissions of Gases and Aerosols from Nature) was online coupled with WRF-Chem to estimate the biogenic emissions. Dust emissions were modelled according to Shaw
10 et al. (2008), with an adjustment made to avoid extremely high desert dust fluxes (Forkel et al., 2015).

For the COSMO-MUSCAT simulations, the physics options were: δ-2-stream for long-wave and short-wave fluxes (Ritter and Geleyn, 1992); prognostic turbulent kinetic energy (TKE) planetary boundary layer (PBL); a multi-layer version of the former two-layer soil model after Jacobsen and Heise (1982); the Tiedtke (1989) mass-flux convection scheme. More details for the physical parametrisation are published in Doms et al. (2011). The cloud microphysics and chemistry options are
15 summarised in Table 1.

COSMO-MUSCAT takes into account ARI following Helmert et al. (2007). Radiative fluxes are computed online with the modified COSMO radiation scheme (Ritter and Geleyn, 1992) by considering variations in the modelled size-resolved

aerosol fields. Radiation effects can influence the COSMO meteorology and feedback on emission and aerosol transport. No COSMO-MUSCAT simulations that took into account ACI were used in this work.

The common physics options for the WRF-Chem simulations were: the Rapid Radiative Transfer Method for Global long-wave and short-wave radiation schemes (Iacono et al., 2008, RRTMG); the Yonsei University (YSU) PBL scheme (Hong
5 et al., 2006); the NOAH land-surface model (Chen and Dudhia, 2001); the Grell 3D ensemble cumulus parameterisation (Grell and Dévényi, 2002). The cloud microphysics and chemistry options are summarised in Table 1. However, a more detailed description of the WRF-Chem simulations can be found in Forkel et al. (2015) and Im et al. (2015a, b).

ARI are treated in WRF-Chem following Fast et al. (2006); Chapman et al. (2009) and Barnard et al. (2010). Each chemical constituent of the aerosol was associated with a complex index of refraction. The overall refractive index for a given size
10 bin was determined by volume averaging. The Mie theory and the summation overall size bins were used to determine the composite aerosol optical properties. Wet particle diameters were used in the calculations. Chapman et al. (2009) also described ACI. These interactions were implemented by linking the simulated cloud droplet number with the short-wave radiation and microphysics schemes. Therefore, the droplet number affected both the calculated droplet mean radius and the cloud optical depth when using the short-wave radiation scheme. One limitation of WRF-Chem in the treatment of aerosol-cloud interactions
15 was that these couplings were not computed in the convective clouds simulated by the cumulus parametrisation (Chapman et al., 2009; Yang et al., 2011; Schultz, 2016; Palacios-Peña et al., 2017).

2.2   Observational Data

The observational data used to evaluate the representation of the aerosol optical properties by the EuMetChem simulations came from different remote-sensing instruments: ground-based data from AERONET and different sensors aboard a satellite:
20 the twin MODIS, OMI and SeaWIFS (Sea-viewing Wide Field-of-view Sensor).

The data used from AERONET were AOD of level 2.0 at different wavelengths from the available European stations during these episodes. The variables used to the satellite-AERONET comparison were AOD at the closer wavelength to satellite retrievals and those used to the models-AERONET comparison were AOD 670nm and AE between 440/870nm. Typically, the total uncertainty for the AOD data under cloud-free conditions is $< \pm 0.01$ for $\lambda > 440$ nm and $< \pm 0.02$ for shorter wavelengths
25 (Holben et al., 1998).

For MODIS, the data used were the level 2 of the Atmospheric Aerosol Product (MxD04_L2) from collection 6 (C6) with a resolution of 10 km. Data came from the two available MODIS platforms, Aqua (MYD04_L2) and Terra (MOD04_L2). The MODIS data were estimated by two different algorithms, Dark Target (DT) and Deep Blue (DB). The DT algorithm variables used were: AOD at 550 nm for both ocean (estimated error (EE) $-0.02 - 10\% \times AOD, +0.04 + 10\% \times AOD$ [http://darktarget.
30 gsfc.nasa.gov/products/ocean]) and land (EE $\pm 0.05 + 15\% \times AOD$ [http://darktarget.gsfc.nasa.gov/products/land-10]); and AE between 550 and 840 nm over the ocean (preliminary EE is 0.45 on pixels with an AOD > 0.2)(Levy et al., 2013). The DB algorithm used were: AOD at 550 nm over land, with an EE approximately $\pm 0.05 + 20\% \times AOD$ (Hsu et al., 2013; Sayer et al., 2013); and AE between 412 and 470 nm over land. Finally, a combined variable of the DT and DB algorithms with a wider

coverage was used. This was AOD at 550 nm for both ocean and land, and its error has not yet been estimated (Levy et al., 2013).

The dialy global level-2G, gridded at $0.25°$, data from the algorithm called Near ultra-violet (UV) of OMI sensor were also used. The used products were AOD at different UV range wavelengths (342.5 nm, 388 nm, 442 nm). The typical values for the retrieval error of the AOD obtained by the non-linear fitting routine were below 0.03 (Team, 2009).

Finally, the daily level 3, gridded at $0.5°$, data from the SeaWIFS sensor were used. As with MODIS, this sensor employed the DB algorithm; consequently the error of these products was the same (Hsu et al., 2013; Sayer et al., 2013). The used products were AOD at different wavelengths over land (412 nm, 490 nm and 670 nm), over oceans (510 nm, 670 nm and 865 nm), and both (550 nm); and AE over land between 412/490 nm and over oceans between 510/670 nm.

2.3 Evaluation methodology

The first evaluation step consisted in establishing the "best" satellite data to assess simulations; for this purpose, different satellite data sets were compared with the AERONET observations by calculating the linear regression and the coefficient of correlation between the daily data. Afterwards, the evaluation of the model outputs against the selected satellite database was done by using classical statistics as the individual model-prediction error or bias (e_i), the mean bias error (MBE), the mean absolute error (MAE), the coefficient of correlation (R) and the coefficient of determination (R^2) according to Willmott et al. (1985), Weil et al. (1992) and Willmott and Matsuura (2005).

Satellite data and model data were bilinearly interpolated to a common working grid, which corresponded with the analysis grid (described above) according to case studies. After the interpolation, and in order to compare with AERONET, the satellite and simulations data were extracted from the cell that covered the corresponding station coordinates of the AERONET station by a nearest neighbour approach. When the models were compared with the satellite data, only those data with matched in time and space with the satellite retrieval was selected. The evaluation was done using both available the daily and hourly data.

Finally, to evaluate whether the inclusion of the radiative feedbacks in the simulations would lead to an improvement in the error of the model, two variables were defined: the *Normalized Improvement of the MAE* (NI_MAE = Eq. 1) and the *Normalized Improvement of the Coefficient of Determination* (NI_R^2 = Eq. 2), both in %. These variables were described so that a positive value would indicate an improvement due to the inclusion of the aerosol radiative feedbacks and a negative value; e.g., worsening. $|e_i|$ was the absolute error of the simulations.

$$NI\_MAE = \frac{1}{n}\left(\frac{\sum_1^n |e_i|_{NRF} - \sum_1^n |e_i|_{ARI/ARI+ACI}}{\sum_1^n |e_i|_{NRF}}\right) \times 100 \tag{1}$$

$$NI\_R^2 = \frac{1}{n}\left(\frac{R^2_{ARI/ARI+ACI} - R^2_{NRF}}{R^2_{NRF}}\right) \times 100 \tag{2}$$

Table 2. Coefficient of correlation of the comparison between satellites and the AERONET AOD data

Case	MOD04_L2		MYD04_L2		OMI		SeaWIFS			
Russian wildfires	Deep Blue	0.82	Deep Blue	0.80	$342.5nm$	0.80				
	Combined	0.66	Combined	0.70	$388nm$	0.76				
	Dark Target	0.76	Dark Target	0.79	$442nm$	0.75				
Dust outbreak	Deep Blue	0.45	Deep Blue	0.64	$342.5nm$	0.26	$412nm\_L$ 0.69		$550nm$	0.74
	Combined	0.77	Combined	0.92	$388nm$	0.26	$490nm\_L$ 0.69		$670nm\_O$ 0.90 $670nm\_L$ 0.53	
	Dark Target	0.77	Dark Target	0.92	$442nm$	0.26	$510nm\_O$ 0.89		$865nm\_O$ 0.90	

L: satellite product over land, *O*: satellite product over ocean

3 Results and discussion

The first part of this section identified the satellite data with the best skill according to the AERONET observations, which were considered referential. Then according to these results, the selected data were used to evaluate the representation of the aerosol optical properties by the different EuMetChem simulations during the Russian heat wave and wildfires episode, described in the second part of this section, and the Sahara desert dust outbreak, described in the third part.

3.1 Satellite-AERONET comparison

The results of the linear regression between the daily mean of the different satellite sensors and the AERONET data are shown in this subsection. Table 2 lists the correlation coefficient values for this statistical analysis. AE was excluded from this analysis as satellite data matching with AERONET sites were not enough, which made the evaluation statistically non-significant.

In general, the best skills of the evaluation of the satellite products against AERONET were found for MODIS, as also indicated in Myhre et al. (2005) or Bibi et al. (2015). Focusing on the results obtained during the Russian wildfires episode, the best skills were obtained for the DB algorithm of MODIS (correlation values over 0.80 both from the Terra and Aqua platforms). Indeed the best AOD estimation was obtained from the DB algorithm of MODIS from the Terra platform, with a R^2 of 0.82. The aerosol products retrieved by the DB algorithm provide useful information about aerosol properties over bright-reflecting land surfaces, such as desert, semiarid and urban regions. In C6, the algorithm was enhanced in order to improve retrievals over regions of mixed vegetated and non-vegetated surfaces (Hsu et al., 2013), and to cover all the cloud-free and snow-free land surfaces (Sayer et al., 2014). This fact could explain the good representation of this satellite product over the

study area. The best estimated variable from the OMI data was AOD at 342.5 nm with a correlation value of 0.80. This could be because organic species, as biomass burning aerosols, present a greater absorption at near-ultraviolet and blue wavelengths (Andreae and Gelencsér, 2006; Sun et al., 2007). For this episode, no data from the SeaWIFS sensor were available.

During the Saharan dust episode, the best skills of the AOD satellite products were found for the combined and DT algorithm products of MODIS from the Aqua platform. The correlation values were much higher than for the other satellite products ($R^2 = 0.92$). The DT algorithm was composed of two different algorithms, which were used to retrieve aerosol information over land (dark at visible and longer wavelengths) and over vegetated/dark-soiled land (dark at visible ones) (Sayer et al., 2014). Its combination with the enhanced DB algorithm in C6 provided expanded coverage over most of Europe, which covered most of the domain. The OMI products estimation presented lower correlation values (0.26) than for the Russian wildfires episode. In the dust episode, visible (550nm in MODIS) wavelengths presents better results that UV wavelengths may be because the dust spectral dependence is greater in visible bands (Müller et al., 2009).

SeaWIFS sensor data were available during the Saharan dust episode. The evaluation of the AOD estimation at $670nm$ showed that this sensor produced better estimations over oceans (correlation values of 0.90) than over land (0.53). SeaWIFS data were partnered with an expanded DB algorithm and the SOAR (SeaWiFS Ocean Aerosol Retrieval) algorithm, with retrievals over oceans and inland water bodies (http://disc.sci.gsfc.nasa.gov/gesNews/swdb_monthly_in_Giovanni). Thus using this combination improved retrievals over oceans, as our results showed.

Therefore, in order to evaluate the representation of aerosol optical properties by the modelling experiments for the Russian wildfires episode, the variables of the DB algorithm from MODIS aboard the Terra Platform were used. For the Saharan dust episode, the combined (DT and DB algorithms) AOD at $550nm$ and the AE of the DT algorithm from MODIS aboard the Aqua platform were selected.

3.2 Evaluation of simulations

Once the most skilful satellite data were defined, these data (together with AERONET observations) were used to establish the biases and errors of the modelling experiments for the episodes covered in this work.

3.2.1 Russian wildfires case

The evaluation of the experiments for the Russian wildfires and heat wave episode are detailed in this subsection. The variables of the DB algorithm from MODIS aboard the Terra Platform were used. In order to show more confident results, a mask was implemented for areas where the satellite observations were higher than the 30% of their maximum. Figure 2 shows the evaluation of AOD at $550nm$ and Figure 4 of AE between $412nm$ and $470nm$. For AERONET data, Figure 3 depicts the results for AOD at $670nm$ and Figure 5 for AE between $440/870nm$.

Figure 2a shows the values of AOD measurements by MODIS. The highest values around 2.6 were found over Russia and its surrounding areas due to emissions from the wildfires. The estimation of the mean bias error (MBE) is shown in Figure 2b. All the WRF-Chem simulations (CS1, CS2, ES1, ES3) and the ensemble mean underestimated AOD over the fire-affected areas (minimum MBE values for NRF in Table 3). Over the rest of the domain, a low overestimation (around 0.4, see the maximum

Table 3. Spatial results from evaluating the NRF simulations

		Russian wildfires episode				
	AOD			AE		
Case	MBE		Max. MAE	MBE		Max. MAE
	Min.	Max.		Min.	Max.	
ENSEMBLE	-1.55	0.37	1.58	-0.79	1.09	1.09
CS1	-1.71	0.40	1.71	-0.93	1.09	1.09
CS2	-1.78	0.16	1.84	-0.96	0.93	0.98
DE3	-0.99	2.61	3.13			
ES1	-1.70	0.34	1.70	-0.94	1	1
ES3	-1.82	0.45	1.82	-0.68	1.24	1.24

		Desert dust outbreak episode				
	AOD			AE		
Case	MBE		Max. MAE	MBE		Max. MAE
	Min.	Max.		Min.	Max.	
ENSEMBLE	-0.40	0.18	0.40	-1	0.77	1.16
CS1	-0.38	0.19	0.38	-0.79	1.16	1.16
DE3	-0.47	0.18	0.47			
ES1	-0.41	0.11	0.41	-1	0.99	1.14
ES3	-0.35	0.40	0.42	-1.43	0.33	1.43

MBE values in Table 3) was obtained with the WRF-Chem simulations and the ensemble mean. For DE3, the underestimation was lower and did not cover such a large area as the rest of the experiments. However, a higher overestimation was found in DE3 over the rest of the domain.

WRF-Chem tended to underestimate high AOD values, so these AOD levels were lower than the MODIS levels, especially over the areas with high aerosol loads. When comparing the WRF-Chem experiments, following a modal aerosol approach (CS1, CS2, ES1) led to lower error values than a sectional approach (ES3). Conversely, COSMO-MUSCAT tended to present higher AOD values than WRF-Chem. Thus its underestimation was lower than WRF-Chem, but sometimes led to a high overestimation of the satellite-observed AOD. As mentioned in the Introduction, the ensemble mean generally presented the best skills, as stated for other variables in different works (e.g. Baklanov et al., 2014; Fernández et al., 2009; Knutti et al., 2010; Kjellström et al., 2011).

Generally for the ARI and ARI+ACI simulations, slightly lower MBE values than NRF were found in all the experiments (e.g., in the ES1 simulations: NRF -1.70; ARI -1.66; ARI+ACI 1.64). However, the MBE for the ensemble did not show this
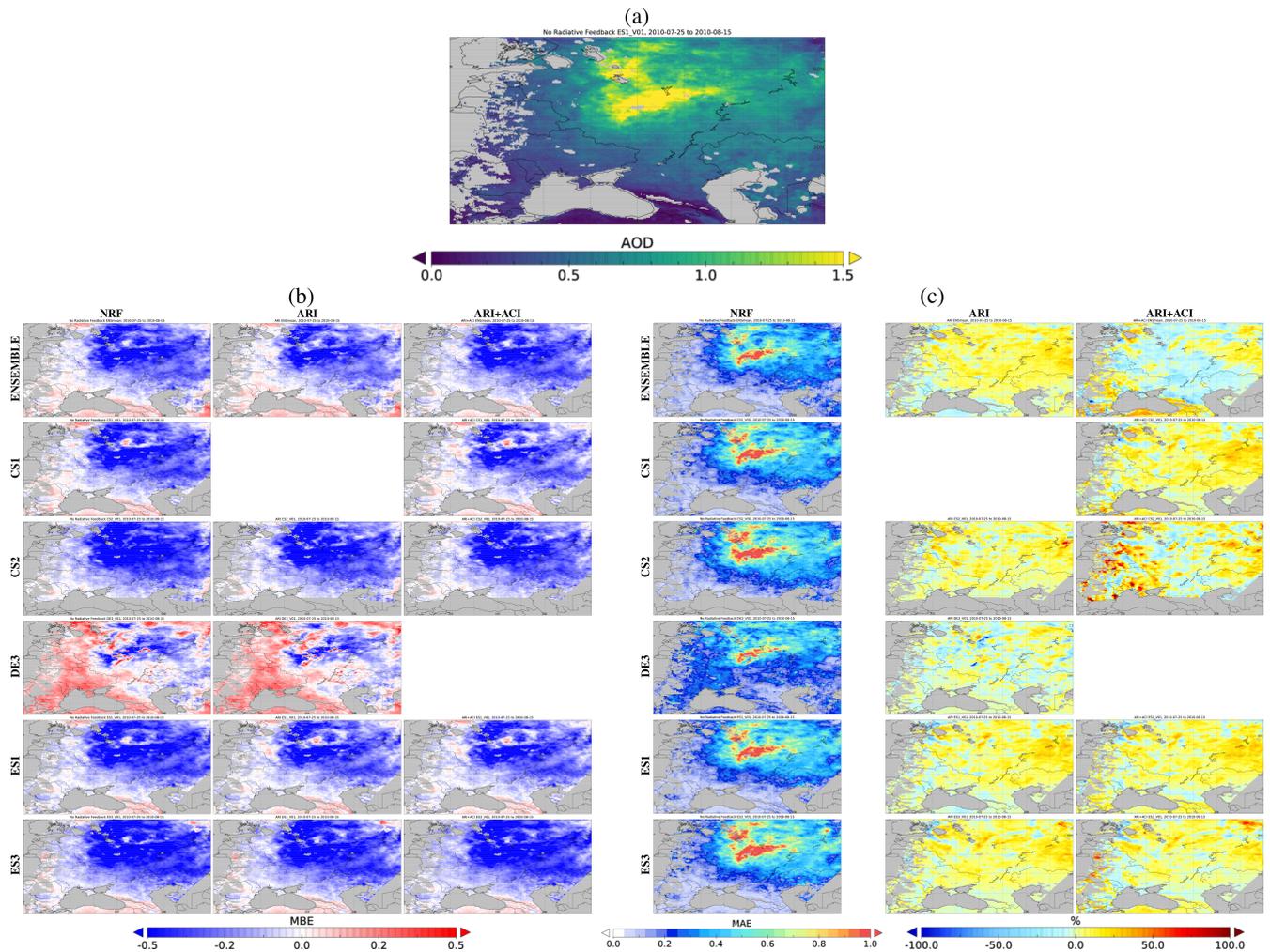
Figure 2. Model-MODIS comparison of AOD at $550nm$ for the Russian wildfires case: (a) Satellite values; (b) MBE for NRF (first column), ARI (second) and ARI+ACI (third); and (c) MAE for the NRF simulations (first column) and their improvements due to ARI (second) and ARI+ACI (third).

improvement with values of -1.55 in NRF scenario, -1.49 in ARI and -1.71 in ARI+ACI. Its analysis should be carefully taken into account because the ARI ensemble mean did not include the CS1 simulation and the ARI+ACI, the DE3 simulation.

The MBE indicates the bias of the model error, that means the tendency to over- or underestimate satellite observations but the mean average error (MAE, Table 3) indicates the magnitude of the model errors. The maximum MAE values over Russia and the surrounding areas were similar or came close to the maximum or minimum MBE values, which indicated that the biases observed above moved in the same way, positive or negative, throughout the simulations period.

As indicated by Palacios-Peña et al. (2017), the AOD discrepancies between simulations and observations can be attributed to the errors in the model estimations of the aerosol dry mass, the fraction of particles for a given mass or water associated

with aerosols. According to the aerosol dry mass estimation, Im et al. (2015b) found an overestimation in the simulated $PM_{2.5}$ levels for the AQMEII phase 2 simulations which are comparable to EuMetChem simulations, and Soares et al. (2015) established a rough overestimation at about 50% of the total biomass burning emissions used here. According to this evidence, an overestimation of the AOD levels over the fires area can be expected for this episode. However, our results showed an

5   AOD underestimation. This behaviour could be due to the understated injection height of the total biomass burning emissions found by Soares et al. (2015). On the other hand, the heat released by the fires, which is responsible for the strong updrafts transporting the emitted tracers rapidly to the free troposphere and even the stratosphere (Freitas et al., 2007), was not taken into account in the simulations and could affect the aerosol vertical distribution. In other words, an inaccurate description in the simulation of the aerosol vertical distribution could affect the representation of the aerosol optical depth. We established

10   this statement following Kipling et al. (2016), which found the extent to which the biomass burning emissions to the free troposphere affected the vertical profile of aerosols, causing changes in AOD and radiative forcing with the HadGEM3-UKCA model.

The estimated NI_MAE (Figure 2c) indicates whether the inclusion of aerosol radiative feedbacks reduces the magnitude of model errors; i.e., , whether the inclusion of ARI or ARI+ACI improves the skills of the models as regards the absolute error.

15   The results indicated a generalized improvement of the MAE over the whole domain. For ARI, the ensemble mean presented an improvement above 30%, but the simulations with the greatest NI_MAE were ES1 ($> 40\%$, over a large area of the domain) and CS2 (NI_MAE up to 90%). These improvements took place especially over the eastern Europe where medium-high AOD values were found. For ARI+ACI over the same areas, the ensemble mean presented an improvement of up to 45%. However in the southwest area of the domain, improvements reached 70%. Over the fire-affected areas, the experiment that led to major

20   improvement was CS1 ($> 50\%$).

The causes for the variation in AOD levels has to be sought in a number of physical mechanisms. Péré et al. (2014) and Baró et al. (2017) found a reduction in the atmospheric boundary layer height and the wind speed over most of eastern Europe when the ARI were taken into account in their simulations to study the same wildfires episode. Due to these effects on meteorology, Zhang et al. (2015) revealed a effect on aerosol concentrations which could explain the observed improvement of AOD over

25   the eastern part of the domain observed herein.

Regarding the determination coefficient (Figure S1, which only show the areas where correlation results were significant at 90%), no clear spatial pattern was associated with this parameter or with its improvement. However, over the whole domain the ensemble mean presented higher values than each specific experiment for R^2, and a clear improvement in the determination coefficient was observed when ARI+ACI were taken into account.

30   The hourly AOD values recorded at 670nm of the different simulations correlated with the AERONET values at this same wavelength. Only the correlation results that were significant at the 95% level in a certain station are shown in Figure 3a. The Model-AERONET R values were > 0.4 in nearly 75% of the stations, and were above 0.5 in nearly 50% of the stations used for all the experiments.

The station where ensemble means showed the best skills was the Toravere station. with R in NRF of 0.68, ARI 0.73 and

35   ARI+ACI 0.73 (Figure 3b). At this station, located in northerly areas, AOD was higher between 25 and 30 July, and between

5 and 10 of August. However, the maximum correlation values among all the experiments were found at the Efoire station (Figure 3c) for the CS2 configuration with correlation coefficient for NRF of 0.83, ARI 0.84 and ARI+ACI 0.82. This can be explained by the enhanced resolution of CS2 around 9.9km. This fine resolution may lead to improvements in the local representation of AOD. It should be highlighted that no clear improvement in the model-AERONET correlations was noted
5 when the aerosol radiative feedbacks were taken into account.

The MODIS AE satellite values are shown in Figure 4a. AE is a variable that provides an idea of particle size in the atmosphere. High AE values indicate fine particles and low values denote coarse particles. High values between 1.7 and 1.8 were found over Russia and the surrounding areas given the fine particles emitted by wildfires. Low AE values were found over the southeast areas of the domain. This fact could be explained by the presence of coarse particles from the mineral dust
10 that came from nearby desert areas.

A general MBE overview (Figure 4b) indicated that all the experiments underestimated high AE values and overestimated low AE ones. Hence the model underestimated the variability of this variable. Over the Russian wildfires area, where the highest AE values were found, simulations underestimated (minimum MBE values for NRF in Table 3) this variable. Over the area with the lowest AE values in the southeast, the model produced an overestimation (maximum MBE values in Table 3).
15 According to these results, and in the same way as AOD, the most skilful simulation was the ensemble mean. ES3, the WRF-Chem simulations using a sectional approach, presented higher AE values than the other experiments and thus, the lowest minimum and highest maximum MBE. No simulations for the DE3 experiments were available for AE.

Figure 4(c) shows the MAE values and the improvements for AE. Maximum MAE values (Table 3) were found over the southeast of the domain. Only ES3 experiment, which used the aerosol sectional approach, presented an improvement to AE
20 representation when ARI+ACI were taken into account (Figure 4c, right column). This improvement, with values around 20%, was found over the areas where the MAE values were lower.

The determination coefficient values (Figure S1, where only the areas where correlation results were significant at 90% are shown) of the evaluation made of AE were worse than they were for AOD. The highest values above 0.4 were observed over the Russian wildfires areas. Values below 0.25 were found for the rest of the domain. Although the minimum MAE values
25 were shown for ES3, this experiment presented a worse determination coefficient over the Russian wildfires area than the other experiments. Despite there being no clear improvement pattern for the determination coefficient, the ARI+ACI scenario showed isolated improvements with values between 50% and 100%.

Regarding the model-AERONET comparison (Figure 5a), the number of stations with significant results for AE data was very limited and substantially lower than for AOD and the results show lower correlation coefficient values. Between 10% and
30 40% of the used stations obtained model-AERONET correlations above 0.5, depending on the model used. The largest number (> 40%) of stations with correlation coefficients >0.5 was found for the ensemble mean. The station where the ensemble mean had the most skills was Bucharest with correlation coefficient for NRF of 0.72, ARI 0.69 and ARI+ACI 0.74 (Figure 5b). For this variable, no station had higher correlation coefficient values for any specific experiment. As for AOD, the inclusion of aerosol radiative feedbacks did not improve the AE representation.

3.2.2 Saharan desert dust outbreak case

This section describes the evaluation of the representation of AOD and AE by the simulations done for the Saharan dust episode (Figures 6 to 9). In this case, CS2 simulations were not available and the combined (DT and DB algorithms) AOD at $550nm$ (Figure 6) and the AE of the DT algorithm (Figure 8) from MODIS aboard the Aqua platform were used. As with the Russian wildfires episode, the mask of 30% of the maximum of observations was implemented. The model-AERONET comparisons of AOD at $670nm$ (Figure 7) and AE between $440/870nm$ (Figure 9) were also shown.

Figure 6 shows the AOD values measured by MODIS. Over the southeast area of the domain, values of >0.6 were observed due to transported dust. This value was not very high for a dust outbreak, but was caused by wet deposition (rain during the episode) and was explained in the event description in the section "Model simulations".

In Figure 6b, all the experiments underestimated high AOD values over the southeast area of the domain. The MBE values over this area were around -0.3 for DE3, but were lower, around -0.2, for the other experiments (WRF-Chem simulations). However small areas with a higher underestimation were found over this zone (minimum MBE values for NRF in Table 3). Over the rest of the domain, minor overestimations were modelled with MBE values around 0.1. Conversely, small sporadic areas with high overestimations were found (maximum MBE values for NRF in Table 3).

In order to consider the magnitude of these under/overestimations, the MAE is depicted in Figure 6c (left column). Over the southeast of the domain, the MAE gave values around 0.2 for all the experiments, except for DE3, which showed a MAE value >0.3. The most relevant improvements to this statistic (Figure 6c, centre and right column) were observed for the ARI+ACI ensemble mean, with improvements up to 25%, over desert areas. Smaller areas over the Mediterranean Sea with improvements up to 75% were found for the ES3 simulation. These improvements could be explained by a better representation of the coarse mode in the sectional approach.

The determination coefficient values (Figure S2) were higher than for the Russian wildfires episode. These values came close to 1 for all the experiments over most of the domain. No improvement in the representation of R^2 when ARI and ACI were taken into account was observed. For both variables AOD and AE, the established requirement of showing only the areas where correlation results were significant at 90% was implemented.

Both WRF-Chem and COSMO-MUSCAT tended to underestimate high AOD values. COSMO-MUSCAT was the model with the highest underestimation. Regarding the WRF-Chem simulations, ES3, which used the sectional approach, presented higher AOD values and thus, lower minimum and higher maximum MBE values, than the other experiments. According to Im et al. (2015b), this underestimation can be explained by a systematic underestimation for all the models of the PM_{10} levels with the highest underestimation for the Mediterranean Region. Moreover, this major underestimation can be attributed to natural emissions, such as the desert dust studied in this episode.

Regarding the model-AERONET comparison (Figure 7a), the ensemble mean showed the best skills at the Toulon station with correlation values of 0.98, 0.98 and 0.96 for NRF, ARI and ARI+ACI, respectively (Figure 7b). According to these values, it was not possible to establish an improvement when aerosol radiative feedbacks were taken into account. Besides, more than 60% of the stations used gave correlation values above 0.5 for all the experiments. Around 80% of the used stations

presented this behaviour for ES1. For these experiments, the highest correlation values were above 0.90 at stations like Toulon (Figure 7b), Efoire (Figure 7c) or Thessaloniki (not shown).

Although the number of data per station was smaller for this episode than for the Russian wildfires episode, all the shown stations complied with the established requirement of the correlation results, which were significant at the 95% level (explained in the previous section).

For AE, the MODIS satellite values are shown in Figure 8a. Values between 0 and 1 were found over the southeast area of the domain due to coarse dust particles. When the bias was analysed (Figure 8b, Table 3), ES3, which uses the aerosol sectional approach, displayed a different behaviour from the rest of the experiments. ES3 underestimated low AE levels over the area affected by the Saharan dust outbreak, with a minimum MBE for NRF of -1.43. An improvement to the AE representation was also achieved for ARI+ACI, with a minimum MBE error of -1.31 (this could also be observed for the NI_MAE, explained below). Small overestimations of the values with a maximum MBE of 0.33 were found for NRF. The other experiments exhibited similar behaviour as they did for the Russian wildfires episode. Low values over the southeast of the domain were overestimated and high values were underestimated.

MAE (Figure 8c, left column) estimates the magnitude of errors. Maximum MAE errors (Table 3) were found over the dust outbreak-affected area. Regarding improvements to this statistical figure shown in Figure 8c (centre and right column), the ensemble mean presented improvements of up to 50% over most of the domain when ARI were included. However, when considering ARI+ACI, ES3 had the highest improvement values with values of up to 75% over the dust outbreak-affected area. This high improvement could be caused because the inclusion of ARI+ACI affects relative humidity (Baró et al., 2017) modifying the size of the particles due to the hygroscopic growth and hence may improve the representation of the particles size distribution.

As for AOD, the determination coefficient values (Figure S2) were higher than for the Russian wildfires episode and came close to 1 over most of the domain. All the experiments provided these R^2 values, with no significant improvement to this statistical figure when ARI and ACI were taken into account.

Finally, the ensemble and CS1 presented the best skills when they were compared with AERONET (Figure 9a). Athens (Figure 9b) was the station where the ensemble mean presented its best skill with correlation coefficients of 0.96, 0.91 and 0.92 for NRF, ARI and ARI+ACI respectively; and Helgoland (Figure 9c), done by CS1 simulations with correlation values for NRF of 0.96 and ARI+ACI 0.95. The latter station did not present any strong dust outbreak influence. It was not possible to establish a significant improvement when aerosol radiative feedbacks were taken into account. Notwithstanding, more than 40% of the used stations presented correlation values above 0.5 for all the experiments. For the ensemble mean, correlation values above 0.5 were found from 70 to 80% of the used stations.

4 Summary and Conclusions

This work attempts to identify the skill of a set of satellite observations and online coupled models to represent the aerosol optical properties during two important episodes: a biomass burning episode and a dust outbreak episode that occurred over

Europe in 2010. The first task consisted in identifying the skills of different satellite observations compared with AERONET. In this sense, MODIS showed the highest correlation coefficient with the ground-based observations. The DB algorithm of MODIS provided the best skilful products for the Russian wildfires/heat wave case. This was because the algorithm was enhanced over regions of mixed vegetated and non-vegetated surfaces, like the study area. The SOAR algorithm also provided

5   a high quality product over oceans when it was used for the SeaWIFS sensor during the dust outbreak episode because SOAR was developed to obtain retrievals over oceans and inland water bodies. However, the most skilful AOD product for the dust outbreak episode was that provided by MODIS combined DB and DT algorithm. This combined product provided an expanded coverage over bright-reflecting land surfaces, such as desert, semiarid and urban regions, which was the case of this dust outbreak. For AE, there were not enough data to estimate how good its representation was.

10   Using the best-available satellite information, a set of regional online chemistry-climate/meteorology models over Europe were evaluated for the two aforementioned episodes. The results for the Russian wildfires episode indicated that AOD was better represented by the models than AE. AOD was generally underestimated by all simulations. This behaviour could be explained by a misunderstanding in the aerosol vertical distribution due to the understated injection height of the total biomass burning emissions or because the heat released by the fires was not taken into account, which could affect the representation of

15   AOD. The underestimation by WRF-Chem was higher than by COSMO-MUSCAT, and the latter simulated higher AOD values than all the WRF-Chem configurations. Among the WRF-Chem experiments, the modal aerosol approach (CS1, CS2 and ES1) presented lower error values than the sectional approach (ES3). The inclusion of ARI+ACI led to an improvement above 30% to the MAE of the ensemble mean. The modification on these AOD levels could be explained by a reduction in the atmospheric boundary layer height and the wind speed over most of eastern Europe when the ARI were taken into account. It could affect the

20   aerosol concentrations and modify the AOD representation. The ensemble mean was more skillful than the single simulations compared with the satellite products. However, at a local level, the fire episode simulation at a fine resolution (CS2) presented the highest model-AERONET correlations (around 0.8). No clear improvement to the model-AERONET correlations was found when ARI and ACI were included.

Therefore, In the Russian wildfires case, the MBE, MAE and R^2 values were lower for AE than for AOD. All the experiments under-

25   estimated the variability of AE (i.e., low values were represented by too high simulated values and high values by lower values than actually observed). According to all the calculated statistics, the most skillful experiment was the ensemble mean. When considering individual models, the MBE indicated that ES3, the only WRF-Chem simulation that used a sectional approach, presented higher AE values than the other experiments, and thus resulted in a lower underestimation. Despite these higher values, ES3 was the only experiment to present improvements (around 20%) when ARI+ACI were taken into account. For

30   R^2 no clear improvement was noted when aerosol radiative feedbacks were included. For AE, the highest model-AERONET correlations were found for the ensemble mean and they were lower (around 0.7) than for AOD.

The Saharan desert dust outbreak case results indicated that high AOD values, which were found over the southeast of the domain (dust-affected area), were underestimated by all the models and configurations. According to the results from Im et al. (2015b), this underestimation can be explained by an underestimation in the PM_{10} levels, attributed mainly to natural

35   emissions, such as desert dust. DE3 presented the most pronounced underestimation. The ES3 experiments, which used the

sectional approach, estimated higher AOD values than the other simulations. The most obvious improvement to the skill from including ARI went to the ensemble mean, with an improvement up to 25%. However, we can find a better improvement due to the inclusion of ARI+ACI over the Mediterranean Sea with values around 75%, which can be explained by a better representation of the coarse mode in the sectional approach when the radiative feedbacks were taking into account. R^2 for AOD during the dust outbreak episode obtained values close to 1 over most of the domain. Although this statistical figure showed better results for this episode than for the Russian wildfires episode, no clear improvement for the aerosol radiative feedbacks was obtained. When models and AERONET were compared, the ensemble mean provided the best skill with correlation values above 0.9. More than 60% of the used stations presented correlation values above 0.5 for all the experiments. ES1 showed skills for correlations over 0.5 in 80% of the used stations, sometimes with correlations >0.9.

Finally, the representation of AE in the dust outbreak case presented similar performance to that for the Russian wildfires episode. The modelling experiments underestimated the variability of this variable, with ES3 providing lower AE values than the other experiments. However for ARI+ACI, an improvement by up to 75% was found for the ES3 experiment. This could be due to a better representation of the aerosol size distribution as a result of the hygroscopic growth. The ensemble mean offered the best skill and showed an improvement by up to 50% when ARI were taken into account. The AE determination coefficient was better for this episode than for the Russian wildfires episode. Locally, the model-AERONET comparison indicated that the ensemble mean and CS1 presented the best skills, with correlation values above 0.9. For the ensemble mean, correlation values above 0.5 were found for 70 to 80% of the used stations; this number was lower (around 40%)for the single experiments.

By way of conclusion, the modelling experiments generally underestimated high AOD values. According to Palacios-Peña et al. (2017), AOD discrepancies between simulations and observations can be attributed to errors in the model estimation of the aerosol dry mass, the fraction of particles for a given mass or the water associated with aerosols. A model underestimation of the aerosol dry mass of dust can explain the AOD underestimation during the Saharan desert dust episode because of the underestimation in the PM_{10} levels found by Im et al. (2015b). However, this fact cannot explain the AOD underestimation obtained during the Russian wildfires episode, when a misunderstanding in the aerosol vertical distribution simulation, due to an understated injection height of the total biomass burning emission or because of the heat released by the fires was not taken into account, could affect the AOD representation. Locally, fine resolutions can help improving the representation of this variable. For AE, the aerosol representation by the modal WRF-Chem approach resulted in a substantial underestimation in this variable's variability. The sectional WRF-Chem approach provided high AE values for the case with fine fire particles and low values for the case with coarse particles (dust). Thus the bin representation for different particle sizes strongly influenced the estimation of this variable when a sectional approach was used. Hence the use of a sectional approach improved the size distribution representation in the models, but led to a worse skill to represent AOD.

Albeit sometimes an individual experiment may show better results, as we assumed in the first section, the ensemble mean generally provides the most skillful simulation. Regarding the inclusion of aerosol radiative feedbacks, the observed improvements (especially for a sectional approach) justify not only the inclusion of radiative feedbacks in online coupled models, but also the increase in the computational time to include these effects. Further studies are necessary to improve the model representation of aerosol optical properties.

Acknowledgements. The authors acknowledge Project REPAIR-CGL2014-59677-R of Spanish Ministry of the Economy and Competitiveness and the FEDER European program for support to conduct this research. Work was possible thanks to the fellowship 19677/EE/14 funded by Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia, Programme Jiménez de la Espada de Movilidad, Cooperación e Internacionalización, within the framework of II PCTIRM 2011-2014 and the fellowship FPU14/05505 funded by the Spanish Ministry of Education, Culture and Sport. The authors thank the support from EuMetChem COST ACTION ES1004 and the Air Quality Modelling Evaluation International Initiative (AQMEII). We also thank the researchers and their staff for establishing and maintaining the European AERONET sites used in this research in the network AERONET, and many others who have been involved in the MODIS, OMI and SeaWIFS 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, 1998.

Aminou, D. M., Jacquet, B., and Pasternak, F.: Characteristics of the Meteosat Second Generation (MSG) radiometer/imager: SEVIRI, in: Aerospace Remote Sensing'97, pp. 19–31, International Society for Optics and Photonics, 1997.

Aminou, D. M., Ottenbacher, A., Jacquet, B., and Kassighian, A.: Meteosat Second Generation: on-ground calibration, characterization, and sensitivity analysis of the SEVIRI imaging radiometer, in: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, pp. 419–430, International Society for Optics and Photonics, 1999.

Andreae, M. and Gelencsér, A.: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols, Atmospheric Chemistry and Physics, 6, 3131–3148, 2006.

Annan, J. and Hargreaves, J.: Reliability of the CMIP3 ensemble, Geophysical Research Letters, 37, 2010.

Antuña-Marrero, J. C., Landulfo, E., Estevan, R., Barja, B., Robock, A., Wolfram, E., Ristori, P., Clemesha, B., Zaratti, F., Forno, R., Armandillo, E., Bastidas, A. E., de Frutos Baraja, A. M., Whiteman, D. N., Quel, E., Barbosa, H. M. J., Lopes, F., Montilla-Rosero, E., and Guerrero-Rascado, J. L.: LALINET: The first Latin American-born regional atmospheric observational network., Bulletin of the American Meteorological Society, doi:10.1175/BAMS-D-15-00228.1, http://dx.doi.org/10.1175/BAMS-D-15-00228.1, 2016.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G., Hirtl, M., Joffre, S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A., Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N., Rao, S. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmospheric Chemistry and Physics, 14, 317–398, doi:10.5194/acp-14-317-2014, http://www.atmos-chem-phys.net/14/317/2014/, 2014.

Bangert, M., Nenes, A., Vogel, B., Vogel, H., Barahona, D., Karydis, V., Kumar, P., Kottmeier, C., and Blahak, U.: Saharan dust event impacts on cloud formation and radiation over Western Europe, Atmospheric Chemistry and Physics, 12, 4045, 2012.

Barnard, J. C., Fast, J. D., Paredes-Miranda, G., Arnott, W., 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, 2010.

Baró, R., Lorente-Plazas, R., Montávez, J., and Jiménez-Guerrero, P.: Biomass burning aerosol impact on surface winds during the 2010 Russian heat wave, Geophysical Research Letters, 44, 1088–1094, 2017.

Bellouin, N., Boucher, O., Haywood, J., and Reddy, M. S.: Global estimate of aerosol direct radiative forcing from satellite measurements, Nature, 438, 1138–1141, 2005.

Bibi, H., Alam, K., Chishtie, F., Bibi, S., Shahid, I., and Blaschke, T.: Intercomparison of MODIS, MISR, OMI, and CALIPSO aerosol optical depth retrievals for four locations on the Indo-Gangetic plains and validation against AERONET data, Atmospheric Environment, 111, 113–126, 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 Fourth Assessment Report of the International Panel of Climate Change, pp. 571–657, Cambridge University Press, 2013.

Chapman, E. G., Gustafson Jr, W., 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, 2009.

Chen, F. and Dudhia, J.: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I:
5   Model implementation and sensitivity, Monthly Weather Review, 129, 569–585, 2001.

Chen, S., Zhao, C., Qian, Y., Leung, L. R., Huang, J., Huang, Z., Bi, J., Zhang, W., Shi, J., Yang, L., Li, D., and Li, J.: Regional modeling of dust mass balance and radiative forcing over East Asia using WRF-Chem, Aeolian Research, 15, 15–30, 2014.

Chou, M.-D., Chan, P.-K., and Wang, M.: Aerosol radiative forcing derived from SeaWiFS-retrieved aerosol optical properties, Journal of the Atmospheric Sciences, 59, 748–757, 2002.

10 Chubarova, N., Nezval, Y., Sviridenkov, I., Smirnov, A., and Slutsker, I.: Smoke aerosol and its radiative effects during extreme fire event over Central Russia in summer 2010, Atmospheric Measurement Techniques, 5, 557–568, 2012.

Curci, G., Hogrefe, C., Bianconi, R., Im, U., Balzarini, A., Baró, R., Brunner, D., Forkel, R., Giordano, L., Hirtl, M., et al.: Uncertainties of simulated aerosol optical properties induced by assumptions on aerosol physical and chemical properties: An AQMEII-2 perspective, Atmospheric Environment, 115, 541–552, 2015.

15 Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A., Gordon, H., Muller, J.-P., Myneni, R., Sellers, P., Pinty, B., and Verstraete: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, IEEE Transactions on Geoscience and Remote Sensing, 36, 1072–1087, 1998.

Doms, G., Förstner, J., Heise, E., Herzog, H., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz, J., and Vogel, G.: A description of the non-hydrostatic regional COSMO model, Part II: Physical Parameterization., http://www.cosmo-model.org/, 2011.

20 Dubovik, O., Smirnov, A., Holben, B., King, M., Kaufman, Y., Eck, T., and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, Journal of Geophysical Research: Atmospheres, 105, 9791–9806, 2000.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., and Slutsker, I.: Variability of absorption and optical properties of key aerosol types observed in worldwide locations, Journal of the atmospheric sciences, 59, 590–608, 2002.

25 Easter, R. C., Ghan, S. J., Zhang, Y., Saylor, R. D., Chapman, E. G., Laulainen, N. S., Abdul-Razzak, H., Leung, L. R., Bian, X., and Zaveri, R. A.: MIRAGE: Model description and evaluation of aerosols and trace gases, Journal of Geophysical Research: Atmospheres, 109, 2004.

Fahey, K. M. and Pandis, S. N.: Optimizing model performance: variable size resolution in cloud chemistry modeling, Atmospheric Environment, 35, 4471–4478, 2001.

30 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, 2006.

Fernández, J., Primo, C., Cofino, A., Gutiérrez, J. M., and Rodríguez, M. A.: MVL spatiotemporal analysis for model intercomparison in EPS: application to the DEMETER multi-model ensemble, Climate dynamics, 33, 233–243, 2009.

35 Forkel, R., Balzarini, A., Baró, R., Bianconi, R., Curci, G., Jiménez-Guerrero, P., Hirtl, M., Honzak, L., Lorenz, C., Im, U., Pérez, J., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., and Žabkar, R.: Analysis of the WRF-Chem contributions to AQMEII phase 2 with respect to aerosol radiative feedbacks on meteorology and pollutant distributions, Atmospheric Environment, 115, 630–645, 2015.

Forkel, R., Brunner, D., Baklanov, A., Balzarini, A., Hirtl, M., Honzak, L., Jiménez-Guerrero, P., Jorba, O., Pérez, J., San José, R., Schöder, W., Tsegas, G., Werhahn, J., Wolke, R., and Žabkar, R.: 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.

5 Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M., Andreae, M., Prins, E., Santos, J., Gielow, R., and Carvalho Jr, J.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, Atmospheric Chemistry and Physics, 7, 3385–3398, 2007.

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

10 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, http://www.atmos-chem-phys.net/15/8217/2015/, 2015.

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

Galmarini, S., Hogrefe, C., Brunner, D., Makar, P., and Baklanov, A.: Preface, Atmospheric Environment, 115, 340 – 344, 2015.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G.,

15 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, http://www.atmos-chem-phys.net/17/1543/2017/, 2017.

Garland, R. M., Yang, H., Schmid, O., Rose, D., Nowak, A., Achtert, P., Wiedensohler, A., Takegawa, N., Kita, K., Miyazaki, Y., Kondo, Y., Hu, M., Shao, M., Zeng, L. M., Zhang, Y. H., Andreae, M. O., and Pöschl, U.: Aerosol optical properties in a rural environment near the

20 mega-city Guangzhou, China: implications for regional air pollution, radiative forcing and remote sensing, Atmospheric Chemistry and Physics, 8, 5161–5186, doi:10.5194/acp-8-5161-2008, http://www.atmos-chem-phys.net/8/5161/2008/, 2008.

Grell, G. A. and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophysical Research Letters, 29, 2002.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry

25 within the WRF model, Atmospheric Environment, 39, 6957–6975, 2005.

Guenther, A.: Isoprene and monoterpene emission rate variability: model evaluations and sensitivity analyses, Journal of Geophysical Research, 98, 1993.

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.

30 Haywood, J. M., Francis, P. N., Glew, M. D., and Taylor, J. P.: Optical properties and direct radiative effect of Saharan dust: A case study of two Saharan dust outbreaks using aircraft data, Journal of Geophysical Research: Atmospheres, 106, 18 417–18 430, 2001.

Heinold, B., Helmert, J., Hellmuth, O., Wolke, R., Ansmann, A., Marticorena, B., Laurent, B., and Tegen, I: Regional modeling of Saharan dust events using LM-MUSCAT: Model description and case studies, Journal of Geophysical Research: Atmospheres, 112, 2007.

Helmert, J., Heinold, B., Tegen, I., Hellmuth, O., and Wendisch, M.: On the direct and semidirect effects of Saharan dust over Europe: A

35 modeling study, Journal of Geophysical Research: Atmospheres, 112, 2007.

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

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Monthly Weather Review, 134, 2318–2341, 2006.

Hsu, N., Jeong, M.-J., Bettenhausen, C., Sayer, A., Hansell, R., Seftor, C., Huang, J., and Tsay, S.-C.: Enhanced Deep Blue aerosol retrieval algorithm: The second generation, Journal of Geophysical Research: Atmospheres, 118, 9296–9315, 2013.

Huang, H., Thomas, G., and Grainger, R.: Relationship between wind speed and aerosol optical depth over remote ocean, Atmospheric Chemistry and Physics, 10, 5943–5950, 2010.

Huang, J., Minnis, P., Lin, B., Wang, T., Yi, Y., Hu, Y., Sun-Mack, S., and Ayers, K.: Possible influences of Asian dust aerosols on cloud properties and radiative forcing observed from MODIS and CERES, Geophysical Research Letters, 33, 2006.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, 2008.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J. J., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G., Jose, R. S., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Evaluation of operational on-line-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part I: Ozone, Atmospheric Environment, 115, 404–420, 2015a.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., van der Gon, H. D., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G., Jose, R. S., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Wang, K., Werhahn, J., Wolke, R., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Evaluation of operational online-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part II: Particulate matter, Atmospheric Environment, 115, 421 – 441, 2015b.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmospheric Chemistry and Physics, 13, 4073–4109, doi:10.5194/acp-13-4073-2013, http://www.atmos-chem-phys.net/13/4073/2013/, 2013.

Jacobsen, I. and Heise, E.: A new economic method for the computation of the surface temperature in numerical models, Beitraege zur Physik der Atmosphaere, 55, 128–141, 1982.

Karl, M., Dorn, H.-P., Holland, F., Koppmann, R., Poppe, D., Rupp, L., Schaub, A., and Wahner, A.: Product study of the reaction of OH radicals with isoprene in the atmosphere simulation chamber SAPHIR, Journal of atmospheric chemistry, 55, 167–187, 2006.

Kaufman, Y. J., Tanré, D., and Boucher, O.: A satellite view of aerosols in the climate system, Nature, 419, 215–223, 2002.

Kipling, Z., Stier, P., Johnson, C. E., Mann, G. W., Bellouin, N., Bauer, S. E., Bergman, T., Chin, M., Diehl, T., Ghan, S. J., Iversen, T., Kirkevåg, A., Kokkola, H., Liu, X., Luo, G., van Noije, T., Pringle, K. J., von Salzen, K., Schulz, M., Seland, Ø., Skeie, R. B., Takemura, T., Tsigaridis, K., and Zhang, K.: What controls the vertical distribution of aerosol? Relationships between process sensitivity in HadGEM3–UKCA and inter-model variation from AeroCom Phase II, Atmospheric Chemistry and Physics, 16, 2221–2241, doi:10.5194/acp-16-2221-2016, http://www.atmos-chem-phys.net/16/2221/2016/, 2016.

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G., and Ullerstig, A.: 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations, Tellus A, 63, 24–40, 2011.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in combining projections from multiple climate models, Journal of Climate, 23, 2739–2758, 2010.

5 Kong, X., Forkel, R., Sokhi, R. S., Suppan, P., Baklanov, A., Gauss, M., Brunner, D., Barò, R., Balzarini, A., Chemel, C., et al.: Analysis of meteorology–chemistry interactions during air pollution episodes using online coupled models within AQMEII phase-2, Atmospheric Environment, 115, 527–540, 2015.

Konovalov, I., Beekmann, M., Kuznetsova, I., Yurova, A., and Zvyagintsev, A.: Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region, Atmospheric Chemistry and Physics, 11,
10 10 031, 2011.

Kuenen, J., Visschedijk, A., Jozwicka, M., and Denier Van der Gon, H.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmospheric Chemistry and Physics, 14, 10 963–10 976, 2014.

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

Levy, R. C., Remer, L. A., and Dubovik, O.: Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, Journal of Geophysical Research: Atmospheres, 112, 2007.

Li, S., Wang, T., Xie, M., Han, Y., and Zhuang, B.: Observed aerosol optical depth and Ångström exponent in urban area of Nanjing, China, Atmospheric Environment, 123, 350–356, 2015.

20 Lin, Y.-L., Farley, R. D., and Orville, H. D.: Bulk parameterization of the snow field in a cloud model, Journal of Climate and Applied Meteorology, 22, 1065–1092, 1983.

Makar, P., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, 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., San José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on
25 chemistry, Atmospheric Environment, 115, 499–526, 2015a.

Makar, P., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., 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 1: Effects on weather, Atmospheric Environment, 115, 442 – 469, doi:https://doi.org/10.1016/j.atmosenv.2014.12.003, 2015b.

30 Mishra, A. K., Koren, I., and Rudich, Y.: Effect of aerosol vertical distribution on aerosol-radiation interaction: A theoretical prospect, Heliyon, 1, e00 036, 2015.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes, Monthly Weather Review, 137, 991–1007, 2009.

Müller, T., Schaladitz, A., Massling, A., Kaaden, N., Kandler, K., and Wiednsohler, A.: Spectral absorption coefficients and imaginary parts
35 of refractive indices of Saharan dust during SAMUM-1, Tellus B, 61, 79–95, doi:10.1111/j.1600-0889.2008.00399.x, http://dx.doi.org/10.1111/j.1600-0889.2008.00399.x, 2009.

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., Holben, B. N., Holzer-Popp, T., Ignatov, A., Kahn, R. A., Kaufman, Y. J., Loeb, N., Martonchik, J. V., Mishchenko, M. I., Nalli, N. R., Remer, L. A., Schroedter-Homscheidt, M., Tanré, D.,

Torres, O., and Wang, M.: Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000, Atmospheric Chemistry and Physics, 5, 1697–1719, doi:10.5194/acp-5-1697-2005, http://www.atmos-chem-phys.net/5/1697/2005/, 2005.

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, http://www.atmos-chem-phys.net/17/277/2017/, 2017.

Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmospheric Measurement Techniques, 7, 2389–2409, doi:10.5194/amt-7-2389-2014, http://www.atmos-meas-tech.net/7/2389/2014/, 2014.

Péré, J., Bessagnet, B., Mallet, M., Waquet, F., Chiapello, I., Minvielle, F., Pont, V., and Menut, L.: Direct radiative effect of the Russian wildfires and its impact on air temperature and atmospheric dynamics during August 2010, Atmospheric Chemistry and Physics, 14, 1999–2013, 2014.

Pérez, C., Nickovic, S., Pejanovic, G., Baldasano, J. M., and Özsoy, E.: Interactive dust-radiation modeling: A step to improve weather forecasts, Journal of Geophysical Research: Atmospheres, 111, 2006.

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.

Prijith, S., Aloysius, M., and Mohan, M.: Relationship between wind speed and sea salt aerosol production: A new approach, Journal of Atmospheric and Solar-Terrestrial Physics, 108, 34–40, 2014.

Ramanathan, V., Crutzen, P., Kiehl, J., and Rosenfeld, D.: Aerosols, climate, and the hydrological cycle, science, 294, 2119–2124, 2001.

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 International Panel of Climate Change, pp. 589–662, Cambridge University Press, 2007.

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.

Ritter, B. and Geleyn, J.-F.: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations, Monthly Weather Review, 120, 303–325, 1992.

Roy, S. and Sharp, J.: Why Atmospheric Stability Matters in Wind Assessment, North American Wind Power Available at:< http://nawindpower. com/online/issues/NAW1301/FEAT_06_Why_Atmospheric. ht ml>[Accessed April 2016], 2013.

Sayer, A., Hsu, N., Bettenhausen, C., and Jeong, M.-J.: Validation and uncertainty estimates for MODIS Collection 6 "Deep Blue" aerosol data, Journal of Geophysical Research: Atmospheres, 118, 7864–7872, 2013.

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.

Schaap, M., Roemer, M., Sauter, F., Boersen, G., Timmermans, R., Builtjes, P., and Vermeulen, A.: LOTOS-EUROS: Documentation, TNO report, 2005.

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, 2001.

Schoeberl, M. R., Douglass, A. R., Hilsenrath, E., Bhartia, P. K., Beer, R., Waters, J. W., Gunson, M. R., Froidevaux, L., Gille, J. C., Barnett, J. J., Levelt, P. F., and DeCola, P.: Overview of the EOS Aura mission, IEEE Transactions on Geoscience and Remote Sensing, 44, 1066–1074, 2006.

Schultz, D. M.: Aerosol-radiation-cloud interactions in a regional coupled model: the effects of convective parameterisation and resolution, Atmospheric Chemistry and Physics, 16, 5573, 2016.

Shao, Y., Wyrwoll, K.-H., Chappell, A., Huang, J., Lin, Z., McTainsh, G. H., Mikami, M., Tanaka, T. Y., Wang, X., and Yoon, S.: Dust cycle: An emerging core theme in Earth system science, Aeolian Research, 2, 181–204, 2011.

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, 2008.

Simpson, D., Fagerli, H., Jonson, J., Tsyro, S., Wind, P., and Tuovinen, J.: The EMEP Unified Eulerian Model. Model Description. EMEP MSC-W Report 1/2003, The Norwegian Meteorological Institute, Oslo, Norway, 2003.

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., Vankevich, R., Lotjonen, M., Prank, M., Petukhov, V., Ermakova, T., Koskinen, J., and Kukkonen, J.: An operational system for the assimilation of the satellite information on wild-land fires for the needs of air quality modelling and forecasting, Atmospheric Chemistry and Physics, 9, 6833–6847, 2009.

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, 1990.

Sun, H., Biedermann, L., and Bond, T. C.: Color of brown carbon: A model for ultraviolet and visible light absorption by organic carbon aerosol, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2007GL029797, http://dx.doi.org/10.1029/2007GL029797, l17813, 2007.

Tanré, D., Kaufman, Y., Holben, B. e. a., Chatenet, B., Karnieli, A., Lavenu, F., Blarel, L., Dubovik, O., Remer, L., and Smirnov, A.: Climatology of dust aerosol size distribution and optical properties derived from remotely sensed data in the solar spectrum, Journal of Geophysical Research: Atmospheres, 106, 18 205–18 217, 2001.

Team, O.: Ozone monitoring instrument (OMI) data user's guide, 2009.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Monthly Weather Review, 117, 1779–1800, 1989.

Tilgner, A., Wolke, R., and Herrmann, H.: Capram Modeling Of Aqueous Aerosol And Cloud Chemistry, Simulation and Assessment of Chemical Processes in a Multiphase Environment, pp. 107–122, 2008.

Van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J.: Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application, Ph.D. thesis, University of British Columbia, 2015.

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, 1985.

5   Winker, D. M., Pelon, J. R., and McCormick, M. P.: The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds, vol. 4893, pp. 1–11, 2003.

Wolke, R., Schröder, W., Schrödner, R., and Renner, E.: Influence of grid resolution and meteorological forcing on simulated European air quality: a sensitivity study with the modeling system COSMO–MUSCAT, Atmospheric environment, 53, 110–130, 2012.

Wong, D., Pleim, J., Mathur, R., Binkowski, F., Otte, T., Gilliam, R., Pouliot, G., Xiu, A., Young, J., and Kang, D.: WRF-CMAQ two-way

10   coupled system with aerosol feedback: software development and preliminary results, Geoscientific Model Development, 5, 299, 2012.

Yang, Q., Gustafson Jr, W., Fast, J. D., Wang, H., Easter, R. C., Morrison, H., Lee, Y.-N., Chapman, E. G., Spak, S., and Mena-Carrasco, M.: Assessing regional scale predictions of aerosols, marine stratocumulus, and their interactions during VOCALS-REx using WRF-Chem, Atmospheric Chemistry and Physics, 11, 11 951–11 975, 2011.

Yang, Y., Russell, L. M., Lou, S., Liao, H., Guo, J., Liu, Y., Singh, B., and Ghan, S. J.: Dust-wind interactions can intensify aerosol pollution

15   over eastern China, Nature Communications, 8, 2017.

Zaveri, R. A. and Peters, L. K.: A new lumped structure photochemical mechanism for large-scale applications, Journal of Geophysical Research: Atmospheres, 104, 30 387–30 415, 1999.

Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for simulating aerosol interactions and chemistry (MOSAIC), Journal of Geophysical Research: Atmospheres, 113, 2008.

20   Zhang, B., Wang, Y., and Hao, J.: Simulating aerosol–radiation–cloud feedbacks on meteorology and air quality over eastern China under severe haze conditions in winter, Atmospheric Chemistry and Physics, 15, 2387–2404, 2015.
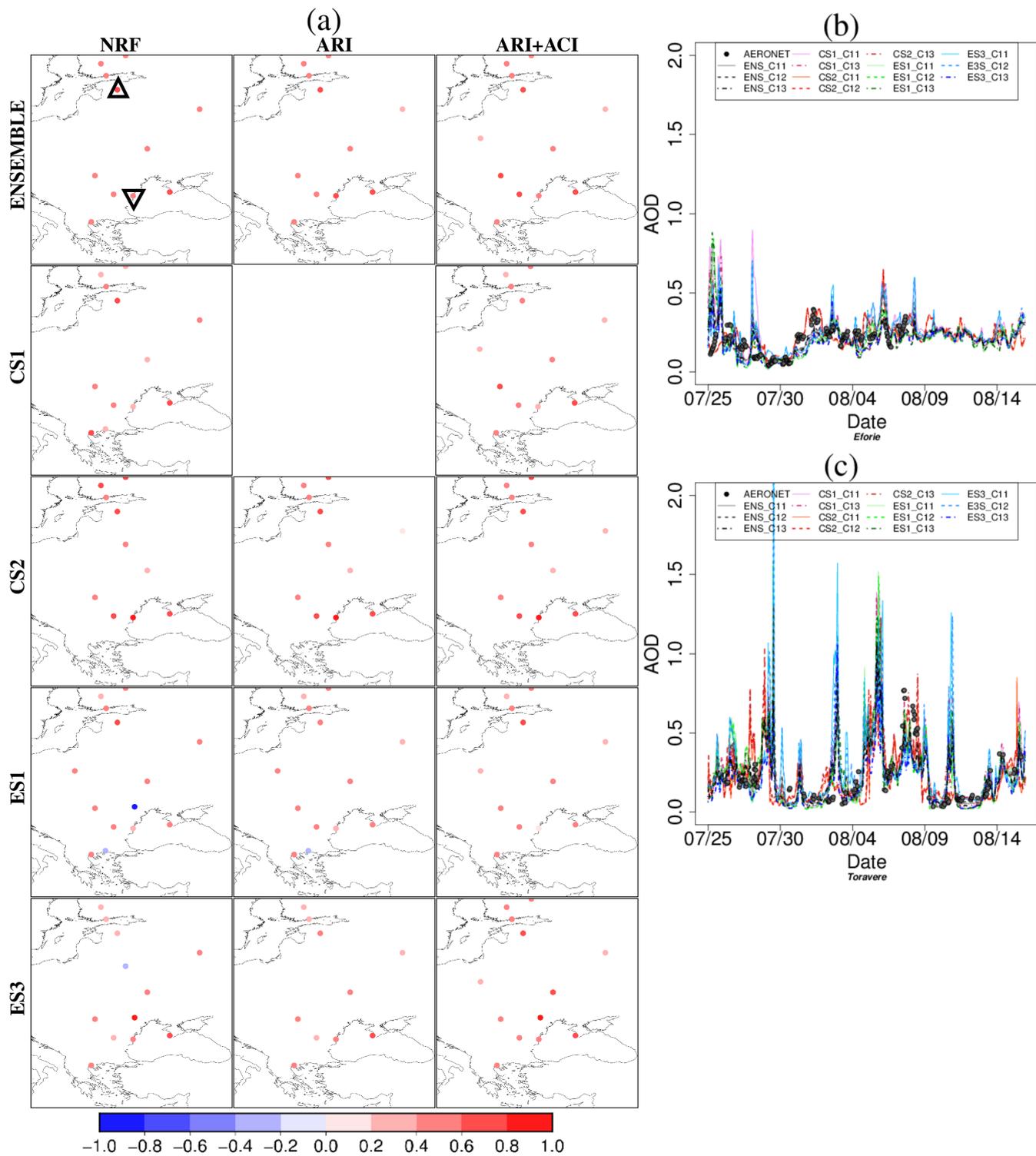
Figure 3. Model-AERONET comparison of AOD at $670nm$ for the Russian wildfires case: (a) temporal correlation values at each AERONET station; and temporal series at the (b) Toravere and (c) Eforie stations.
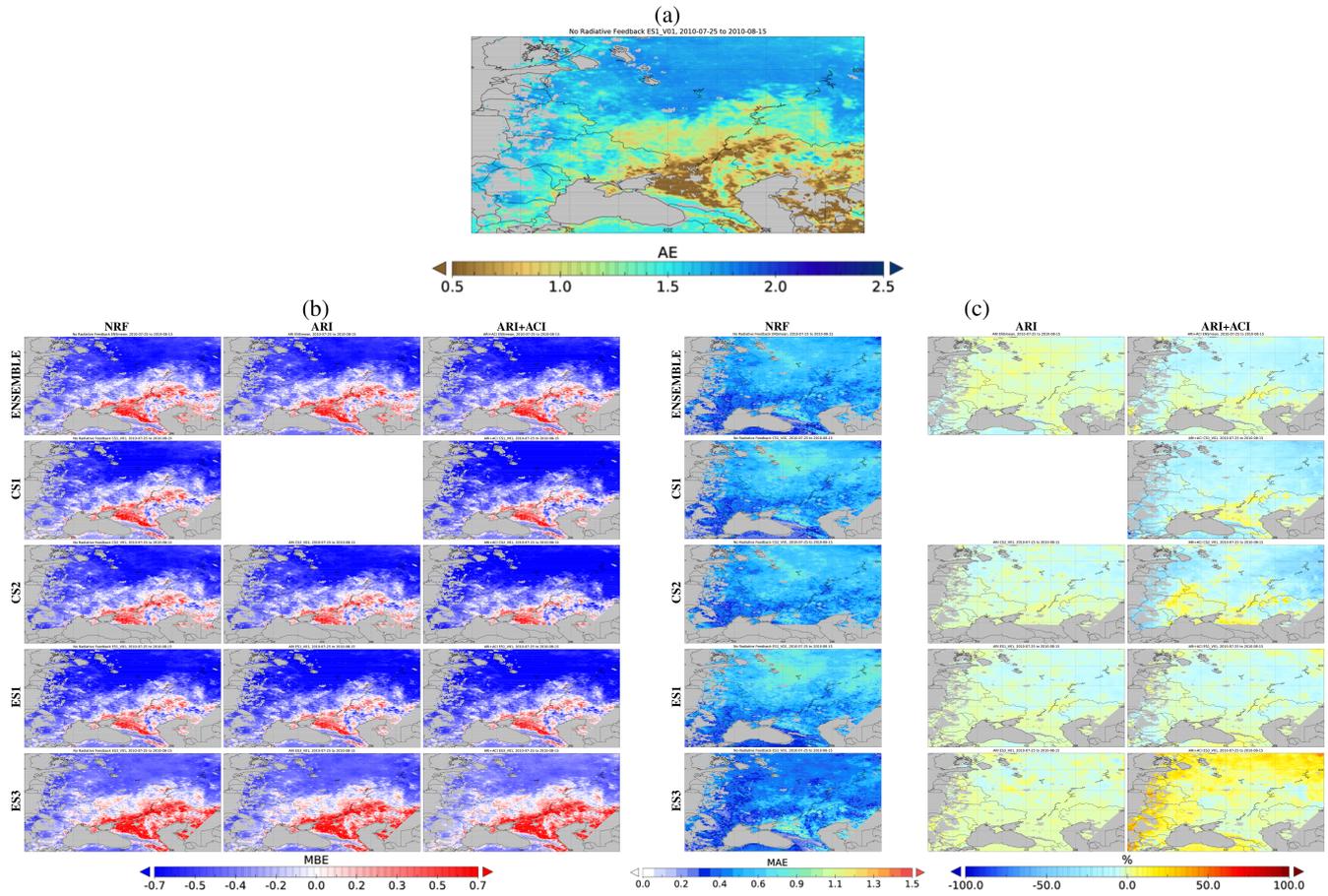
Figure 4. Model-MODIS comparison of AE between $412/470nm$ for the Russian wildfires case: (a) Satellite values; (b) MBE for NRF (first column), ARI (second) and ARI+ACI (third); and (c) MAE for the NRF simulations (first column) and their improvements due to ARI (second) and ARI+ACI (third).
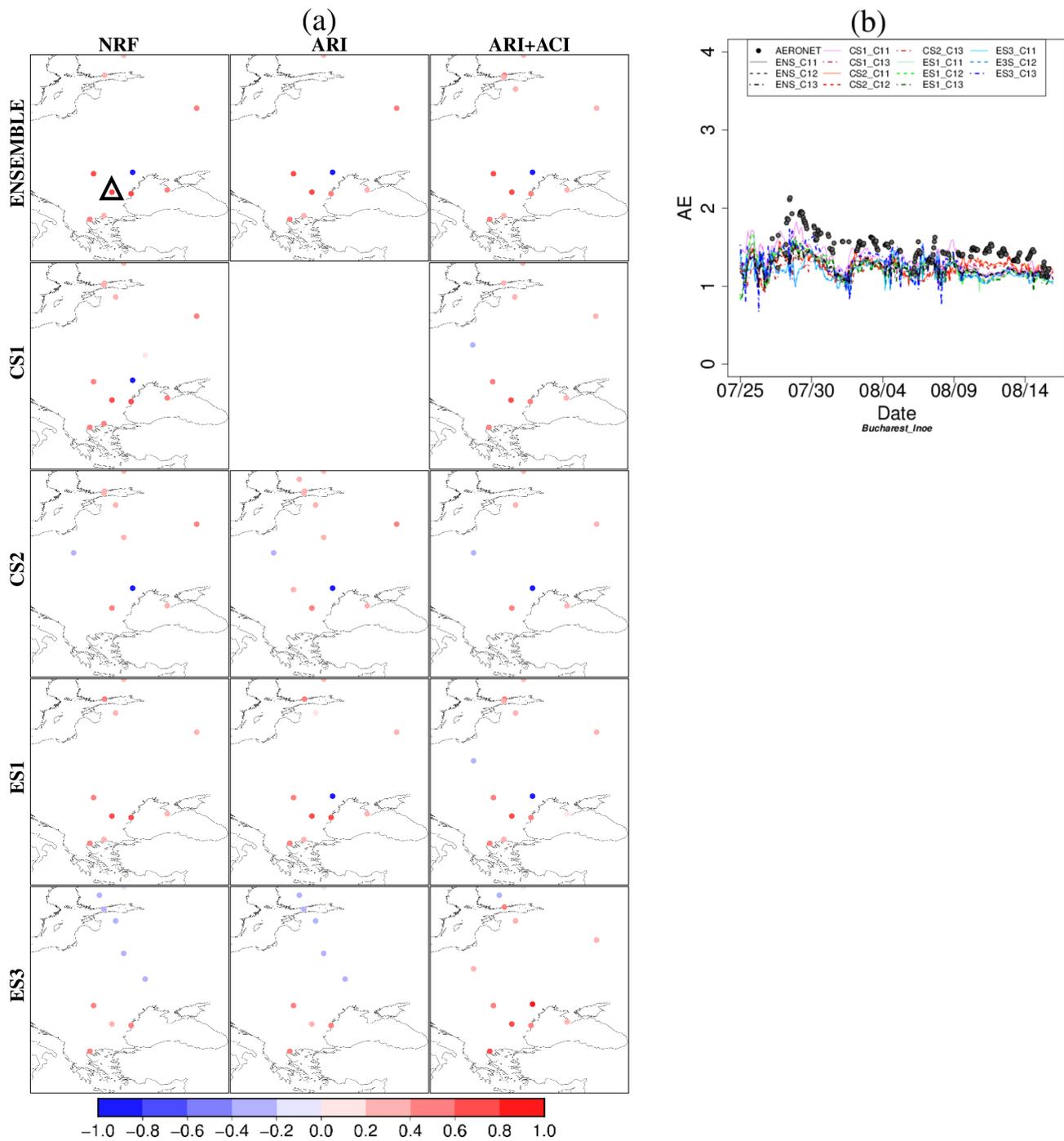
Figure 5. Model-AERONET comparison of AE $440/870nm$ for the Russian wildfires case: (a) temporal correlation values at each AERONET station; and temporal series at (b) the Bucharest station.
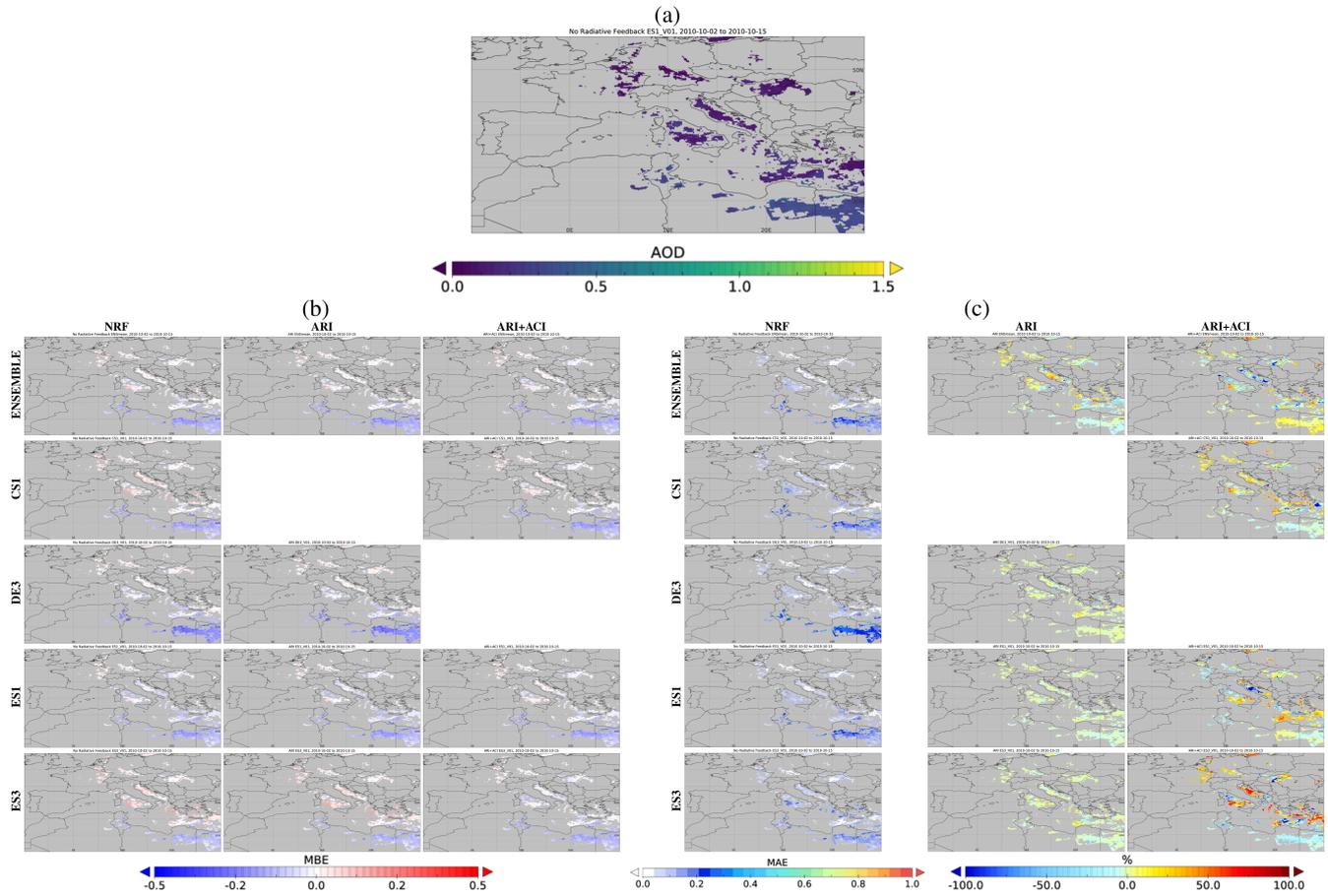
Figure 6. Model-MODIS comparison of AOD at $550nm$ for the Saharan desert dust outbreak case: (a) Satellite values; (b) MBE for NRF (first column), ARI (second) and ARI+ACI (third); and (c) MAE for the NRF simulations (first column) and their improvements due to ARI (second) and ARI+ACI (third).
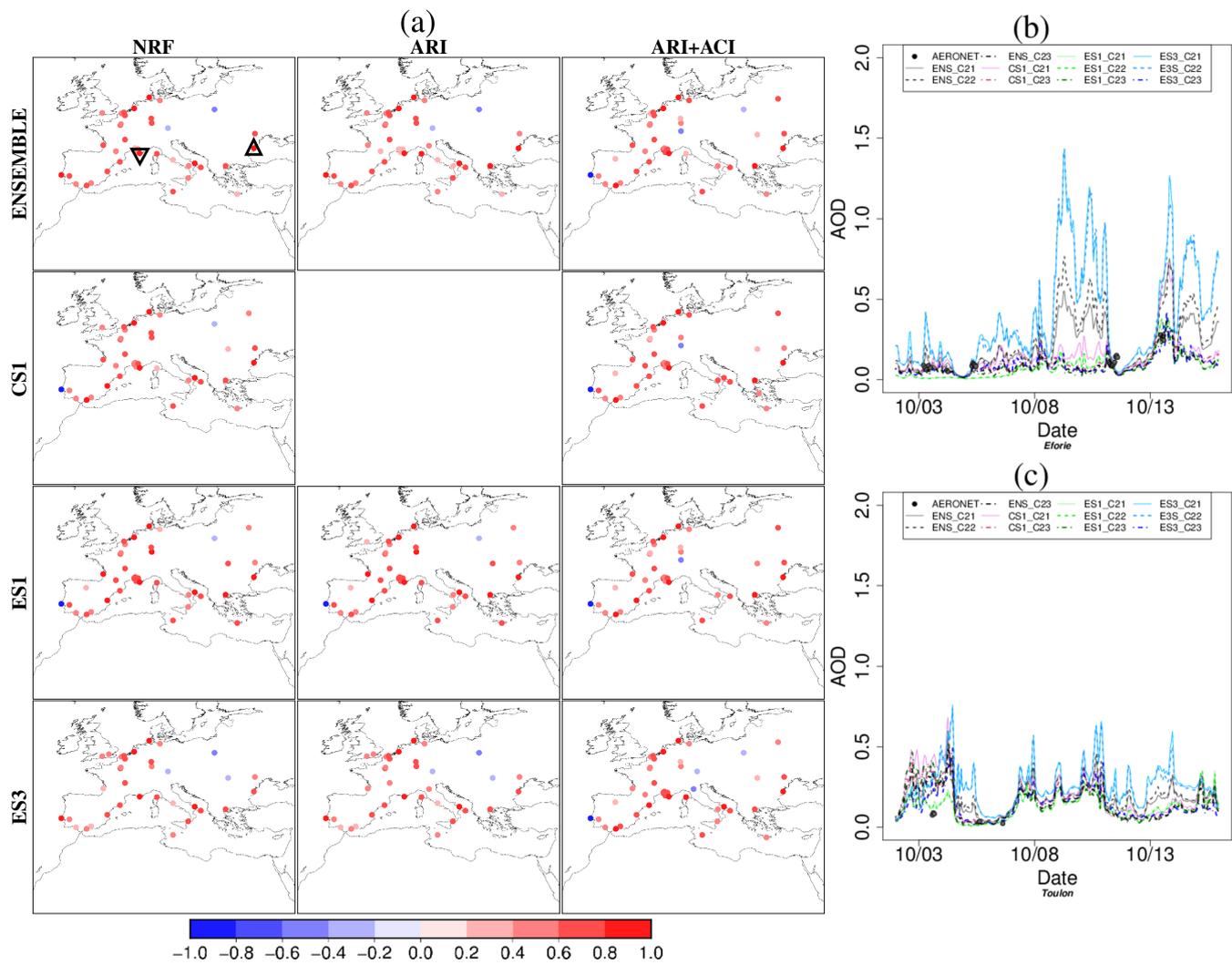
Figure 7. Model-AERONET comparison of AOD at $670nm$ for the Saharan desert dust outbreak case: (a) temporal correlation values at each AERONET station; and temporal series at the (b) Toulon and (c) Eforie stations.
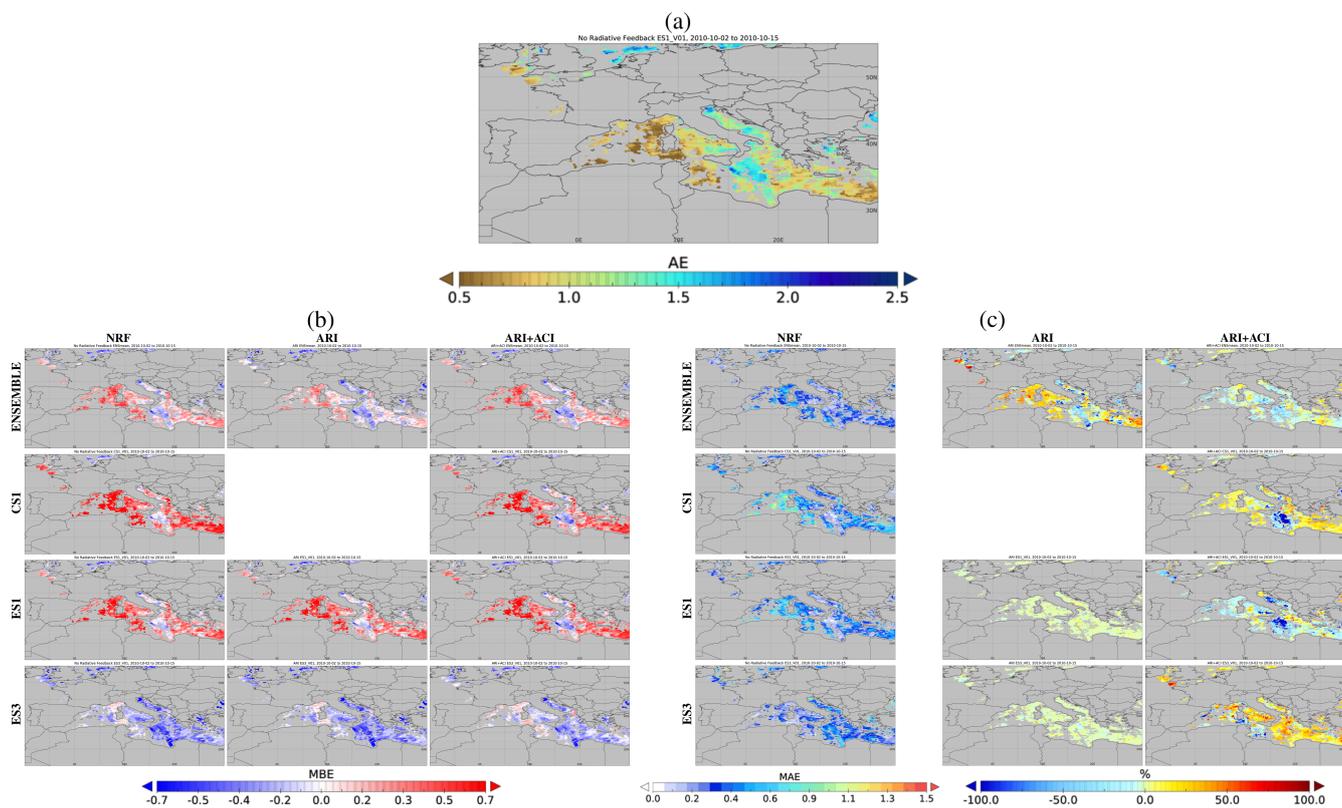
Figure 8. Model-MODIS comparison of AE between $550/860nm$ for the Saharan desert dust outbreak case: (a) Satellite values; (b) MBE for NRF (first column), ARI (second) and ARI+ACI (third); and (c) MAE for the NRF simulations (first column) and their improvements due to ARI (second) and ARI+ACI (third).
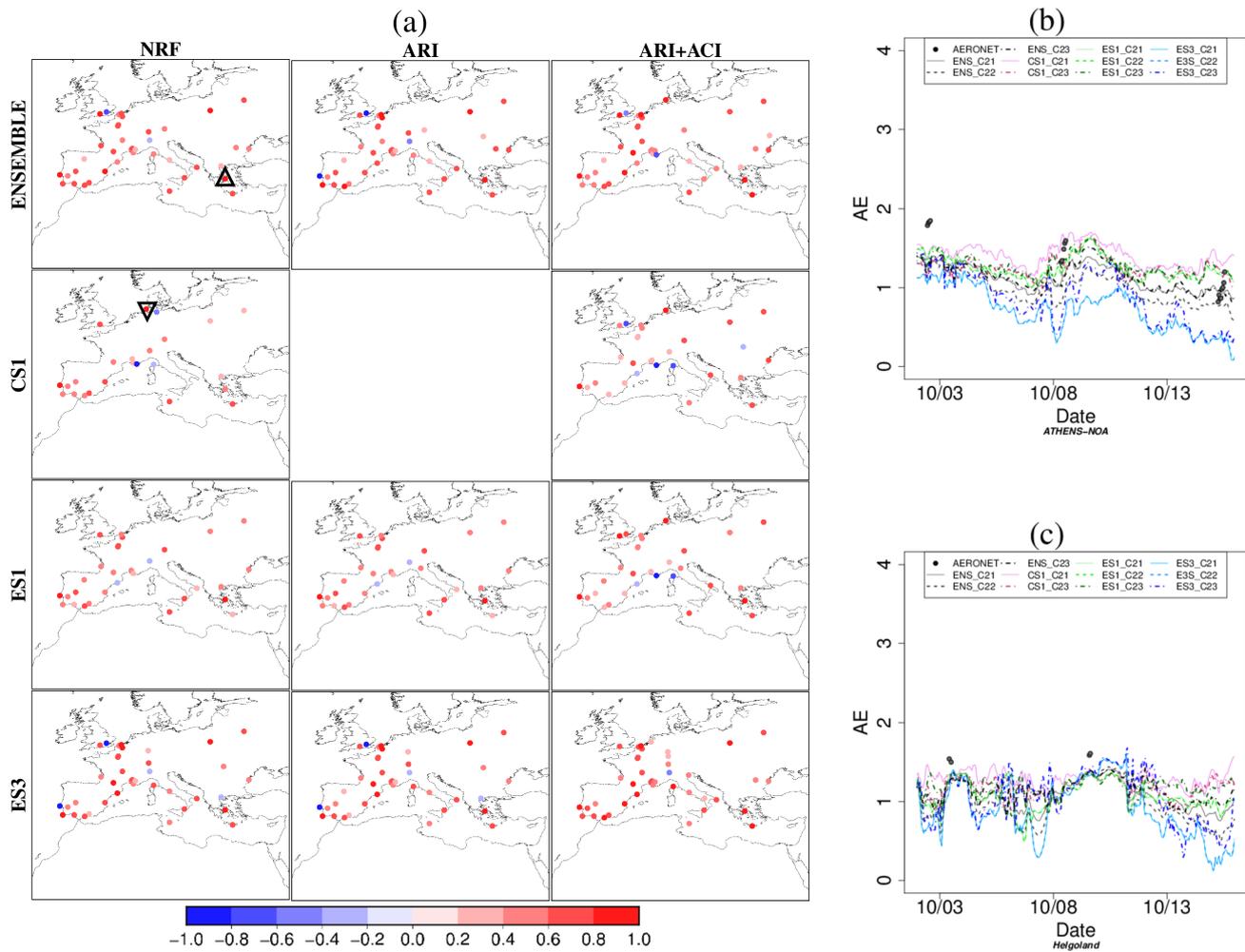
Figure 9. Model-AERONET comparison of AE $440/870nm$ for the Saharan desert dust outbreak case: (a) temporal correlation values at each AERONET station; and temporal series at the (b) Athens and (c) Helgoland stations.