Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2

Dominik Brunner, Nicholas Savage, Oriol Jorba, Brian Eder, Lea Giordano, Alba Badia, Alessandra Balzarini, Rocío Baró, Roberto Bianconi, Charles Chemel, Gabriele Curci, Renate Forkel, Pedro Jiménez-Guerrero, Marcus Hirtl, Alma Hodzic, Luka Honzak, Ulas Im, Christoph Knote, Paul Makar, Astrid Manders-Groot, Erik van Meijgaard, Lucy Neal, Juan L. Pérez, Guido Pirovano, Roberto San Jose, Wolfram Schröder, Ranjeet S. Sokhi, Dimiter Syrakov, Alfreida Torian, Paolo Tuccella, Johannes Werhahn, Ralf Wolke, Khairunnisa Yahya, Rahela Zabkar, Yang Zhang, Christian Hogrefe, Stefano Galmarini

Angaben zur Veröffentlichung / Publication details:

Brunner, Dominik, Nicholas Savage, Oriol Jorba, Brian Eder, Lea Giordano, Alba Badia, Alessandra Balzarini, et al. 2015. "Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2." *Atmospheric Environment* 115: 470–98. https://doi.org/10.1016/j.atmosenv.2014.12.032.

Atmospheric Environment 115 (2015) 470–498

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2

Dominik Brunner [a, *], Nicholas Savage [b], Oriol Jorba [c], Brian Eder [d], Lea Giordano [a], Alba Badia [c], Alessandra Balzarini [e], Rocío Baró [f], Roberto Bianconi [g], Charles Chemel [h], Gabriele Curci [i], Renate Forkel [j], Pedro Jiménez-Guerrero [f], Marcus Hirtl [k], Alma Hodzic [l], Luka Honzak [m], Ulas Im [n, 1], Christoph Knote [l], Paul Makar [o], Astrid Manders-Groot [p], Erik van Meijgaard [q], Lucy Neal [b], Juan L. Pérez [r], Guido Pirovano [e], Roberto San Jose [r], Wolfram Schröder [s], Ranjeet S. Sokhi [h], Dimiter Syrakov [t], Alfreida Torian [d], Paolo Tuccella [i], Johannes Werhahn [j], Ralf Wolke [s], Khairunnisa Yahya [u], Rahela Zabkar [m, v], Yang Zhang [u], Christian Hogrefe [d], Stefano Galmarini [n]

[a] Laboratory for Air Pollution and Environmental Technology, Empa, Dubendorf, Switzerland
[b] Met Office, FitzRoy Road, Exeter EX1 3PB, United Kingdom
[c] Earth Sciences Department, Barcelona Supercomputing Center (BSC-CNS), Barcelona, Spain
[d] Atmospheric Modelling and Analysis Division, Environmental Protection Agency, Research Triangle Park, USA
[e] Ricerca sul Sistema Energetico (RSE S.p.A.), via Rubattino 54, Milano, Italy
[f] University of Murcia, Department of Physics, Physics of the Earth, Campus de Espinardo, Ed. CIOyN, 30100 Murcia, Spain
[g] Enviroware srl, Concorezzo, MB, Italy
[h] Centre for Atmospheric & Instrumentation Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
[i] Department of Physical and Chemical Sciences, Center of Excellence for the Forecast of Severe Weather (CETEMPS), University of L'Aquila, L'Aquila, Italy
[j] Karlsruher Institut für Technologie (KIT), Institut für Meteorologie und Klimaforschung, Atmosphärische Umweltforschung (IMK-IFU), Kreuzeckbahnstr. 19, 82467 Garmisch-Partenkirchen, Germany
[k] Section Environmental Meteorology, Division Customer Service, ZAMG - Zentralanstalt für Meteorologie und Geodynamik, 1190 Wien, Austria
[l] National Center for Atmospheric Research, Boulder, CO, USA
[m] Center of Excellence SPACE-SI, Ljubljana, Slovenia
[n] Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy
[o] Air Quality Research Section, Atmospheric Science and Technology Directorate, Environment Canada, 4905 Dufferin Street, Toronto, Ontario, Canada
[p] Netherlands Organization for Applied Scientific Research (TNO), Utrecht, The Netherlands
[q] Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[r] Environmental Software and Modelling Group, Computer Science School - Technical University of Madrid, Campus de Montegancedo, Boadilla del Monte, 28660 Madrid, Spain
[s] Leibniz Institute for Tropospheric Research, Permoserstr. 15, D-04318 Leipzig, Germany
[t] National Institute of Meteorology and Hydrology, Sofia 1784, Bulgaria
[u] Department of Marine, Earth and Atmospheric Sciences, 2800 Faucette Drive, #1125 Jordan Hall, Campus Box 8208, North Carolina State University, USA
[v] University of Ljubljana, Faculty of Mathematics and Physics, Ljubljana, Slovenia

HIGHLIGHTS

- We evaluate the meteorological performance of coupled chemistry-meteorology models.
- 13 modeling groups from Europe and 4 groups from North America participated.
- Temperature, precipitation and radiation are mostly well simulated.
- Significant biases exist in surface wind speeds and nighttime boundary layer heights.
- Differences between model systems are usually larger than aerosol feedback effects.

* Corresponding author.
E-mail address: dominik.brunner@empa.ch (D. Brunner).
[1] Now at Aarhus University, Department of Environmental Science, Frederiksborgvej 399, 4000, Roskilde, Denmark.

http://dx.doi.org/10.1016/j.atmosenv.2014.12.032

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form
11 December 2014
Accepted 13 December 2014
Available online 15 December 2014

Keywords:
Online-coupled meteorology-chemistry modeling
Model evaluation
Meteorology
AQMEII phase 2

ABSTRACT

Air pollution simulations critically depend on the quality of the underlying meteorology. In phase 2 of the Air Quality Model Evaluation International Initiative (AQMEII-2), thirteen modeling groups from Europe and four groups from North America operating eight different regional coupled chemistry and meteorology models participated in a coordinated model evaluation exercise. Each group simulated the year 2010 for a domain covering either Europe or North America or both. Here were present an operational analysis of model performance with respect to key meteorological variables relevant for atmospheric chemistry processes and air quality. These parameters include temperature and wind speed at the surface and in the vertical profile, incoming solar radiation at the ground, precipitation, and planetary boundary layer heights. A similar analysis was performed during AQMEII phase 1 (Vautard et al., 2012) for offline air quality models not directly coupled to the meteorological model core as the model systems investigated here. Similar to phase 1, we found significant overpredictions of 10-m wind speeds by most models, more pronounced during night than during daytime. The seasonal evolution of temperature was well captured with monthly mean biases below 2 K over all domains. Solar incoming radiation, precipitation and PBL heights, on the other hand, showed significant spread between models and observations suggesting that major challenges still remain in the simulation of meteorological parameters relevant for air quality and for chemistry–climate interactions at the regional scale.

1. Introduction

Air quality models have advanced significantly over the past 20 years driven by the rapid evolution of computer power and by improvements in our understanding of atmospheric processes. Air quality models are increasingly being used not only for research but also in an operational context by national weather centers and environment institutes for air quality prediction, for designing emission control policies, and for environmental impact assessment. A prominent example is the regional model ensemble established in the EU project MACC which provides operational daily air quality forecasts for Europe (Hollingsworth et al., 2008; Huijnen et al., 2010).

The historic separation between the air quality and weather prediction communities led to the separate development of regional atmospheric chemistry and meteorology models. As a consequence, air quality models were mostly driven offline by the output of a separate meteorology model. In the last approximately ten years, supported both by the increased interest of weather centers in air quality issues and by the rapid increase in computer power, a new generation of models has been developed in which the chemical evolution is online coupled to the meteorological simulation. Comprehensive reviews of these coupled model systems developed in North America and Europe have been presented by Zhang (2008) and Baklanov et al. (2014). Online coupled models can account for interactions between chemistry and meteorology, notably for direct effects of aerosols on radiation and for indirect effects of aerosols on clouds (e.g., Bangert et al., 2011; Giorgi et al., 2003; Helmert et al., 2007).

In contrast to offline models, a comprehensive evaluation and intercomparison of this new generation of online coupled models has been missing so far but is urgently needed to build scientific credibility in their use to address a wide range of air quality and climate related questions (Alapaty et al., 2012).

Since 2008, the Air Quality Model Evaluation International Initiative (AQMEII) (Rao et al., 2010) coordinated by the European Joint Research Center (JRC) and U.S. Environmental Protection Agency (EPA), has promoted the evaluation of regional air quality models across Europe and North America. AQMEII has now reached its second phase which is dedicated to the evaluation of online coupled models, as opposed to Phase 1 where, with one exception, only offline models were considered. AQMEII-2 brought together thirteen modeling groups from Europe and four groups from North America. Each group simulated the year 2010 for a domain covering either Europe or North America or both domains.

The purpose of this study is to evaluate the models participating in the AQMEII-2 exercise with respect to the simulation of meteorology. It complements the collective analyses of Im et al. (2015a,b), and Giordano et al. (2015) which are dedicated to the evaluation of ozone, particular matter, and the influence of chemical boundary conditions, respectively. Meteorological parameters are driving chemical processes in numerous ways (Seaman, 2000) and the quality of the meteorological simulation critically affects the predictability of air pollution episodes (Zhang et al., 2007). The analysis of the performance of the AQMEII-2 models with respect to key meteorological variables thus contributes to the understanding of differences in the chemistry modeling reported in the companion studies in this special issue.

A similar analysis was conducted by Vautard et al. (2012) for the meteorological models providing input for the offline chemistry transport models in AQMEII phase 1. Here, we extend their analysis by providing the first comparative evaluation of a large number of online coupled model systems. Eight different model systems built around six different meteorological model cores are considered. Although most of the models accounted for the aerosol direct and some models also for the aerosol indirect effects, the analysis of the feedbacks of these effects onto meteorology is beyond the scope of this study but is addressed in several companion studies (Forkel et al., 2015; Kong et al., 2015; Makar et al., 2015a, b; San Jose et al., 2015). Furthermore, a detailed comparison of bulk aerosol profiles simulated by the models at Aerosol Robotic Network (AERONET) sites and analysis of the influence of different assumptions regarding mixing state, refractive indices, hygroscopic growth and other factors on aerosol optical properties is presented in Curci et al. (2014).

The paper is organized as follows: Section 2 presents an overview of the models operated by the thirteen European and the four North American groups. Section 3 describes the meteorological observations used for the evaluation. Section 4 provides a brief overview of the general weather situation in 2010 to place the simulations conducted in AQMEII-2 in a broader climatological context. Section 5 is the central part of the study presenting the quantitative evaluation for a number of key meteorological variables. Section 6 closes with the summary and conclusions.

2. Online-coupled meteorology and chemistry models

As for AQMEII phase 1, simulations had to be performed for a continental domain covering Europe or North America. In total, 16 groups conducted simulations for Europe and 5 groups for North America. A number of groups shared the same model system but operated the model in different configurations to test the sensitivity to different settings. Overall, eight different model systems were used out of which five were fully integrated online coupled chemistry and meteorology models (WRF-Chem, COSMO-ART, MetUM UKCA-RAQ, NMMB-BSC, GEM-MACH), and three were online access models (COSMO-MUSCAT, RACMO LOTOS-EUROS, WRF-CMAQ) following the model classification scheme of Baklanov et al. (2014). In terms of meteorological core, only six different models were applied: The U.S. Weather Research and Forecasting model *WRF* (Skamarock and Klemp, 2008), the *COSMO* model of the European Consortium for Small-Scale Modeling (Baldauf et al., 2011), the Met Office Unified Model *MetUM* (Brown et al., 2012), the Canadian Global Environmental Multiscale Model *GEM* (Yeh et al., 2002), the Regional Atmospheric Climate Model *RACMO* (van Meijgaard et al., 2012) and the U.S, Nonhydrostatic Multiscale Model on the B-grid *NMMB* (Janjic and Gall, 2012).

An overview of the different models and their configurations with respect to the most relevant meteorological parameterizations is presented in Table 1 for WRF-Chem and in Table 2 or all other models. For an overview of the different models with respect to chemistry we refer to (Im et al., 2015b).

In order to allow the models to respond to aerosol direct and indirect effects while keeping the meteorological simulation close to reality, the simulations were performed in two-day segments. Chemical fields at the end of each segment served as initial conditions for the subsequent segment. For the meteorology, however, each segment was preceded by a spinup which varied between 12 h and 1 day depending on the model. It was recommended not to use any nudging as this would potentially mask any feedback effects. The two WRF-CMAQ models applied over Europe and North America, respectively, deviated from this general recommendation: UK5 performed the simulations in one-day rather than in two-day segments. US6 performed continuous simulations with a weak nudging of upper layer temperature, winds, and water vapor as well as soil moisture and temperature as described in Hogrefe et al. (2015).

An important point to stress is that the results presented here only reflect the model's performance in their configuration used for the AQMEII-2 exercise. These settings may differ significantly from those used by operational weather centers in terms of parameterizations, land-surface treatment (e.g. soil moisture, land use data sets), boundary conditions, horizontal and vertical grid spacing, etc. Nevertheless, the results should reflect some of the fundamental properties of a model system and the parameterizations used, and biases identified here certainly deserve further attention.

3. Meteorological observations

The models participating in the AQMEII-2 exercise were compared against surface observations and, over Europe, against vertical profiles from commercial airliners at the airport of Frankfurt, Germany. The comparative analysis includes the classical meteorological variables 2-m temperature and 10-m wind speed which are available at a large number of sites as well as precipitation and shortwave radiation data and diagnosed PBL heights from a smaller number of stations. An overview of the sites included in the analysis over the two continents is presented in Fig. 1. SYNOP refers to measurements of temperature and wind, IGRA to radiosonde locations at which PBL heights were diagnosed and BSRN/SURFRAD to the sites with radiation measurements. For precipitation, observation-based gridded data sets were used for both continents as described below.

3.1. Temperature and wind speed

Temperature and wind speed data from SYNOP stations in EMSEMBLE with >90% data availability were used for the European surface analyses in domains EU1, EU2 and EU3. The criterion of a data availability of 90% restricted the analysis to those stations with an hourly reporting frequency. Vertical profile information was obtained by using MOZAIC profiles at Frankfurt airport (Marenco et al., 1998). Most observations are in the morning hours with the

Table 1
Overview of meteorology configurations of WRF-Chem models.

Model ID	AT1	DE4	IT1	IT2	ES1	ES3	SI1/SI2	US7	US8
Domain	EU	EU	EU	EU	EU/NA	EU	EU	NA	NA
Group	ZAMG	IMK-IFU	RSE	UNIVAQ	MAR-UMU	UPM	Univ. Ljubljana	NCAR	NCSU
Version	3.4.1	3.4.1	3.4.1	3.4 prerel.	3.4.1	3.4.1	3.4.1	3.4.1	3.4.1
Horiz. resolution	23 km	23 km	23 km	23 km	23 km/36 km	23 km	23 km	36 km	36 km
Nx × Ny	270 x 225	270 x 225	270 x 225	270 x 225	270 x 225/ 161 x 105	270 x 225	270 x 225	161 x 105	148 x 112
Levels	33 (eta)	33 (eta)	33 (eta)	33 (eta)	33 (eta)	33 (eta)	33 (eta)	33 (eta)	34 (eta)
First layer	24 m	24 m	24 m	24 m	24 m	24 m	24 m	24 m	24 m
Meteo IC/BC	ECMWF (oper.)	ECMWF (oper.)	ECMWF (oper.)	ECMWF (oper.)	ECMWF (oper.)	ECMWF (oper.)	ECMWF (oper.)	NCEP-GFS	NCEP-GFS
Microphys.	Morrison	Morrison	Morrison	Morrison	Lin	Morrison	Morrison	Morrison	Morrison
LW radiation	RRTM	RRTM	RRTMG	RRTMG	RRTM	RRTMG	RRTMG	RRTM	RRTMG
SW radiation	RRTMG	Goddard	RRTMG	RRTMG	Goddard	RRTMG	RRTMG	RRTM	RRTMG
Land surface	Noah	Noah	Noah	Noah	Noah	Noah	Noah	Noah	Noah
PBL/turbulence	YSU	YSU	YSU	YSU	YSU	YSU	YSU	MYNN	YSU
Convection	Grell-3D	Grell-3D	Grell-3D	Grell-3D	Grell-3D	Grell-3D	Grell-3D	Grell-3D	Grell-3D
Aerosol feedbacks	Direct &Indirect	Direct &Indirect	no	Direct &Indirect	Direct &Indirect	Direct &Indirect	Direct/no	Direct &Indirect	Direct &Indirect

Microphysics parameterizations: Morrison (Morrison et al., 2009); Lin (Lin et al., 1983).
Land surface parameterizations: Noah (Niu et al., 2011).
Radiative transfer: RRTM (Mlawer et al., 1997); RRTMG (Iacono et al., 2008); Goddard (Chou and Suarez, 1994).
PBL/turbulence schemes: YSU (Hong et al., 2006); MYNN (Nakanishi and Niino, 2006).
Convection schemes: Grell-3D (Grell and Dévényi, 2002).

Table 2
Overview of meteorology configurations of other models.

Model ID	CH1	DE3	NL2[a]	UK4	UK5	US6	CA2/f	ES2a/b
Domain	EU	EU	EU	EU	EU	NA	NA	NA
Group	Empa	IFT	TNO/KNMI	UKMO	HERTS	EPA	Env. Canada	BSC
Met Model	COSMO	COSMO	RACMO	METUM	WRF	WRF	GEM	NMMB
Version	4.23	4.27	2-LE	8.3	3.3	3.4.1		1.2
Chem Model	ART	MUSCAT	LOTOS-EUROS	UKCA	CMAQ 5.0.1	CMAQ 5.0.1	MACH	BSC-CTM
Horiz. resolution	0.22°	0.25°	0.22° 0.5° × 0.25°	0.22°	18 km	12 km	15 km	0.20°
N x x Ny	270 x 225	166 x 164	306 x 200 140 x 160	244 x 238	300 x 300	459 x 299	348 x 465	311 x 251
Levels	40 (z-hyb)	40 (z-hyb)	40 (eta) 5 (z)	38 (z-hyb)	35 (eta)	35 (eta)	58	24/48 (σ-hyb)
First layer	20 m	20 m	10 m 25 m	20 m	19 m	19 m	21 m	45/25 m
Meteo input	ECMWF (oper.)	GME (reanal.)	ECMWF (reanal.)	Met Office	ECMWF (oper.)	NCEP-GFS	CMC-reg_OA	NCEP-FNL
Microphys.	Kessler-type bulk	Kessler-type bulk	Tiedtke, Tompkins Neggers	Wilson & Ballard	Morrison	Morrison	Milbrandt-Yau double moment	Ferrier
LW radiation	∂-2-stream	∂-2-stream	RRTMG	Edwards −Slingo	RRTMG	RRTMG	CDK	RRTMG
SW radiation	∂-2-stream	∂-2-stream	RRTMG/McRad	Edwards −Slingo	RRTMG	RRTMG	CDK	RRTMG
Land surface	TERRA-ML	TERRA-ML	Hurk/Balsamo	MOSES-2	Noah	Pleim-Xiu	ISBA	LISS
PBL/turbulence	Prognostic TKE	Prognostic TKE	Lenderink/Siebesma	Lock	ACM2	ACM2	Moist TKE	MYJ 2.5
Convection	Tiedtke	Tiedtke	Nordeng/DualM	Gregory	Kain−Fritsch	Kain −Fritsch	Kain−Fritsch	BMJ
Aerosol feedbacks	Direct	Direct	Direct & Indirect	Direct &Indirect	Direct	Direct	No/Direct& Indirect	No

Microphysics parameterizations: WSM6 (Hong and Lim, 2006); Kessler bulk type (Doms et al., 2011); (Wilson and Ballard, 1999); Ferrier (http://www.emc.ncep.noaa.gov/mmb/mmbpll/eta12tpb/); Morrison (Morrison et al., 2009); Tiedtke/Tompkins/ECMWF/Neggers (Neggers, 2009; Tiedtke, 1993; Tompkins et al., 2007) [http://www.ecmwf.int/research/ifsdocs/CY33r1/PHYSICS/IFSPart4.pdf]; Milbrandt−Yau double moment scheme (Milbrandt and Yau, 2005).
Land surface parameterizations: TERRA-ML (Grasselt et al., 2008); LISS (Vukovic et al., 2010); MOSES-2 (Essery and Clark, 2003); Hurk/Balsamo (Balsamo et al., 2009; Van den Hurk et al., 2000); ISBA (Noilhan and Mahfouf, 1996).
Radiative transfer: ∂-2-stream (Ritter and Geleyn, 1992); Dudhia (Dudhia, 1989); RRTM (Mlawer et al., 1997); RRTMG (Iacono et al., 2008); Edwards and Slingo (Edwards and Slingo, 1996); McRad (Clough et al., 2005; Morcrette et al., 2008); CDK (Li and Barker, 2005).
PBL/turbulence schemes: YSU (Hong et al., 2006); Prognostic TKE (Doms et al., 2011); MYJ (Janjić, 1994); ACM2 (Pleim, 2007); Lock (Lock et al., 2000); Lenderink/Siebesma (Lenderink and Holtslag, 2004; Siebesma et al., 2007).
Convection schemes: Kain−Fritsch (Kain, 2004); Tiedtke (Tiedtke, 1989); BMJ/Betts−Miller−Janjic(Janjić, 1994); Gregory (Gregory and Rowntree, 1990); Nordeng update of Tiedtke scheme (Nordeng, 1994); DualM (Neggers et al., 2009).
[a] NL2 is an online access model with different resolutions for the meteorology (RACMO) and chemistry transport (LOTOS-EUROS). Resolutions are therefore provided separately for both (meteo/chemistry).

majority of profiles being between 06 and 13 UTC with maxima in the frequency of profiles at 07 and 12 UTC. Unfortunately, data coverage of MOZAIC was much lower in 2010 as compared to other years with observations available only during fall and winter.

3.2. PBL heights

For planetary boundary layer (PBL) heights a data set recently compiled by Seidel et al. (2012) based on the radiosonde networks over North America and Europe is used. In this data set, PBL heights were determined from vertical profiles of wind, temperature and humidity using the bulk Richardson method of Vogelezang and Holtslag (1996) suitable for both stable and convective boundary layers. A critical bulk Richardson number $Ri_{b,crit}$ of 0.25 was applied for both stable and unstable conditions. A value of 0.25 is frequently used although a wide range of values between 0.1 and 1 has been proposed in the literature in a quest for a universal value (Richardson et al., 2013).

PBL heights in the models were mostly also derived using a bulk Richardson approach, but there are subtle differences between the methods. COSMO-ART, for example, uses the diagnosed 2-m temperature as reference and applies the no-slip condition (i.e., reference wind = 0) identical to the way the radiosonde profiles of Ri_b are computed (Seidel et al., 2012). However, COSMO-ART assigns the PBL height to the first model level at which Ri_b exceeds a value of 0.22 in unstable and 0.33 in stable conditions without vertical interpolation (Szintai et al., 2009). The Yonsei University PBL scheme (YSU) (Hong et al., 2006) used in all WRF-Chem models except US7 estimates PBL height based on a value of $Ri_{b,crit}$ of 0.0 in the case of unstable conditions. In stable conditions, a value of 0.25 is used over land and the minimum of 0.3 and 0.16 $(10^{-7}Ro)^{-0.18}$ over ocean, where Ro is the surface Rossby number (Hong, 2010). PBL height is a key variable in the NL2 model since the PBL height diagnosed by the meteorology component (RACMO) is subsequently used to define the vertical grid in the chemistry-transport component (LOTOS-EUROS). RACMO defines the PBL height following the method of Troen and Mahrt (1986) which also uses a $Ri_{b,crit}$ of 0.25. The NMMB derives the top of the PBL from the profile of prognostic turbulent kinetic energy (TKE). The PBL top is assumed to be located where a floor value of TKE < 0.01 m^2 s^{-2} is reached. For the Mellor−Yamada−Nakanishi−Niino (MYNN) scheme (Nakanishi and Niino, 2004, 2006) the PBL height is diagnosed from the profile of virtual potential temperature for neutral/unstable conditions and from the profile TKE under stable conditions. In the Met Office Unified Model (UK4) the boundary layer depth is calculated as the maximum of two parameters: z_{loc} and z_h. The value z_{loc} is the height at which $Ri_b > 0.25$ and z_h is the top of the surface-based mixed layer, found by adiabatic parcel ascent but

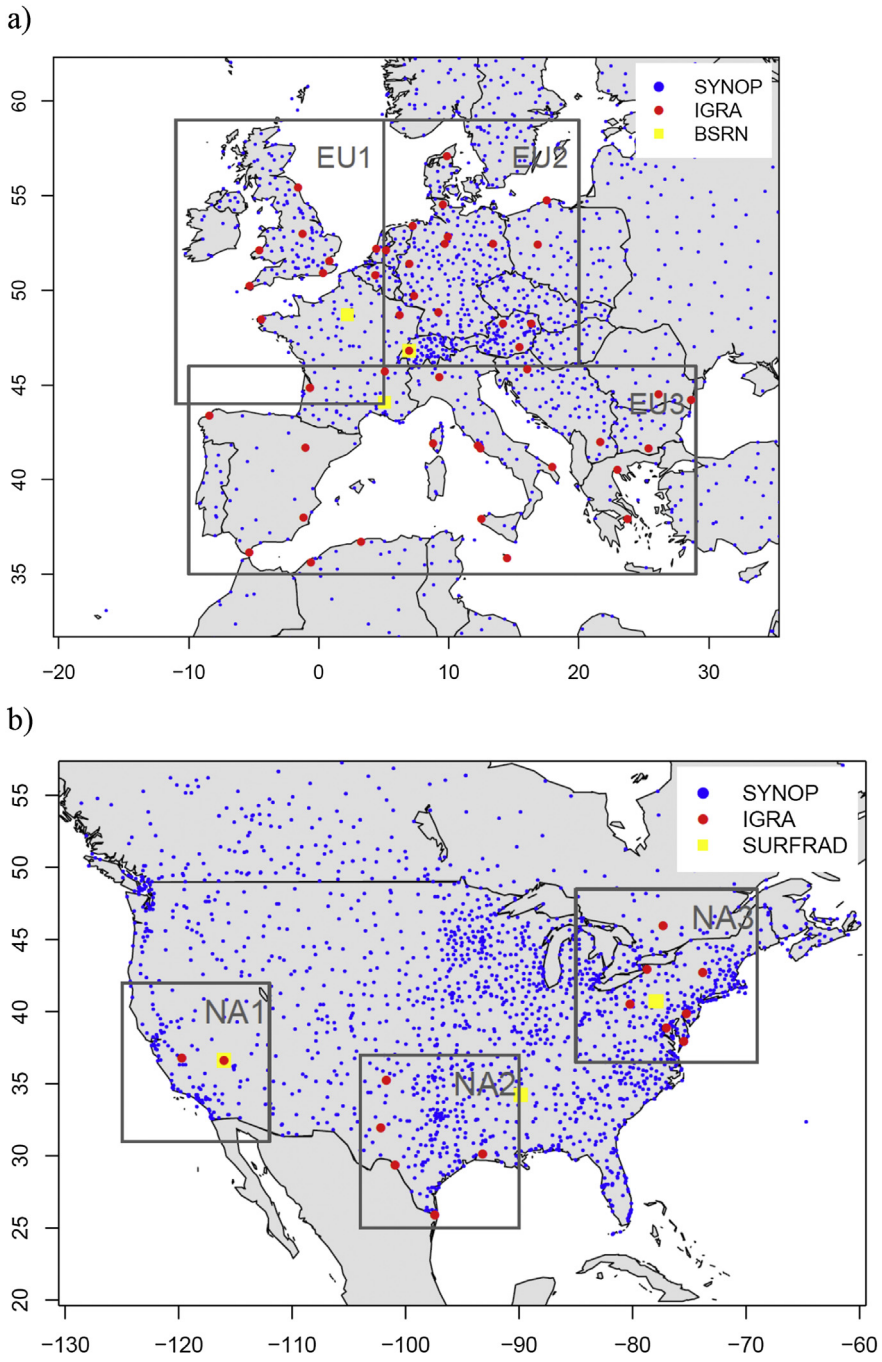

Fig. 1. Map of (a) the three European and (b) three North American domains selected for the comparative analysis. The European domains are EU1: 11°W–05°E/44°N–59°N; EU2: 05°E–20°E/46°N–59°N; EU3: 10°W–29°E/35°N–46°N. The North American domains are NA1: 31°N–42°N, 125°W–112°W; NA2: 25°N–37°N, 104°W–90°W; NA3, 36.5°N–48.5°N, 85°W–69°W. The locations of SYNOP surface meteorology sites, IGRA radiosonde sites, and BSRN/SURFRAD radiation sites are indicated as symbols.

reset to the lifting condensation level in cumulus capped layers and so this value will tend to be higher than the depth derived from the observations under unstable conditions.

Over Europe, the radiosonde data set is based on comparatively coarse resolution profiles (on average 16 levels below 500 hPa) from the Integrated Global Radiosonde Archive (IGRA) maintained by NOAA (Durre and Yin, 2008), which limits the quality of the PBL height estimates particularly for stable nocturnal boundary layers. A comparison with high-resolution soundings indicates that for low PBL heights <500 m this introduces a relative uncertainty of up to 80%, while for PBL heights >1000 m the uncertainty is usually less than 20% (Seidel et al., 2012). A further important limitation is that radiosondes are only launched twice a day at the synoptic hours 00 UTC and 12 UTC. Over Europe, this roughly corresponds to a midnight and a noontime profile. Over North America, the timing is more problematic with respect to model evaluation since the soundings are performed during the transition periods of the PBL evolution in the morning and evening hours.

3.3. Precipitation

The European precipitation simulations have been evaluated

against gridded daily accumulated precipitation provided by the E-OBS data set (Haylock et al., 2008). E-OBS itself is based on observations from approximately 200 sites in Europe collected for the European Climate Assessment (Klein Tank et al., 2002) and interpolated to a regular grid by Kriging as described in Haylock et al. (2008). For North America, the precipitation simulations were evaluated on a monthly basis, using PRISM (Parameter-elevation Regressions on Independent Slopes Model), a robust, high resolution gridded data set developed by the PRISM Climate Group, Oregon State University [http://prism.oregonstate.edu] (Di Luzio et al., 2008). The PRISM data, which uses a 4 km grid, was aggregated up to the common 0.25° grid used by all modeling groups, allowing direct comparisons.

3.4. Radiation

Over Europe, incoming short-wave radiation of the models was compared with measurements from the Baseline Surface Radiation Network (BSRN; www.bsrn.awi.de) which provides highly accurate measurements though only at a few sites. For the comparison, only three BSRN sites (Palaiseau, France; Payerne, Switzerland; Carpentras, France) were available fulfilling the criteria that the station should be located in one of the three subdomains and should be included in the list of receptor sites for which meteorological output was generated by the models. The three stations are located in the domains EU1, EU2 and EU3 respectively, and below 500 m a.s.l. Although the number of observations is limited, the quality of the measurements is high and provides good insights in the models' treatment of the radiative transfer. Unfortunately, the sites are located rather close to each other (see Fig. 1a) and thus do not well represent the spread of climatic conditions across Europe.

Over North America, observations from the NOAA surface radiation network (SURFRAD; www.esrl.noaa.gov/gmd/grad/surfrad/index.html) at the sites Desert Rock (Nevada), Goodwin Creek (Mississippi), and Penn State University (Pennsylvania), were available in the three subdomains NA1, NA2 and NA3, respectively. These sites provide a good representation of different climatic conditions in North America.

4. General weather situation in 2010

Since model performance is analyzed for a single year only in this study, it is useful to put the year 2010 into perspective with the long-term climatological mean weather. This helps to understand to what extent the air pollutant concentrations in different periods of the year corresponded to typical or rather untypical weather and to explain certain phenomena observed in 2010. A similar analysis has been performed by Vautard et al. (2012) for the year 2006. Fig. 2 shows the seasonal mean anomalies of temperature at 850 hPa and geopotential at 500 hPa in the year 2010 with respect to the

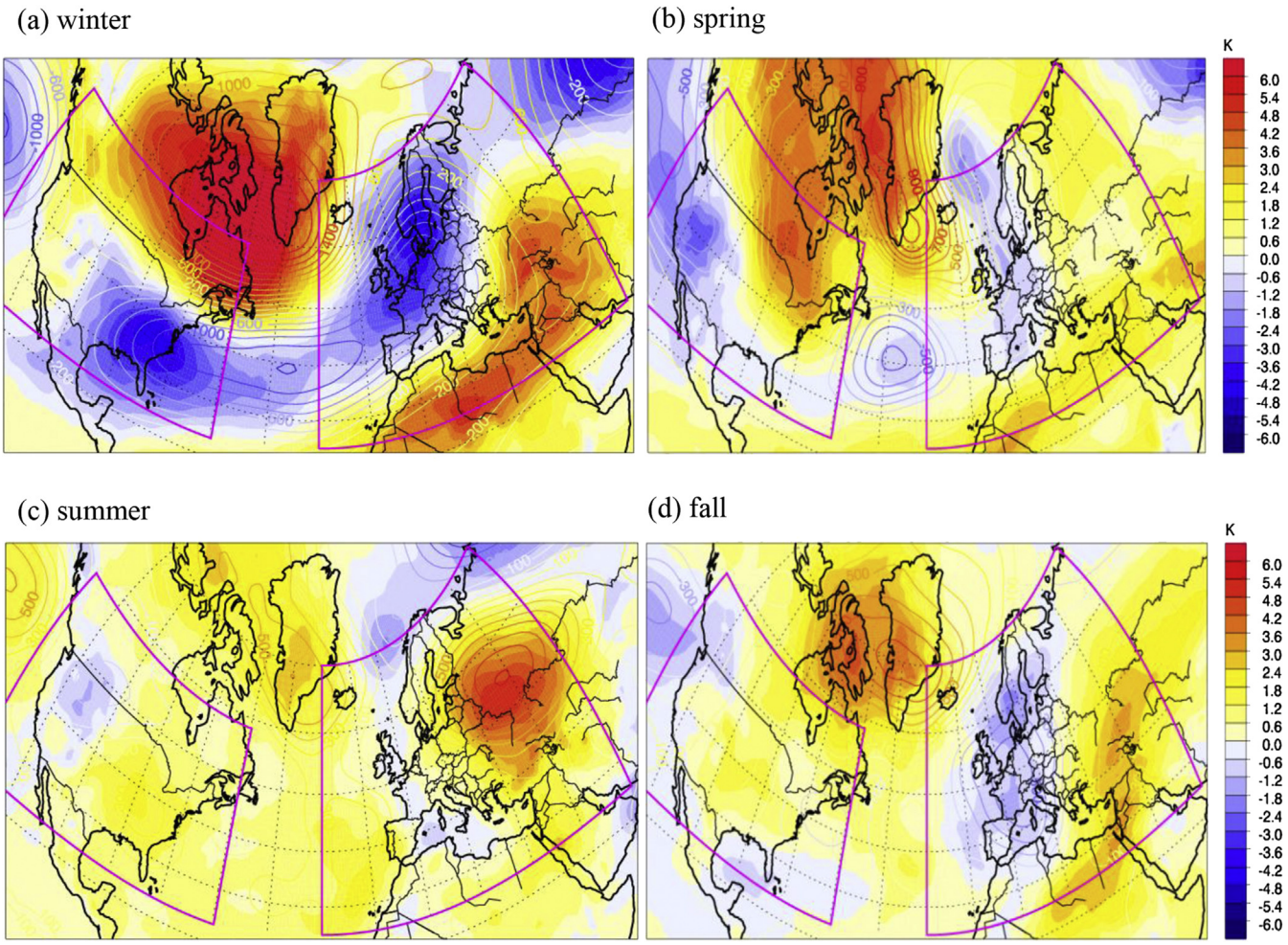

Fig. 2. Seasonal anomalies of temperature at 850 hPa (filled contours) and geopotential at 500 hPa (line contours) in 2010 with respect to the 30-year climatology 1980–2009. Purple lines show the EU and NA model domains. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1980–2009 mean based on the ERA-Interim reanalysis of the European Centre for Medium Range Weather Forecasts (ECMWF).

The winter 2009–2010 was outstanding in many regions of the northern hemisphere. It featured the largest negative index of the North Atlantic Oscillation (NAO) in a century, probably related to an unusual coincidence of a strong El Niño and an easterly phase of the stratospheric Quasi-Biennial Oscillation (Fereday et al., 2012). Accordingly, the figure displays very strong anomalies in winter with negative departures over the east coast of the United States and north western Europe and a positive anomaly extending from Greenland to Canada, a pattern consistent with the negative phase of the NAO. The eastward displacement of the climatological Icelandic low favored the advection of Arctic air into Europe giving rise to unusually cold and wet weather in most parts of the continent. It caused extreme winter precipitation over the Iberian peninsula (Vicente-Serrano et al., 2011) and a series of severe winter storms (Daisy, Xynthia, Andrea) particularly affecting the UK, Spain, and the western provinces of France. As a consequence of the frequent passage of low pressure systems across Europe, no significant winter-smog episodes, which typically develop in stagnant high pressure situations, were recorded in Europe during this winter.

Another remarkable feature of the year 2010 was the strong high pressure anomaly and associated heat wave over Russia which triggered the most severe wildfires on record to that time producing extreme levels of CO and PM_{10} concentrations in Moscow and other regions (Konovalov et al., 2011). An analysis of the influence of aerosol direct and indirect effects on AQMEII-2 model simulations during the Russian forest fire episode is presented in Kong et al. (2015). Over the western and southern parts of Europe, 2010 was a rather mild year in terms of forest fires, except for Portugal, where large forest areas were burned during the first half of August (JRC, 2011) consistent with the positive anomaly in Fig. 2c. In terms of ozone concentrations over Europe, summer 2010 was comparable to the previous three years. The warmest period lasted from 24 June to 22 July during which 85% of the exceedances of the information threshold (a 1-h average ozone concentration of 180 $\mu g/m^3$) in 2010 were reported (EEA, 2011).

Major episodes of Saharan dust outbreaks to Europe occurred in the periods 17–18/2, 15–19/4, 13–15/5, 2–15/10, 25–28/11, 20–26/12, supported by a pattern of low pressure anomalies over western Europe in winter, spring and fall but not in summer.

Meteorological conditions were also unique across much of North America, as the continent was influenced by the historically strong NAO mentioned earlier, a persistent Bermuda High impacting eastern sections of the domain and a rapid transition from a strong El Niño phase early in the year to a strong La Niña later in the year. While conditions are summarized briefly below, further details of the meteorological impact on the NA domain can be found in (Stoeckenius et al., 2015), who contrast conditions in 2006 (for Phase 1) to 2010 (Phase 2). Canada experienced its warmest year on record (+3.1 °C). Although much of the anomaly occurred in northern Canada, outside of the modeling domain, it was prevalent throughout each season as winter (+3.9 °C), spring (+4.1 °C), summer (+1.3 °C) and fall (+2.4 °C), each experienced well above normal temperatures. These anomalies coincide well with the seasonal anomalies of the 850 hPa temperature and 500 hPa Geopotential heights shown in Fig. 2. Despite the record warm temperatures, annual precipitation was near normal across much of the Country (2.0% above normal). Seasonally, spring (−1%), summer (+5%) and fall (+5%) were close to normal. Winter, conversely was considerably drier than normal as precipitation was 22% below normal. Spatially, Saskatchewan and Manitoba were considerably wetter than normal (+20%), while British Columbia, and parts of Ontario were drier than normal (−20%).

The United States also experienced above normal temperatures (+0.5 °C) for 2010, with the anomalies greatest over the Great Lake and New England States. The only area of the contiguous U. S. not experiencing warm conditions was in the southeast, where temperatures were below normal. Seasonally, the NAO resulted in below normal temperatures across most of the U. S, during the winter (−0.3 °C), the exception being several States along the Canadian border. Spring (+0.7 °C), summer (+1.0 °C) and fall (+0.8 °C) were each above normal which is attributed to the strengthening La Niña and a persistently westward extension of the Bermuda High. Annual precipitation across the U. S. was slightly above normal (4.5%), with winter (11.9%) and summer (+10.7%) considerably wetter and spring (0.0%) and summer (+1.2%) near normal. Spatially, the southeast and southern plains were the only areas to experience below normal precipitation, with the remainder of the Nation slightly or much above normal.

Despite the often anomalous meteorological conditions experienced by NA domain, air quality was fairly consistent with previous years. The most notable exceptions being an increase in summer ozone concentrations across the eastern United States associated with the westward extension of the Bermuda High discussed above and several $PM_{2.5}$ exceedances associated with brief stagnation episodes in the San Joaquin Valley of California and the upper midwest during the winter.

5. Quantitative model evaluation

Model output was generated in several different ways and submitted to the ENSEMBLE web-based evaluation system (Bianconi et al., 2004; Galmarini et al., 2012) hosted by the Joint Research Centre where also all observation data required for the comparative analysis were stored. Model output was generated as hourly gridded fields on a common grid of 0.25° × 0.25° resolution covering either Europe or North America, as hourly time series at prescribed measurement station locations (receptor points), and as vertical profiles above a few selected sites.

In this study we mostly rely on the receptor output generated at a large number of meteorological surface measurement sites (see Fig. 1) as well as model profiles interpolated to the location of Frankfurt airport.

The models are primarily evaluated in terms of an operational analysis. As defined by Dennis et al. (2010) an operational analysis aims at determining whether model estimates are in agreement with the observations in an overall sense. The analysis concentrates on the capability of the models to reproduce observed seasonal and diurnal variations. To calculate diurnal averages, all data (models and observations) were adjusted to local solar time at each receptor point separately. In addition to presenting graphs of mean seasonal and diurnal cycles, model performance is quantified in terms of the following performance metrics computed from daily mean values: Mean bias (*MB*), i.e. model mean − measured mean value, centered (unbiased) root mean square error $CRMSE(=\sqrt{RMSE^2 - MB^2})$, and Pearson linear correlation coefficient *r*. Computing these performance metrics from daily mean rather than hourly values removes the influence of diurnal variability and instead emphasizes the capability of the models to reproduce day-to-day (and seasonal) variations.

For both continents, the comparative analysis is performed for three subdomains separately which were selected to cover a range of different climatic conditions. The domains are outlined in Fig. 1. Domain EU1 includes the British Isles and western France and is identical to domain EU1 defined in Vautard et al. (2012). It is characterized by a mixed maritime − continental climate influenced by the North Atlantic. Domain EU2 extends somewhat further to the north but less to the east compared to the domain EU2 selected in Vautard et al. (2012). It represents a rather

continental climate with a strong seasonal cycle. Domain EU3 covers southern Europe and is representative of a Mediterranean climate. The Alps and Pyrenees are important barriers to the flow separating the climate of EU3 from those in EU1 and EU2. A comparatively large domain was selected for EU3 to compensate for the lower density of observations in this region.

The North American domains NA1-NA3 are identical to those defined in Vautard et al. (2012). Region NA1 covers the southwestern US to the west of the Rocky Mountain barrier. It is characterized by high solar radiation and low relative humidity. Domain NA2, the Texas area, has also high solar radiation but in contrast to NA2 is characterized by a hot and humid climate. NA3 covers the northeastern US and parts of Canada and includes the largest urban centers in eastern NA (New York, Philadelphia Toronto, Montreal). It is characterized by a continental climate with a strong seasonal cycle.

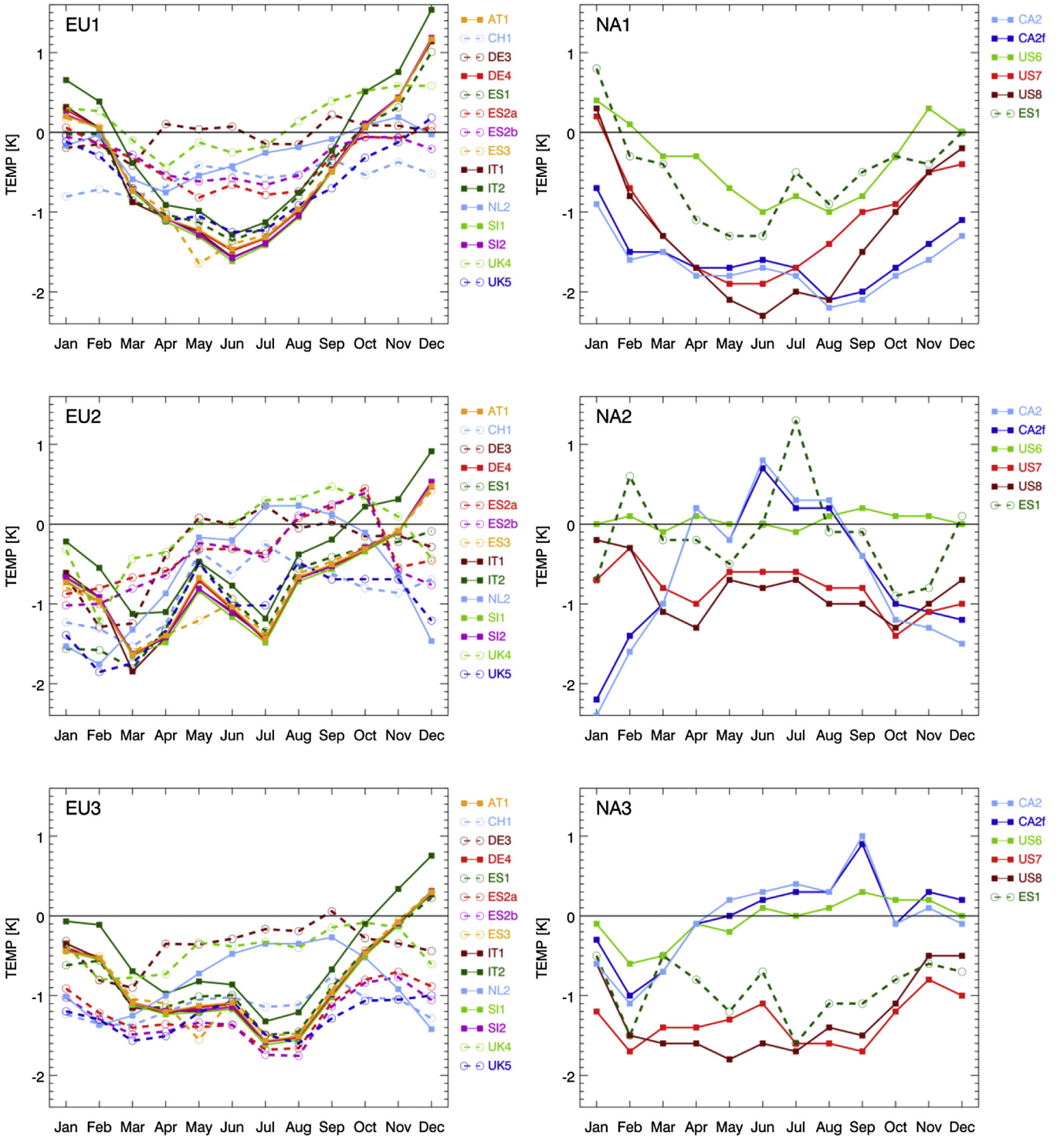

Fig. 3. Seasonal cycle of monthly mean 2 m temperature biases in European domains EU1–EU3 and North American domains NA1-NA3.

5.1. Temperature

Temperature is of prime importance for atmospheric chemistry as it controls the rate of chemical reactions and also alters the gas-particle phase partitioning, thus altering the aerosol concentrations. Furthermore, temperature affects biogenic VOC and emissions. Fig. 3 shows the model biases in monthly mean 2-m temperatures. In domain EU1 the WRF-based models all exhibit a pronounced seasonal cycle in temperature bias with under-predictions in the spring and summer months (up to 1.5 K) and

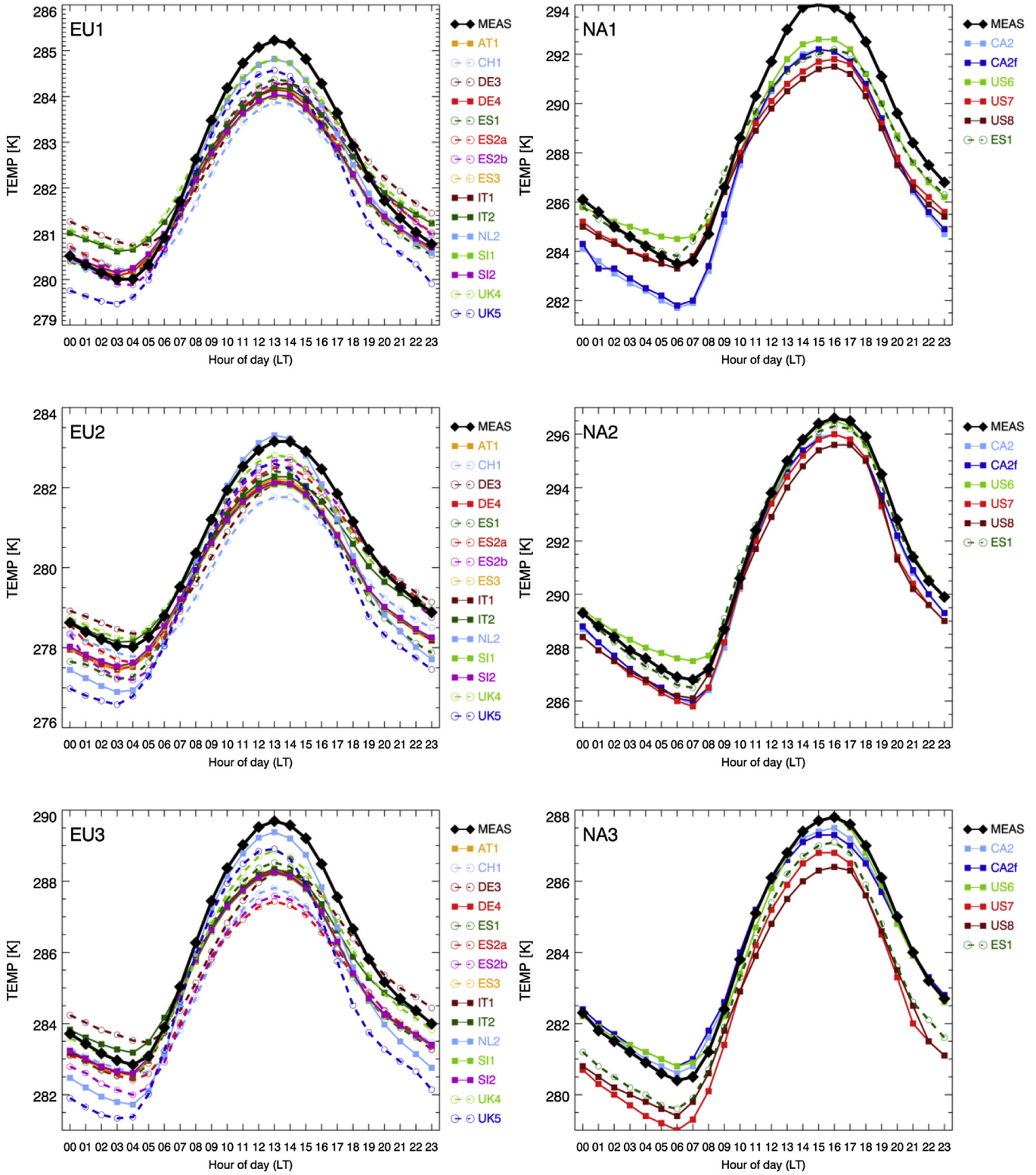

Fig. 4. Mean diurnal cycle of 2 m temperature in European domains EU1–EU3. Measured values are shown in black. Temperature ranges are different for Europe and North America. Vertical axis ranges have been optimized for each domain separately.

smaller biases in fall and winter. In the final two months of the year, the WRF-Chem models develop a positive bias, in particular IT2. DE3 and UK4 show the smallest overall biases. ES2a/b and CH1 have persistent negative biases of up to 0.8 K NL2 shows a negative bias from March to August but closely traces the observations in the remaining months. In domain EU2 with its more continental characteristics, the models again show mostly negative biases but here they are the greatest in the early months of the year with values between −0.5 K and −2 K in February and March. Domain EU3 is more consistent with EU1 with the largest negative biases in most models seen in the summer. Again, DE3 and UK4 have the smallest overall biases while the WRF-based models underpredict 2-m temperatures by up to 1.5 K except for winter. CH1, ES2a/b and UK5 have persistent negative biases between about 1 and 1.5 K. The WRF-Chem models SI2 and SI1 with and without direct aerosol effects, respectively, differ only marginally indicating that direct aerosol effects have little impact on mean temperatures in this model. Model IT2 is somewhat warmer than the other versions of WRF-Chem in all three domains. With respect to meteorology, the configurations of the WRF-Chem models AT1 and DE4 are identical to those of SI1 and SI2 (Table 1), but AT1 and DE4 additionally consider aerosol indirect effects. Including these effects slightly reduces the negative temperature biases in summer in all three domains (by less than 0.2 K), but these improvements are smaller than the differences from other models.

For the NA domain, all models (except US6, at +0.1 K) underpredict surface temperatures when examined over the full year. Across the full domain, this underprediction ranges from −0.2 K for ES1, −0.5 K for both CA2 simulations to −0.9 K for US6 and US7. When examined monthly and among the three subdomains, considerable variability is revealed in Fig. 3. For NA1, the underprediction reveals a seasonality similar to the results obtained for Europe, with the most pronounced bias occurring during the warmer months from April through August, with several models (CA2, CA2f, US7 and US8) underpredicting temperatures by more than 1.5 K US6 and ES1 also underpredict during this time period, though not as badly. The underprediction is mitigated during the cooler months, with several models actually overpredicting during January. Biases associated with NA2 are more complex, with little continuity among the models. US7 and US8 underpredict throughout the year, while CA2 and CA2f underpredict during the cooler months and overpredict during the warmer months. Conversely ES1 and US6 oscillate monthly, with US6 exhibiting the smallest bias. And finally, for NA3, the models cluster into two bias regimes. US7, US8 and ES1 are consistently biased low (between 1 and 1.5 K), while US6 and CA2 and CA2f produce mixed results with negative biases at the beginning of the year, especially February. These three models then produce smaller, slightly positive biases throughout the remainder of the year. The generally better performance of US6 as compared to the other models is likely a result of the weak nudging of temperatures at upper levels above the PBL.

Fig. 4 shows the annual mean diurnal cycle of temperature in the models and the observations. In the domain EU1 the models generally underestimate the amplitude of the diurnal cycle. A large group of models agree well with observations at night but begin to diverge from the observations from about 06:00 and all models have a negative temperature bias at noon. UK4 and NL2 have the smallest noontime bias but UK4 has a positive bias while NL2 is neutral at night. IT2 and DE3 have a similar positive bias to UK4 at night but have larger negative biases at noon. UK5 has a more consistent bias than the other models and generally reproduces the amplitude of the diurnal cycle better, with zero or negative bias at all hours. The picture is similar in domains EU2 and EU3 in that all models except NL2 still have a negative noontime bias, but for this domain the UK4 and IT2 models agree best overnight, with the amplitude of the diurnal cycle also most closely matched by UK4.

NL2 and UK5 have the largest diurnal cycle in EU3 and the two COSMO models CH1 and DE3 the smallest. This underestimation of the diurnal amplitude by CH1 and DE3 is also seen in the other domains. The noontime biases are quite large in some models in EU3 reaching values of about −2 K in the models ES2a/b and CH1. The timing of the diurnal cycle agrees well between the models and observations although time shifts of the order of 1 h are clearly present, for example with respect to the timing of the early morning temperature rise.

These issues in capturing the magnitude and timing of the diurnal cycle of temperature are due to a combination of factors including the representation of the boundary layer evolution (see Sect. 5.3) or the calculation of radiation (e.g., the frequency of radiation updates). Several models such as CH1 and DE3 that underestimate radiation most likely due to excessive cloudiness particularly during summer (see Sect. 5.5), also tend to underestimate the amplitude of the diurnal cycle and peak temperatures at noontime. Comparing the results of the WRF-Chem models SI1, SI2, AT1 and DE4 suggests that the influence of direct and indirect radiative effects of aerosols on annual mean diurnal temperatures is only minor over Europe. It would be useful to further investigate both the causes and impacts of these errors on air quality and other meteorological variables with some additional sensitivity studies.

For the NA domain, the diurnal temperature range appears to be handled better by the models. The phase of the diurnal cycle and differences in amplitude between the different domains are well captured. However, all six models display a general negative or cold bias, similar to the results over Europe. In fact, with only a few exceptions, each of the models underpredict uniformly across most hours of the day. The underprediction is worst for NA1, especially during the early afternoon hours. The major exception to the uniform underprediction is seen with US6, which slightly overpredicts temperatures in the early morning hours, most notably in NA1 and NA2. The Canadian model CA2/CA2f shows the largest underprediction in NA1 but shows the highest temperatures of all models in NA3 where it closely follows the observations. Differences between the simulations with (CA2f) and without (CA2) aerosol feedbacks are small.

Table 3 gives statistics of daily mean temperatures versus observations for the EU1 and EU2 domains and Table 4 for EU3. Correlations are high at all sites but this is probably mainly driven by the fact that the models correctly represent the seasonal cycle. The mean biases are fairly small i.e. less than 1 K for almost all models in all 3 domains but the centered root mean square error (CRMSE) is larger, especially in EU2 and EU3 where the error is greater than 2 K for all models (Table 5).

Fig. 5 shows profiles of mean bias, CRMSE, and correlation calculated using data from 757 MOZAIC profiles at Frankfurt airport for the EU domain, and from 120 profiles at Toronto airport for the NA domain. The model data are interpolated onto the aircraft trajectories at heights of 0, 100, 250, 500, 750, 1000, 2000, 3000, 4000, 5000, 6000, 7500 and 8500 m above ground. Above the surface the models have negative biases of around 2 K both at Frankfurt and Toronto. These biases are rather consistent among models except for CH1, which has lower biases above 6 km and CA2/CA2f which has a large negative bias of up to −6 K at the upper levels. The CRMSE are highly variable from level to level and are typically around 2 K. The models DE3 and CA2/CA2f have the largest CRMSE values. This may point to deficiencies in the global meteorological data sets driving these models at the domain boundaries.

As for the surface observations, all models have high correlation coefficients (>0.93, generally > 0.96) at all model levels. The high correlations are due to the strong seasonal cycle of temperature which is well represented by all models.

Table 3
Statistics of model performance for 2-m temperature (T2) and 10-m wind speed (WS10) for European domains EU1 and EU2. All statistics calculated from daily means.

Model	Mean	Stdev	r	MB	CRMSE	Mean	Stdev	r	MB	CRMSE
	T2 EU1 (K) (237stations)					**WS10 EU1 (m/s) (236 stations)**				
obs	282.4	6.7	–	–	–	4.1	2.2	–	–	–
AT1	281.9	6.0	0.97	−0.4	1.7	5.1	2.3	0.73	1.0	1.7
CH1	281.8	6.7	0.98	−0.6	1.4	4.5	2.3	0.72	0.5	1.7
DE3	282.4	6.7	0.97	0.0	1.5	4.8	2.4	0.70	0.8	1.8
DE4	281.9	6.0	0.97	−0.4	1.7	5.1	2.3	0.73	1.0	1.7
ES1	282.0	6.1	0.97	−0.4	1.6	5.1	2.3	0.73	1.0	1.7
ES2a	282.0	6.4	0.97	−0.4	1.7	5.5	2.5	0.72	1.4	1.8
ES2b	282.1	6.5	0.97	−0.3	1.7	5.4	2.5	0.72	1.4	1.8
ES3	281.9	6.0	0.96	−0.5	1.9	5.1	2.3	0.73	1.0	1.7
IT1	281.9	5.9	0.97	−0.5	1.7	5.1	2.3	0.73	1.0	1.7
IT2	282.3	5.9	0.97	−0.1	1.7	5.5	2.5	0.72	1.4	1.8
NL2	282.2	6.6	0.98	−0.2	1.5	4.7	2.3	0.73	0.6	1.7
SI1	281.9	5.9	0.97	−0.5	1.7	5.1	2.3	0.73	1.0	1.7
SI2	281.9	5.9	0.97	−0.5	1.7	5.1	2.3	0.73	1.0	1.7
UK4	282.5	6.4	0.98	0.1	1.4	4.4	2.1	0.71	0.4	1.7
UK5	281.7	6.2	0.97	−0.6	1.6	5.1	2.3	0.75	1.1	1.6
	T2 EU2 (K) (373stations)					**WS10 EU2 (m/s) (372 stations)**				
obs	280.4	8.8	–	–	–	3.4	2.1	–	–	–
AT1	279.7	8.6	0.95	−0.7	2.8	4.7	2.3	0.54	1.3	2.1
CH1	279.6	9.0	0.96	−0.8	2.6	3.3	1.9	0.58	−0.1	1.9
DE3	280.1	9.1	0.95	−0.3	2.8	3.7	1.9	0.57	0.3	1.9
DE4	279.6	8.6	0.95	−0.8	2.8	4.6	2.3	0.54	1.3	2.1
ES1	279.5	8.9	0.95	−0.9	2.9	4.6	2.3	0.54	1.3	2.1
ES2a	280.0	9.0	0.95	−0.3	2.8	4.9	2.3	0.58	1.5	2.0
ES2b	280.0	9.0	0.95	−0.4	2.9	4.9	2.3	0.58	1.5	2.0
ES3	279.6	8.6	0.95	−0.8	2.9	4.7	2.3	0.54	1.3	2.1
IT1	279.7	8.6	0.95	−0.8	2.8	4.6	2.3	0.54	1.3	2.1
IT2	280.0	8.6	0.95	−0.4	2.8	5.0	2.4	0.54	1.7	2.2
NL2	279.8	9.4	0.95	−0.6	3.0	3.3	2.0	0.57	−0.1	1.9
SI1	279.6	8.6	0.95	−0.8	2.8	4.6	2.3	0.54	1.3	2.1
SI2	279.7	8.6	0.95	−0.7	2.8	4.6	2.3	0.54	1.3	2.1
UK4	280.3	8.9	0.95	−0.1	2.7	3.2	2.1	0.53	−0.1	2.0
UK5	279.4	9.0	0.95	−1.0	2.8	4.5	2.1	0.6	1.1	1.8

In general it seems unlikely that the magnitude of temperature errors seen here have a dominant impact on the chemistry or aerosols when compared to the magnitude of other errors. The issues seen with the diurnal cycle of temperature in some models are probably related to biases in cloudiness and hence in radiation (see Sect. 5.5) as well as to issues in the representation of the boundary layer structure and its development, which in turn affect the formation and evolution of air pollutants. The ability of models to represent the advective and turbulent transport of pollutants near the surface will have a much stronger direct impact on the concentrations of pollutants seen at the surface. An analysis of model performance with respect to wind speed is presented next.

5.2. Wind

Wind speed and direction control the horizontal transport and thereby the spatial distribution of pollutants. Wind speed is a particularly important parameter as it influences the volume of air into which emissions are diluted, determines the transport time between sources and receptor locations, and also controls the emission of sea salt and dust. In this evaluation we focus on wind speeds at 10 m above surface (WS10) as well as on vertical profiles. Wind direction is not evaluated. While the evaluation of wind directions at single stations (e.g. by comparing wind roses) is useful, a statistical evaluation of wind directions is complicated by the fact that wind direction errors typically become large at low wind speeds (Jiménez and Dudhia, 2013). Furthermore, an average wind direction for a given subregion would provide little useful information and other statistics such as correlations or RMSE would not be useful at all.

Figs. 6 and 7 show the seasonal and diurnal cycles of WS10, respectively. In the European domain EU1, most models overestimate wind speed in all months. The seasonal cycle is well reproduced however, with all models showing maxima in November and March. This is consistent with the findings for the models participating in AQMEII Phase 1 for domain 1 which matches the domain used here. It seems likely that this overestimate of wind speeds will lead to emissions being too strongly diluted and too rapidly transported from polluted centers to rural areas in this region. In domains EU2 and EU3 a similar picture is seen for all models except UK4, CH1, DE3 and NL2 which show better agreement with the monthly mean wind speed and tend to slightly underestimate WS10 in EU3. The WRF-Chem models, on the other hand, form a rather compact cluster with IT2 showing somewhat higher WS10 than the other versions.

A similar picture emerges for NA with a tendency of models to overestimate WS10 but to closely trace the overall seasonal pattern. However, the biases vary significantly between the different evaluation subdomains and tend to be larger in domains NA1 and NA3 with lower average wind speeds than in domain NA2. WRF is known to overpredict at low to moderate wind speeds in all PBL schemes available in WRF (Mass and Ovens, 2011), due in part to unresolved topography such as hills and valleys and other smaller scale terrain features by the default surface drag parameterization and in part to the use of coarse horizontal and vertical resolutions (Cheng and Steenburgh, 2005). Similar large positive biases in WS10 were found previously in WRF simulations for both winter and summer over Europe (Zhang et al., 2013) and the U.S. (Penrod et al., 2014; Yahya et al., 2014). Model US8, on the other hand, closely follows the observations in domains NA1 and NA3 but is biased low in region NA2. The good agreement of US8 in domains NA1 and NA3 is due to the use of a simplified surface drag

Table 4
Statistics of model performance for 2-m temperature (T2) and 10-m wind speed (WS10) for European domain EU3.

Model	Mean	Stdev	r	MB	CRMSE	Mean	Stdev	r	MB	CRMSE
	T2 EU3 (K) (121stations)					WS10 EU3 (m/s) (122 stations)				
obs	286.0	8.4	–	–	–	3.4	2.4	–	–	–
AT1	285.2	8.0	0.96	−0.8	2.5	4.6	2.4	0.58	1.2	2.2
CH1	284.9	8.3	0.95	−1.1	2.5	3.1	1.9	0.55	−0.3	2.1
DE3	285.6	8.4	0.95	−0.4	2.6	3.6	2.0	0.52	0.1	2.2
DE4	285.2	8.0	0.96	−0.8	2.5	4.6	2.4	0.58	1.2	2.2
ES1	285.2	8.1	0.95	−0.8	2.5	4.6	2.4	0.58	1.2	2.2
ES2a	284.8	8.1	0.94	−1.2	2.9	4.6	2.4	0.55	1.2	2.3
ES2b	284.7	8.1	0.94	−1.3	3.0	4.6	2.4	0.55	1.1	2.3
ES3	285.2	8.0	0.95	−0.8	2.6	4.6	2.4	0.57	1.2	2.2
IT1	285.2	8.0	0.95	−0.8	2.5	4.6	2.4	0.57	1.2	2.2
IT2	285.5	7.9	0.95	−0.5	2.5	5.0	2.5	0.57	1.5	2.3
NL2	285.2	8.7	0.95	−0.8	2.7	3.2	2.1	0.53	−0.2	2.2
SI1	285.1	8.0	0.96	−0.9	2.5	4.6	2.4	0.58	1.2	2.2
SI2	285.2	8.0	0.96	−0.8	2.5	4.6	2.4	0.58	1.2	2.2
UK4	285.6	8.2	0.95	−0.4	2.5	3.1	2.1	0.52	−0.3	2.2
UK5	284.7	8.4	0.94	−1.3	2.8	4.3	2.1	0.59	0.9	2.1

Table 5
Statistics of model performance for 2-m temperature (T2) and 10-m wind speed (WS10) for North American domains NA1-NA3.

Model	Mean	Stdev	r	MB	CRMSE	Mean	Stdev	r	MB	CRMSE
	T2 NA1 (K) (121 stations)					**WS10 NA1 (m/s) (121 stations)**				
obs	288.4	5.9	–	–	–	2.89	0.91		–	–
US6	288.0	5.4	0.95	−0.4	0.72	3.59	1.08	0.57	1.22	0.73
ES1	287.9	5.8	0.92	−0.5	0.96	3.80	1.30	0.46	0.90	0.93
US7	287.3	5.7	0.91	−1.1	0.96	4.12	1.13	0.48	1.22	0.72
US8	287.1	5.4	0.92	−1.3	0.96	3.14	0.80	0.46	0.24	0.53
CA2	286.7	5.7	0.90	−1.7	0.61	2.89	0.86	0.44	0.03	0.46
CA2f	286.8	5.6	0.90	−1.6	0.61	2.93	0.86	0.44	0.06	0.48
	T2 NA2 (K) (245 stations)					**WS10 NA2 (m/s) (245 stations)**				
obs	291.3	9.0	–	–		3.48	1.1	–	–	–
US6	291.3	8.8	0.97	0.1	0.88	4.01	1.01	0.74	0.53	0.28
ES1	291.1	9.1	0.96	−0.1	0.70	3.70	1.12	0.60	0.22	0.64
US7	290.5	8.9	0.97	−0.8	0.70	3.91	1.07	0.69	0.43	0.40
US8	290.4	8.8	0.97	−0.8	1.25	2.79	0.79	0.66	−0.69	0.44
CA2	290.6	9.9	0.96	−0.7	1.29	4.24	1.16	0.66	0.76	0.40
CA2f	290.7	9.8	0.96	−0.6	1.21	4.26	1.17	0.65	0.78	0.39
	T2 NA3 (K) (362 stations)					**WS10m NA3 (m/s) (362 stations)**				
obs	283.9	10.1	–	–		3.08	1.01	–	–	–
US6	283.8	10.1	0.98	−0.1	0.77	4.00	1.18	0.67	0.92	0.21
ES1	283.0	9.9	0.96	−0.9	1.39	3.86	1.17	0.52	0.77	0.70
US7	282.5	10.0	0.97	−1.3	1.55	4.07	1.18	0.62	0.99	0.39
US8	282.6	9.7	0.97	−1.3	1.51	3.10	0.93	0.58	0.01	0.26
CA2	283.8	10.5	0.96	−0.0	0.91	3.72	0.94	0.50	0.64	0.36
CA2f	283.9	10.3	0.96	−0.0	0.85	3.76	0.95	0.50	0.67	0.36

parameterization of Mass and Ovens (2011) in the boundary layer physics scheme. This parameterization, however, affects the US8's performance for high wind speeds, leading to the low bias in domain NA2. The two versions of the GEM-MACH model CA2 and CA2f without and with feedbacks, respectively, both overestimate WS10 in domains NA2 and NA3 but not in NA1. Interestingly, model biases are not changing coherently from one domain to the next. The largest biases in NA1 and NA3, for example, are exhibited by models US6, US7, and ES1, while the largest biases in domain NA2 are found for the model CA2/CA2f.

Since emissions and photochemistry are characterized by a pronounced diurnal cycle, it is useful to examine wind speed as a function of the time of day as presented in Fig. 7. The figure shows that for domain EU1 the wind speed is consistently overpredicted by all models at all times of day. The models CH1, DE3, NL2 and UK4 have the smallest overpredictions but they also have a smaller amplitude of the diurnal cycle so while they agree well with the observations around noon they still overestimate at night. The same models having the lowest biases in EU1 are also more closely following the observations in domains EU2 and EU3. NL2, UK4 and CH1 even slightly underpredict WS10 in EU3, while all other models are biased high as in the other domains.

The shape of the diurnal cycles is well captured by most of the models. In all subdomains, the observations show a peak around 14 LT and the models also have a peak between 12 LT and 14 LT. The diurnal cycle is poorly represented by IT2 in all three domains and has a substantially reduced amplitude of the diurnal cycle compared to both the observations and the other models. The IT2 run is based on a pre-release of WRF 3.4 with an older implementation of the YSU scheme not including some bug fixes introduced in version 3.4.1 [http://www2.mmm.ucar.edu/wrf/users/wrfv3.4/updates-3.4.1.html]. This version has obvious deficiencies regarding the representation of the diurnal cycle of WS10.

Tables 3 and 4 show the performance statistics for daily mean WS10 for Europe. The correlation coefficients are 0.69–0.75 in domain EU1 and 0.53–0.64 in the other two domains. The weaker correlations in domains EU2 and EU3 are possibly related to the large number of mountainous areas in EU2 and the fact that a large

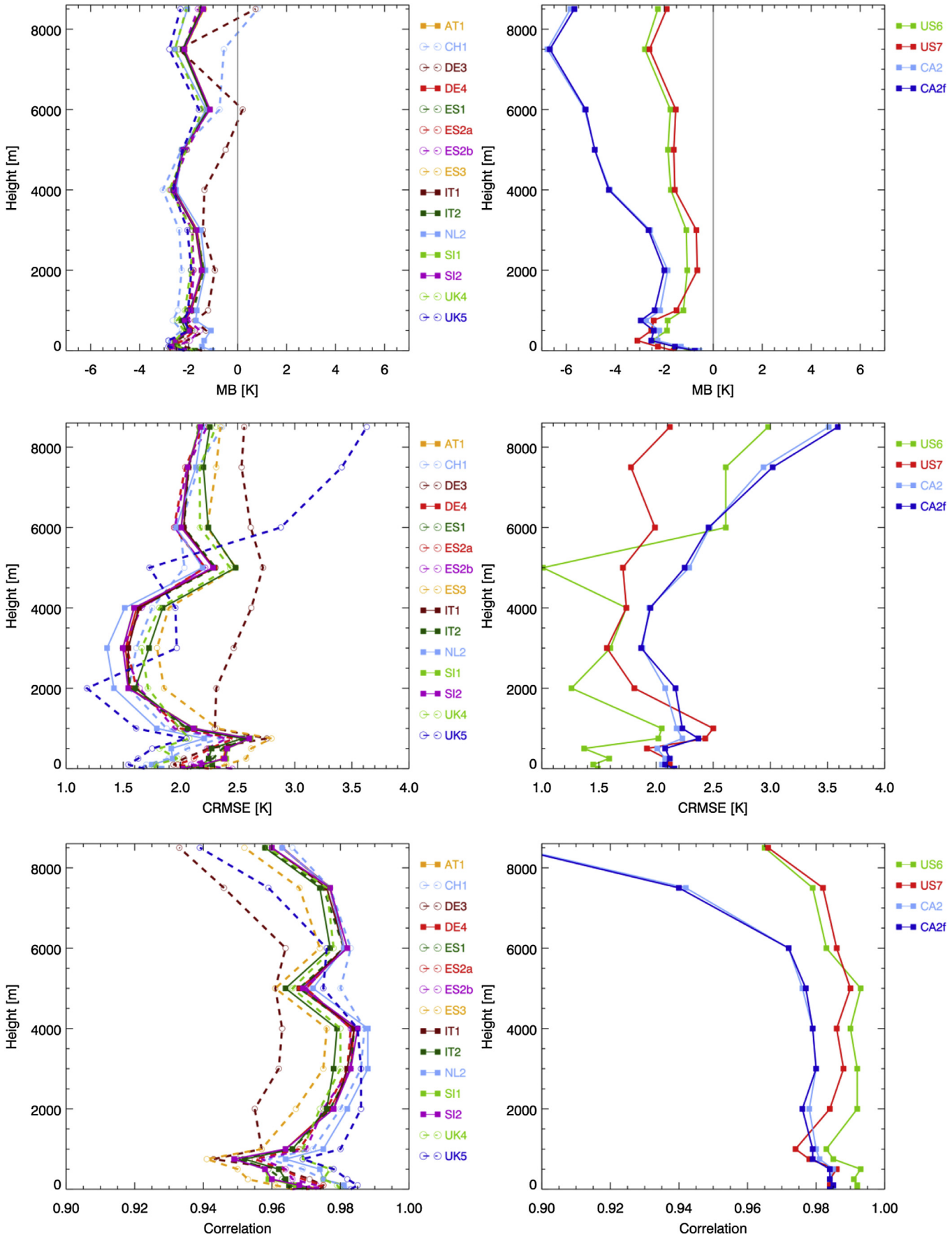

Fig. 5. Vertical profiles of mean bias, CRMSE and correlation for temperature calculated using data from 757 MOZAIC profiles at Frankfurt airport (left) and 120 profiles at Toronto airport (right).

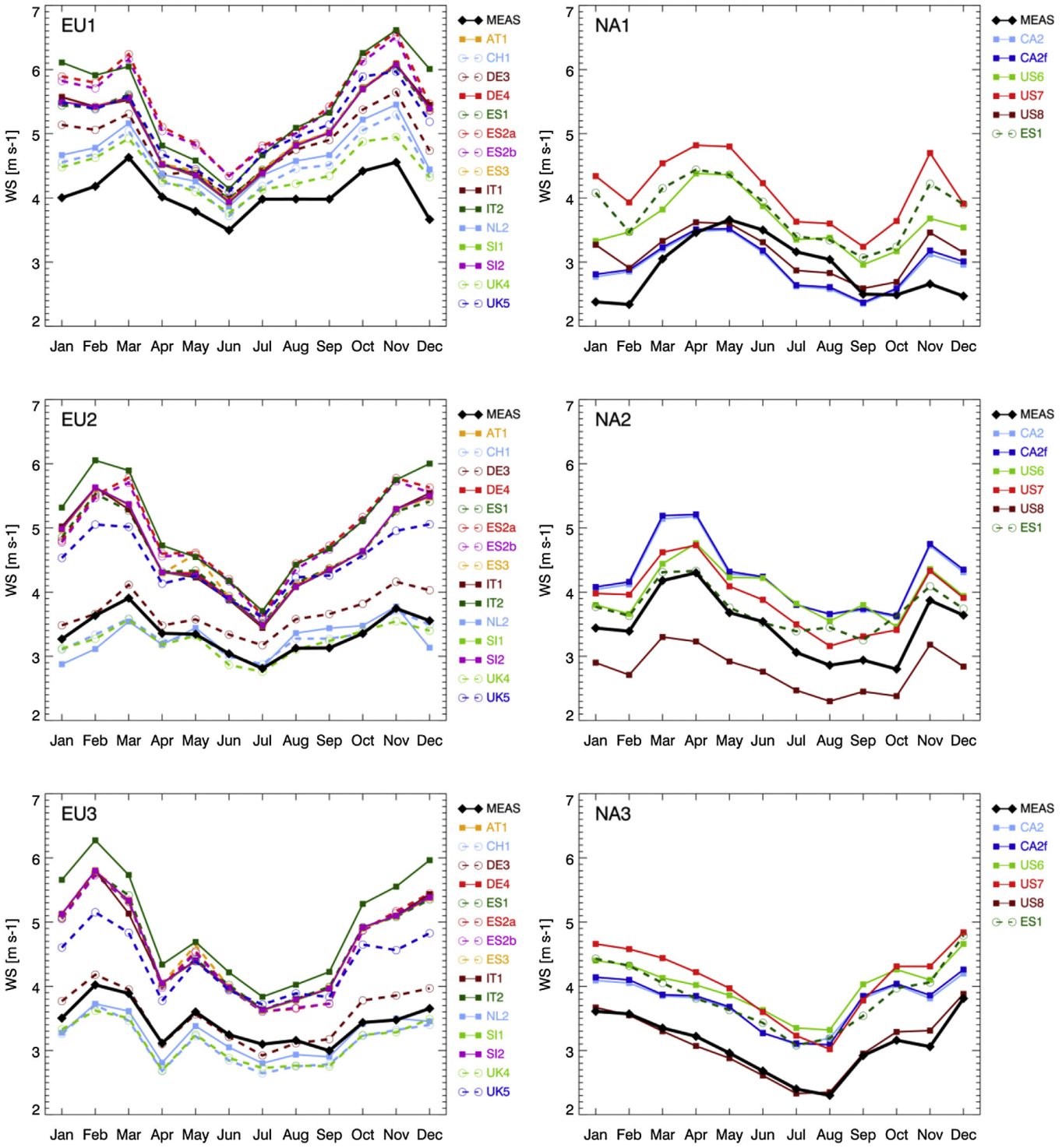

Fig. 6. Season cycle of monthly mean 10 m wind speed (WS10) in European domains EU1–EU3 (left column) and North American domains NA1-NA3 (right column). Measured values are shown in black.

number of sites in EU3 are coastal which are likely to be impacted by sea breeze effects which the models are unable to resolve at the resolutions used in this experiment. In all three domains the models reproduce the wind speed variability well, with the models generally close to the observed variability.

Over NA, the two WRF-Chem models US7 and US8 show a distinctly different behavior not only with respect to mean biases but also with respect to the amplitude and shape of the diurnal profile. While US7 is overestimating WS10 similar to the WRF-Chem models over Europe, US8 tends to underpredict the amplitude of diurnal variations and to be biased low due to the surface drag parameterization of Mass and Ovens (2011). Turning off this parameterization for high winds from mid-morning to late afternoon will help reduce the underpredictions in US8. US7 was the only WRF-Chem model using the more complex MYNN planetary boundary layer scheme, all others were using YSU. Unfortunately,

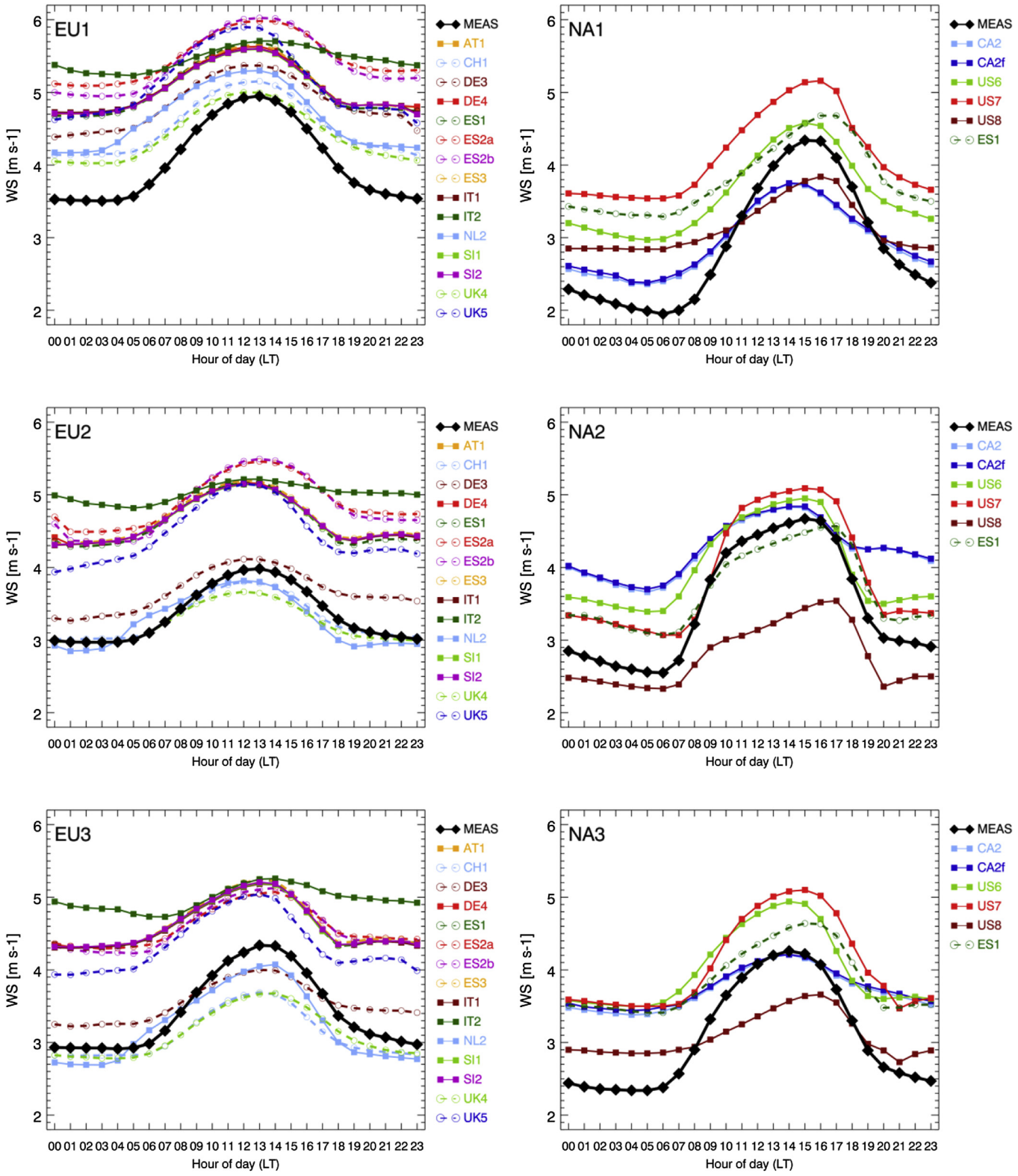

Fig. 7. Mean diurnal cycle of 10 m wind speed in European domains EU1–EU3 (left column) and North American domains NA1-NA3 (right column). Measured values are shown in black.

the effect of using MYNN versus YSU on surface wind speeds cannot be isolated due to additional differences between US7 and US8 regarding the representation of surface drag.

Fig. 8 presents profiles of mean bias, CRMSE and correlation against the MOZAIC data at Frankfurt airport for the EU domain and at Toronto airport for the NA domain, respectively. Consistent with the tendency of the models to overpredict near surface wind speeds, the models show significant positive biases at low

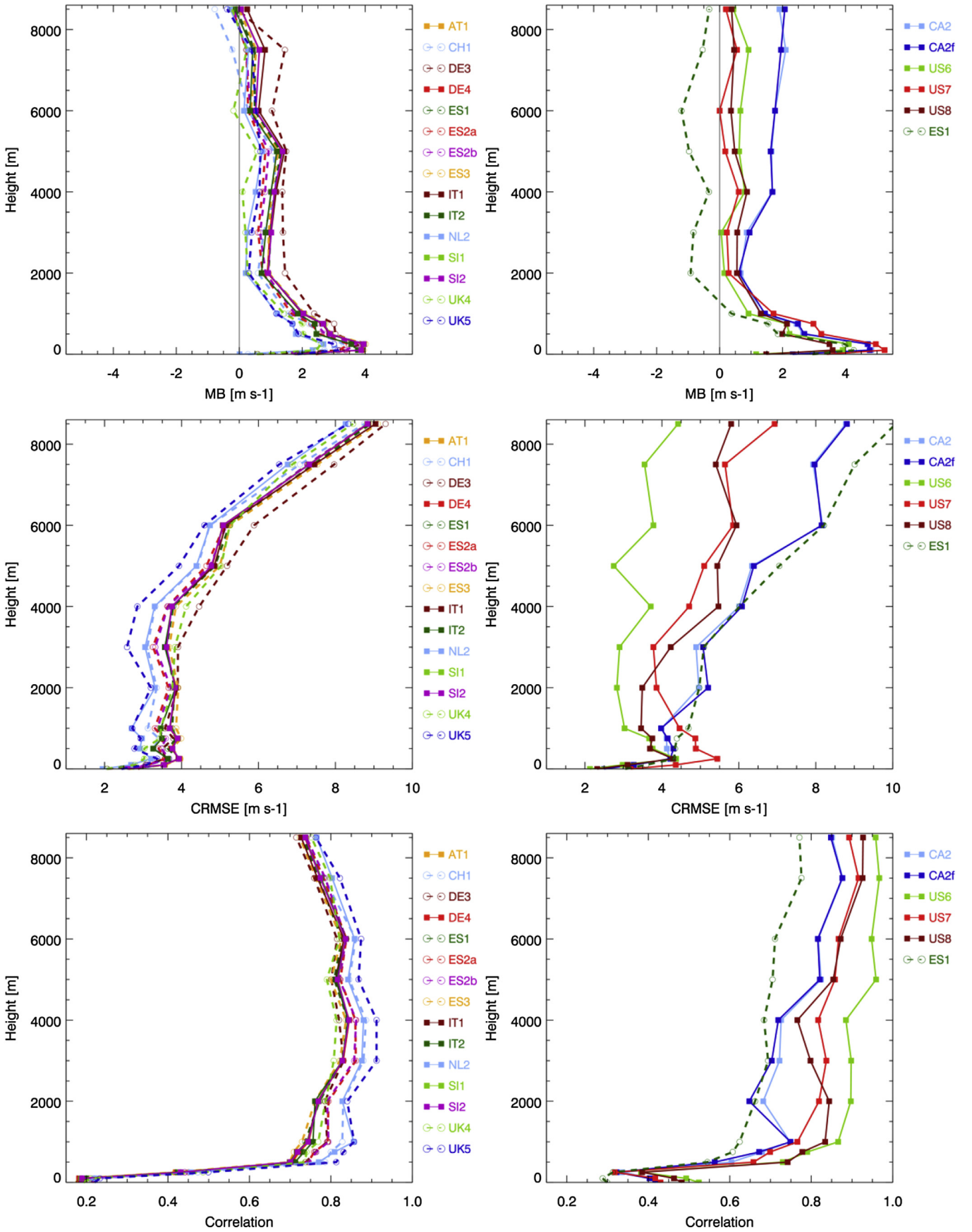

Fig. 8. Vertical profiles of mean bias, CRMSE and correlation for wind speed calculated using data from 757 MOZAIC profiles at Frankfurt airport (left) and 120 profiles at Toronto airport (right).

elevations up to 2000 m above which the mean biases remain rather constant at small values between about 0 and 2 m s^{-1}. ES1 shows a negative bias above 2000 m at Toronto but not at Frankfurt. In contrast to the profiles of bias, CRMSE values remain relatively low (<4 m s^{-1}) up to 4000 m above which all models show an increase in CRMSE due to the general increase in wind speeds (not shown). At Frankfurt, the lowest errors are found in the WRF-CMAQ model UK5 which also shows the highest correlations. This is likely an effect of the different simulation strategy with 1-day instead of 2-day simulation batches which allows UK5 to stay closer to the observations. A similar argument applies to the model US6 at Toronto, which performs best in terms of CRMSE and correlations most likely due to the weak nudging at upper levels. Similar to the profiles of temperature presented in Fig. 5, the models DE3 at Frankfurt and CA2/CA2f at Toronto exhibit lower performance as compared to the other models in particular in the upper troposphere. Over NA, ES1 is performing poorly in terms of CRMSE and correlation pointing to a problem with the meteorology driving this

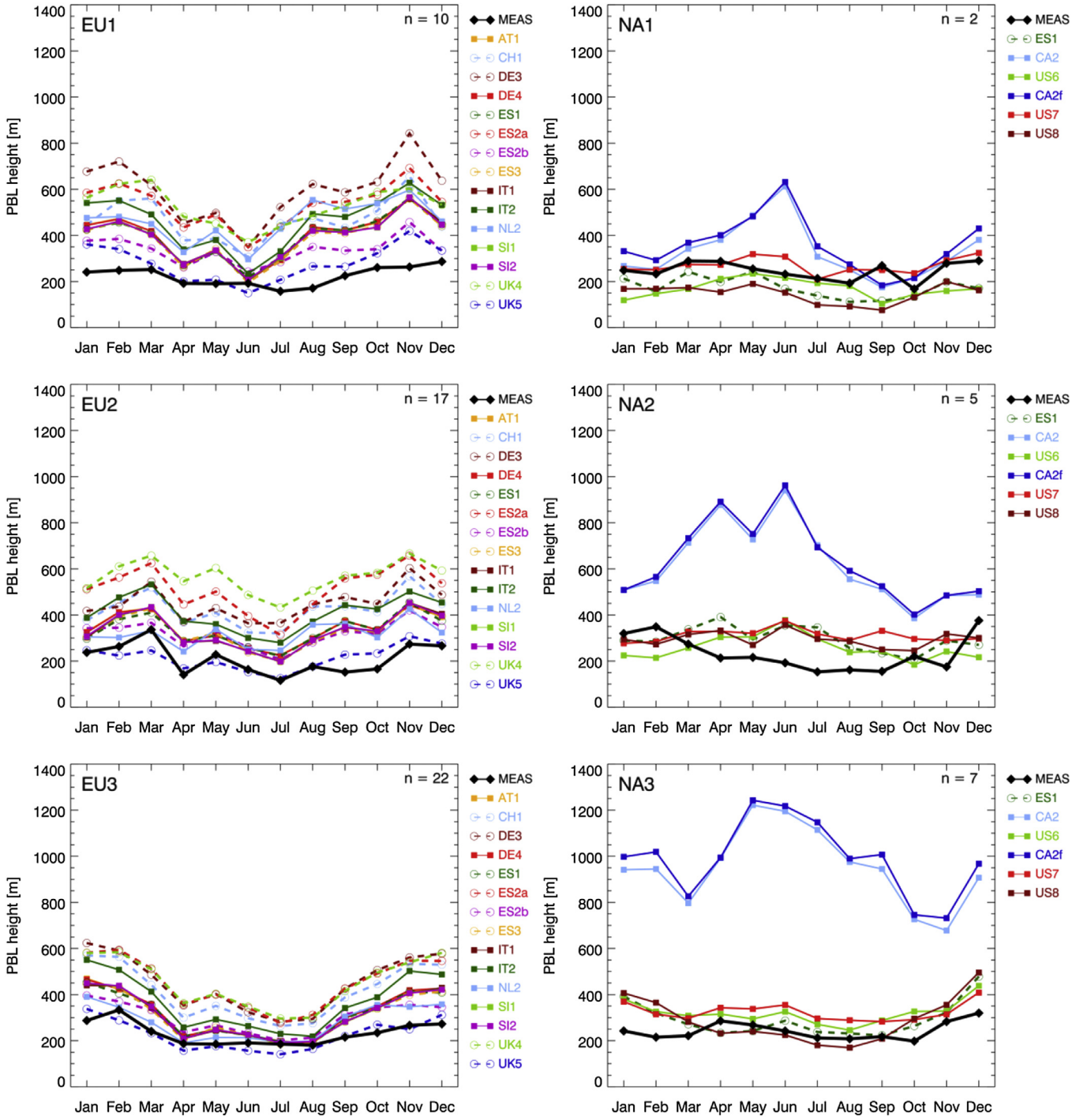

Fig. 9. Seasonal cycle of monthly mean nighttime (00 UTC) PBL heights in domains EU1–EU3 (left column) and early morning (12 UTC) PBL heights in NA1-N3 (right). Measurements are in black. The number of radiosonde sites available per domain is indicated in the top right corner of each panel.

model at the domain boundaries over NA.

In contrast to bias and CRMSE, correlations increase from low values at the surface to remain fairly constant at 0.8 to 0.9 above 100 m. The correlations near the surface are much weaker than seen for the receptors, probably related to a reduced accuracy of the aircraft wind measurements during take-off and landing manoeuvers (Benjamin et al., 1999). The large biases below 1000 m seen for all models both at Frankfurt and Toronto airport are likely a result of errors in the aircraft measurements as well.

The reasons for these differences and impacts on forecast skill for air pollutants (especially under low wind speed conditions) deserve further attention in the future. Some of the differences may be in the way 10-m wind speeds are diagnosed from model level variables and this would have little or no impact on transport but might have an impact on the primary emissions if the models use the diagnosed WS10 to drive emissions parameterizations. Modelers also need to be cautious of tuning their dust and sea salt emissions using observed atmospheric concentrations given the

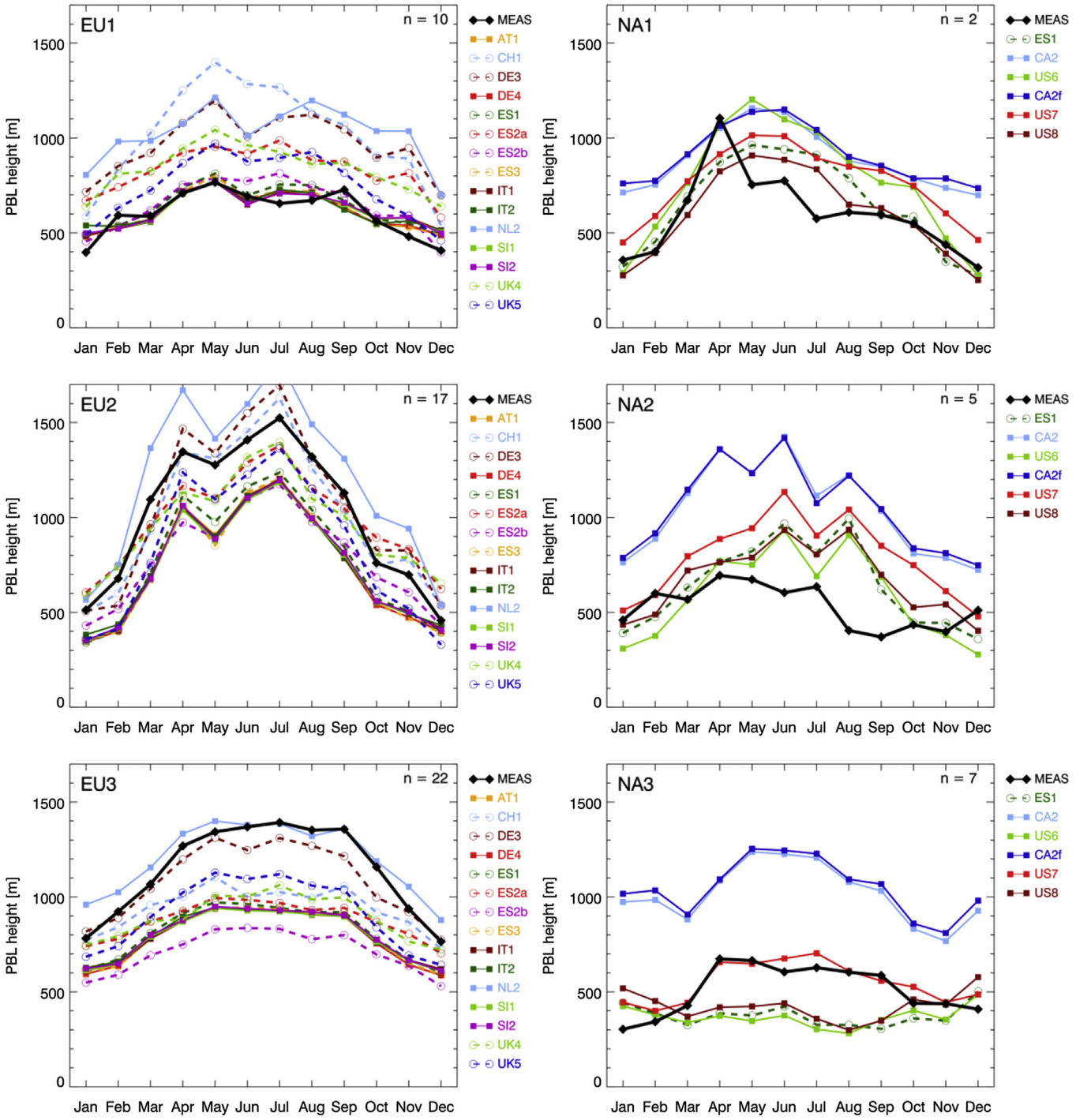

Fig. 10. Seasonal cycle of monthly mean afternoon (12 UTC) PBL heights in domains EU1–EU3 (left column) and evening (00 UTC) PBL heights in NA1-N3 (right). Measurements are in black. The number of radiosonde sites available per domain is indicated in the top right corner of each panel.

large biases in the modeled wind speeds.

5.3. Planetary boundary layer heights

Vertical mixing by atmospheric turbulence controls the dilution of air pollutants released at the surface into the vertical column and thereby critically determines near-surface concentrations. A useful diagnostic for the vertical mixing is the height of the planetary boundary layer (PBL) (or mixing height), which can be determined by a variety of methods mostly relying on vertical profiles of meteorological or chemical parameters measured in situ or by remote sensing (Seibert et al., 2000). Mainly driven by variations in solar radiation heating the Earth's surface and subsequently the atmospheric layers above, the PBL shows a pronounced evolution over the course of the day as well as over the year.

Figs. 9 and 10 present the seasonal evolution of monthly mean PBL heights separately for 00 UTC and 12 UTC due to the availability of radiosonde measurements at these synoptic times (see Sect. 3). For Europe this roughly corresponds to a midnight and a noon sounding. For North America this corresponds to a morning and an evening sounding.

Over Europe the dominant features are generally well reproduced by the models including (i) larger contrasts between day and night over the continental domain EU2 compared to domains EU1 and EU3 with maritime influence, (ii) much larger amplitudes of the seasonal cycle of the afternoon PBL heights over EU2 than over EU1 and EU3, (iii) opposing seasonal cycles between night and day with lower nighttime but higher daytime PBL heights in summer, and (iv) highest PBL heights over EU2 and lowest over EU1. The most striking difference between observations and models is the general overestimation of the nighttime PBL heights by all models except for UK5, which accurately tracks the seasonal evolution particularly in domains EU2 and EU3.

For daytime the models exhibit a much more variable performance over the different domains. The WRF-Chem models, for example, closely match the observed PBL heights over EU1 but are 20%–30% too low over the other domains. Over the continental domain EU2 the models UK4, ES2a, DE3 and CH1 have the smallest biases. NL2 somewhat overpredicts PBL heights in this domain. Much less consistent behavior is found for the two domains EU1 and EU3 with maritime influence. While most models overestimate PBL heights in EU1, they are all too low in EU3. A closer inspection revealed that the problem is related to the fact that in both domains the majority of radiosonde sites are located near the coast. The models are too coarse to represent the strong contrast between PBL heights over land and sea induced by the land-sea breeze circulation (McKendry, 1989). 14 out of 22 sites in domain EU3 are near the coast, and when limiting the analysis to these sites the discrepancy between models and observations becomes even larger: The measurements show a strong seasonal cycle with PBL heights in summer (~1600 m) about double as high as in winter (~800 m) while the models show almost no seasonal variation and a strong underestimation in summer. For these mixed land-sea grid cells the models appear to simulate a PBL that is more representative of a maritime environment than of the convective PBL over land encountered by the radiosondes. Also in EU1 the majority of sites are located near the coast, but here the effect appears to be almost opposite to EU3. A possible explanation may be that the land-sea breeze systems are likely stronger in the Mediterranean domain EU3 than along the Atlantic coasts in EU1. Limiting the analysis to sites further inland would leave only few sites for the analysis, especially in domain EU1. In domain EU3 there are 8 sites remaining at which the models tend to overestimate the summertime PBL heights, in strong contrast to the 14 coastal sites.

This analysis indicates that the results for domains EU1 and EU3 are severely impeded by the comparatively low horizontal resolution of the simulations and the problematic placement of many radiosonde sites near coastlines. It further suggests that much higher resolution models are needed to simulate air quality in coastal areas (Chemel and Sokhi, 2012) and that model evaluations with air pollution monitors in these areas are problematic. Due to the underestimation of daytime PBL heights in EU3, the models will tend to overestimate the concentrations of primary pollutants near the coast.

Table 6 presents the performance statistics with respect to daily 00 UTC (night) and 12 UTC (midday) soundings in the continental domain EU2 which contains the largest number of radiosonde sites suitable for the evaluation. Overall, the models successfully reproduce the seasonal and day-to-day variability in PBL heights at the individual sites during daytime as indicated by correlation coefficients between 0.4 and 0.5. Much lower correlations between 0.2 and 0.3 are found for nighttime and CRMSE are as large as the mean values. This rather poor performance is most likely not only due to model errors but also due to the low vertical resolution of the radiosonde data over Europe. In addition, there are significant biases in nighttime PBL heights mostly in the range of 100–300 m (50–150% in relative terms). Unfortunately, the study of Seidel et al. (2012) does not reveal whether the low resolution is not only producing enhanced scatter but also a bias in shallow PBL heights.

The models ES2a and ES2b only differ in vertical resolution. The overestimation of the nighttime PBL height is much more severe in ES2a which has only half the number of vertical levels as compared to ES2b. A low vertical resolution in the model simulation thus seems to produce higher rather than lower PBL heights, indicating that the found overestimation by the models can not only be ascribed to the low resolution of the radiosonde data.

Over North America the picture is less coherent than over Europe most probably due to the more problematic timing of the soundings during transition periods of PBL evolution in the morning and evening hours. Furthermore, the two models CA2 and CA2f show a distinctly different behavior than the other models overestimating PBL heights significantly at the selected times of the day. The NA domains are less affected by the problem of sites being located along coastlines but they suffer from relatively poor statistics, particularly in domain NA1 where only two sites are available. All models are overestimating both the morning and evening PBL heights in domain NA2 during summer, but most models show a reasonably good performance in other domains and other times of the year.

Performance metrics for daily PBL heights at 12 UTC and 00 UTC are shown in Table 7 for domain NA3. This domain has the largest number of radiosonde sites and with its continental climate is most similar to EU2 for which statistics are presented in Table 6. Despite the large bias, model CA2/CA2f shows similar performance in terms of correlation as the other models suggesting that day-to-day variability is well captured. Correlations for both the morning and evening PBL heights over NA3 are mostly higher than the correlations over the European domain EU2 at night but lower than the correlations over EU2 during the day. Interestingly, the models fall into two groups with respect to their behavior for evening PBL heights: While CA2/CA2f and US7 have a positive bias and rather high correlation coefficients (0.36–0.43), the models ES1, US6 and US8 have negative biases and very low correlations between 0.13 and 0.17. The latter models appear to predict a too early collapse of the daytime PBL in the evening over NA3. Similar to the situation over EU at night, the early morning PBL heights are significantly overpredicted by all models. Although US6 is the same model with the same PBL scheme (ACM2) as UK5, US6 does not perform better than other models as UK5 did over Europe. This different behavior may be due to the different land surface schemes (Pleim-Xiu in US6

Table 6
Statistics of model performance for PBL height at the sites in domain EU2 for 12 UTC (daytime) and 00 UTC (nighttime).

Model	Mean	Stdev	r	MB	CRMSE	Mean	Stdev	r	MB	CRMSE
	PBL height EU2 (m), 12 UTC (17 stations)					PBL height EU2 (m), 00 UTC (17 stations)				
obs	1020.6	636.0	–	–	–	209.8	300.0	–	–	–
AT1	746.5	461.8	0.46	−274.1	588.5	340.5	302.1	0.26	130.7	365.6
CH1	1014.1	573.2	0.48	−6.5	619.9	422.4	376.4	0.26	212.7	416.2
DE3	1056.7	594.4	0.50	36.2	617.9	448.9	361.0	0.27	239.1	401.8
DE4	741.9	461.5	0.46	−278.6	588.2	340.0	302.0	0.27	130.2	364.7
ES1	779.3	462.6	0.49	−241.3	575.0	333.4	296.7	0.26	123.7	362.6
ES2a	985.0	480.1	0.42	−35.6	614.6	510.3	353.7	0.25	300.5	401.8
ES2b	785.8	425.8	0.47	−234.8	577.4	316.4	244.3	0.26	106.7	334.9
ES3	744.9	457.8	0.47	−275.6	583.6	330.1	297.9	0.26	120.3	362.5
IT1	751.9	467.1	0.45	−268.6	595.8	330.4	299.1	0.27	120.6	362.4
IT2	755.7	455.1	0.45	−264.9	592.9	408.9	338.9	0.26	199.1	389.0
NL2	1210.8	615.6	0.51	190.2	620.3	315.1	305.8	0.24	105.3	372.8
SI1	743.8	462.1	0.45	−276.8	593.9	330.5	298.7	0.27	120.7	361.9
SI2	751.6	466.2	0.45	−267.0	594.7	330.8	298.7	0.27	121.1	361.9
UK4	964.8	504.2	0.38	−55.8	643.3	563.8	446.9	0.21	354.0	483.8
UK5	839.5	516.3	0.52	−181.1	571.5	214.7	204.1	0.24	4.9	318.8

Table 7
Statistics of model performance for PBL height at the sites in domain NA3 at 12 UTC (morning) and 00 UTC (evening).

Model	Mean	Stdev	r	MB	CRMSE	Mean	Stdev	r	MB	CRMSE
	PBL height NA3 (m), 12 UTC (7 stations)					PBL height NA3 (m), 00 UTC (7 stations)				
obs	244.2	131.0	–	–	–	510.5	201.2	–	–	–
CA2	950.0	386.4	0.24	705.7	376.9	1016.4	308.2	0.38	505.9	296.8
CA2f	987.3	372.8	0.23	743.0	364.9	1046.0	295.0	0.36	535.5	290.6
ES1	290.1	177.1	0.33	45.8	182.3	373.0	194.2	0.13	−137.5	260.5
US6	318.9	195.3	0.36	74.7	191.8	366.1	209.3	0.17	−144.4	264.2
US7	323.0	134.4	0.34	78.8	152.5	548.6	185.2	0.43	38.1	206.9
US8	286.1	191.7	0.39	41.9	185.1	423.2	213.8	0.13	−87.3	273.3

versus Noah in UK5; see Table 2) and different simulation strategies (weak nudging above PBL in US6 versus 1-day simulation batches in UK5) employed in the two models.

Fig. 11 presents the diurnal cycle of PBL heights in summer. Note that although measurements were taken only at 00 UTC and 12 UTC, the conversion to local time may produce measurements at several different hours each being represented by a different set of sites. Over Europe, the general pattern of highest amplitudes in domain EU2 and lowest in EU1 is followed by all models. There are significant time shifts in the diurnal evolution, e.g. UK5 simulating an earlier collapse in the evening as compared to other models. UK4 has a higher nighttime PBL not only compared to the measurements but also compared to the other models and exhibits a comparatively slow decrease in the evening. NL2 has a particularly broad daytime peak and tends to develop the highest PBL heights of all models. Most models, in particular the WRF-Chem models underestimate the height of the afternoon PBL. The two COSMO-based models perform well with respect to the afternoon PBL but are too high at night.

Over North America the models CA2 and CA2f show again a different behavior than the other models with a much broader daytime peak. The significant overestimation of both the morning and evening PBL heights indicates that the diurnal profile is indeed too broad in CA2/CA2f. The two WRF-Chem models US7 and US8 show a similar behavior in the morning but US7 exhibits a much later collapse of the convective PBL in the evening by more than 1 h. US6 has the highest daytime PBL over domain NA1, which is also much higher than the observations in the evening.

An evaluation of PBL heights and of its diurnal evolution is crucial due to the fundamental importance of vertical mixing for the dilution of air pollutants emitted at the surface. However, this evaluation is severely impeded both by the low vertical resolution of the radiosonde profiles (particularly critical at night) as well as by the fact that PBL heights are diagnosed in different ways by the individual models. We therefore recommend that in future model evaluations PBL heights should be diagnosed from vertical profiles of wind and temperature applying the same approach to all models rather than relying on PBL heights reported by the models individually. This would also allow applying the same method to the models as to the observed profiles. Furthermore, in addition to evaluating PBL heights which are only an indirect measure of vertical mixing it would be useful to include an idealized, Radon-like tracer with a fixed emission rate at the surface and a prescribed atmospheric lifetime as has been done in other model evaluation studies (Brunner et al., 2005).

5.4. Precipitation

Wash-out of water soluble species by precipitation is an important sink of pollutants from the atmosphere. Pollutants are scavenged in-cloud by cloud droplets growing to sizes large enough to form precipitation, or below-cloud by precipitation falling through layers of air below the cloud. We use here the set of available precipitation observations from the E-OBS database described in Sect. 3 to evaluate European domain results. Fig. 12 shows the accumulated monthly precipitation [cm] averaged over all the stations located in the EU1, EU2 and EU3 domains where E-OBS measurements are available. The monthly variability of the accumulated precipitation is very well captured by the models in the three domains. In the EU1 domain, all the models reproduce the peak precipitation in August and the minimum in April. The variability among models is around 4 cm for most of the months but increases to 9 cm in August when the ES3 model reports an accumulated precipitation of 18 cm and CH1 of 9 cm. In general, the ES3

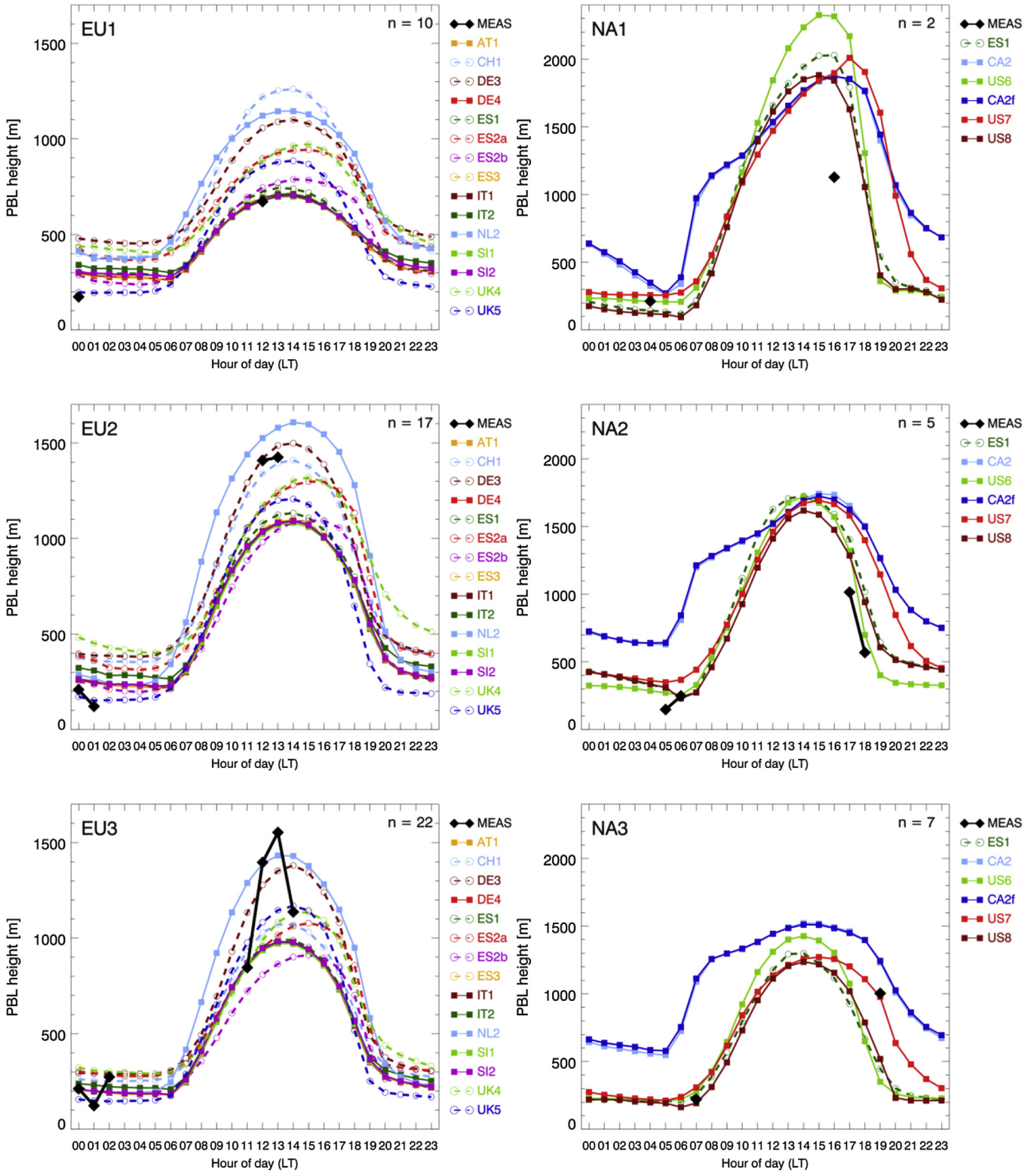

Fig. 11. Annual mean diurnal cycle of PBL heights in summer (June–August) in European domains EU1–EU3 (left column) and North American domains NA1-NA3 (right). Measurements are shown in black. The number of sites available per subdomain is indicated in the top right corner of each panel.

model is producing the largest amounts of precipitation for the whole year while ES2a/b, CH1 and UK4 are among the models with the lowest accumulated precipitation. The WRF-Chem models show a very similar evolution but ES3 simulates higher precipitation than all other versions. In the EU2 domain, all the models very closely trace the observed accumulated monthly precipitation over the year but generally with a small underestimation. The month of August as well as the winter months show the largest variability among models. The monthly evolution of accumulated precipitation is distinctly different in the Mediterranean domain EU3

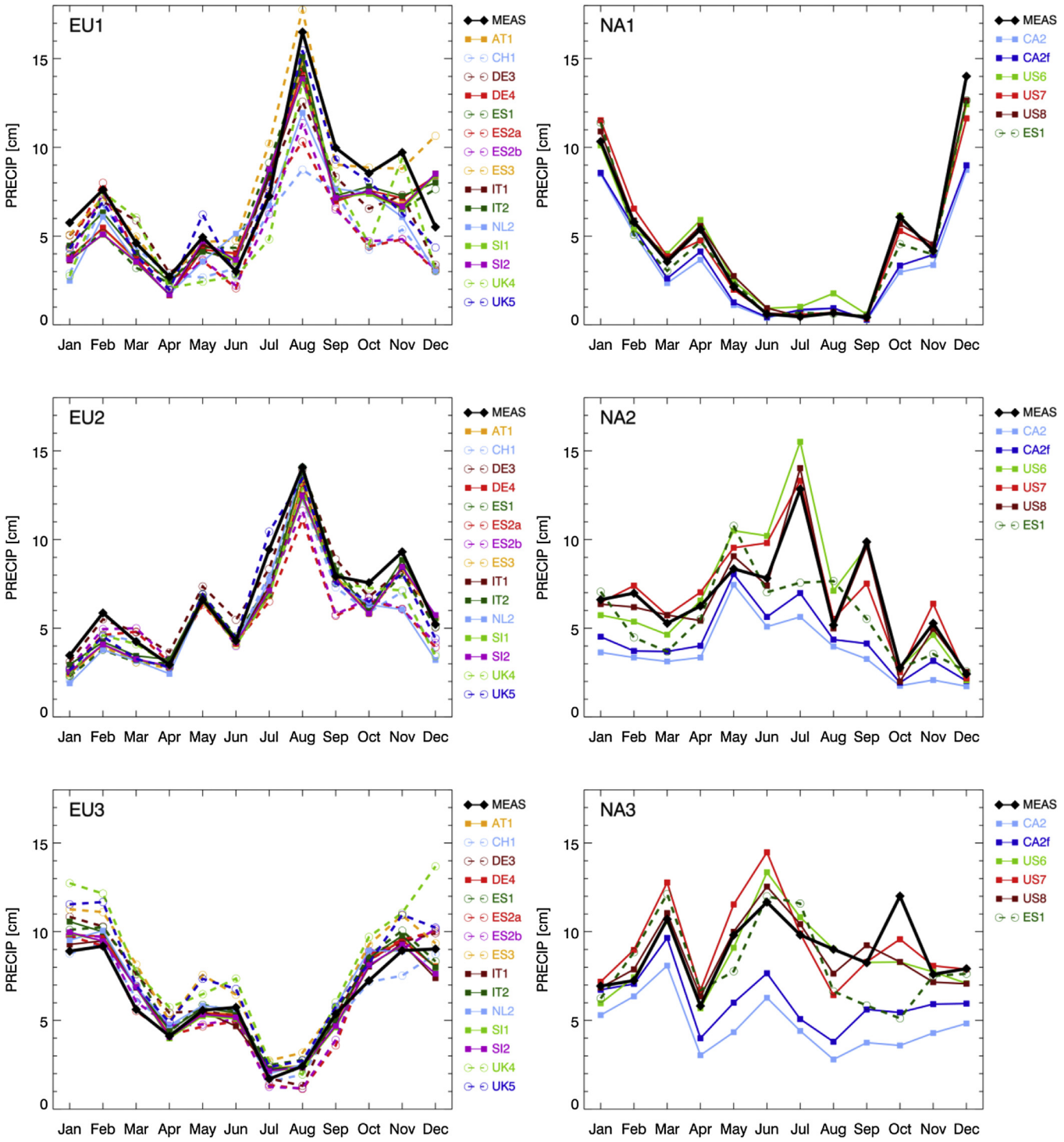

Fig. 12. Mean seasonal cycle of monthly accumulated precipitation in European domains EU1–EU3 (left column) and North American domains NA1-NA3 (right column). Measured values are shown in black.

compared to the other two domains. In EU3, the month with the lowest accumulated precipitation is July with less than 3 cm. Again, all the models closely follow the trend of the measurements, but most of the models slightly overestimate precipitation, especially UK4 during winter. The spread between models remains within a range of 2–3 cm most of the time. The largest differences are found in January, February, November and December.

As with the European domains, monthly precipitation is mostly well simulated by the six models across the three North American domains, as seen in Fig. 12. The variability among models is by far the smallest in the arid NA1 domain. The extremely dry months of June through September (precipitation < 1 cm) are well reproduced by the models. The rapid increase in observed precipitation during fall and winter is captured by each model, though CA2 and CA2f tend to underpredict the increase. As spring arrives, each model simulates the general decrease in precipitation nicely, even

capturing the momentary increase that occurs in April. The performance among the models is much more variable and somewhat poorer in the climatologically wetter NA2 and NA3 subdomains. For NA2, the three US models generally capture the seasonal evolution, indeed even the rapid monthly variability from the wettest month (July) to the driest (October). The ES1 model generally underpredicts, with a few monthly exceptions, while CA2 and CA2f greatly underpredicts except for October and December. The same general model trends are seen in NA3, with the ES1 and each of the US models performing well (with one exception) throughout the year. That exception occurs in October when each of these models miss the transient maximum that occurs in October. In fact, ES1 simulates a local minimum. As with the NA2 domain, CA2 and CA2f both greatly underpredict the magnitude of precipitation, though they do capture the monthly variability better than in NA2.

In summary, there is generally a large consistency among the models and a good agreement with observations. Differences in removal of soluble species between the models is therefore unlikely to be dominated by total precipitation rates but rather by the details of the wet scavenging schemes which may vary considerably between models (Knote and Brunner, 2013). Model CA2/CA2f strongly underestimates precipitation in domains NA2 and NA3. The version with aerosol feedbacks (CA2f) performs significantly better but the improvement is not sufficient to remove the overall low bias of this model. Interestingly, model CA2/CA2f accurately simulates the incoming shortwave radiation (see Sect. 5.5) suggesting that the underprediction of precipitation cannot simply be explained by a bias in cloudiness. The clouds simulated by this model appear to produce too little precipitation but accounting for aerosol direct and indirect effects somewhat improves this deficiency.

Precipitation rates depend on many factors including the influence of lateral boundary conditions (Warner et al., 1997), soil moisture initialization (Moufouma-Okia and Rowell, 2010; Van Weverberg et al., 2010), the treatment of the land surface and soil hydrology (Froidevaux et al., 2013), or the treatment of cloud microphysics (Van Weverberg et al., 2010). The influence of lateral boundary conditions is expected to be largest in regions close to the domain boundaries such as domain EU1. Over Europe, the models indeed show largest scatter in EU1 and the significant differences between the two COSMO models CH1 and DE3 may be due to the fact that they are forced by different global models. The large model-to-model differences over the North American domains NA2 and NA3, on the other hand, are unlikely related to the lateral boundary conditions as these would be expected to influence domain NA1 the strongest. Other factors such as the treatment of the land surface or of cloud microphysics must be dominating here.

5.5. Radiation

Solar radiation is the main energy source that drives all atmospheric processes. It also plays a key role in atmospheric chemistry by its ability to photodissociate a range of chemical species. The amount of incoming shortwave radiation reaching the surface (SWGD, also referred to as global radiation) provides information on the scattering and absorption of radiation by clouds, aerosols and gases in the atmosphere. Most of the models participating in the AQMEII phase 2 exercise accounted for the direct effect of aerosols on radiation and some also for the radiative effects of online simulated ozone. SWGD is thus directly linked to the chemistry in these models although differences between models will likely be dominated by the effects of clouds.

Fig. 13 shows the mean seasonal cycle of SWGD at three representative European stations of the BSRN solar radiation network and three stations of the North American SURFRAD network.

The seasonal cycle of SWGD is well reproduced by most of the European models at the station Palaiseau (EU1 domain). Models ES2a and ES2b show a systematic overestimation likely attributed to the lack of radiative effects of anthropogenic aerosols considered in that model. On the other side, the two COSMO models CH1 and DE3 significantly underestimate SWGD in the months of July and August but not in the other months. Most models show some overestimation during winter. At the station Payerne (EU2 domain) three groups of models can be identified, those that overpredict all year round (ES2ab, ES1, UK5), a cluster of models that nicely reproduce the observations (AT1, IT1, DE4, SI1, IT2) and a last group underestimating SWGD particularly during summer (UK4, DE3, CH1). Deviations from the monthly mean observations can be as large as 10–30%. Similar to domain EU1, the NL2 model overpredicts SWGD in the first half of the year but more closely traces the observations in the remaining months. A notable feature in domain EU2 is the systematic overestimation of SWGD in January and February by all models. At the Carpentras station (EU3 domain) with its Mediterranean climate with a dry summer season, most of the models tend to significantly overestimate the monthly accumulated radiation. The UK4 model closely follows the observations during the whole year, while the two COSMO models CH1 and DE3 are somewhat too low. The largest overestimation by a large group of models including all WRF versions, ES2a/b and NL2, is detected from May to July. The fact that the results of the models SI1 and SI2 with and without considering direct radiative effects differ by no more than 3–4% suggests that the model deficiency is rather due too low cloud cover than to too low aerosol. Also the differences between model IT1 with no aerosol effects and the models AT1, IT2 and DE4 with both direct and indirect effects are much smaller than the deviations from the observed values. Uncertainties in cloud development due to a variety of factors including PBL and convection parameterizations or the treatment of land surface and soil hydrology, thus appear to play a more important role than the effects of aerosols on radiation through direct and indirect interactions.

The seasonal cycles of SWGD are also well simulated by the models at the three NA sites. The SURFRAD site representing NA1, Desert Rock, NV is a high elevation, extremely arid location dominated by continental air masses, resulting in generally clear skies throughout most of the year. This lack of cloud cover lends itself to excellent model performance as seen in Fig. 13, where each model replicates the smooth annual cycle well. The largest discrepancy is found during the period of May through July, where several models (US6, US7 and ES1) slightly overpredict SWGD. NA2 is represented by the SURFRAD site located in Goodwin Creek, MS, which is dominated by maritime tropical air masses originating over the Gulf of Mexico, hence cloud cover is prevalent throughout most of the year. Stratiform clouds, which meteorological models historically simulate well, dominate during the cooler months in this region. This generalization is well supported in this analysis, as each model captures the monthly SWGD quite well from the months of October through March. Convective clouds, conversely, are much more difficult to simulate, as subgrid scale processes dominate their development. Accordingly, model performance is somewhat more variable and slightly degraded during the warmer months at Goodwin Creek. The models performing the worst during this time period include ES1 and US6, which generally overpredict SWGD, especially during June, July and August. This likely indicates an underprediction of convective clouds by these models and the associated attenuation of radiation. The third domain, NA3, is represented by an SURFRAD station located at Penn State University in PA. This site, while moist like Goodwin Creek, experiences a dominance of stratiform clouds throughout most of the year. Accordingly model performance is generally good, though slightly

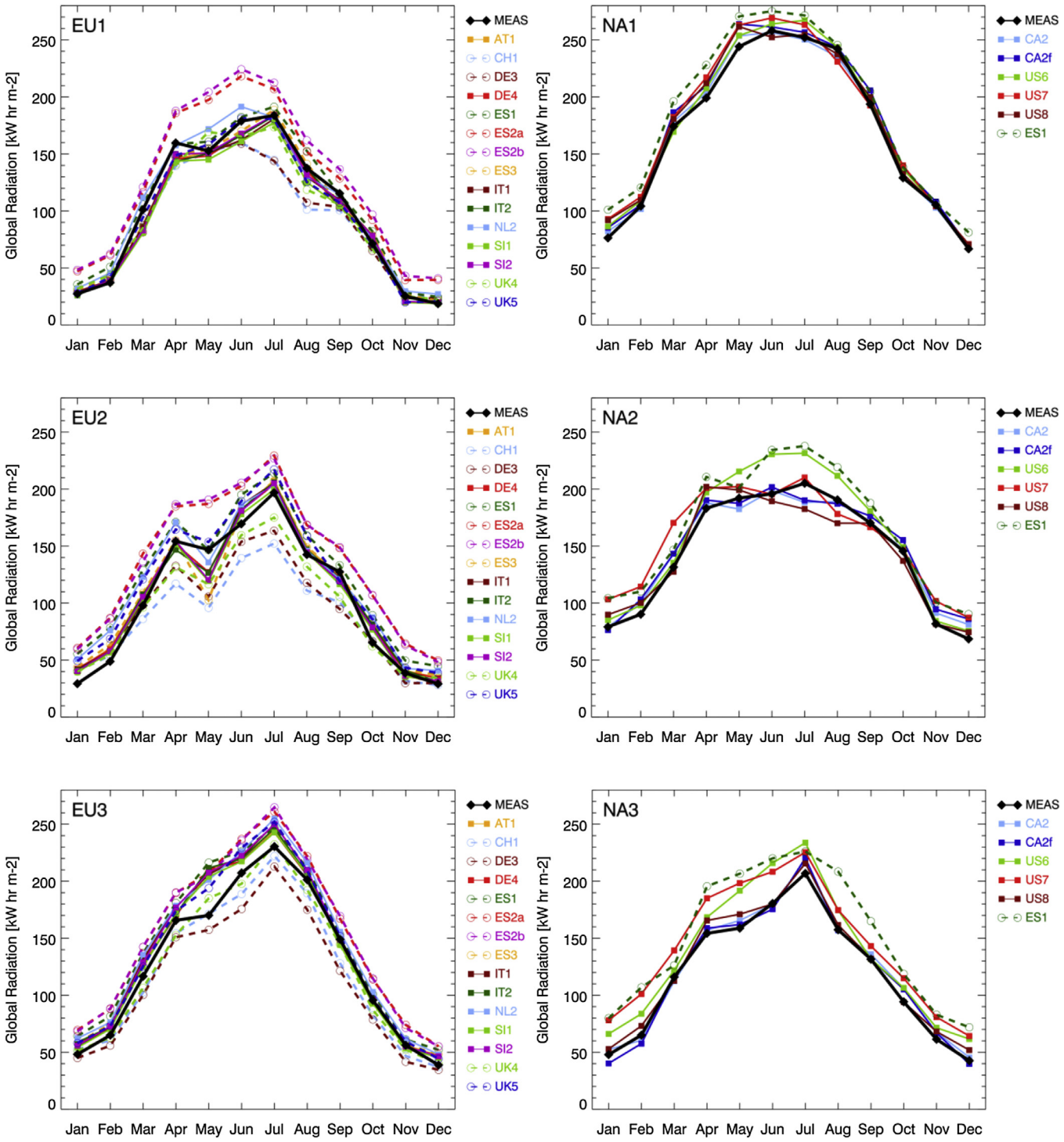

Fig. 13. Mean seasonal cycle of monthly accumulated incoming shortwave radiation (global radiation) at the surface at representative BSRN stations in each domain (Palaiseau, France (EU1); Payerne, Switzerland (EU2); Carpentras, France; (EU3); Desert Rock, NV (NA1); Goodwin Creek, MS (NA2); Penn State, PA (NA3)). Measured values in black.

high. The worst performing models include US6, US7 and ES1, especially during the months of April, May and June. The two WRF-Chem models ES1 and US7 consistently overpredict SWGD in NA3 throughout the year, while the third WRF-Chem model, US8, closely follows the observations. The reasons why US8 performs better than ES1 and US7 remain unclear, in particular since US8 and ES1 have very similar configurations while US7 uses a different PBL scheme (see Table 1). Further investigations will be needed to elucidate these results.

The models' ability to simulate the annual mean diurnal cycle was also examined as shown in Fig. 14. At the European sites there is an obvious1 hour time-shift between the models and the observations suggesting a difference in the way the hourly values were computed or reported. While most models report instantaneous values at the end of the hour, some models (CH1, NL2) report true hourly mean values (mean over the hour before) computed

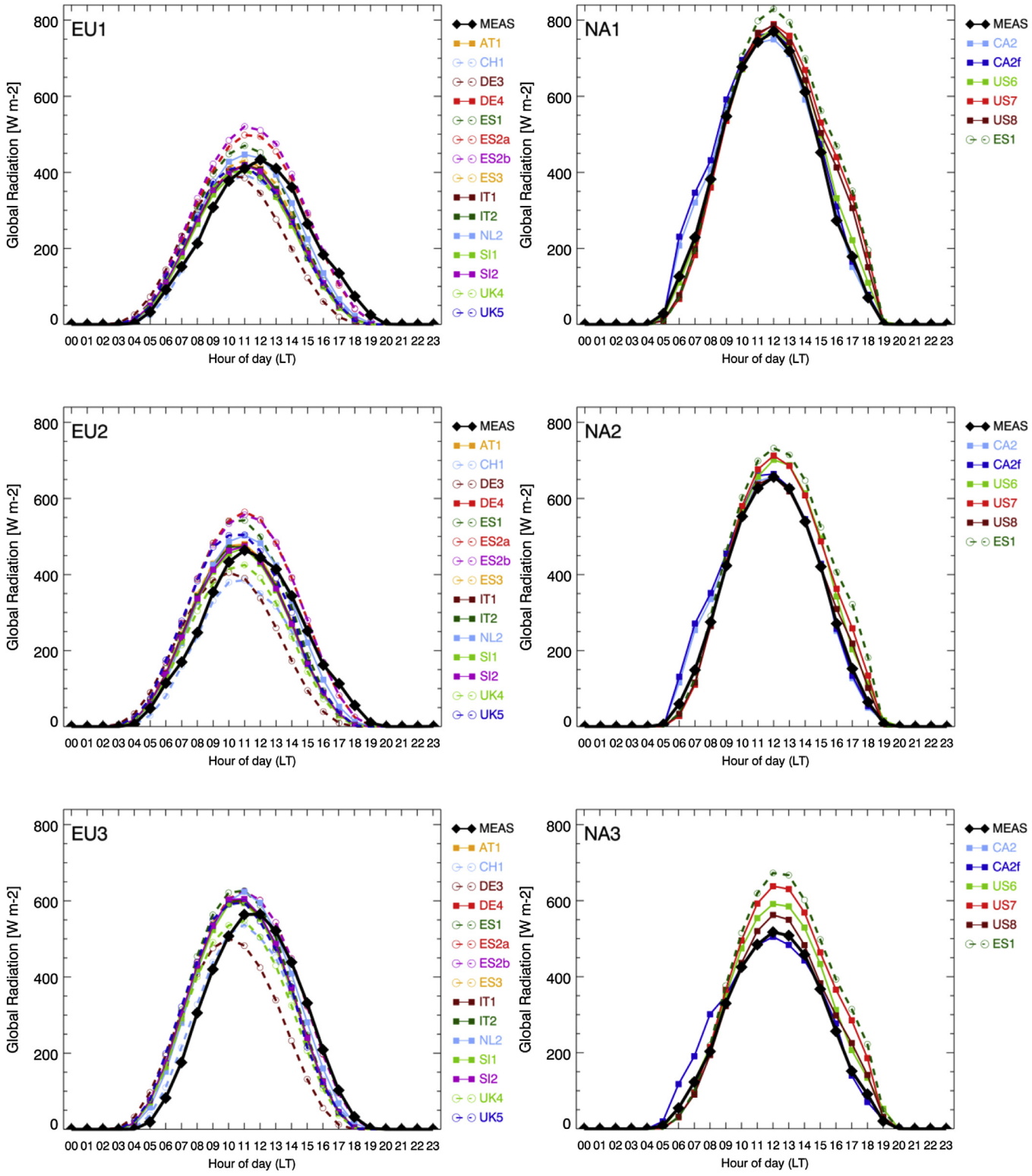

Fig. 14. Annual mean daily cycle of incoming shortwave radiation (global radiation) at the surface at three BSRN stations (Palaiseau, France (EU1); Payerne Switzerland (EU2); Carpentras, France (EU3); Desert Rock, NV (NA1); Goodwin Creek, MS (NA2); Penn State, PA (NA3)) located in EU1, EU2 and EU3 domain. Measured values in black.

from accumulated radiation output, which introduces a 30 min time-shift between model results.

At Palaiseau (EU1 domain) there is a group of models that nicely reproduce the solar radiation hourly cycle (apart from the 1-h shift) with a maximum peak value of about 400 W m^{-2} (AT1, DE4, IT1, SI1/2, NL2). Models ES2ab present the largest overestimation consistent with Fig. 13. At the Payerne station (EU2 domain) the models show a similar behavior with a group of models providing a good estimate of the SWGD daily evolution (AT1, DE4, IT1, SI1/2). DE3, UK4, CH1 are underestimating the peak value by 40 W m^{-2}

while NL2 and UK5 are slightly overpredicting by 20–30 W m^{-2}. ES2a/b are again significantly too high. The WRF-Chem model ES1 has a consistently higher diurnal peak value than the other WRF-Chem models and is significantly overpredicting SWGD in domains NA2 and NA3, similar to ES2a/b. Finally, at the Carpentras station (EU3 domain) most of the models underestimate the maximum peak value around 600 W m^{-2}. The models UK4 and CH1 are the models being able to fit better the observations, while DE3 shows a significant underestimation during summertime. Overall, differences between the models seems to be more relevant regarding cloud development than on aerosol radiation interactions.

Over NA the model's performance is generally good at each of the climatologically disparate SURFRAD locations. There is a slight tendency for each of the models to overpredict the observations, though at different times of the day. Most models (except CA2 and CA2f) overpredict SWGD, especially around local noon and extending into the early afternoon. This may be tied to the issues associated with convective cloud formation mentioned earlier, which peaks during this time frame. Model performance during the morning hours, from sunrise to roughly 10 LT is excellent, with only CA2 and CA2f slightly high for NA1 and NA2 and CA2f slightly high for NA3.

6. Summary and conclusions

This study was devoted to a collective operational evaluation of the meteorology simulated by the coupled chemistry and meteorology models applied in phase 2 of the AQMEII project. As opposed to phase 1 where only offline models were considered, AQMEII-2 focused on the evaluation of the new generation of regional-scale online integrated and online access models which have been developed in Europe and North America over the past approximately ten years (Baklanov et al., 2014). The study complements the collective analyses of Im et al. (2015a,b) and Giordano et al. (2015) which were dedicated to the evaluation of ozone, particular matter, and the influence of chemical boundary conditions, respectively. The meteorological parameters considered in the evaluation are all critically influencing the chemistry: Temperature affects chemical reaction rates, the gas – particle phase partitioning and biogenic emissions. Wind speed influences the volume of air into which emissions are diluted and the transport time between emissions and downwind receptor locations. PBL height is a key measure of turbulent mixing and the corresponding dilution of pollutants in the vertical. Radiation directly affects photochemistry through the photolysis of gases like O_3 and NO_2. Precipitation removes water soluble trace gases and aerosols through wet deposition.

The analysis of these meteorological quantities contributes to the understanding of differences in the chemistry modeling although the connection is rarely straightforward due to multiple factors acting simultaneously.

Consistent with the results of phase 1 presented by Vautard et al. (2012) we found a significant overprediction of 10-m wind speeds by most models, especially the WRF-based models, more pronounced for stable nighttime than for convective daytime conditions. This is expected to lead to too strong dilution of air pollutants over urban source regions but in turn to too rapid transport to rural areas downwind. This effect likely contributes to the frequently reported underprediction of primary air pollutants at urban sites, which is usually attributed to insufficient model resolution.

2-m temperatures were simulated quite accurately by all models with monthly mean biases in the range of −2 K to +1.5 K over all European and North American subdomains investigated. Such small biases are expected to affect ozone concentrations by no more than a few ppb mostly through shifts in the thermal equilibrium of peroxyacetylnitrate (PAN) and changes in isoprene emissions (Sillman and Samson, 1995). Annual mean biases were negative over all domains in Europe and North America, and vertical temperature profiles revealed persistent negative deviations from MOZAIC aircraft measurements throughout the troposphere of the order of −2 K.

Planetary boundary layer heights diagnosed by the individual models were evaluated against PBL heights from radiosondes. The most striking result is a general and often strong (factor 2 and more) overestimation of the nocturnal PBL heights by all models except the WRF-CMAQ model UK5. The WRF-Chem models using the YSU planetary boundary layer scheme tended to underestimate afternoon PBL heights whereas other models showed both over- and underpredictions depending on region. The comparison of PBL heights suffered from a number of problems including different diagnostics applied in different models, the low vertical resolution of the radiosonde data over Europe, the problematic timing of the radiosonde launches over North America during phases of strong PBL evolution in the morning and evening, and the placement of many radiosonde sites over Europe near coastlines where they are influenced by land-sea breezes that cannot be reproduced by the models given their rather coarse resolution. For future model evaluations we therefore recommend to diagnose PBL heights directly from vertical model profiles applying the same approach to all models and the observations. Furthermore, an idealized Radon-like tracer would help diagnose model differences in terms of vertical mixing more directly than it is possible with an evaluation of PBL heights.

The seasonal cycle and month-to-month variability in solar incoming radiation was quite accurately captured by the models with a few exceptions. Over Europe, the model ES2/ES2f tended to overestimate radiation by up to 25% in some months whereas the two COSMO models CH1 and DE3 were frequently biased low by 10–20%. Over North America, the best model performance was obtained for the dry climate over the southeastern US (domain NA1). Solar incoming radiation was accurately simulated by several models also over the more continental climates of NA2 and NA3, whereas other models significantly overestimated the radiation by up to 30% in some months.

Aerosol direct and indirect effects appear to have only a minor influence on the overall radiation biases as suggested by comparing the results of the different WRF-Chem models applied over Europe with and without considering such effects, and the results of CA2 and CA2f over North America. Differences between the models rather seem to be dominated by differences in cloudiness which will depend on many factors such as the treatment of the land surface and soil hydrology, details of planetary boundary layer and convection parameterizations, or the chosen cloud microphysics scheme.

The biases in solar incoming radiation cannot easily be linked to the found biases in temperature. Over Europe, for example, the largest negative biases in 2-m temperatures were observed over the domains EU2 and EU3 where shortwave downward radiation was quite accurately captured or even overpredicted by some models. Over the North American domain NA3, radiation was significantly overestimated by model US7 which, on the other hand, showed one of the largest negative temperature biases over the same region.

Such large differences in solar incoming radiation between models and observations are expected to have a significant impact on the production of ozone and other photooxidants. The underestimation of solar incoming radiation in model CH1, for example, may contribute to the negative O_3 bias reported for this model by Im et al. (2015b).

The seasonal cycle of monthly accumulated precipitation and the large differences between different subdomains were well captured by most of the models. However, significant negative biases of up to 50% were found for model CA2/CA2f over the North American domains NA2 and NA3. Such biases are expected to lead to a proportional underestimation of wet removal of water soluble gases and aerosols.

An important conclusion that may be drawn from this study is that differences between different model systems were typically larger than differences between simulations with the same model including or excluding aerosol feedback effects. This will make it difficult to demonstrate any positive effect of considering feedbacks on numerical weather prediction since the result may depend significantly on the chosen model system. However, this conclusion does not hold for situations with high aerosol loads such as during the Russian forest fires as demonstrated by Kong et al. (2015) and Forkel et al. (2015). Furthermore, Forkel et al. (2015) demonstrated that aerosol indirect effects can be strong over clean areas such as the Atlantic Ocean which, however, was not covered by the sites included in the present evaluation study.

Acknowledgments

We gratefully acknowledge the support of the European groups through COST Action ES1004 EuMetChem. Individual authors of this article were additionally supported by the following projects and grants: Lea Giordano through Swiss State Secretariat for Education, Research and Innovation, project C11.0144. O. Jorba is supported by grant SEV-2011-00067 of Severo Ochoa Program awarded by the Spanish Government. The UPM authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the Centro de Supercomputación y Visualización de Madrid (CESVIMA) and the Spanish Supercomputing Network (BSC). Rahela Žabkar and Luka Honzak were supported by the Centre of Excellence for Space Sciences and Technologies SPACE-SI, which is partly financed by the EU, European Regional Development Fund and Republic of Slovenia, Ministry of Higher Education, Science, Sport and Culture. Y. Zhang acknowledges funding support from the NSF Earth System Program (AGS-1049200) and high-performance computing support from Yellowstone by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and Stampede, provided as an Extreme Science and Engineering Discovery Environment (XSEDE) digital service by the Texas Advanced Computing Center (TACC). The technical assistance of Bert van Ulft (KNMI) and Arjo Segers (TNO) in producing the results of the NL2-model (the RACMO2-LOTOS-EUROS coupled system) is gratefully acknowledged. The University of Hertfordshire acknowledges TRANSPHORM (FP7 project contract number 243406) and ClearFlo (NERC Funded project) for supporting the work on the application of WRF model over European and urban scales respectively. G. Curci and P. Tuccella were supported by the Italian Space Agency (ASI) in the frame of the project PRIMES (contract n. I/017/11/0). The UMU group acknowledges the funding from the project CGL2013-48491-R, Spanish Ministry of Economy and Competitiveness.

The authors would also like to thank numerous data providers: The weather centers ECMWF, NCEP and DWD for meteorological boundary conditions; the ECMW/MACC project & Météo-France/CNRM-GAME for chemical boundary conditions; the MOZAIC Data Centre and its contributing airlines for North American and European aircraft takeoff and landing vertical profiles. Data from meteorological station monitoring networks were provided by NOAA and Environment Canada (for the US and Canadian meteorological network data) and the National Center for Atmospheric Research (NCAR) data support section. The Joint Research Center Ispra/Institute for Environment and Sustainability provided its ENSEMBLE system for model output harmonization and analyses and evaluation. For precipitation data we acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). Solar radiation data were provided by the BSRN network (http://www.bsrn.awi.de/).

References

Alapaty, K., Mathur, R., Pleim, J., Hogrefe, C., Rao, S.T., Ramaswamy, V., Galmarini, S., Schaap, M., Makar, P., Vautard, R., Baklanov, A., Kallos, G., Vogel, B., Sokhi, R., 2012. New directions: understanding interactions of air quality and climate change at regional scales. Atmos. Environ. 49, 419–421.

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., Zhang, Y., 2014. Online coupled regional meteorology chemistry models in Europe: current status and prospects. Atmos. Chem. Phys. 14, 317–398.

Baldauf, M., Seifert, A., Forstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T., 2011. Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Mon. Weather Rev. 139, 3887–3905.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., Betts, A.K., 2009. A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the integrated forecast system. J. Hydrometeorol. 10, 623–643.

Bangert, M., Kottmeier, C., Vogel, B., Vogel, H., 2011. Regional scale effects of the aerosol cloud interaction simulated with an online coupled comprehensive chemistry model. Atmos. Chem. Phys. 11, 4411–4423.

Benjamin, S.G., Schwartz, B.E., Cole, R.E., 1999. Accuracy of ACARS wind and temperature observations determined by collocation. Weather Forecast. 14, 1032–1038.

Bianconi, R., Galmarini, S., Bellasio, R., 2004. Web-based system for decision support in case of emergency: ensemble modelling of long-range atmospheric dispersion of radionuclides. Environ. Modell. Softw. 19, 401–411.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., Shelly, A., 2012. Unified modeling and prediction of weather and climate: a 25-year journey. Bull. Am. Meteorol. Soc. 93, 1865–1877.

Brunner, D., Staehelin, J., Rogers, H.L., Köhler, M.O., Pyle, J.A., Hauglustaine, D.A., Jourdain, L., Berntsen, T.K., Gauss, M., Isaksen, I.S.A., Meijer, E., van Velthoven, P., Pitari, G., Mancini, E., Grewe, V., Sausen, R., 2005. An evaluation of the performance of chemistry transport models – Part 2: detailed comparison with two selected campaigns. Atmos. Chem. Phys. 5, 107–129.

Chemel, C., Sokhi, R., 2012. Response of London's urban heat island to a marine air intrusion in an easterly wind regime. Boundary-Layer Meteorol. 144, 65–81.

Cheng, W.Y.Y., Steenburgh, W.J., 2005. Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the Western United States. Weather Forecast. 20, 812–821.

Chou, M.-D., Suarez, M.J., 1994. An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 84.

Clough, S.A., Shephard, M.W., Mlawer, E.J., Delamere, J.S., Iacono, M.J., Cady-Pereira, K., Boukabara, S., Brown, P.D., 2005. Atmospheric radiative transfer modeling: a summary of the AER codes. J. Quant. Spectrosc. Radiat. Transf. 91, 233–244.

Curci, G., Hogrefe, C., Bianconi, R., Im, U., Balzarini, A., Baró, R., Brunner, D., Forkel, R., Giordano, L., Hirtl, M., Honzak, L., Jiménez-Guerrero, P., Knote, C., Langer, M., Makar, P.A., Pirovano, G., Pérez, J.L., San José, R., Syrakov, D., Tuccella, P., Werhahn, J., Wolke, R., Žabkar, R., Zhang, J., Galmarini, S., 2014. Uncertainties of simulated aerosol optical properties induced by assumptions on aerosol physical and chemical properties: an AQMEII-2 perspective. Atmos. Environ..

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., Irwin, J., Rao, S.T., Scheffe, R., Schere, K., Steyn, D., Venkatram, A., 2010. A framework for evaluating regional-scale numerical photochemical modeling systems. Environ. Fluid Mech. 10, 471–489.

Di Luzio, M., Johnson, G.L., Daly, C., Eischeid, J.K., Arnold, J.G., 2008. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. J. Appl. Meteor. Climatol. 47, 475–497.

Doms, G., Förstner, J., Heise, E., Herzog, H.-J., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz, J.-P., Vogel, G., 2011. A Description of the Nonhydrostatic Regional COSMO Model. Part II: Physical Parameterization.

Dudhia, J., 1989. Numerical study of convection observed during the Winter monsoon experiment using a mesoscale two-dimensional model. J. Atmos. Sci. 46, 3077–3107.

Durre, I., Yin, X., 2008. Enhanced radiosonde data for studies of vertical structure. Bull. Am. Meteorol. Soc. 89, 1257–1262.

Edwards, J.M., Slingo, A., 1996. Studies with a flexible new radiation code. I: choosing a configuration for a large-scale model. Q. J. R. Meteorol. Soc. 122, 689–719.

EEA, 2011. Air pollution by ozone across Europe during summer 2010. In: Agency, E.E., Office for Official Publications of the European, C (Ed.), EEA Technical Report. European Environment Agency, Copenhagen, Denmark, p. 42.
Essery, R., Clark, D.B., 2003. Developments in the MOSES 2 land-surface model for PILPS 2e. Glob. Planet. Change 38, 161–164.
Fereday, D.R., Maidens, A., Arribas, A., Scaife, A.A., Knight, J.R., 2012. Seasonal forecasts of northern hemisphere winter 2009/10. Environ. Res. Lett. 7, 034031.
Forkel, R., Balzarini, A., Baró, R., Curci, G., Jiménez-Guerrero, P., Hirtl, M., Honzak, L., Im, U., Lorenz, C., Pérez, J.L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Žabkar, R., 2015. Analysis of the WRF-Chem contributions to AQMEII phase2 with respect to aerosol radiative feedbacks on meteorology and pollutant distributions. Atmos. Environ. 115, 630–645.
Froidevaux, P., Schlemmer, L., Schmidli, J., Langhans, W., Schär, C., 2013. Influence of the background wind on the local soil moisture–precipitation feedback. J. Atmos. Sci. 71, 782–799.
Galmarini, S., Bianconi, R., Appel, W., Solazzo, E., Mosca, S., Grossi, P., Moran, M., Schere, K., Rao, S.T., 2012. ENSEMBLE and AMET: two systems and approaches to a harmonized, simplified and efficient facility for air quality models development and evaluation. Atmos. Environ. 53, 51–59.
Giordano, L., Brunner, D., Flemming, J., Im, U., Hogrefe, C., Bianconi, R., Badia, A., Balzarini, A., Baro, R., Chemel, C., Curci, G., Forkel, R., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J.J.P., Makar, P.A., Manders-Groot, A., Neal, L., Perez, J.L., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R.S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, K., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Galmarini, S., 2015. Assessment of the MACC reanalysis and its influence as chemical boundary conditions for regional air quality modeling in AQMEII-2. Atmos. Environ. 115, 371–388.
Giorgi, F., Bi, X., Qian, Y., 2003. Indirect vs. direct effects of anthropogenic sulfate on the climate of east Asia as simulated with a regional coupled climate-chemistry/aerosol model. Clim. Change 58, 345–376.
Grasselt, R., Schüttemeyer, D., Warrach-Sagi, K., Ament, F., Simmer, C., 2008. Validation of TERRA-ML with discharge measurements. Meteorol. Z. 17, 763–773.
Gregory, D., Rowntree, P.R., 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. Mon. Weather Rev. 118, 1483–1506.
Grell, G.A., Dévényi, D., 2002. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophys. Res. Lett. 29, 38-31–38-34.
Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J. Geophys. Res. Atmos. 113, D20119.
Helmert, J., Heinold, B., Tegen, I., Hellmuth, O., Wendisch, M., 2007. On the direct and semidirect effects of Saharan dust over Europe: a modeling study. J. Geophys. Res. Atmos. 112, D13208.
Hogrefe, C., Pouliot, G., Wong, D., Alfreida, T., Roselle, S., Pleim, J., Mathur, R., 2015. Annual application and evaluation of the online coupled WRF-CMAQ system over North America under AQMEII phase 2. Atmos. Environ. 115, 683–694.
Hollingsworth, A., Engelen, R.J., Benedetti, A., Dethof, A., Flemming, J., Kaiser, J.W., Morcrette, J.J., Simmons, A.J., Textor, C., Boucher, O., Chevallier, F., Rayner, P., Elbern, H., Eskes, H., Granier, C., Peuch, V.H., Rouil, L., Schultz, M.G., 2008. Toward a monitoring and forecasting system for atmospheric composition: the GEMS project. Bull. Am. Meteorol. Soc. 89, 1147–1164.
Hong, S.-Y., 2010. A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. Q. J. R. Meteorol. Soc. 136, 1481–1496.
Hong, S.-Y., Lim, J.-O.J., 2006. The WRF single-moment 6-class microphysics scheme (WSM6). J. Korean Meteor. Soc. 42, 129–151.
Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Weather Rev. 134, 2318–2341.
Huijnen, V., Eskes, H.J., Poupkou, A., Elbern, H., Boersma, K.F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.H., Zerefos, C., 2010. Comparison of OMI NO2 tropospheric columns with an ensemble of global and European regional air quality models. Atmos. Chem. Phys. 10, 3273–3296.
Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D., 2008. Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. J. Geophys. Res. Atmos. 113, D13103.
Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., Denier van der Gon, H., 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., San Jose, R., 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., Galmarini, S., 2015a. 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. Atmos. Environ. 115, 421–441.
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.P., Makar, P.A., Manders-Groot, A., Neal, L., Pérez, J.L., Pirovano, G., Pouliot, G., San Jose, R., 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., Galmarini, S., 2015b. 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. Atmos. Environ. 115, 404–420.
Janjic, Z., Gall, R.L., 2012. Scientific Documentation of the NCEP Nonhydrostatic Multiscale Model on the B Grid (NMMB). Part 1 Dynamics (NCAR Technical Note).
Janjić, Z.I., 1994. The step-mountain Eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. Weather Rev. 122, 927–945.
Jiménez, P.A., Dudhia, J., 2013. On the ability of the WRF model to reproduce the surface wind direction over complex terrain. J. Appl. Meteor. Climatol. 52, 1610–1617.
JRC, 2011. Forest fires in Europe 2010. In: Centre, E.C.J.R. (Ed.), Joint Research Centre (JRC). Publications Office of the European Union, Luxembourg.
Kain, J.S., 2004. The Kain–Fritsch convective parameterization: an update. J. Appl. Meteorol. 43, 170–181.
Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V., Petrovic, P., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int. J. Clim. 22, 1441–1453.
Knote, C., Brunner, D., 2013. An advanced scheme for wet scavenging and liquid-phase chemistry in a regional online-coupled chemistry transport model. Atmos. Chem. Phys. 13, 1177–1192.
Kong, X., Forkel, R., Sokhi, R.S., Suppan, P., Baklanov, A., Gauss, M., Brunner, D., Barò, R., Balzarini, A., Chemel, C., Curci, G., Jiménez-Guerrero, P., Hirtl, M., Honzak, L., Im, U., Pérez, J.L., Pirovano, G., San Jose, R., Schlünzen, K.H., Tsegas, G., Tuccella, P., Werhahn, J., Žabkar, R., Galmarini, S., 2014. Analysis of meteorology–chemistry interactions during air pollution episodes using online coupled models within AQMEII phase-2. Atmos. Environ. 115, 527–540.
Konovalov, I.B., Beekmann, M., Kuznetsova, I.N., Yurova, A., Zvyagintsev, A.M., 2011. Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region. Atmos. Chem. Phys. 11, 10031–10056.
Lenderink, G., Holtslag, A.A.M., 2004. An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. Q. J. R. Meteorol. Soc. 130, 3405–3427.
Li, J., Barker, H.W., 2005. A radiation algorithm with correlated-k distribution. Part I: local Thermal equilibrium. J. Atmos. Sci. 62, 286–309.
Lin, Y.-L., Farley, R.D., Orville, H.D., 1983. Bulk parameterization of the snow field in a cloud model. J. Clim. Appl. Meteorol. 22, 1065–1092.
Lock, A.P., Brown, A.R., Bush, M.R., Martin, G.M., Smith, R.N.B., 2000. A new boundary layer mixing scheme. Part I: scheme description and single-column model tests. Mon. Weather Rev. 128, 3187–3199.
Makar, P.A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Zakbar, R., Milbrandt, J., Im, U., Galmarini, S., Gravel, S., Zhang, J., Hou, A., Pabla, B., Cheung, P., Bianconi, R., 2015a. Feedbacks between Air Pollution and Weather, Part 2: Effects on Chemistry. Atmos. Environ. 115, 499–526.
Makar, P.A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Zakbar, R., Milbrandt, J., Im, U., Galmarini, S., Gravel, S., Zhang, J., Hou, A., Pabla, B., Cheung, P., Bianconi, R., 2015b. Feedbacks between Air Pollution and Weather, Part 1: Effects on Weather. Atmos. Environ. 115, 442–469.
Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., Cook, T., 1998. Measurement of ozone and water vapor by airbus in-service aircraft: the MOZAIC airborne program, an overview. J. Geophys. Res. Atmos. 103, 25631–25642.
Mass, C., Ovens, D., 2011. Fixing WRF's high speed wind bias: a new subgrid scale drag parameterization and the role of detailed verification. In: 91st AMS Annual Meeting, Seattle, WA, 23–27 Jan 2011, 2011.2019B.2016.
McKendry, I., 1989. Numerical simulation of sea breezes over the Auckland region, New Zealand — Air quality implications. Boundary-Layer Meteorol. 49, 7–22.
Milbrandt, J.A., Yau, M.K., 2005. A multimoment bulk microphysics parameterization. Part I: analysis of the role of the spectral shape parameter. J. Atmos. Sci. 62, 3051–3064.
Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. Atmos. 102, 16663–16682.
Morcrette, J.J., Barker, H.W., Cole, J.N.S., Iacono, M.J., Pincus, R., 2008. Impact of a new radiation package, McRad, in the ECMWF integrated forecasting system. Mon. Weather Rev. 136, 4773–4798.
Morrison, H., Thompson, G., Tatarskii, V., 2009. Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes. Mon. Weather Rev. 137, 991–1007.
Moufouma-Okia, W., Rowell, D.P., 2010. Impact of soil moisture initialisation and lateral boundary conditions on regional climate model simulations of the West African Monsoon. Clim. Dyn. 35, 213–229.
Nakanishi, M., Niino, H., 2004. An improved Mellor–Yamada level-3 model with condensation physics: its design and verification. Boundary-Layer Meteorol. 112, 1–31.

Nakanishi, M., Niino, H., 2006. An improved Mellor–Yamada level-3 model: its numerical stability and application to a regional prediction of advection fog. Boundary-Layer Meteorol. 119, 397–407.
Neggers, R.A.J., 2009. A dual mass flux framework for boundary layer convection. Part II: clouds. J. Atmos. Sci. 66, 1489–1506.
Neggers, R.A.J., Köhler, M., Beljaars, A.C.M., 2009. A dual mass flux framework for boundary layer convection. Part I: transport. J. Atmos. Sci. 66, 1465–1487.
Niu, G.-Y., Yang, Z.-L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. J. Geophys. Res. Atmos. 116, D12109.
Noilhan, J., Mahfouf, J.F., 1996. The ISBA land surface parameterisation scheme. Glob. Planet. Change 13, 145–159.
Nordeng, T.E., 1994. Extended Versions of the Convective Parametrization Scheme at ECMWF and Their Impact on the Mean and Transient Activity of the Model in the Tropics. Technical Momorandum No. 206. ECMWF Research Department, European Centre for Medium Range Weather Forecasts, Reading, UK, p. 41.
Penrod, A., Zhang, Y., Wang, K., Wu, S.-Y., Leung, L.R., 2014. Impacts of future climate and emission changes on U.S. air quality. Atmos. Environ. 89, 533–547.
Pleim, J.E., 2007. A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: application and evaluation in a mesoscale meteorological model. J. Appl. Meteor. Climatol. 46, 1396–1409.
Rao, S.T., Galmarini, S., Puckett, K., 2010. Air Quality Model Evaluation International Initiative (AQMEII): advancing the state of the science in regional photochemical modeling and its applications. Bull. Am. Meteorol. Soc. 92, 23–30.
Richardson, H., Basu, S., Holtslag, A.A.M., 2013. Improving stable boundary-layer height estimation using a stability-dependent critical bulk Richardson number. Boundary-Layer Meteorol. 148, 93–109.
Ritter, B., Geleyn, J.-F., 1992. A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Mon. Weather Rev. 120, 303–325.
San Jose, R., Pérez, J.L., Balzarini, A., Baró, R., Curci, G., Forkel, R., Grell, G., Hirtl, M., Honzak, L., Jiménez-Guerrero, P., Langer, M., Pirovano, G., Tuccella, P., Werhahn, J., Žabkar, R., 2015. Sensitivity of feedback effects in CBMZ/MOSAIC chemical mechanism. Atmos. Environ. 115, 646–656.
Seaman, N.L., 2000. Meteorological modeling for air-quality assessments. Atmos. Environ. 34, 2231–2259.
Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., Tercier, P., 2000. Review and intercomparison of operational methods for the determination of the mixing height. Atmos. Environ. 34, 1001–1027.
Seidel, D.J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A.R., Medeiros, B., 2012. Climatology of the planetary boundary layer over the continental United States and Europe. J. Geophys. Res. Atmos. 117, D17106.
Siebesma, A.P., Soares, P.M.M., Teixeira, J., 2007. A combined Eddy-diffusivity mass-flux approach for the convective boundary layer. J. Atmos. Sci. 64, 1230–1248.
Sillman, S., Samson, P.J., 1995. Impact of temperature on oxidant photochemistry in urban, polluted rural and remote environments. J. Geophys. Res. Atmos. 100, 11497–11508.
Skamarock, W.C., Klemp, J.B., 2008. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. J. Comput. Phys. 227, 3465–3485.
Stoeckenius, T., Zagunis, J., Sturtz, T., Sakulyanontvittaya, T., 2015. A comparison between 2010 and 2006 air quality and meteorological conditions, and emissions and boundary conditions used in simulations of the AQMEII-2 North American domain. Atmos. Environ. 115, 389–403.
Szintai, B., Kaufmann, P., Rotach, M.W., 2009. Deriving turbulence characteristics from the COSMO numerical weather prediction model for dispersion applications. Adv. Sci. Res. 3, 79–84.
Tiedtke, M., 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon. Weather Rev. 117, 1779–1800.
Tiedtke, M., 1993. Representation of clouds in large-scale models. Mon. Weather Rev. 121, 3040–3061.
Tompkins, A.M., Gierens, K., Rädel, G., 2007. Ice supersaturation in the ECMWF integrated forecast system. Q. J. R. Meteorol. Soc. 133, 53–63.
Troen, I.B., Mahrt, L., 1986. A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. Boundary-Layer Meteorol. 37, 129–148.
Van den Hurk, B.J.J., Viterbo, M.P., Beljaars, A.C.M., Betts, A.K., 2000. Offline Validation of the ERA40 Surface Scheme. ECMWF Tech. Report. No., p. 43, 75, Reading, UK.
van Meijgaard, E., van Ulft, L.H., Lenderink, G., de Roode, S.R., Wipfler, L., Boers, R., Timmermans, R.M.A., 2012. Refinement and Application of a Regional Atmospheric Model for Climate Scenario Calculations of Western Europe. Climate changes Spatial Planning Publication, p. 44.
Van Weverberg, K., van Lipzig, N.P.M., Delobbe, L., Lauwaet, D., 2010. Sensitivity of quantitative precipitation forecast to soil moisture initialization and microphysics parametrization. Q. J. R. Meteorol. Soc. 136, 978–996.
Vautard, R., Moran, M.D., Solazzo, E., Gilliam, R.C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A.B., Jericevic, A., Prank, M., Segers, A., Silver, J.D., Werhahn, J., Wolke, R., Rao, S.T., Galmarini, S., 2012. Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations. Atmos. Environ. 53, 15–37.
Vicente-Serrano, S.M., Trigo, R.M., López-Moreno, J.I., Liberato, M.L.R., Lorenzo-Lacruz, J., Beguería, S., Morán-Tejeda, E., El Kenawy, A., 2011. Extreme winter precipitation in the Iberian Peninsula in 2010: anomalies, driving mechanisms and future projections. Clim. Res. 46, 51–65.
Vukovic, A., Rajkovic, B., Janjic, Z., 2010. Land ice sea surface model: short description and verification, paper presented at 2010 International Congress on Environmental Modelling and Software Modelling for Environment's Sake, Int. Environ. Modell. and Software Soc., Ottawa.
Vogelezang, D.H.P., Holtslag, A.A.M., 1996. Evaluation and model impacts of alternative boundary-layer height formulations. Boundary-Layer Meteorol. 81, 245–269.
Warner, T.T., Peterson, R.A., Treadon, R.E., 1997. A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. Bull. Am. Meteorol. Soc. 78, 2599–2617.
Wilson, D.R., Ballard, S.P., 1999. A microphysically based precipitation scheme for the UK meteorological office unified model. Q. J. R. Meteorol. Soc. 125, 1607–1636.
Yahya, K., Zhang, Y., Vukovich, J.M., 2014. Real-time air quality forecasting over the southeastern United States using WRF/Chem-MADRID: multiple-year assessment and sensitivity studies. Atmos. Environ. 92, 318–338.
Yeh, K.-S., Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., Staniforth, A., 2002. The CMC–MRB Global Environmental Multiscale (GEM) Model. Part III: nonhydrostatic formulation. Mon. Weather Rev. 130, 339–356.
Zhang, F., Bei, N., Nielsen-Gammon, J.W., Li, G., Zhang, R., Stuart, A., Aksoy, A., 2007. Impacts of meteorological uncertainties on ozone pollution predictability estimated through meteorological and photochemical ensemble forecasts. J. Geophys. Res. Atmos. 112, D04304.
Zhang, Y., 2008. Online-coupled meteorology and chemistry models: history, current status, and outlook. Atmos. Chem. Phys. 8, 2895–2932.
Zhang, Y., Sartelet, K., Zhu, S., Wang, W., Wu, S.Y., Zhang, X., Wang, K., Tran, P., Seigneur, C., Wang, Z.F., 2013. Application of WRF/Chem-MADRID and WRF/Polyphemus in Europe – Part 2: evaluation of chemical concentrations and sensitivity simulations. Atmos. Chem. Phys. 13, 6845–6875.