Air quality forecasts on a kilometer-scale grid over complex Spanish terrains

M. T. Pay1,*, F. Martínez1, M. Guevara1, and J. M. Baldasano1,2
1Earth Sciences Department, Barcelona Supercomputing Center-Centro Nacional de Supercomputación, Barcelona, Spain
2Environmental Modeling Laboratory, Technical University of Catalonia, Barcelona, Spain
*now at: Laboratoire de Météorologie Dynamique, École Polytechnique, Palaiseau, France

Correspondence to: M. T. Pay (maria.pay@bsc.es)

Received: 8 March 2014 – Published in Geosci. Model Dev. Discuss.: 9 April 2014
Revised: 17 July 2014 – Accepted: 31 July 2014 – Published: 8 September 2014

Abstract. The CALIOPE Air Quality Forecast System (CALIOPE-AQFS) represents the current state of the art in air quality forecasting systems of high-resolution running on high-performance computing platforms. It provides a 48 h forecast of NO2, O3, SO2, PM10, PM2.5, CO, and C6H6 at a 4 km horizontal resolution over all of Spain, and at a 1 km horizontal resolution over the most populated areas in Spain with complex terrains (the Barcelona (BCN), Madrid (MAD) and Andalusia (AND) domains). Increased horizontal resolution from 4 to 1 km over the aforementioned domains leads to finer textures and more realistic concentration maps, which is justified by the increase in NO2/O3 spatial correlation coefficients from 0.79/0.69 (4 km) to 0.81/0.73 (1 km). High-resolution emissions using the bottom-up HERMESv2.0 model are essential for improving model performance when increasing resolution on an urban scale, but it is still insufficient. Decreasing grid spacing does not reveal the expected improvement in hourly statistics, i.e., decreasing NO2 bias by only ∼2 µg m−3 and increasing O3 bias by ∼1 µg m−3. The grid effect is less pronounced for PM10, because part of its mass consists of secondary aerosols, which are less affected than the locally emitted primary components by a decreasing grid size. The resolution increase has the highest impact over Barcelona, where air flow is controlled mainly by mesoscale phenomena and a lower planetary boundary layer (PBL). Despite the merits and potential uses of the 1-km simulation, the limitations of current model formulations do not allow confirmation of their expected superiority close to highly urbanized areas and large emissions sources. Future work should combine high grid resolutions with techniques that decrease subgrid variability (e.g., stochastic field methods), and also include models that consider urban morphology and thermal parameters.

1 Introduction

The World Health Organization (WHO) has recently shown that there is sufficient evidence that particulate matter (PM), ozone (O3) and nitrogen dioxide (NO2) affect human health (WHO, 2013). Although NO2 and PM concentrations improved from 2002 to 2011 in Europe, the situation is still far from matching the WHO air quality guidelines (AQG). The European annual limit values for NO2 (annual) and PM10 (daily) had been exceeded at 42–43 % of the traffic stations in 2011. For the same year, about 33 % of the European urban population was exposed to PM10 concentrations above the daily limit value, and nearly 88 % was exposed to the respective WHO AQG (EEA, 2013). Air pollution legislation for the protection of the increasing city population has recently increased the demand for forecasting systems that can assess and understand air pollution dynamics, alert the population when health-related issues occur, and develop emission abatement plans (EEA, 2011).

When applying an air quality modeling system, defining the grid resolution is an important consideration. The potential benefits of higher-resolution modeling should be weighed against the increased complexity of the inputs, increased CPU time, and disk space requirements. In theory, higher-resolution modeling is expected to yield more accurate forecasts because of better resolved model input fields (topography, land cover and emissions). Furthermore, high resolutions (ranging from 1 to 5 km) are essential for reproducing mesoscale phenomena, e.g., those controlling O3 transport along the mountainous northeastern Mediterranean coast (Fay and Neunhäuserer, 2006; Jiménez et al., 2006). Even on fine scales, the modeled concentrations are not necessarily the best (Mass et al., 2002; Gego et al., 2005; Valari...
Several studies have evaluated the impact of increasing horizontal resolution on different scales over the eastern and southeastern USA using the Community Multiscale Air Quality (CMAQ) model and the Comprehensive Air Quality Model with Extensions (CAMx), which range from 32 to 12 to 4 km (Cohan et al., 2006; Tesche et al., 2006; Queen and Zhang, 2008). They found no significant changes for O\textsubscript{3} and PM (< 5 % on average), and those changes were even smaller at resolutions of between 12 and 4 km (< 3 %). Concerning PM components, Fountoukis et al. (2013) found that increased resolution provides differences mostly for primary PM rather than for secondary PM. Recently, a model intercomparison exercise, named ScaleDep, was performed to determine the effect of grid resolution on air quality modeling performance over Europe on a regional and urban scale (Cuvelier et al., 2013). The exercise, involving five chemical transport models (CTMs) (EMEP, CHIMERE, CMAQ, LOTOS-EUROS and RCGC) running under the same conditions over the full year of 2009 and at four resolutions (56, 28, 14, and 7 km), showed that it is difficult to define a grid size that is adequate for resolving the urban signal under all conditions affecting Europe. Still, a 14-km resolution seems to be a good compromise between background applications and those reproducing most of the urban signals (7-km resolution). However, the ScaleDep exercise did not distinguish between the different topographies or complex meteorological patterns that are characteristic of the Iberian Peninsula.

Few studies have been performed over selected areas in Spain, and of those, the focus has been mainly on O\textsubscript{3} and NO\textsubscript{2}. Vivanco et al. (2008) evaluated the annual impact by increasing the model resolution (to 36, 19, and 7 km) over Madrid for NO\textsubscript{2} and O\textsubscript{3}. They used the WRF-CHIMERE model, disaggregating the EMEP emissions inventory based on land-use information. Their evaluation showed that the model improved more for NO\textsubscript{2} than O\textsubscript{3}, with the most significant improvement achieved when the resolution increases from 36 to 19 km rather than to 7 km, which they linked to increased uncertainty in the emissions data introduced with the disaggregation techniques. Jiménez et al. (2006) used the MM5-CMAQ model along with a bottom-up emissions model EMICAT2000, Air Pollutants Emission from Catalonia during the year 2000 to assess the influence of grid resolution on O\textsubscript{3} (at 8, 4, and 2 km) over the complex terrain of the northeastern Iberian Peninsula (Catalonia) during a summer pollution episode. They found that both an improved performance of the mesoscale phenomena and a better allocation of emissions for the 2-km resolution improve the capability of the model to simulate exceedances of European limit values. An important issue in both studies is the emissions modeling approach (top-down vs. bottom-up) when applying high resolutions on a local scale (< 10 km). As Fountoukis et al. (2013) and Timmermans et al. (2013) demonstrate, in the range of local scales (e.g., the greater Paris area), the grid resolution is not currently the major source of discrepancies in model performance, but instead the predicted concentrations and corresponding gradients are more consistent with observed concentrations when provided by bottom-up emissions inventories rather than by downscaled inventories. If local variation in input data (e.g., emissions patterns or land use) cannot be characterized properly, modeling with a finer grid resolution may not provide any great advantages.

Increasing resolution is a technical challenge, since computational cost markedly increases in inverse proportion to grid spacing. The current progress in computation allows increased model resolution and for multiple spatial scales to be investigated, with the aim of establishing an adequate grid size for forecasting air quality on the local scale. Recently, Colette et al. (2014) evaluated the impact of increasing resolution up to 2 km over the European continent by using the CHIMERE model for an episode of air pollution in 2009. They simulated 2 million grid cells using over 2000 CPUs of a high-performance computing system, which was hosted by the French Computing Centre for Research and Technology (CCRT/CEA).

In terms of computational resources, horizontal resolution is critical to an operational air quality forecast. In Europe, operational air quality systems use resolutions between 12 and 25 km, while applications to a single country can reach resolutions between 4 and 10 km (Zhang et al., 2012). Over Spain, there are three systems providing air quality forecasts running at different horizontal resolutions. The lowest resolution system is the Technical University of Madrid’s OPANA (OPerational Atmospheric Numerical model for urban and regional Areas), running at 27 km × 27 km and based on the MM5/CMAQ/EMIMO models (San José et al., 2009). It is followed by the Spanish meteorological office’s system (AEMET, http://www.aemet.es/es/eltiempo/prediccion/calidad_del_aire), which forecasts at 10 km × 10 km using the HIRLAM-HNR/MOCAGE/GEMS-TNO models. The CALIOPE Air Quality Forecast System (CALIOPE-AQFS; Baldasano et al., 2011; Pay et al., 2012a, and references therein), of the Barcelona Supercomputing Center–Centro Nacional de Supercomputación (BSC-CNS), runs at the highest resolution, 4 km × 4 km, and it is based on the WRFv3.5/CMAQv5.0.1/HERMESv2.0/BSC-DREAM8bv2 models. Moreover, CALIOPE-AQFS provides 1 km × 1 km resolution forecasts for the Madrid and Barcelona metropolitan areas (since 2009) and the Andalusian region (since 2013). Such a resolution has been possible thanks to both the high-performance computing resources at the BSC-CNS and the availability of detailed emissions data covering Spain.

www.geosci-model-dev.net/7/1979/2014/
The present work aims to assess the impact of increasing the horizontal resolution from 4 to 1 km, specifically over areas affected by heterogeneous emission patterns and complex terrains such as the Barcelona and Madrid metropolitan areas (BCN and MAD), together with the Andalusian region (AND). For that purpose, the CALIOPE-AQFS forecasts pollutant concentrations ($O_3$, $NO_2$, and $PM_{10}$) at two horizontal resolutions: first at a 4-km resolution covering Spain (IP4), and second at a 1-km resolution covering the AND, BCN, and MAD domains. Here, a study is performed for the period April 2013, which presents a 7-day air pollution episode. We use observations from routine air quality monitoring networks to evaluate both resolutions.

Section 2 describes the configuration and computational setup of CALIOPE-AQFS. Section 3 quantifies the impact of resolution increase on forecasting hourly concentrations (and exceedances) in terms of pollutant, domain, building density and major emission sources. Section 4 concludes with the main results and some recommendations.

2 Methodology

2.1 Domain and period under study

Figure 1 shows the primary $NO_2$ emission patterns and topographic characteristics of the BCN, MAD and AND domains. BCN is a coastal area characterized by several valleys perpendicular to the coastline and two main mountain...
ranges, one coastal (500 m in height) and one pre-coastal (1000–1700 m in height). These features induce mesoscale phenomena such as sea-breeze and mountain-valley winds. On the other hand, MAD is a continental region with a much simpler topography that includes the Tajo Valley in the south and the mountain range of the Central System in the north, with summits reaching 2500 m.

The urban contribution in BCN (3.1 million inhabitants) is accompanied by industrial and power generation emissions, the road network and the harbor. Meanwhile, the Spanish capital of MAD is mainly influenced by emissions from the urban area (5.8 million inhabitants) and the road network that connects MAD with the surrounding commercial and industrial zones, as well as the urban areas.

AND is the southernmost region in Spain, with complex topography characterized by the large depression of the Guadalquivir Basin (delimited by the Iberian Massif and the Betic Range), which crosses the region from NE to SW over a 60 km stretch. About three quarters of AND has a mountainous orography, including the Sierra Nevada (3481 m). AND includes one of the five biggest cities in Spain, Seville (~700,000 inhabitants), which hosts industrial and electric generation activities around the Algeciras bay, and is affected by dense maritime traffic through the Strait of Gibraltar.

The study is done for April 2013. At the beginning and end of the month, the synoptic circulation was controlled by a low-pressure system located over the south of the British Isles, and affected Western Europe by leading to atmospheric instability over the IP. This pattern is typical of transitional months such as April and November (García-Valero et al., 2012; Valverde et al., 2014), which produce precipitation and low temperatures because of cold and humid winds from the Atlantic Ocean. In contrast, from 12 to 18 April, there was a high pressure system crossing the Iberian Peninsula in the SW–NE direction, transporting dust from the Sahara Desert and increased temperatures up to 25–28 °C, confirmed by the mineral dust forecasts of the BSC-DREAM8b model (http://www.bsc.es/earth-sciences/mineral-dust-forecast-system/bsc-dream8b-forecast). During the latter episode, available air quality stations in the study domains observed several exceedances of the European limit values (8 exceedances of the NO₂ hourly limit value, 25 exceedances of the O₃ information threshold, and 31 exceedances of the PM₁₀ daily limit value).

2.2 CALIOPE-AQFS

CALIOPE-AQFS has provided 48 h air quality forecasts for Europe and Spain since October 2006 (www.bsc.es/caliopo), and has been described and evaluated in detail elsewhere (Baldasano et al., 2008, 2011; Pay et al., 2011, 2012a). Briefly, it integrates a meteorological model (WRF-ARW v3.5; Skamarock and Klemp, 2008), an emissions model (HERMESv2.0; Guevara et al., 2013), a chemical transport model (CMAQv5.0.1; Byun and Schere, 2006; Appel et al., 2013), and a mineral dust atmospheric model (BSC-DREAM8bv2; Pérez et al., 2006; Basart et al., 2012); together, all of these comprise an air quality forecast system.

Figure 1 shows the working domains of the CALIOPE-AQFS. First, CALIOPE-AQFS was run over Europe at a 12 km × 12 km horizontal resolution using initial/boundary conditions from the final analyses of the National Centers of Environmental Prediction (FNL/NCEP). The analyses began at 12:00 UTC, at intervals of 6 h (0.5° × 0.5°) for meteorology. The LMDz-INCA2 global model (3.75° × 2.5°; Szopa et al., 2009) was used for chemistry. Then, CALIOPE-AQFS was run at 4 km × 4 km horizontal resolution (IP4) over the IP using one-way nesting. In the present work CALIOPE-AQFS runs at 1 km × 1 km over the AND, BCN and MAD domains, with a nesting of over the IP4 domain. HERMESv2.0 forecasts anthropogenic emissions for the year 2009 by following a bottom-up methodology (point, linear and area), and biogenic emissions using the MEGANv2.0.4 model (Guenther et al., 2006). Emissions are aggregated into 1 km grids for the AND, BCN and MAD 1 km simulations, and into a 4 km grid for IP4.

Vertically, WRF-ARW is configured with 38 sigma layers up to 50 hPa, with 11 characterizing the planetary boundary layer (PBL). Meanwhile, the CMAQ vertical levels are obtained by collapsing from the 38 WRF layers to a total of 15 layers whose depth increases with height, from the surface up to 50 hPa. Six layers are within the PBL, and the first layer depth is 39 m.

The present WRF setup uses the rapid radiation transfer model (RRTM) and Dudhia for long- and short-wave radiation, respectively, the Kain–Fritsch cumulus parameterization (Kain and Fritsch, 1990), the single-moment 3-class (WSM3) microphysics scheme, and the Yonsei University PBL scheme (YSU). The Noah land-surface model (NoahLSM), based on the US Geological Survey (USGS) land-use data from the year 1993, is used by default in the present WRF configuration.

Currently, a new CMAQ version is being tested in the Community Modeling and Analysis System (CMAS) community, namely CMAQv5.0 (CMAQ, 2014). It includes substantial scientific improvements over version 4.5, and is especially devoted to improving secondary organic aerosol (SOA) formation as well as the dynamic interactions of fine and coarse aerosols. Based on the evaluation results from the previous CMAQ version within CALIOPE-AQFS (4.5 vs. 5.0) (Pay et al., 2012b), CMAQ has been updated to version 5.0.1 using the CB05 chemical mechanism (Yarwood et al., 2005), the AERO5 for aerosol modeling, and the in-line photolysis calculation. Although not used here, future simulations will use the updated version of the aerosol module (AERO6), which includes a significant improvement in the science of the primary organic aerosol aging and the sulfur chemistry.

CALIOPE-AQFS considers desert dust contribution by means of the BSC-DREAM8bv2, which runs offline at a 0.5° × 0.5° resolution covering Europe, North Africa and
Figure 2. Air quality stations for NO\(_2\), O\(_3\) and PM\(_{10}\) in the three domains under study (AND, BCN and MAD) in April 2013. Different types of stations are shown by symbols and color codes. The various symbols represent the major emission type affecting each station (Traffic: triangle; Industrial: square; and Background: circle), while the colors reflect the environment of each station (Urban: red; Suburban: green; and Rural: orange). Cyan dots represent the METAR stations used in Sect. S1 in the Supplement.

2.3 Computational strategy

Running CALIOPE-AQFS at 4 and 1 km is a technical challenge. The simulations are run on the MareNostrum supercomputer (Intel Xeon E5-2670, 16 CPUs and 64 GB RAM memory per node) at BSC-CNS. Table 1 depicts the computational requirements for forecasting air quality at 48 h for each domain. The number of CPUs was chosen to maximize CPU efficiency. Thanks to the parallelization of meteorological and air quality models, MareNostrum uses up to 256 CPUs. Due to the variable nature and complex dependencies, the computational time for forecasting 48 h of air quality fields for the 4 domains is 8–9 h. The most computationally demanding domain is AND, at 1 km resolution (666 \(\times\) 358 cells, 256 CPU max., and 300 min). For the April 2013 simulation, times add up to 2880 CPU hours day\(^{-1}\), or 86 400 CPU hours on a single processor (9.86 years). The storage for the April 2013 output files was 6.13 TB (\(\sim\)200 GB day\(^{-1}\)).

2.4 Evaluating the increase in resolution

A comparison of CALIOPE-AQFS model outputs and measurements was made for gas-phase and aerosol concentrations (O\(_3\), NO\(_2\), and PM\(_{10}\)). Representativeness continues to be a challenge when comparing gridded simulations to observational data at a point in time and space, as modeled concentrations represent a volumetric average over an entire grid cell. Furthermore, the stochastic compound embedded in the observations is not accounted for. Concerning temporal representativeness in the present comparison, both modeled and measured concentrations are averaged.
hourly. CALIOPE-AQFS operationally receives air quality measurements from Spanish administrative networks in near real time (NRT) without any quality control. For the present study, NRT measurements are filtered by removing data before and after measurement interruptions or calibrations. Also, a minimum cut-off threshold of 1 µg m\(^{-3}\) is applied to the observed concentrations in order to avoid unrealistic observations. After filtering, the number of stations is 48/30/36 for O\(_3\), 51/42/42 for NO\(_2\), and 52/15/33 for PM\(_{10}\) in the AND/BCN/MAD domains, respectively.

The meteorological fields are evaluated for wind speed (U10) and wind direction (WD10) at 10 m, and for temperature at 2 m (T2M). The 10 METAR stations are all located at airports (6/2/2 stations in AND/BCN/MAD), and are discussed in Sect. S1 in the Supplement.

Figure 2 shows the location of the air quality and METAR (ME Tecological Aerodrome Report) stations over the respective domains. The spatial representativeness of the air quality network is highly variable. The influence of the station type is based on two classifications of air quality monitoring stations, the environment type (rural, R; suburban, S; and urban, U), and the dominant emissions source (traffic, T; industrial, I; and background, B). These were derived from council decision 97/100/EC (Garber et al., 2002).

The evaluation is based on discrete statistics performed on an hourly basis. We consider the correlation coefficient (r, Eq. A1), mean (absolute, relative, and fractional) biases (MB, MNBE, and MFB; Eqs. A2–A4), and error (MAE, MNGE, and MFE; Eqs. A5–A7). The root mean square error (rmse, Eq. A8) is also calculated, because it intensifies large differences between measured and observed concentrations (Table A1).

In order to evaluate the effect of increased resolution on forecast exceedances and non-exceedances of limit values established by the European legislation, we calculate categorical statistics based on comparisons with fixed concentration thresholds (T). The calculated statistics are accuracy (A, Eq. A9), bias (B, Eq. A10), probability of detection (POD, Eq. A12), critical success index (CSI, Eq. A11), and false alarm ratio (FAR, Eq. A13), whose formulas and descriptions are provided in Table A2 and also in Kang et al. (2005) and Eder et al. (2006). The 2008/50/EC Directive sets an information threshold of 180 µg m\(^{-3}\) for maximum daily O\(_3\) concentrations (Max 1 h O\(_3\)) and a target value of 120 µg m\(^{-3}\) for the maximum daily 8 h running O\(_3\) mean (Max 8 h O\(_3\)), which should not be exceeded more than 25 days per year. It establishes a limit value of 200 µg m\(^{-3}\) for maximum daily NO\(_2\) concentrations (Max 1 h NO\(_2\)), and 50 µg m\(^{-3}\) for the daily PM\(_{10}\) mean (Mean 24 h PM\(_{10}\)), which should not be exceeded more than 35 times per year. Therefore, categorical evaluations will be performed for Max 1 h NO\(_2\), Max 1 h and Max 8 h O\(_3\), and Mean 24 h PM\(_{10}\). Note that mean and maximum concentrations are calculated by considering at least 75 % of the data in the corresponding time base, i.e., values of at least 18 h per day for Mean 24 h, Max 1 h, and Max 8 h, and 6 h for 8 h values, as established by 2008/50/EC.

Table 1. CALIOPE-AQFS computational requirements, in terms of central processor units (CPUs) and computational time (in min), for simulating 48 h air quality forecasts as a function of the domains IP-4 km (D2), AND-1 km (D3), BCN-1 km (D5) and MAD-1 km (D4), all of which are described in Fig. 1. D domains are described in Fig. 1.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Meteorological modeling</th>
<th>Emissions modeling</th>
<th>Air quality modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>128 CPU / 15 min</td>
<td>1 CPU / 4 min</td>
<td>256 CPU / 210 min</td>
</tr>
<tr>
<td>IP-4 km</td>
<td>128 CPU / 80 min</td>
<td>1 CPU / 4 min</td>
<td>256 CPU / 220 min</td>
</tr>
<tr>
<td>AND-1 km</td>
<td>128 CPU / 20 min</td>
<td>1 CPU / 1 min</td>
<td>256 CPU / 150 min</td>
</tr>
<tr>
<td>BCN-1 km</td>
<td>128 CPU / 20 min</td>
<td>1 CPU / 1 min</td>
<td>256 CPU / 150 min</td>
</tr>
<tr>
<td>MAD-1 km</td>
<td>128 CPU / 110 min</td>
<td>1 CPU / 1 min</td>
<td>256 CPU / 110 min</td>
</tr>
<tr>
<td></td>
<td>(399 × 399 cells)</td>
<td>(666 × 358 cells)</td>
<td>(146 × 146 cells)</td>
</tr>
<tr>
<td></td>
<td>(146 × 158 cells)</td>
<td>(146 × 158 cells)</td>
<td>(146 × 158 cells)</td>
</tr>
</tbody>
</table>

3 Concentration maps and spatial representativeness

To analyze the spatial differences between resolutions, Figs. 3, 4, and 5 show the monthly mean concentration maps for April 2013 over the MAD, BCN and AND domains at 4 km (left panels) and 1 km (right panels) for NO\(_2\), O\(_3\), and PM\(_{10}\), respectively.

The maps of NO\(_2\) and PM\(_{10}\) at both resolutions display similar distributions along the MAD and BCN urban plumes. On-road traffic constitutes the main source of primary pollutants in MAD and BCN. HERMESv2.0 estimates that 75 and 59 % of NO\(_x\) emissions are produced by on-road traffic in both domains, respectively. Consequently, when the resolution increases, the monthly mean O\(_3\) concentration maps are almost identical, although the NO\(_x\) titration effect on O\(_3\) is significant along highways and major point sources. In AND, NO\(_2\) and O\(_3\) concentrations are also conserved between resolutions along the shipping route crossing the Strait of Gibraltar towards the Mediterranean Sea.

However, the definitions of NO\(_2\) concentrations along highways connecting the biggest cities to the rest of the country and industrial sectors are more easily identified at 1 km simulations than at 4 km, especially along those roads from/to Barcelona (e.g., the AP7 Mediterranean highway and the C32, which connect the harbor and the airport) and Madrid (the A-2 and A-6 in the north, and the A-3, A-4 and A-5 in the south). In the same way, 1 km O\(_3\) maps are more textured than those at 4 km along highways, because the titration effect is more significant at 1 km, due to less dilution.
within grid cells. The titration effect of NOx on O3 over the main sources is more forceful in BCN than in MAD, given that BCN has a larger concentration gradient resulting from complex topography and recirculation flows that accumulate pollutants.

The improvement in the definition along roads in AND is lower than that observed in the MAD and BCN domains, due to the fact that the AND domain is bigger and displays lower traffic emission sources than the MAD or BCN domains. Regarding PM10, the main component in AND is the desert dust (~40% in both resolutions) from North Africa. This is because there were two episodes on 14–19 and 25–26 April that affected the IP, as shown by the S–N PM10 gradient (Fig. 5e and f). The desert dust is transported from long-range simulations with BSC-DREAM8bv2.

Over complex terrains, the 1 km simulation produces more realistic looking NO2 concentration maps because of its more detailed topographic information. For instance, the BCN 1 km simulation displays the lowest NO2 concentrations (<10 µg m⁻³) along the coastal chain (500 m in height) and pre-coastal chain (1000–1700 m in height), except for the city’s urban hill, where concentrations reach 20–40 µg m⁻³. In contrast, the 4-km simulation provides smoother NO2 concentrations without any concentration gradient. Thus, the 1-km simulation generates slightly higher O3 background concentrations than the 4-km simulation along...
the BCN pre-coastal chain (66–70 vs. 70–74 µg m\(^{-3}\)), as well as across the Iberian Massif (AND), where the O\(_3\) displays a significant structure due to the higher-resolution topography that follows the basin and the areas of high-density on-road traffic.

Figures 3, 4, and 5 include dots corresponding to mean concentrations at air quality stations that help qualitatively to evaluate the modeled spatial representativeness at both resolutions. Note the strong correlation between NO\(_2\) observations and the 1-km simulation near the primary suburban traffic roads in BCN (e.g., Vilafranca, Igualada, Manresa, and Mataró). Regarding O\(_3\), although observed concentrations depict an overall tendency of the model to underestimate concentrations at both resolutions, the 1-km simulation shows better agreement with measurements at rural background stations (e.g., the El Atazar, San Martín, Villa de Prado, Villarejo and Orosco stations in MAD), and at suburban traffic stations (e.g., Manresa, Igualada and Vilafranca in BCN, with modeled O\(_3\) concentrations of around 54–58 µg m\(^{-3}\) at 1 km, and 60–66 µg m\(^{-3}\) at 4 km). For PM\(_{10}\), comparisons with measurements show that modeled concentrations are underestimated over background areas, mainly outside the urban/suburban areas, as already discussed in Pay et al. (2012a). However, PM\(_{10}\) measurements at the urban/suburban stations of Vilafranca, Sant Celoni, and Mataró in BCN (14–16 µg m\(^{-3}\)) show a higher correlation at 1 km than at 4 km (12–14 µg m\(^{-3}\) vs. 8–10 µg m\(^{-3}\)).
The spatial variability of the increased resolution is analyzed quantitatively by means of concentration maps, shown in Fig. 6 for NO$_2$, O$_3$ and PM$_{10}$ over AND, BCN and MAD. Over all the domains, the correlation improves as a function of the resolution increase for NO$_2$ and O$_3$, indicated by the increase in monthly $r$ from 0.79 (4 km) to 0.81 (1 km) for NO$_2$, and from 0.69 to 0.73 for O$_3$. The slopes improve with the resolution increase, from 0.72 (4 km) to 0.77 (1 km) for NO$_2$, and from 0.50 (4 km) to 0.54 (1 km) for O$_3$. This results from the improved model performance at urban stations, indicating that CALIOPE-AQFS captures the magnitude of the variability between urban regions better at 1 km than at 4 km.

In contrast, for PM$_{10}$, the monthly $r$ decreases from 0.67 to 0.58 when the resolution increases. While the PM$_{10}$ spatial variability over BCN and MAD improves when the resolution increases ($r$ increases by 0.01 and 0.04, respectively), the global correlation coefficient decreases at the AND stations (52/100) by 0.1, from 0.36 (4 km) to 0.26 (1 km). Despite the unfavorable effect of the resolution increase in PM$_{10}$ over AND, the NO$_2$ and O$_3$ concentrations show the highest absolute increases for the spatial $r$ over this domain, from 0.62 (4 km) to 0.71 (1 km) for NO$_2$, and from 0.58 (4 km) to 0.64 (1 km) for O$_3$ (increasing $r$ by 0.09 and 0.06, respectively).
is performed on the temporal basis established by the European legislation, but it uses a T based on the 75th percentile (75p) of the observed concentrations in each case, where $T$ is 71 µg m$^{-3}$ for Max 1 h NO$_2$, 108 (101) µg m$^{-3}$ for Max 1 h (Max 8 h) O$_3$, and 27 µg m$^{-3}$ for Mean 24 h PM$_{10}$.

Overall, CALIOPE-AQFS underestimates exceedances at both resolutions, indicating that errors of missing observed exceedances are not completely resolved by increasing the horizontal resolution ($a < d$). The best performance is found for Max 1 h NO$_2$, where categorical bias ($B$, Eq. A10) improves from 37 % (4 km) to 40 % (1 km), although the tendency to underestimate exceedances remains with the resolution increase ($B < 100\%$).

For NO$_2$ Max 1 h, there are 953 observed exceedances ($b + d$) of the threshold ($T = 47 µg m^{-3}$). Increasing the resolution increases the POD from 49 % (4 km) to 56 % (1 km). As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.

Increasing resolution decreases POD and CSI by 22 and 25 %, respectively, when the resolution is increased from 4 to 1 km. As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.

Increasing resolution decreases POD and CSI by 22 and 25 %, respectively, when the resolution is increased from 4 to 1 km. As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.

Increasing resolution decreases POD and CSI by 22 and 25 %, respectively, when the resolution is increased from 4 to 1 km. As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.

Increasing resolution decreases POD and CSI by 22 and 25 %, respectively, when the resolution is increased from 4 to 1 km. As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.

Increasing resolution decreases POD and CSI by 22 and 25 %, respectively, when the resolution is increased from 4 to 1 km. As with POD, CSI examines the exceedances, but in a more comprehensive way, by considering both false alarms and missing events. POD and CSI increase by 14 and 20 %, respectively, when the resolution is increased from 4 to 1 km. For O$_3$, POD has a relatively low POD value, and it decreases as the resolution increases. Of the 1306 observed exceedances of the 108 µg m$^{-3}$ Max 1 h, CALIOPE-AQFS detected 112 exceedances at 4 km and only 96 at 1 km.
For various reasons, accuracy ($A$) remains almost constant when the resolution increases. For NO$_2$ and O$_3$, this is due to a stable sum of $b$ and $c$, increasing $b$ at the cost of $c$, and vice versa. For NO$_2$, the number of hits ($b$) to forecast Max 1 h at 1 km is higher than 4 km (537 vs. 466), but the number of correct negatives at 1 km is lower than at 4 km (2439 vs. 2517). The resolution increase has the opposite effect on O$_3$, decreasing the number of hits by 14 % for Max 1 h and 33 % for Max 8 h. The increase in the number of correct negatives is less than 2 % in both Max 1 h and Max 8 h O$_3$.

### 4.2 PM$_{10}$ components

The resolution increase has the lowest effect on PM$_{10}$ hourly concentrations and its exceedances ($< 1$ %). PM$_{10}$ components are secondary inorganic aerosols (SIA), which include sulfate ($SO_4 = SO_4^{2-}$), nitrate ($NO_3 = NO_3^-$), ammonium ($NH_4 = NH_4^+$), SOA, elemental carbon (EC), sea salt (SS), desert dust (DD), and primary PM (PPM).

Pay et al. (2012a) evaluated the PM components at some urban and rural background stations in Spain using the CALIOPE-AQFS based on CMAQv4.5, where they showed that the model underestimated SIA by a factor of 2–3. The highest underestimation was found for fine carbonaceous aerosols (a factor of 4), in part related to the state of the science concerning SOA formation pathways. The updated version of CMAQ, v5.0.1 includes scientific improvements to SOA formation and aerosol dynamics, which could improve the modeled PM$_{10}$ performance and its components.

Figure 8a shows that the resolution increase does not significantly change the PM$_{10}$ composition. DD remains the main contributor ($\sim 40–41$ %), followed by PPM (22–24 %), SIA ($\sim 21–22$ %), SS (9–11 %), EC ($\sim 4$ %) and SOA ($\sim 0.6$ %). However, the effect of the increased resolution on
PM\textsubscript{10} component concentrations is different (Fig. 8b), depending on their origin, atmospheric cycle and the way they are modeled. DD concentrations do not change between resolutions, because they are mass conservative when interpolated from 0.5° × 0.5° to 1 km × 1 km.

For SIA, increasing the resolution increases NO\textsubscript{3} and NH\textsubscript{4} concentrations by 4 and ~2 %, respectively, and it decreases SO\textsubscript{4} by ~2 %. The NH\textsubscript{4} increase means there are more primary precursors (H\textsubscript{2}SO\textsubscript{4} or HNO\textsubscript{3} / NO\textsubscript{2}) available to neutralize NH\textsubscript{3} (gas) to NH\textsubscript{4} (aerosol). However, the variability between SO\textsubscript{4} and NO\textsubscript{3} is more difficult to explain, due to the nonlinearity of photochemistry and aerosol formation, which is controlled to some extent by the ISORROPIA thermodynamic equilibrium. Furthermore, the absence of aerosol measurements for April 2013 does not allow us to investigate this further.

The resolution increase results in a decrease in SS of ~16 %, the largest change of all the PM\textsubscript{10} components. CMAQv5.0.1 simulates SS emission as a function of the wind speed and the relative humidity (Gong, 2003; Zhang et al., 2005). Although not shown here, when the resolution increases, the wind speed decreases at the available PM\textsubscript{10} stations by ~1.4/0.4/0.2 m s\textsuperscript{-1} over AND/BCN/MAD, and also over the open ocean.

For primary PM components EC and PPM, increasing resolution results in higher concentrations (by 10 and ~12 %, respectively). As for NO\textsubscript{2}, the 1 km simulation leads to a reduced effect of artificial dilution of emissions in a grid cell, so concentration gradients are stronger than in the 4-km simulation.

### 4.3 Domain

The resolution increase has varying impacts due to differences in geographical location and emissions patterns over the domains (Fig. 7). BCN shows the highest NO\textsubscript{2} bias decrease (73 %) when the resolution increases, but no effect on the correlation (~7 %). However, O\textsubscript{3} shows significant variability over BCN, increasing \( r \) by 4 % and MB by 23 %. To a lesser extent, MB also increases over AND (by 8 %). Meanwhile, the variability over MAD is reduced (bias differences less than 4 %). MB decreases for PM\textsubscript{10} (bias differences less than 1 µg m\textsuperscript{-3}) over the urban domains of MAD (3 %) and BCN (16 %), and increases over AND (7 %).

Figure 9 analyzes the impact of the resolution increase on daily cycles. Although PBL measurements are not available, PBL daily cycles are displayed together in order to find some correlations with the daily pollutant variability. Due to the lamination of PBL growth by the Mediterranean Sea breezes, the PBL reaches its maximum height at midday, being the highest in MAD (1600 m a.g.l.), followed by AND (1000 m a.g.l.) and BCN (900 m a.g.l.).
As shown in Sect. S1 in the Supplement, the pollutant transport in the BCN coastal domain is controlled by mesoscale phenomena such as sea breezes (day) and land breezes (night), which are a result of its complex topography and location (Baldasano et al., 1994; Millán et al., 1997; Gonçalves et al., 2009). The NO\textsubscript{2} daily cycle is highly influenced by traffic emissions (Fig. 9). Both resolutions show the highest underestimations for the morning peak (6 a.m.) (∼20 µg m\textsuperscript{-3}). Although the afternoon peak is well reproduced, there is excessive variability at both resolutions, a result of problems with the wind direction. During the sea breeze period, the mean simulated wind was more easterly than westerly, as registered by measurements (Sect. S1 in the Supplement). Several works indicate that WRF has difficulty reproducing the morning and evening transition over the urban environment, possibly because it does not model the heat retention in cities (Baldasano et al., 2006; Appel et al., 2013). Increasing the resolution increases NO\textsubscript{2} concentrations by 21 %, from 14 µg m\textsuperscript{-3} (4 km) to 17 µg m\textsuperscript{-3} (1 km) during the morning hours after sunrise (5–9 a.m.), and in the evening hours after sunset (5–9 p.m.). This behavior could be explained by PBL variability when increasing the resolution, which decreases PBL height by ∼33 m for these hours.

NO\textsubscript{2} performance impacts the O\textsubscript{3} daily cycles over BCN, showing that 4 and 1 km simulations underestimate maximum O\textsubscript{3} concentrations by ∼20 µg m\textsuperscript{-3} at midday (1–4 p.m.) and overestimate minimum O\textsubscript{3} concentrations by ∼20 µg m\textsuperscript{-3} in the morning hours after sunrise (5–9 a.m.). The resolution increase results in slightly lower O\textsubscript{3} concentrations at night, which is perhaps the result of lower PBL heights in the 1-km simulation during the early morning and late afternoon, when the PBL height tends to be the lowest. During these hours, the titration effect of NO\textsubscript{2} on O\textsubscript{3} is greater, improving the O\textsubscript{3} overestimation of the daily minimum, which allows a slightly higher hourly \( r \) (2 %). However, O\textsubscript{3} underestimation increases in the late afternoon, contributing to an increase in the hourly mean bias from ∼9 µg m\textsuperscript{-3} (4 km) to ∼11 µg m\textsuperscript{-3} (1 km).

In BCN, the PM\textsubscript{10} underestimation is not systematic throughout the daily cycle (Fig. 9), which shows a bias of ∼20 (10) µg m\textsuperscript{-3} at day (night) time. The higher daytime underestimation as compared to the nighttime cannot be explained by the current results, but it could be a result of missing sources (e.g., fugitive agricultural emissions and windblown dust), problems with PBL height overestimation, and an excess dilution of emissions. The resolution increase reduces the bias by ∼1 µg m\textsuperscript{-3} (16 %), especially during early morning and late afternoon, when the highest PBL height variability between resolutions is observed. Although the evaluation of T2M, U10 and WD10 indicates that the resolution increase has a small effect over BCN (Sect. S1 in the Supplement), the reduction of the artificial dilution of NO\textsubscript{2} emissions – together with a lower PBL height at 1 km than at 4 km during the night and early morning – improves NO\textsubscript{2}, O\textsubscript{3} and PM\textsubscript{10} concentrations, which in turn decreases their biases.

For AND, the model underestimates observed NO\textsubscript{2} concentrations throughout the daily cycle by ∼5 µg m\textsuperscript{-3} for both the 4 and 1 km simulations, with the highest underestimation during the morning peak (∼25 µg m\textsuperscript{-3}) and the lowest during the afternoon peak (∼10 µg m\textsuperscript{-3}). The resolution increase reduces the bias from ∼3.5 to ∼2 µg m\textsuperscript{-3} (by 43 %) and increases \( r \) by 7 % (from 0.39 to 0.41). As in BCN, the NO\textsubscript{2} underestimation directly impacts the O\textsubscript{3} daily cycle.
Figure 9. Daily cycles for NO$_2$, O$_3$ and PM$_{10}$ for each study domain at available stations as a function of resolution. No observations of PBL are available. Q1, Q2 and Q3 indicate quartiles for the daily cycle. Bars show Q1 and Q3 at each hour.

(which the resolution increase cannot resolve), increasing the bias by $\sim 1$ µg m$^{-3}$, a phenomenon that is predominant in the morning hours. In the case of PM$_{10}$, the daily cycle indicates that the biases are almost systematic throughout the day ($\sim 22$ µg m$^{-3}$). Increasing the resolution increases the bias by less than 4% in the late afternoon, which is perhaps dominated by the decrease in PBL height. When the resolution is increased, NO$_2$ performs better, partially due to the improved model performance for the temperature and wind speed (Sect. S1), as well as the lower nighttime and higher daytime PBL. Meanwhile, the O$_3$ and PM$_{10}$ performances do not change significantly.

During April 2013, the main flow over MAD was controlled by S–SW synoptic winds channeled by orographic barriers in the NW domain and the Tajo Valley (Valverde et al., 2014). The NO$_2$ daily cycle depicts a high influence of traffic emissions (Fig. 9), showing significant model underestimation at both resolutions for the morning/afternoon peaks ($\sim 15/10$ µg m$^{-3}$). Note that, in terms of the temporal variability of the NO$_2$ concentrations, the model performs well at the afternoon peak, when the air flow is controlled by southeastern winds. The NO$_2$ performance leads to more accurate O$_3$ daily cycles than in AND and BCN, especially in the early morning, when the titration effect of...
Figure 10. Temporal series and daily cycles for NO\textsubscript{2} and O\textsubscript{3} at background (a and b, respectively) and traffic stations (c and d, respectively) for the episode of 12–18 April 2013. Q1, Q2 and Q3 indicate quartiles for the daily cycle. Bars show Q1 and Q3 at each hour.

NO\textsubscript{2} is greater, because the NO\textsubscript{2} morning peak underestimation is smaller when compared to the other domains. Meanwhile, the modeled PM\textsubscript{10} at both resolutions presents a profile controlled by traffic emissions. Observed concentrations display a flatter daily cycle, in which the model underestimation reaches 40 \mu g m^{-3} in the morning. Increasing resolution shows a positive effect for NO\textsubscript{2} and PM\textsubscript{10}, because it increases r by 0.01 and reduces MB and the rmse by 0.1–0.2 \mu g m^{-3}. However, it has the lowest variability when compared to the other domains (less than 5 \% for bias, error and r), likely the result of a relatively simpler topography and meteorological patterns.
4.4 Environment and major sources

Figure 7 shows that the resolution impact also depends on the type of area and the dominant emission source. Theoretically, the meteorological fields of urban areas differ from those of surrounding rural areas, because of their different morphologies (radiation trapping and wind profiles), surface materials (heat storage) and variable energy consumption (heat release).

Increasing resolution reduces the NO$_2$ bias at suburban and urban stations by 1.8–2 µg m$^{-3}$, and to a lesser extent by 1.2 µg m$^{-3}$ at rural stations. The correlation coefficients also improve at suburban stations (from 0.48 to 0.52) and rural stations (from 0.34 to 0.35). This may be due to better allocation of land-use categories (urban vs. rural) and their fraction in a grid cell in the 1 km simulation than in the 4 km simulation. The NO$_2$ bias decreases by 39% (65%) at urban (background) stations, but O$_3$ biases increase by 9% (5%). For PM$_{10}$, the resolution increase does not significantly change as a function of area type, with differences in bias and error of less than ±4% (<0.5 µg m$^{-3}$) between the two resolutions.

The low improvement at urban stations may be because the NoahLSM land-surface model does not consider the effect of urban morphology or thermal parameters in order to model meteorological fields accurately. Modeling air quality on an urban scale over cities requires a description of the heat–momentum exchange between buildings and the lower atmospheric layers. For instance, the impact of using an urban model on meteorological fields over the greater Paris area was studied by Kim et al. (2013) using WRF with the urban canopy model, demonstrating that below 1 km in height, overestimations of wind speed were significantly reduced.

The impact on $r$ of increasing resolution is less than 2% for primary pollutants near important emissions sources. For example, it reduces NO$_2$ biases at traffic (industrial) stations by ~3 µg m$^{-3}$ (2 µg m$^{-3}$), but it increases O$_3$ biases by ~2 µg m$^{-3}$ (1 µg m$^{-3}$). However, the resolution increase in the range of 4–1 km does not exhibit the expected improvement in the hourly statistics that are based on the constraints of the current model formulation. In other words, it cannot resolve the subgrid air quality variability merely by increasing resolution. For instance, although on-road traffic emissions are estimated by following a bottom-up approach along highways and routes, heterogeneity is lost in the CTM volume-averaging process, which artificially dilutes emissions rates over the grid cells. The resolution effect is low at background stations, which generally are not influenced by any single source, but rather by the integrated contribution from all sources upwind of the stations where variations are less than 1% for O$_3$ and PM$_{10}$ (<1 µg m$^{-3}$). However, background NO$_2$ levels increase by ~1 µg m$^{-3}$ (48%) from 4 km to 1 km.

Figure 10 shows the temporal series and daily cycles for NO$_2$ and O$_3$ at traffic and background stations throughout the episode of 12–18 April 2013. At traffic stations, the temporal series show a remarkable O$_3$ daily cycle (observed 25p$=23.2$ µg m$^{-3}$ and 75p$=77.5$ µg m$^{-3}$), due to O$_3$ destruction caused by high NO$_x$ levels (observed 50p$=34.5$ µg m$^{-3}$). In contrast, the NO$_2$-limited regime at background sites (observed 50p$=19$ µg m$^{-3}$) allows for higher O$_3$ concentrations (observed 25p$=38$ µg m$^{-3}$ and 75p$=89$ µg m$^{-3}$) than in high NO$_2$ environments.

During the episode mentioned above, the resolution increase at traffic stations had a positive effect by increasing the $r$ for NO$_2$ (from 0.73 to 0.76) and O$_3$ (from 0.83 to 0.86), and also by decreasing the NO$_2$ mean bias by ~5 µg m$^{-3}$ (from 6 to 1 µg m$^{-3}$). The NO$_2$ daily cycle improves in the morning hours after sunrise, reducing bias by 5–10 µg m$^{-3}$ and contributing to a reduction in O$_3$ overestimations (~5 µg m$^{-3}$). In contrast, at background stations, where the NO$_x$/O$_3$ chemistry is less dominant, the resolution effect is not significant. Such behavior indicates that a finer resolution improves the model performance, because horizontal resolution affects the representation of chemical processes near large emissions sources, such as the formation of O$_3$ and nighttime O$_3$ titration (Mathur et al., 2005).

5 Conclusions

The present work shows the effects of increasing the horizontal resolution from 4 km to 1 km using the CALIOPE-AQFS on pollutant concentrations (NO$_2$, O$_3$, and PM$_{10}$) over three Spanish domains (AND, BCN and MAD) in April 2013.

The global features of concentration maps at both resolutions are quite similar, with zones of high/low concentrations identically located, which is expected, since both simulations are based on the same emissions data set. Further comparisons demonstrate that increasing the resolution provides better-defined and more realistic concentration structures over large emission sources (roads and industries) and complex terrains (more sharply defined orographic hills). The titration effect on O$_3$ by NO$_x$ along highways and major point sources is more evident in 1-km simulations than at 4 km, since NO$_x$ concentrations tend to be higher in the 1 km simulation due to less dilution. This improvement is quantified by an increase in the spatial correlation coefficients of 3% (6%) for NO$_2$ (O$_3$).

However, the resolution increase in the range of 4 to 1 km does not exhibit the expected improvement in hourly statistics for any pollutant. Hourly correlation coefficients do not change significantly, and absolute (relative) errors and biases vary by less than 2 µg m$^{-3}$ (9%). The merit of the resolution increase may be underrated when classical statistics are applied at measurement stations (Mass et al., 2002; Gego et al., 2005). For instance, although the structure of important NO$_2$ urban plume features (>40 µg m$^{-3}$) often becomes
more realistic (stronger and more defined plumes) as resolution increases, statistics are deeply degraded by even small timing and spatial errors.

The resolution increase has a significant impact in reducing NO$_2$ hourly bias (by 42 %, 2 µg m$^{-3}$), without any significant change in the error and $r$ (< 2 %). However, O$_3$ hourly biases increased by less than 1 µg m$^{-3}$. The main differences between resolutions appear at daytime and nighttime traffic peaks, when the mixing height experiences rapid changes, allowing the 1 km simulation to reduce NO$_2$ underestimation slightly in the morning by ∼ 5–10 µg m$^{-3}$. The O$_3$ daily cycles at large emission sources depict a high influence of hourly NO$_2$ concentrations, increasing the hourly O$_3$ bias by ∼ 3 µg m$^{-3}$. This behavior is controlled by the daytime O$_3$ underestimation and, to a lesser extent, by the nighttime overestimation. The resolution increase reduces the O$_3$ overestimations at night by ∼ 5 µg m$^{-3}$, partly because of higher nocturnal NO$_2$ concentrations.

Increasing the horizontal resolution improves the model’s ability to forecast 75p exceedances in the observed maximum 1 h concentrations. The number of hits that forecast 75p exceedances in the observed Max 1 h NO$_2$ increases from 466 to 537 over 953 exceedances, while the FAR for Max 1 h O$_3$ exceedances is reduced by 37 %.

The grid effect is less pronounced for PM$_{10}$ than for NO$_2$ and O$_3$. When the resolution increases, the low gradient of PM$_{10}$ mean concentrations (< 0.1 µg m$^{-3}$) is the result of compensating for biases of PM$_{10}$ components, which is controlled mainly by the PPM and EC increase, as well as by the SS decrease.

BCN is the domain where the resolution increase has the largest effect, with changes in bias (error) of 16–73 % (< 5 %), followed by AND with 4–43 % (< 5 %) and MAD < 3–5 % (< 1 %). In BCN, as in the western Mediterranean basin, the transport of O$_3$ and its precursors is governed by mesoscale circulations. In that sense, the resolution increase has a large impact over BCN, where induced mesoscale phenomena control the air flow. Conversely, synoptic transport is more prominent in MAD and AND. Increasing the resolution to 1 km over rural areas (Mass et al., 2002) could contribute to improving the resolution to 1 km over rural areas (Mass et al., 2002) could contribute to improving the resolution to 1 km over urban areas, because a finer spatial structure is expected but unresolved. While the work presented here demonstrates the feasibility of modeling on fine scales, there remain many challenges. First, there is a loss of subgrid emissions heterogeneity. Emission inputs to CTM are an average rate, which accounts for the volume-averaged quantity of mass released per unit of time. No other information regarding emissions allocation (e.g., point, linear or per area) is considered; for instance, a large amount of mass can be emitted by a small portion of the grid surface or by several sources scattered around it (Galmarini et al., 2008; Cassiani et al., 2010; Ching and Majeed, 2012). Despite the fact that emissions are estimated by following a bottom-up approach emissions model, emissions heterogeneity is lost in the volume-averaging process performed within CTM. The loss is even higher when resolution decreases (from 1 km to 4 km). Second, there is a low degree of complexity in flow and dispersion details on urban scales, where most of the pollutants come from street canyons and/or tree canopies, and where they are transported until mixing conditions allow the pollutants to disperse above these urban canopy levels (Kim et al., 2013; Ching, 2013). Third, the USGS land-use data used in the WRF model are based on the 1993 data, and urban changes in MAD and BCN over the last 20 years are significant.

Since temperature and wind speed are very sensitive to the ratio of building width to road width, future improvement for fine-scale modeling should focus on using an urban canopy model that considers effects on the transfer of energy and momentum between urban structures and the lower atmosphere. This is crucial for modeling meteorology and air quality on fine scales in urban environments. However, it requires an urban canopy scheme and a canopy parameter database (urban fraction, building height and area). Furthermore, in order to gain any benefits from increasing resolution, the meteorological modeling should be updated to include a better description of land use, instead of relying on USGS data from the year 1993. To this end, the Coordination of Information on the Environment (CORINE) provides a high-resolution (100 m) land-use database, which was developed by the European Environmental Agency and updated to the year 2006 (CLC2006) (EEA, 2007). This could be implemented in the WRF model following the methodology described in Pineda et al. (2004).

www.geosci-model-dev.net/7/1979/2014/
Appendix A: Statistical parameters

Table A1. Definition of the discrete statistics used in the evaluation, where $C_m(x,t)$ and $C_o(x,t)$ are the modeled and observed concentrations at a location ($x$) and time ($t$); $N$ is the number of pairs of data. $C_m$ and $C_o$ are the modeled and observed mean concentrations over the whole period, respectively.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient</td>
<td>$r = \frac{\sum_{i=1}^{N} (C_m(x,t) - \bar{C}_m)(C_o(x,t) - \bar{C}<em>o)}{\sqrt{\sum</em>{i=1}^{N} (C_m(x,t) - \bar{C}<em>m)^2 \sum</em>{i=1}^{N} (C_o(x,t) - \bar{C}_o)^2}}$ (A1)</td>
</tr>
<tr>
<td>Mean bias MB</td>
<td>$MB = \frac{1}{N} \sum_{i=1}^{N} (C_m(x,t) - C_o(x,t))$ (A2)</td>
</tr>
<tr>
<td>Mean normalized bias error</td>
<td>$MNBE = \frac{1}{N} \sum_{i=1}^{N} \frac{(C_m(x,t) - C_o(x,t))}{C_o(x,t)} \cdot 100$ (A3)</td>
</tr>
<tr>
<td>Mean fractional bias</td>
<td>$MFB = \frac{1}{N} \sum_{i=1}^{N} \frac{(C_m(x,t) - C_o(x,t))}{(C_m(x,t) + C_o(x,t))/2} \cdot 100$ (A4)</td>
</tr>
<tr>
<td>Mean absolute error MAE</td>
<td>$MAE = \frac{1}{N} \sum_{i=1}^{N}</td>
</tr>
<tr>
<td>Mean normalized gross error</td>
<td>$MNGE = \frac{1}{N} \sum_{i=1}^{N} \frac{</td>
</tr>
<tr>
<td>Mean fractional error MFE</td>
<td>$MFE = \frac{1}{N} \sum_{i=1}^{N} \frac{</td>
</tr>
<tr>
<td>Root mean squared error rmse</td>
<td>$\sqrt{\frac{1}{N} \sum_{i=1}^{N} (C_m(x,t) - C_o(x,t))^2}$ (A8)</td>
</tr>
</tbody>
</table>

Table A2. Definition of the categorical statistics used in the evaluation. Exceedance analysis is based on a comparison with a fixed threshold concentration ($T$), where $a$ is the number of false alarms, $b$ is the number of hits, $c$ is the number of correct negatives, and $d$ is the number of misses.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy A</td>
<td>$A = \left( \frac{b+c}{a+b+c+d} \right) \cdot 100$ (A9)</td>
</tr>
<tr>
<td>Bias B</td>
<td>$B = \left( \frac{a+c}{a+b+c+d} \right) \cdot 100$ (A10)</td>
</tr>
<tr>
<td>Critical success index CSI</td>
<td>$CSI = \left( \frac{b}{a+b+c+d} \right) \cdot 100$ (A11)</td>
</tr>
<tr>
<td>Probability of detection POD</td>
<td>$POD = \left( \frac{b}{a+b} \right) \cdot 100$ (A12)</td>
</tr>
<tr>
<td>False alarm ratio FAR</td>
<td>$FAR = \left( \frac{a}{a+b} \right) \cdot 100$ (A13)</td>
</tr>
</tbody>
</table>
The Supplement related to this article is available online at doi:10.5194/gmd-7-1979-2014-supplement.

Acknowledgements. Spanish administrations “Generalitat de Catalunya”, “Junta de Andalucía”, and “Comunidad de Madrid” are acknowledged for providing air quality measurements. The CALIOPE-AQFS team (G. Arévalo, K. Serradell, D. Carrió, M. Castrillo, A. Soret, S. Basart and S. Gassó) and F. Benincasa are also thanked for their technical support. This work is funded by the post-doctoral grant held by M. T. Pay in the Beatríu de Pinós program (2011 BP-A 00427), the Andalusian contract (NET838690), and the Severo Ochoa program awarded by the Spanish Government (SEV-2011-00067).

Edited by: A. Lauer

References


