Evaluation of WRF/Chem model (v3.9.1.1) real-time air quality forecasts over the Eastern Mediterranean

George K. Georgiou1, Theodoros Christoudias1, Yiannis Proestos1, Jonilda Kushta1, Michael Pikridas1, Jean Sciare1, Chrysanthos Savvides2, and Jos Lelieveld3,1

1Climate & Atmosphere Research Centre, The Cyprus Institute, Nicosia, Cyprus
2Department of Labour Inspection, Ministry of Labour, Welfare and Social Insurance, Nicosia, Cyprus
3Atmospheric Chemistry Department, Max Planck Institute for Chemistry, Mainz, Germany

Correspondence to: George K. Georgiou (g.georgiou@cyi.ac.cy)

Abstract.

We describe and evaluate a high-resolution real-time air quality forecast system over the Eastern Mediterranean, based on a regional, on-line coupled atmospheric chemistry and aerosol model. The WRF/Chem model is used to perform daily, 3-day forecasts of regulated pollutants (NO2, O3, PM2.5) over the Eastern Mediterranean, with three nested domains with horizontal resolutions of 50km, 10km and 2km, focusing on Cyprus. Natural (dust, sea-salt, biogenic) emissions are calculated online, while anthropogenic emissions are based on the EDGAR-HTAP global emission inventory. A high spatial (1km) and temporal (hourly) anthropogenic emission inventory is used for the island of Cyprus in the innermost domain. The model skill in forecasting the concentrations of atmospheric pollutants is evaluated using measurements from a network of nine ground stations in Cyprus and compared with the forecast skill of the EU Copernicus Atmosphere Monitoring Service - CAMS.

The forecast of surface temperature, pressure, and wind speed is found to be accurate, with minor discrepancies between the modelled and observed 10m wind speed at mountainous and coastal sites attributed to the limited representation of the complex topography of Cyprus. Compared to CAMS, the WRF/Chem model predicts with higher accuracy the NO2 mixing ratios at the residential site with a normalized mean bias of 7% during winter and -44% during summer, whereas the corresponding biases for CAMS are -81% and -84%. Due to the high temporal resolution of the anthropogenic emission inventory, the WRF/Chem model captures more accurately the diurnal profiles of NO2 and O3 mixing ratios at the residential site. Background PM2.5 concentrations influenced by long-range transport are overestimated by the WRF/Chem model during winter (NMB = 54%) whereas the corresponding NMB for CAMS is 11%. Our results support the adoption of regional, on-line coupled air quality models over chemical transport models for real-time air quality forecasts.

1 Introduction

The term air quality is used to describe to what extent the troposphere is contaminated with atmospheric pollutants. High concentrations of atmospheric pollutants can be hazardous to human health. The atmospheric pollutants with the strongest evidence for public health concern, include ozone (O3), nitrogen dioxide (NO2), and particulate matter (PM) (World Health Organization, 2018, (visited on 2020-01-19). Tropospheric ozone is linked to numerous harmful health effects including reduced lung...
function, increased frequency of respiratory symptoms, and development of asthma (Lippmann, 1989; Broeckaert et al., 1999; Brunekreef and Holgate, 2002), while NO₂ is associated with the development of emphysema-like lesions (Wegmann et al., 2005). Human exposure to particulate matter is associated with acute cardiovascular events and atherosclerosis while they also affect the cardiovascular system directly by entering into the systemic circulation (Kampa and Castanas, 2008).

The Eastern Mediterranean and the Middle East (EMME) region is characterized by high background tropospheric ozone concentrations (Lelieveld et al., 2002, 2009; Zanis et al., 2014), since it is affected by polluted air masses from various sources such as the eastern and central Europe, and the Middle East (Lelieveld et al., 2002; Gerasopoulos et al., 2005; Ladstätter-Weißenmayer et al., 2007; Kalabokas et al., 2008; Kanakidou et al., 2011). The island of Cyprus is located in the eastern part of the Mediterranean Sea, adjacent to the Middle East and North Africa. Cyprus is one of the countries in the EMME region that faces challenges with the exceedance of air quality limits and compliance with European regulatory standards. For the year 2017, Cyprus reported the second highest mean annual NO₂ concentrations (13.7 µg/m³) among 41 European countries (European Environment Agency, 2019). The exposure in high NO₂ concentrations in Cyprus is estimated to cause about 240 premature deaths per year. This translates to 0.02% of the population which is among the highest between the 41 European countries (European Environment Agency, 2019). In addition, exposure to PM₂.₅ is responsible for 473 years of life lost (YLL) per 100,000 inhabitants in Cyprus (European Environment Agency, 2019).

The effects of increased concentrations of air pollution on human health highlight the need for real-time air quality forecasting (RT-AQF) with detail in space and time. RT-AQFs can provide the environmental authorities and the general public with information and warning in advance in order to make informed decisions and take actions that will better protect the population from imminent air pollution episodes.

RT-AQF over the Eastern Mediterranean and Cyprus is provided by the Whole Atmosphere Community Climate Model - WACCM (Gettelman et al., 2019), and the Copernicus Atmosphere Monitoring Service (CAMS). WACCM is a global Chemical Transport Model (CTM) driven by meteorological fields from the Goddard Earth Observing System, Version 5 (GEOS-5) model. The model provides daily 10-day air quality forecasts on a horizontal resolution of 0.9° × 1.25° and 6-hour time-step starting at 00:00 TC. CAMS provides daily 4-day-ahead air quality forecasts over Europe on a horizontal resolution of 0.1° and 1-hour time-step, based on an ensemble of 9 state-of-the-art numerical air quality models developed in Europe: CHIMERE from INERIS (France), EMEP from MET Norway (Norway), EURAD-1M from Jülich IEK (Germany), LOTOS-EUROS from KNMI and TNO (Netherlands), MATCH from SMHI (Sweden), MOCAGE from METEO-FRANCE (France), SILAM from FMI (Finland), DEHM from Aarhus University (Denmark), and GEM-AQ from IEP-NRI (Poland).

In this work we describe and evaluate a high-resolution RT-AQF system established in the Eastern Mediterranean and Cyprus, based on a regional, on-line coupled air quality model. We evaluate the skill of the RT-AQF system to forecast the atmospheric concentrations of NO₂, O₃, and PM₂.₅, which are the three regulated by the European Union atmospheric pollutants with the strongest evidence for their effects on human health. Regional air quality models are able to run in very high horizontal resolutions, down to the convection permitting resolution limit of the order of 1 km. This allows for more accurate representation of the topography and the population density, and therefore anthropogenic emissions, which is important in a small country like Cyprus with steep changes in altitude, and urban centres with radial extent of less than 10km.
The manuscript is structured as follows: In Section 2 we describe the domain set-up, the model configuration, and the model input data. We examine the skillfulness of the WRF/Chem model to forecast the basic meteorological parameters (Section 3.1), the concentrations of NO$_2$ (Section 3.2), O$_3$ (Section 3.3), and PM2.5 (Section 3.4). Our conclusions are given in Section 4.

2 WRF-Chem model and observations

2.1 Model configuration

We use the Weather Research and Forecasting model with Chemistry (WRF/Chem) version 3.9.1.1 to perform daily 3-day-ahead meteorological and air quality forecasts for the three winter (January, February, December) and three summer (June, July, August) months of 2020. We use three one-way nested domains (Fig. 1) with horizontal resolutions of 50 km, 10 km, and 2 km. The outermost domain (d01) includes the Black Sea region, the largest part of Europe, and the Middle East and North Africa deserts which have an important contribution to the background concentrations of gas-phase and aerosol pollutants over the EMME region. The second domain (d02) focuses over the Eastern Mediterranean and includes the major urban centres in the Middle East. The third domain (d03) is focused over the island of Cyprus. We use 33 vertical layers, while adaptive time-stepping is used in order to meet the Courant-Friedrichs-Lewy (CFL) stability criterion at each time-step (Jacobson, 2005) and to reduce the simulation times.

Figure 1. Model simulation domains

Georgiou et al. (2017) showed that the Regional Acid Deposition Model, version 2 (RADM2) gas-phase chemistry mechanism produces the lowest mean bias when simulating the CO, NO$_x$, and O$_3$ concentrations over Cyprus compared to other mechanisms. In this set-up we use the Regional Atmospheric Chemistry Mechanism (RACM) (Stockwell et al., 1997) which is an updated version of RADM2. Aerosol inorganic species are simulated using the Modal Aerosol Dynamics Model for Europe (MADE) (Ackermann et al., 1998), while the secondary organic aerosols parameterization based on the volatility basis set (VBS) by Ahmadov et al. (2012) is used for secondary organic aerosols. Detailed description of the model physics configuration is given in Table 1.
Table 1. Gas-phase chemistry, aerosols, and physics parameterizations used in the simulations.

<table>
<thead>
<tr>
<th>Process</th>
<th>Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas-phase chemistry</td>
<td>RACM (Stockwell et al., 1997)</td>
</tr>
<tr>
<td>Aerosols</td>
<td>MADE/VBS (Ackermann et al., 1998; Ahmadov et al., 2012)</td>
</tr>
<tr>
<td>Cloud microphysics</td>
<td>Morrison double moment (Morrison et al., 2005)</td>
</tr>
<tr>
<td>Longwave &amp; Shortwave radiation</td>
<td>RRTMG (Mlawer et al., 1997)</td>
</tr>
<tr>
<td>Cumulus parameterization</td>
<td>Grell 3D (Grell, 2002)</td>
</tr>
<tr>
<td>Photolysis</td>
<td>Fast-J</td>
</tr>
<tr>
<td>Land-surface physics</td>
<td>Noah Land Surface Model (Chen and Dudhia, 2001)</td>
</tr>
<tr>
<td>Planetary Boundary Layer</td>
<td>Yonsei University (Hong et al., 2006)</td>
</tr>
</tbody>
</table>

Initial and boundary conditions for meteorology are provided by the Global Forecast System (GFS) every six hours on a horizontal grid resolution of 0.25° × 0.25° (06:00 initialization time). Boundary conditions for the gas-phase species and aerosols are provided by the Whole Atmosphere Community Climate Model (WACCM; Gettelman et al. (2019)). The WACCM model output datasets are available on a horizontal grid resolution of 1° × 1° and interpolated in space every six hours to our model domain. The dust component from the boundary conditions is not taken into account in our simulations, since the Middle Eastern and North Africa deserts are included in our model domain.

Biogenic emissions are generated on-line by the WRF/Chem model based on weather and land use data, using the Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1) by Guenther et al. (2012). Dust emissions are simulated using the Georgia Tech/Goddard Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) model (Ginoux et al., 2001). The dust emission flux \( F_p \) in the GOCART model is calculated as:

\[
F_p = CSs_p u_{10m}^2 (u_{10m} - u_t)
\]  

(1)

where \( S \) is a source function which defines the potential dust source regions, \( s_p \) is the fraction of each size class of dust, \( u_{10m} \) is the horizontal wind speed at 10m, and \( u_t \) is the threshold velocity above which dust emission occur. \( C \) is an empirical proportionality constant originally set equal to 1 \( \mu g s^{-2} m^{-5} \). The value of \( C \) estimated by Ginoux et al. (2001) was initially based on regional data over North America. Zhao et al. (2010) evaluated the performance of the WRF/Chem model for different values of \( C \). They found that for \( C = 0.4 \mu g s^{-2} m^{-5} \), the WRF/Chem simulated mean AOD over the Sahel region was consistent with measurements from the Dust and Biomass burning Experiment (DABEX) campaign. Therefore, as that is the most prevalent source of dust emissions in the EMME region, in the following simulations, a value of \( C \) equal to 0.4 \( \mu g s^{-2} m^{-5} \) is used.

Anthropogenic emissions for the first and second domain of the simulations are based on the EDGAR-HTAP Version 2 emission inventory (LRTAP-Wiki, 2014). Anthropogenic emissions for the first and second domain of the simulations are based on the EDGAR-HTAP Version 2 emission inventory (LRTAP-Wiki, 2014) and were interpolated in time and space to produce daily emissions for the first and second domain using the anthro_emiss utility (Kumar, 2018, visited on 2020-01-19). For the innermost domain a high-resolution emission inventory developed by Georgiou et al. (2020) is used. This emission inventory
uses the total reported emissions of CO, NOx, NMVOC, SO2, and PM for the year 2013 on a 1 km×1 km resolution which is upscaled to the resolution of the innermost domain of the simulations (2 km) using a nearest-neighbour grid-point attribution algorithm, while diurnal, weekly, and monthly emission cycles are applied to each species according to Schaap et al. (2005) and the predominant emission activity per season. There are two operational power generation stations in the southern part of Cyprus NOx emissions from these two stations account for about 27% of the total NOx emitted from the part of the island which controlled by the Republic of Cyprus. At these locations, the emission factors for power generation are applied for all species. Two additional power generation stations, one of them located very close to the city of Nicosia, are operational in the northern part of the island and are not included in the high-resolution emission inventory. For these areas, the emission inventory takes into account the emissions from the EDGAR-HTAP global emission inventory, resulting in underestimation of the total NOx emissions. Emissions from road transport dominate the NOx emissions (47%) while there is also important contribution from industrial processes (19%). Industrial processes are the main local source of PM (38%) followed by agriculture (19%). Since 2013, no important changes were observed in the NOx and PM2.5 total emissions. Specifically, PM2.5 emissions increased from 0.97 kt in 2013 to 1.0t kt in 2019, whereas total NOx emissions decreased from 15.36 kt in 2013 to 14.04 kt in 2019. Georgiou et al. (2020) showed that using the updated emission inventory in hindcasting mode resulted in reduction of the normalized mean bias between the modelled and observed NOx mixing ratios at the residential sites (from 67% to 29% and from 51% to 10% for the winter and summer, respectively). In line with this, the overestimation in O3 mixing ratios was reduced from 45% to 28% during the winter and from 25% to 19% during summer. Finally, taking into account the diurnal variability in the emission inventory was found to be crucial for the simulation of the daily profiles of NOx and O3 at residential sites.

2.2 Observational data

The modelled concentrations of the air pollutants from the 1st day of forecast is compared against hourly observational data from nine air quality monitoring ground stations, provided by the Cyprus Department of Labour Inspection (DLI) for the three winter (January, February, December) and three summer (June, July, August) months of 2020. During these periods, there were no restrictions in place due to the COVID-19 pandemic. The station network consists of background (Figure 2, green circle), residential (Figure 2, cyan circle), traffic (Figure 2, yellow circles), and industrial (Figure 2, red circles) stations which span the southern part of the island of Cyprus and apart from the concentrations of the air pollutants, also measure various meteorological variables. The characteristics and the monitored pollutants of each station are shown in Table 2.

3 Results and discussion

3.1 Meteorology

The model skill in forecasting basic meteorological variables is evaluated by comparing the first day of forecast from the WRF/Chem model to hourly measurements from the nine monitoring stations. Table 3 shows the Pearson’s correlation coeffi-
Figure 2. Left: Air quality monitoring stations locations (from Google Maps). The two power generation stations locations are also shown across the south coastline of Cyprus. Right: Terrain elevation and the locations of the Ayia Marina Background station and the city of Nicosia. The red circle indicates the urban area limits.

Table 2. Monitoring stations.

<table>
<thead>
<tr>
<th>Monitoring site</th>
<th>Abbreviation</th>
<th>Type of zone</th>
<th>Alt. (m)</th>
<th>Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayia Marina</td>
<td>AYMBGR</td>
<td>Background</td>
<td>532</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Larnaca Traffic</td>
<td>LATRA</td>
<td>Traffic</td>
<td>15</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Limassol Traffic</td>
<td>LIMTRA</td>
<td>Traffic</td>
<td>19</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Mari Industrial</td>
<td>MARIND</td>
<td>Industrial</td>
<td>88</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Nicosia Residential</td>
<td>NICRES</td>
<td>Residential</td>
<td>208</td>
<td>T2, WS10, PSFC, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Nicosia Traffic</td>
<td>NICTRA</td>
<td>Traffic</td>
<td>176</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Paphos Traffic</td>
<td>PAFTRA</td>
<td>Traffic</td>
<td>40</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM10</td>
</tr>
<tr>
<td>Paralimni Traffic</td>
<td>PARTRA</td>
<td>Traffic</td>
<td>72</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
<tr>
<td>Zygi Industrial</td>
<td>ZYGIND</td>
<td>Industrial</td>
<td>9</td>
<td>T2, WS10, PSFC, CO, NO2, O3, SO2, PM2.5, PM10</td>
</tr>
</tbody>
</table>

cient (R), mean bias (MB), normalized mean bias (NMB), and root mean squared error (RMSE) for temperature at 2 m ($T_{2m}$), wind speed at 10 m ($WS_{10}$), and surface pressure ($P_{surf}$) averaged over all stations. Modelled $T_{2m}$ is in good agreement with observations (NMB < 11% for both winter and summer). The diurnal cycle of $T_{2m}$ is also reproduced by the model (R > 0.78) during both seasons. Modelled $P_{surf}$ is in very good agreement with observations with a normalized mean bias of less than 1%. The model tends to overestimate $WS_{10}$ by an average of 2.33 m/s during winter and 1.88 m/s during summer. Over-predictions of $WS_{10}$ by the WRF model over the Mediterranean have also been reported in previous studies (Zhang et al., 2021).
Table 3. Pearson’s correlation coefficient (R), mean bias (MB), normalized mean bias (NMB), and root mean squared error (RMSE) of hourly values of temperature at 2 m, wind speed at 10 m, and surface pressure averaged over all stations.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2m (°C)</td>
<td>Psurf (hPa)</td>
<td>WS10 (ms⁻¹)</td>
</tr>
<tr>
<td>R</td>
<td>0.80</td>
<td>0.61</td>
</tr>
<tr>
<td>MB</td>
<td>1.26</td>
<td>3.19</td>
</tr>
<tr>
<td>NMB</td>
<td>0.11</td>
<td>0.00</td>
</tr>
<tr>
<td>RMSE</td>
<td>2.94</td>
<td>6.63</td>
</tr>
</tbody>
</table>

2013; Mar et al., 2016; Georgiou et al., 2017). These biases are mainly attributed to the poor representation of surface drag exerted by the unresolved topography (Zhang et al., 2013).

3.2 Nitrogen Dioxide (NO₂)

The seasonal average NO₂ mixing ratios from the observations and the 1st day of the WRF/Chem forecast for winter and summer are shown in Figure 3, 1st row. During both seasons, the higher NO₂ mixing ratios appear near the urban centres and the power generation stations (Figure 2). NO₂ emitted within the island is shown to affect the eastern part of Cyprus through the prevailing westerly winds which is in agreement with Georgiou et al. (2017).

During both periods, the WRF/Chem model forecasts accurately the background NO₂ mixing ratios with a mean bias of less than 1 ppbV at the Ayia Marina background station (Figure 4, 1st row). Similar performance is achieved by the CAMS model. At the Nicosia residential station, the WRF/Chem model outperforms CAMS during winter and summer. The normalized mean bias from the WRF/Chem model is found to be -7% (about 1 ppbV) during winter and -44% (about 3 ppbV) during summer, whereas the corresponding values from CAMS are -81% (about 14 ppbV) and -84% (about 6 ppbV). Underestimation by both WRF/Chem and CAMS is more evident at the traffic stations. The normalized mean bias from the WRF/Chem model is found to be -31% (about 5 ppbV) during winter and -39% (about 4 ppbV) during summer, whereas the corresponding values for CAMS are -86% (about 13 ppbV) and -78% (about 7 ppbV). These stations, due to their proximity to main traffic roads, often record very high concentrations of pollutants (as shown by the large number of outliers and the large standard deviation), which cannot be reproduced by the atmospheric model. At the industrial stations, the WRF/Chem model tends to overestimate the seasonal average NO₂ mixing ratios by about 4 ppbV during the winter and 7 ppbV during the summer. The biases at these locations can be partly attributed to the fact that in the model, atmospheric pollutants are emitted at the surface, while actual emissions occur at the height of the chimneys which is about 70m above ground. On the other hand, CAMS underestimates NO₂ mixing ratios by about 4 ppbV and 5 ppbV during summer and winter respectively. The statistical metrics for NO₂, as well as O₃ and PM2.5 at the background, residential, traffic, and industrial stations are summarized on Table 4. Detailed metrics for each station are given on Table S1 in the Supplement. Similar results were obtained for the second and third day of forecast as shown on Table S2 in the Supplement. This is attributed to the dependency of the model performance on the emissions.
Table 4. Pearson’s correlation coefficient (R), mean bias (MB), normalized mean bias (NMB), and root mean squared error (RMSE) of hourly values of nitrogen dioxide (NO₂), ozone (O₃), and fine particulate matter (PM2.5) averaged over the background, residential, traffic, and industrial stations during winter and summer for the first day of forecast.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th></th>
<th></th>
<th></th>
<th>Summer</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WRF/Chem</td>
<td>CAMS</td>
<td>WRF/Chem</td>
<td>CAMS</td>
<td>WRF/Chem</td>
<td>CAMS</td>
<td>WRF/Chem</td>
<td>CAMS</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>MB</td>
<td>NMB</td>
<td>RMSE</td>
<td>R</td>
<td>MB</td>
<td>NMB</td>
<td>RMSE</td>
</tr>
<tr>
<td>NO₂ (ppbV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>0.12</td>
<td>-0.62</td>
<td>0.36</td>
<td>2.18</td>
<td>0.39</td>
<td>-0.89</td>
<td>-0.52</td>
<td>1.25</td>
</tr>
<tr>
<td>Residential</td>
<td>0.55</td>
<td>-1.13</td>
<td>-0.07</td>
<td>11.11</td>
<td>0.59</td>
<td>-14.01</td>
<td>-0.81</td>
<td>17.27</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.36</td>
<td>-4.87</td>
<td>-0.31</td>
<td>12.11</td>
<td>0.46</td>
<td>-12.83</td>
<td>-0.86</td>
<td>16.22</td>
</tr>
<tr>
<td>Industrial</td>
<td>0.14</td>
<td>4.22</td>
<td>0.77</td>
<td>12.10</td>
<td>0.32</td>
<td>-3.68</td>
<td>-0.67</td>
<td>5.46</td>
</tr>
<tr>
<td>O₃ (ppbV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>0.16</td>
<td>2.66</td>
<td>0.07</td>
<td>10.18</td>
<td>0.44</td>
<td>-3.09</td>
<td>-0.08</td>
<td>6.57</td>
</tr>
<tr>
<td>Residential</td>
<td>0.49</td>
<td>9.91</td>
<td>0.48</td>
<td>16.55</td>
<td>0.65</td>
<td>14.41</td>
<td>0.70</td>
<td>17.37</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.35</td>
<td>12.72</td>
<td>0.61</td>
<td>18.08</td>
<td>0.53</td>
<td>15.05</td>
<td>0.73</td>
<td>18.05</td>
</tr>
<tr>
<td>Industrial</td>
<td>0.04</td>
<td>5.56</td>
<td>0.18</td>
<td>14.63</td>
<td>0.50</td>
<td>8.15</td>
<td>0.27</td>
<td>10.95</td>
</tr>
<tr>
<td>PM2.5 (µg/m³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>0.27</td>
<td>3.87</td>
<td>0.54</td>
<td>10.15</td>
<td>0.48</td>
<td>0.76</td>
<td>0.11</td>
<td>4.47</td>
</tr>
<tr>
<td>Residential</td>
<td>0.19</td>
<td>5.01</td>
<td>0.28</td>
<td>15.19</td>
<td>0.32</td>
<td>8.76</td>
<td>0.50</td>
<td>15.02</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.29</td>
<td>1.82</td>
<td>0.17</td>
<td>9.69</td>
<td>0.48</td>
<td>-1.76</td>
<td>-0.17</td>
<td>5.66</td>
</tr>
</tbody>
</table>

Although showing slightly higher Pearson’s correlation coefficients between the observed and forecasted NO₂ mixing ratios at some of the stations, as shown by the Taylor diagrams (Figure 5, 1st row), CAMS does not capture the diurnal profile of the wintertime NO₂ mixing ratios at the locations with intense anthropogenic activity (Figure 6, 1st row). In particular, the forecasted NO₂ mixing ratios show very small fluctuations throughout the day, with slight increases during the morning and afternoon hours at the Nicosia residential station and the Larnaca, Limassol, and Nicosia traffic stations. In contrast, the WRF/Chem model captures the peaks which appear in the observations at the residential and traffic stations during the morning and afternoon. These peaks in NO₂ mixing ratios appear at the same time with the peaks in the NOₓ emissions that have been applied for Cyprus (Georgiou et al., 2020). Similar results are obtained for the summer period regarding the skill of the models to forecast the diurnal profile of NO₂. During the summer, at the residential and traffic stations, lower mixing ratios appear throughout the day, although the NOₓ emissions diurnal profile is similar to the winter, with higher emissions during the morning (Georgiou et al., 2020). This can be partly attributed to the boundary layer height and the intense photochemical activity during the summer. These two factors result in a weaker peak in NO₂ mixing ratios at the residential station and the majority of the traffic stations. These patterns are resembled by the WRF/Chem model, which highlights the ability of the model to simulate these characteristics of the concentrations of the atmospheric pollutants near the areas with human presence. The correlation coefficients at the background station are low which is partly attributed to the absence of nearby emission sources. The absence of emissions results in a weak diurnal profile of the concentrations of atmospheric pollutants. Therefore, even small fluctuations in the observed concentrations which are not captured by the model, result in low correlation between the observed and modelled concentrations. The summertime diurnal profiles of the NO₂ mixing ratios are shown in Figure S1 in the supplement.
Figure 3. Average NO$_2$ (1st row) and O$_3$ (2nd row) mixing ratios, and PM$_{2.5}$ (3rd row) concentrations from the 1st day of the WRF/Chem forecast during winter (left) and summer (right). The filled dots indicate the observed values.
Figure 4. Box-and-whisker plots at the background, residential, traffic, and industrial NO2 (1st row), O3 (2nd row), and PM2.5 (3rd row) average mixing ratios in observed data (grey color) and the WRF/Chem (blue color) and CAMS (red color) forecasts during winter (left) and summer (right). The seasonal average mixing ratios are derived from hourly mixing ratios.
Figure 5. Taylor diagrams of NO$_2$ (1st row), O$_3$ (2nd row), and PM2.5 (3rd row) for winter (left) and summer (right).
Figure 6. Diurnal variation of NO$_2$ (1st row), O$_3$ (2nd row), and PM2.5 (3rd row) in observations (grey lines), the WRF/Chem (blue lines) and the CAMS (red lines) forecasts during winter.
3.3 Ozone (O\textsubscript{3})

Observed background O\textsubscript{3} mixing ratios reach up to 41 ppbV during winter and 53 ppbV during summer which is in agreement with Gerasopoulos et al. (2005). In their study, using observational data from 1997 to 2004, a maximum in O\textsubscript{3} mixing ratios (58 ± 10 ppbV) in July and a minimum (36 ± 7 ppbV) in December were reported over the Eastern Mediterranean. During both seasons, the lowest O\textsubscript{3} mixing ratios appear at the locations with intense anthropogenic activity, such as the locations of the power generation stations and the urban centers, and the eastern part of the island (Figure 3, 1st row), which coincide with the highest NO\textsubscript{2} mixing ratios appear (Figure 3, 2nd row).

As shown in Figure 4, 2nd row, wintertime background O\textsubscript{3} mixing ratios are captured by both WRF/Chem (NMB = 7\%) and CAMS (NMB = -8\%). Summertime background O\textsubscript{3} mixing ratios are also captured by the WRF/Chem (NMB = 5\%), but underestimated by CAMS by about 10 ppbV (NMB = -19\%). At the residential and traffic stations, CAMS strongly overestimates O\textsubscript{3} mixing ratios during winter by 70\% ( ~14 ppbV) and 73\% ( ~15 ppbV) respectively. Better performance at these locations is achieved by the WRF/Chem model where overestimation is about 48\% ( ~10 ppbV) at the residential and 61\% ( ~13 ppbV) at the traffic stations. The overestimation in O\textsubscript{3} mixing ratios by both models can be partly attributed at the underestimation of NO\textsubscript{2} mixing ratios seen in Sect. 3.2. During the summer, CAMS underestimates background O\textsubscript{3} mixing ratios by about 10 ppbV (NMB = -19\%). This results in lower normalized mean bias at the residential and traffic stations (-2\% and 17\% respectively) compared to WRF/Chem (15\% and 29\% respectively).

Due to intense photochemical activity and high O\textsubscript{3} production, observed O\textsubscript{3} mixing ratios show a maximum in the afternoon at the locations with anthropogenic activity. This maximum is successfully forecasted by both WRF/Chem and CAMS (Figure 6, 2nd row). The WRF/Chem model though, is able to forecast with more accuracy the decreases in O\textsubscript{3} mixing ratios during the morning and evening hours. These decreases are a result of the increases in NO\textsubscript{2} mixing ratios (Sec. 3.2), which are sufficiently forecasted by the WRF/Chem but not by the CAMS model.

3.3.1 Ozone daily maximum 8-hour average

Figure 7 shows the O\textsubscript{3} daily maximum 8-hour average at the Ayia Marina background station and the urban stations in observation data, the WRF/Chem and the CAMS forecasts during the summer. During this period, 35 exceedances of the limits set by European Air Quality Directives (EU, 2008, (visited 2022-01-19) have been observed at the Ayia Marina background station, highlighting the effect of the high background concentrations on air quality over the Eastern Mediterranean and Cyprus. In the same period 22 exceedances have been successfully predicted by both WRF/Chem and CAMS (Figure 6, 2nd row). The WRF/Chem model though, is able to forecast with more accuracy the decreases in O\textsubscript{3} mixing ratios during the morning and evening hours. These decreases are a result of the increases in NO\textsubscript{2} mixing ratios (Sec. 3.2), which are sufficiently forecasted by the WRF/Chem but not by the CAMS model.

The exceedances of the limit of 60 ppbV at the Nicosia traffic station reached up to 10 during the summer period, while at the other traffic stations, no more than 2 exceedances have been observed. The lower number of exceedances at the traffic stations can be attributed to the O\textsubscript{3} titration by the locally emitted NO. At the Nicosia residential station, 20 exceedances have been
observed. Seven exceedances have been successfully predicted by the WRF/Chem model and only two by CAMS. In addition, the WRF/Chem model predicted exceedances that were not shown in observations, as a result of the 15% overestimation in O$_3$ concentrations (Sec. 3.3). During the winter, no exceedances have been observed or predicted due to the lower background O$_3$ concentrations over the region.
Figure 7. O$_3$ daily maximum 8-hour average at the Ayia Marina background station and the urban stations in observation data (black dots), the WRF/Chem (blue lines) and the CAMS (red lines) forecasts during summer. The green line indicates the limit value set by the European Air Quality Directives (EU, 2008, (visited 2022-01-19).
3.4 Fine Particulate Matter (PM2.5)

During summer, observed background PM2.5 concentrations are higher (11.2 µg/m³) than winter (7.2 µg/m³) which can be partly attributed to the absence of precipitation and the enhanced photochemical conditions which lead to secondary aerosol formation during the summer (Pikridas et al., 2018). In addition, the prevalent Etesian winds transport air masses rich in fine mode anthropogenic pollutants from Turkey or to a lesser extent from Europe (Pikridas et al., 2018). Anthropogenic emissions from Turkey and Central Europe were also found by Sciare et al. (2002) to have an important contribution to PM levels over the EMME region. Furthermore, periods with increased levels of non-sea-salt calcium during the summer reported in the study which were associated with air masses Africa and Central Turkey, highlighting the important contribution of background PM2.5 concentrations during the summer.

As can be seen from Figure 4, 3rd row, the WRF/Chem model overestimates background PM2.5 concentrations during winter by about 4 µg/m³, while CAMS forecasts wintertime background PM2.5 concentrations with more accuracy (≤ 1 µg/m³). This overestimation is mainly due to increased PM2.5 concentrations during the 3-day period from December 12th to December 14th which appear at all stations (Figure 8).

At the traffic stations, WRF/Chem and CAMS underestimate PM2.5 concentrations by 5 and 9 µg/m³ respectively. During the summer, the two models show similar behaviour. In particular, there is an underestimation of background PM2.5 concentrations of about 16% (~1.8 µg/m³). Underestimation is slightly higher at the industrial stations (~ 20%). At the traffic stations underestimation reaches up to 5 µg/m³ (~ 33%). The underestimation of PM2.5 concentrations at the traffic stations can in part be attributed to the lack of road dust re-suspension mechanisms in the models and the fact that these stations are located close to main roads. The increased observed PM2.5 concentrations during the evening hours at these stations (Figure 6) are attributed to emission sources from household heating. These sources are not included in detail in the emission inventory resulting to the underestimation of PM2.5 concentrations during these hours. The low correlation between the modelled and observed background concentrations is partly attributed to the weak diurnal profile of the PM2.5 concentrations and the absence of nearby emission sources.
Figure 8. Observed hourly PM2.5 concentrations (black dots) and daily average PM2.5 concentrations from observations (black lines), WRF/Chem (blue lines), and CAMS (red lines) during winter.
4 Conclusions

The WRF/Chem model has been applied for high-resolution, daily, 3-day-ahead air quality forecasts over the EMME region and Cyprus. The model skill in forecasting the atmospheric concentrations of NO\(_2\), O\(_3\), and PM2.5, as well as the basic meteorological parameters, has been evaluated for three winter (January, February, December) and three summer (June, July, August) months during 2020. The forecast output is compared against observational data from nine ground stations in Cyprus and forecast data from the state-of-the-art CAMS model. The magnitude and diurnal variation of surface pressure and temperature are accurately forecasted. Wind speed at 10m is overestimated by about 2 m/s which may be attributed to the limited representation of the complex topographical features by the model.

WRF/Chem and CAMS are able to forecast with accuracy the background NO\(_2\) mixing ratios during both seasons. NO\(_2\) mixing ratios are strongly overestimated by WRF/Chem at the industrial stations which can be partly attributed to the fact that emissions in the model occur at the surface while actual emissions at these locations occur at a height of about 70m. At the residential and traffic stations, WRF/Chem forecasts the magnitude and diurnal variation the NO\(_2\) mixing ratios with higher accuracy than CAMS, capturing the morning and afternoon peaks during winter and the morning peak during summer. Consequently, the decreases in O\(_3\) mixing ratios caused by the increases in NO\(_2\) mixing ratios are successfully forecasted by WRF/Chem. The increased background O\(_3\) concentrations during the summer result in a large number of exceedances of the O\(_3\) daily maximum average limit at the Ayia Marina background station. A smaller number of exceedances is observed at the urban stations due to the local NO\(_x\) emissions. The WRF/Chem model was found to be more skillful in predicting these exceedances, although there were a number of false exceedance predictions which is attributed to the 15% overestimation of O\(_3\) concentrations predicted by the model.

WRF/Chem and CAMS have similar performance in terms of forecasting the PM2.5 concentrations during the summer, revealing an underestimation at the traffic stations which may be attributed to the missing road dust re-suspension mechanism and the proximity of these stations to the main traffic areas. In addition, emissions from domestic heating which are not accurately represented in the air quality forecasting models, result in underestimation of nighttime PM2.5 concentrations during winter. An overestimation in PM2.5 concentrations is shown by WRF/Chem at the background stations during winter, suggesting there is potential for further improvement, as this overestimation may be due to various factors such as dust emissions and transport, secondary aerosol formation, or long-range transport of PM2.5 of anthropogenic origin. Additional effort is needed to improve the skill of the WRF/Chem model to forecast the PM2.5 concentrations in this domain, by examining and evaluating the model performance for the PM2.5 sub-components, as well as dust which is a major contributor to PM in the region.

Based on the findings of this study, the use of up-to-date and high spatiotemporal resolution anthropogenic emission inventories is recommended for forecasting the concentrations and the diurnal fluctuations of atmospheric pollutants, especially near the urban centers where the majority of population lives. To further improve the forecast capacity of the models, emission heights can be employed for power stations, as well as plume dispersion and chemistry to incorporate the rapid conversion of SO\(_2\) to sulphate (adding to the PM2.5 mass).
Coupled on-line air quality models with nesting for higher horizontal resolution can provide improved real-time air-quality forecasts, at least for short-lived species or species that undergo photochemical reactions, compared to the state-of-the-art global chemical transport models. Adaptive time-stepping can be prescribed in RT-AQF applications with coupled models to reduce the required simulation times while meeting the CFL stability criterion at each time-step. With the advent of efficient processor technology and as computational resources become increasingly available to the modelling community, higher spatial resolutions can be used in regional coupled on-line air quality models which can go down to the convection resolving limit, targeting high population areas; while the use of fine temporal resolution allows for better capturing feedbacks between the meteorological and chemical processes, and diurnal variability.

Finally, a longer evaluation period in future studies covering entire and multiple years can provide a representative larger sample of variability of air-pollution events, and aim to capture extreme but rare events, and the frequent spring and autumn dust episodes in this region. The advances in the predictive skill of real-time air quality forecasting models can help reduce population exposure to air pollutants and associated health risks.

**Code availability.** The WRF/Chem model code is available at https://github.com/wrf-model/WRF (visited on 19/01/2022)

**Data availability.** All raw data can be provided by the corresponding authors upon request

**Author contributions.** GKG, TC, JK designed the experiment. GKG carried out the experiment and performed the simulations. GKG, JK and TC analyzed the data and wrote the manuscript. YP provided technical support for implementing the experiment. CS and MP provided and processed the observational data. JS, JL reviewed and edited the manuscript.

**Competing interests.** The authors declare that no competing interests are present.

**Acknowledgements.** The authors are sincerely thankful to the Cyprus Department of Labour Inspection (DLI) for providing the observational data used for the model evaluation in this study, and would also like to acknowledge the Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus.eu/ (visited on 19/01/2022)) which is operated by the European Centre for Medium-Range Weather Forecasts on behalf of the European Commission as part of the Copernicus Programme for the provision of the data used for model intercomparison. The NCAR Command Language (NCL) version 6.5.0 (http://www.ncl.ucar.edu/ (visited on 19/01/2022)) and the openair R package (Carslaw and Ropkins, 2012) were used to produce the plots in this work. This publication has been produced within the framework of the projects EMME-CARE, which has received funding from the European Union’s Horizon 2020 Research and Innovation Programme (under grant agreement no. 856612) and the Cyprus Government, AQ-SERVE no. INTEGRATED/0916/0016 which is co-financed by the European Union's Horizon 2020 Research and Innovation Programme (under grant agreement no. 856612).
Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation, and ACCEPT which is co-financed by the Norwegian Financial Mechanism (85%) and the Republic of Cyprus (15%) in the framework of the programming period 2014–2021.
References


