Response to reviewer #2 "Analysis of summer O3 in the Madrid air basin with the LOTOS-EUROS chemical transport model" by Miguel Escudero et al.

The authors would like to acknowledge the comments from reviewer #2. We thank the positive comments. All the suggestions demonstrate a good knowledge of the scientific field and that has resulted in a great improvement of the paper after his/her revision.

My only major comment would be that I somehow disagree with the authors that model evaluation is only useful to build confidence in the tools: it is also essential to guide their development. The detailed analysis presented here could thus be more conclusive in pointing specific issues that would deserve higher priority in future model development, i.e. being more specific that pointing to "future research" (as stated p26). Taking the example of vertical resolution, the two configurations tested here are somewhat extreme: going from 5 to 70 layers, which is presumably not realistic in long term simulations used in policy support or air quality forecasting. It would have been useful to know to what extent the 5 layer model captures the episode processes discussed in 3.2, and how a tradeoff could be found.

This is an interesting comment by the reviewer. The authors fully agree with the idea of the results evaluation should serve to trigger the improvements in air quality modelling. In this work, however, we aimed to analyse specific summer episodes of O3 in the MAB with the aid of high resolution simulations performed with LOTOS-EUROS. This means, that the main objective of the paper is to provide a phenomenological interpretation of O3 events in the area. It is also true that we make a detailed evaluation of the best configuration of the model for the specific area and period and, for that, we used, along with standard surface observations, highly resolved vertical profiles for O3.

Regarding the specific question of the reasonable number of vertical levels in the model configuration, the answer is that it is dependent on the objective of the study. In this study the environmental analysis was the main objective and it was logical and feasible from the perspective of CPU time to employ a considerable number of vertical levels because it allowed a better representation of the vertical variability of O3. In other studies in which CPU time is limiting such as air quality forecasting or long term analyses, the reasonable number of levels can be less.

This has been indicated in the conclusions:

"The main objective of the paper is to provide a phenomenological interpretation of O3 events in the area after performing a detailed evaluation of the best configuration of the model for the specific area and period. Regarding the specific question of the reasonable number of vertical levels in the model configuration, it is dependent on the objective of the study. In this study the environmental analysis was the main objective and it was logical and feasible from the perspective of CPU time to employ a considerable number of vertical levels because it allowed a better representation of the vertical variability of O3. In other studies such as air quality forecasting or long term analyses in which CPU time may be large, the reasonable number of levels can be less."

Specific comments:

P2L12 : why using plural here? only O3 is a secondary air pollutant

The phrase has been changed to singular.

P5L14: the air pollution regimes (REC/SAD/NAD) should be introduced here and related to synoptic meteorological situations.

Basic information about these three regimes are provided now in the following manner:

"Under low-gradient synoptic conditions, the combination of the strong convective conditions and the blocking effect of the mountain ranges induces an important vertical development of the boundary layer and mesoscale recirculation. During the night, north-easterly winds prevail over the basin and, after dawn, the eastern slopes of the Guadarrama range are progressively warmed up causing a clockwise turning of wind to an E and S during the day finalising with an SW component in the late afternoon. The drainage flows at night-time re-establish the north-easterlies. These events are commonly referred as recirculation (REC) episodes. The presence of the Azores high or low pressure systems over the Atlantic in front of the Iberian Peninsula generate advection of Atlantic air masses from the north (we will refer to these as Northern advective or NAD events) or from the south (Southern advective or SAD events)."

P6L10: indicate the range explored in terms of O3 dry deposition velocities. Was this impact assessed on the basis of free-tropospheric total ozone burden as in Stevenson et al. or rather surface ozone?

We performed runs multiplying the standard dry deposition velocity (calculated by the resistance approach as detailed in Manders et al., 2016) by a factor of 1.25 and 0.75. Comparisons of modelled surface O3 concentrations revealed minimal effect of this parameter on the results for the specific period. The total O3 burden was not examined since we were just were studying specific summer events in a limited region and the background concentrations were not expected to vary significantly for the aforementioned changes in O3 deposition velocity.

On this respect, we have included the following sentence in section 2.3:

"Initial sensitivity studies were performed with the base configuration (configured similar to the operational forecasts that are part of the CAMS regional ensemble as presented in Marécal et al. (2015)) to test the response of the model to changes in the deposition velocity of O3 because night-time dry deposition has been suggested as a factor that could strongly influence the ability of CTMs to simulate tropospheric O3 (Stevenson et al., 2006; Monks et al., 2015). The standard dry deposition velocity, calculated by the resistance approach (Manders et al., 016), has been multiplied by either a factor 1.25 or 0.75. The results (not shown here) reflected a minimal effect of this parameter on O3 concentrations in the chosen domain and period, and therefore deepening in this direction was discarded."

P7L2: what is the reference year for emissions used?

The reference year for the MACC-III emissions was 2011. This has been indicated in the text.

P9L5: the worsening of correlations when increasing resolution has been documented as "double penalty" in the field of meteorological forecast. It would be worth discussing in more detail the issue here

A brief description of the "double penalty" concept is now provided:

"This is known in meteorological modelling as the Double Penalty issue (Mass et al., 2002) and occurs when evaluating simulations using point observations. The high resolution runs may be penalized twice, for not capturing the occurrence of the event and also for not predicting the right location of the event while a low resolution simulation can only fail predicting the event."

P12L2: it is indeed frustrating that NO2 is not included in the detailed validation. It would also be interesting to see some basing meteorological validation, not clear why WRF outperforms IFS and there could be compensation of errors.

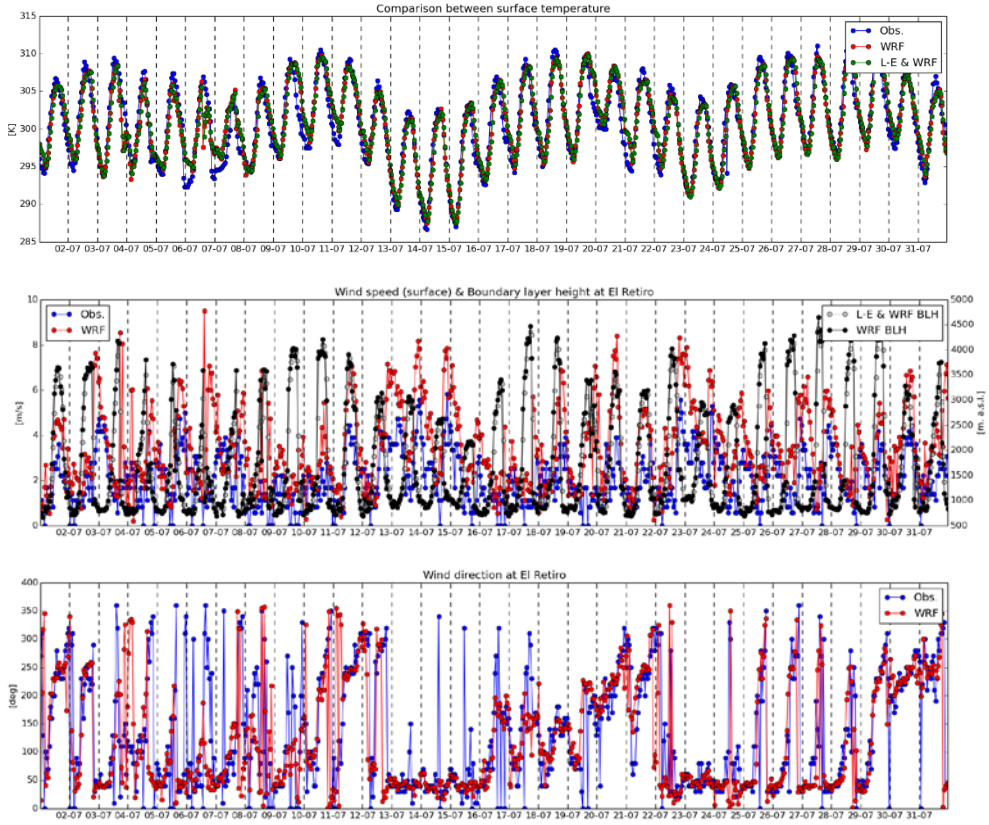
A table showing the Pearson's correlation factor (r) correlation factor for NO2 and NOx was also composed and it is shown below. However, it is fair to indicate that, since O3 was our target pollutant, traffic stations were excluded from the set of stations used for validation so the picture may be a biased. Moreover, the results for the ECMWF_HR_70 were not calculated. As shown in that table, NO2 correlates worse with observations than O3 as generally occurs in CTM's due the high variability. However, it is true that ECMWF runs correlate better than WRF runs probably due to the coarser resolution of the first with respect to the latter. This can be a good example of the "double penalty" situation. These are the results:

		NO_2					NO_x			
STATION	**TYPE**	**ECMWF_5**	**ECMWF_70**	**WRF_5**	**WRF_70**		**ECMWF_5**	**ECMWF_70**	**WRF_5**	**WRF_70**
VILLA DEL PRADO	RURAL	0.481	0.576	0.42	0.437		0.454	0.535	0.348	0.371
SAN MARTIN DE VALDEIGLESIAS		0.405	0.454	0.337	0.358		0.425	0.45	0.344	0.308
EL ATAZAR		0.422	0.524	0.381	0.472		0.446	0.539	0.408	0.476
SAN PABLO DE LOS MONTES		0.463	0.502	0.295	0.423		0.451	0.507	0.304	0.455
CAMPISABALOS		0.328	0.419	0.333	0.36		0.31	0.418	0.314	0.332
GUADALIX DE LA SIERRA		0.509	0.479	0.57	0.408		0.49	0.46	0.583	0.378
ORUSCO DE TAJUNA		0.481	0.471	0.39	0.388		0.156	0.288	0.088	0.046
ALCORCON	URBAN	0.369	0.466	0.407	0.447		0.329	0.422	0.342	0.353
TOLEDO2		0.318	0.33	0.244	0.233		0.274	0.281	0.221	0.196
ENSANCHE DE VALLECAS		0.483	0.474	0.266	0.372		0.421	0.459	0.217	0.268
VILLAVERDE		0.436	0.437	0.322	0.333		0.348	0.412	0.224	0.229
ARTURO SORIA		0.422	0.522	0.319	0.342		0.391	0.526	0.328	0.391
FAROLILLO		0.36	0.427	0.29	0.371		0.298	0.446	0.217	0.409
PLAZA DEL CARMEN		0.219	0.307	0.184	0.204		0.162	0.304	0.131	0.263
GUADALAJARA		0.359	0.394	0.271	0.235		0.363	0.399	0.256	0.221
MOSTOLES		0.404	0.507	0.44	0.401		0.363	0.48	0.382	0.319
ARANJUEZ		0.343	0.283	0.204	0.211		0.322	0.239	0.168	0.159
RETIRO		0.339	0.523	0.232	0.455		0.346	0.575	0.266	0.519
TRES OLIVOS		0.472	0.55	0.378	0.476		0.529	0.593	0.39	0.495
AZUQUECA DE HENARES		0.448	0.425	0.318	0.26		0.434	0.403	0.291	0.256
BARAJAS - PUEBLO		0.595	0.505	0.213	0.416		0.567	0.501	0.176	0.323
RIVAS-VACIAMADRID	SUBURBAN	0.446	0.348	0.322	0.301		0.386	0.363	0.27	0.224
AVILA 2		0.502	0.487	0.499	0.467		0.482	0.471	0.5	0.476
JUAN CARLOS I		0.482	0.467	0.28	0.328		0.469	0.458	0.257	0.314
EL PARDO		0.493	0.433	0.483	0.473		0.509	0.449	0.505	0.49
ALGETE		0.413	0.451	0.328	0.267		0.419	0.456	0.314	0.219
MAJADAHONDA		0.301	0.423	0.352	0.341		0.293	0.434	0.349	0.31
ESTACION DE LA SAGRA (ILLESCAS)		0.322	0.344	0.274	0.198		0.282	0.29	0.232	0.134
TORREJON DE ARDOZ		0.521	0.488	0.466	0.407		0.503	0.493	0.453	0.338
VALDEMORO		0.449	0.422	0.385	0.378		0.422	0.437	0.406	0.327
CASA DE CAMPO		0.243	0.378	0.264	0.39		0.183	0.391	0.177	0.423

Regarding, the evaluation of the meteorological datasets we made a qualitative validation of the degree of concordance between the two meteorological datasets (WRF and ECMWF) and observations (at El Retiro in the centre of Madrid) of temperature, wind direction and wind speed. We concluded that temperature compared almost optimally for both datasets.

Regarding the wind direction was is adequately represented in the two meteorological datasets. Finally, wind speed is occasionally overestimated by the meteorological models although this overestimation is more marked in the case of WRF (see below the plots of the comparison with WRF data were a comparison between the BLH provided by the input meteorological data and the LOTOS-EUROS transformation has been added).

With respect to the validation of meteorological fields in the vertical, Figure S4 in the supplementary information contains the comparison between the vertical fields of RH, T, and wind speed and direction.



P16L25: how do you explain the local minima of O3 around 4km agl?

We do not have a clear answer to this question. However, an analysis of the vertical profiles and concentration maps revealed that the band with high O3 located just below this local minima may correspond to an intrusion of O3 generated southeast from the MAB.

P18, Fig 7: what is driving the sharp horizontal convergence of NO2 between 12 and 18UT?

The convergence of NO2 is probably driven by convection since temperature was very high during the described event. Moreover, the presence of the Guadarrama range also influences. Finally, the presence of the NO2 column in the plots in the centre of the domain is logically associated with the region of high NOx emissions (Madrid metropolitan area).

P20L15-20: this section deserves further discussion in light of Querol et al. 2018 (Section 4), where they analyze the likelihood of impact at the ground of the STE event in relation with diurnal PBL variability.

We have done this. The following lines have been introduced in the paper:

"The actual impact of this stratospheric intrusion on surface levels remains unclear. In the referenced paper, the authors estimate a possible but limited impact of the intrusion on surface levels assuming that the boundary layer could exceed the 3000 ma.s.l. during the day. As shown in Figure 9, LOTOS-EUROS predicts that the maximum altitude of the boundary layer according to LOTOS-EUROS reached its maximum values of 2500–2700 ma.s.l. limited by the wind ventilation so, probably, the impact on the surface should be low (if any) in this case."

This is true for 18/7. Thus, this sentence has been included in the text for clarity.

Figure 6, 8, 10: those figures are very nice and comprehensive. The only missing information is modelled O3 time series. Although model and observations are compared in the in 3.1.1 and 3.1.2, I am missing a visual comparison of time series. The vertical cross sections are difficult to read and would benefit from being consistent with Fig 7&9 (i.e. both could extend to 7km to allow discussing the stratospheric intrusion). The color legend of O3 for the vertical cross section should be consistent with the maps (and the label corrected: it is probably ppb, not ppm).

In these figures our main interest was the environmental interpretation of the O3 episodes with the valuable information provided by the high resolution simulation. This, along with the need of keeping the figure readable, advised against including the O3 time series for comparison especially, taking into account that a complete evaluation of the model performance has been already provided in the first sections.

Nevertheless, in correspondence with the interest of the reviewer, we have added a figure in the supplementary information. In that figure we show the modelled (WRF_70) vs observations time series of July 2016 in four stations representative of the different areas of the MAB: ALG on the central part of the MAB, AZU on the Henares valley at the northeast, ORT on the east and SMV on the southwest.

The objective of Figures 7, 9 and 11 was to show the behaviour of the vertical behaviour of NO2 and O3 in the relevant part of the troposphere (first 5 km) with good resolution. Meanwhile, the objective of the time cross-sections in Figures 6, 8 and 10 was to analyse the time evolution of concentrations so extending the vertical dimension up to 14 km allowed to observe the formation and evolution of, for example, stratospheric intrusions. This is the reason why we have chosen different vertical extent in the cross- sections presented in these two sets of figures. With respect to the different colour legends between the maps and the cross sections was compulsory since the concentration ranges were different in both cases. Finally, we have corrected the label indicated by the reviewer.

Response to reviewer #3 of "Analysis of summer O3 in the Madrid air basin with the LOTOS-EUROS chemical transport model" by Miguel Escudero et al.

The authors would like to acknowledge the comments from reviewer #3. We thank the positive comments. All the suggestions demonstrate a good knowledge of the scientific field and that has resulted in a great improvement of the paper after his/her revision.

Before answering to the reviewer's comments one by one, the authors want to clarify that the main objective of this paper is to perform a detailed analysis of summer O3 events in a topographically complex basin such as the MAB with the valuable help of LOTOS-EUROS CTM. This model was adapted/configured to provide the best representation of the scientific observations resulting from a comprehensive field campaign carried out in Madrid on July 2006. Our main goal was not to realise a deep analysis of the driving factors of LOTOS-EUROS in O3 simulations (physical and chemical schemes, emission patterns…).

We opted to focus our efforts in finding the best configuration for the particular study area and period. In addition to the horizontal resolution, we knew (from previous experimental studies in the region) that the vertical resolution was a key aspect that had not been taken into account so often in previous simulation studies in the Mediterranean. Moreover, the use of two different meteorological input datasets was planned in order to evaluate if there was a relevant difference in the outputs in the two cases. That is a relevant question for CTMs especially when modelling O3.

What is the state of the art terms of fine restitution of the fields of ozone in the Mediterranean? What has been undertaken at this scale so far? What are the current shortcomings in terms of model performance for ozone, what are the locks for the restitution of fine-scale ozone fields, on the horizontal and on the vertical?

The following paragraphs have been added to the introduction:

"Air quality model results vary at different resolutions especially due to the resolution of emissions and the description of the driving meteorology (Fenech et al, 2008). Some authors have found that the impact of higher horizontal resolutions in O3 simulations is more sensitive to the resolution of emissions than to meteorology (Valari and Menut, 2008). Moreover, finer resolution result in less dilution of emissions but also in differences have been found in the O3-NOx interaction (Valari and Menut, 2008, Stock et al., 2014, Schaap et al., 2015).

In the Iberian Peninsula, the use of fine grids (in the order of 1-5 x 1-5 km) has been found beneficial in the context of complex terrains where mesoscale processes acquire importance for interpreting production and transport of O3 (Toll and Baldasano, 2000, Jimenez-Guerrero et al., 2008,). In coastal areas, with complex topography, high resolution simulations have been generally employed with good results (Carvalho et al., 2006, Jimenez et al., 2006, Gonçalves et al., 2009). Moreover fine grids have been recommended for describing O3 variability especially in urban and industrial areas (Jimenez et al, 2006, Baldasano et al., 2011). In general, the use of finer resolution simulations in the Iberian Peninsula generally imply benefits in the O3 description such as improvement in correlation and reduction in bias and errors (Jimenez et al., 2006).

Less importance has been given to the vertical resolution mostly because the vertical O3 profile evaluation of CTM is difficult due to the lack of experimental vertical O3 data. In complex domains in the Iberian Peninsula the models may not reproduce O3 concentrations due to a poor representation of mesoscale flows and layering and accumulation of pollutants (Gonçalves et al., 2009). In general, it has been demonstrated that incrementing vertical resolution would help to resolve meteorological phenomena (Carvalho et al., 2006) and would also offer a more realistic vertical exchange between the boundary layer and the free troposphere (Jimenez et al., 2006)."

What is required to go further in terms of ozone modelling?

From the generic point of view, there are a variety of questions that could be named as next steps for the improvement of CTMs regarding O3 modelling. In general, simulating episodes is challenging due to the impact of multi-driving factors (emissions, meteorology, physical-chemical schemes…). Thus, investigating responses to emission changes (including biogenic emissions that are poorly described) is obviously one of the key aspects but it also relevant to find specific configurations for the CTMs in order to provide the best possible estimates of the formation and transport mechanisms that trigger O3 concentrations especially in summer. For these and other challenges, it is extremely important to evaluate modelling tools with data obtained in campaigns/measurements using state-of-the-art equipment both on the surface and vertically.

More in particular, on the basis of the results obtained in the present work we would like to mention four particular studies/strategies that could favour CTM development:

- **Include vertical measurements of NOx in field campaigns.**
- **Perform measurements of NOy for carrying out NOx/VOC sensitivity studies (according to Sillman's methodology).**
- **Improve temporal resolution in emission models.**
- **Study the vertical exchanges of O3 in detail.**

These arguments are now reflected in the paper in the conclusions section:

"The results from this study can be useful to understand the phenomenology of high O3 episodes in the MAB and to gain knowledge to design appropriate strategies for air-quality management. Further research must be implemented to investigate aspects like the sensitivity to emission reduction scenarios or the role of VOCs with emphasis on the biogenic ones. Moreover, to perform the tasks of validating and optimising CTMs, increasing efforts should be made to conduct more field campaigns in different air basins in the Mediterranean using state-of-the-art equipment to generate data and knowledge about O3 behaviour both on the surface and vertically. Useful parameters to be included in these campaigns are O3, NOx, VOC and, when possible, intermediate products like NOy, HNO3 and H2O2 that, according to previous experience (Sillman, 1995), are key parameters for facing model-based NOx–VOC sensitivity studies and the assessment of emission inventories. Some of these parameters (especially NOx) should also be incorporated in vertical measurements.

In future, similar simulations to the one presented in this study should be performed in the different air basins in the IP where O3 exceedances have been recorded (Querol et al., 2016). CTMs should be configured specifically for each region or air basin to assure the best performance by capturing the influence of topography and local circulations. For such studies,

we highlight the importance of conducting experimental campaigns that can support the necessary model evaluation.

Finally, it should be noted that when running such fine resolutions for real applications it is also important to work on the emission datasets (out of the scope of this work). Increasing the detail in emission inventory (mainly based on a bottom-up approach) could improve the performance of CTMs when assessing sensitivities or emission scenarios. Moreover, improving time resolution in the emission models can be beneficial for simulating O3 episodes."

In what way does this fine-scale study bring new elements, not only to the evaluation of the model itself, but also to research on the subject?

The topographical complexity of the MAB along with the large variability of emissions in a conurbation like Madrid required the use of fine scale configurations. In the vertical, it is clear that a considerable number of layers in the model allow a better representation of the vertical variability of O3 that would have been more difficult to observe and analyse with a coarser resolution (stratospheric intrusion, formation of residual layers, etc). This has been stated by many authors who carried out other fine-scale modelling studies (see discussion about fine resolution studies added to the introduction above). Moreover, the temporal extent of the analysis (1 month) is not common and has offered the opportunity to describe the patterns/scenarios of O3 episodes in the MAB during summer. Finally, it is innovative the discussion of the role of the BLH because this parameter is highly dependent of the model resolution.

Also, why do the authors only focus on model resolution?

In these detailed analysis for regions with complex topography and a great variety of sources, model resolution in the horizontal is an issue. Also vertical resolution was a key aspect on the light of previous studies which studied O3 phenomenology based on experimental campaigns (Querol et al., 2018 among others). In those studies, vertical fumigation was observed to have a key influence on surface O3 concentrations. In consequence, if LOTOS-EUROS was able to adequately represent mechanisms such as vertical exchange, formation of O3 residual layers, high altitude intrusions or fumigation according to what had been observed in the field campaigns, it could be trustfully used for a phenomenological study on O3 in summer in the MAB. We should bear in mind that LOTOS-EUROS in its standard configuration is a computationally cheap CTM (fast and accurate) and, by this study, we demonstrate that it is also able to reproduce complex events with an obvious resolution adaptation.

In the conclusions, there is a lack of discussion about the importance of improvements (improved indicators), with regard to the requirements of enhanced configurations (especially CPU time).

The question of whether it is worthwhile to perform these high resolution simulations (both horizontally and vertically) in terms of CPU time depends on the objectives that you are seeking. In this case, CPU time was increased reasonably when running 70-layer schemes with fine resolution (ECMWF_70 or WRF_70) because our objective was to describe O3 phenomenology in the MAB.

We have added to the conclusions section a discussion about this:

"The main objective of the paper is to provide a phenomenological interpretation of O3 events in the area after performing a detailed evaluation of the best configuration of the model for the specific area and period. Regarding the specific question of the reasonable number of vertical levels in the model configuration, it is dependent on the objective of the study. In this study the environmental analysis was the main objective and it was logical and feasible from the perspective of CPU time to employ a considerable number of vertical levels because it allowed a better representation of the vertical variability of O3. In other studies such as air quality forecasting or long term analyses in which CPU time may be large, the reasonable number of levels can be less."

It would also be important to give a strategy to choose a future configuration: is it reasonable to choose so many vertical levels (70)?

This answer is linked with the previous one. The use of 70 layers allows the model to represent a vertical complexity that would not be possible with the usual number of layers in standard CTM configurations. This was important for this work because we aimed to perform a correct phenomenological study of O3 in a complex region. Is it reasonable to make operational analyses or forecasts with that high number of layers? Probably we cannot afford that but the objective is different in those cases. Just to highlight the relevance of the vertical configuration of the model for O3 simulations, this study has triggered a discussion among the developers of LOTOS-EUROS on incrementing the number of vertical layers in the standard version of the model without putting in risk the computational efficiency of the model. As commented in the previous question, a paragraph discussing about this has been added to the conclusions section.

Also, still for a contextualization (this time of the results and not of the stakes), the interpretations of the episodes (which are precise), lack some context. There has been indeed a lot of studies of ozone formation episodes in the literature. The formation of ozone downwind sources, and its dependence upon wind speed and vertical dilution are known: what exactly are the new knowledge at the end of the study, concerning the phenomenology of episodes and their properties, or concerning the ability of a CTM to simulate them? How can these new knowledge improve air quality management strategies?

This study provides new knowledge on the occurrence of O3 episodes in the MAB. Among these, we could highlight three. Firstly, it is the first time that summer O3 events in the MAB have been classified in detail, taking into account transport features and the relevance of vertical exchange. This is then a necessary step to investigate emission scenarios and control strategies to be applied. Moreover, this study is innovative because it demonstrates the transfer of O3 produced at the MAB towards other air basins under certain meteorological scenarios (see, for example, section 3.2.3). Finally, regarding the ability of CTM to simulate complex O3 events in Southern Europe, this study demonstrate the relevance of horizontal and, especially, vertical resolution in CTM configuration.

Are we just talking about improved scores from one version to another, or do some configurations allow to depict specific features that do not appear in the others?

Apart from the relevant improvement of scores associated with increments in the horizontal resolution, it is clear that, for example, the 5-layer schemes (ECMWF_5 and WRF_5) cannot describe vertical exchange of O3 since they do not have enough resolution. This means that, with low number of layers, it is impossible, for example, to observe the formation of residual layers or the influence of stratospheric intrusions.

Specific comments:

Page 10 – line 5 - "In the plots corresponding to ECMWF_5 and ECMWF_70 runs we observe systematic positive bias especially in the period 14–20 UTC when the formation is strong although it only spiked with low wind speed. This feature was not so marked in the three remaining configurations and, in particular, in the two WRF runs the bias values were randomly distributed around zero." How do the authors explain the mid-day biases of ECMWF compared to WRF? Is it just a problem of resolution of the meteorological calculations, or does it depend on the meteorological model itself?

We suspect that it can be attributed mainly to the resolution given that the ECMWF run performed with higher resolution (ECMWF_HR_70) responded more closely to the WFR runs.

Line 15 page 15: the best overall performance is analyzed on which criteria? Which parameters are used to affirm that the simulation is better? Should it be the restitution of the diurnal peak, the phasing of the morning increase, the total amount on the vertical, or just the indicator average...? In particular, the WRF70 has a strong underestimation of diurnal ozone in Figure 4 at El Pardo and this is still considered as the "best run". What about this feature at other stations?

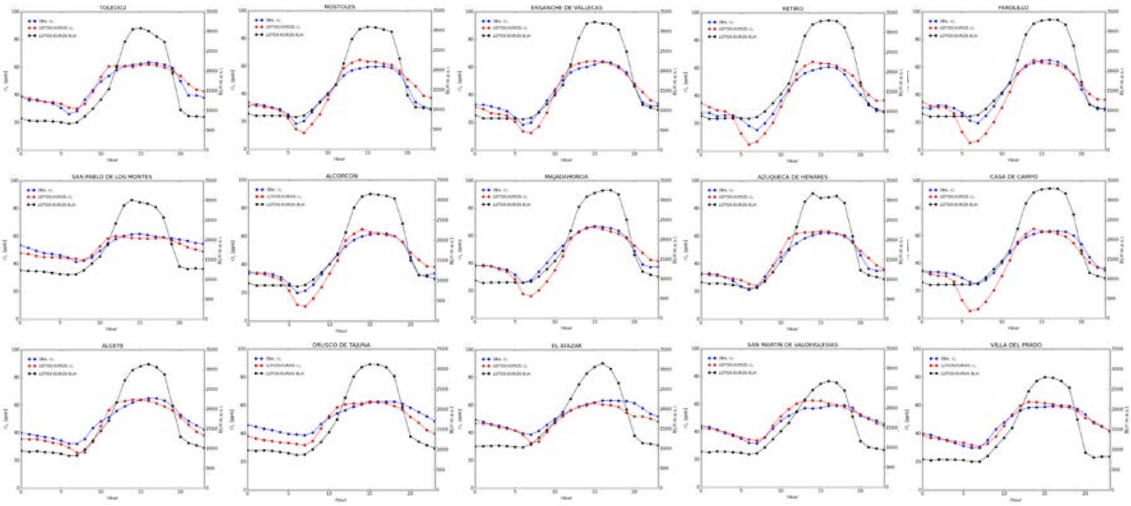
The best overall performance was selected from all parameters obtained in the validation assessment:

i. **From the inspection of r values we concluded that the configurations with higher vertical resolution perform better than the 5-layer schemes.**
ii. **From the analysis of the model bias (Figures 2, 3 and S3), we concluded that clear improvements were observed using finer spatial resolution so either WRF simulations (WRF_5 and WRF_70) or ECMWF_HR_70 were the best option on this regard.**
iii. **The assessment of NMSE (Figure 2) revealed that WRF_70 showed slightly lower errors than the other configurations although differences were not drastic.**
iv. **After the evaluation of mean daily cycles (Figure 4) we found that the five-level configurations (ECMWF_5 and WRF_5) presented a not realistic sharp increment in O3 concentrations from 6 to 7 UTC. Although the timing of the increase was better reproduced in these runs than in the 70-level schemes, the latter were more realistic in the daily evolution of O3 concentrations (without steep increases in the morning).**
v. **From the comparison of vertical profiles, we concluded that obviously high resolution in the vertical was needed to capture the layering of pollutants reflected in the O3 soundings.**

Summarising, the best overall configuration for the analysis of O3 in this region/period were setups with high resolution both in the vertical and in the horizontal, namely ECMWF_HR_70

and WRF_70. In particular, we used WRF_70 for vertical cuts because spatial resolution over the study area was better.

The underestimation of diurnal O3 in the mean daily cycle at El Pardo was not present in other stations on a routine basis. The following pictures present cases in which the mentioned underestimation is not observed.



The difference between what is observed and simulated is not always specified, even if we can guess it. Example: in "Figure 7. Longitudinal and latitudinal vertical crosssections of NO2 and O3 for 16 July 2016". Same for figure 9 and figure 11. Also, for page 17 lines 1 to 10, it is important to specify that it is a vision drawn by the model and not the result of observations.

We have made this clear in those Figures' captions. In page 17 we have added a comment making clear that the interpretation is made by means of the analysis of simulation results:

"From the simulation results, we can interpret the evolution of O3. At 18 UTC of 16 July, we can see how O3 levels increased drastically, and the boundary layer depth grew up to 3500 m a.s.l. aided by convection at 18 UTC. Normally, REC events show higher planetary boundary layer (PBL) heights in the evening. Figure 7 shows how the strong convection during REC events injected ground-level pollutants at high altitudes during the late afternoon 5 and the evening reaching up to 3500 m a.s.l. as illustrated in the NO2 plots. When the night-time stable boundary layer forms after sunset, air masses with high O3 that originated near the surface during the previous day were decoupled and remained in the residual layer at altitudes ranging between 2000 and 4000 m a.s.l. forming reservoir layers (00 and 06 UTC cross-sections in Figure 7) which can fumigate the following day. These reservoir layers can also be observed as a relatively thin band at an altitude of 2000–4000 m a.s.l. during every night of the REC period (Figures 6 and S5)."

Figures 6, 8, 10: The location/typology of the station groups should be mentioned.

This information is presented in Figure 1. Thus, in order to clarify we have added in the figure caption of Figures 6, 8, 10 the reference to Figure 1.

Figure 7: it would be more convenient for the reader to visualize on a map the latitudinal and longitudinal cuts.

Exactly as in the previous comment, the map with the latitudinal and longitudinal cuts is presented in Figure 1. We refer again in the figures' captions to Figure 1.

ANALYSIS OF SUMMER O_3 IN THE MADRID AIR BASIN WITH THE LOTOS-EUROS CHEMICAL TRANSPORT MODEL

Miguel Escudero[1,2], Arjo Segers[2], Richard Kranenburg[2], Xavier Querol[3], Andrés Alastuey[3], Rafael Borge[4], David de la Paz[4], Gotzon Gangoiti[5], and Martijn Schaap[2]

[1]Centro Universitario de la Defensa (CUD) de Zaragoza, Academia General Militar, Ctra. de Huesca s/n, E-50090 Zaragoza, SPAIN
[2]TNO, P.O. Box 80015, 3584 TA, Utrecht, THE NETHERLANDS
[3]Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Jordi Girona 18-26, E-08034 Barcelona, SPAIN
[4]Department of Chemical & Environmental Engineering, Technical University of Madrid (UPM), c/José Gutiérrez Abascal 2, E-28006 Madrid, SPAIN
[5]Escuela Técnica Superior Ingeniería de Bilbao, Departamento Ingeniería Química y del Medio Ambiente, Universidad del País Vasco UPV/EHU, Urkixo Zumarkalea, s/n, Bilbao, E-48013, SPAIN

Correspondence: Miguel Escudero (mescu@unizar.es)

Abstract. Tropospheric O_3 remains a major air-quality issue in the Mediterranean region. The combination of large anthropogenic emissions of precursors, transboundary contributions, a warm and dry aestival climate and topographical features results in severe cases of photochemical pollution. Chemical transport models (CTMs) are essential tools for studying O_3 dynamics and for assessing mitigation measures but they need to be evaluated specifically for each air basin. In this study, we present an optimisation of the LOTOS-EUROS CTM for the Madrid air basin. Five configurations using different meteorological datasets (from the European Centre for Medium Weather Forecast (ECMWF) and Weather Research and Forecasting (WRF)), horizontal resolution and number of vertical levels were compared for July 2016. LOTOS-EUROS responded satisfactorily in the five configurations reproducing observations of surface O_3 with notable correlation and reduced bias and errors. However, the best-fit simulations for surface O_3 were obtained by increasing spatial resolution and using a large number of vertical levels to reproduce vertical transport phenomena and the formation of reservoir layers. Using the optimal configuration obtained in the evaluation, three characteristic

events have been described: recirculation (REC) episodes and northern and southern advection (NAD and SAD, respectively) events. REC events were found to produce the highest O_3 due to the reduced ventilation associated with low wind speeds and the contribution of reservoir layers formed by vertical transport of O_3 formed near the surface in the previous days of the event. NAD events, usually associated with higher wind speeds , present the lowest ground-level O_3 concentrations in the region. During SAD episodes, external contributions along with low wind speeds allow O_3 to increase considerably, but not as much as in REC events because steady southerly winds disperse local emissions and hinder the formation of reservoir layers.

Copyright statement. TEXT

1 Introduction

Ozone (O_3) is formed in the troposphere by the interaction of gaseous precursors like nitrogen oxides (NO_x) and volatile organic compounds (VOCs) in the presence of sunlight. Much attention has been given to this secondary air pollutant in the last decades due to the variety of negative effects on health, ecosystems, crops, climate and materials associated with it (see review by Monks et al. (2015) and references therein).

The oxidative effect of O_3 generates inflammation of airways. Increases in morbidity and mortality and chronic alterations of the cardiovascular and cerebrovascular systems have also been associated with exposure to O_3 (WHO, 2006, 2013a, b). Tropospheric O_3 is also harmful for vegetation, generating leaf symptoms, reduced growth, senescence, defoliation and reducing crop productivity (Paoletti, 2006; WGE, 2013). Damage to construction materials like plastics, surface coatings and rubber due to O_3 has been documented (Lee et al., 1996; Screpanti and Marco, 2009). Moreover, O_3 in the troposphere acts as a greenhouse gas with positive global radiative forcing (IPCC, 2013).

It is estimated that 98% of the urban population in Europe in 2016 were exposed to excessive concentrations of tropospheric O_3 according to the World Health Organization (WHO) guidelines values, a steady proportion since 2000 (EEA, 2018). However, it is the Mediterranean basin where the most acute episodes are registered (Millán et al., 1997, 2000; Sicard et al., 2013; Querol et al., 2016). In the Iberian Peninsula (IP), located at the Western Mediterranean Basin, the intense solar radiation, high temperatures and lack of

precipitation in spring and summer, associated with persistent anticyclonic conditions, favour the formation of O_3 in the area and the accumulation in rural and suburban regions (Escudero et al., 2014, 2016; Querol et al., 2016; Massagué et al., 2019). The emissions of precursors from anthropogenic sources in the Mediterranean basin and the surrounding regions are considerable, especially in some densely populated areas. In addition to that, the amount of biogenic VOCs emitted in southern Europe is considerably higher than in central and northern Europe (Seco et al., 2011). Moreover, during the frequent biomass burning episodes in summer, air-quality problems associated with tropospheric O_3 are aggravated (Tressol et al., 2008). In particular, the complex orography of the IP with mountain ranges running parallel to the coast intersected by river basins that penetrate towards the inner continental areas and elevated plateaus in the centre of the peninsula, help air masses to recirculate and age under the influence of sea and mountain breezes that develop when synoptic circulation is inhibited by the presence of the Azores high (Millán et al., 2000; Gangoiti et al., 2001; Valverde et al., 2016; Querol et al., 2017). Previous studies also suggest that local strategies designed to meet NO_2 ambient air-quality standards may have caused an increase of urban O_3 that, in turn, causes an increase in oxidative capacity of Madrid's atmosphere by increasing OH and NO_3 radicals (Saiz-López et al., 2017).

In recent years, several comprehensive summer campaigns with intensive measurements of surface and vertical profiles of O_3 concentrations and its precursors have been undertaken near the two main conurbations in Spain: Barcelona (2015, 2017 and 2018) and Madrid (2016) (Querol et al., 2017, 2018; Reche et al., 2018; Carnerero et al., 2018). The main objective of these campaigns was to interpret the phenomenology of high O_3 and ultrafine particles' episodes in Spain.

Another essential objective of retrieving data from these intensive campaigns is related to the validation and optimisation of chemical transport models (CTMs). These models constitute an essential tool for analysing O_3 behaviour with high spatial and time resolution, providing air-quality forecasts and supporting the design of policies. This includes the study the NO_x–VOC sensitivity (Sillman, 1999; Sillman and West, 2009) essential for proposing and evaluating potential mitigation measures. Regional CTMs have been used to investigate O_3 pollution in Spain in several studies. Most of these studies aimed to describe short-term (rarely exceeding 5 days) pollution events (Toll and Baldasano, 2000; Jiménez et al., 2005; San José et al., 2005; Jiménez et al., 2006; Carvalho et al., 2006; Valverde et al., 2016; Pay et al., 2018) and, in some cases, to discuss the effectiveness of potential mitigation strategies (Palacios et al., 2002; Soret et al., 2014). Despite these efforts, work is still needed to evaluate the impact of changes in the vertical configuration of CTMs,

especially in the Mediterranean region where the atmospheric dynamics in summer is characterised by complex recirculation processes with effective vertical exchange (Millán et al., 1997, 2000; Gangoiti et al., 2001; Borge et al., 2010; Querol et al., 2017, 2018). The lack of an appropriate representation of the vertical variability of O_3 has been recognised as one of the shortcomings of the CTMs and in consequence a major challenge in the future development of the models (Hess and Zbinden, 2013; Monks et al., 2015). Moreover, it is strongly recommended to combine modelling with observations because this will bring knowledge from both sources together and permit adequate evaluation procedures of the model outputs (Canepa and Builtjes, 2017).

Air quality model results vary at different resolutions especially due to the resolution of emissions and the description of the driving meteorology (Fenech et al., 2018). Some authors have found that the impact of higher horizontal resolutions in O_3 simulations is more sensitive to the resolution of emissions than to meteorology (Valari and Menut, 2008). Moreover, finer resolution result in less dilution of emissions but also in differences have been found in the O_3-NO_x interaction (Valari and Menut, 2008; Stock et al., 2014).

In the Iberian Peninsula, the use of fine grids (in the order of 1-5 x 1-5 km) has been found beneficial in the context of complex terrains where mesoscale processes acquire importance for interpreting production and transport of O_3 (Toll and Baldasano, 2000; Jiménez-Guerrero et al., 2008). In coastal areas, with complex topography, high resolution simulations have been generally employed with good results (Carvalho et al., 2006; Jiménez et al., 2006; **?**) . Moreover fine grids have been recommended for describing O_3 variability especially in urban and industrial areas (Jiménez et al., 2006; Baldasano et al., 2011). In general, the use of finer resolution simulations in the Iberian Peninsula generally imply benefits in the O_3 description such as improvement in correlation and reduction in bias and errors (Jiménez et al., 2006). Less importance has been given to the vertical resolution mostly because the vertical O_3 profile evaluation of CTM is difficult due to the lack of experimental vertical O_3 data. In complex domains in the Iberian Peninsula the models may not reproduce O_3 concentrations due to a poor representation of mesoscale flows and layering and accumulation of pollutants (**?**). In general, it has been demonstrated that incrementing vertical resolution would help to resolve meteorological phenomena (Carvalho et al., 2006) and would also offer a more realistic vertical exchange between the boundary layer and the free troposphere (Jiménez et al., 2006).

Making use of the results on the O_3 episodes phenomenology from the aforementioned field campaign in Madrid in July 2016, we were able to assess and optimise LOTOS-EUROS CTM v2.0 (Manders et al.,

2017) for simulating O_3 in this region. Five configurations combining different meteorological input data and vertical structures were employed after identifying these two aspects as key factors for the capability of the model for reproducing O_3 levels. We simulated the entire month of July 2016 in accordance with the experimental campaign with a spin-up period of 24 h. The aim of this comparison was to elucidate the optimal configurations for operating with LOTOS-EUROS in the region but also to identify relevant factors to set up other CTMs used in this region.

Moreover, employing the optimal configuration of the modelling system, we discuss the phenomenology of tropospheric O_3 in the Madrid air basin (MAB) for the study period. This was done by analysing simulated fields of meteorological variables and pollutants with special emphasis on the vertical variability to test the importance of the up– down transport of O_3 in the region.

2 DATA AND METHODS

2.1 Study area

The Madrid Metropolitan Area (MMA) is a densely populated area with more than 5 million inhabitants. According to Salvador et al. (2015) and Borge et al. (2014), the main sources of pollutants in the region are road traffic, residential heating (which maximize their emissions in winter), a busy airport and minor contributions from industry.

The MMA is located in the centre of the MAB and lies on an elevated plateau ($\sim$700 m above sea level (a.s.l.)) in the middle of the IP (Figure 1). The climate in the area is continental Mediterranean with warm and dry summers and cold and also dry winters. The main orographic features surrounding the basin are, around 120 km to the south of the MMA, the Toledo Mountains (altitudes up to 1600 m a.s.l.) with an E–W axis and the Guadarrama range (maximum heights of 2400 m a.s.l.) which runs diagonally from SW to NE, 50 km to the west and north of the MMA. The Guadarrama range is part of the Central System that extends until the Ebro valley and, together with the western flank of the Iberian range delimits a channel to the NE along the Henares valley (Figure 1). As a result of this configuration, the circulation in the Madrid basin shows a dominant SW–NE direction (Plaza et al., 1997). Under low-gradient synoptic conditions, the combination of the strong convective conditions and the blocking effect of the mountain ranges induces an important vertical development of the boundary layer and mesoscale recirculation. During the night,

north-easterly winds prevail over the basin and, after dawn, the eastern slopes of the Guadarrama range are progressively warmed up causing a clockwise turning of wind to an E and S during the day finalising with an SW component in the late afternoon. The drainage flows at night-time re-establish the north-easterlies. These events are commonly referred as recirculation (REC) episodes. The presence of the Azores high or low pressure systems over the Atlantic in front of the Iberian Peninsula generate advection of Atlantic air masses from the north (we will refer to these as Northern advective or NAD events) or from the south (Southern advective or SAD events).

2.2 The LOTOS-EUROS model

The 3D CTM LOTOS-EUROS v2.0 and its previous versions have been extensively used in the past for air-quality studies, including NO_x (Schaap et al., 2013; Vlemmix et al., 2015), SO_2 (Barbu et al., 2009) and particulate matter (PM) (Schaap et al., 2004; Manders et al., 2009; Timmermans et al., 2017). In particular, tropospheric O_3 has been the scientific target in different studies carried out with successive versions of LOTOS-EUROS. It has been employed in health-related studies (van Zelm et al., 2008) and, more recently, Beltman et al. (2013) applied LOTOS-EUROS to simulate the response of tropospheric O_3 in Europe to a 5% shift from crop- and grassland into poplar plantations used for biomass production, while Hendriks et al. (2016) tested the response to a decarbonisation scenario in the continent. Although LOTOS-EUROS has been generally employed in a continental domain (mainly in Europe), a sensitivity study to regional changes in emissions in three areas of Europe (Poland, the Po valley and Flanders) was also performed by Thunis et al. (2015). In addition, LOTOS-EUROS has also been used in a number of intercomparison studies with other CTMs for the simulation of O_3 (Hass et al., 1997; van Loon et al., 2007; Cuvelier et al., 2007; Vautard et al., 2007; Solazzo et al., 2012; Im et al., 2015) showing a satisfactory performance. Finally, regarding air-quality predictions, LOTOS-EUROS participates in the CAMS (Copernicus Atmosphere Monitoring Service) ensemble (Curier et al., 2012), which offers operational forecasts for NO_2, O_3 and PM.

2.3 Model experimental design

A detailed description of the 2.0 version of LOTOS-EUROS can be found in its reference guide (Manders et al., 2016) where all technical issues (processes, schemes, etc.) are described and referenced (accessible at

www.lotos-euros.nl). In this section, we provide a brief description focusing on the most relevant aspects for this study.

Initial sensitivity studies were performed with the base configuration (configured similar to the operational forecasts that are part of the CAMS regional ensemble as presented in Marécal et al. (2015)) to test the response of the model to changes in the deposition velocity of O_3 because night-time dry deposition has been suggested as a factor that could strongly influence the ability of CTMs to simulate tropospheric O_3 (Stevenson et al., 2006; Monks et al., 2015). The standard dry deposition velocity, calculated by the resistance approach (Manders et al., 2016) by a factor of 1.25 and 0.75. The results (not shown here) reflected a minimal effect of this parameter on O_3 concentrations in the chosen domain and period so deepening in this direction was discarded.

As shown in Table 1, two major aspects were modified in the set of five configurations: the meteorological input data and the vertical structure of the model. We fed LOTOS-EUROS with operational data from the reanalysis of the ECMWF model (Flemming et al., 2009) retrieved with a spatial resolution of 7 x 7 km^2. A second meteorological gridded dataset was obtained with the WRF model (Skamarock et al., 2008) with a resolution of 1 x 1 km^2 over a square domain of approximately 220 km of side centred on the city of Madrid (Figure 1). Data from WRF simulations with similar configurations have been previously used to drive air-quality simulations over the IP and, in particular, in the Madrid area (Borge et al., 2008, 2014). In this case, the WRF model was run on a three -nested-domain configuration as shown in Figure S1. Additional information about the WRF model configuration is provided in Table S1.

For the vertical structure, we compared the standard five-level mixed-layer configuration (Manders et al., 2017) with a hybrid-layer multilevel scheme. This version uses the lowest 70 layers of the 137 hybrid sigma-pressure layers used by ECMWF for the operational meteorological forecasts in 2016. In such a vertical co-ordinate system, model layers are defined by pressure boundaries that follow surface pressure at lower altitudes but slowly evolve into fixed pressure levels in the stratosphere (Eckermann, 2009).

Finally, the MACC III emission inventory (reference year 2011) has been used for all set-ups and initial and boundary concentrations were taken from global simulations produced by and used in CAMS services, as described in Marécal et al. (2015). These include concentrations of the most important trace gases and aerosols.

	ECMWF_5	ECMWF_70	ECMWF_HR_70	WRF_5	WRF_70
Meteorological model spatial resolution	7 x 7 km^2			1 x 1 km^2	
Vertical structure	Mixed-layer (5 levels)	Hybrid-layer (70 levels)		Mixed-layer (5 levels)	Hybrid-layer (70 levels)
LOTOS-EUROS Spatial resolution	25 x 25 km^2		10 x 10 km^2	3 x 3 km^2	
Emission inventory	MACC III				

Table 1. Summary of the LOTOS-EUROS v2.0 model runs and settings performed for this work.

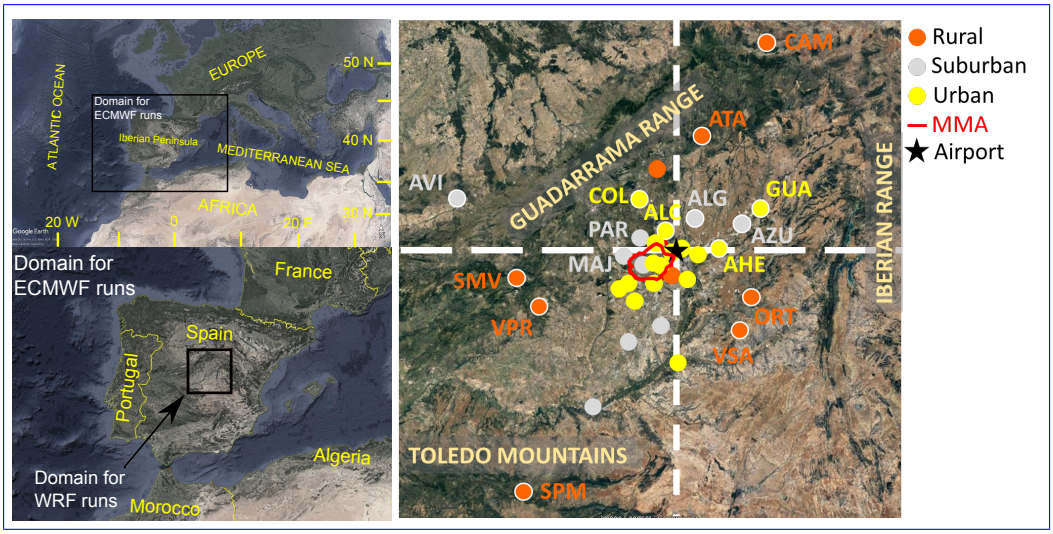


Figure 1. Situation maps of the Iberian Peninsula and the MAB with the indication of the domains used for the simulations with LOTOS-EUROS. In the map located on the right side, the selected set of monitoring stations used for illustrating O_3 variability in different sectors of the region is shown. Stations' codes: Alcalá de Henares-AHE, Alcobendas-ALC, Avila-AVI, El Atazar-ATA, Azuqueca de Henares-AZU, Campisábalos-CAM, Colmenar Viejo-COL, Guadalajara-GUA, Majadahonda-MAJ, Orusco de Tajuña-ORT, San Martín de Valdeiglesias-SMV, San Pablo de los Montes-SPM, El Pardo-PAR, Villa del Prado-VPR, Villarejo de Salvanés-VSA. The white dashed lines indicate the E–W and N–S cross-sections presented in figures from section 3.2.

2.4 Monitoring data

Hourly O_3 data for the simulation period (July 2016) were collected from 35 air-quality monitoring sites (17 urban, 9 suburban and 9 rural) located in the MAB (Table 2 and Figure 1) and its boundary region. Traffic stations were in general discarded for the model evaluation due to their limited spatial representativity although four traffic sites were also employed. In spite of belonging to different air-quality networks (Table 2), all the O_3 monitors are based on ultraviolet photometry according to EN 14625:2012, which is the reference technique for automatic monitoring of O_3 established in the European Directive 2008/50/EC.

Several previous studies used modelling techniques to analyse intense short-term O_3 episodes (Palacios et al., 2002; San José et al., 2005), and, more specifically, to evaluate the impact of specific environmental policies in the Madrid region (Soret et al., 2014) or the influence of sectoral emissions (Borge et al., 2014; Valverde et al., 2016; de la Paz et al., 2016; Pay et al., 2018). The present work contributes with an evaluation and optimisation of LOTOS-EUROS CTM in the region incorporating a high degree of vertical resolution. The extended simulation period of this study (one month) is suitable for characterising the typical O_3 episodes occurring in the area in summer.

3 RESULTS

3.1 Model performance

3.1.1 Ground-level O_3

The results from the outputs from the five runs were assessed following the suggestions provided by Borrego et al. (2008) on the evaluation of the quality of model simulations. These authors stated that Pearson's correlation coefficient (r), the fractional bias (FB) and the normalised mean square error (NMSE) were the most relevant statistical parameters to be analysed for O_3 simulations with CTM. Figure 2 shows a coloured grid indicating the values of r, FB and NMSE obtained in the comparison between the observations from background monitoring stations and the simulated O_3. The colour scale is independent for each parameter and serves as an indication of the agreement between observations and model predictions.

All five configurations presented good correlations with observations with an average r of 0.752 although lower values were obtained for the five-layer schemes (0.695 ± 0.077 and 0.745 ± 0.044) than for the

configurations using the hybrid-layer scheme (0.750 $\pm$ 0.062 - 0.801 $\pm$ 0.034). Among these multiple-layer-scheme simulations, the one showing the best r is ECMWF_70 which was the configuration with the coarser spatial resolution for LOTOS-EUROS among the three. Therefore, regarding the degree of correlation for multilayer configurations, increasing horizontal resolution to very fine grid sizes in the photochemical model does not improve results (providing that no changes are implemented in the emission inventory). This is known in meteorological modelling as the Double Penalty issue (Mass et al., 2002) and occurs when evaluating simulations using point observations. The high resolution runs may be penalized twice, for not capturing the occurrence of the event and also for not predicting the right location of the event while a low resolution simulation can only fail predicting the event.

		r					FBIAS					NMSE				
		ECMWF_5	ECMWF_70	ECMWF_HR_70	WRF_5	WRF_70	ECMWF_5	ECMWF_70	ECMWF_HR_70	WRF_5	WRF_70	ECMWF_5	ECMWF_70	ECMWF_HR_70	WRF_5	WRF_70
VILLA DEL PRADO		0.756	0.796	0.784	0.711	0.739	0.179	0.232	0.039	0.084	0.009	0.066	0.080	0.052	0.043	0.041
SAN MARTÍN DE VALDEIGLESIAS		0.744	0.811	0.736	0.692	0.751	0.157	0.210	0.099	0.089	0.036	0.057	0.066	0.051	0.035	0.029
EL ATAZAR		0.690	0.780	0.741	0.621	0.702	0.043	0.079	0.083	0.014	0.051	0.034	0.031	0.052	0.032	0.032
SAN PABLO DE LOS MONTES	RURAL	0.682	0.737	0.638	0.563	0.638	0.030	0.084	0.063	0.020	0.054	0.031	0.031	0.061	0.023	0.027
CAMPISÁBALOS		0.683	0.747	0.496	0.618	0.684	0.313	0.358	0.168	0.264	0.211	0.140	0.162	0.142	0.116	0.088
GUADALIX DE LA SIERRA		0.721	0.785	0.771	0.759	0.803	0.258	0.293	0.185	0.196	0.154	0.120	0.128	0.094	0.088	0.067
ORUSCO DE TAJUÑA		0.750	0.822	0.785	0.690	0.747	0.029	0.120	0.025	0.005	0.071	0.045	0.040	0.040	0.027	0.034
ALCORCÓN		0.748	0.807	0.801	0.736	0.793	0.176	0.282	0.035	0.042	0.020	0.110	0.127	0.130	0.075	0.075
TOLEDO		0.702	0.738	0.709	0.735	0.771	0.139	0.224	0.072	0.103	0.017	0.081	0.097	0.081	0.059	0.047
ENSANCHE DE VALLECAS		0.767	0.817	0.786	0.711	0.797	0.076	0.236	0.064	0.021	0.024	0.105	0.107	0.107	0.097	0.082
VILLAVERDE		0.747	0.783	0.754	0.715	0.782	0.169	0.328	0.034	0.066	0.048	0.145	0.171	0.171	0.113	0.103
ARTURO SORIA		0.735	0.815	0.763	0.605	0.778	0.231	0.388	0.083	0.066	0.094	0.181	0.207	0.180	0.151	0.118
FAROLILLO		0.732	0.784	0.729	0.650	0.743	0.083	0.243	0.066	0.108	0.070	0.123	0.122	0.169	0.151	0.125
PLAZA DEL CARMEN	URBAN	0.609	0.727	0.714	0.476	0.704	0.373	0.526	0.228	0.183	0.223	0.358	0.389	0.283	0.287	0.230
GUADALAJARA		0.710	0.808	0.764	0.757	0.812	0.227	0.313	0.199	0.175	0.107	0.125	0.141	0.101	0.084	0.060
MOSTOLES		0.774	0.824	0.790	0.786	0.811	0.211	0.317	0.051	0.121	0.027	0.115	0.143	0.128	0.074	0.071
ARANJUEZ		0.769	0.785	0.754	0.793	0.838	0.238	0.331	0.193	0.163	0.119	0.118	0.157	0.104	0.069	0.052
RETIRO		0.737	0.822	0.778	0.598	0.788	0.184	0.342	0.036	0.010	0.031	0.156	0.171	0.156	0.163	0.109
TRES OLIVOS		0.742	0.811	0.744	0.617	0.770	0.120	0.202	0.063	0.084	0.087	0.086	0.089	0.113	0.126	0.088
AZUQUECA DE HENARES		0.771	0.838	0.804	0.773	0.813	0.182	0.269	0.111	0.126	0.051	0.092	0.109	0.072	0.069	0.053
BARAJAS - PUEBLO		0.784	0.809	0.792	0.686	0.793	0.165	0.324	0.034	0.474	0.188	0.131	0.173	0.139	0.487	0.193
RIVAS-VACIAMADRID		0.730	0.768	0.748	0.740	0.781	0.100	0.260	0.088	0.167	0.087	0.128	0.135	0.131	0.107	0.085
JUAN CARLOS I		0.786	0.825	0.765	0.707	0.759	0.006	0.167	0.071	0.188	0.158	0.088	0.079	0.129	0.165	0.138
EL PARDO		0.822	0.831	0.693	0.778	0.763	0.119	0.201	0.066	0.101	0.014	0.065	0.085	0.105	0.075	0.072
ALGETE		0.807	0.864	0.825	0.718	0.787	0.084	0.166	0.038	0.031	0.050	0.047	0.051	0.053	0.042	0.042
MAJADAHONDA	SUBURBAN	0.782	0.832	0.717	0.717	0.734	0.066	0.174	0.019	0.027	0.051	0.065	0.065	0.104	0.064	0.077
ILLESCAS		0.805	0.835	0.792	0.766	0.812	0.212	0.303	0.151	0.160	0.076	0.097	0.127	0.090	0.074	0.055
TORREJON DE ARDOZ		0.784	0.839	0.820	0.761	0.811	0.134	0.293	0.132	0.103	0.033	0.119	0.137	0.102	0.090	0.077
VALDEMORO		0.751	0.763	0.772	0.762	0.812	0.243	0.336	0.136	0.158	0.086	0.128	0.169	0.099	0.084	0.063
CASA DE CAMPO		0.716	0.814	0.737	0.608	0.754	0.027	0.188	0.006	0.164	0.125	0.112	0.079	0.127	0.158	0.119
	AVERAGE	**0.745**	**0.801**	**0.750**	**0.695**	**0.769**	**0.152**	**0.260**	**0.088**	**0.117**	**0.079**	**0.109**	**0.122**	**0.112**	**0.108**	**0.082**
	S. DEVIATION	**0.044**	**0.034**	**0.062**	**0.077**	**0.044**	**0.089**	**0.095**	**0.060**	**0.095**	**0.059**	**0.060**	**0.068**	**0.050**	**0.090**	**0.046**

Figure 2. Values of the Pearson's correlation factor (r), fractional bias (FB) and normalised mean standard error (NMSE) of the comparison of the five LOTOS-EUROS simulations with observations of hourly O_3 data from 35 background monitoring stations located in the MAB. Colours illustrate the model agreement from blue (worst) to red (best).

LOTOS-EUROS tends to moderately overestimate ground-level O_3 concentrations varying widely in the five set-ups. It was clear that deviations with respect to observations declined with the use of a higher number of vertical levels and finer spatial resolution. The absolute values of the averaged FB for WRF_70

(0.079 $\pm$ 0.059) and ECMWF_HR_70 (0.088 $\pm$ 0.060) were substantially lower than those of the other three configurations (0.260 $\pm$ 0.095 – 0.117 $\pm$ 0.095). Clear improvements were observed using finer spatial resolution in LOTOS-EUROS (either WRF simulations or ECMWF_HR_70) while ECMWF_5 and, especially, ECMWF_70 presented systematic but moderate overestimations (Figure S3). In consequence, the best configurations for adjusting the model bias were WRF_70 and ECMWF_HR_70.

A major reason for the overestimation detected for ECMWF runs with coarser spatial resolution was associated with an excessive O_3 formation in the noon hours of the day in situations of low wind speed as shown in Figure 3. This plot shows the correlation between the model bias and the modelled wind speed in the location of El Retiro (see location in Table 2) in Madrid for the five runs. In the plots corresponding to ECMWF_5 and ECMWF_70 runs we observe systematic positive bias especially in the period 14–20 UTC when the formation is strong although it only spiked with low wind speed. This feature was not so marked in the three remaining configurations and, in particular, in the two WRF runs the bias values were randomly distributed around zero. ECMWF_HR_70 run showed a subtler systematic overestimation during daytime but the correlation with low wind speeds was not observed in this case. Analysing the night-time period (0–6 UTC) we detect that the systematic overestimation was only present in the ECMWF_70 execution.

Regarding the simulation errors quantified via the NMSE (Figure 2), the results were satisfactory because the values of this parameter remained low for the five configurations (means ranging between 0.082 $\pm$ 0.046 and 0.122 $\pm$ 0.068 with WRF_70 run showing the best performance). Relevant information extracted from NMSE values in Figure 2 was that the errors of the model were consistently lower in rural areas (0.063 $\pm$ 0.038) than in suburban (0.096 $\pm$ 0.033) and urban (0.135 $\pm$ 0.077) sites. This might be an effect of the simulation of the interaction between O_3 and NO_x, which acquires more relevance as a source of errors in the vicinity of traffic sources. Other authors (de la Paz et al., 2016) suggest that the use of an urban canopy model (not used in this simulation) improves model predictions in densely built areas by reducing the overestimation of wind speed.

Systematic features characterised the modelled mean daily cycles of the 35 stations (Figure 4 shows the plots for the El Pardo site as a typical example). The daily cycles obtained with the simulations performed with the mixed-layer scheme of five levels (ECMWF_5 and WRF_5) presented a sharp increment in O_3 concentrations from 6 to 7 UTC which was not present in the simulations performed with the hybrid-scheme of 70 levels. In these runs, the morning increment of O_3 after the rush hour was delayed one or two hours

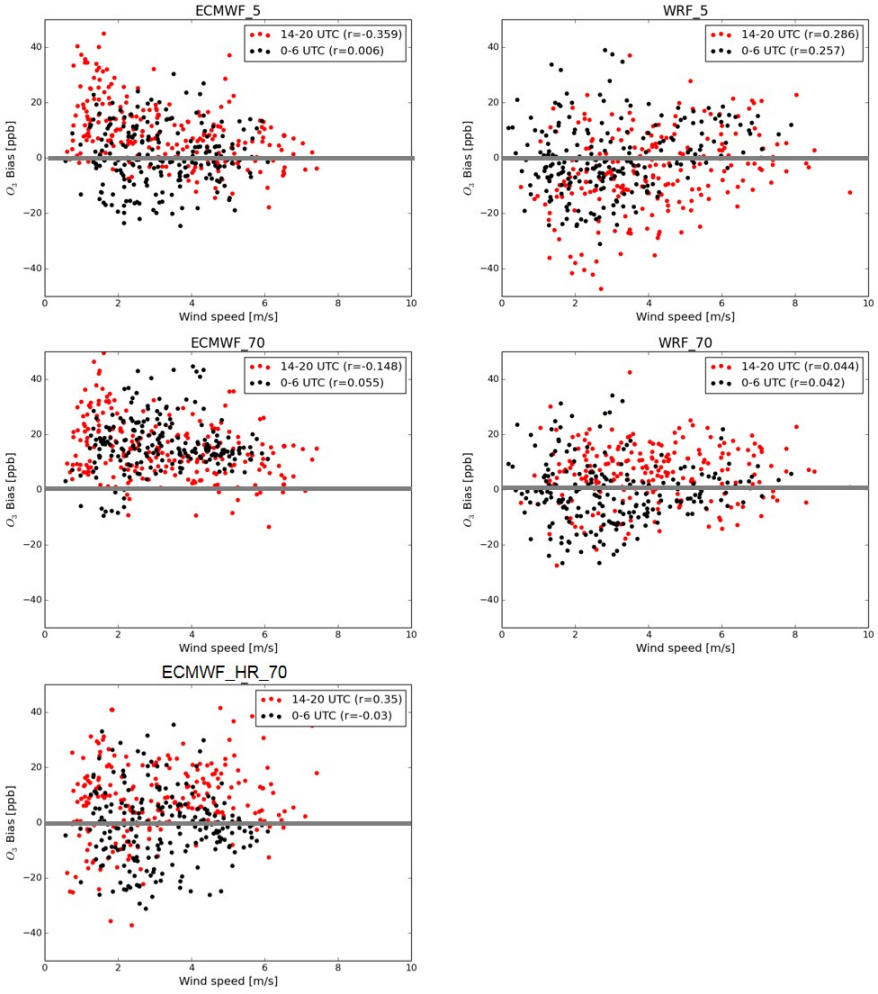


Figure 3. Correlation plots between the LOTOS-EUROS bias of surface O_3 and simulated wind speed at the station of Retiro disaggregated in two day periods: 0–6 and 14–20 UTC.

(WRF_70 and ECMWF_HR_70 runs) and clearly smoothed. The observations reflected that the timing of the increase was better represented in the mixed-layer scheme runs although the increase was excessively abrupt. The occurrence of this steep increase in the concentrations in the executions performed with the

mixed-layer scheme coincided with the first steps of the boundary layer development. In the mixed-layer scheme, a rise of the boundary layer leads immediately to complete mixing of NO_x emitted at the surface over the (increased) boundary layer, and thus limits titration of ozone; in the 70-layer schemes, however, the mixing over the boundary layer seems to take place more gradually. A more extensive validation including other tracers than the chemically active ozone should provide insight into which scheme performs better under which conditions, and preferably lead to better characterization of the vertical diffusion.

3.1.2 O_3 vertical profiles

The evaluation of CTMs in the vertical direction has always been a difficult task due to the small number of high-resolution vertical observations to compare with. In this work, data from O_3 free soundings launched from Madrid airport at 12 UTC every seven days on 6, 13, 20 and 27 July 2016 were used for this purpose. Comparisons between the vertical profiles of modelled O_3 with the five different set-ups and the observations are presented in Figure 5. The corresponding profiles of wind direction and speed (modelled data are taken from the input meteorological datasets) can be found in the supplementary information (Figure S4). As suggested by Querol et al. (2018), the enrichment of O_3 in the lower troposphere during episodes without ventilation is high as a consequence of the intense photochemical formation and the development of convective circulations. This typically results in vertical profiles in which O_3 concentrations are relatively high near the surface up to 2000 m a.s.l. at midday soundings in Madrid, as shown also here in Figures 5a and 5d. During venting episodes, the more intense surface dispersion explains why O_3 vertical profiles have lower values in the mixing layer (Figures 5b and 5c).

The event of 27 July was characteristic of an accumulation scenario with high concentrations near the surface while on 13 July (a typical venting event), O_3 in the lower levels was moderate and increased with altitude as described above. These two profiles were correctly simulated by most runs. However, the modelled profiles for the 6 and 20 July showed overestimation in the lower levels with respect to observations and for the 20 July case, all the high-resolution simulations overestimated the observed values. In Figure S4, we can also check how the input meteorological data used to feed the simulations (ECMWF and WRF fields) closely reproduced the wind profiles obtained during the soundings of those four days both on speed and direction. Because of the complexity of the vertical mechanisms, further research should be conducted to investigate the causes of this mismatch in some events.

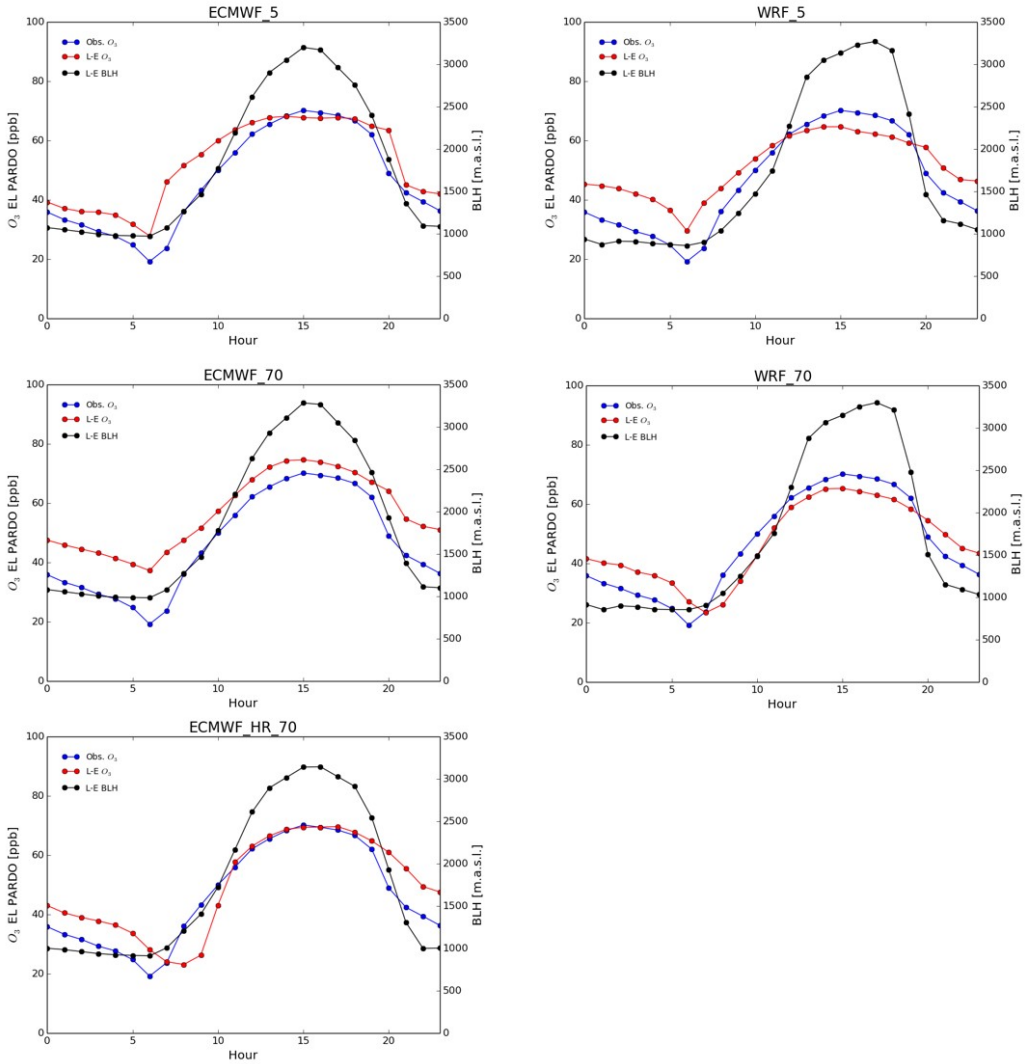


Figure 4. Simulated average diurnal cycles obtained from the five LOTOS-EUROS configurations compared with the mean cycle from the observations in the El Pardo background station. The modelled evolution of the boundary layer height is also shown.

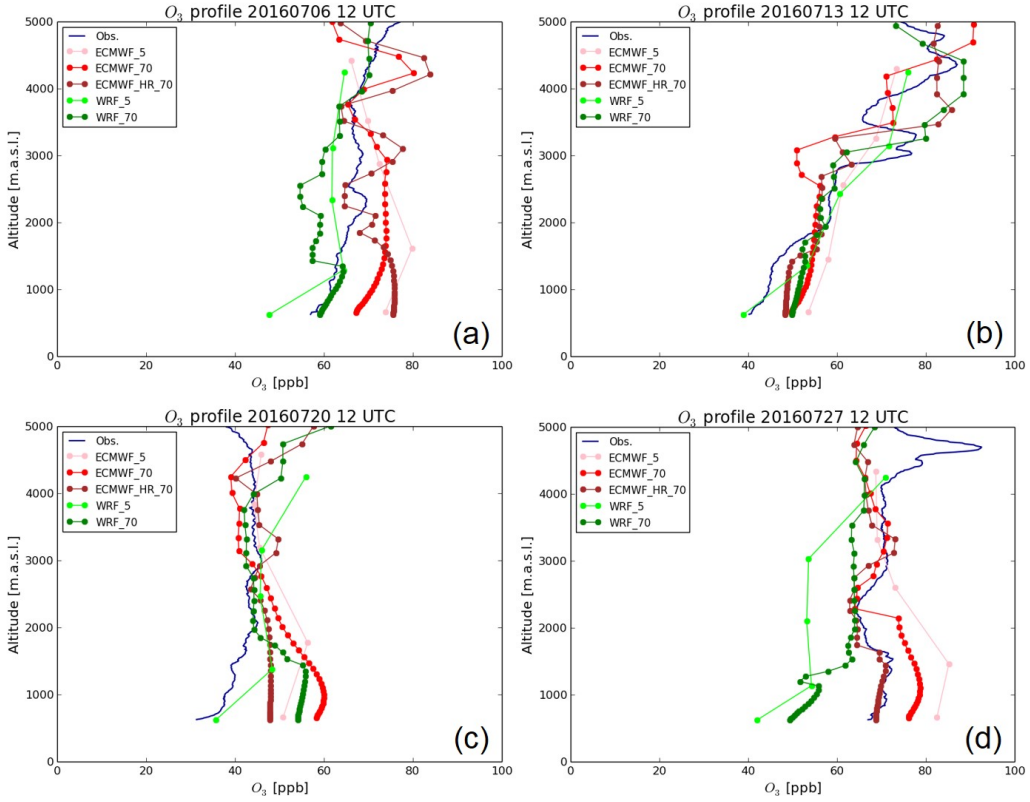


Figure 5. Real and simulated vertical profiles of O_3 for the 6, 13, 20 and 27 July 2016.

From the qualitative perspective, the first obvious conclusion was that a larger number of vertical levels in the model considerably improved the capability for capturing the vertical gradients of O_3 concentrations with the exception of the lowest level on 20 July. However, even in the simplest vertical scheme (five layers), the model was able to reproduce the general vertical trends. A particular meteorological scenario was present during July 13. We can observe two O_3 layers centred around 3000 and 4300 m a.s.l . The two multilayered WRF runs captured these two layers at, approximately, the correct altitudes although the 3000 m layer was not as marked as in the observations. The ECMWF runs also presented these two features but displaced in altitude by around 200–300 m with respect to observations.

In all the WRF_5 runs and on 27 July WRF_70 simulations a steep drop of surface concentrations was noticed. This is probably associated with the emission model configuration, the fine spatial resolution of the runs (3 x 3 km^2) and the vertical mixing in these set-ups. The O_3 soundings were released from Adolfo Suárez Barajas Airport, which is one of the major airports in Europe with more than 53 million passengers and 470 tonnes of goods transported in 2017 (http://www.aena.es/) and significant NO_x emissions. The emission model employed by LOTOS-EUROS allocates at the surface level all the air traffic emissions in the grid cell where the airport is located. These emissions, added to those produced by road traffic, lead to a grid cell that has the highest NO_x emissions in the domain. As a consequence, and given the fine size of the grid cells in the WRF runs, O_3 levels were excessively reduced compared with reality in the aforementioned cases. Regarding ECMWF runs, the coarser spatial resolution did not allow observation of this feature because the modelled emissions from the airport were distributed on a larger surface.

Quantitatively, the model generally showed the ability to reproduce the same order of magnitude of the concentrations observed in the O_3 soundings at all altitude levels. The values of the statistical parameters that indicate the quality of the simulations (FB and NMSE) of the vertical profiles of O_3 are shown in Table 3 where generally satisfactory values can be observed with poorer results especially for July 20. The FB showed a majority of positive values indicating overestimation although, in most cases, it was moderate (the range of averages for the five configurations was –0.5 to 0.11). The NMSE data in Table 3 support this conclusion because the average errors were small (0.023–0.042).

Summarising, the configuration that presented the best overall performance among the five tested in the previous sections was WRF_70, so it was employed for interpreting the variability of O_3 in the MAB during July 2016.

3.2 Interpretation of O_3 in the MAB in July 2016

According to the dominant circulation over the MAB, three different episodes were distinguished and, with the aid of the model outputs, the basic features of the three events were described. Figure S2 shows the location of the selected monitoring stations used to characterise the behaviour of surface O_3 in the different sectors of the MAB.

3.2.1 Recirculation events (REC)

These events correspond to the pattern sketched by Plaza et al. (1997) in which wind direction turns clockwise during the day aided by the effect of the blocking effect of the Guadarrama range while Querol et al. (2018) described that the mixing layer growth at midday was reduced favouring vertical recirculation at the eastern slopes of the Guadarrama range (see section 2.1). In July 2016 four REC periods were identified: 1–6, 8–11, 15–17 and 25–28. To illustrate the main features of REC episodes, the period 15–17 July will be used as an example (Figures 6 and 7). A complete pattern of simulated fields of O_3, wind and relative humidity (RH) for July 2016 can be consulted in Figure S5.

Surface wind speeds registered during REC episodes were weak (Figures 6 and S5) and the change in direction associated with recirculation is observed. However, despite the local circulation, air masses remain inside the basin during REC days aided by a relatively thin mixing layer at 12 UTC (Figure S4).

A stable band of high RH centred at around 4000 m is observed in Figures 6 and S5 which can be associated with the evapotranspiration caused by the intense heating registered during these events. The presence of a high-altitude trough located to the west of the IP during the 3–6 July REC period, induced moist south-westerlies at altitude resulting in the development of convective clouds in the evenings (Figure S5).

Surface O_3 concentrations generally reach high values at the central time of the day during REC episodes (Figure 6). During the 17 REC days registered in July 2016, 37 exceedances of 180 μg/m^3 were recorded in all the monitoring stations in the study area. The sectors suffering the greatest impact of O_3 are the W–N belt of the MMA (MAJ, PAR, COL, ALC and ALG) and more episodically, the Henares valley (AZU, AHE and GUA), the SE of the basin (ORT and VSA) or, when SW winds in the evening reached sufficient intensity (for example on 16 July), the O_3-enriched air masses reach rural stations located in the NE of the basin (ATA and CAM) late in the day. LOTOS-EUROS surface concentration maps show that at 12 UTC O_3 begins to rise, reaching the maximum concentrations around 18 UTC. Time series of O_3 also support this conclusion because the highest O_3 levels are observed in the evening (Figure 6).

To study the three-dimensional variability of O_3 during a typical REC event, Figure 7 presents longitudinal and latitudinal cross-sections of simulated O_3 and NO_2 for four different hours during 16 July. The use of NO_2 plots allows observation of the evolution of fresh emissions from the Madrid conurbation in the course of the day. This figure illustrates the strong photochemical formation of O_3 followed by the accumulation

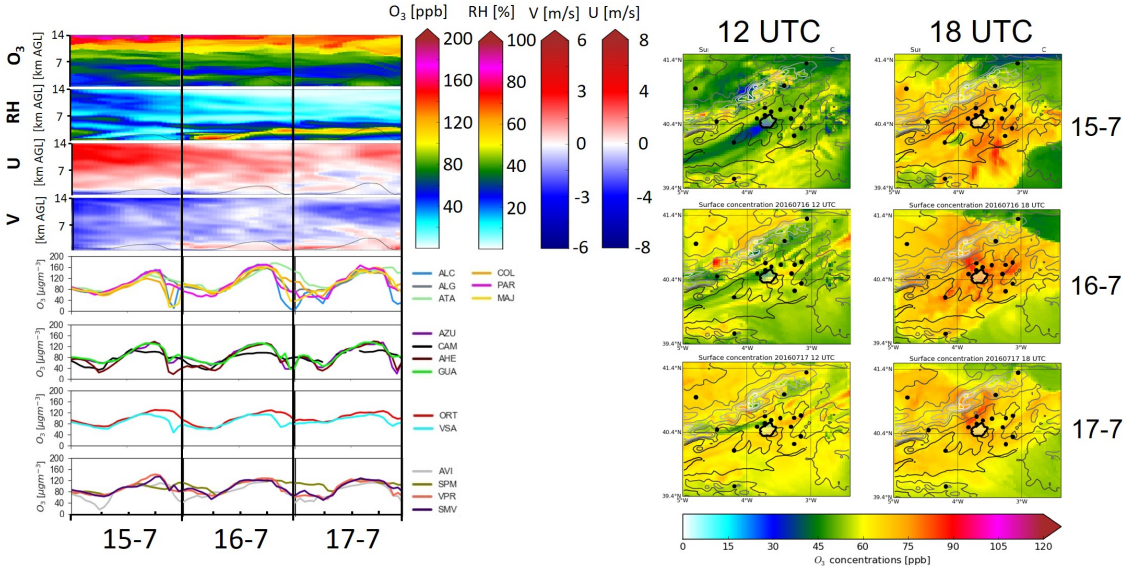


Figure 6. Left: Hourly O_3 concentrations recorded at selected monitoring stations in the MAB and simulated fields of O_3 concentration, relative humidity and U and V components of wind for the period 15–17 July 2016 over the centre of the MAB. Right: Surface O_3 concentration maps obtained from LOTOS-EUROS simulations at 12 and 18 UTC of 15, 16 and 17 July 2016. See Figure 1 to consult the location and typology of the selected stations.

during a typical REC episode. In the first hours of the day, a very shallow boundary layer combined with the stagnant conditions allows locally emitted precursors to accumulate inside the basin. Ozone is then effectively eliminated by titration with NO coming mainly from traffic emissions. This explains the steep drop in O_3 time series during the rush hours (Figure 6) with the exception of rural stations far from the Madrid conurbation like SPM, which showed a more stable behaviour in their O_3 concentrations.

From the simulation resuls, we can interpret the evolution of O_3. At 18 UTC of 16 July, we can see how O_3 levels increased drastically, and the boundary layer depth grew up to 3500 m a.s.l. aided by convection at 18 UTC. Normally, REC events show higher planetary boundary layer (PBL) heights in the evening. Figure 7 shows how the strong convection during REC events injected ground-level pollutants at high altitudes during the late afternoon and the evening reaching up to 3500 m a.s.l. as illustrated in the NO_2 plots. When the

night-time stable boundary layer forms after sunset, air masses with high O_3 that originated near the surface during the previous day were decoupled and remained in the residual layer at altitudes ranging between 2000 and 4000 m a.s.l. forming reservoir layers (00 and 06 UTC cross-sections in Figure 7) which can fumigate the following day. These reservoir layers can also be observed as a relatively thin band at an altitude of 2000–4000 m a.s.l. during every night of the REC period (Figures 6 and S5).

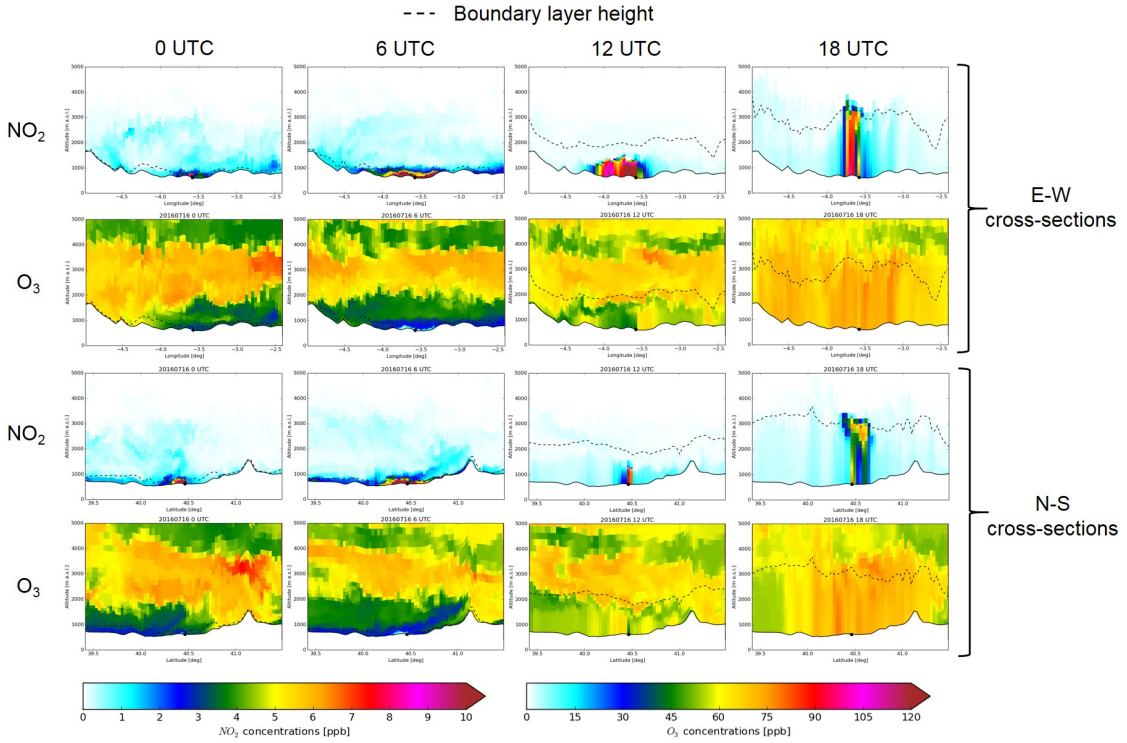


Figure 7. ~~Longitudinal~~ Simulated longitudinal and latitudinal vertical cross-sections of NO_2 and O_3 for 16 July ~~2016.~~2016 (WRF_70 run). See Figure 1 to consult the the latitudinal and longitudinal cuts.

3.2.2 Northern advective events (NAD)

We will refer to NAD events as those during which the dominant situation consisted of the advection of air masses coming from the north over the MAB. During July 2016, the following two periods matched that description: 12–14 and 22–24.

During NAD periods, surface wind is channelled following the NE–SW axis parallel to the Guadarrama range resulting in prevailing north-easterlies in the lowest tropospheric layers while in the upper levels the dominant component is NW (Figures 8, 9 and S5) often associated with the passage of cold fronts from the Atlantic. Winds are generally stronger than in REC events, which implies a renovation of air masses and lower temperatures.

Humidity during NAD events is conditioned by the arrival of air masses off the Atlantic, which are generally moist. During the period 12–14 July, a band of high RH (in the order of 50–60%) that reaches an altitude of approximately 3000 m a.s.l. can be observed (Figure 8). Moreover, a diagonal band of low RH can be observed descending from an altitude of 8000–9000 m a.s.l. at around 10 UTC on 12 July reaching the surface by midday of 14 July. This structure was associated with layers of high O_3 detached from the lower stratosphere as shown in Figure 8. These low-RH stratospheric intrusions were observed also on the 22nd and 23rd of July during the third NAD episode (Figure S5).

The northern winds during NAD episodes push surface air masses with high O_3 towards the south and south-west of the MAB. However, ground-level O_3 concentrations are lower than those observed during REC events. The highest O_3 levels were registered on 12 July at the south-eastern part of the domain (AHE and ORT) with concentrations below 160 μg/m^3. Meanwhile, the maximum concentrations in the rest of stations that day fell below 130 μg/m^3, which is consistent with LOTOS-EUROS surface concentration maps for 12 July (Figure 8). The next two days, winds intensified and veered north resulting in concentrations that did not exceed 130 μg/m^3 along with noteworthy O_3 increments in the rural station of SPM located in the southernmost part of the basin, in accordance with LOTOS-EUROS surface concentration maps of 13 and 14 July (Figure 8). Only one hourly exceedance of 180 μg/m^3 was registered during the six NAD days in all the air-quality stations in the study area in July 2016.

Figure 9 shows how increased advection reduced the residence time of polluted air masses over the region, resulting in lower and less variable O_3 concentrations throughout the day as observed in observations

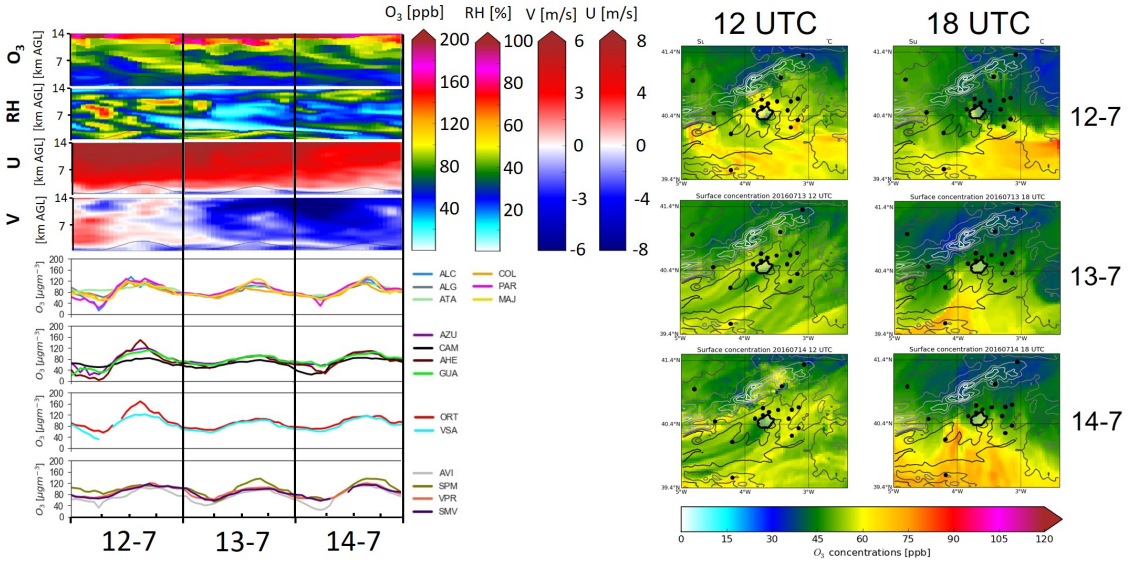


Figure 8. Left: Hourly O_3 concentrations recorded at selected monitoring stations in the MAB and simulated fields of O_3 concentration, relative humidity and U and V components of wind for the period 12–14 July 2016 over the centre of the MAB. Right: Surface O_3 concentration maps obtained from LOTOS-EUROS simulations at 12 and 18 UTC of 12, 13 and 14 July 2016 in the MAB. See Figure 1 to consult the location and typology of the selected stations.

presented in Figure 8. Moreover, the formation of reservoir layers during NAD episodes was less common due to the lower convection, relative to that of REC cases presented before.

It is also remarkable in Figure 9 that above 3000–3500 m a.s.l. O_3 concentrations were very high (in the order of 100 ppb according to the model). This was associated with the stratospheric intrusion of very dry air described above. LOTOS-EUROS reproduced this stratospheric intrusion that was detected from data obtained with free and tethered O_3 soundings for the same period during a field campaign (Querol et al., 2018). The actual impact of this stratospheric intrusion on surface levels remains unclear~~because vertical cross-sections do not allow concluding if O_3 from this layer was effectively transported to the lowest tropospheric levels . In this particular case,~~ . In the referenced paper, the authors estimate a possible but limited impact of the intrusion on surface levels assuming that the boundary layer could exceed the 3000

m a.s.l. during the day. As shown in Figure 9, LOTOS-EUROS predicts that the maximum altitude of the boundary layer according to LOTOS-EUROS reached its maximum values of ~~2500–2700 m a~~2500–2700 ma.s.l. ~~(Figure 9) also~~ limited by the wind ventilation so, probably, the impact on the surface should be low (if any) in this case.

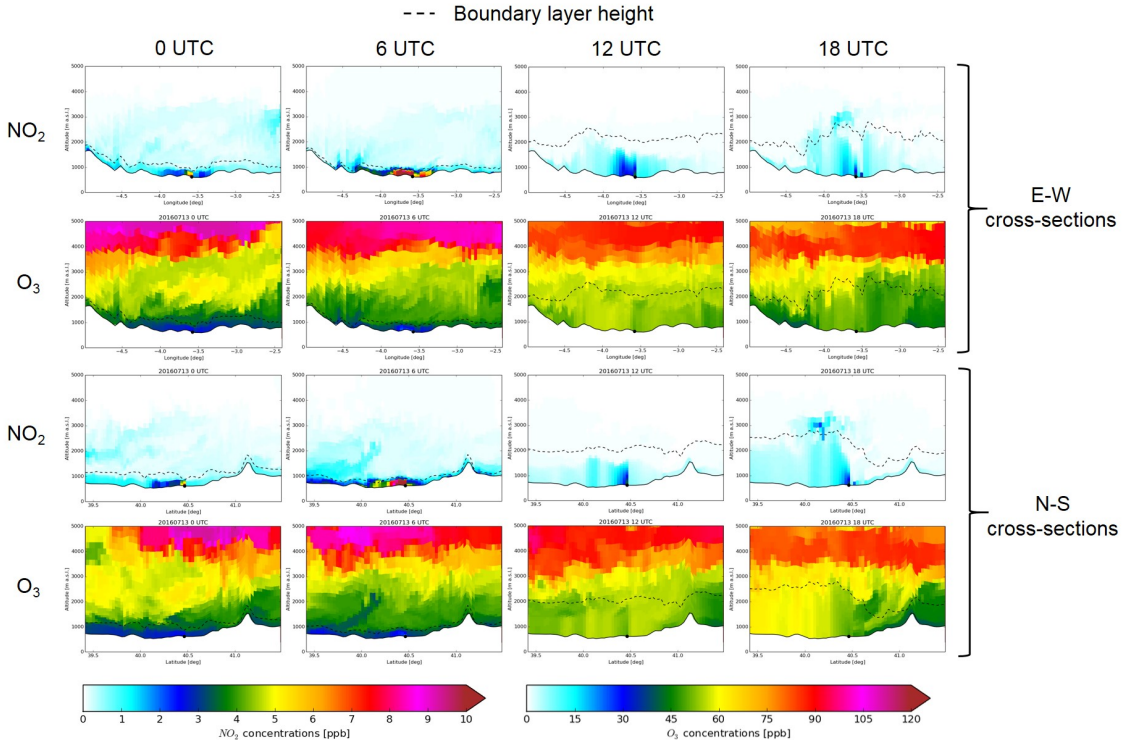


Figure 9. ~~Longitudinal~~ Simulated longitudinal and latitudinal vertical cross-sections of NO_2 and O_3 for 13 July ~~2016.~~2016 (WRF_70 run). See Figure 1 to consult the the latitudinal and longitudinal cuts.

3.2.3 Southern advective events (SAD)

Southern transport implies the arrival of warm air masses, sometimes, coming from northern Africa (maximum temperatures at El Retiro during SAD events in the study period varied between 35.1 and 38.1 °C while

during NAD periods they ranged from 27.1 to 35.1 °C). In July 2016, two SAD periods were observed: 18–21 and 29–31. The first of these two periods has been chosen to illustrate the main features of SAD episodes (Figures 10–12).

SAD events are characterised by constant southerly winds at the surface and at altitude as shown in the simulated fields of U and V (Figures 10 and S5). Because the southern coast of the IP is a densely populated area (especially in summer due to the strong touristic pressure) and anthropogenic emissions of O_3 precursors are high, including industrial emissions around the cities of Huelva, Seville and Algeciras, regional contribution of external O_3 may acquire importance at the basin during SAD events.

High RH values in the middle troposphere are observed during SAD periods (Figures 10 and S5) where relevant increments were registered in the period 19–20 July. Although southerly winds are often associated with rain in the MAB, only small amounts of precipitation were collected during this period.

Analysing the 18–21 July case as a typical example of a SAD episode, we observe in the O_3 time series that concentrations increased in the entire basin earlier in the day than in the other two scenarios. While ground-level O_3 peaks around 18 UTC for REC conditions, daily maxima are found around 12 UTC under SAD patterns, as confirmed in Figure 10. After the intense photochemical formation observed at midday, O_3 is transported towards the NE in the afternoon. This explains the concentration peaks registered at the rural stations located on that side of the basin (ATA and CAM) for example on July 19 or 29 (no figure shown of this day). Other days such as in July 18, the model still produces a peak in the mid/late afternoon and the shift of the O_3 daily maxima is biased by the spatial distribution of stations, which miss the plume that has been advected NW at 18 UTC. SAD episodes are then periods in which O_3 produced at the MAB can be exported towards the north of the IP. Less often, if winds blow from the SE in SAD events, O_3 can be transported across the Guadarrama range towards the NW as happened on 18 July (Figure 10).

The effect of southern winds ventilating the area and some weak rainfall events limited the increase of O_3 although considerable levels were recorded in monitoring stations during SAD events. In general, O_3 concentrations during SAD periods are slightly lower than during REC events (with specific exceptions like 19 July) but higher than in NAD episodes. In the stations located in the basin hourly concentrations rarely exceeded 180 μg/m^3 during SAD days (twice in July 2016 in the seven SAD days) but concentrations above 120 μg/m^3 were more frequent (1083 in total in all the stations in the study area or 155 per day) especially in the stations like ATA and CAM located on the NE of the basin (see 19 July in Figure 10). This proportion

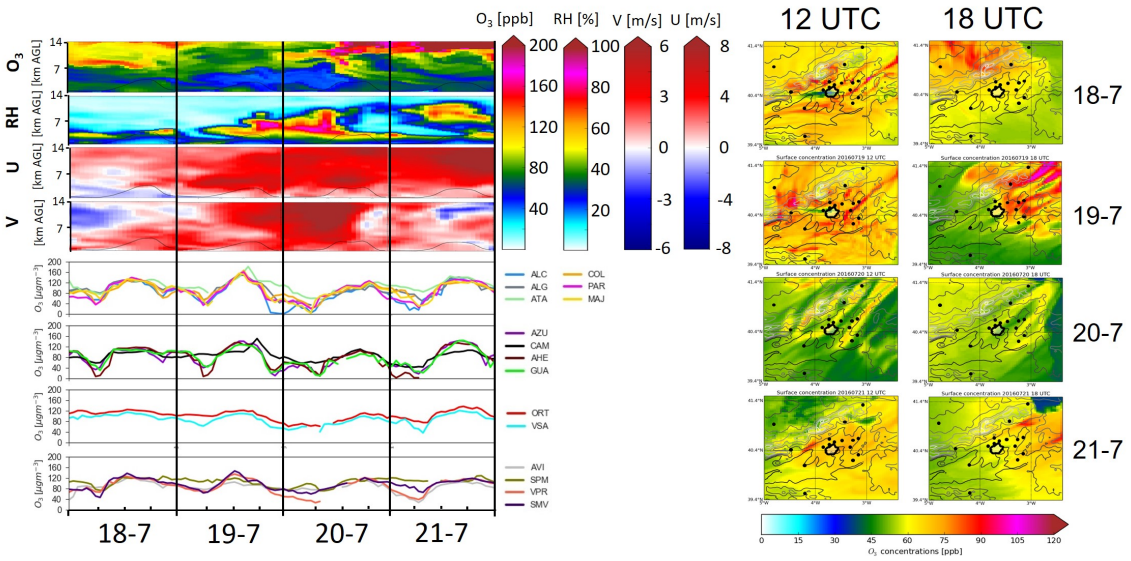


Figure 10. Left: Hourly O_3 concentrations recorded at selected monitoring stations in the MAB and simulated fields of O_3 concentration, relative humidity and U and V components of wind for the period 18–21 July 2016 over the centre of the MAB. Right: Surface O_3 concentration maps obtained from LOTOS-EUROS simulations at 12 and 18 UTC of 18, 19, 20 and 21 July 2016 in the MAB. See Figure 1 to consult the location and typology of the selected stations.

of records above 120 μg/m^3 is higher than during NAD events (average of 56 per day) and lower than the rate registered during REC events (203 per day).

Vertical cross-sections of O_3 and NO_2 on two consecutive days (18 and 19 July) from a SAD period have been used to illustrate the different behaviours observed (Figures 11 and 12). The intense accumulation of NO_2 observed in the 00 and 06 UTC plots points out that on these days, ventilation was not as effective as in NAD events. As a consequence, the O_3 daily cycles showed a considerable drop associated with titration in the morning rush hour unlike on NAD days and closer to the situation of REC episodes (Figure 10). Likewise, for NAD events, vertical mixing is limited as shown in the NO_2 vertical cross-sections of 18 and 19 July preventing the formation of reservoir layers during SAD events. The higher O_3 registered on 19 July

seems to be related to the fact that a deeper boundary layer (maximum heights above 4000 m a.s.l. on 18 July, 3200 m a.s.l. on 19 July) allowed larger dilution, lowering surface concentrations.

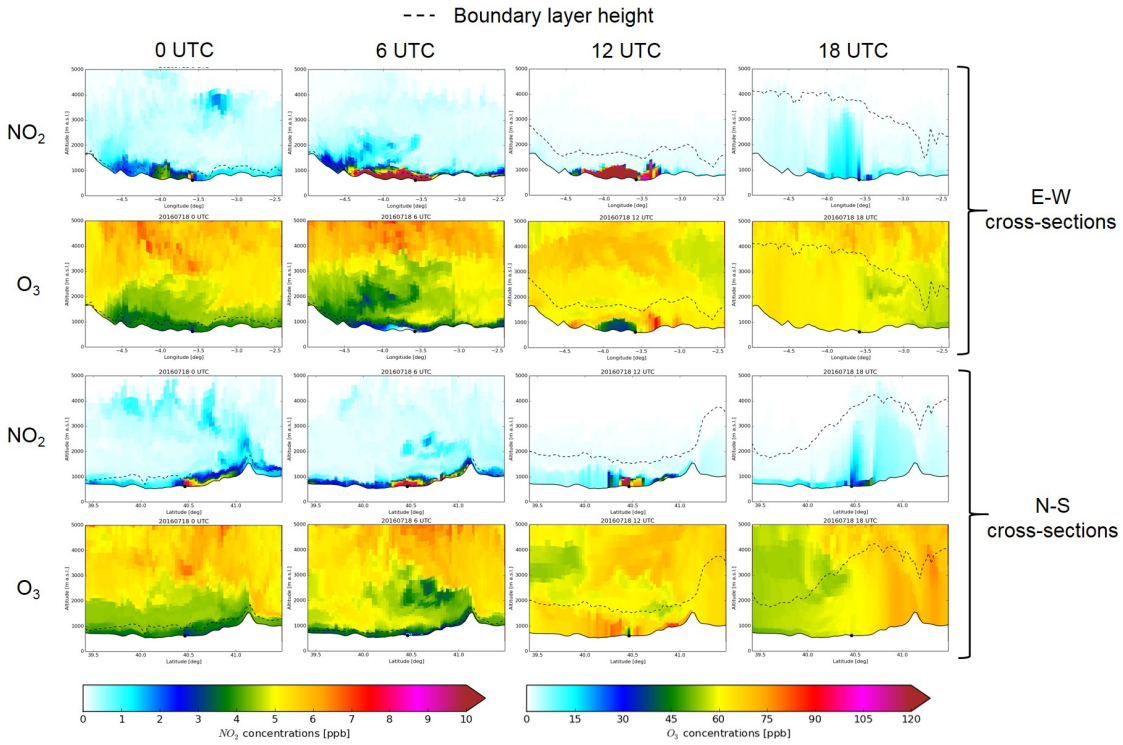


Figure 11. ~~Longitudinal~~ Simulated longitudinal and latitudinal vertical cross-sections of NO_2 and O_3 for 18 July ~~2016.~~2016 (WRF_70 run). See Figure 1 to consult the the latitudinal and longitudinal cuts.

3.2.4 The role of the Boundary Layer Height (BLH)

Figure 13a presents a comparison between observed (data from Querol et al. (2018)) and simulated (WRF_70 configuration) BLH at 12 UTC. We can observe that the model tends to overestimate the BLH at midday although the general trends are captured. In particular, the gradual decrease in the 12 UTC BLH from 11 to 14 July allowed O_3 to accumulate smoothly in the basin, which was described in the aforementioned work,

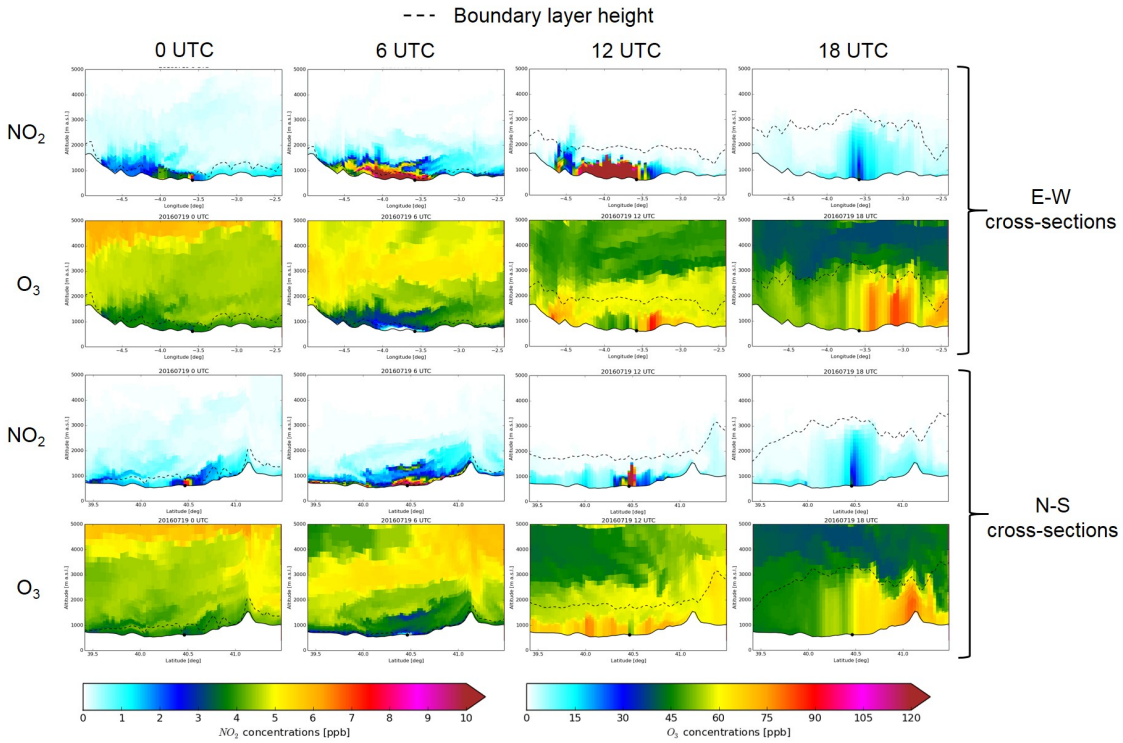


Figure 12. ~~Longitudinal~~ Simulated longitudinal and latitudinal vertical cross-sections of NO_2 and O_3 for 19 July ~~2016.~~2016 (WRF_70 run). See Figure 1 to consult the the latitudinal and longitudinal cuts.

is also observed in the simulated data. The overestimation is slight on most days although larger differences are observed in certain periods (5–10 and 29–31 July).

Querol et al. (2018) describe lower midday BLH in O_3 accumulation episodes (equivalent to REC events described here) than in venting episodes (NAD or SAD). The simulations with LOTOS-EUROS confirm that finding as observed in Figure 13b. The mixing layer is deeper on 13 July (NAD) than on 16 July (REC) from 0 to 16 UTC which includes the period of most effective photochemical formation of O_3. This allows a more effective formation of reservoir layers during REC events that fumigate to the surface as the diurnal

convective circulation develops. After 16 UTC the BLH during the REC event grows higher than during the NAD day due to the larger convection in the second scenario.

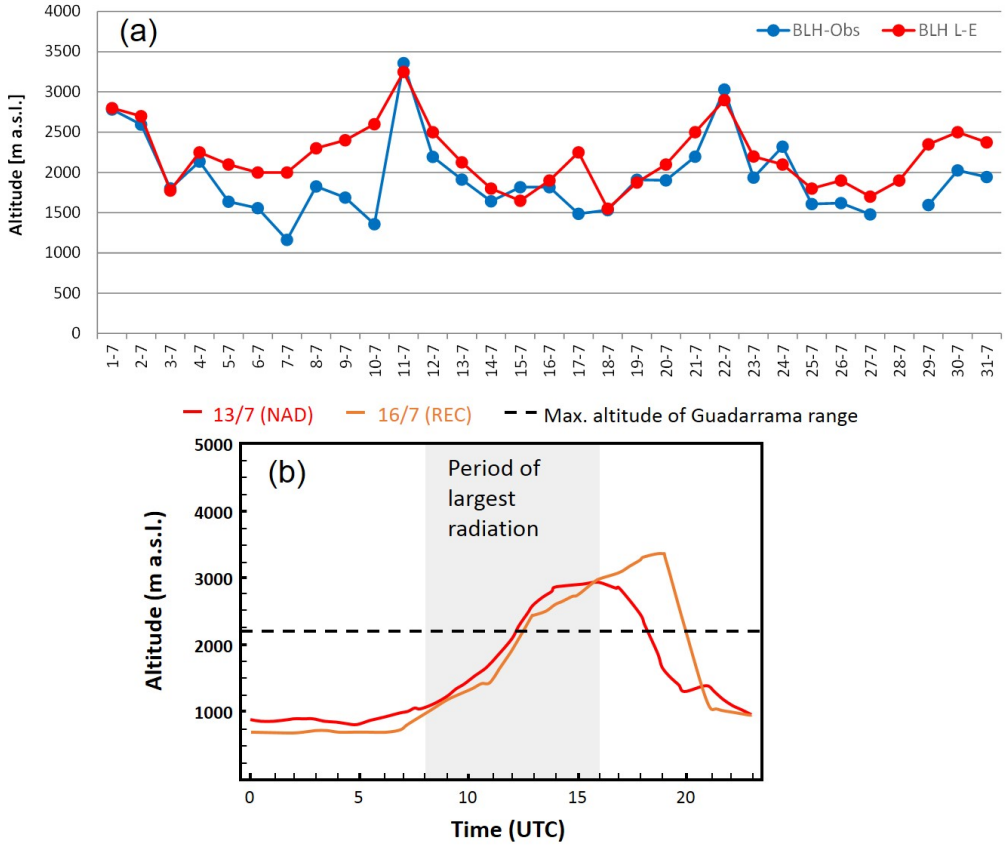


Figure 13. (a) Time series of the estimated and modelled midday BLH over Madrid airport for July 2016. Estimations calculated from the daily AEMET radio-soundings using the simple parcel method. (b) Modelled BLH for the 13 (NAD event) and the 16 (REC event) July 2016.

4 Conclusions

Evaluation of a CTM is a basic tool for the analysis and forecasting of photochemical processes that give rise to high concentrations of tropospheric O_3 that frequently occur in the Mediterranean in summer. A preliminary requirement for the application of CTM for policy decisions is that they could reproduce adequately the processes and mechanisms identified by the field campaigns and reasonably reproduce the observations of the monitoring stations, especially during acute O_3 episodes.

In this work, we present the results obtained from a simulation exercise (July 2016) performed with the LOTOS-EUROS CTM over the MAB, representative of summer conditions. Five configurations with different combinations of spatial resolution (25 x 25 and 3 x 3 km^2), input meteorological data (ECMWF 7 x 7 km^2 for the IP and WRF, 1 x 1 km^2 for the MAB) and vertical structures (mixed-layer scheme with five altitude levels and hybrid-layer scheme with 70 altitude levels) for model evaluation and optimisation.

The main objective of the paper is to provide a phenomenological interpretation of O_3 events in the area after performing a detailed evaluation of the best configuration of the model for the specific area and period. Regarding the specific question of the reasonable number of vertical levels in the model configuration, it is dependent on the objective of the study. In this study the environmental analysis was the main objective and it was logical and feasible from the perspective of CPU time to employ a considerable number of vertical levels because it allowed a better representation of the vertical variability of O_3. In other studies such as air quality forecasting or long term analyses in which CPU time may be large, the reasonable number of levels can be less.

Our results show that the LOTOS-EUROS model performs in a satisfactory manner in the five set-ups. However, regarding surface O_3, it is clear that the model benefits from finer spatial resolutions in the horizontal and also from the use of multilayered vertical schemes. As a result, WRF_70 and ECMWF_HR_70 were the optimal configurations.

Using multilayered 70 level set-ups, LOTOS-EUROS was able to reproduce the vertical gradients of O_3 in the Madrid basin, although in some cases the model presented an overestimation in the lower levels with respect to observations. In most cases, the model was also able to reproduce features like fine O_3 layers. The performance of LOTOS-EUROS was partly successful, differentiating the vertical structure of O_3 under distinct meteorological conditions so further research is needed to improve CTMs' performances in this

particular aspect with, for example, comparisons with data from O_3 soundings under different meteorological scenarios.

Therefore, the modelling system is suitable to be employed for the interpretation of O_3 variability in the region. In light of the present study, we suggest using vertical schemes of CTMs with a sufficient number of levels for capturing O_3 variability in the simulations of summer episodes in the Mediterranean region.

Employing the WRF_70 configuration of LOTOS-EUROS which has shown the best performance simulating surface and vertical concentrations of O_3 in the MAB, we interpreted the variability of O_3 in the region. Three episode types have been identified regarding the dominating circulation. Two of them are associated with advection, either from the north (NAD) or from the south (SAD), while the third is associated with local/regional recirculation of air masses (REC). REC events are characterised by low winds that veer during the day from NE to SW following the axis of the Guadarrama range. These stagnant conditions combined with the strong insolation and temperature registered during REC events favour the strong photochemical production and accumulation of O_3 over the MAB, likely to exceed 180 μg/m^3 during the afternoon and evening . Moreover, the strong convection helps to form reservoir layers located at 2000–4000 m a.s.l. which contribute to increasing surface O_3 in the following days when the upper limit of the boundary layer reaches those altitudes.

Marked differences have been found between the two venting episodes. During SAD episodes, winds are weak and external O_3 contributions from the south and south-east of the IP can be relevant, while in the case of NAD events, winds were generally stronger, favouring ventilation. As a consequence, surface O_3 during NAD events did not grow excessively (except in specific cases when wind speed is low). During SAD conditions, higher base O_3 concentrations ($>$120 μg/m^3) were registered but the 180 μg/m^3 threshold is exceeded rarely in both episode types. One of the factors for this is the existence of a steady wind direction avoiding an effective accumulation of O_3 in reservoir layers. Both NAD and SAD events are associated with O_3 exportation to other air basins on the IP like the Ebro valley (to the NE) and the Tagus and Guadiana valleys (to the SW).

Intrusions of stratospheric O_3 have been observed with LOTOS-EUROS simulations in the form of bands with a high concentration of O_3 and very low humidity. It is unclear whether these intrusions have an impact at the surface and, if so, what is the exact contribution to the O_3 observed there. Specific model-based analyses of these episodes should be performed to evaluate their actual impact on surface O_3 in the MAB.

The results from this study can be useful to understand the phenomenology of high O_3 episodes in the MAB and to gain knowledge to design appropriate strategies for air-quality management. Further research must be implemented to investigate aspects like the sensitivity to emission reduction scenarios or the role of VOCs with emphasis on the biogenic ones. Moreover, to perform the tasks of validating and optimising CTMs, increasing efforts should be made to conduct more field campaigns in different air basins in the Mediterranean using state-of-the-art equipment to generate data and knowledge about O_3 behaviour both on the surface and vertically. Useful parameters to be included in these campaigns are O_3, NO_x, VOC and, when possible, intermediate products like NO_y, HNO_3 and H_2O_2 that, according to previous experience (Sillman, 1995), are key parameters for facing model-based NO_x–VOC sensitivity studies and the assessment of emission inventories. Some of these parameters (especially NO_x) should also be incorporated in vertical measurements.

In future, similar simulations to the one presented in this study should be performed in the different air basins in the IP where O_3 exceedances have been recorded (Querol et al., 2016). CTMs should be configured specifically for each region or air basin to assure the best performance by capturing the influence of topography and local circulations. For such studies, we highlight the importance of conducting experimental campaigns that can support the necessary model evaluation.

Finally, it should be noted that when running such fine resolutions for real applications it is also important to work on the emission datasets (out of the scope of this work). Increasing the detail in emission inventory (mainly based on a bottom-up approach) could improve the performance of CTMs when assessing sensitivities or emission scenarios. Moreover, improving time resolution in the emission models can be beneficial for simulating O_3 episodes.

Author contributions. Dr. Escudero conceived the presented idea, carried out the simulations with LOTOS-EUROS, collected and treated experimental data used for model evaluation and wrote the manuscript. Drs. Segers, Kranenburg and Schaap supervised LOTOS-EUROS simulations, modified the model code for performing runs with the different configurations and contributed in the post-processing of model outputs. Drs. Borge and de la Paz performed and validated simulations with WRF. Drs. Querol, Alastuey and Gangoiti contributed especially in the interpretation of O_3 phenomenology. All authors discussed the results and contributed to the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the Ministry of Economy, Industry and Competitiveness and FEDER funds through the project HOUSE (CGL2016-78594-R), the Ministry of Agriculture, Fishing, Food and Environment, the Madrid City Council, the Madrid Regional Government and by the Department of Research, Innovation and University of the Aragón Regional Government and the European Social Fund (project E23_17D). The study was also partially supported by the scientific programme TECNAIRE-CM funded by the Directorate General for Universities and Research of the Greater Madrid Region (S2013/MAE-2972). The authors gratefully acknowledge air-quality data provision by the following entities: MITECO, Madrid City Council, AEMET and the Autonomous Communities of Madrid, Castilla León and Castilla La Mancha. Dr Escudero received a grant from the José Castillejo programme of the Ministry of Education and Science of Spain (ref. CAS17/00108) for a 6-month research visit at TNO.

References

Baldasano, J., Pay, M., Jorba, O., Gassó, S., and Jiménez-Guerrero, P.: An annual assessment of air quality with the CALIOPE modeling system over Spain, Science of The Total Environment, 409, 2163 – 2178, https://doi.org/https://doi.org/10.1016/j.scitotenv.2011.01.041, http://www.sciencedirect.com/science/article/pii/S0048969711000787, 2011.

Barbu, A., Segers, A., Schaap, M., Heemink, A., and Builtjes, P.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, Atmospheric Environment, 43, 1622 – 1631, https://doi.org/https://doi.org/10.1016/j.atmosenv.2008.12.005, http://www.sciencedirect.com/science/article/pii/S1352231008011266, 2009.

Beltman, J. B., Hendriks, C., Tum, M., and Schaap, M.: The impact of large scale biomass production on ozone air pollution in Europe, Atmospheric Environment, 71, 352 – 363, https://doi.org/https://doi.org/10.1016/j.atmosenv.2013.02.019, http://www.sciencedirect.com/science/article/pii/S1352231013001209, 2013.

Borge, R., Alexandrov, V., del Vas, J. J., Lumbreras, J., and Rodríguez, E.: A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula, Atmospheric Environment, 42, 8560 – 8574, https://doi.org/https://doi.org/10.1016/j.atmosenv.2008.08.032, http://www.sciencedirect.com/science/article/pii/S1352231008007954, 2008.

Borge, R., López, J., Lumbreras, J., Narros, A., and Rodríguez, E.: Influence of boundary conditions on CMAQ simulations over the Iberian Peninsula, Atmospheric Environment, 44, 2681 – 2695, https://doi.org/https://doi.org/10.1016/j.atmosenv.2010.04.044, http://www.sciencedirect.com/science/article/pii/S1352231010003511, 2010.

Borge, R., Lumbreras, J., Pérez, J., de la Paz, D., Vedrenne, M., de Andrés, J. M., and Rodríguez, M. E.: Emission inventories and modeling requirements for the development of air quality plans. Application to Madrid (Spain), Science of The Total Environment, 466-467, 809 – 819, https://doi.org/https://doi.org/10.1016/j.scitotenv.2013.07.093, http://www.sciencedirect.com/science/article/pii/S0048969713008759, 2014.

Borrego, C., Monteiro, A., Ferreira, J., Miranda, A., Costa, A., Carvalho, A., and Lopes, M.: Procedures for estimation of modelling uncertainty in air quality assessment, Environment International, 34, 613 – 620, https://doi.org/https://doi.org/10.1016/j.envint.2007.12.005, http://www.sciencedirect.com/science/article/pii/S0160412007002346, assessment of Urban and Regional Air Quality and its Impacts, 2008.

Canepa, E. and Builtjes, P. J.: Thoughts on Earth System Modeling: From global to regional scale, Earth-Science Reviews, 171, 456 – 462, https://doi.org/https://doi.org/10.1016/j.earscirev.2017.06.017, http://www.sciencedirect.com/science/article/pii/S0012825216304688, 2017.

Carnerero, C., Pérez, N., Reche, C., Ealo, M., Titos, G., Lee, H.-K., Eun, H.-R., Park, Y.-H., Dada, L., Paasonen, P., Kerminen, V.-M., Mantilla, E., Escudero, M., Gómez-Moreno, F. J., Alonso-Blanco, E., Coz, E., Saiz-López, A., Temime-Roussel, B., Marchand, N., Beddows, D. C. S., Harrison, R. M., Petäjä, T., Kulmala, M., Ahn, K.-H., Alastuey, A., and Querol, X.: Vertical and horizontal distribution of regional new particle formation events in Madrid, Atmospheric Chemistry and Physics, 18, 16 601–16 618, https://doi.org/10.5194/acp-18-16601-2018, https://www.atmos-chem-phys.net/18/16601/2018/, 2018.

Carvalho, A., Carvalho, A., Gelpi, I., Barreiro, M., Borrego, C., Miranda, A., and Pérez-Muñúzuri, V.: Influence of topography and land use on pollutants dispersion in the Atlantic coast of Iberian Peninsula, Atmospheric Environment, 40, 3969 – 3982, https://doi.org/https://doi.org/10.1016/j.atmosenv.2006.02.014, http://www.sciencedirect.com/science/article/pii/S135223100600207X, 2006.

Curier, R., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport model and in-situ measurements, Atmospheric Environment, 60, 217 – 226, https://doi.org/https://doi.org/10.1016/j.atmosenv.2012.06.017, http://www.sciencedirect.com/science/article/pii/S1352231012005687, 2012.

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet, B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Carnavale, C., Coppalle, A., Denby, B., Douros, J., Graf, A., Hellmuth, O., Hodzic, A., Honoré, C., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi, E., Moussiopoulos, N., Pertot, C., Peuch, V., Pirovano, G., Rouil, L., Sauter, F., Schaap, M., Stern, R., Tarrasón, L., Vignati, E., Volta, M., White, L., Wind, P., and Zuber, A.: CityDelta: A model intercomparison study to explore the impact of emission reductions in European cities in 2010, Atmospheric Environment, 41, 189 – 207, https://doi.org/https://doi.org/10.1016/j.atmosenv.2006.07.036, http://www.sciencedirect.com/science/article/pii/S1352231006008053, 2007.

de la Paz, D., Borge, R., and Martilli, A.: Assessment of a high resolution annual WRF-BEP/CMAQ simulation for the urban area of Madrid (Spain), Atmospheric Environment, 144, 282 – 296, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.08.082, http://www.sciencedirect.com/science/article/pii/S1352231016306884, 2016.

Eckermann, S.: Hybrid σ–p Coordinate Choices for a Global Model, Monthly Weather Review, 137, 224–245, https://doi.org/10.1175/2008MWR2537.1, https://doi.org/10.1175/2008MWR2537.1, 2009.

EEA: Air quality in Europe-2017 report, Tech. Rep. EEA Report, No 12/2018, Luxembourg, https://doi.org/10.2800/777411, https://www.eea.europa.eu/publications/air-quality-in-europe-2017, 2018.

Escudero, M., Lozano, A., Hierro, J., del Valle, J., and Mantilla, E.: Urban influence on increasing ozone concentrations in a characteristic Mediterranean agglomeration, Atmospheric Environment, 99, 322 – 332, https://doi.org/http://dx.doi.org/10.1016/j.atmosenv.2014.09.061, 2014.

Escudero, M., Lozano, A., Hierro, J., Tapia, O., del Valle, J., Alastuey, A., Moreno, T., Anzano, J., and Querol, X.: Assessment of the variability of atmospheric pollution in National Parks of mainland Spain, Atmospheric Environment, 132, 332 – 344, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.03.006, http://www.sciencedirect.com/science/article/pii/S1352231016301765, 2016.

Fenech, S., Doherty, R. M., Heaviside, C., Vardoulakis, S., Macintyre, H. L., and O'Connor, F. M.: The influence of model spatial resolution on simulated ozone and fine particulate matter for Europe: implications for health impact assessments, Atmospheric Chemistry and Physics, 18, 5765–5784, https://doi.org/10.5194/acp-18-5765-2018, https://www.atmos-chem-phys.net/18/5765/2018/, 2018.

Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, Geoscientific Model Development, 2, 253–265, https://doi.org/10.5194/gmd-2-253-2009, https://www.geosci-model-dev.net/2/253/2009/, 2009.

Gangoiti, G., Millán, M. M., Salvador, R., and Mantilla, E.: Long-range transport and re-circulation of pollutants in the western Mediterranean during the project Regional Cycles of Air Pollution in the West-Central Mediterranean Area, Atmospheric Environment, 35, 6267 – 6276, https://doi.org/https://doi.org/10.1016/S1352-2310(01)00440-X, http://www.sciencedirect.com/science/article/pii/S135223100100440X, 2001.

Hass, H., Builtjes, P., Simpson, D., and Sternii, R.: Comparison of model results obtained with several european regional air quality models, Atmospheric Environment, 31, 3259 – 3279, https://doi.org/https://doi.org/10.1016/S1352-2310(97)00066-6, http://www.sciencedirect.com/science/article/pii/S1352231097000666, eUMAC: European Modelling of Atmospheric Constituents, 1997.

Hendriks, C., Forsell, N., Kiesewetter, G., Schaap, M., and Schöpp, W.: Ozone concentrations and damage for realistic future European climate and air quality scenarios, Atmospheric Environment, 144, 208 – 219, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.08.026, http://www.sciencedirect.com/science/article/pii/S1352231016306197, 2016.

Hess, P. G. and Zbinden, R.: Stratospheric impact on tropospheric ozone variability and trends: 1990-2009, Atmospheric Chemistry and Physics, 13, 649–674, https://doi.org/10.5194/acp-13-649-2013, https://www.atmos-chem-phys.net/13/649/2013/, 2013.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorbá, O., Knote, C., Kuenen, J. J., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G.,

Jose, R. S., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Evaluation of operational on-line-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part I: Ozone, Atmospheric Environment, 115, 404 – 420, https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.09.042, http://www.sciencedirect.com/science/article/pii/S1352231014007353, 2015.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University press, 2013.

Jiménez, P., Parra, R., Gassó, S., and Baldasano, J. M.: Modeling the ozone weekend effect in very complex terrains: a case study in the Northeastern Iberian Peninsula, Atmospheric Environment, 39, 429 – 444, https://doi.org/https://doi.org/10.1016/j.atmosenv.2004.09.065, http://www.sciencedirect.com/science/article/pii/S1352231004009458, 2005.

Jiménez, P., Lelieveld, J., and Baldasano, J. M.: Multiscale modeling of air pollutants dynamics in the north-western Mediterranean basin during a typical summertime episode, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006516, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006516, 2006.

Jiménez-Guerrero, P., Jorba, O., Baldasano, J. M., and Gassó, S.: The use of a modelling system as a tool for air quality management: Annual high-resolution simulations and evaluation, Science of The Total Environment, 390, 323 – 340, https://doi.org/https://doi.org/10.1016/j.scitotenv.2007.10.025, http://www.sciencedirect.com/science/article/pii/S0048969707010959, 2008.

Lee, D. S., Holland, M. R., and Falla, N.: The potential impact of ozone on materials in the U.K., Atmospheric Environment, 30, 1053 – 1065, https://doi.org/https://doi.org/10.1016/1352-2310(95)00407-6, 1996.

Manders, A., Schaap, M., and Hoogerbrugge, R.: Testing the capability of the chemistry transport model LOTOS-EUROS to forecast PM10 levels in the Netherlands, Atmospheric Environment, 43, 4050 – 4059, https://doi.org/https://doi.org/10.1016/j.atmosenv.2009.05.006, http://www.sciencedirect.com/science/article/pii/S1352231009004269, 2009.

Manders, A. M. M., Segers, A., and Jonkers, S.: LOTOS-EUROS v2.0 Reference guide, Tech. Rep. TNO 2016 R10898, TNO, 2016.

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier van der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J. P., Segers, A. J., Timmermans, R. M. A., Visschedijk, A. J. H., Wichink Kruit, R. J., van Pul, W. A. J., Sauter, F. J., van der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A. C., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., and Schaap, M.: Curriculum

vitae of the LOTOS–EUROS (v2.0) chemistry transport model, Geoscientific Model Development, 10, 4145–4173, https://doi.org/10.5194/gmd-10-4145-2017, https://www.geosci-model-dev.net/10/4145/2017/, 2017.

Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Denier van der Gon, H. A. C., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, A., van Velthoven, P., van Versendaal, R., Vira, J., and Ung, A.: A regional air quality forecasting system over Europe: the MACC-II daily ensemble production, Geoscientific Model Development, 8, 2777–2813, https://doi.org/10.5194/gmd-8-2777-2015, https://www.geosci-model-dev.net/8/2777/2015/, 2015.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A.: Does increasing horizontal resolution produce a more skillful forecasts?, Bulletin of the American Meteorological Society, 83, 407–430, https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2, https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2, 2002.

Massagué, J., Carnerero, C., Escudero, M., Baldasano, J. M., Alastuey, A., and Querol, X.: 2005–2017 ozone trends and potential benefits of local measures as deduced from air quality measurements in the north of the Barcelona Metropolitan Area, Atmospheric Chemistry and Physics Discussions, 2019, 1–28, https://doi.org/10.5194/acp-2019-33, https://www.atmos-chem-phys-discuss.net/acp-2019-33/, 2019.

Millán, M. M., Salvador, R., Mantilla, E., and Kallos, G.: Photooxidant dynamics in the Mediterranean basin in summer: Results from European research projects, Journal of Geophysical Research: Atmospheres, 102, 8811–8823, https://doi.org/10.1029/96JD03610, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03610, 1997.

Millán, M. M., Mantilla, E., Salvador, R., Carratalá, A., Sanz, M. J., Alonso, L., Gangoiti, G., and Navazo, M.: Ozone Cycles in the Western Mediterranean Basin: Interpretation of Monitoring Data in Complex Coastal Terrain, Journal of Applied Meteorology, 39, 487–508, https://doi.org/10.1175/1520-0450(2000)039<0487:OCITWM>2.0.CO;2, https://doi.org/10.1175/1520-0450(2000)039<0487:OCITWM>2.0.CO;2, 2000.

Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, Atmospheric Chemistry and Physics, 15, 8889–8973, https://doi.org/10.5194/acp-15-8889-2015, https://www.atmos-chem-phys.net/15/8889/2015/, 2015.

Palacios, M., Kirchner, F., Martilli, A., Clappier, A., Martín, F., and Rodríguez, M.: Summer ozone episodes in the Greater Madrid area. Analyzing the ozone response to abatement strategies by modelling, Atmospheric Environment, 36, 5323 – 5333, https://doi.org/https://doi.org/10.1016/S1352-2310(02)00590-3, http://www.sciencedirect.com/science/article/pii/S1352231002005903, 2002.

Paoletti, E.: Impact of ozone on Mediterranean forests: A review, Environmental Pollution, 144, 463 – 474, https://doi.org/http://dx.doi.org/10.1016/j.envpol.2005.12.051, 2006.

Pay, M. T., Gangoiti, G., Guevara, M., Napelenok, S., Querol, X., Jorba, O., and Pérez García-Pando, C.: Ozone source apportionment during peak summer events over southwestern Europe, Atmospheric Chemistry and Physics Discussions, 2018, 1–43, https://doi.org/10.5194/acp-2018-727, https://www.atmos-chem-phys-discuss.net/acp-2018-727/, 2018.

Plaza, J., Pujadas, M., and Artíñano, B.: Formation and Transport of the Madrid Ozone Plume, Journal of the Air & Waste Management Association, 47, 766–774, https://doi.org/10.1080/10473289.1997.10463938, http://dx.doi.org/10.1080/10473289.1997.10463938, 1997.

Querol, X., Alastuey, A., Reche, C., Orío, A., Pallarés, M., Reina, F., Diéguez, J., Mantilla, E., Escudero, M., Alonso, L., Gangoiti, G., and Millán, M.: On the origin of the highest ozone episodes in Spain, Science of The Total Environment, 572, 379 – 389, https://doi.org/https://doi.org/10.1016/j.scitotenv.2016.07.193, http://www.sciencedirect.com/science/article/pii/S0048969716316473, 2016.

Querol, X., Gangoiti, G., Mantilla, E., Alastuey, A., Minguillón, M. C., Amato, F., Reche, C., Viana, M., Moreno, T., Karanasiou, A., Rivas, I., Pérez, N., Ripoll, A., Brines, M., Ealo, M., Pandolfi, M., Lee, H.-K., Eun, H.-R., Park, Y.-H., Escudero, M., Beddows, D., Harrison, R. M., Bertrand, A., Marchand, N., Lyasota, A., Codina, B., Olid, M., Udina, M., Jiménez-Esteve, B., Sóler, M. R., Alonso, L., Millán, M., and Ahn, K.-H.: Phenomenology of high-ozone episodes in NE Spain, Atmospheric Chemistry and Physics, 17, 2817–2838, https://doi.org/10.5194/acp-17-2817-2017, https://www.atmos-chem-phys.net/17/2817/2017/, 2017.

Querol, X., Alastuey, A., Gangoiti, G., Perez, N., Lee, H. K., Eun, H. R., Park, Y., Mantilla, E., Escudero, M., Titos, G., Alonso, L., Temime-Roussel, B., Marchand, N., Moreta, J. R., Revuelta, M. A., Salvador, P., Artíñano, B., García dos Santos, S., Anguas, M., Notario, A., Saiz-Lopez, A., Harrison, R. M., Millán, M., and Ahn, K.-H.: Phenomenology of summer ozone episodes over the Madrid Metropolitan Area, central Spain, Atmospheric Chemistry and Physics, 18, 6511–6533, https://doi.org/10.5194/acp-18-6511-2018, https://www.atmos-chem-phys.net/18/6511/2018/, 2018.

Reche, C., Moreno, T., Amato, F., Pandolfi, M., Pérez, J., de la Paz, D., Diaz, E., Gómez-Moreno, F., Pujadas, M., Artíñano, B., Reina, F., Orio, A., Pallarés, M., Escudero, M., Tapia, O., Crespo, E., Vargas, R., Alastuey, A., and Querol, X.: Spatio-temporal patterns of high summer ozone events in the Madrid Basin, Central Spain, At-

mospheric Environment, 185, 207 – 220, https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.05.002, http://www.sciencedirect.com/science/article/pii/S1352231018302991, 2018.

Saiz-López, A., Borge, R., Notario, A., Adame, J. A., de la Paz, D., Querol, X., Artíñano, B., Gómez-Moreno, F. J., and Cuevas, C. A.: Unexpected increase in the oxidation capacity of the urban atmosphere of Madrid, Spain, Scientific Reports, 7, 45 956, https://doi.org/10.1038/srep45956, 2017.

Salvador, P., Artíñano, B., Viana, M., Alastuey, A., and Querol, X.: Multicriteria approach to interpret the variability of the levels of particulate matter and gaseous pollutants in the Madrid metropolitan area, during the 1999–2012 period, Atmospheric Environment, 109, 205 – 216, https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.03.008, http://www.sciencedirect.com/science/article/pii/S1352231015002228, 2015.

San José, R., Stohl, A., Karatzas, K., Bohler, T., James, P., and Pérez, J.: A modelling study of an extraordinary night time ozone episode over Madrid domain, Environmental Modelling & Software, 20, 587 – 593, https://doi.org/https://doi.org/10.1016/j.envsoft.2004.03.009, http://www.sciencedirect.com/science/article/pii/S136481520400091X, 2005.

Schaap, M., van Loon, M., ten Brink, H. M., Dentener, F. J., and Builtjes, P. J. H.: Secondary inorganic aerosol simulations for Europe with special attention to nitrate, Atmospheric Chemistry and Physics, 4, 857–874, https://doi.org/10.5194/acp-4-857-2004, https://www.atmos-chem-phys.net/4/857/2004/, 2004.

Schaap, M., Kranenburg, R., Curier, L., Jozwicka, M., Dammers, E., and Timmermans, R.: Assessing the Sensitivity of the OMI-NO_2 Product to Emission Changes across Europe, Remote Sensing, 5, 4187–4208, https://doi.org/10.3390/rs5094187, http://www.mdpi.com/2072-4292/5/9/4187, 2013.

Screpanti, A. and Marco, A. D.: Corrosion on cultural heritage buildings in Italy: A role for ozone?, Environmental Pollution, 157, 1513 – 1520, https://doi.org/http://dx.doi.org/10.1016/j.envpol.2008.09.046, special Issue Section: Ozone and Mediterranean Ecology: Plants, People, Problems, 2009.

Seco, R., Peñuelas, J., Filella, I., Llusià, J., Molowny-Horas, R., Schallhart, S., Metzger, A., Müller, M., and Hansel, A.: Contrasting winter and summer VOC mixing ratios at a forest site in the Western Mediterranean Basin: the effect of local biogenic emissions, Atmospheric Chemistry and Physics, 11, 13 161–13 179, https://doi.org/10.5194/acp-11-13161-2011, https://www.atmos-chem-phys.net/11/13161/2011/, 2011.

Sicard, P., De Marco, A., Troussier, F., Renou, C., Vas, N., and Paoletti, E.: Decrease in surface ozone concentrations at Mediterranean remote sites and increase in the cities, Atmospheric Environment, 79, 705–715, 2013.

Sillman, S.: The use of NO_y, H_2O_2, and HNO_3 as indicators for ozone-NO_x-hydrocarbon sensitivity in urban locations, Journal of Geophysical Research: Atmospheres, 100, 14 175–14 188, https://doi.org/10.1029/94JD02953, https://agupubs.pericles-prod.literatumonline.com/doi/abs/10.1029/94JD02953, 1995.

Sillman, S.: The relation between ozone, NO_x and hydrocarbons in urban and polluted rural environments, Atmospheric Environment, 33, 1821 – 1845, https://doi.org/https://doi.org/10.1016/S1352-2310(98)00345-8, http://www.sciencedirect.com/science/article/pii/S1352231098003458, 1999.

Sillman, S. and West, J. J.: Reactive nitrogen in Mexico City and its relation to ozone-precursor sensitivity: results from photochemical models, Atmospheric Chemistry and Physics, 9, 3477–3489, https://doi.org/10.5194/acp-9-3477-2009, https://www.atmos-chem-phys.net/9/3477/2009/, 2009.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, Tech. rep., NCAR Tech. Note NCAR/TN-475+STR, https://doi.org/10.5065/D68S4MVH, 2008.

Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., van der Gon, H. D., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jeričević, A., Kraljević, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S., and Galmarini, S.: Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII, Atmospheric Environment, 53, 60 – 74, https://doi.org/https://doi.org/10.1016/j.atmosenv.2012.01.003, http://www.sciencedirect.com/science/article/pii/S1352231012000064, 2012.

Soret, A., Guevara, M., and Baldasano, J.: The potential impacts of electric vehicles on air quality in the urban areas of Barcelona and Madrid (Spain), Atmospheric Environment, 99, 51 – 63, https://doi.org/https://doi.org/10.1016/j.atmosenv.2014.09.048, http://www.sciencedirect.com/science/article/pii/S1352231014007419, 2014.

Stevenson, D. S., Dentener, F. J., Schultz, M. G., Ellingsen, K., van Noije, T. P. C., Wild, O., Zeng, G., Amann, M., Atherton, C. S., Bell, N., Bergmann, D. J., Bey, I., Butler, T., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Krol, M. C., Lamarque, J.-F., Lawrence, M. G., Montanaro, V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Shindell, D. T., Strahan, S. E., Sudo, K., and Szopa, S.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006338, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006338, 2006.

Stock, Z. S., Russo, M. R., and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, Atmospheric Chemistry and Physics, 14, 3899–3912, https://doi.org/10.5194/acp-14-3899-2014, https://www.atmos-chem-phys.net/14/3899/2014/, 2014.

Thunis, P., Pisoni, E., Degraeuwe, B., Kranenburg, R., Schaap, M., and Clappier, A.: Dynamic evaluation of air quality models over European regions, Atmospheric Environment, 111, 185 – 194, https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.04.016, http://www.sciencedirect.com/science/article/pii/S1352231015300194, 2015.

Timmermans, R., Kranenburg, R., Manders, A., Hendriks, C., Segers, A., Dammers, E., Zhang, Q., Wang, L., Liu, Z., Zeng, L., van der Gon, H. D., and Schaap, M.: Source apportionment of PM2.5 across China using LOTOS-EUROS, Atmospheric Environment, 164, 370 – 386, https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.06.003, http://www.sciencedirect.com/science/article/pii/S1352231017303758, 2017.

Toll, I. and Baldasano, J.: Modeling of photochemical air pollution in the Barcelona area with highly disaggregated anthropogenic and biogenic emissions, Atmospheric Environment, 34, 3069 – 3084, https://doi.org/https://doi.org/10.1016/S1352-2310(99)00498-7, http://www.sciencedirect.com/science/article/pii/S1352231099004987, 2000.

Tressol, M., Ordonez, C., Zbinden, R., Brioude, J., Thouret, V., Mari, C., Nedelec, P., Cammas, J.-P., Smit, H., Patz, H.-W., and Volz-Thomas, A.: Air pollution during the 2003 European heat wave as seen by MOZAIC airliners, Atmospheric Chemistry and Physics, 8, 2133–2150, https://doi.org/10.5194/acp-8-2133-2008, https://www.atmos-chem-phys.net/8/2133/2008/, 2008.

Valari, M. and Menut, L.: Does an Increase in Air Quality Models' Resolution Bring Surface Ozone Concentrations Closer to Reality?, Journal of Atmospheric and Oceanic Technology, 25, 1955–1968, https://doi.org/10.1175/2008JTECHA1123.1, https://doi.org/10.1175/2008JTECHA1123.1, 2008.

Valverde, V., Pay, M. T., and Baldasano, J. M.: Ozone attributed to Madrid and Barcelona on-road transport emissions: Characterization of plume dynamics over the Iberian Peninsula, Science of The Total Environment, 543, 670 – 682, https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.11.070, http://www.sciencedirect.com/science/article/pii/S0048969715310500, 2016.

van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P., Christensen, J., Cuvelier, C., Graff, A., Jonson, J., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmospheric Environment, 41, 2083 – 2097, https://doi.org/https://doi.org/10.1016/j.atmosenv.2006.10.073, http://www.sciencedirect.com/science/article/pii/S1352231006011046, 2007.

van Zelm, R., Huijbregts, M. A., den Hollander, H. A., van Jaarsveld, H. A., Sauter, F. J., Struijs, J., van Wijnen, H. J., and van de Meent, D.: European characterization factors for human health damage of PM10 and ozone in life cycle impact assessment, Atmospheric Environment, 42, 441 –

453, https://doi.org/https://doi.org/10.1016/j.atmosenv.2007.09.072, http://www.sciencedirect.com/science/article/pii/S1352231007008667, 2008.

Vautard, R., Builtjes, P., Thunis, P., Cuvelier, C., Bedogni, M., Bessagnet, B., Honoré, C., Moussiopoulos, N., Pirovano, G., Schaap, M., Stern, R., Tarrasón, L., and Wind, P.: Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the CityDelta project, Atmospheric Environment, 41, 173 – 188, https://doi.org/https://doi.org/10.1016/j.atmosenv.2006.07.039, http://www.sciencedirect.com/science/article/pii/S1352231006007862, 2007.

Vlemmix, T., Eskes, H. J., Piters, A. J. M., Schaap, M., Sauter, F. J., Kelder, H., and Levelt, P. F.: MAX-DOAS tropospheric nitrogen dioxide column measurements compared with the LOTOS-EUROS air quality model, Atmospheric Chemistry and Physics, 15, 1313–1330, https://doi.org/10.5194/acp-15-1313-2015, https://www.atmos-chem-phys.net/15/1313/2015/, 2015.

WGE: Ozone pollution: Impacts on ecosystem services and biodiversity. ICP Vegetation Programme Coordination Centre, Tech. rep., Bangor, UK, http://icpvegetation.ceh.ac.uk/publications/documents/ICPVegetationozoneecosystemservicesandbiodiversityreport2013_FULL.pdf, 2013.

WHO: Air Quality Guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global Update 2005, WHO Regional Office for Europe, Copenhagen, Denmark, 2006.

WHO: Review of evidence on health aspects of air pollution, REVIHAAP project: technical report, Tech. rep., Copenhagen, Denmark, http://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-report-final-version.pdf?ua=1, 2013a.

WHO: Health Risks of Air Pollution in Europe-HRAPIE Project: Recommendations for Concentration-Response Functions for Cost–Benefit Analysis of Particulate Matter, Ozone and Nitrogen Dioxide, Tech. rep., Copenhagen, Denmark, http://www.euro.who.int/__data/assets/pdf_file/0017/234026/e96933.pdf?ua=1, 2013b.

Name	Network	Type	Lat. (°)	Long. (°)	Alt. (m a.s.l.)
VILLA DEL PRADO	MAC	RURAL	40.25	–4.27	469
SAN MART'IN DE VALDEIGLESIAS	MAC	RURAL	40.37	–4.40	707
EL ATAZAR	MAC	RURAL	40.91	–3.47	995
SAN PABLO DE LOS MONTES	EMEP	RURAL	39.55	–4.35	917
CAMPISÁBALOS	EMEP	RURAL	41.27	–3.14	1360
GUADALIX DE LA SIERRA	MAC	RURAL	40.78	–3.70	852
ORUSCO DE TAJUÑA	MAC	RURAL	40.29	–3.22	795
ALCORCÓN	MAC	URBAN	40.34	–3.83	693
TOLEDO	CLM	URBAN	39.87	–4.02	500
ENSANCHE DE VALLECAS	MCO	URBAN	40.37	–3.61	630
VILLAVERDE	MCO	URBAN	40.35	–3.71	593
ARTURO SORIA	MCO	URBAN	40.44	–3.64	698
FAROLILLO	MCO	URBAN	40.39	–3.73	625
PLAZA DEL CARMEN	MCO	URBAN	40.42	–3.70	657
GUADALAJARA	CLM	URBAN	40.63	–3.17	620
MÓSTOLES	MAC	URBAN	40.32	–3.88	650
ARANJUEZ	MAC	URBAN	40.04	–3.59	512
RETIRO	MCO	URBAN	40.41	–3.69	672
TRES OLIVOS	MCO	URBAN	40.50	–3.69	715
AZUQUECA DE HENARES	CLM	URBAN	40.57	–3.26	600
BARAJAS-PUEBLO	MCO	URBAN	40.47	–3.58	631
RIVAS-VACIAMADRID	MAC	SUBURBAN	40.36	–3.54	610
JUAN CARLOS I	MCO	SUBURBAN	40.47	–3.61	669
EL PARDO	MCO	SUBURBAN	40.52	–3.77	700
ALGETE	MAC	SUBURBAN	40.59	–3.50	721
MAJADAHONDA	MAC	SUBURBAN	40.45	–3.87	722
ILLESCAS	CLM	SUBURBAN	40.12	–3.83	548
TORREJÓN DE ARDOZ	MAC	SUBURBAN	40.46	–3.48	581
VALDEMORO	MAC	SUBURBAN	40.19	–3.68	610
CASA DE CAMPO	MCO	SUBURBAN	40.42	–3.75	645
ÁVILA	CL	SUBURBAN	40.66	–4.70	1150
ALCOBENDAS	MAC	URBAN	40.54	–3.64	671
COLMENAR VIEJO	MAC	URBAN	40.67	–3.77	905
ALCALÁ DE HENARES	MAC	URBAN	40.48	–3.38	589
VILLAREJO DE SALVANÉS	MAC	URBAN	40.17	–3.28	761

Table 2. Details of the air-quality monitoring stations selected for this study. With the exception of the last five, which are classified as traffic sites, all the stations included in this table are located in background locations. Network codes: MCO, Madrid Council; CL, Castilla y León region; CLM, Castilla La Mancha region; MAC, Madrid region and EMEP.

	ECMWF_5	ECMWF_70	ECMWF_HR_70	WRF_5	WRF_70
Fractional Bias (FB)					
06-July-2016	0.10	0.10	0.13	–0.08	–0.04
13-July-2016	0.05	0.05	0.05	0.01	0.08
20-July-2016	0.17	0.25	0.16	0.13	0.22
27-July-2016	0.08	0.05	–0.03	–0.24	–0.17
Normalised mean square error (NMSE)					
06-July-2016	0.024	0.015	0.030	0.009	0.007
13-July-2016	0.015	0.024	0.015	0.002	0.016
20-July-2016	0.057	0.116	0.040	0.026	0.086
27-July-2016	0.021	0.011	0.005	0.071	0.039

Table 3. FB and NMSE of the comparisons between LOTOS-EUROS simulations of vertical profiles of O_3 and data from four O_3 soundings performed in Madrid in July 2016.