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.

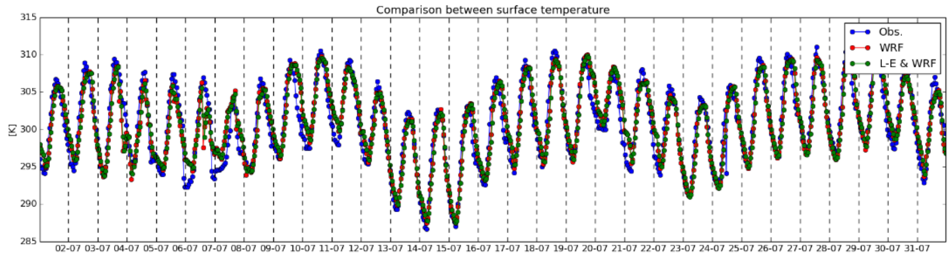
A table showing the Pearson's correlation factor (r) correlation factor for NO2 and NOx was also composed and it is shown below. However, it is fair to indicate that, since O3 was our target pollutant, traffic stations were excluded from the set of stations used for validation so the picture may be a 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:

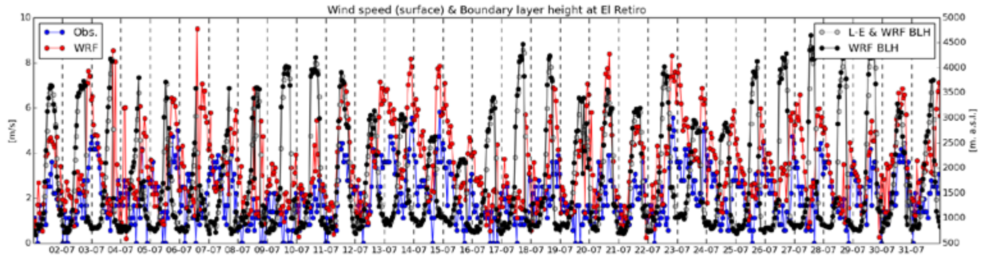
		NO_2				NO_x			
STATION	TYPE	ECMWF_5	ECMWF_70	WRF_5	WRF_70	ECMWF_5	ECMWF_70	WRF_5	WRF_70
VILLA DEL PRADO	RURAL	0.481	0.576	0.42	0.437	0.454	0.535	0.348	0.371
SAN MARTIN DE VALDEIGLESIAS		0.405	0.454	0.337	0.358	0.425	0.45	0.344	0.308
EL ATAZAR		0.422	0.524	0.381	0.472	0.446	0.539	0.408	0.476
SAN PABLO 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

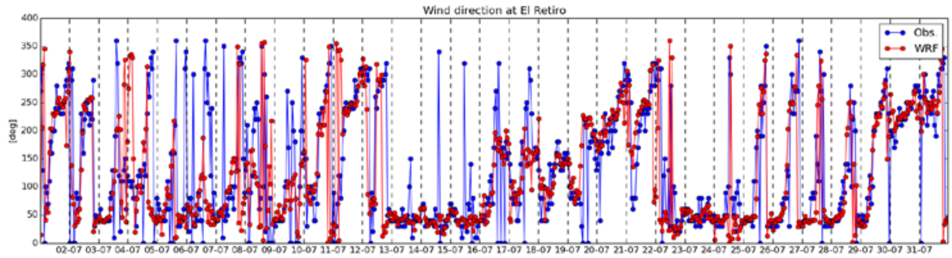
Regarding, the evaluation of the meteorological datasets we made a qualitative validation of the degree of concordance between the two meteorological datasets (WRF and ECMWF) and observations (at El Retiro in the centre of Madrid) of temperature, wind direction and wind speed. We concluded that temperature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