Reviewer #1

General comments

Throughout the manuscript, there is thorough discussion of both transport and chemistry, but none of deposition. Deposition is notably missing from both the Introduction and Discussion sections. Could deposition play a role in explaining the discrepancies between the MRE and observations?

We agree that deposition is an important process. Additional tests performed currently in the C-IFS system indicate that surface O3 is sensitive to the dry deposition mechanism. In Figure 1 is shown how three different dry deposition schemes affect the annual cycle of surface ozone over Europe in three, 1-year-long sensitivity experiments. These results indicate that the deposition can perhaps contribute to improvement of surface ozone seasonality, but cannot completely fix the spring ozone maximum problem over north Europe (left plot). Clearly, by improving the dry deposition scheme, the bias is decreasing, mostly over southern Europe in summer (right plot). The following text has been added in the revised manuscript

"*Ongoing work on the impact of dry deposition on surface ozone indicates that the new on-line dry depositions schemes currently tested in the C-IFS system improve the surface ozone positive bias, appearing mostly over southern Europe in summer, but cannot completely tackle the spring ozone maximum problem over north Europe (J. Flemming, personal communication, 2015).*"

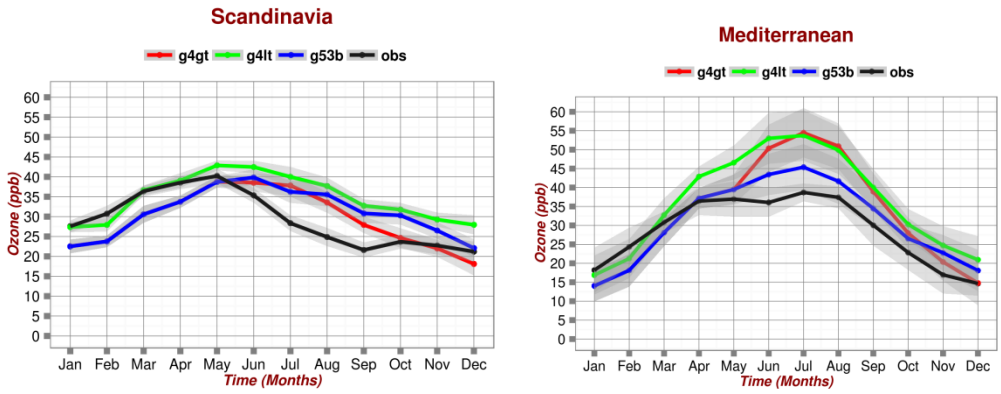
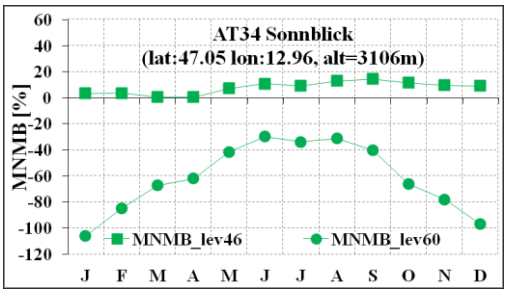


Figure 1. Annual cycle of surface O3 for Scandinavian (left) and the Mediterranean (right) stations. Different colors indicate different deposition schemes tested in C-IFS.

The model is sampled at vertical levels other than the surface to match altitude with observing sites, but this will also impact deposition. This issue should be discussed and the offset between surface and above-surface grid boxes should probably be evaluated.

In the following plot it is shown the modified normalized mean bias (mnmb) of the Austrian station Sonnblick (altitude = 3,106 m), evaluated i) with surface model data (Lev60, dots) ii) with data from level 46 (squares). Following the objective methodology described in section

2.2, it is shown that the bias is reduced, when adjusting o3 concentrations using atmospheric pressure as the correction criterion. The impact of deposition affects directly only the lowest model level (L60) and indirectly the higher leves of ABL. We have seen that concentrations of the lowest model level are not representative for stations with higher altitude.



The failure of the MRE to capture the spring peak in ozone that is noted by the authors requires further exploration, but it is fine with me for the authors to present it as a question for future work. One puzzling aspect is that many models do capture a springtime maximum in ozone, and particularly with the assimilation of column ozone observations, I would have expected long-range transport contributions to spring ozone [Parrish et al., 2013] to be captured.

As mentioned in Section 4.1 the shape of the observed ozone annual cycle (based on the ozonesondes) in lower free troposphere at 700 hPa and for the middle troposphere at 500 hPa is reproduced rather well by the MRE. The reasonable reproduction of the shape of the observed ozone seasonal cycle by MRE in the middle and lower free troposphere is consistent with transport processes from the lower stratosphere and the upper troposphere, as well as long-range transport being resolved adequately by the MRE. Hence, it is rather at near surface and within the boundary layer that MRE fails to capture well the spring peak. This mismatch could be related to a) overestimated photochemical ozone production within the atmospheric boundary layer, b) deposition, c) insufficient entrainment and mixing from the lower free troposphere into the atmospheric boundary layer.

Here we discuss that overestimated local photochemical ozone production at near surface may actually mask the contribution of transport on the seasonal ozone cycle. We also discuss the role of deposition in MRE, indicating that the use of an online deposition scheme reduces the positive bias in summer at southern Europe. This can contribute to the improvement of near surface ozone seasonality, but cannot completely fix the spring ozone maximum problem over north –Europe.

Some further justification is required in defining the subregions that Europe is broken up into. The authors argue that "Overall, the annual cycles of the observed data reflect the specific subregional characteristics: : :". However, there are three counterarguments to this:

1) For some regions, the seasonal observed cycle varies substantially within the subregion. For example, in the Scandinavian subregion, the sites in the Baltic states and Denmark peak in the summer, while those on the Fennoscandian peninsula peak in the spring. This could be

complicating the analysis of the offset in seasonality between modeled and observed cycles in the Scandinavian subregion.

2) All of the modeled seasonal cycles shown in Fig. 4 look much more sinusoidal than the observed seasonal cycles, so while the model is doing a reasonable job of the capturing the magnitude of the annual mean and seasonal amplitude, the shape and phase of the seasonality are not captured.

In response to this comment, we performed a separate analysis for the Baltic (5 stations over Latvia, Estonia, Denmark) and the Fennoscandian (15 stations over Sweden, Norway, Finland) regions. Figure 3 shows the annual cycles for surface O3 over the Baltic (left) and the Fennoscandian (right) region, exhibiting both an observational spring maximum. This analysis justifies the grouping of all stations in a common subregion (denoted as SC), with similar seasonal characteristics.

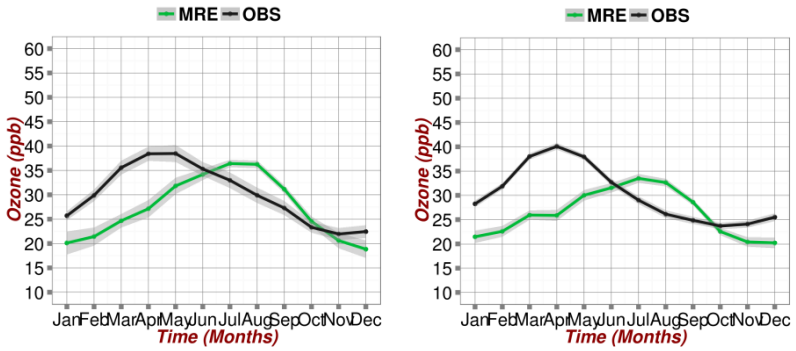


Figure 3. **Annual cycle of surface O3 for the Baltic (left) and the Fennoscandian (right) stations.**

3) The Mediterranean sites are broken into continental and coastal sites, but the other regions are not. There is likely a distinction in the observed seasonal and diurnal cycles between coastal and continental sites for the British Isles and Central Europe

Indeed the geographical convenience was our initiative to split the regions and the fact the similar regions have been used in previous climate-oriented studies (Christensen, J. H. and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Clim. Change, 81, 7–30, doi:10.1007/s10584-006-9210-7, 2007).

It has not be used a statistical cluster analysis to objectively discriminate regions with distinct ozone characteristics. This is part of our on-going work. However, in the case of ozone, there are a number of difficulties for an objective way to discriminate regions with distinct characteristic from station data. This is because even within a small region with similar large scale ozone features, the stations may differ significantly in terms of the ozone behavior depending on the distance from sea, the elevation and the distance for pollution sources. This becomes even worse for regions with small number of stations.

A nice example is highlighter in the case of Mediterranean, with a small number of stations and with different ozone characteristics of the maritime rural EMEP stations from the continental rural classified AIRBASE stations. In our analysis we did a geographical compromise, when calculating ozone averages.

Specific comments:

- 1078L9: "Annual overall error" is a vague term in the abstract.
  Done. Corrected to 'fractional gross error'.

- 1080L7-11: Discussion of sources, chemistry, transport, but no discussion of deposition.
  Done. Dry deposition was added as a removal process.

- 1081L19-21: In addition to stratospheric and column ozone, the MRE also appears to assimilate satellite observations of other relevant gases (CO, NO2) that will impact ozone chemistry [Inness et al., 2013].
  That is correct. This note was added in the sentence.
  "*The impact of assimilation on near surface ozone is only the "residual" of correcting the stratospheric and total ozone column, plus the assimilation of other relevant gases that impact ozone chemistry (CO, NO2) (Inness et al., 2013).*"

- 1081L22-28: While the explanation of the configuration for the control run is clear, I am unclear on what is meant by the "control run is not a "clean" control analysis experiment"
  Done. A proper explanation was provided.

- 1082L10-12: Is there a literature reference for the choice to use background stations for comparison to coarse-resolution model output?
  A reference has been added in the revised manuscript. Schaap et al., 2014, Atm Env., text from Section 3 "Model performance evaluation":
  "*…As it is fitted to catch background air pollution patterns with stations at a considerable distance from source areas in rural or remote regions, this network (EMEP) is appropriate to evaluate regional scale models performance with coarse resolutions (50 -150 km^2)…*"

- 1087L8-11: Why does assimilation make the seasonal cycle worse in some areas?
  We attribute the deterioration of R to an inherent problem of the data assimilation procedure, related to the MLS bias correction, described in detail in the paper of Inness et al., 2013. The bias correction of MLS data, has caused drifts in the tropospheric ozone concentrations between August 2004 and December 2007, an issue which have been tracked down and alleviated after year 2008 of the MRE. Comments are inserted in the revised manuscript.

  Figure 4 shows that temporal correlation of the MRE increases after bias correction (2008-2012).

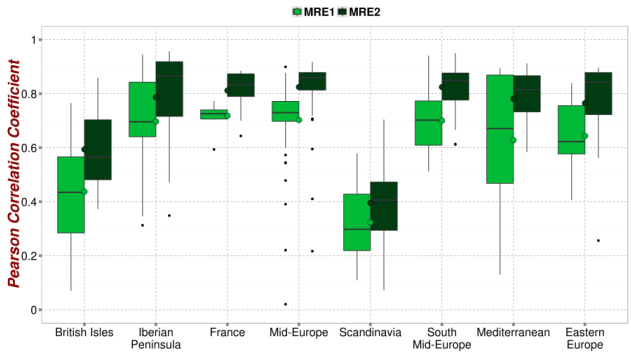


Figure 4: Annual Whisker plots for surface temporal correlation for MACC reanalysis averaged over 2003-2007 with bias correction (MRE1, light green) and over 2008-2012 without bias correction (MRE2, dark green).

- 1092L12: "Other PAN homologues (PANs)" I believe should be abbreviated APNs (standing for acyl peroxy nitrates).
  Done

- 1106: Figure 2 caption. Describe the box and whisker structure in the figure caption in addition to its description in the text on page 1086.
  Done

- L1110: Figure 6. If possible, color coding the shaded envelopes to be consistent with the line colors would help to improve the readability.
  Done. A new figure 6 is provided.

- L1112: Figure 7. 24 subplots is too much for one figure! The profiles become very hard to read when that small.
  Done. A new Figure 7 is provided.

- 1113: Figure 8 caption. Change "near surface ozone at 700 hPa" to "lower tropospheric ozone at 700 hPa" to distinguish from the "near surface" observations discussed throughout the rest of the manuscript.
  Done

Technical corrections:

- 1079L12: Change "year-long experience" to "many years of"
  Done
- 1080L18: "(even at near surface)" change to "even near the surface"
  Done
- 1086L4-5: The line indicating the median in Fig. 2 is horizontal, not vertical.
  Done
- 1090L9: "and the fail in MRE: : :" change "fail" to "failure"
  Done
- 1090L10: Add "It" before "Is known that: : :"
  Done

Reviewer #2

Main comments

1. It is not clear to me why the authors do not make more use of the CTRL simulations in their comparisons. While the CTRL simulation does not extend to 2012 like the MRE, comparison of the MRE statistics for 2003-2012 (table 1) and 2003-2010 (table 2) suggests little difference when the extra 2 years are included. If the authors restricted all their analysis to the common 2003-2010 period, they could add the CTRL results to the seasonal cycle plots. I would be interested in seeing this, particularly as the CTRL simulation's seasonal cycle agrees better with observations (at least marginally). Might an extended comparison of MRE against CTRL hint at further drivers of observation/MRE discrepancies?
   This paper was thought to be an extended evaluation of the reanalysis product with respect to near surface ozone. For anyone wishing to have this paper as a reference for the evaluation of the ECWMF reanalysis product, we believe it would be better to include the whole period of the reanalysis (2003-1012).
   In the current manuscript we provide the basics of the comparison between the ctrl and the reanalysis (Fig 3 and Table 2). The extended report on the impact of assimilation on surface ozone from 2003 to 2010 is available as a VAL technical report, Deliverable D84.2 "Validation report on the Comparison of surface ozone in the global (2003- 2010) and regional reanalysis (2011) over Europe".
   The basic findings of our analysis are robust and do not depend on the length of the period examined.

2. The language is often imprecise when discussing the comparisons. For example, what are "acceptable temporal correlations" (P1085, L20)? What does "reasonably well" (P1091, L2) mean? In addition, for the discussion on CTRL vs MRE, if r = 0.74 to 0.49 is "slightly reduced" (P1087, L10), what should we make of the bias improvements discussed on P1087, L1-5?
   In response to this comment we modified the manuscript accordingly:
   i) "acceptable temporal correlations" the expression was removed and the discussion on the correlations was rewritten.
   ii) The sentence was rewritten: "*Comparison with ozonesonde measurements at different locations (Fig. 7) indicate that MRE ozone profiles reproduce the basic structure of the profile, overestimating in most cases ozone below the 850hPa*."
   iii) The word "slightly" is deleted and the reduction of R in the assimilated experiment is discussed more thoroughly.

3. The font size on the figures is too small and there are often too many panels to give a useful picture of what's going on (esp. fig 5 and fig 7). In addition, the authors could consider plotting the biases and correlations on maps like Figure 1 (e.g. coloring the dots by the r and FGE values).
   New plots are provided in the revised manuscript with larger fonts.

Specific comments (including technical corrections)

- P1078, L24: "and MACC-II: : :"
  Done

- P1079, L8: Define ECMWF (you do it for all other acronyms)

Done

- P1079, L12: year-long -> long
  Done

- P1079, L20: "(AQME)"
  Done

- P1079, L25-: Near surface ozone is not very important as a GHG
  We remove the sentence referring to tropospheric ozone and its role as GHG.

- P1080, L2-: Mention chemical loss and deposition
  The sentence "*It can be destroyed photochemically or by dry deposition at the surface*" was added.

- P1080, L7: "soil, vegetation"
  Done

- P1080, L25: "Monks, 2000"
  Done

- Section 2.1: Would be useful to know which species/observations are assimilated which are relevant for ozone
  The ozone data assimilated in MRE are listed in Table 1 of Inness et al. (2013). A comment has been inserted in the revised manuscript.

- P1081, L18: Define "variational bias correction"
  In the variational scheme biases are estimated during the analysis by including bias parameters in the control vector. The bias corrections are continuously adjusted to optimize the consistency with all information used in the analysis.

- P1081, L19 (and several other places): Mind that -> Note that (former sounds like an admonishment)
  The sentence is re-written: "*The assimilation correction on ozone is due to the stratospheric and total ozone column*."

- P1082, L9: What are "types 1-3"? Need more detail
  Joly and Pech (2012) used Linear Discriminant Analysis on the pollution measurements of the AIRBASE network to discriminate the rural stations from the most polluted, urban and traffic stations. This statistical process is specific for its measured pollutant and using nine percentiles from 10% to 90% as fixed arbitrary thresholds, ten classes have been defined. The first three classes for the case of O3 discriminate reasonably the stations with rural characteristics that are the most representative of the large scale. 'Types 1-3' are corrected to 'classes 1-3' in the manuscript.

- P1082, L14: "corresponding observational data" (data can be from a model too)
  Done

- P1082, L26: Is there any rationale for these regions beyond geographical convenience? E.g. can you demonstrate that similar processes control ozone in each region. This is important for making inferences about the regional statistics.
  We agree with the reviewer that this is an important issue, when averaging stations within a region. Our regional selection is arbitrary. Indeed the geographical convenience was our initiative to split the regions and the fact the similar regions have been used in previous climate-oriented studies (*Christensen, J. H. and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Clim. Change, 81, 7–30, doi:10.1007/s10584-006-9210-7, 2007*).
  It has not be used a statistical cluster analysis to objectively discriminate regions with distinct ozone characteristics. This is part of our on-going work. However, in the case of ozone, there are a number of difficulties for an objective way to discriminate regions with distinct characteristic from station data. This is because even within a small region with similar large scale ozone features, the stations may differ significantly in terms of the ozone behavior depending on the distance from sea, the elevation and the distance for pollution sources. This becomes even worse for regions with small number of stations. A nice example is the case of Mediterranean with a small number of stations and with different ozone characteristics of the maritime rural EMEP stations from the continental rural classified AIRBASE stations. In our analysis we did a geographical compromise, when calculating ozone averages.

- P1083, L1 (and Fig 1): Could you indicate the region codes on the figure? Additionally, using the full name for the region in the text makes for easier reading (there are some instances where just "BI" etc are used)
  Done: Region codes are explicitly indicated in Figure 1 caption.

- P1083, L16-21: This paragraph could do with re-phrasing and making into <1 sentence.
  The sentence was split as shown below:
  "*We have also to take into consideration that the NOx observations are affected strongly by local emissions. Furthermore there are known issues with interference by oxidized nitrogen compounds (e.g. HNO3, PAN and other organic nitrates) for ground-based NO2 measurements by most commercially available NO2 instruments using molybdenum converters, hence leading to an overestimation of NOx concentrations (Steinbacher et al., 2007).*"

- P1085, L3: Be clear that you're using R to refer to the seasonal cycle, rather than correlating the whole time series (or time series of DJFs etc). The latter might be interesting though to investigate interannual variability.
  Done. "Seasonality" was corrected to "interannual variability".

- P1085, L5: SD -> standard deviation (at least the first time)
  Done

- P1086, L3-6: This information can go in the caption
  Done

- P1086, L11-: This is repeating the point in L8 (i.e. not "On the other hand")
  Done. The first sentence was deleted.

- P1087, L27: however, -> but
  Done

- P1088, L3 (and throughout): depicted -> found
  Done

- P1088, L14: revealing -> causing (?)
  "Revealing" changed to "causing"

- P1089, L18: "captures the shape: : :cycles quite well, but with a: : :"
  Done

- P1090, L14 (and throughout): You've changed from "ozone" to "O3". I much prefer the former for easier reading.
  Done

- P1091, L16: suggests that -> is consistent with (and then "being resolved: : :")
  Done

- P1092, L5-: Whether an environment is "NOx-limited" will also depend on the mix of VOCs (their reactivity, propensity to form NOy etc), and presumably the VOC mix differs across Europe.
  We agree with the reviewer that the split between NOx-sensitive and VOC-sensitive conditions is correlated with the ratio of reactivity-weighted VOC mixture to NOx and this ratio differs across Europe. The sentence was modified accordingly.

  "*In global scale, nitrogen oxides (NOx) are the limiting precursors for O3 production throughout most of the troposphere, and also directly influence the abundance of the hydroxyl radical concentration in the troposphere (e.g. Crutzen, 1988).*
  At regional scale for rural environments with NOx values less than a few parts per billion by volume, O3 formation is NOx limited (Liu et al., 1987) and therefore almost independent of hydrocarbon concentrations, depending of course on the ratio of reactivity-weighted VOC mixture to NOx which may differ from region to region across Europe (Beekmann and Vautard, 2010)."

  Beekmann, M. and Vautard, R.: A modelling study of photochemical regimes over Europe: robustness and variability, Atmos. Chem. Phys., 10, 10067-10084, doi:10.5194/acp-10-10067-2010, 2010.

  P1093, L1 -: Do you know that the NOx diurnal cycle is all chemical, with no transport? In general, these arguments might be more convincing if you were able to demonstrate them with (e.g.) a box model
  We agree with the reviewer that this is an important issue. We used a box model with the CBIV chemical mechanism to calculate ozone production efficiencies for typical summer conditions using initial conditions for NOx and other gaseous species from the MACC model at BI, IP and ME. We calculated that 3 to 4 molecules of O3 are produced for every molecule of NOx oxidised at BI and ME, and up to 5 pbbv at IP. The above values agree well with ozone production efficiency estimates from previous studies for summer at rural semi-polluted sites with NOx more than a few ppbv in Europe and US (Chin et al., 1994; Derwent and Davis, 1994; Rickard et al.,

2002). Furthermore, the box model estimated that ozone production efficiency values are comparable to the near surface ratio ΔO3/ ΔNOx (ΔO3 increased over the day; ΔNOx decreased over the day) shown in Figure 10, which is roughly 3 for BI, 3.5 for ME and 10 for IP. Additionally, we have also estimated MRE ΔO3/ΔNOx ratio values at 925 hPa (above near surface but within the atmospheric boundary layer) being roughly 3.5 for BI, 3 for ME and 4 for IP, which is in good agreement with the box model calculations.

Nevertheless, diurnal meteorological patterns of wind speed and boundary layer height, that lead to higher dilution of primary pollutants at daytime than at nighttime, may also contribute for the diurnal pattern of NOx in Figure 10 (see Figure 1, following the diurnal cycle of wind speed and boundary layer height). This is supported from the fact that CO which is a species with much longer chemical lifetime than NOx has a similar diurnal pattern with NOx (not shown in the manuscript but shown in the Figure below).

The text in Section 4.2 was modified accordingly and the relevant references were added (see below). The figure shown below was added as Figure S1 in the revised manuscript.

References

Chin, M., Jacob, D. J., Munger, J. W., Parrish, D. D., and Doddridge, B. G., 1994: Relationships of ozone and carbon monoxide over North America, J. Geophys. Res. 99, 14,565–14,573.

Derwent, R. G. and Davis, T. J., 1994: Modelling the impact of NOx or hydrocarbon control on photochemical ozone in Europe, Atmos. Environ. 28, 2039–2052.

Rickard, A. R., Salisbury, G., Monks, P. S., Lewis, A. C., Baugitte, S., Bandy, B. J., Clemitshaw, K. C., and Penkett, S. A.: Comparison of measured ozone production efficiencies in the marine boundary layer at two European coastal sites under different pollution regimes, J. Atmos. Chem., 43, 107–134, 2002.

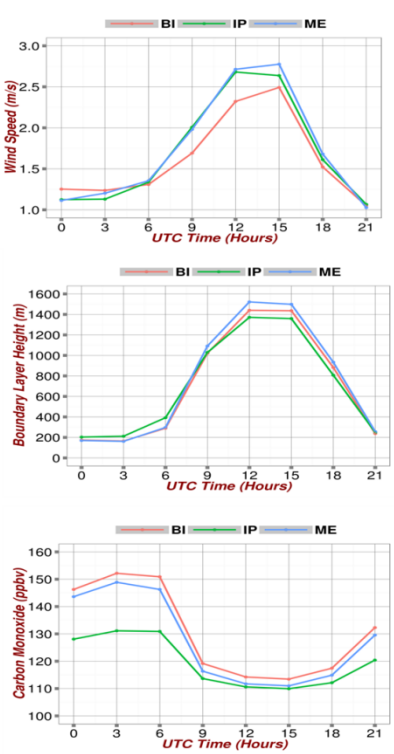


Figure 1. Mean summer diurnal cycle of wind speed (m/s), boundary layer height (m) and near surface carbon monoxide (ppbv) for the sub-regions BI, IP and ME based on MRE.

- P1093, L24: ": : :adequately capture the seasonality,: : :"
  Done

- P1095, L27: Final paragraph should be aligned to the left margin.
  Done

- Table 2: The FGE, MNMB and R should be centered above the MRE and CTRL columns.
  This comment will be delivered to the Editorial Office, since all Tables and Figures are edited.

Reviewer #3

General comments

1) All along the manuscript, the authors should be more precise, both in their qualification of the results and the terms they use in general. For instance, in the abstract, what is 'the annual overall error' accounting for? What is the value of the 'average correlation' (p 1086, L1) etc : : : There are many points like these, I willl go back to these in the specific comments

We improved the language in the revised manuscript, when communicating findings on scores and model skill.

2) The use of the CTRL simulation was very promising but is finally disappointing because too short. In particular an explanation of the drop of the correlation from CTRL to MRE in Mediterranean marine stations and in Scandinavia would be expected in the discussion part. Finally, either the CTRL simulation should not be used at all, or compared to MRE all along the manuscript, with, if necessary, an adaptation of the time period to be analysed.

This paper was thought to be an extended evaluation of the reanalysis product with respect to near surface ozone. For anyone wishing to have this paper as a reference for the evaluation of the reanalysis product, we believe it would be better to include the whole period of the reanalysis (2003-1012).
In the current manuscript we provide the basics of the comparison between the ctrl and the reanalysis (Fig 3 and Table 2). An extended report focusing on the impact of assimilation on surface ozone for the 2003-2010 is available as a VAL technical report, Deliverable D84.2 "Validation report on the Comparison of surface ozone in the global (2003- 2010) and regional reanalysis (2011) over Europe". The basic findings of our analysis are robust and do not depend on the length of the period examined.

The lower temporal coefficients (R) in the MRE are discussed in the revised manuscript. We attribute the deterioration of R the data assimilation procedure, related to the MLS bias correction, described in detail in the paper of Inness et al., 2013. The bias correction of MLS data, has caused drifts in the tropospheric ozone concentrations between August 2004 and December 2007, an issue which have been tracked down and alleviated after year 2008 of the MRE. After 2008 R appears to be improving.

Comments are inserted in the revised manuscript. Figure 1 shows that temporal correlation of the MRE increases after bias correction (2008-2012).

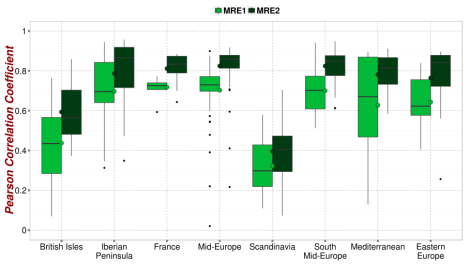


Figure 1: Annual Whisker plots for surface temporal correlation for MACC reanalysis averaged over 2003-2007 with bias correction (MRE1, light green) and over 2008-2012 without bias correction (MRE2, dark green).

3) Even if the paper describing precisely the MRE is referred to I would like the assimilation process to be described more precisely. In particular, the time-steps of the assimilation, its

vertical extent, and the chemical species that are assimilated. Only one sentence (p1087, L17-18) mentions that point: this is not enough. Moreover, even if it is implicit, the authors should explicitly mention that observations they use for this evaluation are independent from the assimilated ones.

More details are provided in the assimilation procedure and Table 1 is added in the revised manuscript.

The observations used for this evaluation are independent from the assimilated ones. A comment is inserted in section 2.2

4) Figures and captions are generally too small (however they are in general very informative).
Done. Figures are improved in the revised manuscript.

5) I would appreciate a conclusion that would give more perspectives to this work.
We have re-written conclusions with more concrete directions to future work.

Specific comments

- p1078 L9-10 : define the "annual overall error" and "on average" (spatial, temporal, both?)
  The annual overall error is the fractional gross error calculated on an annual basis and it is corrected in the revised manuscript. It is "on average".. "over Europe" i.e. a spatial average.

- p1082, L20-25 : can you give an estimation of the impact of taking real-altitude of the station instead of surface ?
  In Figure 2 it is shown the modified normalized mean bias (mnmb) of the Austrian station Sonnblick (altitude = 3,106 m), evaluated i) with surface model data (Lev60, dots) ii) with data from level 46 (squares). Following the objective methodology described in section 2.2, it is shown that the bias is reduced, when adjusting o3 concentrations using atmospheric pressure as the correction criterion.

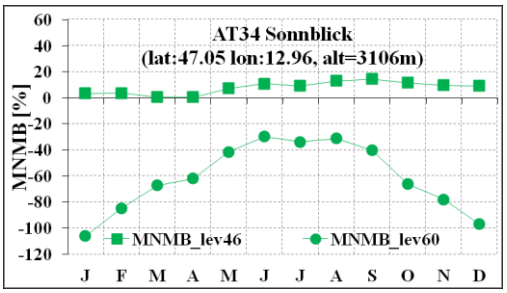


Figure 2 Normalized Mean Bias for the Sonnblick station evaluated with model data from the surface model level and upper vertical model levels

- p1083, L28 : the precision of the ozonesondes is no more referred to hereafter. In 4.1, you should recall it to the reader and comment the results correspondingly.
  The following sentence was added in the revised manuscript:

"It should be also considered that the range of the % biases in the troposphere are comparable with the respective precision of electrochemical concentration cell ozonesonde measurements."

- P1085, L8-9 : "the confidence interval : : : subregion" : this sentence is unclear to me.
  "The confidence interval for each month was derived using the values of the diurnal range for the stations that reside in the same subregion."
  The grey areas in Figure 6 show the 95% confidence interval of the mean sub-regional diurnal range, derived when averaging diurnal ranges of all stations within the sub-region.

- p1085-86, L20-1 : I do not agree for having British Isles and Scandinavia at the same level. Their correlations are really different. I would put together BI and MDm (0.51 and 0.54) and separate Sc (0.26). This is implicitly what you mention later. (L26-27), so this sentence could finally be removed.
  Done

- P1086,L10-11: I suppose the numbers you give (40:28% etc..) correspond to the mean value of the data. It should be specified, since the median could also be taken into account and give significantly different results.
  The numbers refer to the FGE which is introduced I section 2.3 (Eq 2).

- P1087, L9-10 : for a correlation that drops from 0.74 to 0.49, 'a slightly reduction' is not an appropriate description. Moreover, how do you explain that drop?
  We delete the word "slight". The temporal correlation over Scandinavia is low, because the MRE cannot capture the spring maximum. Moreover, MLS bias correction in the assimilation procedure has caused drifts in tropospheric ozone (a detailed explanation of the technical problem can be found in Inness et al., 2013). This issue was tracked down and alleviated after year 2008 of the MRE.

  We attribute the deterioration of R to this inherent problem of the data assimilation procedure. Comments are inserted in the revised manuscript. Figure 1 shows that temporal correlation increases after bias correction (2008-2012).

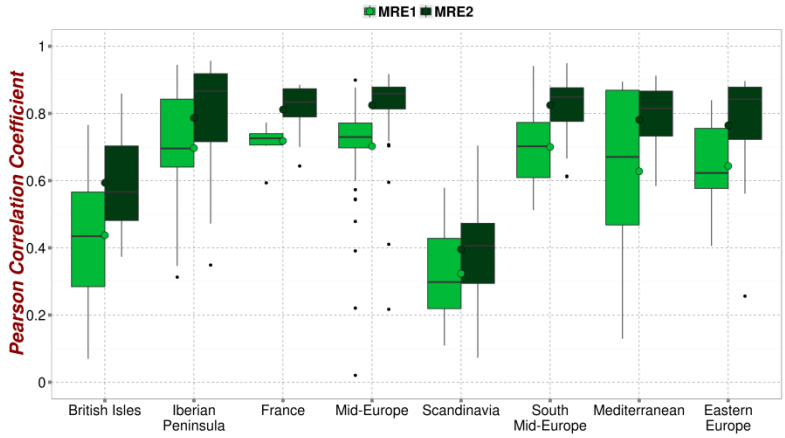


Figure 1: Annual Whisker plots for surface temporal correlation for MACC reanalysis averaged over 2003-2007 with bias correction (MRE1, light green) and over 2008-2012 without MLS bias correction (MRE2, dark green).

- P1087,L23-25 : The terms “cycles have differences in the shape”, although it is true, are too imprecise. These differences should be numerically estimated through correlation, to make sure the analysis is objective.
  The two tables (Figure 3) provide the correlation of the diurnal cycles (left) and the annual cycles (right). The diurnal cycle is well reproduced, while problems in the correlation of the annual cycles are discussed in the manuscript.

R_daily_cycle	
subregion	R
BI	0.93
IP	0.98
FR	0.99
ME	0.97
SC	0.95
SME	0.92
MDc	0.98
MDm	0.98
EA	0.96

R_annual_cycle	
subregion	R
BI	0.67
IP	0.96
FR	0.91
ME	0.89
SC	0.32
SME	0.89
MDc	0.96
MDm	0.99
EA	0.83

Figure 3. Correlation of diurnal cycles (left) annual cycles (right)

- P1088, L8-12: you should mention that this point will be discussed later.
  Done

- P1089, L18-19: “the MRE captures quite well: : :” : once again, this should be more precise.
  The Table with R_diurnal cycle (provided above) justifies the fact the MRE reproduces the diurnal cycle. The text has been modified “*The MRE reproduces the diurnal cycle, but exhibits positive bias in summer (except for the Mediterranean marine region),…*”

- P1090, L12 : the word “ozone” is lacking between “near surface” and “between”.
  Done

- P1090, L13 : “It” is lacking before “is known”. P1091, 4.1 : this subsection would really benefit from a comparison to CTRL simulation.
  Done

- P1093,L1-4 : “The amplitude: : :, which indicates that we have more intense local oxidation”. I find this interpretation too rapid. I agree that photochemical processes will play an important role. But a too active convection, or a bad representation of emissions could for instance lead to the same behaviour.
  We agree with the reviewer that this is an important issue. We used a box model with the CBIV chemical mechanism to calculate ozone production efficiencies for typical summer conditions using initial conditions for NOx and other gaseous species from the MACC model at BI, IP and ME. We calculated that 3 to 4 molecuels of O3

produced for every molecule of NOx oxidised at BI and ME and up to 5 pbbv at IP. The above values agree well with ozone production efficiency estimates from previous studies for summer at rural semi-polluted sites with NOx more than a few ppbv in Europe and US (Chin et al., 1994; Derwent and Davis, 1994; Rickard et al., 2002).
Furthermore, the box model estimated ozone production efficiency values comparable to the MRE ratio ΔO3/ ΔNOx (ΔO3 increased over the day; ΔNOx decreased over the day) shown in Figure 10, which is roughly 3 for BI, 3.5 for ME and 10 for IP. Additionally, we have also estimated MRE ΔO3/ΔNOx ratio values at 925 hPa (above near surface but within the atmospheric boundary layer) being roughly 3.5 for BI, 3 for ME and 4 for IP, being in good agreement with the box model calculations.

Nevertheless, the diurnal meteorological patterns of wind speed and boundary layer height that lead to higher dilution of primary pollutants at daytime than at nighttime, may also contribute for the diurnal pattern of NOx in Figure 10 (see Figure 5 of this review, the diurnal cycle of wind speed and boundary layer height). This is supported from the fact that CO, which is a species with much longer chemical lifetime than NOx, has a similar diurnal pattern with NOx (Figure S1 in the revised manuscript).

The text in Section 4.2 was modified accordingly and the relevant references were added (see below). Figure 5 of this review is added as Figure S1 in the revised manuscript.

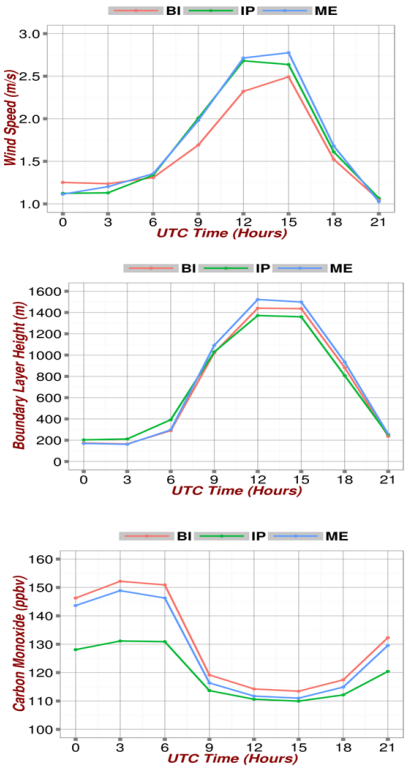


Figure 5. Mean summer diurnal cycle of wind speed (m/s), boundary layer height (m) and near surface carbon monoxide (ppbv) for the sub-regions BI, IP and ME based on MRE.

References

Chin, M., Jacob, D. J., Munger, J. W., Parrish, D. D., and Doddridge, B. G., 1994: Relationships of ozone and carbon monoxide over North America, J. Geophys. Res. 99, 14,565–14,573.

Derwent, R. G. and Davis, T. J., 1994: Modelling the impact of NOx or hydrocarbon control on photochemical ozone in Europe, Atmos. Environ. 28, 2039–2052.

Rickard, A. R., Salisbury, G., Monks, P. S., Lewis, A. C., Baugitte, S., Bandy, B. J., Clemitshaw, K. C., and Penkett, S. A.: Comparison of measured ozone production efficiencies in the marine boundary layer at two European coastal sites under different pollution regimes, J. Atmos. Chem., 43, 107–134, 2002.

- P1094,L12 : a word is missing at the end of the line.
  The sentence was corrected.

Reviewer #4

Assimilation

The most significant and preoccupying result of the paper regard lower performances of MRE compared to CTRL. From Table 1 it is clear that FGE is better in CTLR for 5 out of 9 regions, and monthly correlation is better for 9 out of 9 regions, while the reanalysis outweight the control only for MNMB (in 7 out of 9 regions).The degradation of model performances when implementing the assimilation is a strong concern. A few hypothesis are indeed mentioned in the paper with regard to the larger weight of assimilation in the stratosphere / upper troposphere (P1087L20), but it does not explain why would it be detrimental at the surface.

We believe that an important outcome of this evaluation study is the confirmation that the modeling system provides an adequate representation of near surface O3. Model weaknesses are identified and suggestions for future improvements are provided. It has been demonstrated that the assimilation (correction of stratospheric and total O3 column) in many cases corrects the surface O3 bias. The lower temporal coefficients (R) in the control run versus the MRE, is not an inherent problem of the modeling system. We attributed the deterioration of R the data assimilation procedure, related to the MLS bias correction, described in detail in the paper of Inness et al., 2013. The bias correction of MLS data, has caused drifts in the tropospheric ozone concentrations between August 2004 and December 2007, an issue which have been tracked down and alleviated after year 2008 of the MRE. After 2008, R improves.

Comments are inserted in the revised manuscript. Figure 1 shows that temporal correlation of the MRE increases after bias correction (2008-2012).

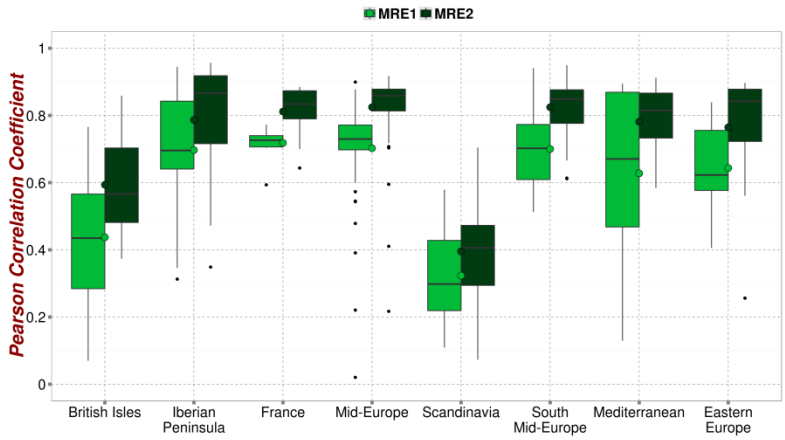


Figure 1: Annual Whisker plots for surface temporal correlation for MACC reanalysis averaged over 2003-2007 with bias correction (MRE1, light green) and over 2008-2012 without bias correction (MRE2, dark green).

A few statements also need to be modified in order to better reflect that assimilation is not improving the overall model performances:

P1078 L12: "assimilation reduces the bias in near surface ozone" is not fully correct.

We modified the text in the revised manuscript to be in agreement with Table 2 (Table 3 in the revised manuscript).

"*Assimilation reduces the bias in near surface ozone in most of the European subregions, with the exception of the British Isles (13% in the MRE and 7% in the CTRL) and the Iberian Peninsula (15% in the MRE and 10% in the control)* "

P1087 L1: reorganise the whole paragraph to start by making the case that assimilation is detrimental overall, before going into the exceptions where it improves model performances.

We do not believe the impact of assimilation is overall detrimental, on the contrary. We expanded the discussion on the impact of assimilation more in detail and we discuss more thoroughly the issue of MLS bias correction and the misbehavior that has caused in surface ozone, an issue that has been alleviated after the year 2008 in the reanalysis. The manuscript is modified accordingly.

P1087L14: it is not objective to discuss only the improvements brought about by assimilation in Fig 3 while it is clear from that Figure that assimilation can also degrade performances in many instances.

We agree with the reviewer: The issue of the deterioration in the temporal correlation is more thoroughly discussed in the revised manuscript.

P1081 L15 : please add a couple of sentence to explain which type of observation and chemical compounds are assimilated. It is not satisfactory to limit to an external reference, especially given that this reference is not available (even on GMDD) to date.

Table 1 has been added in the revised manuscript, along with some additional information in section 2.1 with respect to assimilation.

P1083: please confirm that none of the measurement used for validation are assimilated.

We do confirm. A note has been made in section 2.2

Springtime ozone maximum

The potential processes contributing to this springtime peak are introduced too late in the paper (Section 4 P1090 L15). Given the importance of this feature throughout the article they should be presented in the introduction (P1080 L 24-25), also discussing how the model is expected to capture these processes.

We agree with the reviewer, and so we move this part of the discussion in the introduction.

P1080 L25 : I am struggling with the logical link with the previous sentence, I don't see how assimilation can help in better understanding processes, please explain.

This sentence is removed and a new small paragraph is added in the introduction section, to summarize the structure of the presented work.

P1087L22: include CTRL in this section and corresponding Figures. The difference in temporal coverage is not a good enough reason, it would not be a problem if this figure would be limited to 2003-2010. It seems that other reviewers are sharing that concern.

In the current manuscript we provide the basics of the comparison between the ctrl and the reanalysis (Fig 3 and Table 2). We provide a document for further reading, which is an extended report on the impact of assimilation on surface ozone for the 2003-2010, available as a VAL technical report, Deliverable D84.2 "Validation report on the Comparison of surface ozone in the global (2003- 2010) and regional reanalysis (2011) over Europe".

We note that the basic findings of our analysis are robust and do not change, when the analysis is limited to the 2003-2010. We prefer therefore, to have the analysis cover the whole time period (2003-2012).

L1089 L8: a link to the following discussion section should be included here since possible causes of the failure to capture the early springtime peak will be given there. There is also a

risk of inconsistency when mentioning the findings of Inness (ACP 2013) here whereas the present paper seems to point out new causes for this model caveat.
A link to the following section is provided : "*This shortfall of MRE to capture the early spring peak has been also noted by Inness et al. (2013) and it is further discussed in the following sections*"

P1092 L3: The potential role of the loose coupling between ABL and FT is very interesting. Please include a detailed formulation on how turbulent mixing at the top of the ABL is handled in the model and how it could be improved.

It is not so easy to reach any definite conclusions on the issue of coupling ABL and FT. Transport in the MRE is a mixture between the MOZART diffusion scheme and the IFS diffusion scheme. On top of it, there is data assmilation and the prescribed vertical correlations (see Inness et al. 2013), which also has an influence on the profiles.

The vertical transport is treated more consistently in the new on-line integration (C-IFS) (presented by Flemming et al., 2015) than in the coupled system IFS-MOZART, which was used in the MACC reanalysis.

P1093L7: the more intense oxidation from NOx to NOz in BI and ME is also interesting, what could be the reason for this? You may consider adding a comparison with the model indicators proposed by Beekmann and Vautard (ACP 2010) in order to better document chemical regimes. Using such indicators would also allow drawing more robust conclusion than leaving production and loss analyses to further work in the conclusion (P1096 L1).
The differences in local photochemical ozone production at BI and ME versus IP are consistent with the chemical regime indicator analysis for near surface ozone over Europe by Beekmann and Vautard (2010), who defined three particular regions: a) the region in North-Western Europe with a pronounced VOC sensitive regime (1W–6 E, 50 N–53 N), b) the Mediterranean region (6W–20 E, 38 N–43 N) with an average NOx sensitive chemical regime and c) Northern-Eastern Germany (9 E–14E, 50 N–54 N) which is a transition region between both regimes. Comparing this chemical regime analysis with our selected sub-regions BI, ME and IP we note that BI and ME sub-regions are a mixture of a VOC sensitive regime and a NOx sensitive regime while IP is a NOx sensitive regime.

We added this discussion in the manuscript.

Seasonal cycles

It is surprising that the present paper is limited to comparison of monthly values and daily cycles, while the model is available on a 3-hr basis (P1082 L14). Validation of daily ozone variability was presumably deliberately left aside of this paper. Please explain why.
The metrics were calculated twice. The first time with monthly mean values and the second time using daily mean values. The metrics are most of the times improved when the monthly values are used, but the differences do not alter our main findings, and therefore we prefer to keep the manuscript concise, including only the monthly analysis.
The results based on daily values are presented in the technical report Deliverable D84.2 "Validation report on the Comparison of surface ozone in the global (2003- 2010) and regional reanalysis (2011) over Europe" in Table 1 (provided as Supplementary file in this response)

Table 1. Mean values of basic statistics parameter for each region. Mean Bias (MB), Root Mean Square Error (RMSE), Modified Normalized Mean Bias (MNMB), Fractional Gross Error (FGE), Pearson Correlation Coefficient (r). Top and bottom panel illustrates results that were calculated from daily and monthly values respectively.

	MRE					sa				
Region	MB	RMSE	MNMB	FGE	r	MB	RMSE	MNMB	FGE	r
AL	1.19	14.63	-0.01	0.42	0.48	1.40	14.50	0.01	0.41	0.50
BI	4.98	11.06	0.14	0.31	0.55	-1.43	9.29	-0.10	0.31	0.60
EA	0.01	13.68	-0.06	0.39	0.45	-3.81	13.85	-0.27	0.49	0.53
FR	3.01	12.07	0.05	0.33	0.56	0.87	10.67	-0.05	0.32	0.62
IP	7.57	14.42	0.18	0.33	0.36	5.42	12.66	0.14	0.30	0.38
MD	0.30	13.58	0.02	0.35	0.39	1.20	15.17	0.04	0.39	0.41
ME	1.86	12.08	0.05	0.38	0.61	-2.46	11.48	-0.20	0.44	0.68
SC	-4.17	13.09	-0.20	0.42	0.34	-8.63	13.19	-0.39	0.48	0.47
	MRE					sa				
Region	MB	RMSE	MNMB	FGE	r	MB	RMSE	MNMB	FGE	r
AL	0.59	12.12	-0.04	0.34	0.56	-1.58	11.67	-0.01	0.32	0.58
BI	4.99	9.34	0.13	0.25	0.49	0.94	7.19	-0.08	0.23	0.59
EA	-0.14	10.55	-0.05	0.29	0.58	0.34	11.36	-0.25	0.40	0.62
FR	2.94	9.29	0.05	0.24	0.69	0.92	7.83	-0.04	0.23	0.74
IP	7.43	12.07	0.18	0.28	0.42	0.74	10.05	0.13	0.24	0.46
MD	-0.08	10.62	0.01	0.27	0.44	0.69	13.37	0.03	0.34	0.46
ME	1.81	8.95	0.03	0.26	0.71	0.76	8.87	-0.18	0.33	0.77
SC	-4.24	10.97	-0.18	0.34	0.26	-12.64	11.99	-0.37	0.43	0.40

P1085L20: it is non-standard to compute correlations on the basis of monthly values for surface data, please repeat throughout the text that monthly statistics are used to avoid confusion.

We add in section 2.3 that R is calculated out of mean monthly data.

P1086 L1: consider transposing Figure 2. The text discusses the amplitude of seasonal scores which would be much easier to grasp with one panel per region instead of one panel per season.

We prefer to keep fig in the current orientation format to be in accordance with the rest of the figures.

P1086 L26: it is likely that the correlation is influenced by the amplitude of the cycle, please compare with a rank correlation

In our manuscript we calculated Spearman correlation coefficient. Additionaly, we provide R Kendall as a metric for rank correlation. Results are very similar (Figure 2 of this review).

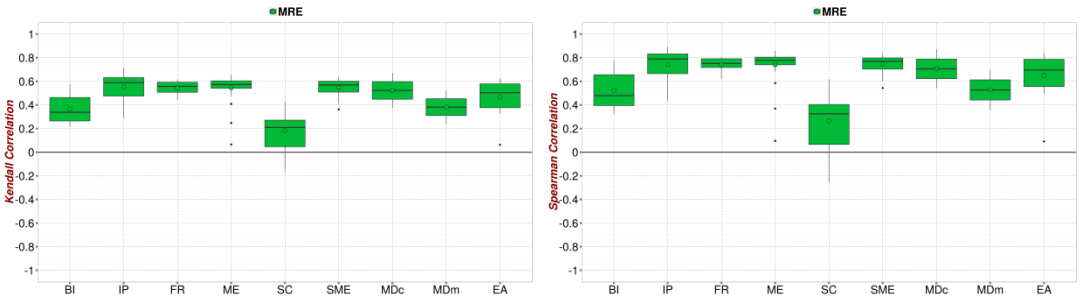


Figure 2. Calculation of R_kendal (left) and R_spearman (right) for the different subregions

Minor remarks

- P1079 L20 : the correct acronym is "AQMEII"
  Done
- P1079L21 : "it is therefore useful"
  Done
- P1080L8 : "of these precursors"
  Done
- P1080L17 : Vestreng et al. 2009 does not address trends in peak ozone
  The sentence was slightly changed, to cite correctly the work of Vestreng, et al.
  "*Although a number of measures aimed at reducing NOx and VOC emissions have been effective in reducing concentration of precursor species (Vestreng et al., 2009) and peak ozone values in Europe (EMEP/CCC-Report 1/2005*) …."

- P1081 L24: define what would be a "clean" control
  Done. A full definition of the clean control is now provided, according to the Inness et al, 2013 paper.

- P1082 L10: what is the temporal coverage of the stations selected here, did you limit the study to stations covering the whole period?
  Yes, as mentioned in the same section, we used only stations that fulfill the criteria of the Jolly-Peuch classification and are available for the whole time period examined, with 75% data availability for near surface ozone.

- P1086 L5: the coloured point is not next to the boxplot for Fig 2.
  The sentence is corrected to: "The colored point on each box indicate…"

- P1095 L23: wrong indentation.
  Done

Evaluation of near surface ozone over Europe from the MACC reanalysis

E. Katragkou[1], P. Zanis[1], A. Tsikerdekis[1], J. Kapsomenakis[2], D. Melas[3], H. Eskes[4], J. Flemming[5], V. Huijnen[4], A. Inness[5], M.G. Schultz[6], O. Stein[6], C.S. Zerefos[2]

[1]{Department of Meteorology and Climatology, School of Geology, Aristotle University of Thessaloniki, Thessaloniki, Greece}

[2] {Research Centre for Atmospheric Physics and Climatology, Academy of Athens, Athens, Greece}

[3]{Laboratory of Atmospheric Physics, School of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece}

[4]{KNMI, De Bilt, the Netherlands}

[5] {ECMWF, Reading, UK}

[6] {Forschungszentrum Jülich, Jülich, Germany}

Correspondence to: E. Katragkou (katragou@auth.gr)

Abstract

This work is an extended evaluation of near surface ozone as part of the global reanalysis of atmospheric composition, produced within the European Funded project MACC (Monitoring Atmospheric Composition and Climate). It includes an evaluation over the period 2003-2012 and provides an overall assessment of the modelling system performance with respect to near surface ozone for specific European subregions. Measurements at rural locations from the European Monitoring and Evaluation Program (EMEP) and the European Air Quality Database (AirBase) were used for the evaluation assessment. The ~~annual overall~~ fractional gross error of near surface ozone reanalysis is on average 24% over Europe, the highest found over Scandinavia (27%) and the lowest over the Mediterranean marine stations (21%). Near surface ozone shows mostly a negative bias in winter and a positive bias during warm months. Assimilation reduces the bias in near surface ozone in most of the European subregions - with the exception of the British Isles and the Iberian Peninsula- and its impact is mostly notable in winter. With respect to the seasonal cycle, the MACC reanalysis reproduces the

photochemically driven broad spring-summer maximum of surface ozone of central and south Europe. However, it does not capture adequately the early spring peak and the shape of the seasonality at northern and north-eastern Europe. The diurnal range of surface ozone, which is as an indication of the local photochemical production processes, is reproduced fairly well, with a tendency for a small overestimation during the warm months for most subregions (especially in central and southern Europe). Possible reasons leading to discrepancies between the MACC reanalysis and observations are discussed.

1 Introduction

The European projects MACC (Monitoring Atmospheric Composition and Climate) ~~,~~ and MACC-II (Interim Implementation) were established under the umbrella of the European Copernicus programme, formerly known as GMES (Global Monitoring for Environment and Security), to build and demonstrate a core capability for providing a comprehensive range of services related to the chemical and particulate composition of the atmosphere (Hollingsworth et al. 2008; Flemming et al., 2009; Inness et al., 2013). Within MACC operational forecasts of atmospheric composition on global (Stein et al., 2012) and regional scale are produced. Furthermore, the MACC reanalysis (Inness et al., 2013) provides global atmospheric composition fields which can be used to serve as boundary conditions for regional air quality models over Europe and world-wide.

The MACC global model used for both reanalysis and forecasts consists of the European Center for Medium-Range Weather Forecasts' (ECMWF)~~s'~~ Integrated Forecast System (IFS) coupled to the MOZART-3 (Kinnison et al., 2007) chemistry transport model. The ECMWF modelling system makes use of its data-assimilation capabilities to combine observations of atmospheric composition with the numerical model in order to produce a reanalysis of atmospheric composition (Inness et al., 2009; Inness et al., ~~2014~~2015). ECMWF has ~~year-long~~ many years of experience in producing reanalysis products, starting from ERA-40 (Dethof and Holm, 2004) and continuing with ERA-Interim (Dragani, 2010, 2011).

Evaluation of MACC data is being done on a regular basis (Eskes et al., ~~2014~~2015) and specifically for trace gases in the global troposphere (e.g. Stein et al., 2014) and the stratosphere (e.g. Lefever et al., 2014). The global reanalysis products are mostly used as a reference dataset for specific case studies (e.g. Knowland et al, 2014) or as boundary conditions for international activities, like the Air Quality Modelling Evaluation International Initiative-AQMEII (Air Quality Modelling Evaluation International Initiative) starting from

phase I (e.g. Schere et al., 2012) up to its current phase III. It is ~~useful,~~ therefore useful to have a systematic analysis on a key atmospheric species of the global reanalysis product i) as a reference for those wishing to use it in their studies ii) as a general assessment of the system performance, identifying potential issues needing further improvement.

In this work special emphasis is given on the evaluation of near surface ozone over Europe for the whole reanalysis period produced within MACC (2003-2012). ~~Tropospheric ozone is an important trace gas controlling the oxidation capacity of the atmosphere (Penkett, 1988; Crutzen, 1988) and acting as a greenhouse gas in terms of radiative forcing at the Earth's surface (IPCC, 2007). Furthermore n~~Near surface ozone is one of the main pollutants affecting both human health and vegetation (Fuhrer and Booker, 2003; Scebba et al., 2005; Schlink et al., 2006). Sources of tropospheric ozone can be either the stratosphere-troposphere transport or the photochemical production through oxidation of VOCs (volatile organic compounds) and CO in the presence of adequate NOx (NOx=NO2+NO) concentrations (Lelieveld and Dentener, 2000). It can be destroyed photochemically or by dry deposition at the surface. Ozone precursors have natural as well as anthropogenic sources, the most important of which are emissions from soil, ~~/~~vegetation and fossil fuel combustion. Ambient ozone concentrations depend strongly on availability and relative abundance of those ~~gaseous~~ precursors but they are also modulated by the meteorological conditions (Davies et al., 1992; Bloomfield et al., 1996; Baertsch-Ritter et al., 2004; Hegarty et al., 2007; Kalabokas et al., 2008).

The issue of the short-term and long-term ozone variability is complex, being related to changes of anthropogenic and natural emissions, meteorological conditions, atmospheric boundary layer mixing processes and stratosphere-troposphere exchange. Although a number of measures aimed at reducing NOx and VOC emissions have been effective in reducing concentration of precursor species (Vestreng et al., 2009) and peak ozone values in Europe (EMEP/CCC-Report 1/2005; ~~Vestreng et al., 2009~~), there are many studies suggesting that background tropospheric ozone levels (even ~~at~~ near the surface) are increasing (Chevalier et al., 2007; Ordóñez et al., 2007; Hess and Zbinden, 2013; Wilson et al., 2012; Akritidis et al., 2014). Parrish et al., (2012) reported a slower rate of increase over the last decades at European sites, to the extent that at present O3 is decreasing at some sites, mostly in summer.

Furthermore, although the current consensus view is that photochemistry is the major contributor to the observed background ozone levels in the troposphere, there is still no

consensus as to the mechanisms that lead to the formation of the spring ozone maximum observed in certain locations of the northern hemisphere, distant from nearby pollution sources (Crutzen et al., 1999; Lelieveld and Dentener, 2000; Monks ~~et al.~~, 2000; Zanis et al., 2007). ~~Hence the evaluation of the MACC near surface ozone is essential in order to assess the performance of the assimilated global reanalysis system with regard to a key near surface pollutant.~~ The spring ozone maximum observed in certain locations of the northern hemisphere, distant from nearby pollution sources, has mainly two contributions; i) the stratosphere to troposphere transport (STT) (Stohl et al., 2003 and references therein) and ii) ozone production in the troposphere on a hemispherical scale, related to photochemical processing of precursor tropospheric trace gases (CO, NOx, VOCs) built up in winter (Penkett and Brice, 1987) and the longer lifetime of ozone during winter that allows anthropogenically produced ozone to accumulate (Lie et al., 1987; Yienger et al., 1999).

In this paper we evaluate near surface ozone of the MACC reanalysis over Europe from 2003 to 2012. We provide an overall assessment of the model performance, putting special emphasis on the reproduction of annual and diurnal cycles. When possible, we provide potential explanations for model inabilities to reproduce specific observational characteristics of certain subregions and finally we suggest points of future work.

2 Methodology

2.1 Global model

The IFS includes greenhouse gases (Engelen et al., 2009) and aerosols (Benedetti et al., 2009; Morcrette et al., 2009). In MACC, the MOZART-3 chemistry transport model has been coupled to the IFS to provide chemical tendencies for ozone, carbon monoxide, nitrogen oxides, and formaldehyde (Flemming et al., 2009), while chemical data assimilation for these species takes place in IFS (Inness et al., 2009; Inness et al., ~~2014~~2015). MOZART-3 as used in the MACC reanalysis system is described in Stein et al. (2012; 2013).

A data assimilation system for aerosol, greenhouse gases and reactive gases is in place based on ECMWF's 4D-VAR data assimilation system. The fields of MACC reanalysis (hereafter MRE) are available globally at a horizontal resolution of ~80 Km (T159 spectral resolution) and 60 hybrid sigma-pressure levels from the surface up to 0.1 hPa. More details on the CTM and the IFS configurations and the data assimilation system are provided by Inness et al.

(~~2014~~2015) and references therein. A combination of profile and total column ozone retrievals was assimilated in MRE, namely GOME, MIPAS, MLS, OMI, SBUV/2, SCIAMCHY (Table 1) using ECMWF's 4D-Var assimilation algorithm (Courtier et al., 1994). For a more detailed description of the assimilation setup see Inness et al. (2013). It should be noted that no tropospheric ozone data were assimilated, so that the impact of the assimilation on near surface ozone comes from the residual of assimilating stratospheric and total column ozone

Since several satellite instruments are used to assimilate one parameter in the data assimilation system, a bias correction method is applied to the data to account for the instrumental inconsistencies. In MRE a variational bias correction scheme for radiance data has been extended to atmospheric composition data (Inness et al., 2013). In the variational scheme biases are estimated during the analysis by including bias parameters in the control vector. The bias corrections are continuously adjusted to optimize the consistency with all information used in the analysis. ~~Mind that T~~the impact of assimilation on near surface ozone is only the "residual" of correcting the stratospheric and total ozone column, plus the assimilation of other relevant gases that impact ozone chemistry (CO, NO2) (Inness et al., 2013).

To investigate the impact of assimilation on key atmospheric species, a control run was also performed (herafter CTRL), using the same reanalysis settings without assimilation. As explained in Inness et al. (2013) (section 2.5), it would have been computationally too expensive to produce a control analysis experiment that was identical to the MACC reanalysis, but did not actively assimilate observations of reactive gases. Instead, a MOZART-3 stand-alone run was carried out that applied the same settings (model code, resolution, emissions) as MOZART in the MACC reanalysis. The meteorological data for the stand-alone run were taken from the reanalysis, but the control run had free-running chemistry. The results from this control run can be used to detect the impact of the assimilation of GRG observations in the MACC reanalysis. Since the meteorological input data were derived from interpolation of archived 6-hourly output from the MACC reanalysis, and not through hourly exchange as in the reanalysis, the stand-alone run was not a completely clean control run. However, these differences would be small.~~The control run is not a "clean" control analysis experiment, but a MOZART-3 stand alone simulation with the reanalysis settings (Inness et al., 2013). It has a free running chemistry while meteorology is~~

~~taken from the 6-hourly reanalysis output. In this work we use the comparison of the CTRL and the MRE to identify the general trend on the impact of assimilation on near surface ozone.~~ The comparison between the MRE and the CTRL is confined to the time period 2003-2010, when both time series are available.

2.2 Observations

Measurements from ground based European stations were used for the evaluation of modelled surface ozone, from the European Monitoring and Evaluation Programme (EMEP) and the European Environment Agency databases (AirBase) covering the time period from 2003 to 2012. The observations used for this evaluation are independent from the assimilated ones. EMEP is appropriate to evaluate coarse resolution simulations, as it is fitted to catch background air pollution patterns with stations at a considerable distance from source areas in rural or remote regions (Schaap et al., 2015). Only background rural stations have been used from the AirBase database for comparisons with the coarse resolution model surface ozone. These include stations ~~type~~ class 1-3 according to the Joly-Peuch classification methodology for surface ozone (Joly and Peuch, 2012). There is a total of 138 stations included in the current analysis, fulfilling the above-mentioned criteria. This selection ensures that all stations are adequate for comparisons with coarse resolution (80 km) model data.

Observed data from the EMEP and AirBase database were available in hourly resolution, while model values were available in 3-hourly intervals. The corresponding observational data were extracted with a 3-hourly interval, to be comparable with modelled time-series. The modelled data were extracted from the coupled system by means of interpolating surface ozone into each station location. Different model levels were used for comparison with ground based stations. The rationale behind the selection of different model level selection instead of extracting time series from the first model level (surface) is that in coarse resolution grids, areas with anomalous terrain (e.g. mountainous areas) are represented with an average elevation, which is less than the actual station elevation. Based on the difference between the actual station altitude and the average grid-cell elevation, the corresponding model level is selected, using atmospheric pressure as the correction criterion. We have used only those stations that fulfil the criteria of 75% data availability for near surface ozone.

In order to acquire a more detailed view of model performance, eight European subregions have been defined as shown in Figure 1. These regions fit data coverage and avoid

overlapping between each subregion. The eight European subregions are: the British Isles (BI), France (FR), Iberian Peninsula (IP), East Europe (EA), Middle Europe (ME), Mediterranean (MD), South Middle Europe (SME) and Scandinavia (SC). Furthermore, the Mediterranean region was further split into the continental part (MDc) and the marine part (MDm), according to their spatial location (coastal or interior continental), since each type of station has different characteristics.

Additional NO and NO_2 data are included in the analysis, in order to assess the potential of the photochemical ozone production. The NO and NO_2 were extracted from EMEP and AirBase. Unfortunately the number of EMEP stations that provide NO and NO_2 measurements – besides ozone~~O3~~ – for the whole reanalysis period (2003-2012) is limited (30 stations). After application of the station type classification for ~~O3~~ ozone and the data availability criteria, only 3 subregions with both O3 and NOx measurements remained, namely the British Isles (BI) with 10 stations, Iberian Peninsula (IP) with 8 stations and Middle Europe (ME) with 12 stations. The plots referring to ozone and nitrogen-species comparison correspond to a smaller number of the common stations mentioned above, always being a subset of the total.

We have also to take into consideration that the NOx observations are affected strongly by local emissions. Furthermore ~~, while~~ there are known issues with interference by oxidized nitrogen compounds (e.g. HNO3, PAN and other organic nitrates) ~~such as nitric acid (HNO3), peroxyacetyl nitrate (PAN), and other organic nitrates~~ for ground-based NO2 measurements by most commercially available NO2 instruments using molybdenum converters, hence leading to an overestimation of NOx concentrations (Steinbacher et al., 2007).

Ozonesondes are used to validate ozone MRE profiles into the troposphere at 6 European stations: Haute-Provence (43.9N, 5.7E), Hohenpeissenberg (47.8N, 11E), Legionowo (52.4N, 20.9E), Payerne (46.8N, 6.9E), Sodankyla (67.4N, 26.6E) and Uccle (50.8N, 4.3E). The sondes used for the validation come from Network for the Detection of Atmospheric Composition Change (NDACC; ftp://ftp.cpc.ncep.noaa.gov/ndacc/station). The precision of electrochemical concentration cell ozonesondes in the troposphere is between −7% and +17% below 200 hPa (Komhyr et al., 1995).

2.3 Metrics and intercomparison methodology

For the current evaluation study we use statistical metrics to quantify the bias, gross error and temporal correlation of the model with regards to observational surface ozone. Comparisons of the diurnal ranges and cycles are also performed, as indices of photochemical processes. As is also discussed by Savage et al. (2013), spatial and temporal variations in chemical composition, including tropospheric ozone, can be large, while also differences between model and observed values are frequently much larger in magnitude than usual for meteorological variables. Therefore, mean error and root mean square error, even though being important metrics for estimating model errors, are not optimal when assessing model performance at different chemical regimes as found over Europe.

Based on the evaluation guidelines and previous work within GEMS/MACC (Seigneur et al., 2010; Elguindi et al., 2010; Ordonez et al., 2010; Eskes et al., ~~2014~~2015) we use the Modified Normalized Mean Bias (MNMB) as a measure of the bias of modelled versus observed values. This metric treats over- and underprediction in a symmetric manner ranging between -2 and 2, in contrast to normalized mean bias that can grow to very high values much greater than unit. The MNMB is calculated from equation (1) as follows:

$$MNMB = \frac{2}{N}\sum_{i}^{N}\frac{f_i - o_i}{f_i + o_i} \quad (1)$$

where f_i and o_i are the mean monthly modelled and observed values, respectively and N the sample size. Seasonal averages are calculated as: winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

Furthermore as a measure of the overall model error we use the Fractional Gross Error (FGE) calculated from equation (2), with its values ranging between 0 and 2. The advantage of this measure is the linear dependence on the departure, which makes this measure less sensitive to outliers and tails in the distribution as compared to the more standard root-mean square.

$$FGE = \frac{2}{N}\sum_{i}^{N}\left|\frac{f_i - o_i}{f_i + o_i}\right| \quad (2)$$

The Pearson correlation (R) is used for the quantification of the temporal agreement (~~seasonality~~interannual variability), between the mean monthly observational and simulated data, where σ_f and σ_o in equation (3) denote the standard deviation –of the modelled and observed values, respectively:

$$R = \frac{\frac{1}{N}\sum_i (f_i - \bar{f})(o_i - \bar{o})}{\sigma_f \sigma_o} \quad (3)$$

The annual cycle of the diurnal range was calculated from the mean diurnal cycle of each station. The confidence interval for each month was derived using the values of the diurnal range for the stations that reside in the same subregion.

In the following section we present a thorough evaluation of surface ozone covering the years from 2003 to 2012, including the three basic validation metrics, analysis of diurnal/annual cycles and diurnal ranges. Seasonal averages are calculated as: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Additionally, surface ozone data are discussed along with nitrogen oxides, wherever data allows comparisons, in order to characterize different chemistry regimes above Europe, with respect to photochemical production.

3 Evaluation of the 2003-2012 MACC reanalysis near surface ozone

3.1 Validation metrics

The annual statistics of surface ozone are shown in Table ~~1~~2**.** The FGE for the whole reanalysis period (2003-2012) ranges mostly from 21% in Mediterranean marine stations to 27% in Scandinavia. ~~The MACC reanalysis has generally a small MNMB (<15%) and acceptable temporal correlations (0.6 to 0.7), with the exception of the British Isles and Scandinavia which score below average R values.~~ Figure 2 shows the basic validation metrics on a seasonal basis for the MACC reanalysis. ~~Box and whisker plots summarize the following details: the bottom and top of the box are the first and third quartiles (Q1 or 25th percentile and Q3 or 75th percentile) and the vertical line in the box is the median (Q2 or 50th percentile). The colored points next to each box indicate the mean value. Some European sub-regions have a strong seasonal variability with respect to FGE like East Europe, where the FGE ranges from 10 to 40%, while some others exhibit a rather constant overall error throughout the year (e.g. IP). More precisely, in East Europe and Scandinavia surface ozone has larger error in winter/spring-time (40/28% and 30/30% for the two regions respectively) than in summer/autumn (10/17% and 20/18%). On the other hand, regions like the~~ Iberian Peninsula ~~or~~ and Mid-Europe have a more stable performance with respect to FGE, with an average 20% for all seasons. All other regions have errors ranging from 10 to 30% depending

on season. A more thorough analysis on the seasonal behavior of surface ozone is provided in the following section.

The seasonal MNMB in Fig. 2 (middle panel) is close to zero for most subregions. The final MRE surface ozone product, exhibits its highest MNMB for Scandinavia and East Europe in winter (-20%). In summer the MNMB is mostly positive and remains <± 20% for most sub-regions, with the exception of British Isles (+30%). Transitional season (spring/autumn) biases follow the patterns of the preceding season (winter/summer), since the atmospheric trace gases need some time to adjust from the winter to the summer-time chemistry regime.

Figure 2 (bottom panel) shows the temporal correlation of the 2003-2012 near-surface ozone timeseries, build upon mean monthly values, and therefore providing a clue on the representation of ozone seasonality. The lowest correlation is found over Scandinavia (0.26), followed by the British Isles (0.51) and the Mediterranean marine stations (0.54). All other regions have correlations ≥ 0.7.

To investigate the impact of assimilation on near surface ozone we compare the MRE and CTRL simulations with the observations. Table ~~2~~ 3 shows the annual statistics of the MRE and the CTRL simulation. The greatest improvement in the MACC reanalysis because of the assimilation is noted over Scandinavia, where the annual FGE is reduced from 40% to 27%, East Europe (FGE drops from 38% to 25%), Mediterranean continental stations (from 43% to 29%) and Mid Europe (from 31% to 24%). In the same areas the MNMB is also reduced by up to 23% (SC). In France and the Iberian Peninsula there seems to be a small increase in the FGE (~~by~~ 6 and 8% respectively) and a small change in the MNMB (reduced to zero in FR and increased by 5% in IP). Over South Mid-Europe and the Mediterranean marine stations the change in FGE and MNMB is negligible on an annual basis.

The ~~annual~~ temporal correlation of monthly mean timeseries from 2003 to 2010 is ~~slightly~~ reduced in the MRE, especially over the Mediterranean marine stations (drops from 0.74 to 0.49) and Scandinavia (from 0.39 to 0.23). The temporal correlation over Scandinavia is very low, because the MRE cannot capture the spring maximum, as it will be shown in section 3.2. Moreover, the issue of the MLS bias correction in the assimilation procedure has caused drifts in the tropospheric ozone concentrations between August 2004 and December 2007 (a detailed explanation of this issue can be found in Inness et al., 2013). The problem was tracked down and alleviated after year 2008 of the MRE. The deterioration of the temporal correlation in the MRE in comparison to the control simulation can be attributed to the

assimilation procedure followed up to MRE year 2008. Calculation of temporal correlation coefficients before (2003-2007) and after (2008-2012) indicates that R increases in all subregions after removal of MLS bias correction.

Figure 3 shows the comparison of the seasonal FGE, MNMB and R for the MRE and the CTRL near surface ozone over the different European subregions for the common time period 2003-2010. On a seasonal basis ~~(Fig 3)~~ the greatest improvement due to assimilation is seen during the winter months, when the CTRL suffers from the largest negative bias. In summer the impact of assimilation is smaller, eventually because near surface ozone is largely controlled by the photochemical processes. ~~Mind that T~~the assimilation correction on ozone is due to the stratospheric and total ozone column. More results on the impact on tropospheric ozone from assimilation in the stratosphere can be found in Lefever et al. (2014).

3.2 Annual cycle of near surface ozone

The average 2003-2012 observed and MRE annual cycle of near surface ozone is shown in Figure 4. With the only exception of the Mediterranean region (MDc and MDm), the modeled annual cycles of ozone have differences in the shape from the observed ones. The most striking disagreement is seen over Scandinavia (SC), where the MRE captures the annual range (13 ppb: the monthly maximum minus the monthly minimum of the year), but ~~however,~~ completely fails to reproduce surface ozone seasonality. While observations indicate a clear spring maximum (40 ppb), a characteristic ozone behavior in very clean and remote atmospheres in the northern hemisphere (Volz and Kley, 1988), no indication of spring ozone maximum is evident in the MRE surface ozone; on the contrary, a clear lower maximum (35 ppb) is ~~depicted~~ found in late summer.

Over the British Isles (BI) we also note striking differences in the shape of the annual cycle. Specifically, there is disagreement a) in the "timeliness" of the early spring maximum, which is ~~depicted~~ seen in April for observed ozone and the late spring-early summer for the MRE, and b) in the annual ozone range, which is overestimated by about 7 ppb. The overestimation occurs mainly during the summer/autumn season. We should note that, even though the MRE near surface ozone at SC and BI does not capture the observed spring maximum peaking in April, this spring ozone maximum is better ~~depicted~~ seen in the lower free troposphere at 850 hPa and 700 hPa vertical levels of MRE (not shown here).

In Mid-Europe (ME), the observational broad spring-summer maximum (April – July) is captured by the MRE, with a month's time-lag (May to August) ~~revealing~~ causing an underestimation in MRE of 2-3 ppbv from January to April and an overestimation from May to November (Fig 4). ~~Mind that~~ T~~t~~he highest overestimation (ranging from 5 ppbv to 9 ppbv) in MRE is seen during the warm months from June to September. This behavior results to an overestimated annual amplitude in MRE in comparison to observations.

Over the Iberian Peninsula (IP) there is an agreement in the seasonal cycle of MRE near surface ozone with observations, with a broad spring-summer maximum but MRE misses the April peak shown in observations. The amplitude of the MRE annual cycle is also overestimated by roughly 4 ppbv in comparison to observations, mostly stemming from the MRE summer O3 overestimation, with the MRE June-maximum reaching up to50 ppbv, while the observed to 40 ppbv. We should also take into consideration that the seasonal cycle of MRE at 700 hPa shows a broad spring-summer maximum with a peak in April as in near surface observations (~~not shown here~~discussed in Section 4.1).

A similar pattern of differences between MRE and observations are ~~depicted~~ found for France (FR), South Mid-Europe (SME) and Eastern Europe (EA) although over EA the differences are smaller.

Overall, the annual cycles of the observed data reflect the specific subregional characteristics, namely the broad spring-summer maximum at Mediterranean (MDc and MDm) and South Mid-Europe (SME), the broad spring-summer maximum peaking in April at Eastern Europe (EA), Mid-Europe (ME), France (FR) and Iberian Peninsula (IP) and the early spring maximum over northern latitudes at Scandinavia (SC) and British Isles (BI). MRE near surface ozone reproduces fairly well the photochemically driven broad spring-summer maximum of surface ozone of the sub-regions at central and south Europe, however, fails to capture the early spring peak in most of these subregions. This shortfall of MRE to capture the early spring peak has been also noted by Inness et al. (2013) and it is further discussed in the following sections. Furthermore, there is generally a tendency for overestimating the annual amplitude in MRE in comparison to observations.

Factors improving ozone seasonality could be emission strengths and temporal profiles and dry deposition (Val Martin et al., 2014). Ongoing work on the impact of dry deposition on surface ozone indicates that the new on-line dry depositions schemes currently tested in the C-IFS system improve the surface ozone positive bias, appearing mostly over southern Europe

in summer, but cannot completely tackle the spring ozone maximum problem over north Europe (J. Flemming, personal communication, 2015).

3.3 Diurnal cycle of near surface ozone

Figure 5 depicts the mean 2003-2012 diurnal cycle of near surface ozone for each season for the selected European regions. All diurnal cycles have the expected behavior with sharply increasing ozone concentrations during the daytime hours (from 5:00-6:00 UTC in summer and 1-2 hours later in winter to 15:00-16:00 UTC) and decreasing afterwards. The diurnal cycles are more pronounced in the summer season and south Europe due to the more intense photochemistry. The MRE reproduces the diurnal cycle ~~captures quite well the shape of the diurnal cycles with~~ but exhibits positive bias in summer (except for the Mediterranean marine region), which may be persisting during the whole day (BI, SME, IP, ME) or occur mostly during daytime (EA, FR, MDc). In winter there is small negative bias in all regions, except for MDc (positive bias) and BI (zero bias). The transitional seasons have diurnal cycles that share both winter and summertime characteristics: the spring diurnal bias resembles winter with respect to bias, but has the enhanced photochemical diurnal cycle of summer, though not fully developed.

Figure 6 shows the annual cycle of the diurnal range of near surface ozone over the different European subregions. The diurnal range of ozone is a good indication of the potential for the local diurnal ozone build up through photochemical production processes (Zanis et al., 2000). There is generally a good agreement with observations, suggesting that MRE reproduces adequately the observed diurnal ozone range with a tendency for a small overestimation during the warm months for the subregions of central and south Europe. More specifically, over SME, FR and MDc the diurnal range is overestimated during the whole year but, to a lesser extent in colder months, while over EA, ME, BI and SC the overestimation is smaller and restricted during the summer. Hence the diurnal range is overestimated more at the southern regions (SME, FR and MDc) than at the northern regions (EA, ME, BI and SC) and more during the warm months than during the cold months.

4 Discussion

In this section we discuss possible reasons for the differences revealed in the shape of the annual cycle of near surface ozone between observations and MRE and the failure in MRE to capture the early spring peak in most of the subregions. ~~Is known that the spring O3~~

~~maximum observed in certain locations of the northern hemisphere, distant from nearby pollution sources, has mainly two contributions; i) the stratosphere to troposphere transport (STT), (Stohl et al., 2003 and references therein) and ii) ozone production in the troposphere on a hemispherical scale, related to photochemical processing of precursor tropospheric trace gases (CO, NOx, VOCs) built up in winter (Penkett and Brice, 1987) and the longer lifetime of ozone during winter that allows anthropogenically produced ozone to accumulate (Lie et al., 1987; Yienger et al., 1999). Here w~~We discuss possible contributions from the above mentioned processes based on the comparison of MRE ozone profiles with available ozonesonde measurements, as well as on NOx versus O3 annual and diurnal cycles.

4.1 Ozone profiles

Comparison with ozonesonde measurements at different locations (Fig. 7) indicate that MRE ozone profiles ~~reproduce~~ reproduce the basic structure of the profile, overestimating in most cases ozone below the 850hPa. ~~reasonably well the observed ozone profiles for all seasons~~. We note positive and negative biases depending on the location and the altitude, but there is a tendency for a larger positive bias during summer and autumn for most locations below 850 hPa, while the % biases in the middle and upper troposphere are generally smaller. This is in agreement with the study of Inness et al. (2013), who, analyzing MACC reanalysis over the time period (2003-2010), reported a negative bias with respect to ozonesondes above 650 hPa and the largest positive bias below 800 hPa. It should be also considered that the range of the % biases in the troposphere are comparable with the respective precision of electrochemical concentration cell ozonesonde measurements.

Furthermore, the shape of the observed ozone annual cycle (based on the ozonesondes) in lower free troposphere at 700 hPa is reproduced rather well by the MRE (Fig. 8). The course of the annual cycle is also reproduced for the middle troposphere at 500 hPa (not shown here). Despite the biases, the reasonable reproduction of the shape of the observed ozone seasonal cycle by MRE in the middle and lower free troposphere ~~suggests that~~ is consistent with transport processes from the lower stratosphere and the upper troposphere as well as long-range transponrt being are resolved adequately by the MRE.

4.2 NO_x versus O_3 annual and diurnal cycles

According to the analysis of ozone profiles (see Section 4.1) we may assume that assimilation in MRE leads to a reasonable representation of the ozone annual cycles at the middle and upper troposphere, thus mediating for a realistic contribution of STT. It could be hence speculated that differences in the shape of the seasonal cycle of near surface ozone between observations and the MRE could be also linked to the potential of photochemical ozone production and the strength of the exchange between the lower free troposphere and the atmospheric boundary layer (ABL). Two tentative explanations could be provided on the mismatch between model and observations: a) inadequate seasonality/emission strengths in surface emissions of precursor species (some issues discussed in Stein et al., 2014) and b) a loose coupling of the free troposphere to the ABL, which would be responsible for the entrainment of the assimilated free tropospheric O3 into the ABL.

In global scales ~~N~~nitrogen oxides (NO_x) are the limiting precursors for O3 production throughout most of the troposphere, and also directly influence the abundance of the hydroxyl radical concentration in the troposphere (e.g. Crutzen, 1988). At regional scale for rural environments with NO_x values less than a few parts per billion by volume, O3 formation is NO_x limited (Liu et al., 1987) and therefore almost independent of hydrocarbon concentrations, depending of course on the ratio of reactivity-weighted VOC mixture to NOx, which may differ from region to region across Europe (Beekmann and Vautard, 2010). Emissions of NO_x occur primarily as NO, followed by oxidation to NO2 while O3 is photochemically produced as NO_x are consumed in favor of their atmospheric oxidation products NO_z (Liu et al., 1987; Zanis et al., 2007). NO_z comprises mostly of peroxyacetylnitrate (PAN) and nitric acid (HNO_3), along with HNO_4, N_2O_5, NO_3 and other ~~PAN~~ Acyl-peroxy nitrates (APNs) ~~homologues (PANs)~~ and organic nitrates (Emmons et al., 1997). The lifetime of NO_x before photochemical conversion to NO_z is less than a day in summer at mid-latitudes (Logan, 1983).

Here, in order to assess the potential of the photochemical ozone production related to NO_x emissions, we have looked at the annual cycle of NO_x versus the respective annual cycle of O_3, as well as the summertime diurnal cycle of O3 along with the diurnal cycle of NO_x at the different sub-regions of our domain. As mention in Section 2.2, after our station-filtering only 3 sub-regions remained, with a considerable number of stations having both O3 and NO_x measurements; the British Isles (BI), Iberian Peninsula (IP) and Mid-Europe (ME).

Figure 9 shows the annual cycle of O3 and NO_x for BI, IP and ME. At the BI the NO_x levels are overestimated in MRE throughout the year by up to 2 ppbv in comparison to the observations while ozone is overestimated from May to November. The overestimation of NO_x concentrations at MRE may partially account for the positive ozone bias during the warm period of the year, through overestimated photochemical ozone production. At IP and ME, NO_x levels are systematically underestimated in MRE throughout the year, and still ozone is overestimated in MRE – especially during the warm part of the year – despite the NO_x underestimation.

Figure 10 shows the average diurnal cycle of O_3 and NO_x during summer for BI, IP and ME. Discarding any biases in the level of O_3 and NO_x concentrations, it is shown that O_3 builds up during the daytime, while NO_x is consumed in both MRE and observations~~–~~. This daytime NO_x decrease can be attributed~~,~~ to chemical loss through ~~presumably due to~~ oxidation to NO_z. Nevertheless, diurnal meteorological patterns of wind speed and boundary layer height, that lead to higher dilution of primary pollutants at daytime than at nighttime, may also contribute ~~for~~to the diurnal pattern of NOx in Figure 10 (see Figure S1 in the supplementary material). This is supported ~~from~~by the fact that CO in MRE, which is a species with much longer chemical lifetime than NO_x, has a similar diurnal pattern with NO_x.

Based on the diurnal amplitudes of O_3 (ΔO_3 increased over the day) and NO_x (ΔNO_x decreased over ~~-~~the day) shown in Figure 10, we have calculated the~~ir~~ ratio $\Delta O_3/\Delta NO_x$ values for both MRE and observations. The $\Delta O_3/\Delta NO_x$ ratio values for near surface based on MRE are estimated roughly to 3 for BI, 3.5 for ME and 10 for IP. The respective $\Delta O_3/\Delta NO_x$ values based on the observed diurnal amplitudes are roughly 10 for BI, 6 for ME and 10 for IP. Additionally, we have also estimated $\Delta O_3/\Delta NO_x$ ratio values based on MRE at 925 hPa (above near surface but within the atmospheric boundary layer) being roughly 3.5 for BI, 3 for ME and 4 for IP. These ratio values reflect the ozone production efficiency, if we assume that daytime NOx loss is through oxidation to NO_z. In order to compare these $\Delta O_3/\Delta NO_x$ ratio values with theoretical calculations of ozone production efficiency, a zero dimension box model with the CBIV chemical mechanism was implemented to calculate ozone production efficiencies for typical summer conditions using initial conditions for NOx and other gaseous species from MRE at BI, IP and ME. These box model calculations indicated that 3 to 4 molecules of O_3 produced for every molecule of NO_x oxidised at BI and ME and up to 5 pbbv at IP. The above values agree well with ozone production efficiency estimates from previous

studies for summer at rural semi-polluted sites with NO_x more than a few ppbv in Europe and US (Chin et al., 1994; Derwent and Davis, 1994; Rickard et al., 2002). The $\Delta O_3/\Delta NO_x$ ratio values based on MRE are comparable with the box model calculated ozone production efficiency values.

The amplitude of the diurnal cycle of NO_x is much stronger in the MRE, than at observations for BI and ME, which indicates that in MRE we presumably have a more intense local oxidation from NO_x to NO_z. This more intense local oxidation from NO_x to NO_z at BI and ME can lead to higher local photochemical ozone production, which may account for the slightly higher amplitude of the diurnal cycle of O_3 for the MRE than the observations (by roughly 2 pppv at BI and 1 ppbv at ME) and partially for the generally higher O_3 levels of the MRE compared to the observed. The differences in local photochemical ozone production at BI and ME versus IP are consistent with the chemical regime indicator analysis for near surface ozone over Europe by Beekmann and Vautard (2010), who defined three particular regions: a) the region in North-Western Europe with a pronounced VOC sensitive regime (1W–6 E, 50 N–53 N), b) the Mediterranean region (6W–20 E, 38 N–43 N) with an average NO_x sensitive chemical regime and c) Northern-Eastern Germany (9 E–14E, 50 N–54 N) which is a transition region between both regimes. Comparing this chemical regime analysis with our selected sub-regions BI, ME and IP, we note that BI and ME sub-regions are a mixture of a VOC sensitive regime and a NOx sensitive regime, while IP is a NO_x sensitive regime.

In the case of IP, the amplitude of the diurnal cycle of NO_x is similar for both observations and MRE, while the amplitude of the diurnal cycle of O_3 is slightly underestimated in the MRE, indicating that local photochemical ozone production is captured adequately or slightly underestimated. Nevertheless, the ozone levels are generally overestimated for the MRE, implying other processes than local photochemistry as a reason for the positive bias.

5 Summary and Conclusions

In the current work we evaluate the MACC-II reanalysis (MRE) near surface ozone for the time period 2003-2012 using rural stations of the EMEP and AirBase monitoring networks. Overall, the evaluation of MRE near surface ozone with station based observations shows a negative bias in winter over northern Europe and generally positive bias during warm months. With respect to the seasonal cycle, MRE reproduces the photochemically driven broad spring-

summer maximum of near surface ozone at central and south Europe. However, it does not ~~capture~~ adequately capture the shape of the seasonality with a characteristic early spring maximum at northern and north-eastern Europe. The diurnal range of surface ozone, which is as an indication of the local photochemical production processes, is reproduced fairly well in the MACC reanalysis, with a tendency for a small overestimation during the warm months for the subregions of central and south Europe. Comparison of MRE ozone profiles with ozonesonde profiles revealed reasonable reproduction of the shape of the observed ozone seasonal cycle in the middle and lower free troposphere, despite the biases. This suggests that transport processes from the lower stratosphere and the upper troposphere are resolved acceptably by MRE with the aid of the assimilation.

More specifically, the characteristics of near surface ozone in the MACC reanalysis 2003-2012 can be summarized as follows for the different sub-regions:

a) At British Isles and Scandinavia, the observed near surface spring ozone maximum peaking in April is not reproduced by MRE. However, this spring ozone maximum is better ~~depicted~~ seen in the lower free troposphere (at 850 hPa and 700 hPa) implying adequate vertical transport within the free troposphere, ~~of the assimilated~~ as was also indicated by the good comparison with ozonesonde data. The possibility insufficient entrainment and mixing from the lower free troposphere into the atmospheric boundary layer should be further investigated. MRE diurnal range of near surface ozone compares relatively well with the observed diurnal range with a slight overestimation during summer. Analysis of the average MRE diurnal cycle of O_3 versus NO_x during summer for the BI could possibly indicate among other reasons, more intense local oxidation from NO_x to NO_z than the observed and a systematic positive bias in NO_x which can lead to higher local photochemical ozone production.

b) The ozone summer maximum of the Mediterranean area is captured by the MRE, with a slight overestimation during summer and autumn for the continental stations (MDc). The MRE near surface ozone diurnal range compares well with the observed one throughout the year for the marine stations (MDm) and is slightly overestimated during the warm months for the continental stations (MDc). This implies that part of the MRE overestimation of near surface in summer and autumn for MDc may be associated to an overestimation of local photochemical production. Zanis et al.

(2014) also noted for the Mediterranean an overestimation of near surface ozone during summer by another global chemistry–climate model, due to overestimated photochemical ozone production within the atmospheric boundary layer.

c) In East Europe, Mid-Europe, South Mid-Europe and France, MRE near surface ozone reproduces the photochemically driven broad spring-summer maximum, but fails to capture the early spring peak in April. Furthermore, there is a slight shift of the seasonal cycle towards summer in MRE compared to observations, with a tendency for an underestimation of ozone levels in cold months (from January to April) and an overestimation in summer and autumn. The diurnal range of near surface ozone in the MRE is overestimated during summer. This maybe implies an overestimated local photochemical ozone production, which can partially account for the summer overestimated MRE near surface ozone levels (similarly to MDc). Further analysis of the average diurnal cycle of O_3 versus NO_x during summer for Mid-Europe, gives some indication for more intense local oxidation from NO_x to NO_z for the MRE than the observations, which can lead to higher local photochemical ozone production despite the systematic negative bias in NO_x.

d) At the Iberian Peninsula there is a positive bias throughout the year and the MRE does not capture the April peak shown in the observed seasonal cycle. The MRE diurnal range compares relatively well with the observed diurnal range, maybe indicating that local photochemical production is captured adequately throughout the year. This is also supported from the analysis of the average diurnal cycle of O_3 versus NO_x during summer. The seasonal cycle of MRE at 700 hPa shows a broad spring-summer maximum with a peak in April as in near surface observations. This feature could possibly indicate a loose coupling of the free troposphere with atmospheric boundary layer.

Our analysis suggests that in ~~ordet~~ order to ~~understanf~~ understand better the behaviour of near surface ozone, further analysis is needed for firm conclusisons, including model diagnostics for photochemical production and loss terms, as well as the mixing between ABL and free troposphere. Improvement in the dry-deposisiont scheme –which is fixed in the current implementation – would also contribute to improvement of model performance (bias/seasonality) with respect to near surface ozone.

Acknowledgements

MACC II is funded by the European Union's Seventh Framework Programme (FP7) under Grant Agreement no. 283576. We thank the European Environmental Agency (AirBase) and the European Monitoring and Evaluation Programme (EMEP) for providing access to European O3 and NOx observations

References

Akritidis D., Zanis, P., Pytharoulis, I., Karacostas, Th.: Near-surface ozone trends over Europe in RegCM3/CAMx simulations for the time period 1996-2006, Atmospheric Environment, 97, 6-18, 2014.

Baertsch-Ritter, N., Keller, J., Dommen, J., and Pr´evˆot, A. S. H.: Effects of various meteorological conditions and spatial emissionresolutions on the ozone concentration and ROG/NOx limitationin the Milan area (I), Atmos. Chem. Phys., 4, 423–438, doi:10.5194/acp-4-423-2004, 2004.

Bhartia, P. K. and Wellemeyer, C.: TOMS-V8 total O3 algorithm, in: OMI Ozone Product ATBD Volume II, NASA Goddard Space Flight Center, Greenbelt, MD, USA, 2002.

Bhartia, P. K., McPeters, R. D., Mateer, C. L., Flynn, L. E., and Wellemeyer, C., Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, J. Geophys. Res., 101, 18793–18806, 1996.

Beekmann, M. and Vautard, R.: A modelling study of photochemical regimes over Europe: robustness and variability, Atmos. Chem. Phys., 10, 10067-10084, doi:10.5194/acp-10-10067-2010, 2010.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., Suttie, M., and the GEMS-AER team: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.Crutzen, P. J.: Tropospheric ozone: An overview, in: Tropospheric Ozone, edited by: I.S.A. Isaksen, D. Reidel Publ. Co., 3–32, 1988.

Bloomfield, P., Royle, J. A., Steinberg, L. J., and Yang, Q.: Accounting for meteorological effects in measuring urban ozone levels and trends, Atmos. Environ., 30(17), 3067–3077, 1996.

Carli, B., Alpaslan, D., Carlotti, M., Castelli, E., Ceccherini, S., Dinelli, B. M., Dudhia, A., Flaud, J. M., H¨opfner, M., Jay, V., Magnani, L., Oelhaf, H., Payne, V., Piccolo, C., Prosperi, M., Raspollini, P., Ridolfi, M., Remedios, J., and Spang, R.: First results from

MIPAS/ENVISAT with operational level 2 code, Adv. Space Res., 33, 1012–1019, doi:10.1016/S0273-1177(03)00584-2, 2004.

Chevalier, A., Gheusi, F., Delmas, R., Ordonez, C., Sarrat, C., Zbinden, R., Thouret, V., Athier, G., Cousin, J.M.: Influence of altitude on ozone levels and variability in the lower troposphere: a ground-based study for western Europe over the period 2001-2004, Atmos. Chem. Phys. 7, 4311-4326, 2007.

Chin, M., Jacob, D. J., Munger, J. W., Parrish, D. D., and Doddridge, B. G.: Relationships of ozone and carbon monoxide over North America, J. Geophys. Res. 99, 14,565–14,573, 1994.

Courtier, P., Th´epaut, J.-N. and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. Roy. Meteor. Soc., 120, 1367–1388, 1994.

~~Crutzen, P. J.: Tropospheric ozone: An overview, in: Tropospheric Ozone, edited by: I.S.A. Isaksen, D. Reidel Publ. Co., 3–32, 1988.~~

Crutzen, P. J., Lawrence, M. G., and Poeschl, U.: On the background photochemistry of tropospheric ozone, Tellus, 51A-B, 123–146, 1999.

Derwent, R. G. and Davis, T. J.: Modelling the impact of NOx or hydrocarbon control on photochemical ozone in Europe, Atmos. Environ. 28, 2039–2052, 1994.

Dethof, A. and Holm, E. V.: Ozone assimilation in the ERA-40 re-analysis project, Q. J. Roy. Meteor. Soc., 130, 2851–2872, 2004.

Dragani, R.: On the quality of the ERA-Interim ozone reanalyses: comparisons with in situ data, ERA Report Series, 2, available at: http://www.ecmwf.int/publications/library/do/references/list/782009 (last access: 29 November 2012), 2010.

Dragani, R.: On the quality of the ERA-Interim ozone reanalyses: comparisons with satellite data, Q. J. Roy. Meteor. Soc., 137, 1312–1326, doi:http://dx.doi.org/10.1002/qj.82110.1002/qj.821, 2011.

Elguindi, N., Clark, H., Ordóñez, C., Thouret, V., Flemming, J., Stein, O., Huijnen, V., Moinat, P., Inness, A., Peuch, V.-H., Stohl, A., Turquety, S., Athier, G., Cammas, J.-P., and Schultz, M.: Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC, Geosci. Model Dev., 3, 501-518, doi:10.5194/gmd-3-501-2010, 2010.

EMEP/CCC-Report 1/2005, March 2005. The development of European surface ozone. Implications for a revised abatement policy. In: Solberg, Sverre, Lindskog, Anne (Eds.), A Contribution from the EU Research Project NEPAP U-103003.Emmons, L. K., Carroll, M. A., Hauglustaine, D. A., Brasseur, G. P., et al.: Climatologies of NOx and NOy: A comparison of data and models, Atmos. Environ., 31(12), 1851–1904, 1997.

Engelen, R. J., Serrar, S., and Chevallier, F.: Four-dimensional data assimilation of atmospheric CO2 using AIRS observations, J. Geophys. Res., 114, D03303, doi:10.1029/2008JD010739, 2009.

Eskes, H. J., van der A, R. J., Brinksma, E. J., Veefkind, J. P., de Haan, J. F., and Valks, P. J. M.: Retrieval and validation of ozone columns derived from measurements of SCIAMACHY on Envisat, Atmos. Chem. Phys. Discuss., 5, 4429– 4475, doi:10.5194/acpd-5-4429-2005, 2005.

Eskes H., V. Huijnen, A. Arola, A. Benedictow, A. Blechschmidt, E. Botek, O. Boucher, I. Bouarar, S. Chabrillat, E. Cuevas, R. Engelen, H. Flentje, A. Gaudel, J. Griesfeller, L. Jones, J. Kapsomenakis, E. Katragkou, S. Kinne, B. Langerock, M. Razinger, A. Richter, M. Schultz, M. Schulz, N. Sudarchikova, V. Thouret, M. Vrekoussis, A. Wagner, and C. Zerefos, Validation of reactive gases and aerosols in the MACC global analysis and forecast system, Geosci. Model Dev. Discuss., 8, 1117-1169, 2015 ~~Submitted in Geoscientific Model Development Discussions, this issue, 2014~~

Flemming, J., Dethof, A., Moinat, P., Ordonez, C., Peuch, V.-H., Segers, A., Schultz, M., Stein, O., van Weele, M.: Coupling global atmospheric chemistry transport models to ECMWF Integrated Forecasts System for forecast and data assimilation within GEMS, in: Integrated Systems of Meso-Meteorological and Chemical Transport Models, edited by: Baklanov, A., Mahura, A., and Sokhi, R., Springer-Verlag, Berlin Heidelberg, doi:10.1007/978-3-642-13980-2 10, 2011.Fuhrer, J. and Booker, F.: Ecological issues related to ozone: agricultural issues, Environ. Int. 29(2–3), 141–154, 2003.

Flemming J., Inness, A. Flentje, H. Huijnen, V. Moinat, P. Schultz, M. G.. Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, Geosci. Model Dev., 2, 253–265, doi:10.5194/gmd-2-253-2009, 2009.

Hegarty J., Mao, H., and Talbot, R.: Synoptic controls on summertime surface ozone in the northeastern United States, J. Geophys. Res., 112, D14306, doi:10.1029/2006JD008170, 2007.

Hess, P.G., Zbinden, R.: Stratospheric impact on tropospheric ozone variability and trends: 1990e2009. Atmos. Chem. Phys. 13, 2013.

Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J.-J., Rayner, P., Peuch, V. H., Rouil, L., Schultz, M. G., Simmons, A. J., and 5 The GEMS Consortium: toward a monitoring and forecasting system for atmospheric composition: the GEMS project, B. Am. Meteorol. Soc., 89, 1147–1164, 2008.

Inness, A., Flemming, J., Suttie, M., and Jones, L.: GEMS data assimilation system for chemically reactive gases, European Centre for Medium-Range Weather Forecasts (ECMWF), Technical Memoradum No. 587, 2009.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, a. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M. and Zerefos, C.: The MACC reanalysis: an 8 yr data set of atmospheric composition, Atmos. Chem. Phys., 13(8), 4073–4109, doi:10.5194/acp-13-4073-2013, 2013.

Inness, A., Blechschmidt, A., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Errera, Q., Flemming, J., Gaudel, A., Huijnen, V., Jones, L., Kapsomenakis, J.,Keppens, A., Lambert, J.-C., Langerock, B., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G.,Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation experiments of satellite retrievals of O3, CO and NO2 with Composition IFS~~, Geosci. Model Dev. Discus., 2014.~~ Atmos. Chem. Phys., 15, 5275–5303, 2015

~~IPCC, 2007 Climate change 2007: the physical science basis. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 996, 2007.~~

Joly, M. and Peuch, V.-H., Objective classification of air quality monitoring sites over Europe (2012), Atmospheric Environment, 47, pp. 111-123.

Kalabokas, P. D., Mihalopoulos, N., Ellul, R., Kleanthous, S., and Repapis, C. C.: An investigation of the meteorological and photochemical factors influencing the background rural and marine surface ozone levels in the Central and Eastern Mediterranean, Atmos. Environ., 42, 7894–7906, doi:10.1016/j.atmosenv.2008.07.009, 2008.

Knowland K.E., R. M. Doherty, and K. I. Hodges, The effects of springtime mid-latitude storms on trace gas composition determined from the MACC reanalysis. Atmos. Chem. Phys. Discuss., 14, 27093-27141, 2014

Komhyr,W. D., Barnes, R. A., Borthers, G. B., Lathrop, J. A., Kerr, J. B., and Opperman, D. P.: Electrochemical concentration cell ozonesonde performance evaluation during STOIC 1989, J. Geophys. Res., 100, 9231–9244, 1995.

Lefever K., R. van der A, F. Baier, Y. Christophe, Q. Errera, H. Eskes, J. Flemming, A. Inness, L. Jones, J.-C. Lambert, B. Langerock, M. G. Schultz, O. Stein, A. Wagner, and S. Chabrillat, Copernicus atmospheric service for stratospheric ozone: validation and intercomparison of four near real-time analyses, 2009–2012, Atmos. Chem. Phys. Discuss., 14, 12461-12523, 2014

Lelieveld, J. and Dentener, F.: What controls tropospheric ozone, J. Geophys. Res., 105(3), 3543–3563, 2000.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., M¨alkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, IEEE T. Geosci. Remote, 44, 1093–1101, 2006.

Liu, S. C., Trainer, M., Fehsenfeld, F. C., Parrish, D. D., Williams, E. J., Fahey, D. W., Hübler, G., and Murphy, P. C.: Ozone Production in the Rural Troposphere and the Implications for Regional and Global Ozone Distributions, J. Geophys. Res., 92(D4), 4191–4207, 1987.

Monks, P. S.: A review of observations and origins of the spring ozone maximum, Atmos. Environ., 34, 3545–3561, 2000.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J.W., Razinger, M., Schulz, M., Serrar, S., Simmons, A.J., Sofiev, M., Suttie, M., Tompkins, A.M., Untch, A., Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: Forward modeling (2009) Journal of Geophysical Research: Atmospheres, 114 (6)

Ordóñez, C., Brunner, D., Staehelin, J., Hadjinicolaou, P., Pyle, J.A., Jonas, M., Wernli, H., Prevot, A.S.H.: Strong influence of lowermost stratospheric ozone on lower tropospheric background ozone changes over Europe. Geophys. Res. Lett. 34, L07805. , 2007.

Ordóñez C., N. Elguindi, O. Stein, V. Huijnen, J. Flemming, A. Inness, H. Flentje, E. Katragkou, P. Moinat, V-H. Peuch, A. Segers, V. Thouret, M. van Weele, C. S. Zerefos, J-P. Cammas, and M. G. Schultz, Global model simulations of air pollution during the 2003 European heat wave, Atmospheric Chemistry and Physics, 10, 789–815, 2010

Parrish, D.D., Lamarque, J.-F., Naik, V., Horowitz, L., Shindell, D.T., Staehelin, J., Derwent, R., Cooper, O.R., Tanimoto, H., Volz-Thomas, A., Gilge, S., Scheel, H.-E., Steinbacher, M., Fröhlich, M., Long-term changes in lower tropospheric baseline ozone concentrations: Comparing chemistry-climate models and observations at northern midlatitudes (2014) Journal of Geophysical Research: Atmospheres, 119 (9), pp. 5719-5736.

Penkett, S. A. and Brice, K. A.: The spring maximum in photooxidant in the Northern hemisphere troposphere, Nature, 319, 655–657, 1986.

Rickard, A. R., Salisbury, G., Monks, P. S., Lewis, A. C., Baugitte, S., Bandy, B. J., Clemitshaw, K. C., and Penkett, S. A.: Comparison of measured ozone production efficiencies in the marine boundary layer at two European coastal sites under different pollution regimes, J. Atmos. Chem., 43, 107–134, 2002.

~~Penkett, S. A.: Indications and causes of ozone increase in the troposphere, in: The changing atmosphere, edited by: Rowland, F.S. and Isaksen, I. S. A., J. Wiley & Sons, 91, 1988.~~

Savage, N. H., Agnew, P., Davis, L. S., Ordóñez, C., Thorpe, R., Johnson, C. E., O'Connor, F. M., and Dalvi, M.: Air quality modelling using the Met Office Unified Model (AQUM OS24-26): model description and initial evaluation, Geosci. Model Dev., 6, 353-372, doi:10.5194/gmd-6-353-2013, 2013.

Siddans, R., Reburn, W. J., Kerridge, B. J., and Munro, R.: Height resolved ozone information in the troposphere and lower stratosphere stratosphere from GOME, Technical report, British Atmospheric Data Centre (BADC), available at: http://cedadocs.badc.rl.ac.uk/97/ (last access: 29 November 2012), 2007.

Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J.M., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyiri, A., Pay, M.T., Rouïl, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., Wind, P., Performance of European chemistry transport models as function of horizontal resolution (2015) Atmospheric Environment, 112, pp. 90-105

Scebba, F., Giuntini, D., Castagna, A., Soldatini, G., and Ranieri, A.: Analysing the impact of ozone on biochemical and physiological variables in plant species belonging to natural ecosystems, Environ. Exp. l Bot., 235–246, 2005.

Schere, K., Flemming, J., Vautard, R., Chemel, C., Colette, A., Hogrefe, C., Bessagnet, B., Meleux, F., Mathur, R., Roselle, S., Hu, R.-M., Sokhi, R.S., Rao, S.T., Galmarini, S. Trace gas/aerosol boundary concentrations and their impacts on continental-scale AQMEII modeling domains (2012) Atmospheric Environment, 53, pp. 38-50.

Schlink U., Herbarth, O., Richter, M., Dorling, S., Nunnari, G., Cawley, G., and Pelikan, E.: Statistical models to assess the health effects and to forecast ground-level ozone, Environ. Modell. Soft., 21(4), 547–558, 2006.

Stein, O.: Model documentation of the MOZART CTM as implemented in the GEMS system, available at: http://gems.ecmwf.int/do/get/PublicDocuments/1531/1172 (last access: 29 November 2012), 2009.

Stein, O., Flemming, J., Inness, A., Kaiser, J. W., and Schultz, M. G.: Global reactive gases forecasts and reanalysis in the MACC project, J. Integr. Environ. Sci., 9, 57–70, doi:10.1080/1943815X.2012.696545, 2012.

Stein O., M. G. Schultz, I. Bouarar, H. Clark, V. Huijnen, A. Gaudel, M. George, and C. Clerbaux, On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, Atmos. Chem. Phys., 14, 9295–9316, 2014

Steinbacher, M., Zellweger, C., Schwarzenbach, B., Bugmann, S., Buchmann, B., Ordóñez, C., Prevot, A. S. H., and Hueglin, C.: Nitrogen oxide measurements at rural sites in Switzerland: Bias of conventional measurement techniques, J. Geophys. Res., 112, D11307, doi:10.1029/2006JD007971, 2007.

Stohl, A., Bonasoni, P., Cristofanelli, P., Collins, W., Feichter, J., Frank, A., Forster, C., Gerasopoulos, E., Gäggeler, H., James, P., Kentarchos, T., Kreipl, S., Kromp-Kolb, H., Krüger, B., Land, C., Meloen, J., Papayannis, A., Priller, A., Seibert, P., Sprenger, M., Roelofs, G. J., Scheel, E., Schnabel, C., Siegmund, P., Tobler, L., Trickl, T., Wernli, H., Wirth, V., Zanis, P., and Zerefos, C.: Stratosphere-troposphere exchange – a review, and what we have learned from STACCATO, J. Geophys. Res., 108(D12), doi:10.1029/2002JD002490, 2003.

Valmartin, M., Heald, C.L., Arnold, S.R. Coupling dry deposition to vegetation phenology in the Community Earth System Model: Implications for the simulation of surface O3 (2014) Geophysical Research Letters, 41 (8), pp. 2988-2996.

Valcke, S. and Redler, R.: OASIS4 User Guide (OASIS4 0 2), PRISM–Support Initiative, Technical Report No 4, available at: http://www.prism.enes.org/Publications/index.php (last access: 29 November 2012), 2006.

Vestreng, V., Ntziachristos, L., Semb, A., Reis, S., Isaksen, I.S.A., Tarrason, L.: Evolution of NOx emissions in Europe with focus on road transport control measures. Atmos. Chem. Phys. 9, 1503e1520, 2009.

Volz, A., Kley, D., Evaluation of the Montsouris series of ozone measurements made in the nineteenth century, Nature, 332 (6161), pp. 240-242, 1988

Waters, J.W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Van Snyder, W., Tope, M. C.,Wagner, P. A., andWalch, M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, IEEE Trans. Geosci. Remote, 44, 1075–1092, 2006.

Wilson, R.C., Fleming, Z.L., Monks, P.S., Clain, G., Henne, S., Konovalov, I.B., Szopa, S., Menut, L.: Have primary emission reduction measures reduced ozone across Europe? An analysis of European rural background ozone trends 1996-2005. Atmos. Chem. Phys. 12, 437e454. http://dx.doi.org/10.5194/acp-12-437-2012, 2012.

Yienger, J. J., Klonecki, A. A., Levy II, H., Moxim, W. J., and Carmichael, G. R.: An evaluation of chemistry's role in the winter-spring ozone maximum found in the northern midlatitude free troposphere, J. Geophys. Res., 104(D3), 3655–3667, 1999.

Zanis P., P.S. Monks, E. Schuepbach, and S.A. Penkett, The role of in-situ photochemistry in the control of ozone during spring at the Jungfraujoch Observatory (3,580 m asl) – Comparison of model results with measurements, Journal of Atmospheric Chemistry, 37(1), 1-27, 2000.

Zanis, P., Ganser, A., Zellweger, C., Henne, S., Steinbacher, M., and Staehelin, J.: Seasonal variability of measured ozone production efficiencies in the lower free troposphere of Central Europe, Atmos. Chem. Phys., 7, 223–236, doi:10.5194/acp-7-223-2007, 2007.

1 Zanis P., P. Hadjinicolaou, A. Pozzer, E. Tyrlis, S. Dafka, N. Mihalopoulos, J. Lelieveld,
2 Summertime free tropospheric ozone pool over the Eastern Mediterranean/Middle East,
3 Atmospheric Chemistry and Physics, 14, 115–132, 2014.
4

Table 1: Ozone satellite retrievals that were assimilated in the MACC reanalysis. PROF denotes profile data, TC total columns, PC partial columns, and SOE solar elevation. PC SBUV/2 data consist of 6 layers between the surface and 0.1 hPa. NRT (near-real time) data are available within a few hours after the observation was made, and are being used in operational forecast systems. For periods towards the end of the MACC reanalysis period, NRT data were used for some of the species when no offline products were available.

Sensor	Satellite	Provider	Version	Period	Type	Data usage criteria	Reference
GOME	ERS-2	RAL		20030101-20030531	O_3 PROF	Used if SOE>15° and 80°S<lat<80°N	Siddans et al. 2007
MIPAS	ENVISAT	ESA		20030127-20040326	O_3 PROF	All data used	Carli et al. 2004
MLS	AURA	NASA	V02	20040808-20090315, NRT data from 20090316	O_3 PROF	All data used	Waters et al. 2006
OMI	AURA	NASA	V003	From 20041001, NRT data 20070321-20071231	O_3 TC	Used if SOE >10°	Bhartia et al. 2002; Levelt et al. 2006
SBUV/2	NOAA-16	NOAA	V8	From 20040101	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SBUV/2	NOAA-17	NOAA	V8	From 20030101	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SBUV/2	NOAA-18	NOAA	V8	From 20050604	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SCIAMACHY	ENVISAT	KNMI		From 20030101	O_3 TC	Used if SOE>6°	Eskes et al. 2005

Table ~~1~~2. Annual statistics of near surface ozone for the MACC reanalysis (2003-2012) over the different European subregions. FGE and MNMB are expressed in %.

Region	FGE	MNMB	R
BI	23	12	0.51
IP	25	14	0.72
FR	26	-2	0.73
ME	22	3	0.74
SC	27	-13	0.26
SME	24	2	0.74
MD_c	24	20	0.71
MD_m	21	-12	0.54
EA	25	-9	0.66

Table ~~2~~3. Annual statistics of near surface ozone for the MACC reanalysis (MRE) and the control run (CTRL) over the different European subregions for the common period from 2003 to 2010. FGE and MNMB are expressed in %.

Region	FGE		MNMB		R	
	MRE	CTRL	MRE	CTRL	MRE	CTRL
BI	24	22	13	-7	0.51	0.59
IP	25	17	15	10	0.70	0.79
FR	28	22	0	-5	0.73	0.79
ME	24	31	4	-17	0.73	0.80
SC	27	40	-12	-35	0.23	0.39
SME	25	22	3	-5	0.73	0.78
MDc	29	43	26	42	0.71	0.74
MDm	21	19	-10	-12	0.49	0.74
EA	25	38	-8	-28	0.64	0.70

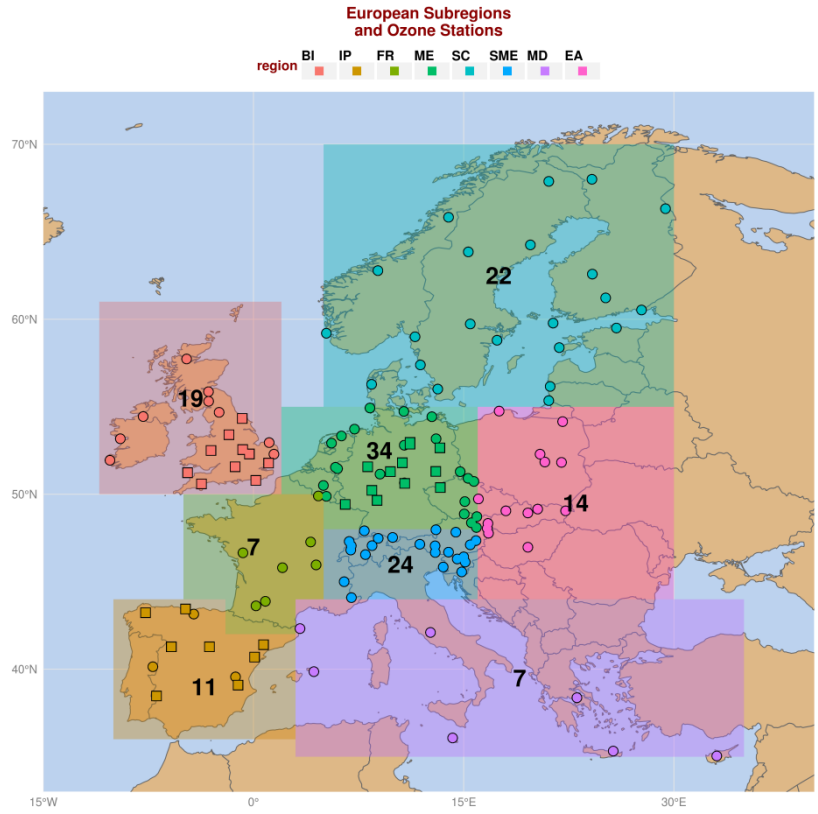


Figure 1. The European subregions that were used in the analysis and the corresponding EMEP and AIRBASE stations. The numbers denote the number of stations taken into consideration for every subregion. ~~For details see text.~~ The subregions are: the British Isles (BI), France (FR), Iberian Peninsula (IP), East Europe (EA), Middle Europe (ME), Mediterranean (MD), South Middle Europe (SME) and Scandinavia (SC).

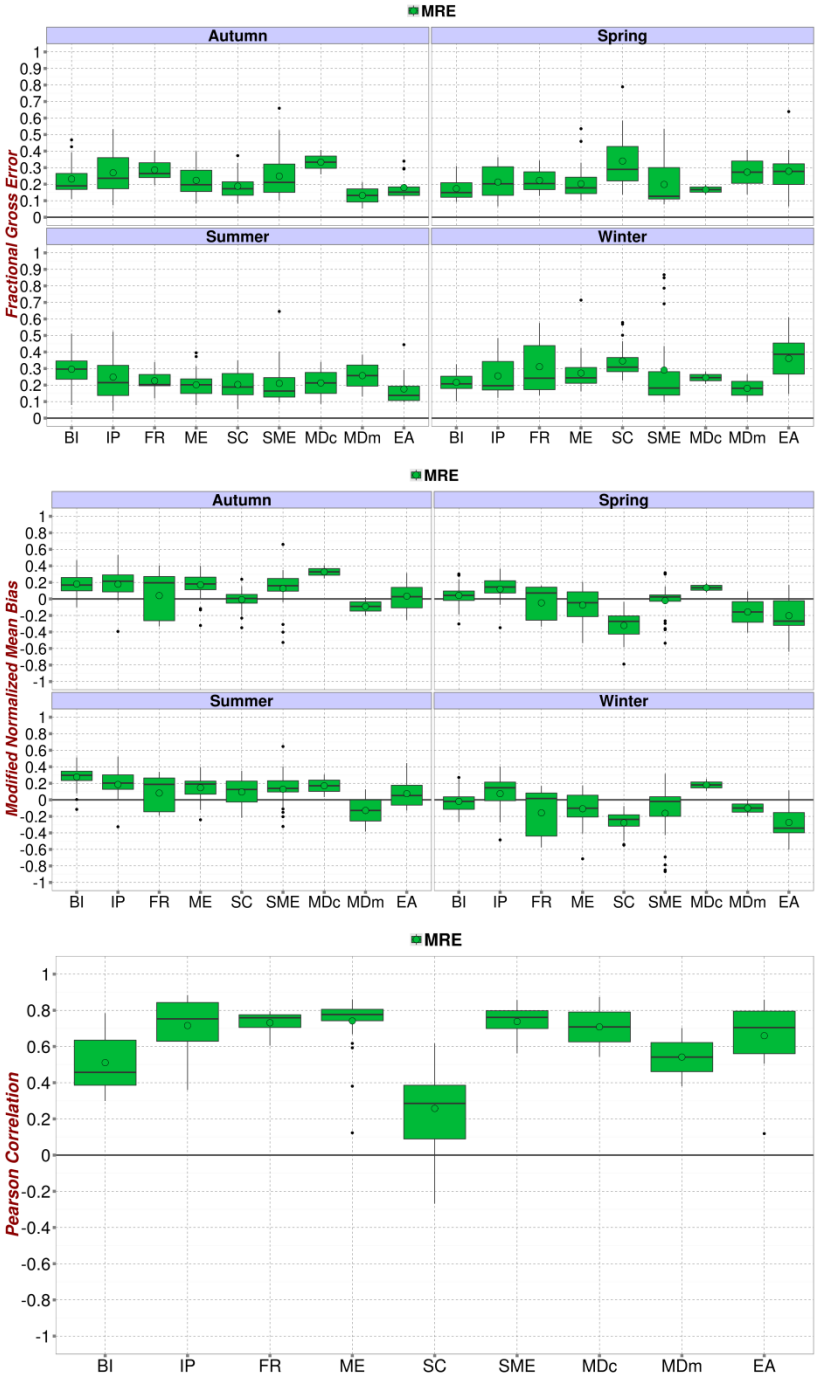


Figure 2. Average 2003-2012 seasonal FGE (top), MNMB (middle) and annual R (bottom) of near surface ozone for the different European subregions of the MACC reanalysis. The color dots correspond to means. The bottom and top of the box are the first and third quartiles (Q1 or 25th percentile and Q3 or 75th percentile) and the vertical horizontal line in the box is the median (Q2 or 50th percentile). The colored points on each box indicate the mean value.

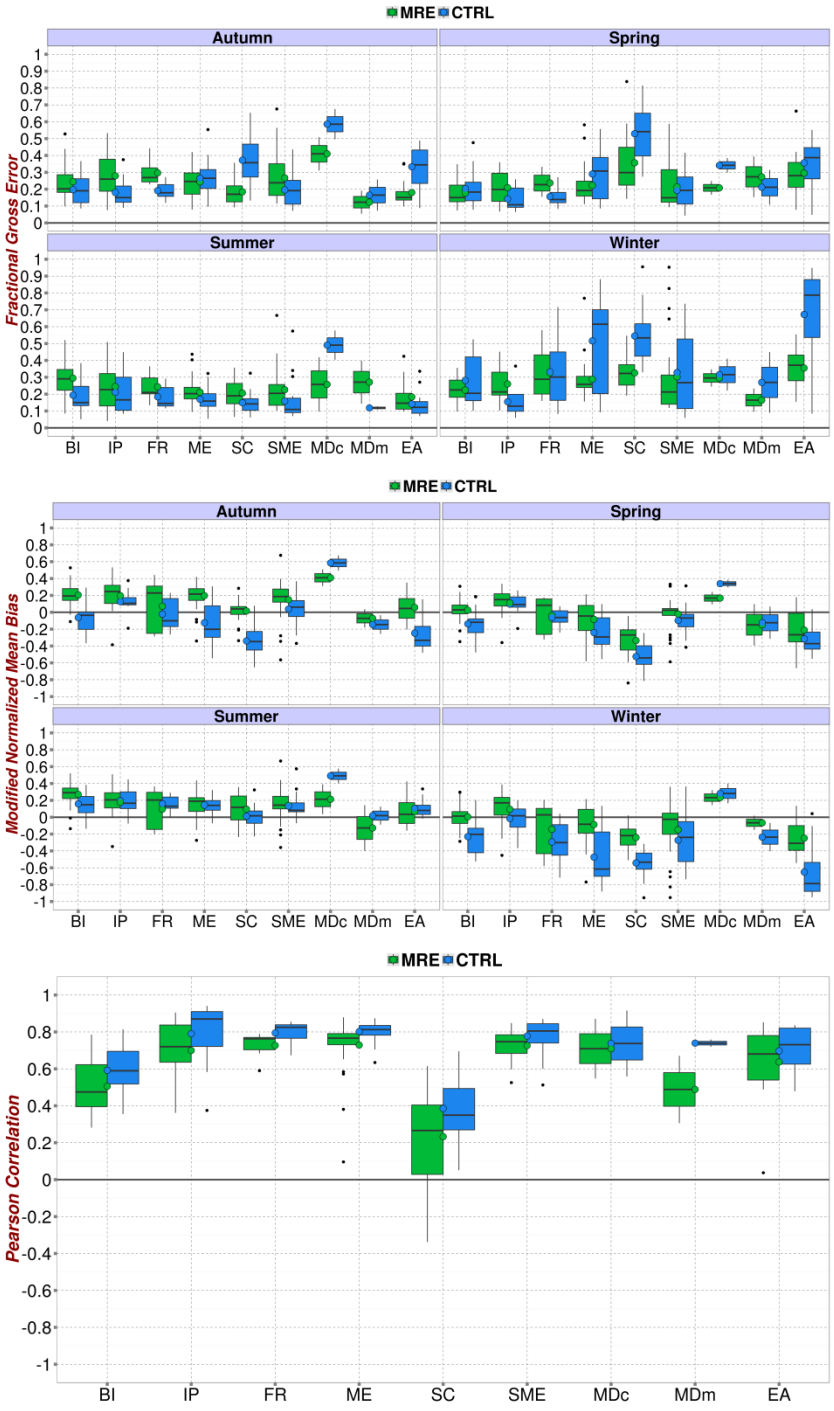


Figure 3 Average 2003-2010 seasonal FGE (top), MNMB (middle) and annual R (bottom) of near surface ozone for the different European subregions of the MACC reanalysis (green) and the control run (blue).

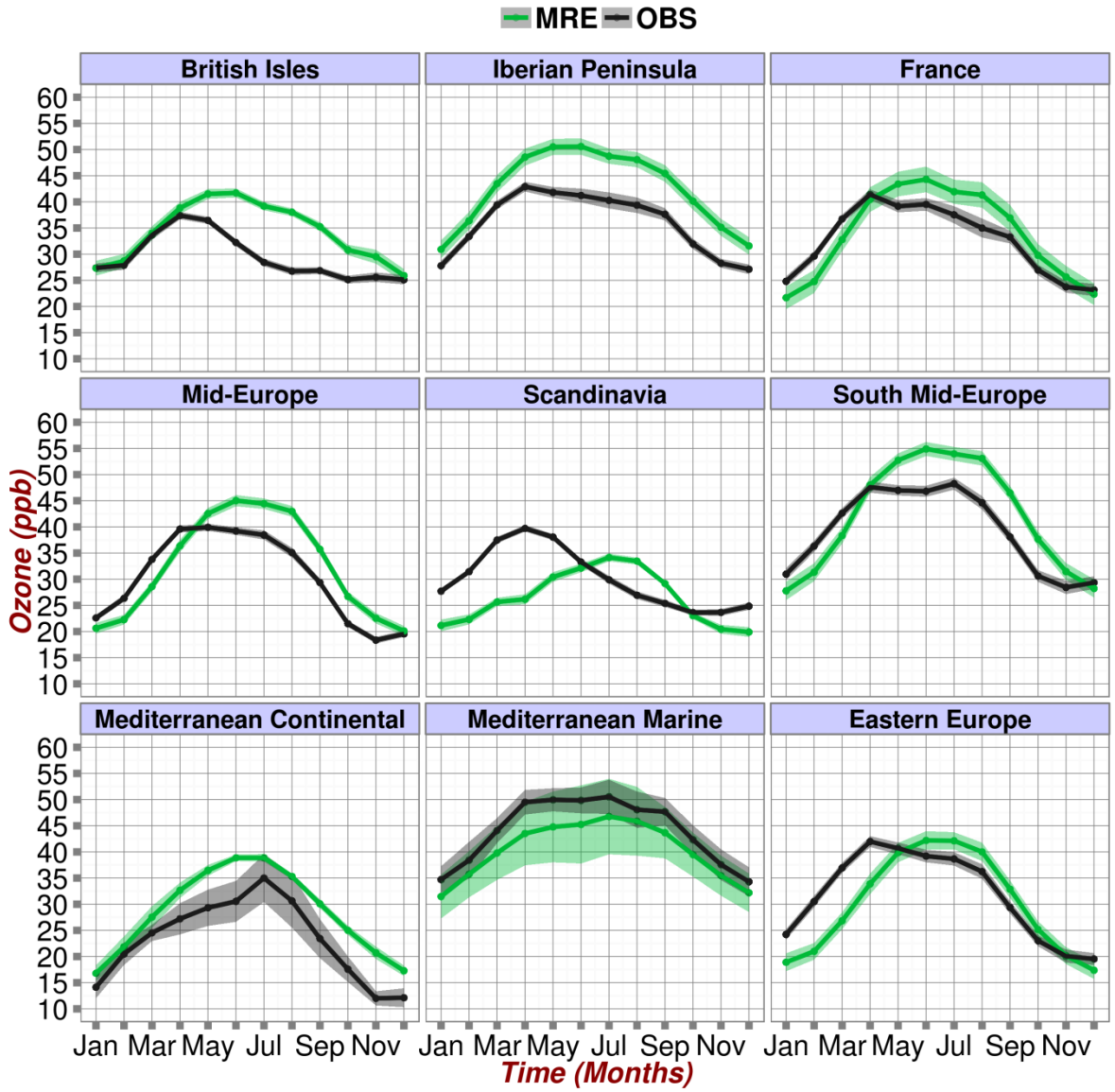


Figure 4. Mean 2003-2012 annual cycle of near surface ozone for the different European subregions of the MACC reanalysis and observations. The shading areas denote 95% confidence interval of the mean values.

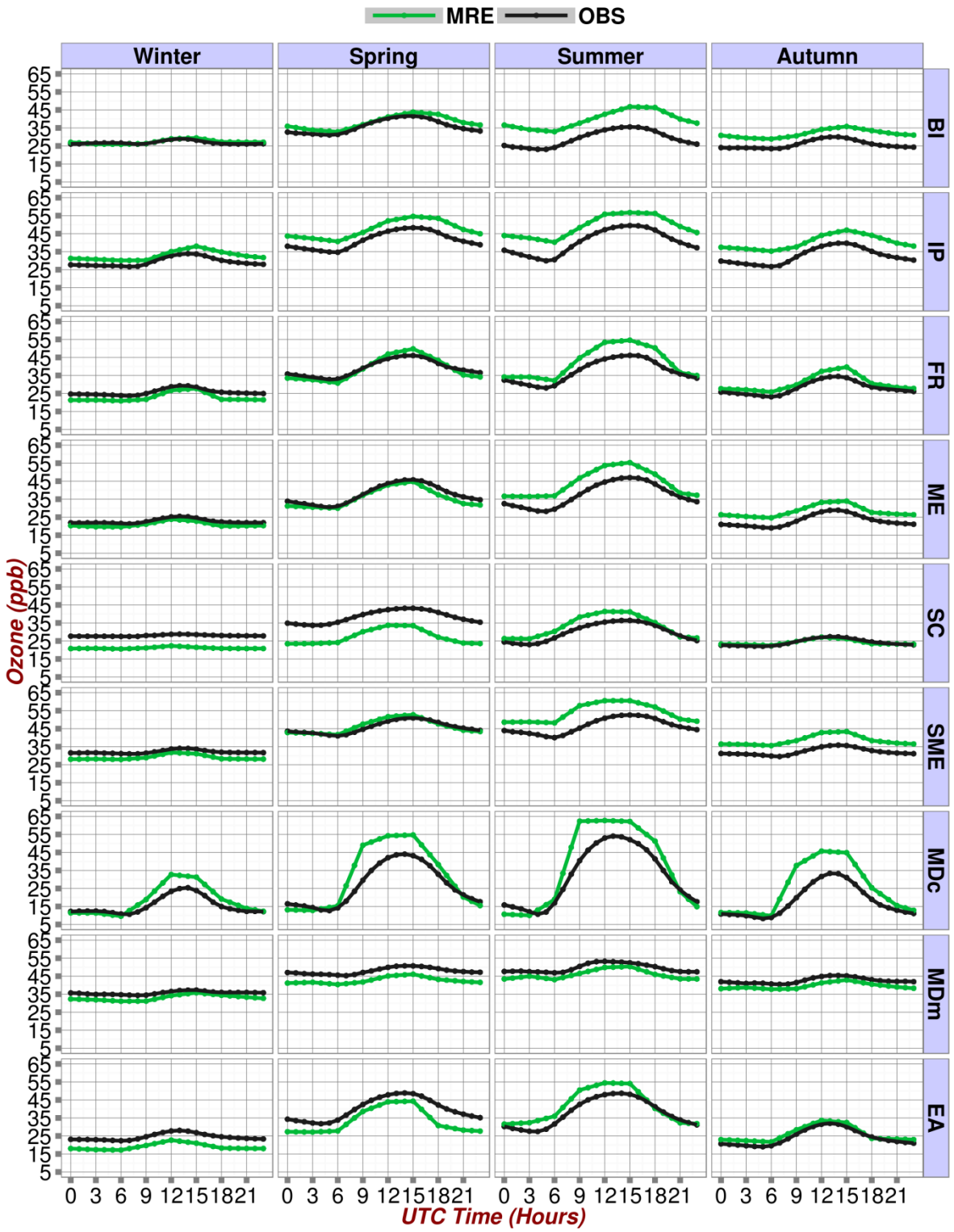


Figure 5. Mean 2003-2012 diurnal cycle of near surface ozone for the different European subregions based on MRE (green line) and observations (black line) calculated for winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

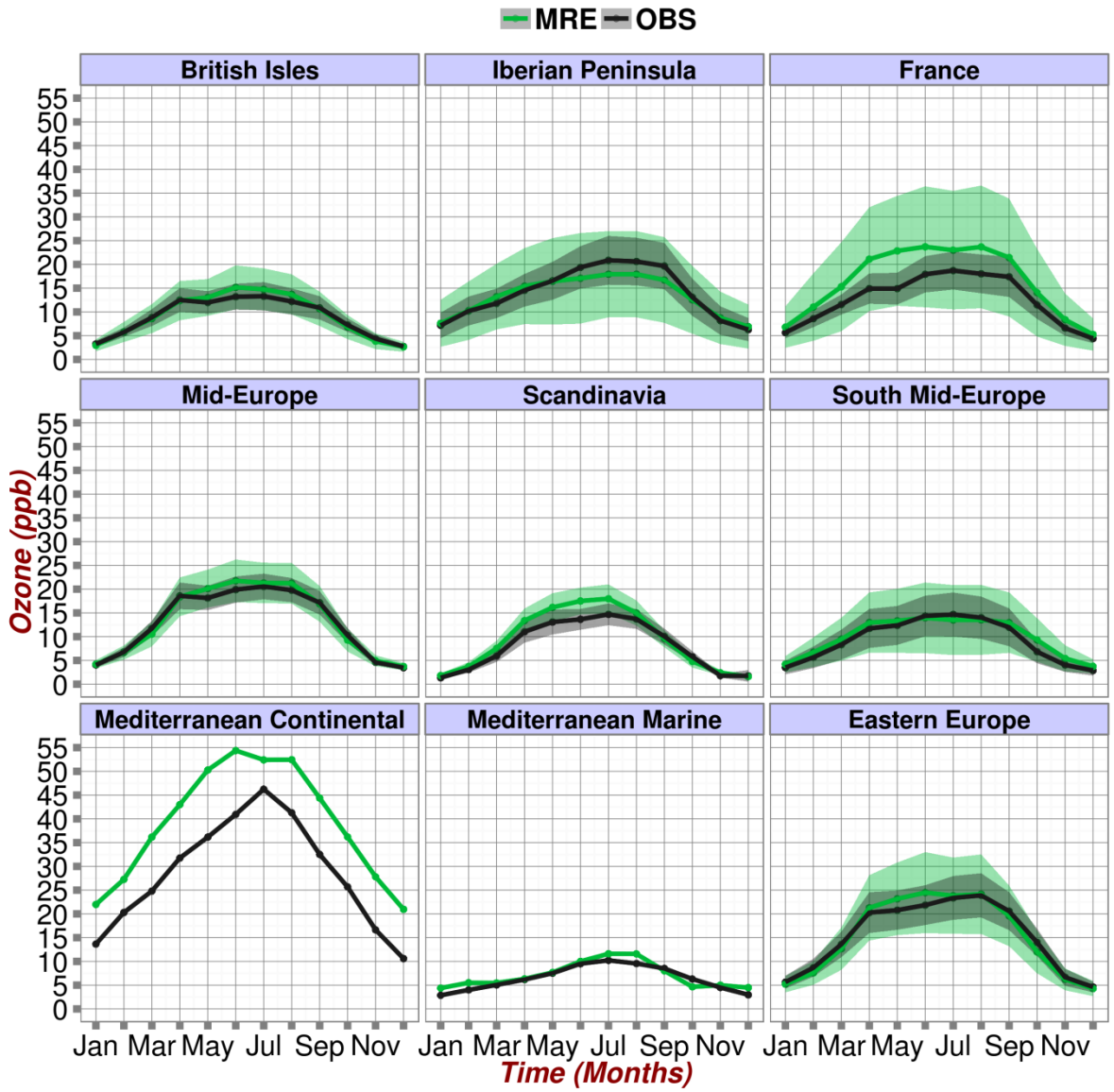
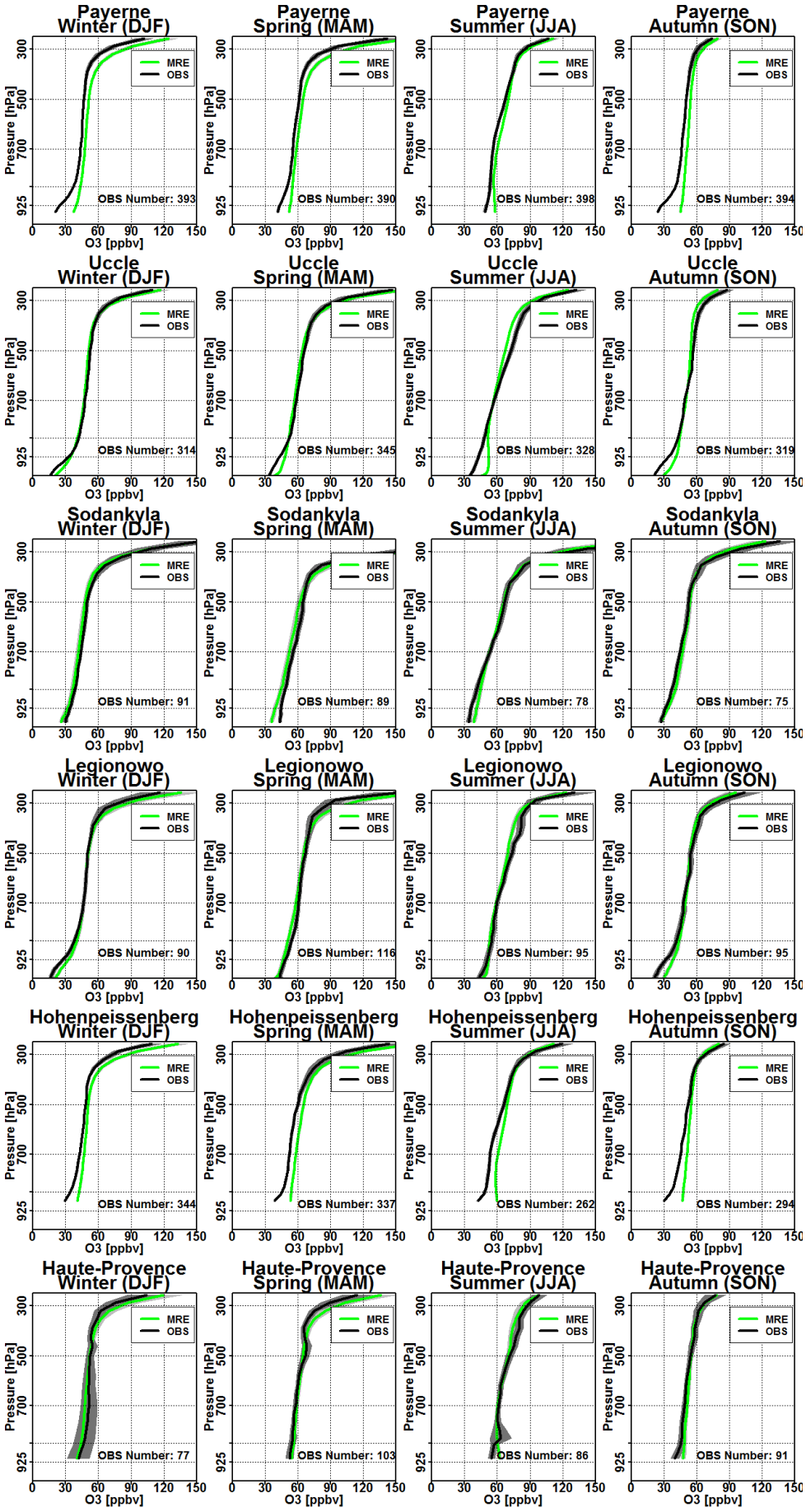


Figure 6. Annual cycle of the diurnal range of near surface ozone for observations (black line) and MRE (green line) averaged over the time period 2003-2012 for the different European subregions. Shading areas denote the 95% confidence interval of the mean values.


Payerne
Winter (DJF)
Payerne
Spring (MAM)
Payerne
Summer (JJA)
Payerne
Autumn (SON)
Uccle
Winter (DJF)
Uccle
Spring (MAM)
Uccle
Summer (JJA)
Uccle
Autumn (SON)
Sodankyla
Winter (DJF)
Sodankyla
Spring (MAM)
Sodankyla
Summer (JJA)
Sodankyla
Autumn (SON)
Legionowo
Winter (DJF)
Legionowo
Spring (MAM)
Legionowo
Summer (JJA)
Legionowo
Autumn (SON)
Hohenpeissenberg
Winter (DJF)
Hohenpeissenberg
Spring (MAM)
Hohenpeissenberg
Summer (JJA)
Hohenpeissenberg
Autumn (SON)
Haute-Provence
Winter (DJF)
Haute-Provence
Spring (MAM)
Haute-Provence
Summer (JJA)
Haute-Provence
Autumn (SON)
MRE
OBS
Pressure [hPa]
O3 [ppbv]
OBS Number: 393
OBS Number: 390
OBS Number: 398
OBS Number: 394
OBS Number: 314
OBS Number: 345
OBS Number: 328
OBS Number: 319
OBS Number: 91
OBS Number: 89
OBS Number: 78
OBS Number: 75
OBS Number: 90
OBS Number: 116
OBS Number: 95
OBS Number: 95
OBS Number: 344
OBS Number: 337
OBS Number: 262
OBS Number: 294
OBS Number: 77
OBS Number: 103
OBS Number: 86
OBS Number: 91

1

Figure 7. Mean 2003-2012 ozone profiles based on MRE near surface ozone (green line) and ozonesonde measurements (black line) at the stations of Sodankyla (67.4N, 26.6E), Legionowo (52.4N, 20.9E), Uccle (50.8N, 4.3E), Hohenpeissenberg (47.8N, 11E), Payerne (46.8N, 6.9E), and Haute-Provence (43.9N, 5.7E). The shading areas denote 95% confidence interval of the mean values.

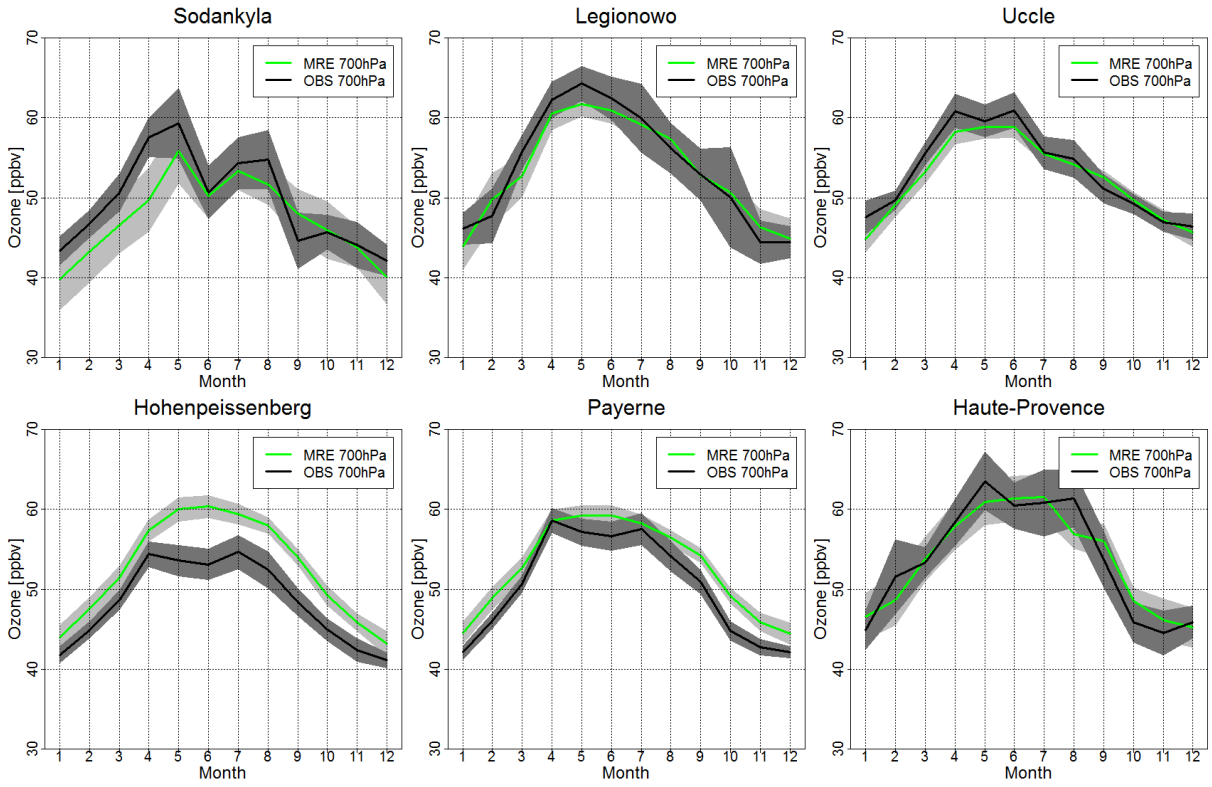


Figure 8. Mean 2003-2012 annual cycle of ~~near surface ozone~~ lower tropospheric ozone at 700 hPa based on MRE (green line) and ozonesonde measurements (black line) at the stations of Sodankyla (67.4N, 26.6E), Legionowo (52.4N, 20.9E), Uccle (50.8N, 4.3E), Hohenpeissenberg (47.8N, 11E), Payerne (46.8N, 6.9E), and Haute-Provence (43.9N, 5.7E). The shading areas denote 95% confidence interval of the mean values.

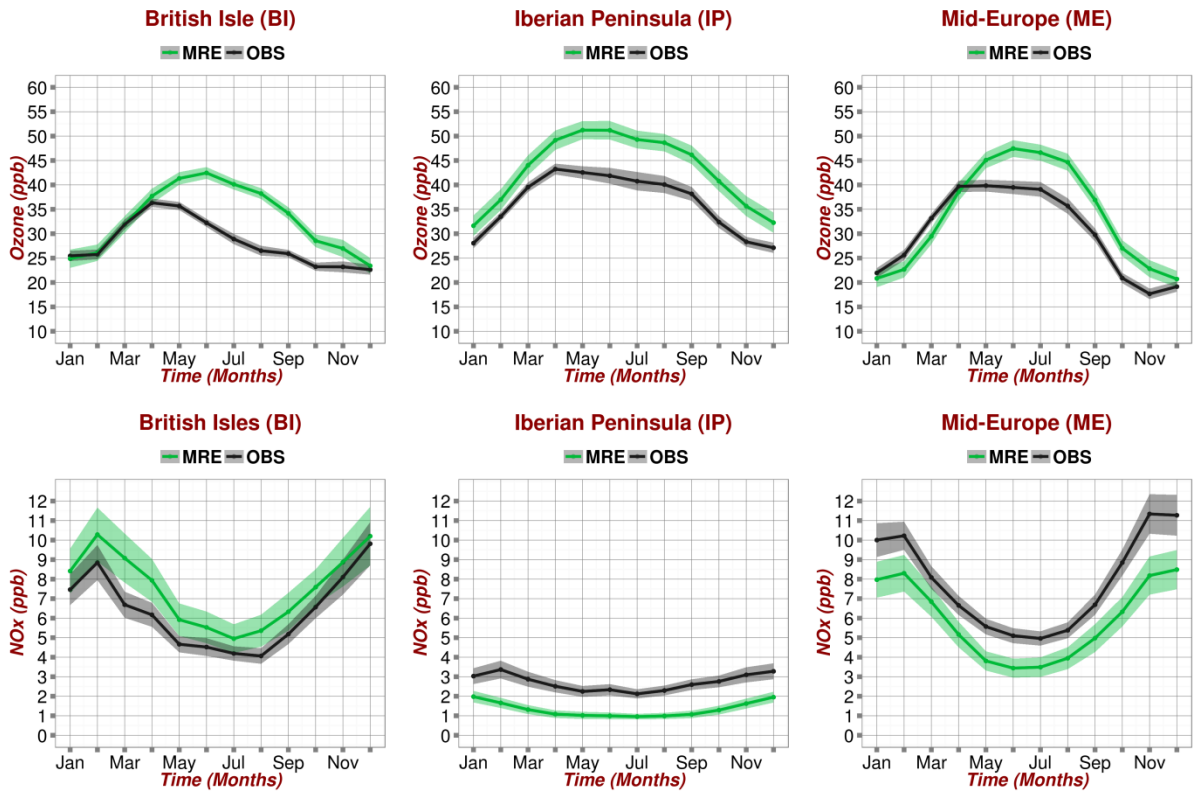


Figure 9. Mean annual cycle of near surface O3 (top panel) and NOx (bottom panel) based on observations (solid black line) and MRE (green line) for the subregions BI, IP, ME over the period 2003-2012.

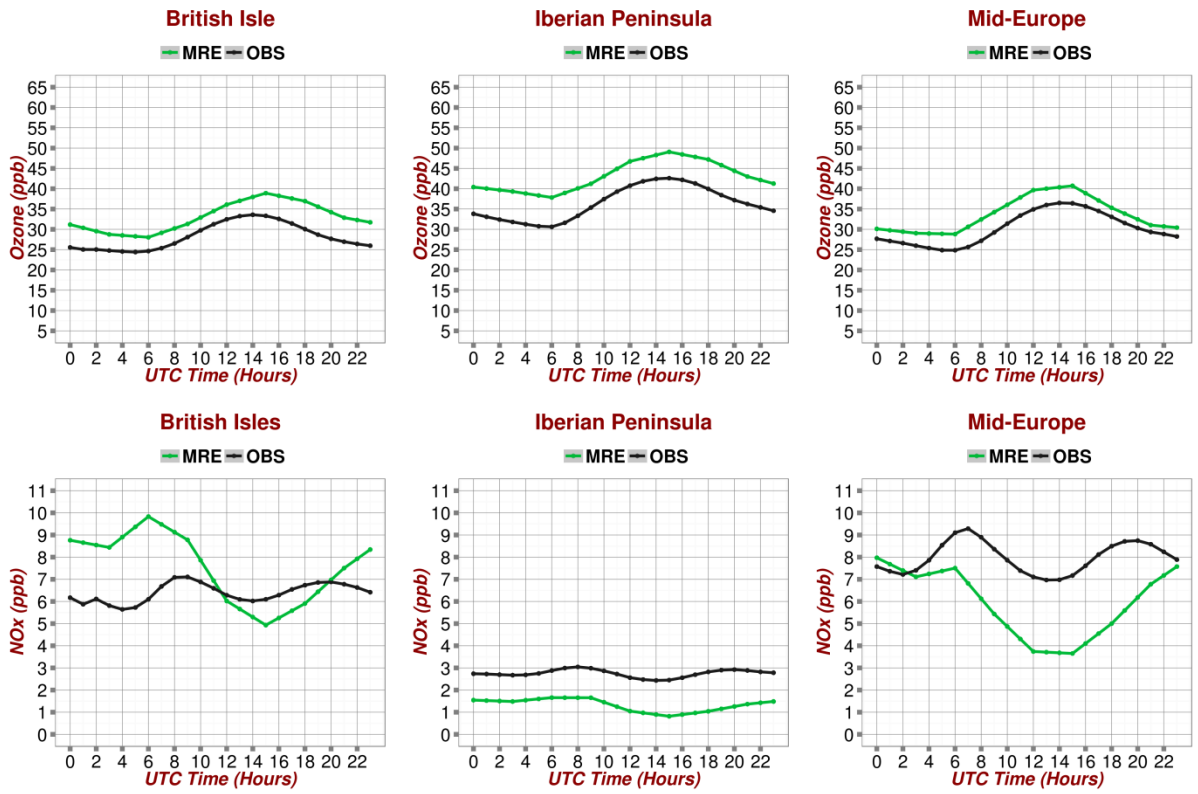


Figure 10. Mean diurnal cycle of near surface O3 (top panel) and NOx (bottom panel) based on observations (solid black line) and MRE (green line) for the subregions BI, IP, ME during summer over the period 2003-2012.

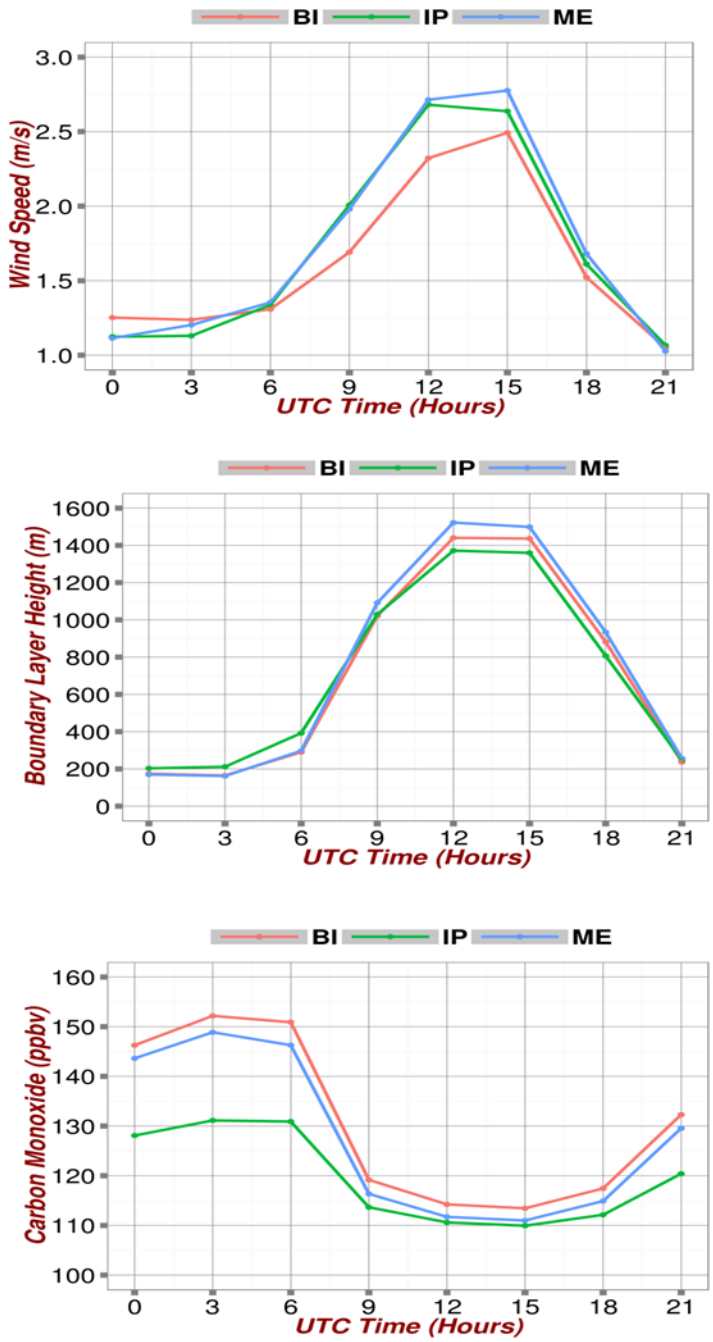


Fig S1 Diurnal meteorological patterns of wind speed (upper panel) and boundary layer height (middle panel) and Carbon monoxide (bottom panel).