General author's comment, responses to referees and marked-up manuscript version

"A multi-model comparison of meteorological drivers of surface ozone over Europe" (acp-2017-787)

Dear Editor and Referees,

We are pleased to see that the manuscript **"*A multi-model comparison of meteorological drivers of surface ozone over Europe*" (acp-2017-787)** has received great attention during the open discussion phase and we truly appreciate the constructive comments and suggestions from all the anonymous referees as well as the short comment and recommendations published by Dr. Carlos Ordóñez.

We are very happy to hear that in general the referees point out the valuable contribution of our manuscript to the air quality community. We also note the concerns from anonymous referee #2, who is less convinced regarding the novelty of the manuscript. In response to all comments, we have carefully revised the manuscript and have made major changes in the manuscript. Thanks to the referees' comments we consider that the paper is now improved, especially regarding the balance of the sections. We hope to have adequately addressed the referees' concerns and that they find the revised version of the manuscript suitable for publication.

In this letter we would like to provide a brief summary of the major changes in the revised version that have addressed similar issues, since some concerns were common between the four reviews of our manuscript. In the following, we list the main changes in the manuscript, which include:

- A substantial rewriting and reorganization of the introduction, removing redundant information and paragraphs not relevant for the paper.
- A deeper review of the existing literature to better clarify the key novelty of our study and highlight its main contribution to the air quality community.
- A more detailed discussion regarding our definition of the seasons. Accordingly new figures in the supplementary material are provided.
- A new detailed table with the main characteristics of the models is added.
- An additional time series analysis regarding the issue in the Balkans region is now added in the supplementary material (SC, Carlos Ordónez).
- References and figures as well as the supplementary material have been accordingly modified.

We believe that all of these changes have lead to a significantly improved version of the paper, which in our opinion can certainly benefit the air quality modelling community.

Finally, we will be happy to clarify any question that might remain open and we would be glad to see the manuscript considered for publication.

Bellow, we provide a point-by-point response to the referees' comments (in ***bold italics***) following by our responses (in blue Font). Furthermore, a marked-up manuscript version showing the changes made is included.

Author's response to Referee #1

Anonymous Referee #1-Received and published: 21 May 2018

General comments

The authors discuss an important question in air quality modeling, i.e. the capability of model results in reproducing several features of ozone time series. To this end, they compare observations with several state-of-the-art models. I think this issue is important and worthy of discussion, but I also think that the paper needs major revisions before it can be accepted for publication. First of all, I suggest to better balance the length of all sections; the introduction is too long but fails to highlight what has already been previously achieved and what are the main advances of this paper. What is really new with this work?

We are very grateful to the referee for the constructive comments and suggestions, which lead to improvement of this manuscript. We have carefully revised the manuscript and made the suggested changes.
The balance in some parts of manuscript, particularly the introduction, is a common concern from the referees. As stated in the general author's response, the revised version of the manuscript has significantly been improved according to the referees' suggestions.

A second major concern is about the use of observations. As I can guess, the authors interpolate observations over a regular grid, but this introduces an additional problem. What is the representativeness of area-averaged observations? Usually, the support of point observations is much more limited than 1x1 grid cells. How the authors address this issue? Why not use a much simpler approach consisting of the comparison between observations and interpolated model values? As I can guess, Airbase observations contain several different station types (e.g. remote, suburban, urban, etc.). How are they treated?

We would like to emphasise that we did not interpolate observations. As stated in the manuscript (see section 2.1) we have used a gridded dataset of MDA8 O3 provided by Schnell et al. (2014). This product was generated with objective mapping algorithm by merging thousands of stations from EMEP and Airbase. This dataset has been originally used by Schnell et al. (2014) to evaluate global air quality models and in Otero et al. (2016) to examine the influence of synoptic and meteorological conditions on MDA8 O3. Recently, Órdoñez et al. (2017) and Carro-Calvo et al. (2017) have been used this product to assess the impacts of high-latitude blocks and subtropical ridge on ozone and to examine a spatial clustering of meteorological drivers over Europe, respectively. Here, we have used this product for the first time to evaluate a set of regional CTMs over Europe.

A third comment concern the use of multiple models. The authors use a suite of model values, but they do not refer to any ensemble. I'm curious to know if an ensemble treatment may help in this case.

The CTMs included in this study did not use the same meteorological input: three of the models (CHIMERE, EMEP and MINNI) used WRF, while LOTOS was driven by

RACMO2 and MACTH by HIRLAM. Since we aim to identify potential deficiencies in the inputs of the CTMs, we consider that the use of an ensemble would not bring more insights to this study. Then, we decide not to include an ensemble mean.

Finally, I also suggest a deeper analysis of the regression model. The authors implicitly assume a homoscedastic behaviour. Is this supported by data? Due to the large interval of values and intrinsic periodicities in time series, I think that the variance cannot be assumed independent on model values. I suggest investigating on data properties, as well as on independence between the beta's values between different models and areas.

As referee#1 suggests, we have analysed the results from the statistical models (CTMs-based and observation-based MLR) during the model development. We follow a similar strategy to that of Otero et al. (2016), which included an analysis of multicollinearity between predictors and a further examination of the residuals. Figures 1 and 2 show the MLR diagnostics regarding the homoscedastic behaviour for both seasons, AMJ and JJA, respectively. In general, we did not observe from these plots a serious indication or pattern of heteroscedasticity. Additionally, we examined the distribution of the standardised residuals. Figures 3, 4 show the histograms of standardised residuals for both seasons AMJ and JJA, respectively. In addition, the normal probability plots (QQ-plots) of standardised residuals are depicted in figures 5 and 6. The diagnostic plots, did not reveal serious evidence that the residuals are not normally distributed. Therefore, we can conclude that generally, the MLRs were properly developed.

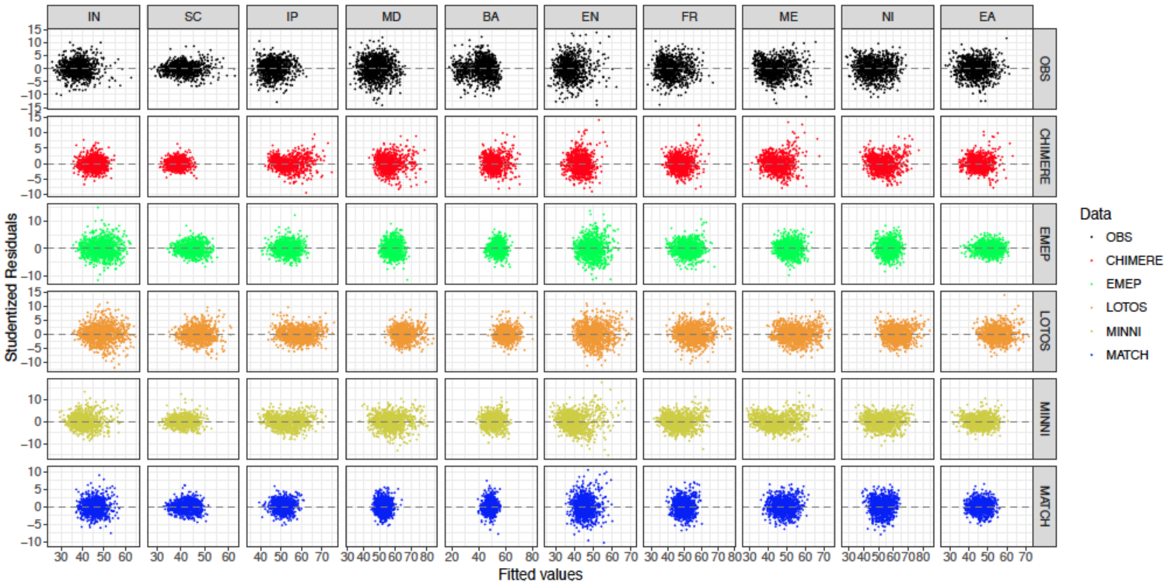


Figure 1. Standardized residuals versus predicted values for the observed-based and CTMs-based MRL in each region during springtime (AMJ).

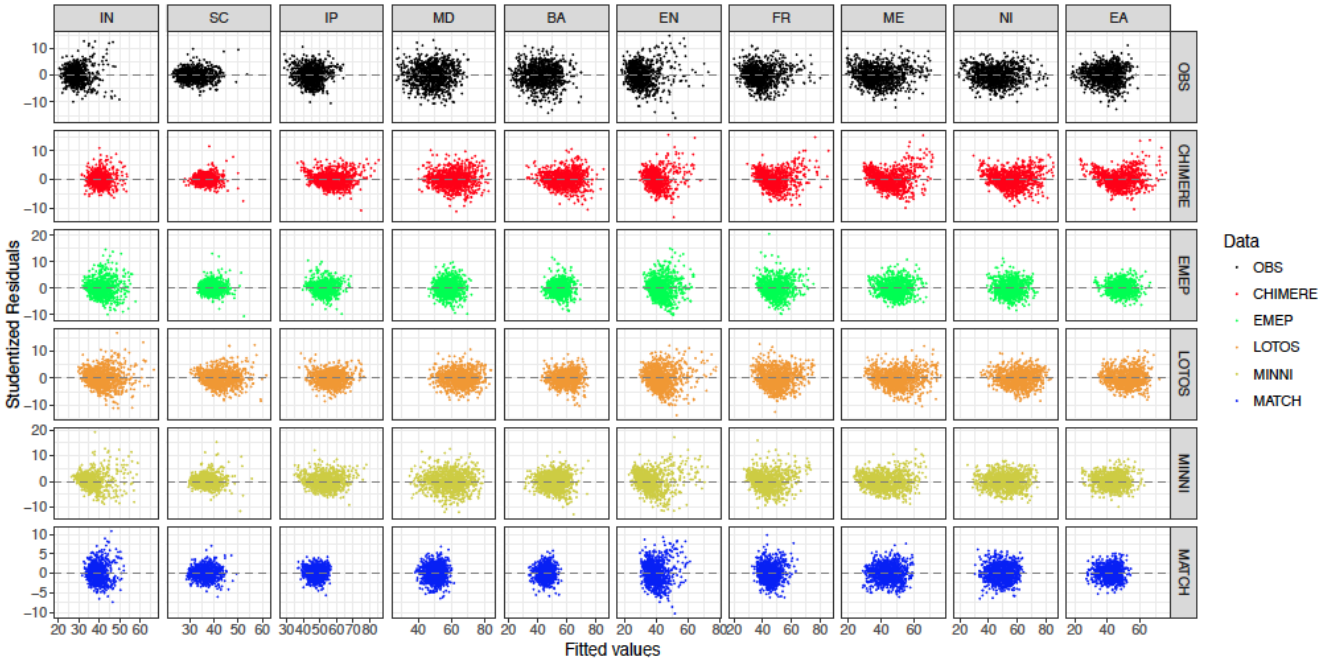


Figure 2. Standardized residuals versus predicted values for the observed-based and CTMs-based MRL in each region during summertime (JAS).

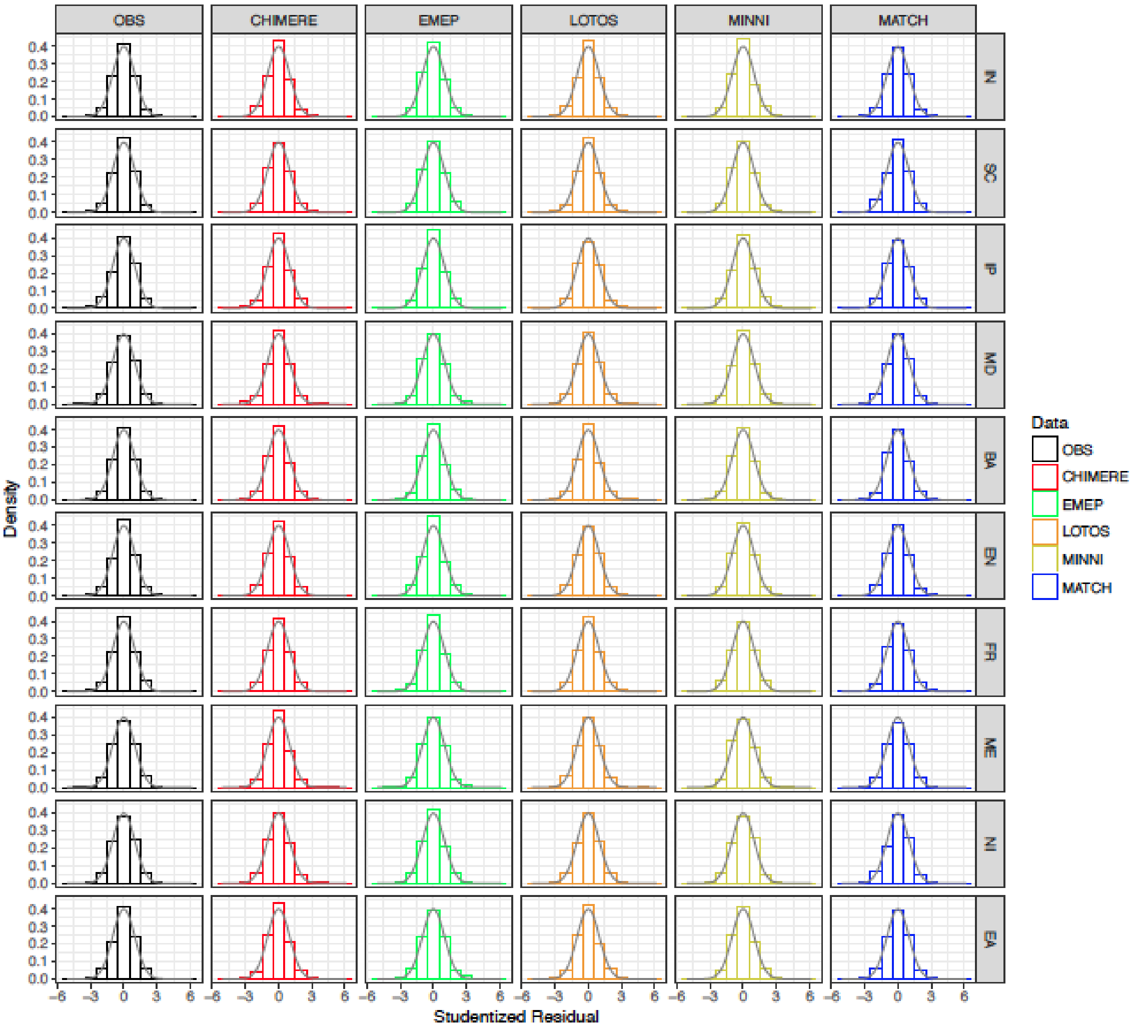


Figure 3. Histograms of standardized residuals for the observed-based and CTMs-based MRL in each region during springtime (AMJ).

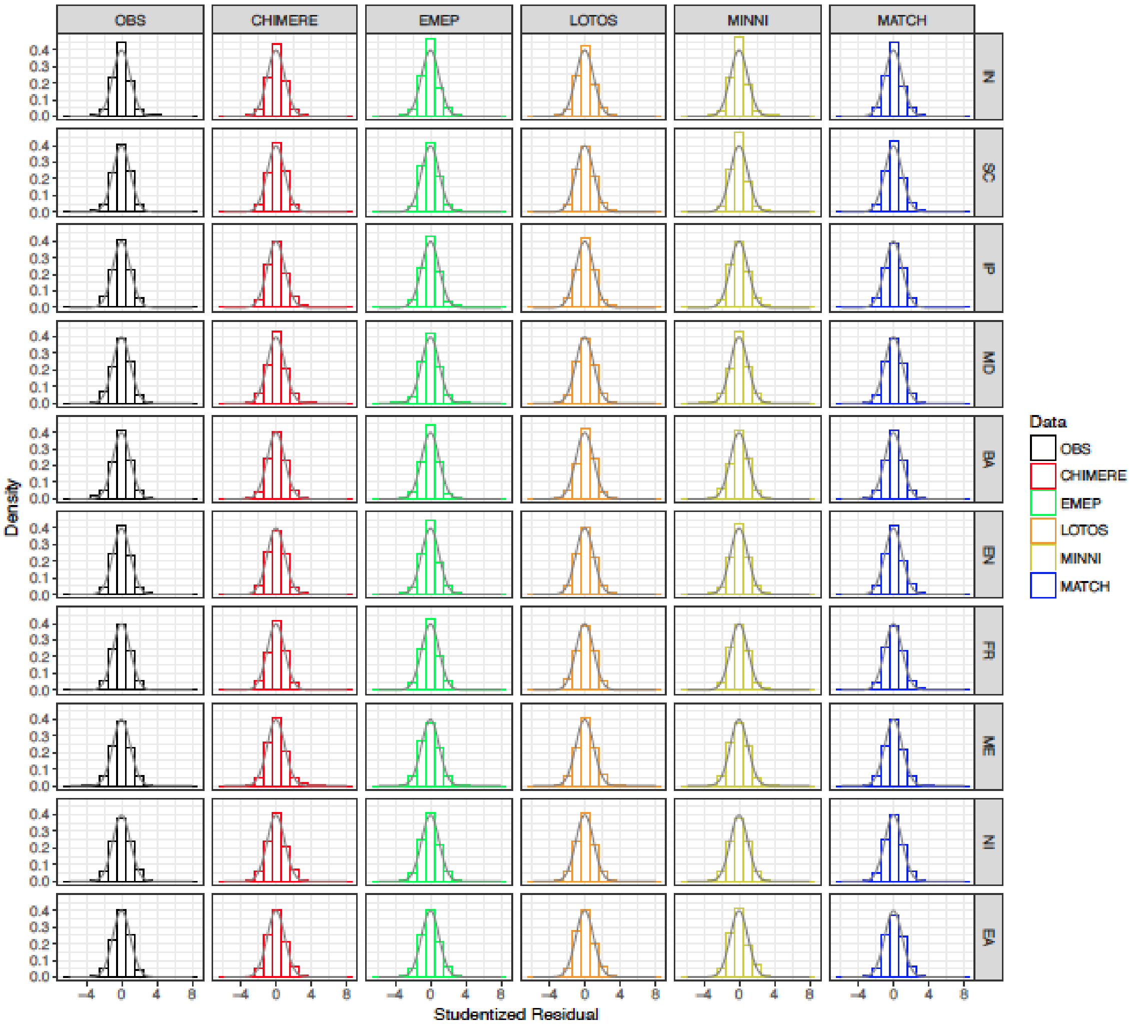


Figure 4. Histograms of standardized residuals for the observed-based and CTMs-based MRL in each region during summertime (JAS).

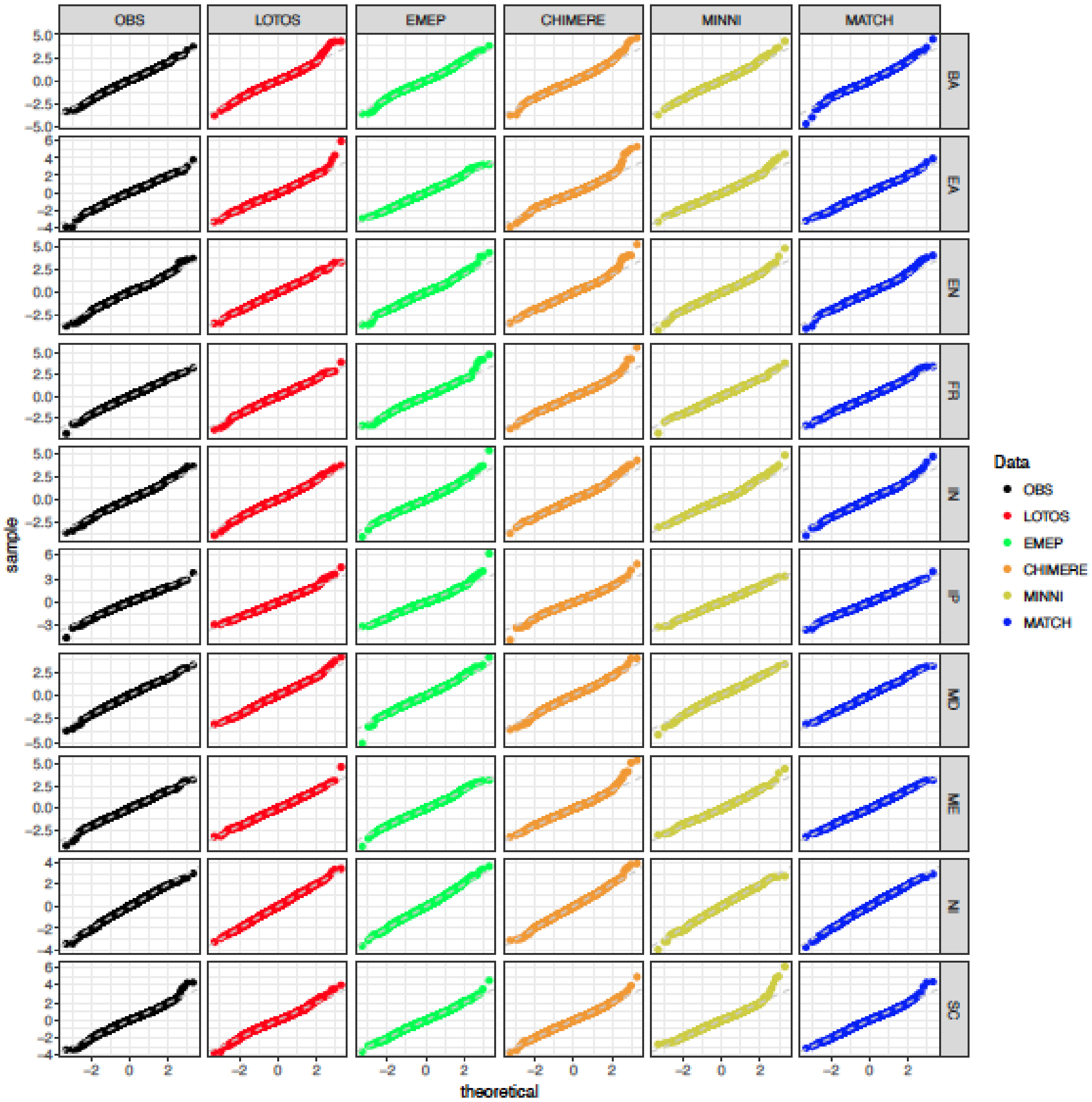


Figure 5. Normal probability plots (QQ-plots) of standardized residuals for the observed-based and CTMs-based MRL in each region during springtime (AMJ).

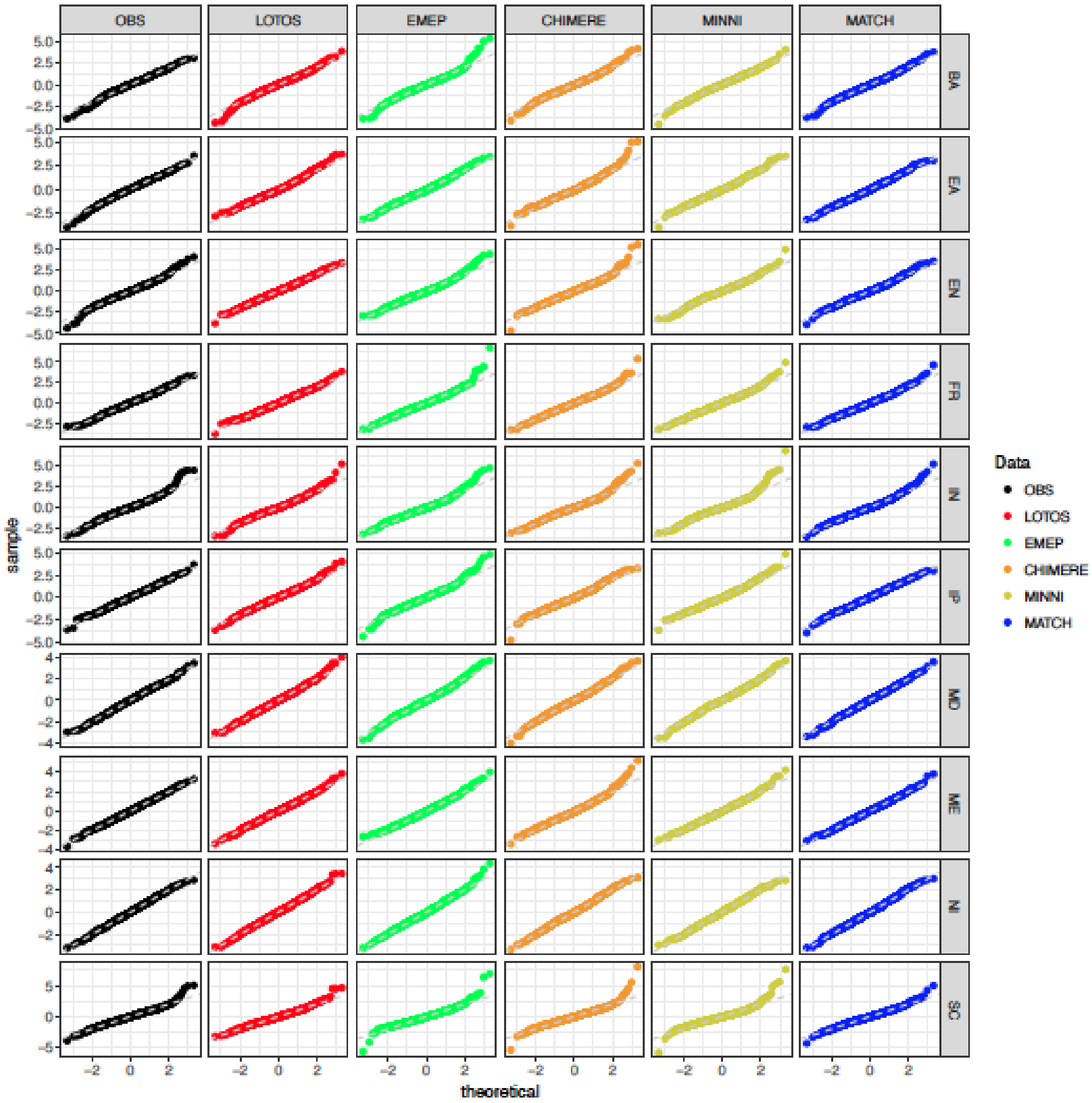


Figure 6. Normal probability plots (QQ-plots) of standardized residuals for the observed-based and CTMs-based MRL in each region during summertime (JAS).

Typos: line 23: "ENE" should be "ENEA" line 161: "van Lon" should be "van Loon"

Thank you. This has been corrected in the revised version.

References

Carro-Calvo, L., C. Ordóñez, R. García-Herrera, J. L. Schnell: Spatial clustering and meteorological drivers of summer ozone in Europe, Atmos. Environ., 167, 496- 510,https://doi.org/10.1016/j.atmosenv.2017.08.050, 2017.

Ordóñez, C., Barriopedro, D., García-Herrera, R., Sousa, P. M., and Schnell, J. L.: Regional responses of surface ozone in Europe to the location of high-latitude blocks and subtropical ridges, Atmos. Chem. Phys., 17, 3111-3131, https://doi.org/10.5194/acp-17-3111-2017, 2017.

Otero, N., Sillmann, J., Schnell, J.L., Rust, H.W., Butler, T., 2016. Synoptic and meteorological drivers of extreme ozone concentrations over Europe. Environ. Res. Lett. 11 (2), 24005. http://dx.doi.org/10.1088/1748-9326/11/2/024005.

Schnell, J. L., Holmes, C. D., Jangam, A., and Prather, M. J.: Skill in forecasting extreme ozone pollution episodes with a global atmospheric chemistry model, Atmos. Chem. Phys., 14, 7721–7739, doi:10.5194/acp-14-7721-2014, 2014.

Author's response to Referee #2

Anonymous Referee #2 ***Received and published: 18 May 2018***

The paper presents an analysis of a suite of chemical transport models applied to simulate ozone levels over Europe. I value very high the community effort of gathering around a common exercise and I am aware of the amount of time required to run, collect, harmonise and analyse model's results. I have to note, however, that the comments I posted for the 'quick' assessment of the paper have not been addressed:
'The methodology presented is sound and entails massive amount of analysis work. Though, I fail to see the fruits of such analysis! I found the paper unbalanced in several parts, with a very long and qualitative introduction and scarcity of quantitative results.
What is the main message of the paper? What the advancement? What is novel (in? The results) with respect to other existing multi model comparison activities? Again, the authors conclude with speculation as to why certain models behave in some way rather than in another?. But where is the quantitative, supporting argument? I would invite the authors to deeply analyse existing results from the literature (e.g. Vautard et al., 2012 Brunner et al, 2015; Makar et al. 2015, among many others) and reconsider their contribution in light of the novelty it might bring.'

We would like to thank referee #2 for the helpful and constructive comments to improve the quality of the manuscript. We have carefully studied the comments and suggestions and we have made a major revision of the original manuscript. We hope that the revisions are acceptable and that our responses adequately address the comments.
Following the referee's comment, we have modified the current version of the manuscript in order to improve the quality and clarify the novelty and contribution of our study.

As the referee points out, in the recent literature several intercomparison and evaluation exercises of chemistry-transport models (CTMs) have contributed to better understand model uncertainties related the parameterization of processes and the quality of input data (e.g. Bessagnet et al. 2016, Solazzo et al. 2017, Vautard et al. 2012). Previous intercomparison studies have evaluated the meteorological input data with different configurations used by air quality models (e.g. Vautard et al., 2012 and references therein). Most of these exercises were designed to cover short-time periods (e.g. one year, Vauvard et al. 2012, Brunner et al. 2015), while only a few numbers of studies included longer periods (e.g. Vautdard et al., 2006, Wilson et al., 2012), but only using one model. As stated in Colette et al., (2017), the design of EURODELTA-Trends (EDT) exercise serves to examine the long-term evolution of air quality and its drivers in Europe. In this context, the EDT experiment provided us a great opportunity to assess long-term air quality performance and the role of meteorological driving factors.

We understand the referee's concern regarding the novelty of this manuscript with respect to the previous cited studies, since we address a common objective by investigating the role of the meteorological input used by CTM. However, we believe that our study contributes to the existing multi-model comparison exercises in several unique ways:

- We have used an interpolated product (Schnell et al., 2015) that offers a good opportunity to directly compare model outputs, rather than using a reduced number of stations within a smaller spatial coverage.

- Our study presents a multi-model evaluation over the whole Europe analysing the model performance in 10 different subregions, while most of the cited studies focused on a smaller number of European regions.

- A common approach to evaluate the performance of the meteorological inputs that drive air quality models consists of comparing directly with observational datasets through standard statistical metrics (e.g. bias, correlations, RMSE, among others). However, our primary goal was not to evaluate the models' skill to reproduce observed parameters, but rather we aimed to provide an alternative method to examine whether the current air quality models are capable to reproduce the observed meteorological sensitivities of ozone that have been reported in a wide number of statistical modelling studies. Thus, systematic differences in the ozone response to the most important meteorological factors can provide a valuable diagnostic tool to identify potential deficiencies in the inputs and parameters of air quality models. Only a couple of studies have addressed this issue but only using one model (e.g. Davis et al. 2011, Fix et al. 2017). To our knowledge this is first study that presents a multi-model evaluation of ozone sensitivities to the main meteorological key drivers over Europe.

- One of the main outcomes from this study points out the limitations in the CTMs to reproduce observed relationships between ozone and some meteorological key drivers. In particular we found that CTMs underestimate the strength of ozone-relative humidity relationship, which might be related to the dry deposition schemes (as suggested in previous works, e.g. Kavassalis and Murphy, 2017). Due to the strict requirements of the EDT exercise in terms of input data, most of the differences in the model outputs can be attributed to the model formulation and set-up. Furthermore, it is beyond of the scope of this study to fully diagnose the issues within each model, which would require further sensitivity simulations and model set-up.

After re-reading the paper I still fail to see what is the key novelty and scientific advancement brought by this paper. All findings (sensitivity to season, regions, etc..) are documented by dozens of papers. The authors apply a, perhaps, different methodology that converges, nonetheless, at conclusions similar to those already known since several years. In my view the paper, in its current shape, is too qualitative and lacks a clear message that stands out and justifies a publication. In light of the poorly exposed scientific significance I advise the editor to ask the authors to significantly review the paper before it can be considered suitable for publication.

We appreciate the referee's comment and we agree that the manuscript needed to clarify the novelty and the contribution to the existing literature. We have carefully addressed these suggestions that are now reflected in the revised version of the manuscript. As the other referees pointed out, we do believe that our study represents a valuable contribution to the ACP community.

References

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Fi- nardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J.-C., Fagerli, H., Gonzalez, L., Menut, L., Prévôt, A. S. H., Roberts, P., and White, L.: Presentation of the EURODELTA III intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmos. Chem. Phys., 16, 12667–12701, doi:10.5194/acp-16-12667-2016, 2016.

Brunner, D., Jorba, O., Savage, N., Eder, B., Makar, P., Giordano, L., Badia, A., Balzarini, A., Baro, R., Bianconi, R., Chemel, C., Forkel, R., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Im, U., Knote, C., Kuenen, J. J. P., Makar, P. A., MandersGroot, A., Neal, L., Perez, J. L., Pirovano, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Werhahn, K., Wolke, R., van Meijgaard, E., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2, Atmos. Environ., 115, 470–498, 2015.

Colette, A., Andersson, C., Manders, A., Mar, K., Mircea, M., Pay, M.-T., Raffort, V., Tsyro, S., Cuvelier, C., Adani, M., Bessagnet, B., Bergström, R., Briganti, G., Butler, T., Cappelletti, A., Couvidat, F., D'Isidoro, M., Doumbia, T., Fagerli, H., Granier, C., Heyes, C., Klimont, Z., Ojha, N., Otero, N., Schaap, M., Sindelarova, K., Stegehuis, A. I., Roustan, Y., Vautard, R., van Meijgaard, E., Vivanco, M. G., and Wind, P.: EURODELTA-Trends, a multi-model experiment of air quality hindcast in Europe over 1990–2010, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2016-309, accepted, 2017a.

Davis, J., Cox, W., Reff, A., Dolwick, P.: A comparison of Cmaq-based and observation-based statistical models relating ozone to meteorological parameters. Atmospheric Environment 45, 3481e3487. http://dx.doi.org/10.1016/ J.Atmosenv.2010.12.060, 2011.
Fix, M. J., D. Cooley, A. Hodzic, E. Gilleland, B. T. Russell, W. C. Porter, and G. G. Pfister, 2018: Observed and predicted sensitivities of extreme surface ozone to meteorological drivers in three US cities. Atmospheric Environment, 176, 292-300, doi:10.1016/j.atmosenv.2017.12.036.
Kavassalis, S. C., and J. G. Murphy: Understanding ozone-meteorology correlations: A role for dry deposition, Geophys. Res. Lett., 44, 2922–2931, doi:10.1002/2016GL071791, 2017.

Makar, P.A., Gong,W., Hogrefe, C., Zhang, Y., Curci, G., Zakbar, R., Milbrandt, J., Im, U., Galmarini, S., Gravel, S., Zhang, J., Hou, A., Pabla, B., Cheung, P., Bianconi, R., 2015b. Feedbacks between Air Pollution and Weather, Part 1: Effects on Weather. Atmos. Environ. 115, 442e469.

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Cameron-Smith, P., Bergmann, D., Zeng, G., Plummer, D. A., Sudo, K., Nagashima, T., Shindell, D. T., Faluvegi, G., and Strode, S. A.: Use of North American and European air quality networks to evaluate global chemistry–climate modeling of surface ozone, Atmos. Chem. Phys., 15, 10581-10596, doi:10.5194/acp-15-10581-2015, 2015.

Vautard, R., Moran, M.D., Solazzo, E., Gilliam, R.C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A.B., Jericevic, A., Prank, M., Segers, A., Silver, J.D.,Werhahn, J.,Wolke, R., Rao, S.T., Galmarini, S., 2012. Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations. Atmos. Environ. 53, 15e37.

Wilson, R. C., Fleming, Z. L., Monks, P. S., Clain, G., Henne, S., Konovalov, I. B., Szopa, S., and 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, 437–54, https://doi.org/10.5194/acp-12- 437-2012, 2012.

Author's response to Referee #3

Anonymous Referee #3   Received and published: 15 May 2018

GENERAL REMARKS

A strong connection exists between air quality - in particular the surface ozone concentration and accompanying meteorological conditions; hot, sunny, stable conditions favour formation of ozone while turbulent and cloudy conditions are associated with low ozone concentrations. It is crucial that air quality models correctly represent this connection. This paper uses data from observations and a number of state-of-science air quality models to investigate this issue. Using simple multiple linear regression models the relationship between ozone and a number of key meteorological quantities is analysed and compared to analogous MLR based on observations. The authors find that model performance varies with variable and geographic region analysed. Model performance with respect to temperature is commonly good while more limited performance is found in case of the ozone-relative humidity relationship for all models. It is concluded that model resolution, boundary conditions and the parameterization of ozone dry deposition can have an important impact on model performance in addition to the meteorological variables under investigation. This paper addresses an important question in air quality modelling. Obviously, if models fail to adequately represent the relationship that exists between meteorology and atmospheric chemistry/composition then air quality assessments and mitigation strategies based on those will be flawed. The use of multi-model datasets and the relatively simple approach of MLR seem appropriate and serve their purpose. Tables and Figures are used appropriately throughout the text and support well the findings. The conclusions drawn at the end are somewhat sparse and limited but interesting and useful. The only real discrepancy that exists in this study is the definition of the spring and summer seasons. I understand that shifting these back by one month may have benefits with respect to the ozone chemistry in the models but springtime is springtime for a good reason. The meteorological variability in springtime (March, April, May) has a profound impact on atmospheric chemistry and so has the comparative stability with its hot, dry and sunny conditions in summer (in general). I am not sure it is such a good idea to give up on the definition of the seasons but I am not going to make this a make-or-break condition for the paper to be published because I can foresee and endless discussion on the pros and cons with potentially little impact on the study at hand.
In my opinion the paper represents an important contribution to understanding model performance and potential discrepancies. However, I do not consider it a scientific milestone. Overall, I have read this manuscript with interest. If some minor issues and typing errors are corrected I believe the paper can be published.

Firstly, we would like to thank the referee for acknowledging the contribution of the manuscript and the careful reading of the manuscript.

We understand the referee's concern regarding the use of two seasons that differ from the meteorological seasons. As stated in the manuscript, we have selected two three month-periods, namely April-May-June (AMJ) and July-August-September (JAS) in order to examine separately the role of the meteorological parameters during the early and late parts of the European "ozone season", which typically lasts from April to September.

Following the recommendation of Referee #3, we have performed a sensitivity test using the meteorological seasons (March-April-May, MAM and June-July-August, JJA) to address this concern. The main results obtained from this analysis are shown in Figs. 1-5, and summarised below. Furthermore, we have stressed the choice of the seasons in the revised version of the manuscript including Figs. 4 and 5 in the Supplement.

Figure 1 illustrates the seasonal cycle of daily averages of MDA8 O3 for each month over the whole period of study. As depicted in Figure 1, the high ozone mixing ratios are typically observed between April-September, beginning midway through the meteorological spring season (March-April-May), and extending beyond the meteorological summer season (June-July-August). Then, we consider that our choice of 3-month periods covering the whole ozone season is particularly interesting to examine the impact of individual meteorological parameters when ozone shows high levels.

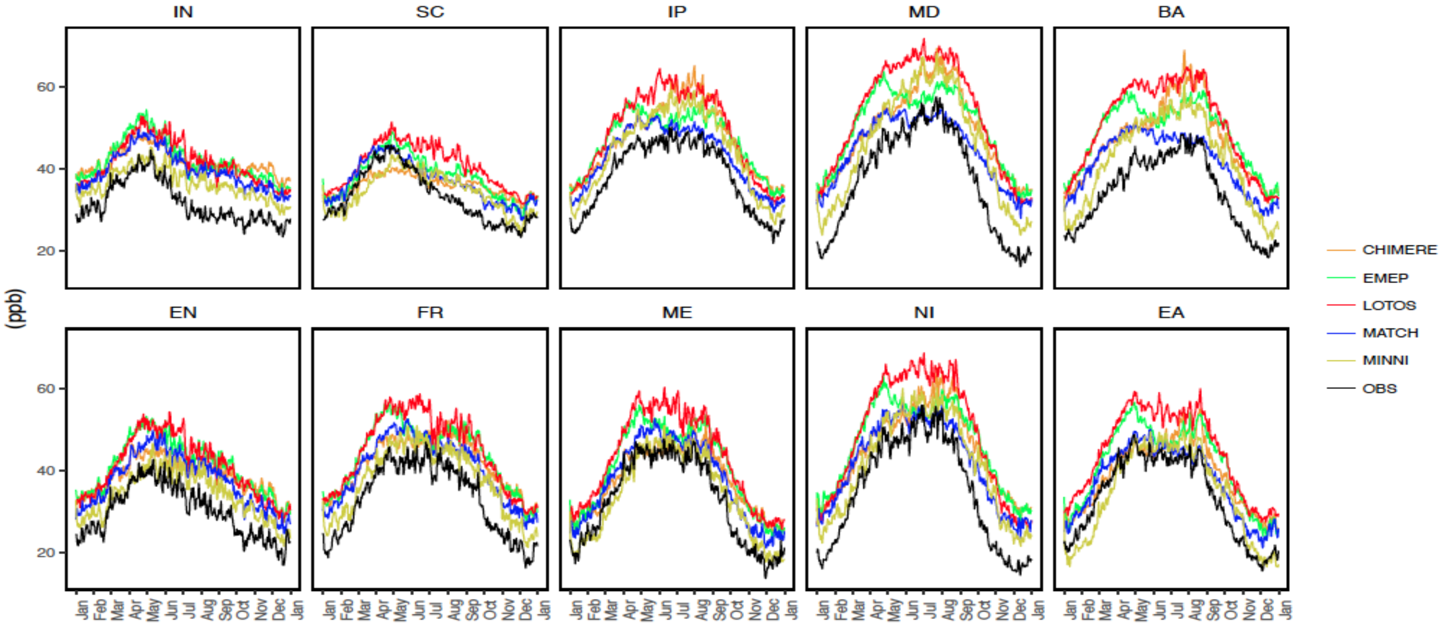


Figure1. Time series of daily averages of MDA8 O3 during the all period of study (2000-2010) at each subregion.

Figures 2 and 3 show the relative importance of individual predictors in the MLRs developed at the internal and external regions for both meteorological seasons (MAM, JJA) respectively. As shown in the manuscript (see Figures 6 and 7) the influence of the meteorological factors is stronger in the internal regions than in the external regions. When comparing the individual drivers' contribution in each meteorological season (.g. top, MAM and bottom, JJA figures 2, 3) we observe relatively small differences in the influence of meteorological parameters on ozone variability between internal and external regions.

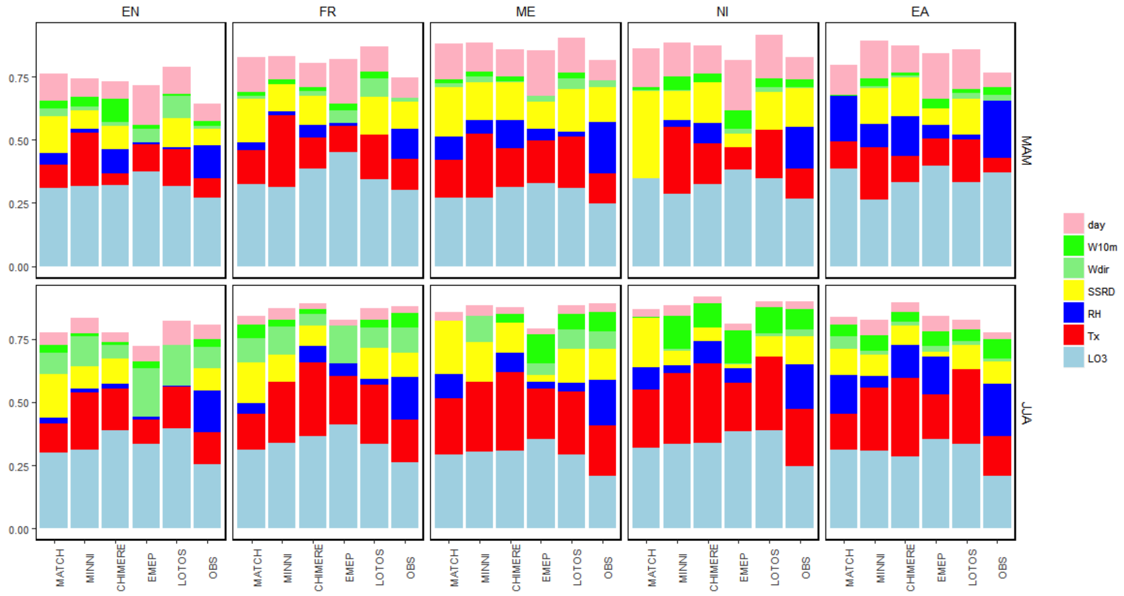


Figure2. Contribution of each predictor to the total explained variance for each CTM-based and observation for the internal regions: England (EN), France (FR), Mid-Europe (ME), North Italy (NI) and East-Europe (EA).

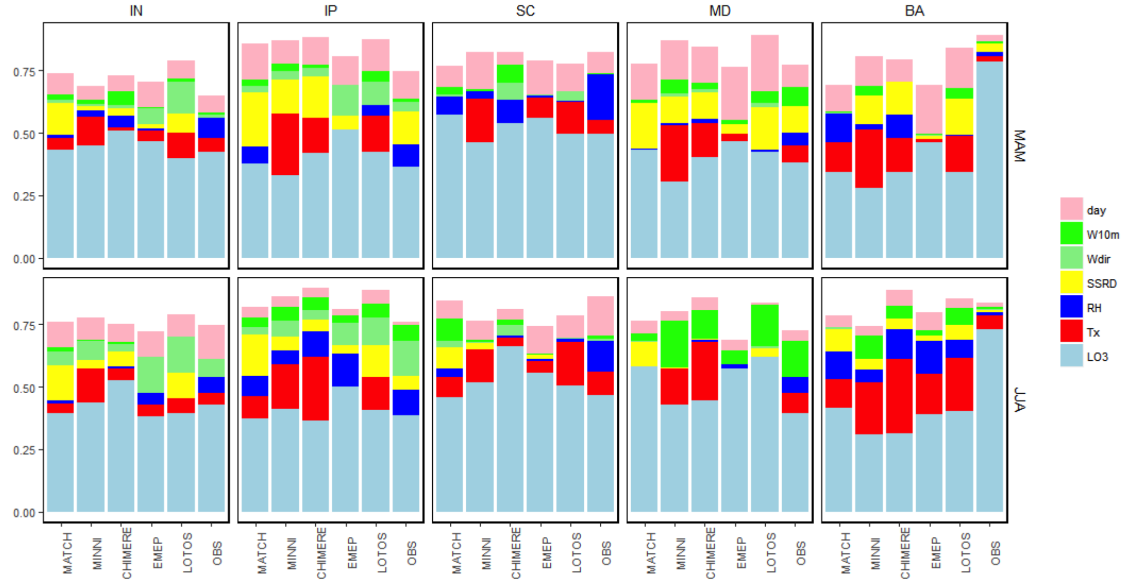
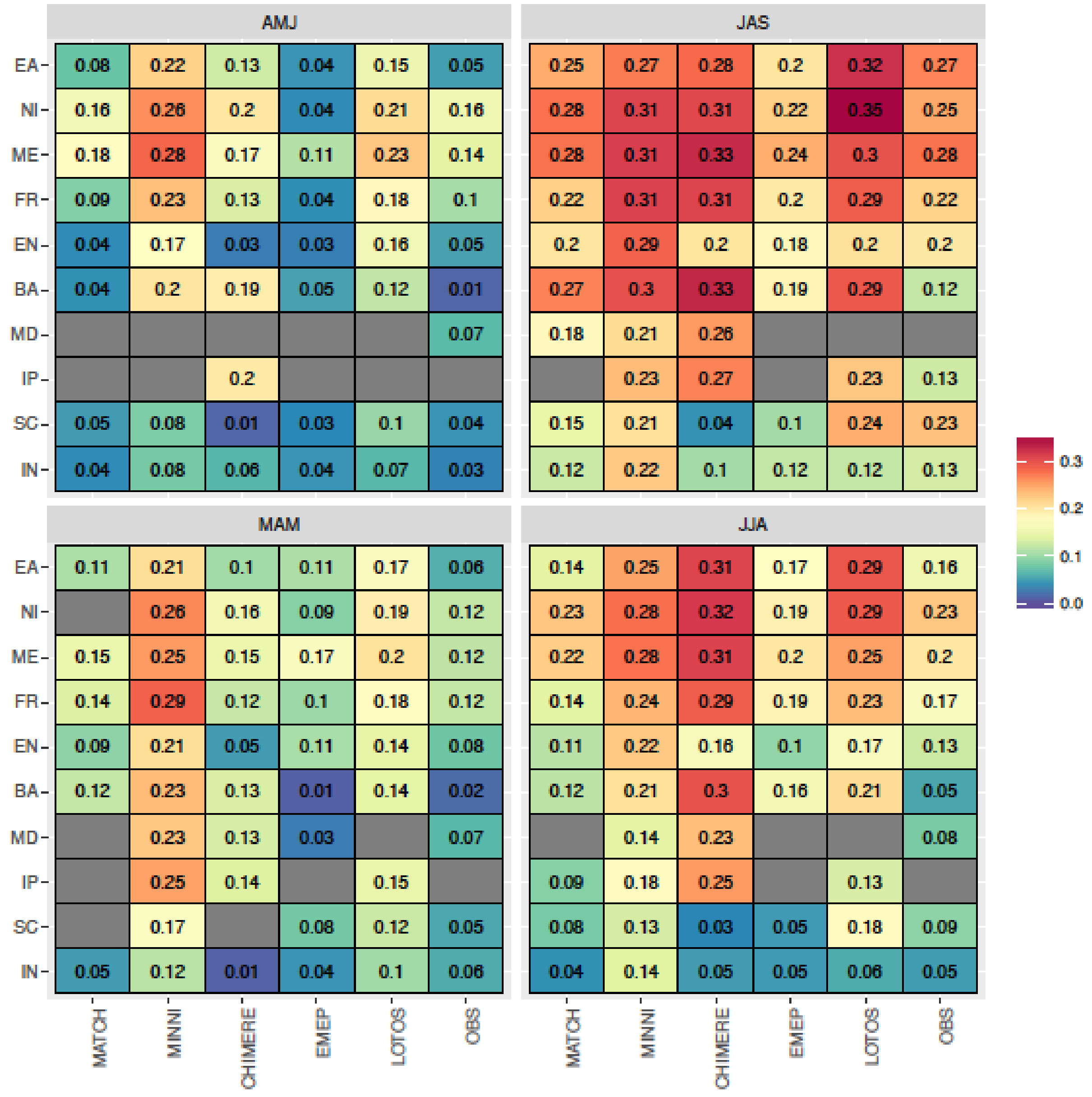


Figure3. Contribution of each predictor to the total explained variance for each CTM-based and observation for the external regions: Inflow (IN), Iberian Peninsula (IP), Scandinavia (SC), Mediterranean (ME) and Balkans (BA).

Similarly, as described in the manuscript we also found notable differences between the re-defined seasons (AMJ, JAS) in most of the regions. In particular in the case of maximum temperature (Tx) and relative humidity (RH), two key-driving factors, it can be shown the different effects, in terms of their individual contribution to the total explained variance, when using both season definitions (MAM, JAS and AMJ, JAS). Figures 4 and 5 show the values of the relative importance of Tx and RH respectively for the seasons AMJ, JAS (top of figures 4,5) and MAM, JJA (bottom of figures 3,4). In general, we notice a stronger influence of maximum temperature during the warmer months, referred to as JAS, than during the meteorological season, JJA (Figure 4). We also observe a larger contribution of relative humidity in AMJ in most of regions, especially in the internal regions (e.g. EA, FR, NI, EN) (figure 5). Therefore, our choice of the seasons allow us to better understand the different meteorological impacts on ozone variability when considering the early and late parts of the European ozone season (April-September). Furthermore, we consider that the contrast of drivers' contribution between seasons and regions is an interesting contribution of this study.



AMJ

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.08	0.22	0.13	0.04	0.15	0.05
NI	0.16	0.26	0.2	0.04	0.21	0.16
ME	0.18	0.28	0.17	0.11	0.23	0.14
FR	0.09	0.23	0.13	0.04	0.18	0.1
EN	0.04	0.17	0.03	0.03	0.16	0.05
BA	0.04	0.2	0.19	0.05	0.12	0.01
MD						0.07
IP			0.2			
SC	0.05	0.08	0.01	0.03	0.1	0.04
IN	0.04	0.08	0.06	0.04	0.07	0.03

JAS

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.25	0.27	0.28	0.2	0.32	0.27
NI	0.28	0.31	0.31	0.22	0.35	0.25
ME	0.28	0.31	0.33	0.24	0.3	0.28
FR	0.22	0.31	0.31	0.2	0.29	0.22
EN	0.2	0.29	0.2	0.18	0.2	0.2
BA	0.27	0.3	0.33	0.19	0.29	0.12
MD	0.18	0.21	0.26			
IP		0.23	0.27		0.23	0.13
SC	0.15	0.21	0.04	0.1	0.24	0.23
IN	0.12	0.22	0.1	0.12	0.12	0.13

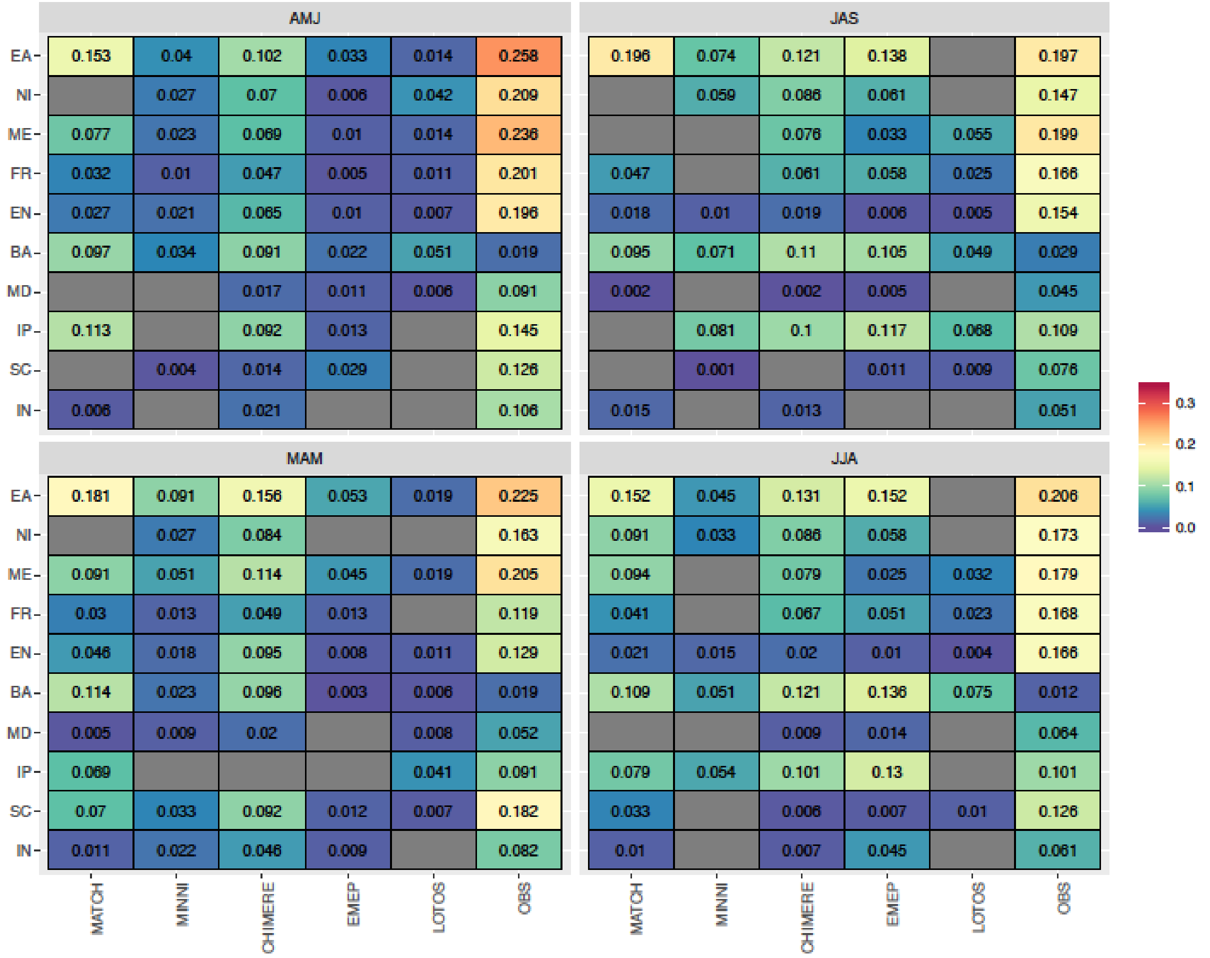
MAM

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.11	0.21	0.1	0.11	0.17	0.06
NI		0.26	0.16	0.09	0.19	0.12
ME	0.15	0.25	0.15	0.17	0.2	0.12
FR	0.14	0.29	0.12	0.1	0.18	0.12
EN	0.09	0.21	0.05	0.11	0.14	0.08
BA	0.12	0.23	0.13	0.01	0.14	0.02
MD		0.23	0.13	0.03		0.07
IP		0.25	0.14		0.15	
SC		0.17		0.08	0.12	0.05
IN	0.05	0.12	0.01	0.04	0.1	0.06

JJA

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.14	0.25	0.31	0.17	0.29	0.16
NI	0.23	0.28	0.32	0.19	0.29	0.23
ME	0.22	0.28	0.31	0.2	0.25	0.2
FR	0.14	0.24	0.29	0.19	0.23	0.17
EN	0.11	0.22	0.16	0.1	0.17	0.13
BA	0.12	0.21	0.3	0.16	0.21	0.05
MD		0.14	0.23			0.08
IP	0.09	0.18	0.25		0.13	
SC	0.08	0.13	0.03	0.05	0.18	0.09
IN	0.04	0.14	0.05	0.05	0.06	0.05

Figure 4. Values of relative importance (contribution to the total explained variance) of maximum temperature in the seasons AMJ, JAS and MAM and JJA.



AMJ

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.153	0.04	0.102	0.033	0.014	0.258
NI		0.027	0.07	0.006	0.042	0.209
ME	0.077	0.023	0.069	0.01	0.014	0.236
FR	0.032	0.01	0.047	0.005	0.011	0.201
EN	0.027	0.021	0.065	0.01	0.007	0.196
BA	0.097	0.034	0.091	0.022	0.051	0.019
MD			0.017	0.011	0.006	0.091
IP	0.113		0.092	0.013		0.145
SC		0.004	0.014	0.029		0.126
IN	0.006		0.021			0.106

JAS

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.196	0.074	0.121	0.138		0.197
NI		0.059	0.086	0.061		0.147
ME			0.076	0.033	0.055	0.199
FR	0.047		0.061	0.058	0.025	0.166
EN	0.018	0.01	0.019	0.006	0.005	0.154
BA	0.095	0.071	0.11	0.105	0.049	0.029
MD	0.002		0.002	0.005		0.045
IP		0.081	0.1	0.117	0.068	0.109
SC		0.001		0.011	0.009	0.076
IN	0.015		0.013			0.051

MAM

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.181	0.091	0.156	0.053	0.019	0.225
NI		0.027	0.084			0.163
ME	0.091	0.051	0.114	0.045	0.019	0.205
FR	0.03	0.013	0.049	0.013		0.119
EN	0.046	0.018	0.095	0.008	0.011	0.129
BA	0.114	0.023	0.096	0.003	0.006	0.019
MD	0.005	0.009	0.02		0.008	0.052
IP	0.069				0.041	0.091
SC	0.07	0.033	0.092	0.012	0.007	0.182
IN	0.011	0.022	0.046	0.009		0.082

JJA

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
EA	0.152	0.045	0.131	0.152		0.206
NI	0.091	0.033	0.086	0.058		0.173
ME	0.094		0.079	0.025	0.032	0.179
FR	0.041		0.067	0.051	0.023	0.168
EN	0.021	0.015	0.02	0.01	0.004	0.166
BA	0.109	0.051	0.121	0.136	0.075	0.012
MD			0.009	0.014		0.064
IP	0.079	0.054	0.101	0.13		0.101
SC	0.033		0.006	0.007	0.01	0.126
IN	0.01		0.007	0.045		0.061

Figure 5. Values of relative importance (contribution to the total explained variance) of relative humidity in the seasons AMJ, JAS and MAM and JJA.

SPECIFIC COMMENTS

L214: something is missing in this sentence after "observed"; please correct.
L214 has been changed.
L217: insert "on" after "meteorological influence".
L217 changed.
L271: should read "... defined from a climatology of observational data ..." or "... defined from climatologies of observational data ..."
L271 corrected.
L272: "including" instead of "included"
L272 corrected.
L305: "emission densities" instead of "emissions densities"
L305 corrected.
L352: nothing wrong but IMO the sentence would read more easily this way: "the domains covered by observations and CTMs do not coincide exactly"
Thanks for this suggestion. L352 has been modified.
L438: better: "... shows ozone peak concentrations in ..." or "... in the EMEP model ozone concentrations peak in April while ..."
L438 modified.
L467: "products" instead of "product"
L476 corrected.

L552: "mentioned" instead of "mentions"

L552 removed.

L569-571: it appears to me that these two sentences contradict each other; please clarify.

L569-571 modified.

L617: insert "the" after "While"

L617 modified

L658: "associated with" instead of "associated to"

L658 modified

L667: insert "a" after "show" to read "... observasions show a lower ..."

L667 modified

L698: insert "back" after "brought" to read "... are brought back the following ..."

L698 modified

L718: insert "be" after "partly" to read "... could partly be explained ..."

L718 modified

L722: insert "to" after "attributed" to read "... could be attributed to other ..."

L722 modified

Author's response to Referee #4

Anonymous Referee #4 ***Received and published: 16 May 2018***

General comments:

Otero et al. evaluate the ability of a suite of state-of-the-art regional air quality models in their ability to reproduce the observed relationship between meteorological variables and surface ozone over Europe. They use a multiple linear regression approach, harnessing their previous experience using MLR from an observation-only standpoint. Their results are very relevant given the simulations for CMIP6 are beginning, providing context for future simulations based on how well models represent the ozone meteorology relationship in present day. Although some of their results are mostly speculative, I understand that it is difficult to fully diagnose the potential issues within each model without further sensitivity simulations. I would current form is much more qualitative. The paper is well written but requires minor revision prior to publication in ACP.

We would like to thank the referee for the careful reading, and acknowledging the usefulness of the manuscript. We appreciate the insightful comments that have helped to improve the manuscript.

The introduction is quite long and contains information that is not highly relevant to the paper, which seems, at least in part, due to some excessive self-citation. I suggest trimming some of the more basic background details as well as some reorganization for clarity. For example, paragraphs 2-5 could be condensed into a single "how meteorology affects AQ" paragraph. Although general numbers can be gleaned from the figures, the results are very lacking in the amount of quantitative statistics portrayed in the main text. Much of the discussion is very surface level and doesn't really provide the reader with new information other than the models are different from one another and observations. In the entire results section, there are only five or six specific numbers quoted. Obviously numbers shouldn't just be reported for the sake of reporting, but I'm certain the reader could benefit from more. It would be helpful if the authors could put some of the results in the context of the differences between the models; i.e., "model A,B, and C may show X due to their representation of Y". This may require a bit of digging into the additional model diagnostics (e.g., biogenic emission rates) but it would help to provide some additional insight.

Thank you for your suggestions to improve the balance of the manuscript. This has been a common concern from others referees and as mentioned in the general author's response, the revised version of the manuscript has significantly been improved following the referees' suggestions.

We understand the referee's comment regarding the quantitative discussion. However we would like to highlight that in this case it is difficult to provide a full diagnostic about the specific issues of the CTMs without more simulations and sensitivity tests in the model set-up, which is beyond of the scope of this study. Considering the strict requirements of the input data in the EDT exercise, most of the discrepancies among the models found in this analysis can be attributed to the model formulation of the physical and chemical processes. Therefore, we believe that our study provides further insights and indications for future improvement and model development.

Specific comments:

Page 2, lines 84-85: referred *TO* as *THE* climate penalty
L84, 85 have been modified.
Page 2, line 89: remove "the", *concluded that climate change*
L89 have been modified.
Page 3, lines 116-117: "meteorological dependence: : :" this sentence is oddly worded
L117 has been modified.
Page 3, lines 118-119: Here and elsewhere, when describing meteorological relationships with ozone, it should be worded "the ozone–relative humidity relationship" or "the relationship between ozone and relative humidity"
Thanks you for this comment. L118, 630, 660, 668 have been changed.
Page 3, lines 126-127: The statement about wind speed is abrupt and out of place.
Thanks you for this comment. L186-187 have been modified.
Page 4, lines 186-187: Which studies?
L186,187 have been modified.
Page 4, line 196: Mid-Atlantic U.S. states?
It is Eastern U.S. L196 has been slightly modified.
Page 5, line 217: influence *ON* MDA8 O3
L217 has been corrected.
Page 5, line 217 and elsewhere: Curious, but why not use subscripts for O3?
Thank for this comment. All subscript are now used in the text.
Table 1: This could be expanded so the reader does not have to refer elsewhere.
Thank for this comment. Table 1 has been modified.
Page 8, line 415: *MDA8*
L415 has been corrected.
Page 9, lines 433-435: The sentence: "Models show discrepancies: : :" is extremely vague.
L433-435 have been modified.
Page 10, line 485: the phrase "that, in general show lower values of R2 in JAS than in AMJ" is unnecessary since "with the exception" implies the reverse of what was said previously.
L485 has been modified.
Page 10, line 486: No need to say "certain regions such as", just say which ones.
L486 has been corrected.
Page 11, line 518: Please elaborate on "non-local processes".
L518 has been modified.
Page 12, line 595: due to *THE* effect
L518 has been corrected.
Page 14, lines 656-659: Please elaborate, are you saying that thunderstorms are not represented well? What part of the meteorological-chemistry relationship?
Here, we are arguing that the contribution of the relative humidity in the MLR might represent the impact of afternoon thunderstorms or moister conditions that reduce high levels of ozone, which might be poorly captured by CTM-based MLRs (compared with the influence of relative humidity in the observed-based MLR).

Page 14, lines 660-662: Sentence is awkwardly worded, also provide a reference if you say "the documented" anything.
L660-662 have been modified.
Page 15, line 723: *processes*
L723 has been corrected.

Author’s response to Short Comment from Dr.C. Ordóñez

In different parts of this manuscript the authors mention that the largest discrepancies between modelled and observed MDA8 O3 are found for the Balkans. They attribute that to the low number of stations interpolated into the 1x1 degree grid cells in the dataset by Schell et al. (2014). I agree that is problem for the Balkans and for other "external regions", as correctly indicated by the authors, but our experience also shows that there are some inhomogeneities in that ozone dataset over the Balkans.

First, in Ordóñez et al. (2017) we examined the impact of high-latitude and subtropical anticyclones on surface ozone. That work found (i) upward ozone trends in that dataset over the Balkans and (ii) did not establish a clear impact of anticyclonic systems on ozone over the same region. Consequently, we omitted the Balkans from our regional analyses. Later on, Carro-Calvo et al. (2017) carried out a much more detailed evaluation of the quality of the ozone dataset before analysing the synoptic drivers of summer ozone in Europe. Figure S1 in the supplement of that paper displays the regions with some inhomogeneities prior to 2004. There is a small region with inhomogeneities over Scandinavia which should hardly affect your results and a much larger area covering most of your Balkan region. In Carro-Calvo et al. (2014) we decided to remove all O3 data over those regions before 2004.Note that both Ordóñez et al. (2017) and Carro-Calvo et al. (2017) used a longer ozone dataset created by Jordan Schnell for a 15-year period (1998-2012). However, I have had a quick look at the shorter dataset used here and still see very low ozone mixing ratios in the Balkans during the first years. That might at least partly explain the high model biases (Figure 2) and low correlations (Figure 3) reported by this manuscript for that region. That could also have important implications for the results of the multiple linear regression models. As an example, the observation-based models suggest a very strong impact of ozone persistence in the Balkans, while that impact is not so strong for modelled ozone (Figure 7).
I would recommend the authors to plot the full time series of MDA8 O3 (daily values) averaged over that region and see if there is any break-point (I guess that around 2004) with a clear shift in the data. Then I would remove the data before that break-point and repeat all the analyses for that region.

We would like to thank Dr. Carlos Ordóñez for the thorough revision of our manuscript and his useful comments.

We agree with you regarding the problem in the Balkans region. As we indicated in the manuscript, the evaluation in this region is particularly complicated due to the reduced number of stations. In addition, we think that the larger differences in this region (as showed in others so-called “*external*” regions) might be due to the impact of boundary conditions and the difficulty in capturing dynamical processes (such as recirculation patterns). Following your suggestion we have analysed the full time series of MDA 8 O_3 spatially averaged over each region. However, we did not observe a clear break point in the time series corresponding to the Balkans region (see revised version of the supplementary material). Therefore, we decided to include the whole dataset in our analysis. We agree that further analysis should investigate this issue in order to clarify and better understand the differences of this dataset over this particular region.

In the last paragraph of page 14, the authors speculate on the reasons for the relatively low skill of the models in northern Europe: "Moreover, in the case of the external regions of northern Europe, it could also be explained due to the dominance of transport processes such as the stratospheric-tropospheric exchange or long-range transport from the European continent, rather than local meteorology, particularly in AMJ (Monks, 2000, Tang et al. 2009, Andersson et al. 2009)". According to the results of Carro-Calvo et al. (2017), I believe that is the case not only for spring but also for the summer months (JJA in that paper).

Thank you for this comment. We agree that even in summer the meteorology has a smaller influence in the so-called *external* regions when compared with the meteorological impact in the *internal* regions (see Figs. 6 and 7 in the manuscript).

Finally, I have read the manuscript with interest. I understand the reviewers' concerns but still think that some of the findings will be relevant for the community. As pointed out by one of the reviewers, "it is difficult to fully diagnose the potential issues within each model without further sensitivity simulations". The analysis of the results for the ensemble mean/median, as suggested by another reviewer, will not be sufficient to understand all the reasons for those discrepancies. However, that could help summarise some of the results and identify the meteorological drivers and processes (e.g. relative humidity, dry deposition?) which should be investigated in more detail in the future, through (i) careful evaluation of model parameterisations and (ii) sensitive simulations.

We also understand the suggestion of the referee regarding the use of an ensemble mean. We would like to emphasize that our main goal is to assess the discrepancies between the CTMs using specific meteorological drivers. Therefore, as Carlos Ordóñez points out in this comment, we do not consider that the use of an ensemble mean can provide more insights within our evaluation.

Having that in mind, I am confident this manuscript will be a good contribution to the field. Some of its findings will hopefully raise our awareness about some processes which need to be better investigated in air quality models.

Finally, we appreciate the positive feedback acknowledging the contribution of our study.

A multi-model comparison of meteorological drivers of surface ozone over Europe

Noelia Otero[1,3], Jana Sillmann[2], Kathleen A. Mar[1], Henning W. Rust[3], Sverre Solberg[4], Camilla Andersson[5], Magnuz Engardt[5], Robert Bergström[5], Bertrand Bessagnet[6], Augustin Colette[6], Florian Couvidat[6], Cournelius Cuvelier[7], Svetlana Tsyro[8] , Hilde Fagerli[8], Martijn Schaap[9,3], Astrid Manders[9], Mihaela Mircea[10], Gino Briganti[10], Andrea Cappelletti[10], Mario Adani[10], Massimo D'Isidoro[10], María-Teresa Pay[11], Mark Theobald[12], Marta G. Vivanco[12] , Peter Wind[8,13], Narendra Ojha[14], Valentin Raffort[15] and Tim Butler[1,3]

[1]Institute for Advanced Sustainability Studies e.V., Potsdam, Germany
[2]CICERO Center for International Climate Research, Oslo, Norway
[3]Freie Universität Berlin, Institut für Meteorologie, Berlin, Germany
[4]Norwegian Institute for Air Research (NILU), Box 100, 2027 Kjeller, Norway
[5]SMHI, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[6]INERIS, Institut National de l'Environnement Industriel et des Risques, Verneuil en Halatte, France
[7]European Commission, Joint Research Centre (JRC), Ispra, Italy
[8]MET Norway, Norwegian Meteorological Institute, Oslo, Norway
[9]TNO, Netherlands Institute for Applied Scientific Research, Utrecht, the Netherlands
[10]ENE-National Agency for New Technologies, Energy and Sustainable Economic Development, Bologna, Italy
[11]Barcelona Supercomputing Center, Centro Nacional de Supercomputación, Jordi Girona, 29, 08034 Barcelona, Spain
[12]CIEMAT, Atmospheric Pollution Unit, Avda. Complutense, 22, 28040 Madrid, Spain
[13]Faculty of Science and Technology, University of Tromsø, Tromsø, Norway
[14]Max-Planck-Institut für Chemie, Mainz, Germany
[15]CEREA, Joint Laboratory Ecole des Ponts ParisTech – EDF R&D, Champs-Sur-Marne, France

Abstract. The implementation of European emission abatement strategies has led to significant reduction in the emissions of ozone precursors during the last decade. Ground level ozone is also influenced by meteorological factors such as temperature, which exhibit interannual variability, and are expected to change in the future. The impacts of climate change on air quality are usually investigated through air quality models that simulate interactions between emissions, meteorology and chemistry. Within a multi-model assessment, this study aims to better understand how air quality models represent the relationship between meteorological variables and surface ozone concentrations over Europe. A multiple linear regression (MLR) approach is applied to observed and modelled time series across ten European regions in springtime and summertime for the period of 2000-2010 for both models and observations. Overall, the air quality models are in better agreement with observations in summertime than in springtime, and particularly in certain regions, such as France, Mid-Europe or East-Europe, where local meteorological variables show a strong influence on surface ozone concentrations. Larger discrepancies are found for the southern regions, such as the

Balkans, the Iberian Peninsula and the Mediterranean basin, especially in springtime. We show that the air quality models do not properly reproduce the sensitivity of surface ozone to some of the main meteorological drivers, such as maximum temperature, relative humidity and surface solar radiation. Specifically, all air quality models show more limitations to capture the strength of the ~~relationship~~ ozone-relative humidity relationship detected in the observed time series in most of the regions, in both seasons. Here, we speculate that dry deposition schemes in the air quality models might play an essential role to capture this relationship. We further quantify the relationship between ozone and maximum temperature ($m_{o3\text{-}T}$, climate penalty) in observations and air quality models. In summertime, most of the air quality models are able to reproduce reasonably well the observed climate penalty in certain regions such as France, Mid-Europe and North Italy. However, larger discrepancies are found in springtime, where air quality models tend to overestimate the magnitude of observed climate penalty.

1. Introduction

Tropospheric ozone is recognised as a threat to human health and ecosystem productivity (Mills et al. 2007). ~~Moreover, ozone is an important greenhouse gas (IPCC, 2013).~~ It is produced by photochemical oxidation of carbon monoxide and volatile organic compounds (VOCs) in the presence of nitrogen oxides ($NOx=NO+NO_2$) (Jacob and Winner, 2009). While it is an important pollutant on a regional scale, due to the long-range transport effect it may also influence air quality on a hemispheric scale (~~Monks et al., 2015,~~ Hedegaard et al, 2013, Monks et al., 2015). ~~Moreover, its strong relationship with temperature represents a major concern, since under a changing climate the efforts on new air pollution mitigation strategies might be insufficient. This effect, referred~~ to ~~as~~ the ~~climate penalty (Wu et al., 2008), is expected to play an important role on future air quality (Hendriks et al. 2016). Therefore it is essential to better understand the potential implications of climate change on pollutant levels. In a comprehensive review of the existing literature about the robustness of climate penalty on Europe, Colette et al. (2015) concluded that the climate change might act against mitigation measures.~~

Previous studies have shown that the reduction of emissions of ozone precursors, ~~NOx and VOCs,~~ lead to a decrease in tropospheric ozone concentrations in Europe (Solberg et al. 2005, Jonson et al. 2006). However, there is also a large year-to-year variability due to weather conditions (Andersson et al. 2007).

Ozone variability is strongly related to meteorological conditions. ~~There is a~~ S~~strong~~ignificant correlations between ozone and temperature ~~that~~ have~~s~~ been associated with the temperature-dependent lifetime of peroxyacetyl nitrate (PAN), and also due to the temperature dependence of biogenic emission of isoprene (Sillman and Samson, 1995).

Substantial increases in surface ozone have been associated with high temperatures and stable anticyclonic, sunny conditions that promote ozone formation (Solberg et al. 2008). ~~Ozone peak concentrations are also affected by closing of the plants' stomata at very high temperatures (Hodnebrog et al. 2012).~~ Moreover, its strong relationship with temperature represents a major concern, since under a changing climate the efforts on new air pollution mitigation strategies might be insufficient. This effect, referred to as the climate penalty (Wu et al., 2008), is expected to play an important role in future air quality (Hendriks et al. 2016). Similarly, increasing solar radiation leads to high levels of ozone, though with a weak correlation (Dawson et al. 2007) and it has been suggested that it could reflect in part the association of clear sky with high temperatures (Ordónez et al., 2005). ~~Several studies have assessed the model dependence of ozone on temperature (e.g. Steiner et al. 2006, Rasmussen et al. 2013). Recently, Coates et al. (2016) used a box model to investigate the influence of temperature and NOx on ozone production. Their analysis suggested that reductions in NOx would be required to offset additional ozone increase due to increasing temperatures under a warmer climate. An extensive review about the impacts of temperature on ozone production can be found in Pusede et al. (2015).~~

~~Relative h~~Humidity influences photochemistry through reactions between water vapor and ~~the~~ atomic oxygen ~~radical~~ (Vautard et al. 2012). ~~Previous studies have shown the importance of relative humidity on ozone pollution episodes (Camalier et al. 2007, Davies et al. 2011). Regional studies reported a negative relationship between ozone and relative humidity (Dueñas et al. 2002, Elminir 2005, Demuzere et al., 2009). Some authors attributed this negative correlation to the photolysis of ozone and subsequent loss of O1(D) to H2O (Jacob and Winner). H~~High levels of humidity are ~~usually~~ normally related ~~with~~ to enhanced cloud cover and thus reduced photochemistry (Dueñas et al. 2002, Camalier et al. 2007)~~. Andersson and Engardt (2010) highlighted the importance of including meteorological dependence for dry deposition of ozone to vegetation, also incorporating soil moisture dependence~~.The relationship between ozone and relative humidity can be also explained by dry deposition through stomatal uptake: under low levels of humidity plants close their stomata, which reduce the biogenic uptake (Hodnebrog et al. 2012, Kavassalis and Murphy, 2017). –High wind speed is usually correlated with low ozone concentrations due to enhanced advection and deposition, although the processes involved are complex and studies from different regions reported weak or insignificant correlations (Dawson et al., 2007, Jacob and Winner, 2009).

~~With a simple modelling approach, Kavassalis and Murphy (2017) found that the relationship~~ between ~~ozone~~ and ~~relative humidity was well captured by the inclusion of the vapour pressure deficit-dependent dry deposition, indicating the relevance of detailed dry deposition schemes in the CTMs.~~

~~Increasing solar radiation leads to an increase of ozone, though with a weak effect (Dawson et al. 2007) and it has been suggested that it could reflect in part the association of clear sky with high temperatures (Ordónez et al., 2005). Then, changes in cloud cover can also affect the photochemistry of ozone production and loss (Jacob and Winner, 2009). Additionally, low wind speed is usually associated with high ozone pollution levels (Jacob and Winner, 2009).~~

~~The influence of climate change on ozone and its precursors can involve multiple processes (Colette et al, 2015). A common approach to study the impact of climate change on air quality requires the use of air quality models that aim to represent~~

~~dynamic and chemical processes in the atmosphere. The relevance of climate change for future European air quality has been assessed in several studies that also reflect differences depending on the modelling system and future emissions scenarios adopted for each study (e.g. Lagner et al. 2005, Meleux et al. 2007, Anderson and Engardt, 2010).~~

~~Air quality models can be divided into two categories: offline chemistry transport models (CTMs) in which the model chemistry runs using meteorological data as input, and online models that allow coupling and integration of chemistry with some of the physical components to various degrees (Baklanov et al. 2014). Differences between offline and online modelling approaches can be fairly small or significant, depending on the level of the model complexity and simulated variables (Zhang, 2008).~~ Chemistry transport models (CTMs) are one of the most common tools to investigate the impacts of climate change on air quality (Jacob and Winner, 2009, Colette et al. 2015). ~~The large number and complex interactions between meteorology and chemistry in the atmosphere influence the ability of the model to represent observed situations (Kong et al. 2014).~~ Due to assumptions, parametrizations and simplifications of processes, the models themselves are subject to large uncertainties (Manders et al. 2012), which have been reflected in some regional differences in the magnitude of surface ozone response to projected climate change (Andersson and Engardt, 2010). Thus, model biases when compared to observations still remain a concern, especially in terms of the response of air quality under future climate (Fiore et al. 2009, Rasmussen et al. 2012). Comparisons between model outputs and measurements of available observational dataset are essential to evaluate the models ability to reproduce observations. Discrepancies in the outputs of CTMs can be due chemical and physical processes, fluxes (emissions, deposition and boundary fluxes) and meteorological processes (Vautard et al. 2012, Bessagnet et al. 2016). ~~Comparisons between model outputs and measurements of available observational dataset assess the reliability of air quality models, and they are essential to quantify the models ability to reproduce observations.~~ In particular, q~~Q~~uantification and isolation of the effects of meteorology on ozone is a challenge, due to the complex interrelation between ozone, meteorology, emissions and chemistry (Solberg et al. 2015). Thus, evaluating air quality models with respect the meteorological inputs is important given that meteorology drives numerous chemical processes (Vautard et al. 2012). A number of studies have evaluated the performance of the meteorological models that drive CTMs by comparing them with observations of weather parameters relevant for air quality (Smyth et al., 2006, Vautard et al. 2012, Brunner et al. 2014, Makar et al. 2015, Bessagnet et al. 2016).

Capturing~~There is a large number of representative studies in the literature that have established the relationship between surface ozone concentrations and meteorological variables using statistical modelling techniques (e.g. Bloomfield et al. 1996, Chaloukau et al 2003, Barrero et al. 2005, Ordóñez et al., 2005, Camalier et al., 2007, Seo et al., 2014, Porter et al. 2015, Otero et al., 2016). Most of these works examined the impact of meteorology on ozone pollution levels through observational datasets. Only a few studies (e.g. Davis et al. 2011), to our knowledge, have been used statistical modelling to assess the examined the statistical statistical relationship between surface ozone and meteorological parameters from models.~~ observed sensitivities of ozone to meteorological factors is required to assess the confidence in the models and their ability to reproduce the observed relationships between pollutants and meteorology and better understand potential impacts under climate change. However, only a few studies have used model simulations to analyse ozone sensitivities to meteorological parameters. Davis et al. (2011) evaluated the performance of the Community Multiscale Air Quality (CMAQ) model to reproduce the ozone sensitivities to meteorology across Eastern US. Their results showed that the model underestimated the observed ozone sensitivities to temperature and relative humidity. Recently, Fix et al. (2017) examined the capability of the NRCM-Chem model to capture the meteorological sensitivities of high/extreme ozone. Overall, they found substantial differences between the modelled and the observed sensitivities of high levels of ozone to meteorological drivers that were not consistent between the three regions of study. Due to the complex interactions and processes, estimating the ozone sensitivities to key meteorological variables remains a challenge. Thus, we aim to examine~~Our study evaluates~~ ~~–~~the capabilities of a set of CTMs to reproduce the observed ozone responses to meteorological variables. To our knowledge, this is the first multi-model evaluation that compares ~~the~~ observed and modelled meteorological sensitivities of ozone over Europe using a set of regional air quality models.

~~The EURODELTA project was initiated by the Task Force on Measurement and Modelling and the Joint Research Centre of the European Commission to provide a benchmark for the EMEP model in order to assess its relevance for policy support (Colette et al.2017a). These multi-model exercises contribute to further improving modelling techniques and understanding the associated uncertainties in the models performance. Previous exercises have evaluated the performance of chemistry transport models for future European air quality (e.g. van Lon et al. 2007, Thunis et al. 2008). Recently, Bessagnet et al. (2016) presented an intercomparison and evaluation of chemistry transport model performance with a joint analysis of some meteorological fields. They highlighted the limitations of models to simulate meteorological variables, such as wind speed and planetary boundary layer height. Particularly, in the case of ozone, they showed the importance of boundary conditions on model calculations. Within this framework, the ongoing Eurodelta-Trends (EDT) exercise (Colette et al. 2017a) builds upon this tradition and focuses on the context of air quality trends modelling.~~ The~~is~~ EURODELTA-Trends (EDT) exercise has been designed to better understand the evolution of air pollution and assess the efficiency of mitigation strategies for improving air quality.~~on the EDT exercise~~ The EDT ~~project~~exercise ~~will~~ allows the evaluation of the skill of regional air quality models and quantification of the role of the different key driving factors of surface ozone, such as emissions changes, long-range transport and meteorological variability (more details on the EDT exercise can be found in Colette et al. 2017a). Earlier phases of EURODELTA and other relevant modelling exercises, such as AQMEII (Air Quality Model Evaluation International Initiative, Rao et al. 2009) covered a short period of time of one year, while only a few studies assessed long-term air quality but limited to one model (Vautard et al., 2006, Jonson et al. 2006, Wilson et al. 2012), or utilised climate data rather than reanalysed meteorology (e.g. Simpson et al., 2014; Colette et al., 2015). The EDT exercise presents a multi-model hindcast of air quality over 2 decades (1990-2010), ~~which~~ and thus ~~and provides~~offers a ~~great~~ good opportunity to evaluate the role of driving meteorological factors on ozone variability.

~~its drivers over the last two decades (1990-2010) by the use of state-of-the-art air quality models. The EDT project will allow the evaluation of the skill of regional air quality models and quantification of the role of the different key driving factors of surface ozone, such as emissions changes, long-range transport and meteorological~~

~~variability.~~ ~~One of the main goals of the EDT project is to assess the efficiency of mitigation strategies for improving air quality (more details can be found in Colette et al. 2017a).~~

~~Davis et al. (2011) developed regression models to analyse the observed and modelled relationship between meteorology and surface ozone across the Eastern of U.S. They found that the Community Multiscale Air Quality (CMAQ) model did not capture the effect of temperature and relative humidity on daily maximum 8-h ozone and it generally underestimated the observed sensitivities to both meteorological variables, especially in the northeast. Rasmussen et al. (2012) examined the ozone-temperature relationship in a coupled chemistry-climate model and they found that the model underestimated the effect of temperature on ozone over the Mid-Atlantic. Lemaire et al. (2016) proposed a combined statistical and deterministic approach to assess the air quality response to projected climate change. Based on a data set from a deterministic climate and chemistry models, they identified the two major drivers of surface ozone over eight European regions, selected from a set of potential predictors that reached the highest correlations with ozone. Afterwards they built statistical models consisting of generalized linear models, which could be used to predict air quality.~~

~~Given that meteorology plays an essential role for surface ozone concentrations, it might be a considerable source of uncertainties in model outputs.~~ T~~herefore, t~~he present study provides a novel and ~~, thus, aims to provide a~~ simple method to evaluate the performance of air quality models~~examine the influence of meteorological variability on modelled surface ozone concentrations over Europe.~~ ~~. Therefore, this study offers a method of model evaluation capable of understanding the discrepancies between air quality models and observations~~ in terms of ~~representing the relationship to~~ meteorological sensitivities of ozone.~~input variability.~~ Specifically, our analysis focuses on the European ozone season (April to September) over the years 2000-2010. The choice of this period is mainly motivated by the availability of the observational dataset from Schnell et al. (2014, 2015) (see section 2.1). Within the EDT framework, a recent report has presented the main findings on the long-term evolution of air quality (Colette et al. 2017b). Part of these results was obtained from the analysis of the 1990s (1990-2000) and 2000s (2000-2010) separately. ~~Consistently,~~ W~~w~~e ~~decided to~~ focus on the second decade (2000-2010), for which the interpolated dataset of observed ~~on~~ maximum daily 8-hourly mean ozone (MDA8 O_3) used in this study was available. Similarly to Otero et al. (2016), we apply a multiple linear regression approach to examine the meteorological influence on MDA8 O_3 ~~O3~~. Statistical models are developed separately for observational datasets and air quality models, with the primary focus on examining ~~the~~ both observed and simulated relationships between MDA8 O_3~~O3~~ and ~~potential~~ meteorological drivers ~~in the air quality models and comparing these with the corresponding relationships determined from observed data~~.~~. Therefore, this study offers a method of model evaluation capable of understanding the discrepancies between air quality models and observations in terms of representing the relationship to meteorological input variability.~~

The present paper is structured as follows. Section 2 describes the observational data as well as the air quality models studied here. The methodology and the design of the statistical models are introduced in section 3. Section 4 discusses the results and the summary and conclusions are discussed in section 5.

2. **Data**

2.1. **Observations**

This study uses gridded MDA8 O_3~~O3~~ concentrations created with an objective-mapping algorithm developed by Schnell et al. (2014). They applied a new interpolation technique over hourly observations of stations from the European Monitoring and Evaluation Programme (EMEP) and the European Environment Agency's air quality database (AirBase) to calculate surface ozone averaged over 1° by 1° grid cells (see Schnell et al., 2014, 2015). ~~Recently,~~ Otero et al. (2016) used this dataset for examining the influence of synoptic and local meteorological conditions over Europe.~~).~~ ~~A complete description of this process can be found in Schnell et al. (2014, 2015). The gridded dataset covers a total of 15-years (1998-2012), but here we use a common period of 11-years for both observations and CTMs (2000-2010).~~ This interpolated product offers a possibility to establish a direct comparison between observations and CTMs. However, it must be acknowledged that for some areas with a low number of stations (i.e. the southeastern or northeastern European regions) the values interpolated into the 1°x1° degree grid cells may not be representative of such large scales. Recently, Ordóñez et al. (2017) and Carro-Calvo et al. (2017) have used this product to assess the impact of high-latitude and subtropical anticyclonic systems on surface ozone and the synoptic drivers of summer ozone respectively. They reported inhomogeneities during some years for specific grid-cells (e.g. in the Balkans and Sweden), which were excluded from their analysis. However, we did not observe a clear shift when analysing the spatial averages of the time series of the MDA8 O_3 for those particular regions (e.g. Balkans and Scandinavia) (Figs. S1, S2)~~). A complete description of this process can be found in Schnell et al. (2014, 2015). The gridded dataset covers a total of 15-years (1998-2012), but here we use a common period of 11-years for both observations and CTMs (2000-2010).~~, Therefore our analysis includes the whole dataset.

This study investigates the ~~observed~~ influence of observed meteorological variables on MDA8 O_3~~O3~~, based on the ERA-Interim reanalysis product provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) at 1°x1° resolution (Dee et al. 2011). Meteorological reanalyses products are essentially model simulations constrained by observations and they have been widely validated against independent observations. Daily mean values are calculated as the mean of the four available time steps at 00, 06, 12, and 18UTC for 10m wind speed components (u and v) and 2m relative humidity. Maximum temperature is approximated by the daily maximum of those time steps, while daily mean surface solar radiation is obtained from the 3-hourly values provided for the forecast fields.

2.2. **Chemistry Transport Models (CTMs)**

A set of state-of-the-art air quality models participating in the EDT exercise is used here: LOTOS-EUROS (Schaap et al., 2008, Manders et al. 2017), EMEP/MSC-W (Simpson et al., 2012), CHIMERE (Mailer et al., 2017), MATCH (Robertson et al., 1999) and~~,~~ MINNI (Mircea et al., 2016)~~ and WRF-Chem (Grell et al. 2005, Mar et al. 2016)~~. The domain of the CTMs extends from 17°W to 39.8°E and from 32°N to 70°N and it follows a regular latitude-longitude projection of 0.25x0.4 respectively. The main features of the CTM setup are largely constrained by the EDT experimental protocol (e.g. meteorology, boundary conditions, emissions, resolution, see Colette et al. 2017a for further details). For instance, the boundary conditions were defined from a climatology of observational data for most of the experiments of the EDT exercise (including~~ed~~ the data used here). However, the representation of physical and chemical processes and the vertical distribution differ in the CTMs, as well as the vertical distribution of model layers (including altitude of the top layer and derivation of surface concentration at 3m height in the case of EMEP, LOTOS-EUROS and MATCH). Moreover, there were no specific constrains imposed on biogenic emissions (including soil NO emissions), which are represented by most of the models using an online module (Colette et al. 2017a). Since we aim here to compare the modelled relationship between meteorology and surface ozone, prescribing common features in the CTMs is particularly an advantage to identify potential sources of discrepancies.

~~Only one of the participating CTMs included online coupled chemistry/meteorology (WRF-Chem), while all the rest of the models used are offline.~~ The CTMs were forced by regional ~~climate~~ numerical weather model simulations using boundary conditions from the ERA-Interim global reanalysis (Dee et al., 2011). Most of the~~se~~ ~~offline~~ CTMs used the same meteorological input data, with a few exceptions. Three of them (EMEP, CHIMERE and MINNI) used input meteorology from the Weather Research and Forecast Model (WRF) (Skamarock et al. 2008). LOTOS-EUROS and MATCH used the input meteorology produced by RACMO2 (van Meijgaard, 2012) and HIRLAM (Dahlgren et al. 2016), respectively. Unlike the rest of the regional ~~climate~~ weather models, RACMO2 used in the EDT exercise excluded nudging towards ERA-Interim, which might have some impact in the meteorological fields generated by RACMO2. ~~As mentioned, WRF-Chem couples the meteorology simulations online with chemistry. The meteorology used to drive WRF-Chem (initial and lateral boundary conditions and the application of limited four-dimensional data assimilation; see Colette et al GMD 2017a) is the same WRF meteorology from Skamarock et al. (2008) used as input for the EMEP, CHIMERE, and MINNI runs.~~ A summary of ~~Table 1 summarises~~ the CTMs and the corresponding sources of meteorological input data with some of the main characteristics are~~is~~ given in Table 1.~~used here.~~ ~~It is important to highlight that though WRF-Chem is not strictly a CTM, in order to avoid confusion with the statistical models developed in this study, we refer to all the air quality models considered (offline and online models) as CTMs hereafter.~~ As with the observations, CTMs and their meteorological counterpart were interpolated to a common grid with 1° x 1° horizontal resolution. The use of a coarser resolution could have an impact in some regions with a complex orography where airflow is usually controlled by mesoscale phenomena (e.g. see-breeze and mountain-valley winds) or in regions characterized by high emission~~s~~ densities (Schaap et al., 2015, Gan et al. 2016). In such cases the use of a finer grid could be beneficial to capture the variability of local processes.

A set of meteorological parameters was selected from the meteorological input data for the regression analyses. Similarly to the procedure with ERA-Interim, daily means are obtained from the available time steps every 3 hours in the case of WRF and RACMO2, and every 6 hours for HIRLAM for the following variables: 10m wind speed components, 2m relative humidity and surface solar radiation. Maximum temperature is also approximated by the daily maximum of those time steps.

3. Multiple l~~L~~inear regression model

Summertime usually brings favourable conditions for high ~~tropospheric~~ near-surface ozone concentrations, such as air stagnation due to high-pressure systems, warmer temperatures, higher UV radiation, and lower cloud cover (Dawson et al. 2007). ~~As stated above, the impact of meteorology on ozone concentration has been addressed through a wide variety of statistical methods in the literature.~~ This study attempts to better understand how CTMs represent the ~~influence of meteorology~~meteorological sensitivities ~~on~~ of ozone. To this aim, we use a multiple linear regression approach that can provide useful information of sensitivities in the distribution of ozone concentration as a whole (Porter et al., 2015).

A total of five meteorological predictors (Table 2) are selected based on the existing literature that has shown their strong influence on ozone pollution~~.~~ (e.g. Bloomfield et al. 1996, Barrero et al. 2005, Camalier et al. 2007, Dawson et al. 2007, Andersson and Engardt, 2010, Rasmussen et al. 2012, Davis et al. 2011, Doherty et al., 2013, Otero et al. 2016). Moreover, it has been shown that the occurrence of air pollution episodes might increase when the pollution levels of the previous day are higher than normal (Ziomas et al. 1995). Then, apart from the meteorological predictors, we add the effect of the lag of ozone (MDA8 from the previous day) in order to examine the role of ozone persistence. Additionally, we include harmonic functions that capture the effect of seasonality as in Rust et al. al (2009) and Otero et al. (2016), which is referred as “day” in the MLRs (see Table 2).

For this study, we divide the European domain into 10 regions: England (EN), Inflow (IN), Iberian Peninsula (IP), France (FR), Mid-Europe (ME), Scandinavia (SC), North Italy (NI), Mediterranean (MD), Balkans (BA) and Eastern Europe (EA). These regions are based on those defined in the recent ETC/ACM Technical Paper (Colette et al. 2017b). For our study, we further subdivide the original Mediterranean region (MD) into a region covering the Balkans (BA), due to the strong influence of the ozone persistence on MDA8 O_3~~O3~~ over this particular region as noted previously in Otero et al. (2016). Figure 1 shows the spatial coverage of each region and Table 3 lists their coordinates. As shown in Otero et al. (2016), the relative importance of predictors in the MLRs shows distinct seasonal patterns. ~~Then~~Here, multiple linear regression models (MLR, hereafter) are developed for each region for two seasons: springtime (April-May-June, AMJ) and summertime (July-August-September, JAS). These seasons differ from the meteorological definition, but cover the period when surface ozone typically reaches its highest concentrations (i.e. April-September). Additionally, we analysed the impact of the seasons’ definition by performing sensitivity tests using the meteorological seasons (i.e. March-May-April, MAM and June-July-August, JJA). As shown in Figs. S3 and S4 (see Supplement), we found a stronger impact of some relevant key driving factors of ozone (e.g. temperature and relative humidity) when using the seasons defined above (AMJ and JAS) than when using the meteorological seasons. Therefore, we consider that our choice of 3-month periods that cover the whole ozone season is particularly useful for examining the impact of individual meteorological parameters when ozone levels are highest. Since the domains covered by observations and CTMs do not coincide exactly, ~~the observations did not cover exactly the whole European domain as CTMs,~~ we applied an observational-mask to use the same number of grid-cells for CTMs and observations. Data used to estimate parameters of the MLR were spatially averaged over each region. Thus, we compare MLRs developed separately for CTMs and observations ~~at~~ for each region and season. The observational dataset contains the gridded ~~MDA8O3~~ MDA8 O_3 and the meteorology input from ERA-Interim, while the dataset for the CTMs contains the ~~MDA8O3~~ MDA8 O_3 from each one of them along with the corresponding meteorological input (~~e.g.~~ LOTOS and RACMO2, CHIMERE and WRF, MATCH and HIRLAM) (see T~~t~~able 1).

A MLR is built to describe the relationship between MDA8 O_3~~O3~~ (predictand) and a set of covariates (or predictors) describing seasonality, ozone persistence and the influence of meteorological fields (T~~t~~able 2). A data series y_t, t= 1,..N (e.g. observations or CTM simulations) for a given region and season is conceived as a Gaussian random variable Y_t with varying mean μ_t and homogeneous variance σ^2. The mean μ_t is described as a linear function of the covariates, i.e.

$$Y_t \sim \mathcal{N}(\mu_t, \sigma^2),$$
$$\mu_t = \beta_0 + \beta_{sin} sin\left(\frac{2\pi}{365.25} d_t\right) + \beta_{cos} cos\left(\frac{2\pi}{365.25} d_t\right) + \beta_{lag} y_{t-1} + \sum_{K=1}^{K} \beta_k \, x_{t,k} \quad (1)$$

with t indexing daily values and d_t referring to the day in the year associated with the index t. β_0 is a constant offset, β_{sin} and β_{cos} are the first order coefficient of a Fourier series (e.g. Rust et al. 2009, 2013, Fischer et al. 2017), β_{lag} describes the persistence with respect to the previous day concentration y_{t-1} ; if t is the first day in the late summer season (JAS, July 1st), y_{t-1} is the concentration of June 30th. Further regression coefficients β_k describe the linear relation to potential meteorological drivers (see table 2). For covariates standardized to unit variance, the regression coefficients (β) are standardised coefficients giving the change in the predictand with the covariate in units of covariate standard deviation.

Following the same strategy as used in Otero et al. (2016), the MLRs are developed through several common steps: 1) starting with the full set of potentially useful components in the predictor, a stepwise backward regression using the Akaike Information Criterion (AIC) as a selection criterion removes successively those components in the predictor, which contribute least to the model performance; and 2) a multi-collinearity index known as variance inflation factor (VIF, Maindonald and Braun 2006) is used to detect multi-collinearity problems in the predictor (i.e. high correlations between two or more components in the predictor). Components with a VIF above 10 are left out of the predictor (Kutner et al 2004).

The statistical performance of each MLR (built separately from observations and CTMs) is assessed through the adjusted coefficient (R^2) and the root mean square error (RMSE). The R^2 estimates the fraction of total variability described by the MLR and the RMSE gives the average deviation between model and observation obtained in the MLR. We also examine the relative importance of the individual components in the predictor. According to the method proposed by Lindeman et al (1980), the relative importance of each predictor is estimated by its contribution to the R^2 coefficient (Grömping 2007). We assess the sensitivities of ozone to the predictors through the standardised coefficients obtained from the regression. These coefficients indicate the changes in the ozone response to the changes in the predictors, in terms of standard deviation. Thus, for every standard deviation unit increase (decrease) of a specific predictor, the predictand (MDA8 O_3~~O3~~) will increase (decrease) the amount indicated by its coefficient in standard deviation units,. The use of standardised coefficients allows us to establish a direct comparison in the influence of individual predictors. The effect of seasonality introduced by the harmonic functions (namely, “day”, table 2) is kept in the MLRs (Eq. 1) for its usefulness in improving the power of the regression analysis, however further explanation about the effect of the predictors focuses on the rest of the variables.

4. Results and discussion

4.1. CTM performance by region

We compare the seasonal cycle of observations and CTM results through the time series of daily averaged values of MDA~~O~~8 O_3~~O3~~ from observations and CTMs for the whole period (i.e. April-September, 2000-2010) spatially averaged over each region. Furthermore, correlation coefficients between both CTMs and observations at each region and season are used to quantify the CTM performance.

4.1.1. Seasonal cycle of MDA8 O_3~~O3~~

We examine the ozone seasonal cycle represented by both the observational and modelled dataset. Figure 2 depicts daily averages during 2000-2010 of MDA8 O_3~~O3~~ at each region for the CTMs and observations. In general, all CTMs are biased high compared with observations. CTM results are visually closer to observations in the northwestern regions (i.e. IN, EN and FR), while the spread becomes larger over the southern and southeastern regions (i.e. BA, NI, MD). The IN, EN and SC regions show the highest observed concentrations in the starting months (AMJ), which is not generally well captured by most of the CTMs, ~~and they~~which show a more flat timeline (e.g. LOTOS, MATCH, CHIMERE~~ or WRF-Chem~~). For example, in the SC region, some of the CTMs underestimate the ozone concentrations in AMJ (i.e. ~~WRF-Chem,~~ CHIMERE and MINNI). The rest of the regions show the highest observed concentrations in JAS, which is generally overestimated by the CTMs. Models show discrepancies in the ozone seasonal cycle –when compared to each other and when compared against~~ to~~ observations. ~~, and in some regions we find substantial differences.~~ For example, we observed ~~Larger ~~substantial ~~discrepancies are found~~differences in the southern regions, such as IP, MD and BA, where the models show a considerable spread. ~~There~~In those regions, the CTMs ~~are not able to capture the variability of MDA8 O_3O3 and they ~~exhibit a different behaviour when compared to each other. For instance, the EMEP model shows ozone peak concentrations ~~a peak of ozone levels ~~in April, while CHIMERE and MINNI show a peak in July. Overall LOTOS shows a relatively constant positive bias in all regions, more evident in the MD and NI regions. ~~WRF-Chem tends to underestimate the ozone concentrations at the start of the seasonal period in some regions (e.g. SC, ME, EN, or EA).~~

~~CTM assessments have been presented in early EURODELTA exercises, although with a different set up for different purposes, which makes it difficult to establish a direct comparison on the performance of the models. For instance, Colette et al. (2017b) reported systematic differences among some models (i.e. CHIMERE, EMEP and LOTOS) when examining the long-term mean ozone concentration during the whole period of 1990-2010. Bessagnet et al. (2016) showed that most of the models in their study, (e.g. CHIMERE, LOTOS, or MINNI among others) overestimated the ozone concentrations in the selected study period. Specifically, they found a larger spread during nighttime than daytime, which was suggested to be related to the vertical mixing,~~

~~given that most of the models shared the same meteorology but different vertical resolution and boundary conditions.~~

4.1.2. Correlation coefficients between modelled and observed time series

The correlation coefficients between the observed and modelled values of MDA8 O3 at each region and in each season are shown in Fig. 3. Overall, MDA8 O_3~~O3~~ from the CTMs is better correlated with observations in JAS than in AMJ in the regions ME, NI, EA and EN. As expected from inspection of the average time series (Fig. 2), the lowest correlations between models and observations are found in BA, especially in AMJ for all models. In particular, EMEP is negatively correlated with observations over this region. As mentioned above, the larger discrepancies between CTMs and observations found over BA might be attributed to a low density of observation sites from which the interpolated dataset is derived, resulting in a lower quality or higher uncertainties of such products (Schnell et al. 2014). The highest correlations in AMJ are obtained at the following regions: ME; FR; NI; and EN for most of the models, except for EMEP for which the highest correlation with observations was found in IN and SC. ~~The WRF-Chem model also shows a different behaviour in terms of the correlation coefficient with higher values in NI, MD and IP, and very low and negative correlations (-0.02) in SC.~~ In general, the models that are most closely correlated with observations are MATCH, MINNI and CHIMERE, while LOTOS ~~and WRF-Chem~~ shows the lowest correlations, which ~~In the case of LOTOS, it~~ could be partially due to the use of a different set-up of the RACMO2 model, without nudging towards ERA-Interim (section 2.2). These correlations reflect the patterns represented by the seasonal cycle described above.

4.2. MLR performance

Figures 4 and 5 depict the statistical performance of each MLR in terms of R^2 and RMSE (respectively) at the different regions for both seasons, AMJ and JAS. The R^2 values indicate that all MLRs models (based on both observations and CTMs) are able to explain more than 60% of the MDA8 O_3~~O3~~ variance in all regions. Overall, the MLRs show a stronger fit in JAS than in AMJ in most of the regions~~, with the exception of SC and IN that, in general show lower values of R^2 in JAS than in AMJ~~ (Fig. 4). The MLRs appear to perform better in ~~certain~~ regions such as NI, ME, FR or EA, while the poorest statistical performance is found in IN and EN. The results obtained from the CTM-based MLRs show a similar performance to the observation-based MLRs in most of the regions. The lowest RMSE values for most of the MLR are found in SC ranging between 1 and 3 ppb, while EN shows the largest RMSE values~~, especially for the MLR built from WRF-Chem (Fig. 5)~~. The MLRs from MATCH and CHIMERE show the lowest RMSE values (1-3ppb) suggesting the best statistical fit from a predictive point of view.

Both R^2 and RMSE metrics indicate that the statistical performance of MLRs for observations and CTMs show distinct variations between seasons and regions. Overall, better performances are found in JAS and in some regions (i.e. ME, NI, or FR) where MLRs are able to describe more than the 80% of the variance in CTMs and observations. This could be attributed to the major role of meteorology in summer influencing local photochemistry processes of ozone production, while in spring long range transport plays a stronger role (Monks, 2000, Tarasova et al. 2007). As it includes the bias, the RMSE reveals more differences among the MLRs when compared to each other (e.g. larger errors for ~~WRF-Chem or~~ LOTOS when compared to MATCH or CHIMERE). However, it is interesting that in general all MLRs show a similar tendency when evaluating the statistical performance, which indicate that observations-based and CTMs-based MLRs present a similar statistical performance for modelling MDA8 O_3~~O3~~. The ability of the CTMs to reproduce the influence of meteorological drivers on MDA8 O_3 ~~O3~~ is discussed in more detail below.

4.3. Effects of drivers o~~f~~n ozone concentrations

The analysis of the influence of the predictors in the MLRs reveals distinctive regional patterns in both observation-based and CTM-based MLRs. In agreement with Otero et al. (2016), here we also find that the regions geographically located towards the interior (including central, western and eastern regions) appear to be more sensitive to the meteorological predictors, especially in JAS. On the contrary, ~~a minor~~less meteorological contribution is found in the regions over the northernmost and southernmost ~~edges~~part of the domain, implying that non-local processes (e.g. long-range transport) play a stronger role here. Considering such similarities, in the following, the regions: EN, FR, ME, NI and EA are referred as the internal regions, while the rest of the regions: IN, SC, IP, MD and BA, are referred as the external regions (see Fig. 1).

4.3.1 Relative importance

Figure 6 depicts the relative importance of the predictors for the observation-based and CTM-based MLRs in the internal regions (Fig. 1). Here, a larger meteorological influence (i.e., the predictors other than LO3 and day) can be seen in JAS compared to AMJ in all of these regions. In general, the dominant meteorological drivers from the observation-based MLRs in these internal regions are RH and Tx. The contribution of RH is evident in AMJ (e.g. ME, or EA), while Tx is clearly dominant in JAS. SSRD is also a key driver of MDA8 O_3 ~~O3~~ and generally, the wind factors (W10m and Wdir) appear to have a minor contribution.

Despite the CTM-based MLRs being able to capture the meteorological predictors, we observe discrepancies among the internal regions when compared to the observation-based MLR. The inter-model differences in terms of the relative importance of predictors are greater in AMJ than in JAS. For instance, the contribution of the LO3 is overestimated by most of CTMs~~, specifically WRF-Chem that shows a larger sensitivity to LO3 in both seasons over all of these regions. Similarly, EMEP also shows a larger contribution of LO3 than the rest of the CTMs, particularly in AMJ~~. Substantial differences are found in the influence of RH when comparing the observation-based and the CTMs-based models. The CTMs do not capture the relative importance of the RH well, especially in AMJ. In general, the CTMs driven by WRF meteorology show a slightly larger contribution of RH in most of the cases, although we notice that there are also some differences among the models that share the same meteorology. CTMs do capture the relative importance of Tx in all regions, but overall they overestimate it, as they also show for SSRD. Here, we find discrepancies when comparing the contribution of predictors in the statistical models from CTMs driven by the same meteorology (e.g. EMEP ~~and WRF-Chem~~ when compared to CHIMERE and MINNI). Such differences among the models using the same meteorology point out that the model setup (e.g.

number of vertical levels, depth of first layer) and model parameterizations (e.g chemistry/physical processes) have a larger influence in the model performance than the meteorological processes.
~~The largest differences among the CTMs are found for WRF-Chem, which tends to underestimate the contribution of the meteorological drivers in most of the regions. Interestingly, as mentions in Section 2, this is the only online coupled model participating in EDT.~~

Figure 7 presents the relative importance of individual predictors in the MLRs developed at the external regions (Fig. 1) for both seasons. The observation-based MLRs show that the main driving factor is LO3 in AMJ, while the effect of meteorological drivers becomes stronger in JAS. RH presents a larger contribution in some regions (e.g. IN, IP or SC) in AMJ and Tx in JAS (e.g. IN, IP, SC and BA). The contribution of wind components, Wdir and W10m, is mainly reflected in both seasons in the western regions (i.e. IN and IP) and in MD, respectively.

Overall, all CTMs show this tendency, although there are substantial differences when comparing the individual drivers' contribution in the observation-based and CTM-based MLRs, particularly in AMJ (Fig. 7). CTMs do not capture the contribution of LO3 reflected by the observation-based MLRs. As in the previous analysis (section 4.1) the largest discrepancies are found in BA~~, –~~. In this region,~~where~~ the observation-based MLR shows that most of the variability of ozone would be explained by LO3~~. On the contrary the~~, while the CTM-based MLRs underestimate the contribution of LO3 and overestimate the meteorological ~~effect in terms of larger~~ contribution of Tx, SSRD and RH (e.g. LOTOS, CHIMERE and MINNI). The contribution of RH is, again, underestimated by the CTMs in most of the regions, (except in BA). On the contrary, the relative importance of SSRD is overestimated in some regions (e.g. IP, IN or MD) and Tx (IN, SC), in particular for the CTMs driven by WRF. Overall, CTMs show the observed contribution of W10m and Wdir in both seasons, although with some inconsistences among the regions and CTMs.

Our results indicate that the relative importance of meteorological factors is stronger in the internal regions (Fig.6) than in the external regions (Fig.7), which could be partially attributed to a larger variability of most of the meteorological fields in internal regions (Fig. S5~~1~~). The external regions are also more likely to be influenced by the lateral boundary conditions applied by each CTM. In addition, in some external regions (e.g. IP or MD), as mentioned in section 2, the use of a coarser grid in some regions might be insufficient to capture mesoscale processes, such as land-sea breezes, which also control MDA8 O_3 ~~O3~~ concentrations (Millán et al. 2002). Moreover, we observe that meteorology becomes more important in summer, when local photochemistry processes are dominant. In general, CTMs show this tendency, but limitations to reproduce the effect of some meteorological drivers are found. Specifically, while CTMs tend to overestimate the contribution of Tx, and SSRD, they underestimate the relative importance of RH, which is also reflected in the correlations coefficients between predictand the predictors (Figs. S6~~2~~, S7~~3~~).

4.3.2 Sensitivity of ozone to the drivers

We assess the sensitivities of MDA8 O_3 ~~O3~~ to the drivers through their standardised coefficients obtained in the MLR (Section 3). These coefficients provide further information about the changes of MDA8 O3 due to the effect of each driver. Figures 8 and 9 depict the values of the main driving factors obtained in the MLR for the internal and the external regions (respectively): LO3, Tx and RH. Similarly to those patterns described by the relative importance of drivers, we observe that the ozone response to LO3 is stronger in AMJ than in JAS: the corresponding standardised coefficients are always positive and generally higher in AMJ. The observed sensitivities to LO3 are smaller in the internal regions (Fig. 8), being particularly dominant in the external regions (Fig. 9). Overall, most of the CTMs reflect a similar tendency. However, there are evident differences ~~among~~ between observations and CTMs when comparing the values of the standardised coefficients, specifically in some regions such as BA or MD. When comparing the ozone responses of the CTMs to LO3, we observe that in most of the regions MATCH and MINNI show values closest to observations~~, while WRF-Chem shows a large sensitivity to LO3~~, which ~~.~~ is consistent with the results described at the beginning of this section (4.1.2).

Correlations between MDA8 O_3 ~~O3~~ and Tx are strong, especially in the internal regions in JAS (Fig. ~~S2~~S6). Overall, we show that the CTMs appear to capture the observed effect of Tx better in JAS than in AMJ in most of the regions. The highest sensitivities to Tx are found in ~~some~~ internal regions such as ME, NI, FR and EN, which is also shown in the CTMs (Fig. 8). However, we see that most of the CTMs tend to overestimate the effect of Tx~~.~~ ~~Moreover,~~and distinct sensitivities to Tx are also ~~shown by~~found for those models that share the same meteorology (i.e. CHIMERE, EMEP and~~,~~ MINNI~~ and WRF-Chem~~). In particular, the MINNI and CHIMERE models show higher Tx sensitivities when compared to the rest of the CTMs. While the MINNI model presents the highest sensitivities to Tx in spring, ~~specifically~~ especially in EN and FR, EMEP shows smaller values and it underestimates the correlations between Tx and MDA8 O_3 (Figs. ~~S2~~S6, ~~S3~~S7).

The slope of the ozone-temperature relationship ($m_{O3\text{-}T}$) has been used in several studies to assess the ozone climate penalty (eg. Bloomer et al., 2009, Steiner et al., 2010, Rasmussen et al., 2012, Brown-Steiner et al. 2015) in the context of future air quality. Thus, we additionally analyse the ~~relationship~~ ozone-temperature relationship in order to provide insight into the ability of CTMs to reproduce the observed $m_{O3\text{-}T}$. Similarly ~~as in~~to previous work (Brown-Steiner et al. 2015), the slopes are obtained from a simple linear regression using only Tx (without the influence from other predictors) and they are used to quantify such relationship in both seasons, AMJ and JAS.

Figures 10 and 11 illustrate the $m_{O3\text{-}T}$ for the internal and the external regions respectively. The observed $m_{O3\text{-}T}$ is larger in JAS than in AMJ. In AMJ, it ranges between -0.45 and 1.15 ppbK^{-1} with the largest values found in ME, NI and MD. In JAS, the observed climate penalty is of the order of 1-2.7 ppbK^{-1} with the largest values in EN, FR, ME, NI, and MD. CTMs show a better agreement with observations in JAS than in AMJ. CTMs tend to overestimate the climate penalty in AMJ in most of the regions, with some exceptions, such as EMEP and MATCH that systematically underestimate the slopes. Also, CTMs are generally better in simulating the observed $m_{O3\text{-}T}$ in the internal regions compared to the $m_{O3\text{-}T}$ in the external regions, where in general CTMs appear to overestimate the climate penalty in both seasons. Using this metric, we identify some regions particularly sensitive to temperature, with larger values of $m_{O3\text{-}T}$ (e.g. EN, ME, FR, NI or MD). Through a multi-model assessment, Colette et al. (2015) showed a significant summertime climate penalty in southern, western and central European regions (e.g. EA, IP, FR, ME or MD) in the majority of the future climate scenarios used. Our study shows that most of the CTMs confirm the observed climate penalty in JAS in such regions in the near present, although we found that most of the CTMs overestimate the climate penalty in AMJ, especially in the external regions.

We see a stronger effect of RH in AMJ than in JAS in the observations ~~compared with the CTMs (Figs. 8 and 9),~~ with the greatest impact in the internal regions (e.g. EA, ME, NI, FR and EN), which is not well represented by the CTMs (Figs. 8 and 9). ~~The CTMs show this tendency slightly in some regions (e.g. ME, FR or EN), but differences become evident when compared to the observed values and overall they underestimate the effect of RH.~~ As mentioned, CTMs underestimate the strength of the ~~relationship~~ correlations between ozone and relative humidity~~ RH~~ (Figs. S6~~2~~, S7~~3~~). This general lack of sensitivity to RH could also partially explain the tendency for all CTMs to show a high bias in simulated ozone compared with observations (Fig. 2). Among the possible reasons for this inconsistency, we hypothesize that it can be related to the fact that ozone removal processes can be associated ~~to~~ with higher relative humidity levels during thunderstorm activity on hot moist days, which might not be well captured by CTMs. As previous studies pointed out (e.g. Andersson and Engardt, 2010~~, Kavassalis and Murphy, 2017~~), the~~Furthermore, the documented~~ impacts of ozone dry deposition suggest that it may also play a role in explaining the problems that CTMs show to reproduce the observed ~~relationship~~ ozone-relative humidity relationship. With a simple modelling approach, Kavassalis and Murphy (2017) found that the relationship between ozone and relative humidity was better captured by the inclusion of the vapour pressure deficit-dependent dry deposition, pointing out the relevance of detailed dry deposition schemes in the CTMs.

High SSRD levels favour photochemical ozone formation and it is usually positively correlated to ozone. In this case, CTMs also present some limitations to capture this effect and they overestimated the sensitivities of ozone to SSRD (Figs. S8, S9). For example, the observations show a lower and surprisingly negative effect of SSRD. Although the correlations between SSRD and ozone are positive (see Figs. S6, S7), the presence of other predictors in the regression may reverse the sign of the estimated coefficient. The CTMs show a stronger sensitivity of ozone to SSRD and they overestimate its influence on surface ozone. Similarly, the sensitivities to Wdir and W10m are also overestimated by the CTMs, especially in AMJ (Figs. S8, S9).

Our analysis suggests that CTMs present more limitations to reproduce the influence of meteorological drivers to MDA8 O_3 ~~O3~~ concentrations in the external regions than in the internal regions, particularly in AMJ. Moreover, we find the largest discrepancies in BA, where models show the poorest seasonal performance and correlation coefficients (Figs. 2 and 3, respectively), probably due a low quality of the observational dataset.

Furthermore, LO3 is the main driver over most of the external regions and explains a large proportion to the total variability of MDA8 O_3~~O3~~, while meteorological factors play a smaller influence. Lemaire et al. (2016) found a very low performance (based on R^2) over the British Isles, Scandinavia and the Mediterranean using a different statistical approach that only included two meteorological drivers. They attributed this low skill to the large influence over those regions of long-range transport of air pollution (Lemaire et al. 2016). Our results confirm the small influence of the meteorological drivers over those regions and the strong influence of the ozone persistence. Moreover, in the case of the external regions of northern Europe, it could also be explained due to the dominance of transport processes such as the stratospheric-tropospheric exchange or long-range transport from the European continent, rather than local meteorology, particularly in AMJ (Monks, 2000, Tang et al. 2009, Andersson et al. 2009).

Previous work ~~pointed out~~suggested that local sources of NOx and biogenic VOC (ozone precursors) are important factors of summertime ozone pollution in the Mediterranean basin (Richards et al. 2013). Moreover, some studies suggested that the local vertical recirculation and accumulation of pollutants play an important role in ozone pollution episodes in this region: during the nighttime the air masses are held offshore by land-sea breeze, creating reservoirs of pollutants that are brought back the following day (Millán et al. 20002, Jiménez et al. 2006, Querol et al. 2017). All of these factors (e.g. local emissions as well as local and large-scale processes) control the ozone variability, which might explain the smaller influence of local meteorological factors shown in this study over the Mediterranean basin when compared to meteorological influence in the internal regions. Thus, we may hypothesize that the strong impact of LO3 observed in the external regions over southern Europe (i.e. IP, MD, BA) could be partially due to the role of vertical accumulation and recirculation of air masses along the Mediterranean coasts as a result of the mesoscale phenomena, which is enhanced by the complex terrains that surround the Basin. Other important factor for the strong impact of LO3 observed is the slow dry deposition of ozone on water that would favour the ozone persistence in southern Europe.

Overall we conclude that CTMs capture the effect of meteorological drivers better in the internal regions (EN, FR, ME, NI and EA), where the influence of local meteorological conditions is stronger. The major effect of meteorological parameters found in the internal European regions might be also attributed to the fact that overall the variability of meteorological conditions is larger in those regions (Fig. S5). We also find differences among the CTMs driven by the same meteorology. As mentioned in the introduction, Bessagnet el al. (2016) suggested that the spread in the model results could be partly explained by the differences in the vertical ~~diffusion coefficient and~~turbulent mixing in the planetary boundary layer, differently diagnosed in each of the CTMs. Our results also indicate that even though models share the same meteorology (considering the prescribed requirements defined by the EDT exercise) they show discrepancies when compared to each other, which could be attributed to other sources of uncertainties (such as physical and chemical internal processes in the CTMs). The NMVOC and NO_x emissions from the biosphere are critical in the ozone formation. Since biogenic emissions were not specifically prescribed, which have a strong dependence on temperature and solar radiation, discrepancies in the CTMs performances, (e.g. different sensitivities to Tx) might be expected. Furthermore, we notice that the CTMs do not reproduce consistently the regional ozone-temperature relationship, which is a key factor when assessing the impacts of climate change on future air quality.

5. Summary and conclusions

The present study evaluates the ~~capability~~ capabilities of a set of Chemical Transport Models (CTMs) to ~~represent the regional relationship between~~capture the observed meteorological sensitivities of daily maximum 8-hour average ozone (MDA8 O_3~~O3~~)

~~and meteorology~~ over Europe. Our ~~results show~~study reveals systematic differences between the CTMs in reproducing the seasonal cycle when compared to observations. In general, ~~they~~ CTMs tend to overestimate the MDA8 O_3 ~~O3~~ in most of the regions. In the western and northern regions (i.e. Inflow, England and Scandinavia), some models did not capture the high ozone levels in spring (e.g. CHIMERE and~~,~~ MINNI ~~and WRF-Chem~~), while in ~~other~~ some southern regions (e.g. Iberian Peninsula, Mediterranean and Balkans) they overestimated the ozone levels in summer (e.g. LOTOS, CHIMERE). Of the CTMs, MATCH and MINNI were the most successful in capturing the observed seasonal cycle of ozone in most regions. All CTMs revealed limitations to reproduce the variability of ozone over the Balkans region, with a general overestimation of the ozone concentrations, considerably larger during the warmer months (July, August). As reflected in the results, a limitation of the interpolated observational product used here is that in some regions (e.g. southern Europe) it has a lower quality due to a reduced number of stations (section 2.1).

The MLRs performed similarly for most of the CTMs and observations, describing more than 60 % of the total variance of MDA8 O_3~~O3~~. Overall, the MLRs perform better in JAS than in AMJ, and the highest percentages of described variance were found in Mid Europe and North Italy. This could be attributed to local photochemical processes being more important in JAS, and is consistent with a relatively stronger influence of long-range transport in AMJ.

The effects of predictors revealed spatial and seasonal patterns, in terms of their relative importance in the MLRs. Particularly, we noticed a larger local meteorological influence in the regions located towards the interior of Europe, here termed as the internal regions (i.e. England, France, Mid-Europe, North Italy and East-Europe). A minor local meteorological contribution was found in the ~~rest of the~~ remaining regions, referred as the external regions (i.e. Inflow, Iberian Peninsula, Scandinavia, Mediterranean and Balkans). The CTMs are in better agreement with the observations in the internal regions than in the external regions, where they were not as successful in reproducing the effects of the ozone drivers. Overall, the different behaviour in the MLRs developed in the external regions could be attributed to (i) a larger influence of dynamical processes rather than local meteorological processes (e.g. long range transport in the northern regions) (ii) a stronger impact of the boundary conditions (iii) the use of a coarser grid that might be insufficient to capture mesoscale processes that also influence MDA8 O_3 ~~O3~~ (e.g. sea-land breezes in the southern regions).

We found substantial differences in the sensitivities of MDA8 O_3 to the different meteorological factors among the CTMs, even when they used the same meteorology. As Bessagnet et al. (2016) point out, the differences amongst CTMs could be partly attributed to some other diagnosed model variables (e.g. vertical ~~diffusion coefficient~~turbulent mixing and boundary layer height, as well as vertical model resolution). To assess the effect of such potential sources of uncertainties, further investigations would be required. Moreover, variations in the sensitivity of ozone to meteorological parameters could depend on differences in the chemical and photolysis mechanisms and the implementation of various physics schemes, all of which differ between the CTMs (see Colette et al. 2017a). Specifically, the discrepancies found in the sensitivities of MDA8 O_3~~O3~~ to maximum temperature might be also attributed to biogenic emissions not prescribed in the models. This was particularly reflected in the analysis of the slopes ozone-temperature ($m_{O3\text{-}T}$) to assess the climate penalty, which differed between CTMs and regions when compared to the observations in both seasons. Most of the CTMs confirm the observed climate penalty in JAS, but with larger discrepancies in the external regions than in the internal regions. Furthermore, CTMs tend to overestimate the climate penalty in AMJ (particularly in the external regions).

Our results have shown discrepancies in the observed and simulated ozone sensitivities to relevant meteorological parameters for ozone formation and removal processes. In particular, we found ~~have shown~~ that CTMs tend to overestimate the influence of maximum temperature and surface solar radiation in most of the regions, both strongly associated with ozone production. None of the CTMs captured the strength of the observed relationship between ozone and relative humidity appropriately, underestimating the effect of relative humidity, a key factor in the ozone removal processes. We speculate that ozone dry deposition schemes used by the CTMs in this study may not adequately represent the relationship between humidity and stomatal conductance, thus underestimating the ozone sink due to stomatal uptake. Further sensitivity analyses would be recommended for testing the impact of the current dry deposition schemes in the CTMs.

Data availability

The data are available upon request from the corresponding author.

Acknowledgments

We acknowledge Jordan L. Schnell for providing the interpolated dataset of MDA 8 O_3~~O3~~. Modelling data used in the present analysis were produced in the framework of the EURODELTA-Trends Project initiated by the Task Force on Measurement and Modelling of the Convention on Long Range Transboundary Air Pollution. EURODELTA-Trends is coordinated by INERIS and involves modelling teams of BSC, CEREA, CIEMAT, ENEA, IASS, JRC, MET Norway, TNO, SMHI. The views expressed in this study are those of the authors and do not necessarily represent the views of EURODELTA-Trends modelling teams. The MATCH participation was partly funded by the Swedish Environmental Protection Agency through the research program Swedish Clean Air and Climate (SCAC) and NordForsk through the research programme Nordic WelfAir (grant no. 75007).

List of Tables:

Table 1. Summary of the chemistry-transport models used in the study and the main characteristics (adapted from Colette et al. 2017).

Comment [NF1]: Table 1 has been changed with additional model characteristics information

Model	**Meteorological driver**	**Research group**	**Vertical layers (vl)** **Vertical extent (ve)** **Surface concentration (sc)** **Depth first layer (dl)**	**Biogenic VOC**	**Dry deposition (dd)** **Stomatal resistence (sr)**	**Land use database(lu** **Advection scheme (ad)** **Vertical diffusion (vd)**
CHIMERE	WRF (common driver)	INERIS	**vl:** 9 sigma **ve:** surface to 500 hPa **sc:** First model level **dl:** 20m	MEGAN model v2.1 with highresolution spatial and temporal leaf area index (LAI; Yuan et al., 2011) and recomputed emissions factors based on the land use (Guenther et al., 2006)	**dd:** Resistance model (Emberson et al., 2000a, b) **sr:** Emberson et al. (2000a, b)	**lu:** GLOBCOVER (24 classes) **ad:** van Leer (1984) **vd:** vertical diffusion coefficient (Kz) approach following Troen and Mahrt (1986)
EMEP	WRF (common driver)	MET Norway	**vl:** 20 sigma **ve:** surface to 100 hPa **sc:** Downscaled to 3 m **dl:** 90m	Online emissions based upon maps of 115 species from Koeble and Seufert (2001), and hourly temperature and light using Guenther et al. (1993, 1994). See Simpson et al. (1995, 2012)	**dd:** Resistance model for gases (Venkatram and Pleim, 1999); for aerosols: Simpson et al. (2012) **sr:** DO3SEEMEP: Emberson et al. (2000a, b), Tuovinen et al. (2004), Simpson et al. (2012)	**lu:** CCE/SEI for Europe, elsewhere GLC2000 **ad:** Bott (1989) **vd:** Kz approach following O’Brien (1970) and Jeriˇceviˇc et al. (2010)
LOTOS-EUROS	RACMO2	TNO	**vl:** 5 (4 dynamic layers and a surface layer) **ve:** 5000 m **sc:** Downscaled to 3 m **dl:** 25m	Based upon maps of 115 species from Koeble and Seufert (2001), and hourly temperature and light (Guenther et al., 1991, 1993). See Beltman et al. (2013)	**dd:** Resistance model, DEPAC3.11 for gases, Van Zanten et al. (2010) and Zhang et al. (2001) for aerosols **rs:** Emberson et al. (2000a, b)	**lu:** Corine Land Cover 2000 (13 classes) **ad:** Walcek (2000) **vd:** Kz approach Yamartino et al. (2004)
MATCH	HIRLAM EURO4M	SMHI	**vl:** 39 hybrid levels of the meteorological model layers **ve:** surface to ca. 5000 m (4700–6000 m) **sc:** Downscaled to 3 m **dl:** ca. 60m	Online emissions based on Simpson et al. (2012), dependent on hourly temperature and light	**dd:** Resistance model depending on aerodynamic resistance and land use (vegetation). Similar to Andersson et al. (2007) **sr:** Simple, seasonally varying, diurnal variation of surface resistance for gases with stomatal resistance (similar to Andersson et al., 2007 and Simpson et al., 2012)	**lu:** CCE/SEI for Europe **ad:** Fourth-order massconserved advection scheme based on Bott (1989) **vd:** Implicit mass conservative Kz approach (see Robertson et al., 1999); Boundary layer parameterisation as detailed in Robertson et al. (1999) forms the basis for vertical diffusion and dry deposition
MINNI	WRF (common driver)	ENEA/Arianet S.r.l	**vl:** 16 fixed terrain-following layers **ve** 10 000m **sc:** First model level **dl:** 40m	MEGAN v2.04 (Guenther et al., 2006)	**dd:** Resistance model based on Wesely (1989) **sr:** Wesely (1989)	**lu:** Corine Land Cover 2006 (22 classes) **ad:** Blackman cubic polynomials (Yamartino,1993) **vd:** Kz approach following Lange (1989)

Predictor	Definition
LO3	Lag of MDA8 O_3 (24 h)
Tx	Maximum temperature
RH	Relative humidity
SSRD	Surface solar radiation
Wdir	Wind direction
W10m	Wind speed
day	$\sin(2\pi d_t/365.25)$, $\cos(2\pi d_t/365.25)$

Table 2. List of the predictors used in the multiple linear regression analysis: meteorological parameters, lag of MDA8 O_3 ~~O3~~ (24h, previous day) and the seasonal cycle components.

Region	Acronym	Coordinates (longitude, latitude)
England	EN	5W-2E, 50N-55N
Inflow	IN	10W-5W, 50N-60N, and 5W-2E, 55N-60N
Iberian Peninsula	IP	10W-3E, 36N-44N
France	FR	5W-5E, 44N-50N
Mid-Europe	ME	2E-16E, 48N-55N
Scandinavia	SC	5E-16E, 55N-70N
North Italy	NI	5E-16E, 44N-48N
Balkans	BA	18E-28E, 38N-44N
Mediterranean	MD	3E-18E, 36N-44N
Eastern Europe	EA	16E-30E, 44N-55N

Table 3. List of the regions with the short name and the coordinates.

List of Figures:

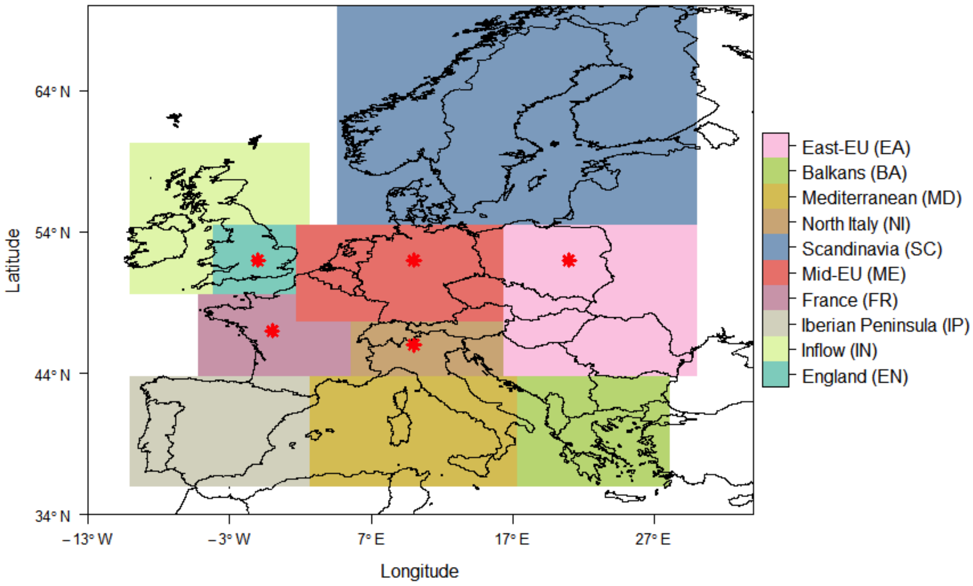


Figure 1. Map of the regions considered in the study. Regions indicated with a black star are referred to the internal regions in the text. The rest of regions are referred to the external regions of the European domain.

Comment [NF2]: Figure 1 has been changed, only the legend has been modified (NorthIt, has been replaced by North Italy).

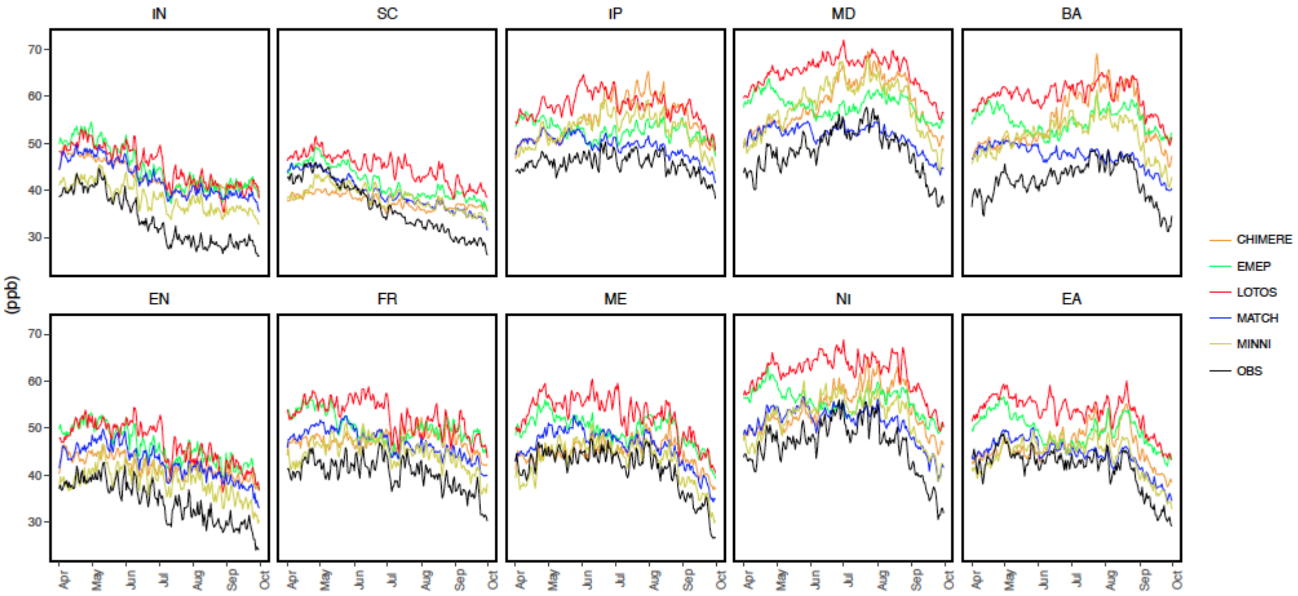
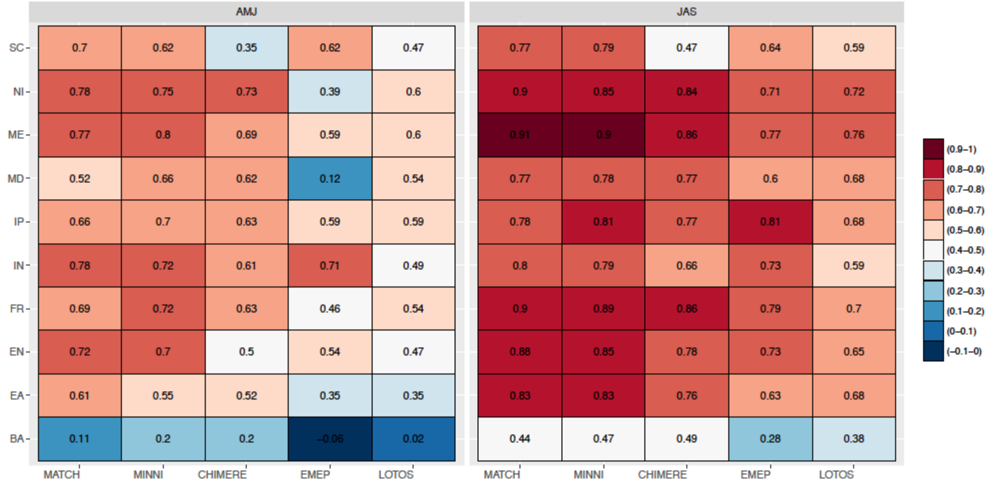


Figure 2. Time series of daily averages of MDA8 O_3~~O3~~ during the ozone season (April-September) for the period of study (2000-2010) at each subregion.



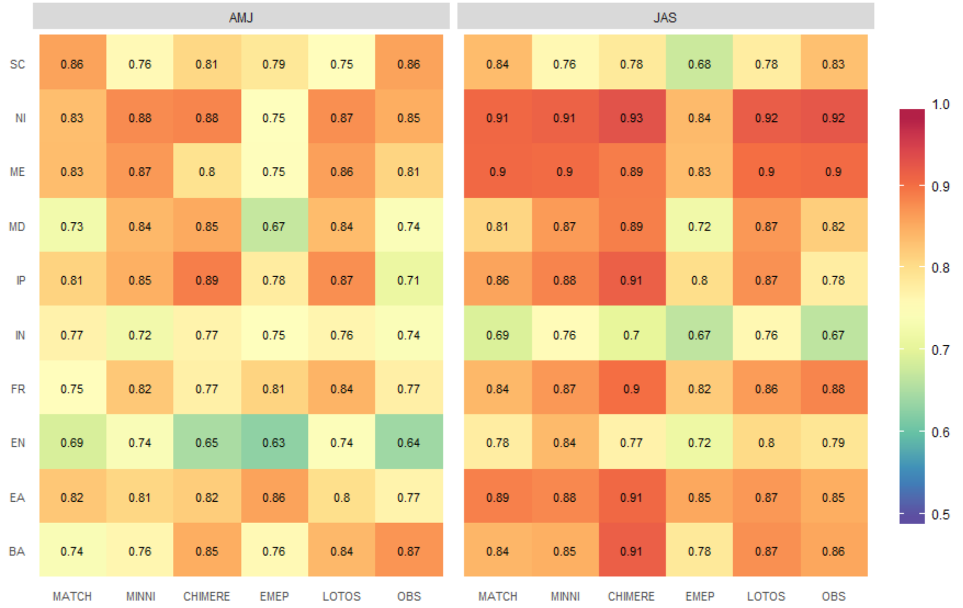
AMJ

	MATCH	MINNI	CHIMERE	EMEP	LOTOS
SC	0.7	0.62	0.35	0.62	0.47
NI	0.78	0.75	0.73	0.39	0.6
ME	0.77	0.8	0.69	0.59	0.6
MD	0.52	0.66	0.62	0.12	0.54
IP	0.66	0.7	0.63	0.59	0.59
IN	0.78	0.72	0.61	0.71	0.49
FR	0.69	0.72	0.63	0.46	0.54
EN	0.72	0.7	0.5	0.54	0.47
EA	0.61	0.55	0.52	0.35	0.35
BA	0.11	0.2	0.2	−0.06	0.02

JAS

	MATCH	MINNI	CHIMERE	EMEP	LOTOS
SC	0.77	0.79	0.47	0.64	0.59
NI	0.9	0.85	0.84	0.71	0.72
ME	0.91	0.9	0.86	0.77	0.76
MD	0.77	0.78	0.77	0.6	0.68
IP	0.78	0.81	0.77	0.81	0.68
IN	0.8	0.79	0.66	0.73	0.59
FR	0.9	0.89	0.86	0.79	0.7
EN	0.88	0.85	0.78	0.73	0.65
EA	0.83	0.83	0.76	0.63	0.68
BA	0.44	0.47	0.49	0.28	0.38

Figure 3. Correlation coefficients between observed and modelled MDA8 O_3~~O3~~ for spring (AMJ) and summer (JAS) for the period of study (2000-2010) at each region (rows) and model~~s~~ (columns, ordered by highest correlation values).



AMJ

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
SC	0.86	0.76	0.81	0.79	0.75	0.86
NI	0.83	0.88	0.88	0.75	0.87	0.85
ME	0.83	0.87	0.8	0.75	0.86	0.81
MD	0.73	0.84	0.85	0.67	0.84	0.74
IP	0.81	0.85	0.89	0.78	0.87	0.71
IN	0.77	0.72	0.77	0.75	0.76	0.74
FR	0.75	0.82	0.77	0.81	0.84	0.77
EN	0.69	0.74	0.65	0.63	0.74	0.64
EA	0.82	0.81	0.82	0.86	0.8	0.77
BA	0.74	0.76	0.85	0.76	0.84	0.87

JAS

	MATCH	MINNI	CHIMERE	EMEP	LOTOS	OBS
SC	0.84	0.76	0.78	0.68	0.78	0.83
NI	0.91	0.91	0.93	0.84	0.92	0.92
ME	0.9	0.9	0.89	0.83	0.9	0.9
MD	0.81	0.87	0.89	0.72	0.87	0.82
IP	0.86	0.88	0.91	0.8	0.87	0.78
IN	0.69	0.76	0.7	0.67	0.76	0.67
FR	0.84	0.87	0.9	0.82	0.86	0.88
EN	0.78	0.84	0.77	0.72	0.8	0.79
EA	0.89	0.88	0.91	0.85	0.87	0.85
BA	0.84	0.85	0.91	0.78	0.87	0.86

Figure 4. Coefficients of determination (R^2) for each CTM-based (ordered as in Fig.3) and observation-based MLR in spring (AMJ) and summer (JAS).

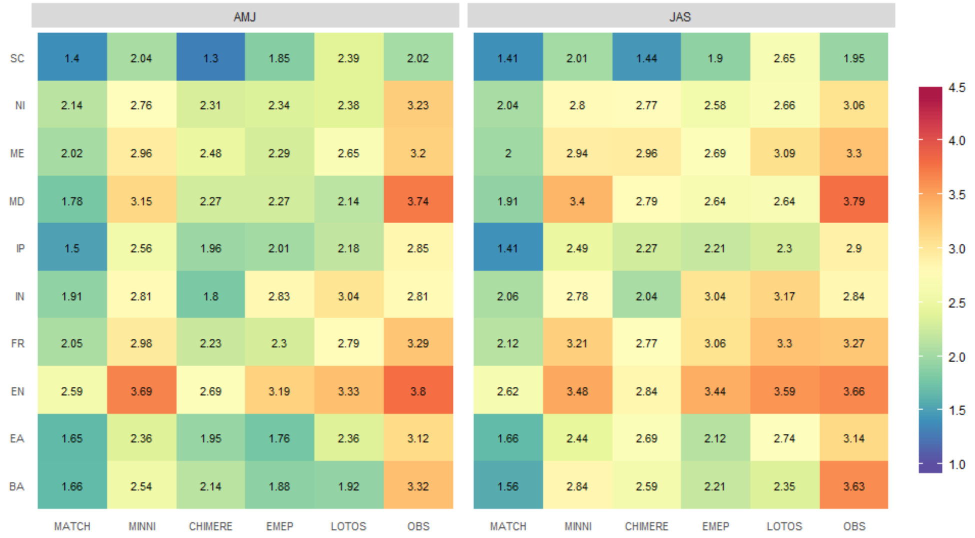


Figure 5. Root mean square errors (RMSE) for each CTM-based (ordered as in Fig.3) and observation-based MLR at each region, in spring (AMJ) and summer (JAS).

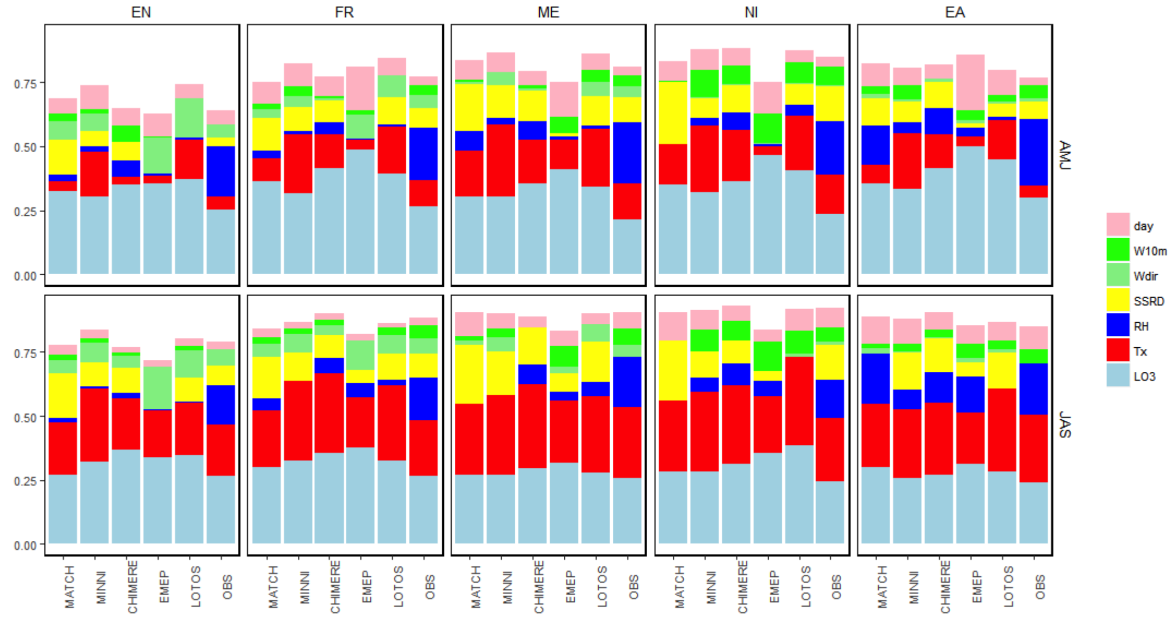


Figure 6. Proportion of each predictor to the total explained variance for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top) and JAS (bottom) for the internal regions: England (EN), France (FR), Mid-Europe (ME), North Italy (NI) and East-Europe (EA).

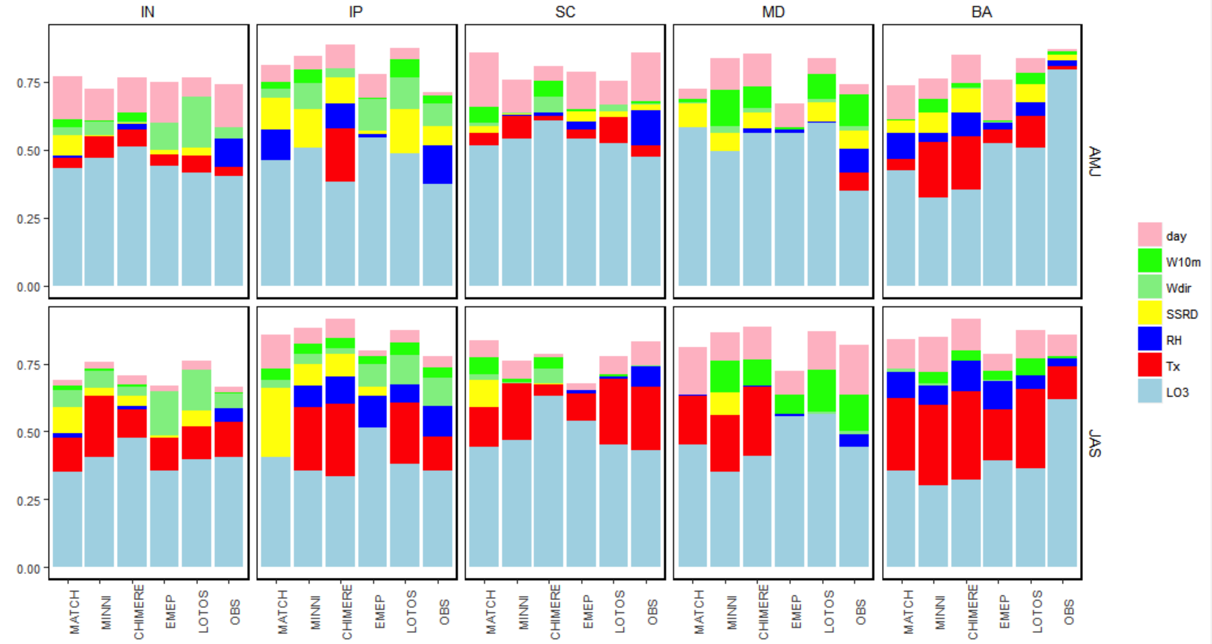


Figure 7. Proportion of each predictor to the total explained variance for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top) and JAS (bottom) for the external regions: Inflow (IN), Iberian Peninsula (IP), Scandinavia (SC), Mediterranean (ME) and Balkans (BA).

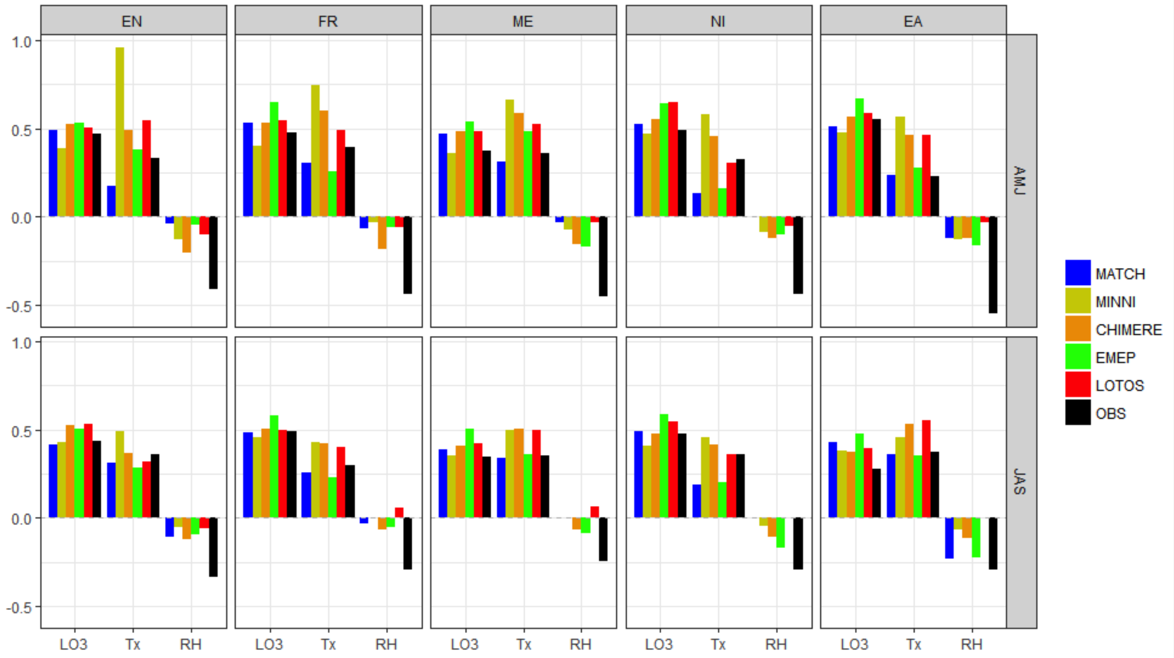


Figure 8. Standardised coefficients values of the main key-driving factors (LO3, Tx and RH) for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top) and JAS (bottom) and for the internal regions: England (EN), France (FR), Mid-Europe (ME), North Italy (NI) and East-Europe (EA).

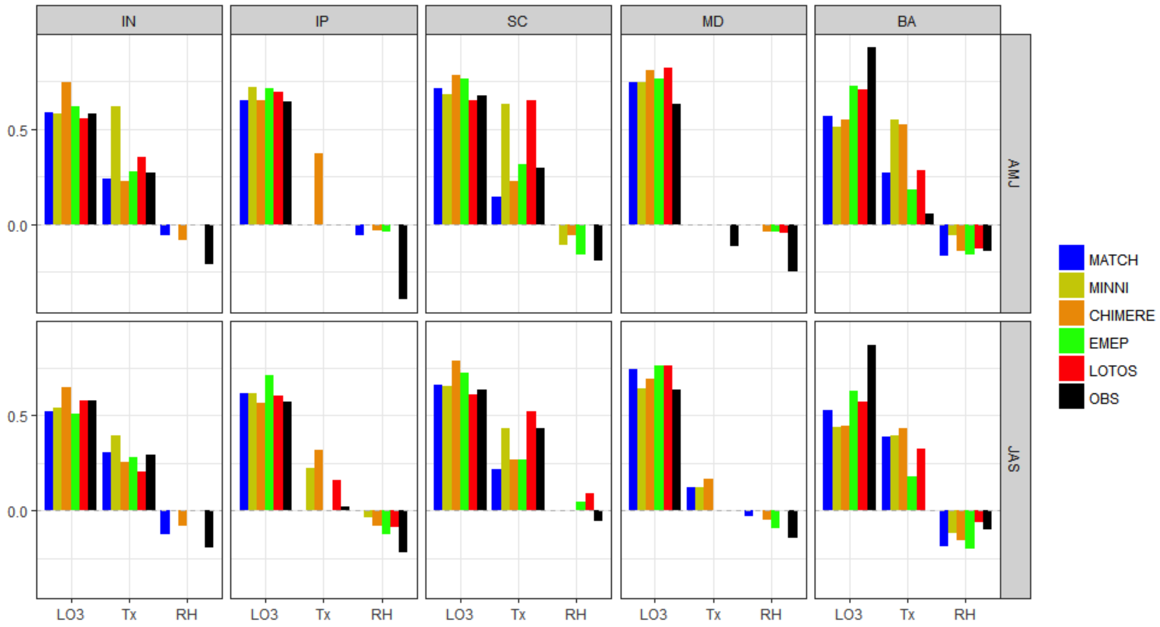


Figure 9. Standardised coefficients values of the main key-driving factors (LO3, Tx and RH) for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top) and JAS (bottom) and for the external regions: Inflow (IN), Iberian Peninsula (IP), Scandinavia (SC), Mediterranean (ME) and Balkans (BA).

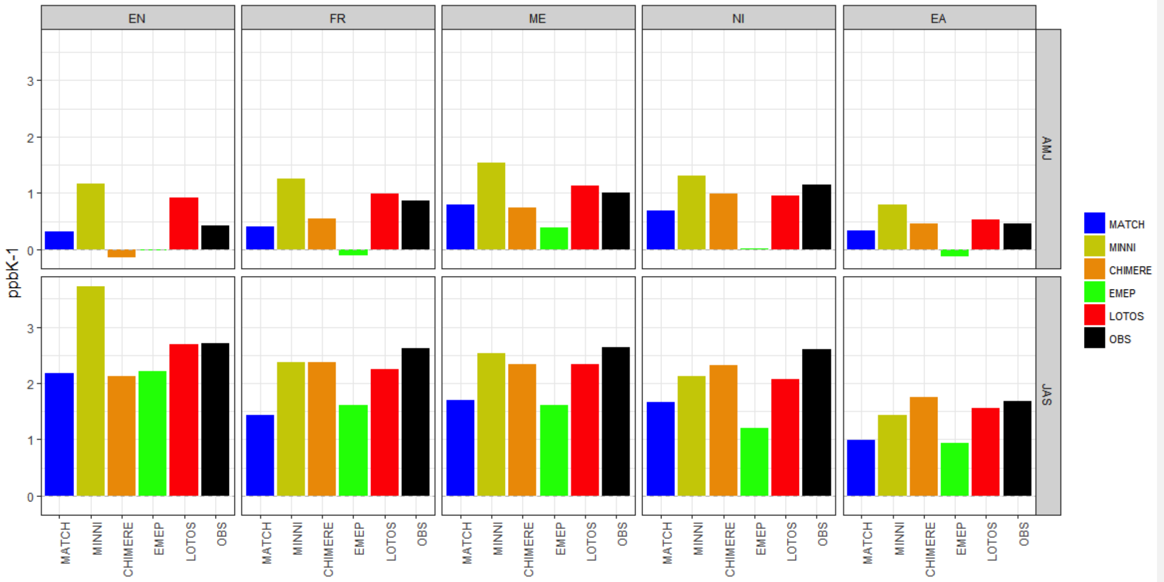


Figure 10. Slopes ($m_{O3\text{-}T}$; $ppbK^{-1}$) obtained from a simple linear regression to estimate the relationship ozone-temperature for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top)

and JAS (bottom) and for the internal regions: England (EN), France (FR), Mid-EU (ME), North Italy (NI), East-EU (EA).

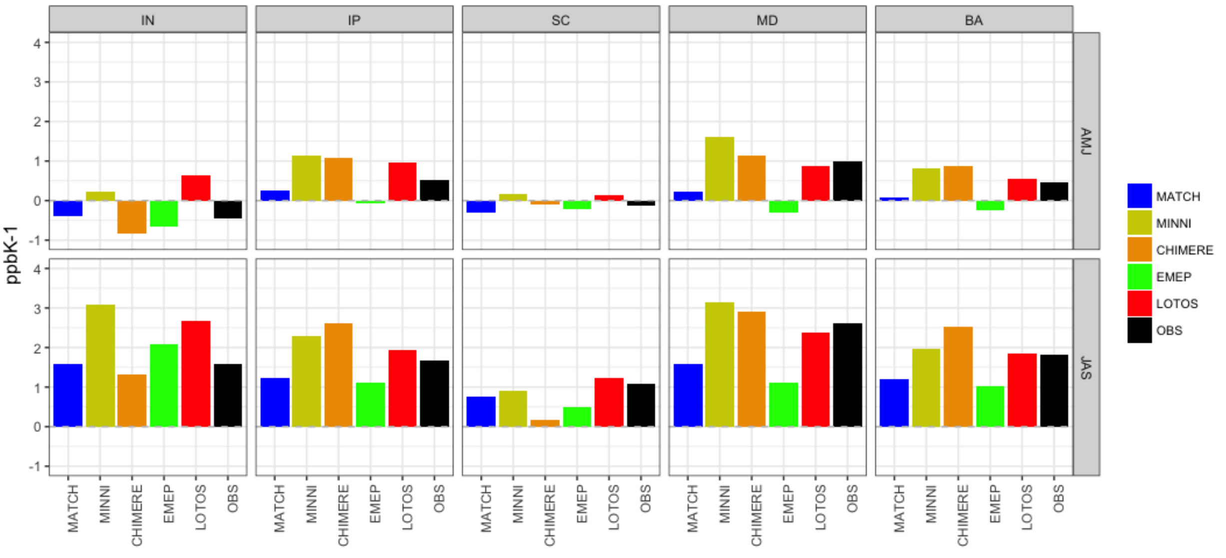


Figure 11. Slopes (m_{O3-T}; ppbK^{-1}) obtained from a simple linear regression to estimate the relationship ozone-temperature for each CTM-based (ordered as in Fig.3) and observation-based MLR in AMJ (top) and JAS (bottom) and for the external regions: Inflow (IN), Iberian Peninsula (IP), Scandinavia (SC), Mediterranean (ME) and Balkans (BA).

Comment [NF3]: Figure 11 has been changed switching the order of SC and IP (wrong in the previous version).

References

Andersson, C., J. Langner, and R. Bergström: Interannual variation and trends in air pollution over Europe due to climate variability during 1958–2001 simulated with a regional CTM coupled to the ERA-40 reanalysis, Tellus, Ser. B, 59, 77–98, 2007.

Andersson C., Bergström R., Johansson C., Population exposure and mortality due to regional background PM in Europe–Long-term simulations of source region and shipping contributions, Atmospheric Environment, 43, 22, 3614-3620, http://dx.doi.org/10.1016/j.atmosenv.2009.03.040, 2009.

Andersson, C. and Engardt, M.: European ozone in a future climate: Importance of changes in dry deposition and isoprene emissions, J. Geophys. Res.-Atmos., 115, D02303, doi:10.1029/2008jd011690, 2010.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G., Hirtl, M., Joffre, S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A., Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N., Rao, S. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmos. Chem. Phys., 14, 317–398, doi:10.5194/acp-14-317-2014, 2014.

Barrero M A, Grimalt J O and Canton L.: Prediction of daily ozone concentration maxima in the urban atmosphere Chemometr. Intell. Lab. Syst. 80 67–76, 2005.

Beltman, J. B., Hendriks, C., Tum, M., and Schaap, M.: The impact of large scale biomass production on ozone air pollution in Europe, Atmos. Environ., 71, 352–363, 2013.

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Fi- nardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J.-C., Fagerli, H., Gonzalez, L., Menut, L., Prévôt, A. S. H., Roberts, P., and White, L.: Presentation of the EURODELTA III intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmos. Chem. Phys., 16, 12667–12701, doi:10.5194/acp-16-12667-2016, 2016.

Bloomfield P J, Royle J A, Steinberg L J and Yang Q: Accounting for meteorological effects in measuring urban ozone levels and trends Atmos. Environ. 30 3067–77, 1996 .

Bloomer, B. J., Stehr, J. W., Piety, C. A., Salawitch, R. J., and Dickerson, R. R.: Observed relationships of ozone air pollution with temperature and emissions, Geophys. Res. Lett., 36, L09803, doi:10.1029/2009gl037308, 2009.

Bott, A.: A Positive Definite Advection Scheme Obtained by Nonlinear Renormalization of the Advective Fluxes, Mon. Weather Rev., 117, 1006–1015, 1989.

Brown-Steiner, B., Hess, P. G., and Lin, M. Y.: On the capabilities and limitations of GCCM simulations of summertime regional air quality: A diagnostic analysis of ozone and temperature simulations in the US using CESM CAM-Chem, Atmos. Environ., 101, 134–148, doi:10.1016/j.atmosenv.2014.11.001, 2015.

Brunner, D., Jorba, O., Savage, N., Eder, B., Makar, P., Giordano, L., Badia, A., Balzarini, A., Baro, R., Bianconi, R., Chemel, C., Forkel, R., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Im, U., Knote, C., Kuenen, J. J. P., Makar, P. A., MandersGroot, A., Neal, L., Perez, J. L., Pirovano, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Werhahn, K., Wolke, R., van Meijgaard, E., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2, Atmos. Environ., 115, 470–498, 2015.

Camalier, L., Cox, W., and Dolwick, P.: The effects of meteorology on ozone in urban areas and their use in assessing ozone trends, Atmos. Environ., 41, 7127–7137, 2007.

Carro-Calvo, L., C. Ordóñez, R. García-Herrera, J. L. Schnell: Spatial clustering and meteorological drivers of summer ozone in Europe, Atmos. Environ., 167, 496-510,https://doi.org/10.1016/j.atmosenv.2017.08.050, 2017.

Chaloulakou A, Saisana M and Spyrellis N: Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens Science of the Total Environment 313 1–13, 2003.

Coates, J., Mar, K. A., Ojha, N., and Butler, T. M.: The influence of temperature on ozone production under varying NOx conditions – a modelling study, Atmos. Chem. Phys., 16, 11601-11615, https://doi.org/10.5194/acp-16-11601-2016, 2016.

Colette, A., Andersson, C., Baklanov, A., Bessagnet, B., Brandt, J., Christensen, J., Doherty, R., Engardt, M., Geels, C., Giannakopoulos, C., Hedegaard, G., Katragkou, E., Langner, J., Lei, H., Manders, A., Melas, D., Meleux, F., Rouil, L., Sofiev, M., Soares, J., Stevenson, D., Tombrou-Tzella, M., Varotsos, K., and Young, P.: Is the ozone climate penalty robust in Europe?, Environ. Res. Lett., 10, 084015, doi:10.1088/1748-9326/10/8/084015, 2015.

Colette, A., Andersson, C., Manders, A., Mar, K., Mircea, M., Pay, M.-T., Raffort, V., Tsyro, S., Cuvelier, C., Adani, M., Bessagnet, B., Bergström, R., Briganti, G., Butler, T., Cappelletti, A., Couvidat, F., D'Isidoro, M., Doumbia, T., Fagerli, H., Granier, C., Heyes, C., Klimont, Z., Ojha, N., Otero, N., Schaap, M., Sindelarova, K., Stegehuis, A. I., Roustan, Y., Vautard, R., van Meijgaard, E., Vivanco, M. G., and Wind, P.: EURODELTA-Trends, a multi-model experiment of air quality hindcast in Europe over 1990–2010, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2016-309, accepted, 2017a.

Colette, A., Solberg, S., Beauchamp, M., Bessagnet, B., Malherbe, L., and Guerreiro, C.: Long term air quality trends in Europe: Contribution of meteorological variability, natural factors and emissions, ETC/ACM, Bilthoven, 2017b.

Comrie A. C.: Comparing neural networks and regression models for ozone forecasting J. Air Waste Manage. Assoc. 47 653–63, 1997.

Dahlgren, P., Landelius, T., Kållberg, P., and Gollvik, S.: A high-resolution regional reanalysis for Europe. Part 1: Three-dimensional reanalysis with the regional HIgh-Resolution Limited-Area Model (HIRLAM), Quarterly Journal of the Royal Meteorological Society, 142, 2119-2131, 10.1002/qj.2807, 2016.

Davis, J., Cox, W., Reff, A., Dolwick, P.: A comparison of Cmaq-based and observation-based statistical models relating ozone to meteorological parameters. Atmospheric Environment 45, 3481e3487. http://dx.doi.org/10.1016/ J.Atmosenv.2010.12.060, 2011.

Dawson, J.P., Adams, P.J., Pandis, S.N.: Sensitivity of ozone to summertime climate in the Eastern USA: a modeling case study. Atmos. Environ. 41, 1494– 1511, 2007.

Dawson, J. P., Racherla, P. N., Lynn, B. H., Adams, P. J., and Pan- dis, S. N.: Simulating present-day and future air quality as cli- mate changes: model evaluation, Atmos. Environ., 42, 4551– 4566, doi:10.1016/j.atmosenv.2008.01.058, 2008.

Dee D P et al. :The ERA-Interim reanalysis: configuration and performance of the data assimilation system Quart. J. R. Meteorol. Soc. 137 553–97, 2001.

Doherty, R. M., Wild, O., Shindell, D. T., Zeng, G., MacKenzie, I. A., Collins, W. J., Fiore, A. M., Stevenson, D. S., Dentener, F. J., Schultz, M. G., Hess, P., Derwent, R. G., and Keating, T. J.: Impacts of climate change on surface ozone and intercontinental ozone pollution: a multi- model study, J. Geophys. Res.-Atmos., 118, 3744–3763, doi:10.1002/jgrd.50266, 2013.

Emberson, L. D., Ashmore, M. R., Simpson, D., Tuovinen, J.-P., and Cambridge, H. M.: Towards a model of ozone deposition and stomatal uptake over Europe, Norwegian Meteorological Institute, Oslo, Norway, 57, 2000a.

Emberson, L. D., Ashmore, M. R., Simpson, D., Tuovinen, J.-P., and Cambridge, H. M.: Modelling stomatal ozone flux across Europe, Water Air Soil Pollut., 109, 403–413, 2000b.

Elminir H.K.: Dependence of urban air pollutants on meteorology, Science of The Total Environment, 350,225-237, http://dx.doi.org/10.1016/j.scitotenv.2005.01.043, 2005.

Fischer M., Rust H.W., and Ulbrich U.: Seasonality in extreme precipitation – using extreme value statistics to describe the annual cycle in german daily precipitation. Meteorol. Z. accepted, 2017.

Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., Textor, C., Schulz, M., Doherty, R. M., Horowitz, L. W., MacKenzie, I. A., Sanderson, M. G., Shindell, D. T., Stevenson, D. S., Szopa, S., Van Dingenen, R., Zeng, G., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Collins, W. J., Duncan, B. N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S., Hauglustaine, D., Holloway, T., Isaksen, I. S. A., Jacob, D. J., Jonson, J. E., Kaminski, J. W., Keating, T. J., Lupu, A., Marmer, E., Montanaro, V., Park, R. J., Pitari, G., Pringle, K. J., Pyle, J. A., Schroeder, S., Vivanco, M. G., Wind, P., Wojcik, G., Wu, S., and Zuber, A.: Multimodel estimates of intercontinental source-receptor relationships for ozone pollution, J. Geophys. Res.-Atmos., 114, D04301, doi:10.1029/2008jd010816, 2009.

Fix, M. J., D. Cooley, A. Hodzic, E. Gilleland, B. T. Russell, W. C. Porter, and G. G. Pfister, 2018: Observed and predicted sensitivities of extreme surface ozone to meteorological drivers in three US cities. Atmospheric Environment, 176, 292-300, doi:10.1016/j.atmosenv.2017.12.036.

Gan C., Hogrefe C., Mathur R., Pleim J., Xing J., Wong D., Gilliam R., Pouliot G., Wei C.: Assessment of the effects of horizontal grid resolution on long-term air quality trends using coupled WRF-CMAQ simulations, Atmospheric Environment, 132, 207-216,1352-2310,https://doi.org/10.1016/j.atmosenv.2016.02.036., 2016.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled “online” chemistry within the WRF model, Atmospheric Environment, 39, 6957-6975, http://dx.doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Grömping U.: Estimators of relative importance in linear regression based on variance decomposition Am. Stat. 61 139–47, 2007.

Guenther, A., Zimmerman, P., Harley, P., Monson, R., and Fall, R.: Isoprene and monoterpene rate variability: model evaluations and sensitivity analyses, J. Geophys. Res., 98, 12609–12617, 1993.

Guenther, A., Zimmerman, P., andWildermuth, M.: Natural volatile organic compound emission rate estimates for US woodland landscapes, Atmos. Environ., 28, 1197–1210, 1994.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, doi10.5194/acp-6-3181-2006, 2006.

Hedegaard, G. B., Christensen, J. H., and Brandt, J.: The relative importance of impacts from climate change vs. emissions change on air pollution levels in the 21st century, Atmos. Chem. Phys., 13, 3569-3585, doi:10.5194/acp-13-3569-2013, 2013.

Hendriks C., Forsell N., Kiesewetter G., Schaap M., Schöpp W.: Ozone concentrations and damage for realistic future European climate and air quality scenarios, Atmospheric Environment, 144, 208-219, http://dx.doi.org/10.1016/j.atmosenv.2016.08.026, 2016.

Hodnebrog, Ø., Solberg, S., Stordal, F., Svendby, T. M., Simpson, D., Gauss, M., Hilboll, A., Pfister, G. G., Turquety, S., Richter, A., Burrows, J. P., and Denier van der Gon, H. A. C.: Impact of forest fires, biogenic emissions and high temperatures on the elevated Eastern Mediterranean ozone levels during the hot summer of 2007, Atmos. Chem. Phys., 12, 8727-8750, https://doi.org/10.5194/acp-12-8727-2012, 2012

Hogrefe, C., Biswas, J., Lynn, B., Civerolo, K., Ku, J. Y., Rosenthal, J., Rosenzweig, C., Goldberg, R., and Kinney, P. L.: Simulating regional-scale ozone climatology over the eastern United States: model evaluation results, Atmos. Environ., 38, 2627–2638, 2004.

IPCC, Climate Change 2013, the Physical Science Basis. Working Group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2013.

Jacob, D. J. and Winner, D. A.: Effect of climate change on air quality, Atmos. Environ., 43, 51–63, doi:10.1016/j.atmosenv.2008.09.051, 2009.

Jericevic, A., Kraljevic, L., Grisogono, B., Fagerli, H., and Vecenaj, Ž.: Parameterization of vertical diffusion and the atmospheric boundary layer height determination in the EMEP model, Atmos. Chem. Phys., 10, 341–364, https://doi.org/10.5194/acp-10-341-2010, 2010.

Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in European ozone levels?, Atmos. Chem. Phys., 6, 51–66, doi:10.5194/acp-6-51-2006, 2006.

Kavassalis, S. C., and J. G. Murphy: Understanding ozone-meteorology correlations: A role for dry deposition, Geophys. Res. Lett., 44, 2922–2931, doi:10.1002/2016GL071791, 2017.

Koeble, R. and Seufert, G.: Novel Maps for Forest Tree Species in Europe, A Changing Atmosphere, 8th European Symposium on the Physico-Chemical Behaviour of Atmospheric Pollutants,Torino, Italy, 2001.

Kong, X., Forkel R., Sokhi R. S., Suppan P., Baklanov A., Gauss M., Brunner D., Barò R., Balzarini A., Chemel C., Curci G., Jiménez-Guerrero P., Hirtl M., Honzak L., Im U., Pérez J. L., Pirovano G., San Jose R., Schlünzen K. H. , Tsegas G., Tuccella P., Werhahn J., Žabkar R., Galmarini S.: Analysis of Meteorology-Chemistry Interactions During Air Pollution Episodes Using Online Coupled Models within AQMEII Phase-2, Atmospheric Environment, http://dx.doi.org/10.1016/j.atmosenv.2014.09.020., 2014.

Kutner M. H., Nachtsheim C. J. and Neter J.: Applied Linear Regression Models 4th Ed (Boston, MA: McGraw-Hill Irwin) 2004.

Lange, R.: Transferrability of a three-dimensional air quality model between two different sites in complex terrain, J. Appl. Meteorol., 78, 665–679, 1989.

Langner, J., Bergstro ̈m R. and Foltescu V.: Impact of climate change on surface ozone and deposition of sulphur and nitrogen in Europe, Atmos. Environ., 39, 1129–1141, doi:10.1016/j.atmosenv.2004. 09.082, 2005.

Lelieveld, J., and P. J. Crutzen: Influence of cloud and photochemical processes on tropospheric ozone, Nature, 343, 227–233, 1990.

Lemaire, V. E. P., Colette, A. and Menut, L.: Using statistical models to explore ensemble uncertainty in climate impact studies: the example of air pollution in Europe, Atmos. Chem. Phys., 16, 2559-2574, doi:10.5194/acp-16-2559-2016, 2016.

Lindeman R.H. , Merenda P. F. and Gold RZ: Introduction to Bivariate and Multivariate Analysis. Scott, Foresman, Glenview, IL, 1980.

Linderson M. L.: Objective classification of atmospheric circulation over southern Scandinavia Int. J. Climatol. 21 155–69, 2001.

Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Létinois, L., Markakis, K., and Meleux, F.: CHIMERE-2017: from urban to hemispheric chemistry-transport modeling, Geosci. Model Dev., 10, 2397-2423, https://doi.org/10.5194/gmd-10-2397-2017, 2017.

Maindonald J, Braun J: Data analysis and graphics using R: an example-based approach.Cambridge (United Kingdom), Cambridge University Press, 2006.

Makar, P.A., Gong,W., Hogrefe, C., Zhang, Y., Curci, G., Zakbar, R., Milbrandt, J., Im, U., Galmarini, S., Gravel, S., Zhang, J., Hou, A., Pabla, B., Cheung, P., Bianconi, R., 2015b. Feedbacks between Air Pollution and Weather, Part 1: Effects on Weather. Atmos. Environ. 115, 442e469.

Manders, A. M. M., van Meijgaard, E., Mues, A. C., Kranenburg, R., van Ulft, L. H., and Schaap, M.: The impact of differences in large-scale circulation output from climate models on the regional modeling of ozone and PM, Atmos. Chem. Phys., 12, 9441–9458, doi:10.5194/acp-12-9441-2012, 2012.

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier van der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J., Segers, A. J., Timmermans, R. M. A., Visschedijk, A., Wichink Kruit, R. J., Van Pul, W. A. J., Sauter, F. J., van der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., and Schaap, M.: Curriculum Vitae of the LOTOS-EUROS (v2.0) chemistry transport model, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-88, in review, 2017.

~~Mar, K. A., Ojha, N., Pozzer, A. and Butler, T. M.: Ozone air quality simulations with WRF-Chem (v3.5.1) over Europe: model evaluation and chemical mechanism comparison, Geosci. Model Dev., 9, 3699- 3728, 10.5194/gmd-9-3699-2016, 2016.~~

Meijgaard, E. v., van Ulft, L. H., Lenderink, G., de Roode, S. R., Wipfler, L., Boers, R., and Timmermans, R. M. A.: Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe, KvR 054/12, 44, 2012.

Meleux, F., Solmon, F., and Giorgi, F.: Increase in summer European ozone amounts due to climate change, Atmos. Environ., 41, 7577–7587, doi:10.1016/j.atmosenv.2007.05.048, 2007.

Millán, M. M., Sanz, M. J., Salvador, R., and Mantilla, E.: Atmospheric dynamics and ozone cycles related to nitrogen deposition in the western Mediterranean, Environ. Pollut., 118, 167–186, 2002.

Mills, G., Hayes, F., Jones, M. L. M., and Cinderby, S.: Identifying ozone-sensitive communities of (semi-)natural vegetation suitable for mapping exceedance of critical levels, Environ. Pollut., 146, 736–743, doi:10.1016/j.envpol.2006.04.005, 2007.

Mircea, M., Grigoras, G., D'Isidoro, M., Righini, G., Adani, M., Briganti, G., Ciancarella, L., Cappelletti, A., Calori, G., Cionni, I., Cremona, G., Finardi, S., Larsen, B. R., Pace, G., Perrino, C., Piersanti, A., Silibello, C., Vitali, L., and Zanini, G.: Impact of grid resolution on aerosol predictions: a case study over Italy, Aerosol and Air Quality Research, 1253–1267, doi: 10.4209/aaqr.2015.02.0058, 2016.

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

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

O'Brien, J. J.: A note on the vertical structure of the eddy Exchange coefficient in the planetary boundary layer, J. Atmos. Sci., 27, 1213–1215, 1970.

Ordóñez, C., Mathis, H., Furger, M., Henne, S., Hoglin, C., Staehelin, J., Prevot, A.S.H.: Changes of daily surface ozone maxima in Switzerland in all seasons from 1992 to 2002 and discussion of summer 2003. Atmos. Chem. Phys. 5, 1187– 1203, 2005.

Ordóñez, C., Barriopedro, D., García-Herrera, R., Sousa, P. M., and Schnell, J. L.: Regional responses of surface ozone in Europe to the location of high-latitude blocks and subtropical ridges, Atmos. Chem. Phys., 17, 3111-3131, https://doi.org/10.5194/acp- 17-3111-2017, 2017.

Otero, N., Sillmann, J., Schnell, J.L., Rust, H.W., Butler, T., 2016. Synoptic and meteorological drivers of extreme ozone concentrations over Europe. Environ. Res. Lett. 11 (2), 24005. http://dx.doi.org/10.1088/1748-9326/11/2/024005.

Porter W C, Heald C L, Cooley D and Russell B 2015 Investigating the observed sensitivities of air quality extremes to meteorological drivers via quantile regression Atmos. Chem. Phys. Discuss. 15 10349–66, 2015.

Pusede S E et al. : On the temperature dependence of organic reactivity, nitrogen oxides, ozone production, and the impact of emission controls in San Joaquin Valley, California Atmos. Chem. Phys. 14 3373–95, 2014.

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

Rasmussen, D. J., Fiore, A. M., Naik, V., Horowitz, L. W., McGinnis, S. J., and Schultz, M. G.: Surface ozone-temperature relationships in the eastern US: A monthly climatology for evaluating chemistry-climate models, Atmos. Environ., 47, 142–153, doi:10.1016/j.atmosenv.2011.11.021, 2012.

Robertson, L., Langner, J., and Engardt, M.: An Eulerian Limited-Area Atmospheric Transport Model, Journal of Applied Meteorology, 38, 190-210, 1999.

Rust H, Maraun D. and Osborn T.: Modelling seasonality in extreme precipitation Eur. Phys. J. Special Topics 174 99–111, 2009.

Rust H. W., Vrac M., Sultan B., and Lengaigne M.: Mapping weather-type influence on Senegal precipitation based on a spatial-temporal statistical model. J. Climate, 26:8189–8209. ISSN 0894-8755. URL http://dx.doi.org/10.1175/JCLI-D-12-00302.1.1. 2013

Schaap, M., Timmermans, R. M. A., Roemer, M., Boersen, G. A. C., Builtjes, P., Sauter, F., Velders, G., and Beck, J.: The LOTOS-EUROS model: description, validation and latest developments, International Journal of Environment and Pollution, 32, 270-290, 2008.

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, Atmospheric Environment, 112, 90-105, http://dx.doi.org/10.1016/j.atmosenv.2015.04.003, 2015.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR, 2008.

Schnell, J. L., Holmes, C. D., Jangam, A., and Prather, M. J.: Skill in forecasting extreme ozone pollution episodes with a global atmospheric chemistry model, Atmos. Chem. Phys., 14, 7721–7739, doi:10.5194/acp-14-7721-2014, 2014.

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Cameron-Smith, P., Bergmann, D., Zeng, G., Plummer, D. A., Sudo, K., Nagashima, T., Shindell, D. T., Faluvegi, G., and Strode, S. A.: Use of North American and European air quality networks to evaluate global chemistry–climate modeling of surface ozone, Atmos. Chem. Phys., 15, 10581-10596, doi:10.5194/acp-15-10581-2015, 2015.

Seo, J., Youn, D., Kim, J. Y., and Lee, H.: Extensive spatiotemporal analyses of surface ozone and related meteorological variables in South Korea for the period 1999–2010, Atmos. Chem. Phys., 14, 6395-6415, https://doi.org/10.5194/acp-14-6395-2014, 2014.

Sillman, S. and Samson, P.J: Impact of temperature on oxidant photochemistry in urban, polluted rural and remote environments. Journal of Geophysical Research 100: doi: 10.1029/94JD02146. issn: 0148-0227, 1995.

Simpson, D., Guenther, A., Hewitt, C., and Steinbrecher, R.: Biogenic emissions in Europe 1. Estimates and uncertainties, J. Geophys. Res., 100, 22875–22890, 1995.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmos. Chem. Phys., 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012

Simpson, D., Andersson, C., Christensen, J. H., Engardt, M., Geels, C., Nyiri, A., Posch, M., Soares, J., Sofiev, M., Wind, P., and Langner, J.: Impacts of climate and emission changes on nitrogen deposition in Europe: a multi-model study, Atmos. Chem. Phys., 14, 6995-7017, https://doi.org/10.5194/acp-14-6995-2014, 2014.

Smyth, S., Yin, D., Roth, H., Jiang, W., Moran, M.D., Crevier, L.P., 2006. The impact of GEM and MM5 meteorology on CMAQ air quality modeling results in eastern Canada and the northeastern United States. Journal of Applied Meteorology 45, 1525e1541. doi:10.1175/JAM2420.1.

Solberg, S., R. G. Derwent, Ø. Hov, J. Langner, and A. Lindskog: European abatement of surface ozone in a global perspective, Ambio, 34, 47–53, 2005.

Solberg, S., Hov, Ø., Sovde, A., Isaksen, I. S. A., Coddeville, P., De Backer, H., Forster, C., Orsolini, Y., and Uhse, K.: European surface ozone in the extreme summer 2003, J. Geophys. Res.Atmos., 113, D07307, doi:10.1029/2007jd009098, 2008.

Solberg, S., Colette, A., and Guerreiro, C. B. B.: Discounting the impact of meteorology to the ozone concentration trends. ETC/ACM, NILU, INERIS. https://doi.org/10.13140/rg.2.2.15389.92649, 2016.

Steiner, A.L., Tonse, S., Cohen, R.C., Goldstein, A.H., Harley, R.A.: Influence of future climate and emissions on Regional air quality in California. Journal of Geophysical Research-Atmospheres 111, 2006.

Vautard, R., Honore, C., Beekmann, M., and Rouil, L.: Simulation of ozone during the August 2003 heat wave and emission control scenarios, Atmos. Environ., 39, 2957–2967, doi:10.1016/j.atmosenv.2005.01.039, 2005.

Tang L., Chen D.L., Karlsson P.E., Gu Y.F., Ou T.H. Synoptic circulation and its influence on spring and summer surface ozone concentrations in Southern Sweden Boreal Environment Research, 14, 889-902, 2009.

Tarasova, O. A., Brenninkmeijer, C. A. M., Jöckel, P., Zvyagintsev, A. M., and Kuznetsov, G. I.: A climatology of surface ozone in the extra tropics: cluster analysis of observations and model results, Atmos. Chem. Phys., 7, 6099–6117, doi:10.5194/acp-7-6099-2007, 2007.

Thompson M L, Reynolds J, Cox L H, Guttorp P and Sampson P D: A review of statistical methods for the meteorological adjustment of tropospheric ozone Atmos. Environ. 35 617–30, 2001.

Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer: Sensitivity to surface evaporation, Bound.-Lay. Meteorol., 37, 129–148, 1986.

Tuovinen, J.-P., Ashmore, M., Emberson, L., and Simpson, D.: Testing and improving the EMEP ozone deposition module, Atmos. Environ., 38, 2373–2385, 2004.

van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmospheric Environment, 41, 2083-2097, 2007.

Vautard, R., Moran, M.D., Solazzo, E., Gilliam, R.C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A.B., Jericevic, A., Prank, M., Segers, A., Silver, J.D.,Werhahn, J.,Wolke, R., Rao, S.T., Galmarini, S., 2012. Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations. Atmos. Environ. 53, 15e37.

van Leer, B.: Multidimensional explicit difference schemes for hyperbolic conservation laws, in: Computing Methods in Applied Sciences and Engineering VI, edited by: Lions, R. G. A. J. L., Elsevier, Amsterdam, 1984.

Van Zanten, M. C., Sauter, F. J.,Wichink Kruit, R. J., Van Jaarsveld, J. A., and Van Pul, W. A. J.: Description of the DEPAC module: 75 Dry deposition modelling with DEPAC_GCN2010, Bilthoven, the Netherlands, 2010.

Walcek, C. J.: Minor flux adjustment near mixing ratio extremes for simplified yet highly accurate monotonic calculation of tracer advection, J. Geophys. Res., 105, 9335–9348, 2000.

Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmos. Environ., 23, 1293–1304, 1989.

Wilson, R. C., Fleming, Z. L., Monks, P. S., Clain, G., Henne, S., Konovalov, I. B., Szopa, S., and 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, 437–54, https://doi.org/10.5194/acp-12- 437-2012, 2012.

Wu, S., Mickley, L. J., Leibensperger, E. M., Jacob, D. J., Rind, D., and Streets, D. G.: Effects of 2000–2050 global change on ozone air quality in the United States, J. Geophys. Res.-Atmos., 113, D18312, doi:10.1029/2007JD009639, 2008.

Yamartino, R. J., J. Flemming, and R.M. Stern: Adaptation of analytic diffusivity formulations to Eulerian grid model layers of finite thickness. In 27th ITM on Air Pollution Modelling and its Application. Banff, Canada. 2004.

Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W.: Reprocessing the MODIS Leaf Area Index Products for Land Surface and Climate Modelling, Remote Sens. Environ., 155, 1171–1187, https://doi.org/10.1016/j.rse.2011.01.001, 2011.

Zhang, L., Gong, S., Padro, J., and Barrie, L.: A size-segregated particle dry deposition scheme for an atmospheric aerosol module, Atmos. Environ., 35, 549–560, 2001.

Zhang, Y.: Online-coupled meteorology and chemistry models: his- tory, current status, and outlook, Atmos. Chem. Phys., 8, 2895– 2932, doi:10.5194/acp-8-2895-2008, 2008.