

## Reply to reviewers

### Referee #1

#### General comments:

My major concern is related to a lack of statistics showing the significance and robustness of the results.

The authors present almost solely average plots of certain quantities but draw several times the conclusion that the frequency of events was increased/decreased (I give an example below, e.g. section 3.2.1). How do they get to these conclusions? The figure in the supplement is the only one showing at least a standard deviation for some stations. Though a frequency analysis would be the appropriate analysis method. What causes the standard deviation - extreme events or a difference in the frequency of the occurrence of certain events? Were extreme events filtered for the average plots?

*The data were analysed in more detail than we could fit in the article. 15 locations were analysed in detail, 7 of them are presented here. In Figures 3, 5 and 6 we focus on extreme situations that are particularly relevant for air quality: the number of hot days, the number of wet days and the number of days with low wind speed, and differences therein between the models and differences within a model due to climate change. These show substantial differences between the models for the present-day climate, and smaller differences due to climate change. In addition, frequency plots were added to the supplementary material for temperature, ozone concentration and PM10 concentration for two stations. No filtering took place for extreme events.*

One of the main objectives of this article is to point out differences between global model meteorological fields and the impact of the related uncertainties on air quality predictions. However, the authors do not show or mention a single time if the differences are even statistically significant. I highly recommend to add at least a simple significance test to the analysis of the model simulations.

*A significance test was done for daily maximum temperature, summer mean ozone daily maximum and annual mean PM10 concentration, significances are indicated in the text and tables in the supplementary material.*

The authors write several times sentences similar like the following: 'For RLE ECHAM, the difference in concentration between future and present-day is smaller than the interannual variability in North Europe and about equal to the interannual variability in South Europe.

Though, I have difficulties seeing immediately how the authors come to this conclusion. Therefore I would suggest to show the interannual variability!

*The interannual variability is shown for a number of locations in the supplementary material, but we now explicitly added the significance to the discussion of the differences, see previous point.*

Vast parts of the discussion focus on the uncertainties of the regional model even though this is not the objective of the article. I would suggest to focus the discussion more on the objective as given in the abstract: .... The present study illustrates the impact of this uncertainty on air quality

*The first part of the discussion has been rewritten to better focus the discussion, stronger conclusions were added to the ends of the paragraphs in the discussion.*

#### Specific comments:

##### Section 2.1:

I am a bit surprised by a (rather long) timestep of 15 minutes for a horizontal resolution of 0.44°. Don't you get interferences with the CFL criterion?

*No, since RACMO uses a semi-Lagrangian advection scheme, as mentioned in the text, allowing time step sizes beyond the CFL criterion.*

##### Section 2.2:

Could the authors add a reference for MACC?

*A reference to the MACC website for regional air quality is added*

I find the structure of this section a bit confusing. Could the authors shift the description, which gas phase mechanism and which aerosol module they are using to the beginning of the section (right after mentioning which species are included)?

*Done, and shifted the domain description to be more consistent with the RACMO section.*

A reference for CBMIV and EQSAM is missing.

*Reference to EQSAM added, but for the modified CBMIV some reaction rates are specific to LOTOS-EUROS, the version used here is not documented explicitly and no literature reference is available.*

### **Section 3.1:**

This section is a summary of Manders et al. 2011 and therefore it does not belong to the chapter 'results'.

As the objective of this article is not a future climate projection but an assessment of climate projection uncertainties I do not necessarily expect the authors to present a more in depth analysis of the hindcast for proving a general good performance of the applied model system. However, when comparing model simulations driven by ERA-interim data with ECMWF analysis data - can't we expect a good correlation, simply because these two data sets are not fully independent? Shouldn't the model simulations (the hindcast / 'present day climate') be evaluated against a totally independent data set (e.g. the CRU data set)?

*The paragraph was moved to the introduction section and modified slightly. The idea was to investigate whether RACMO would correctly downscale the ERA-Interim data, so a good correlation with ECMWF analysis was hoped for (and found indeed). This good correlation could then be used to verify the general behavior of the climate models, so there is no direct need for an independent data set for the present application. Since we are interested in differences between the models, we do not need an extensive assessment of the model performance against observations.*

C4363

### **Section 3.2.1:**

Can the authors explain in more detail how they come to the conclusion that a southward shift of the average low pressure centre leads to a more frequent occurrence of stagnant weather conditions?

*This is not what is stated in the paper. Both RLE\_MIROC and RLE\_ECHAM show a southward shift, but the spatial pattern in shift of the horizontal pressure gradient is different, causing the differences in circulation, as pointed out in the paper. The text is left as it is.*

How can you draw a conclusion from the AVERAGE (difference) in the mslp on the FREQUENCY of the occurrence of certain weather conditions? Does the southward shift of the low pressure centre in RLE-MICRO result from more frequently occurring southward shifts of the low pressure system (compared to RLE-ERA) or were 'extreme events' responsible for the difference in the average mslp? Were extreme events somehow filtered when calculating the average mean sea level pressure?

*The average mslp provides direct information on the major circulation. This includes the locations of passages of frontal systems on the one hand and anti-cyclones with prevailing stable conditions on the other hand, which so to say form the extreme events.*

p.12255, l. 24: Maybe this sentence can be rephrased as using the expression 'overestimate' might give the wrong impression that RLE-ERA is closer to reality than RLE-MICRO.

*RLE\_ERA is by its construction closer to reality, since it is driven by reanalysis meteorology instead of a free-running global climate model which has not been constrained by observations. This difference is stated more explicitly in the text.*

p. 12255, l.21: ECHAM instead of ECHAM5 ?  
*It should read ECHAM everywhere, is changed now.*

### Section 3.2.2:

The authors switched from present tense to past tense in the first sentences of this section. Can they consistently use present tense throughout the paper?

*Changed*

p. 12256, l.10: Actually the annual average daily maximum temperature is higher at ALL stations listed in S1 when using ERA compared to ECHAM.

*Yes, but the difference in North-West Europe is larger than in Southern Europe, which is the statement that we wanted to make.*

l. 14: Can you add a reference from which you draw this conclusion?

*This can be concluded from Fig. S1 (added to text)*

Can you add a plot or table showing the interannual variability for all three scenarios?

*See figure S1. We did not include further pictures except for the pdfs in the supplementary material, since there are already many plots. But we took the point of adding significance information. Note that we have two climate scenarios and a reference run for the present-day climate. (RLE\_ERA is not a scenario!).*

l. 18-21 : 'Also the seasonal cycle is weaker.....'

The formulation of this conclusion is from my point of view misleading as it appears to be valid for all of Europe even though only results for two stations are shown in Fig.4. I am also not sure if results from Madrid can be extrapolated to all of 'Southern Europe'. Again, I think it would be helpful to know the interannual variability for interpreting the results.

*Results were analysed at more location, as stated more clearly now in Section 2.3, but could not be shown all to keep the paper concise.*

Differences in number of wet days: Is the number of wet days related to e.g. the number of frontal passages simulated differently in each global model? Or to the duration of cyclones over Central Europe, a feature that can be simulated differently in the global models?

*There can be many reasons for the apparent difference in number of wet days. Different circulation, different parameterization of convective processes, different parameterization for the formation and release of precipitation, etc. They all contribute to a different number of wet days.*

### Section 4.1 :

p.12259, l. 20: (Figs 7, S1): (Figs 7 and FigXXXX/Tab.XXXX in S(supplement)1)

?

Fig.7: I would find it convenient if the figure caption explained as precisely as the text in section 4.1. what I can see in the figure (e.g. 'June-July-August' instead of 'summer', 'average daily maximum' instead of 'average O<sub>3</sub> summer maximum'). Units are also missing.

*Captions are improved*

I assume surface values were analysed? This should be mentioned somewhere!

*This is now made explicit in section 2.3.*

p.12260, l3: presumably not only southern Europe, SOA formation should contribute to the PM load over large parts of Europe

*The reference to southern Europe is now only made for dust, for SOA references to the literature were added (Gelencsér et al 2007, Bergström et al 2012).*

Fig.8: Units are missing. The figure caption is confusing. Please write more precisely which differences exactly I can see in each panel.

*Caption is improved*

p.12260, l19: 'Over sea....' This sentence seems incomplete. Or at least I don't

C4365

understand its meaning.  
*Sentence has been rephrased.*

p.12260, l27: The patterns of differences...: Have you analysed a correlation between the number of calm days and PM concentrations?  
*No, but we have analysed correlations between daily average PM concentrations and daily average wind speeds, both using observations and RLE-results (Manders et al 2011), this is now added to the text.*

p.12261, l22: 'magnitude' instead of 'amount' ?

*Sentence adapted*

p.12261, l38, 'nex' -! 'next'

*Changed*

## **Section 4.2**

p.12263, l13: 'The reason is thats' -! 'The reason is that...'

*Changed*

l14: 'formation' rather than 'concentration' ?

*It favours higher concentrations through more formation. Changed concentration to concentrations.*

How do you know that O<sub>3</sub> formation is VOC limited in the Netherlands and NO<sub>x</sub> limited on the Iberian peninsula? As far as I understand, the authors assume that nitrate titration leads to a destruction of O<sub>3</sub> concentrations in Madrid - then they should clearly write that this is an assumption and not a result from their analysis.

*It is now explicitly stated in the text that these are assumptions*

How do you come to the conclusion that wind speeds are higher for temperatures around 12\_C?

*This was found in analysis of the meteorology, but this picture is not shown in the paper. This is now indicated in the text.*

How well can you simulate a change in the mixing height with a very coarse (vertical) model resolution and only 5 model layers? Which parts of the results presented on p.12264 are based on assumptions and which parts are based on the analysis of you simulations?

*RACMO2 has 40 pressure levels in the vertical, which gives a good resolution of the mixing layer. The mixing layer height as modeled by RACMO2 is used by LOTOS-EUROS. LOTOS-EUROS has dynamical layers, the lowest dynamical layer being the mixing layer, which is placed above a 25m surface layer to avoid too strong dilution of surface emissions. The upper layers are reservoir layers. In this way, concentrations can be modeled quite reliably in a very efficient way, as appears from model intercomparison studies and validation (see references in Model section). All results that are used as an argument in the speculation are a result of the analysis of the model results, speculations are indicated by 'may'.*

'In Vredepeel, ammonium and nitrate concentrations are higher' -! are higher than what?

*Higher than other analysis locations. Added to the text.*

Seasonal cycle of PM: How is this related to the seasonal cycle of emissions?

And to what extent is it related to meteorology?

*We have taken the emissions as 'given' and focus on changes in the seasonal cycle due to climate change. The effect is different per location and depends also the local composition of PM, as indicated in the text. In fact, the seasonal cycle in emissions is indirectly determined by meteorology but at present average annual profiles are used, uniform over the domain per emission source (e.g. road transport, weak seasonal cycle, domestic heating, stronger seasonal cycle). This discussion goes beyond the scope of this paper.*

## **Section 5 :**

'Two long-term...' -! 'Three long-term...' (ERA, ECHAM, MICRO)?

*The ERA simulation is not as long-term as the ECHAM and MIROC simulations, and is therefore treated separately. This is now made more explicit in the text.*

'Changes for PM10 are smaller than the interannual variability.' What exactly is

the conclusion? Differences are (statistically) not significant?  
*Significance is now indicated in the text.*

p.12266, l.23: 'fot' –! 'for'  
*changed*

How do you come to the conclusion that ECHAM5 and MICRO are well-performing?  
*Added a reference to literature in which climate models are compared and this conclusion was drawn (Van Ulden and van Oldenborgh 2006)*

p.12268, l.14: '...between of meteorology' –! '...between meteorology'  
*changed*

It would be very helpful for the reader if the panels in all figures would be labelled with a), b), c) and it might be helpful to refer to the figures and corresponding panels in the text (e.g. Fig 3a) instead of Fig.3)  
Figure S1: This is the only figure showing standard deviations to indicate the interannual variability. However, the quality of this figure is bad and should be improved (even though it is only part of the supplement). The labels are too small, the colours of the bars are too similar, standard deviations for the black bars are almost not visible.

*Figure labeling and figure S1 have been improved.*