Dear editor, colleagues,

Below we address the reviewer comments and questions raised during the open discussion of the paper "Dynamic model evaluation for secondary inorganic aerosol and its precursors over Europe between 1990 and 2009".
We would like to thank the reviewers for their comments. We think that their comments have improved the paper considerably.
We have listed all reviewer comments below. Answers are provided in Italics.

1. To Anonymous Reviewer#1:

General comments:

(1) One minor comment I have is the inconsistent use and spelling out of acronyms throughout. For example, LOTOS-EUROS is not spelled out until page 4651 and IIASA is not spelled out until page 4654; both acronyms are used earlier in the paper. Other acronyms, such as the RAINS and GAINS models (page 4654) are not spelled out. Please review the entire manuscript for consistency, and spell out all acronyms that might not be common for an international audience.
- *The manuscript has been reviewed for acronyms, which are all spelled out now.*

(2) I am somewhat concerned about the comparisons across the three time ranges: 1990-2009, 1995-2009, and 2000-2009. These time ranges can have very different numbers of sites located in very different regions of Europe, and it is not clear to me how much of the differences in trends or model performance is due to changes in emissions and how much is due to the available data in a given time frame. The 2000-2009 has the additional complication of, perhaps, not being long enough to infer statistically significant trends.
- *We would like to thank the reviewer for this comment. The sections 3.1.2 and 3.2 of the "Results" part of the manuscript have been revised with special attention to the concerns named above. In the course of the latter two Tables (Table 5 (see comment (11)) and Table S2 (see below)) have been revised/added in which the evaluation statistics/observed and modelled trends for different time ranges are presented for the same set of stations per component. Further details on the revision are partly given below following the "Specific comments".*

Table S2. Number of stations and derived observed and modelled absolute (µg m^{-3} a^{-1}) and relative (% a^{-1}) median trends for SO_2, NO_2 and SO_4^{2-} at those stations that passed the selection criteria presented in Section 2.2.1 for all three time periods.

Period	Evaluation	SO_2	NO_2	SO_4
all	number of stations	15	33	11
1990-2009	Observed abs. median trend	-0.41	-0.37	-0.11
	Modelled abs. median trend	-0.36	-0.48	-0.06
	Observed rel. median trend	-4.52	-1.89	-3.10
	Modelled rel. median trend	-4.16	-2.46	-2.33
1995-2009	Observed abs. median trend	-0.28	-0.34	-0.09
	Modelled abs. median trend	-0.19	-0.58	-0.05
	Observed rel. median trend	-5.16	-1.69	-3.37
	Modelled rel. median trend	-4.83	-2.70	-2.33
2000-2009	Observed abs. median trend	-0.10	-0.14	-0.05
	Modelled abs. median trend	-0.09	-0.35	-0.03
	Observed rel. median trend	-4.46	-1.07	-3.07
	Modelled rel. median trend	-4.69	-2.40	-2.10

Specific comments:

(3) Page 4654 line 12: The authors use the word “pops” here. Do they mean “persistent organic pollutants (POPs)?”
- The paragraph containing line 12 has been rewritten as Reviewer#2 has asked for a more detailed description of the lateral boundary conditions (→ See response to Reviewer#2, comment (4)).

(4) Page 4654-4655 and Figure 1: Showing emissions normalized to 1990 is fine, but it might be helpful to include the actual emissions in tons, perhaps in the supplemental section. It is not clear to me how substantial emissions from shipping are compared to total emissions in this region. Total emissions are generally decreasing (Figure 1a) but shipping emissions behave somewhat differently (Figure 1b).
- A table containing the total emissions (land-based and shipping emissions separately) of the relevant components has been added to the Supplement (Table S1) and is referred to in the text in Section 2.1.2.

Table S1. Annual total emissions of SO_2, NO_x and NH_3 for the years 1990, 1995, 2000, 2005 and 2010 of the EU-27+ member States and from International shipping (sum of annual total emissions from the Baltic Sea, the North-East Atlantic Ocean, the North Sea, the Mediterranean Sea and the Black Sea).

Emissions (kton a^{-1})	Year	EU-27+ member States	International shipping
SO_2	1990	24518	1931
	1995	15710	2190
	2000	10621	2399
	2005	8130	2714
	2010	4393	2549
NO_x	1990	17393	2774
	1995	16254	3146
	2000	12416	3510
	2005	11352	3876
	2010	8898	4271
NH_3	1990	4725	-
	1995	4133	-
	2000	4126	-
	2005	3933	-
	2010	3820	-

(5) Also, the section on the source apportionment module needs to be clarified. What does the 10 kton threshold apply to? A single source, such as a large power plant? An entire source sector, such as all power plants? Or something else?
- *To better describe the set-up of the performed source apportionment exercise we have put these additional sentences in the text for further explanation:*
"Note that the 10 ktons are chosen arbitrary as tracking any other fraction of the emissions would give the same results due to the labelling approach used (Kranenburg et al., 2013). In practice, for each year the 10 ktons are normalized to the total emissions. The obtained fraction is applied to all emissions in the country and allocated to the respective label."

(6) Page 4657: I want to make sure I understand the three selection criteria. I interpret criterion 1 to mean that a complete year means 75% valid data. I interpret criterion 2 to mean that 80% of the 10, 15, or 20 years need to be considered complete to be included in the analysis. Is this correct?
- *Yes, that is correct.*

(7) Also, I don't understand why the highest variability is expected at the beginning of each time period; why is this? Couldn't an "unusual" year occur somewhere in the middle or end of the range?
- *The reason why we needed the first year of each time period to be covered is that we also address relative trends. Hence, we have adapted the corresponding sentence in the manuscript to:*
"As we also address relative trends within this study we consider it important to have the first year of each time period covered. Hence, only stations that could provide the requested 75% data coverage for the first year of the time period were included in the corresponding subset."

(8) Finally, I am a little uncomfortable with the visual screening (criterion 3). It seems to me that "outliers" can occur, and I am not sure why sites showing "constant values" over time are necessarily wrong. This visual screening removes more than half of the available SO2 sites.
- *We understand the concern of the reviewer that the visual screening removes a large number of stations from the dataset. However, the examination of the observed time series has revealed the necessity of a visual screening of the data. In fact, we have performed a very basic screening of the data for obvious errors like e.g. a station that shows constant values (e.g. 2.57 µg/m^3) for each day/hour over a time period of months or years. Two examples of erroneous or not-usable measurement time series are presented below in the Figure 1 and Figure 2 showing SO_2 concentration observations (red dots) compared to modelled concentrations (black line). For comparison Figure 3 shows observed and modelled SO_2 concentrations at a station that has passed the visual screening.*

Figure 1

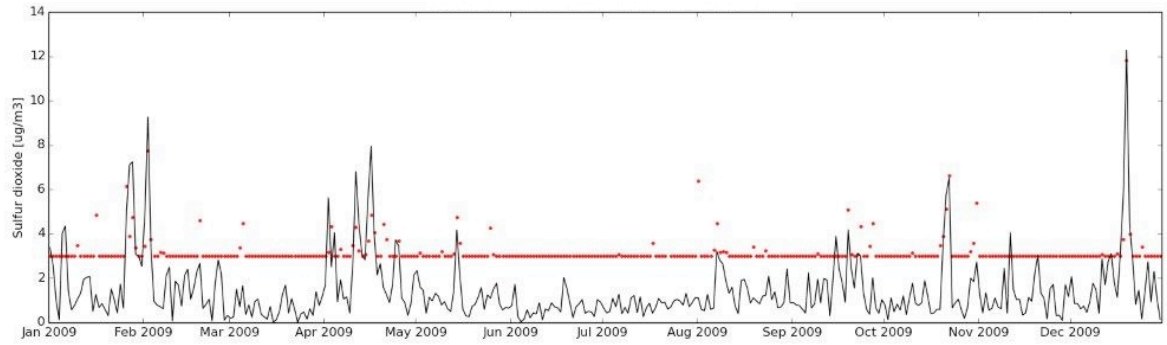


Figure 2

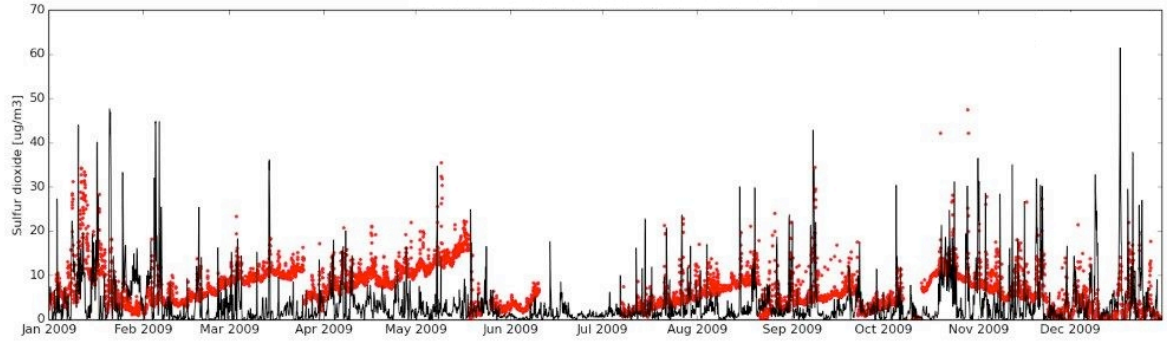
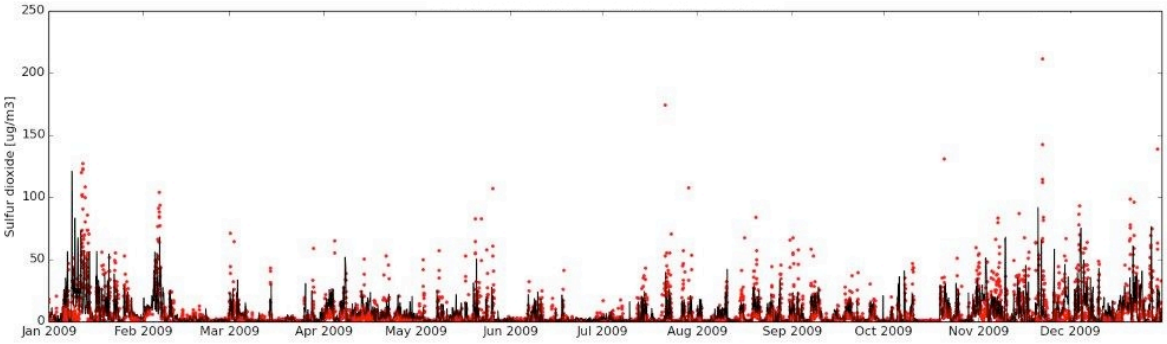


Figure 3



(9) Page 4658: There are so few TNO3 and TNH4 sites across the model domain. This limitation needs to be stressed, and I don't know how this impacts the results pertaining these two parameters. This comment also applies to the discussion on page 4665 (lines 11-21).

- *We have adapted the text by emphasizing that the results for TNO_3 and TNH_4 rely on a small set of stations with most of them located in northern Europe at the corresponding text passages throughout the manuscript.*

(10) Page 4661-4662: Does RACMO2 assimilate surface meteorological data for nudging? If so, were any of the surface meteorological stations used? If they were, it would not be surprising that there was good correlation for some of the meteorological fields, since this would not be an independent verification of the model.

- There is no nudging performed in the RACMO model runs. We have added the following sentence to Section 2.1.2.:

"Nudging of meteorological data has not been performed for the model runs and RACMO2 is only constrained by the lateral boundary conditions."

(11) Page 4664 lines 15-23: The authors discuss different correlation coefficients in different time periods, but I am not sure if these differences are statistically significant. This is especially true for SO4, which changed from 05 to 0.56. A test of significance might be needed here. Also, the reasons for this apparent change in model performance seems to border on opinion. I think the last sentence in this paragraph should be re-written.

- As mentioned above, section 3.1.2 has been revised following the concerns of the reviewer. Please note that during the revision process we have identified a copy-paste error within Tables 4 and 5 (see below) of the manuscript. The numbers of the correlation coefficients have changed and the text has been adapted accordingly. We very much regret this inconvenience.

- The last sentence in the paragraph has been re-written to: "We speculate that the models' underestimation of SO_2 and SO_4^{2-} concentrations in the 1990s could be connected to the lack of a good representation of the change in emission structures in the power sector in eastern and parts of central Europe as a consequence of the fall of the Berlin wall and political changes associated with the liberalisation of the Eastern Bloc's authoritarian systems as discussed below."

Table 4. Statistical comparison between measured and modelled concentrations using daily observations. The number of considered stations, mean correlation, observed mean, RMSE and bias are given for each component and each time period.

Period	Evaluation	SO_2	NO_2	SO_4	TNO_3	TNH_4
1990-2009	number of stations	23	37	15	9	7
	mean correlation	0.60	0.65	0.46	0.46	0.48
	observed mean (µg/m^3)	3.86	15.97	2.77	0.56	1.35
	RMSE (µg/m^3)	6.01	8.66	2.86	0.61	1.21
	Bias (µg/m^3)	-0.44	-2.43	-0.88	0.04	0.03
1995-2009	number of stations	40	64	22	9	8
	mean correlation	0.58	0.62	0.40	0.44	0.44
	observed mean (µg/m^3)	4.00	14.19	2.46	0.46	1.17
	RMSE (µg/m^3)	6.49	8.58	2.27	0.54	1.05
	Bias (µg/m^3)	-0.67	-2.58	-0.66	0.12	0.03
2000-2009	number of stations	60	112	28	16	15
	mean correlation	0.45	0.61	0.40	0.48	0.40
	observed mean (µg/m^3)	3.34	14.12	2.16	0.60	1.38
	RMSE (µg/m^3)	5.01	9.37	1.95	0.6	1.18
	Bias (µg/m^3)	-0.69	-3.77	-0.58	0.12	0.21

Table 5. Statistical comparison between measured and modelled concentrations using daily observations at those stations that passed the selection criteria presented in Section 2.2.1 for the 1990-2009 and 2000-2009 time period. The number of considered stations, mean correlation, observed mean, RMSE and bias are given for each component.

Period	Evaluation	SO_2	NO_2	SO_4	TNO_3	TNH_4
all	number of stations	15	33	11	4	3
1990-2009	mean correlation	0.62	0.67	0.47	0.49	0.54
	observed mean ($\mu g/m^3$)	4.19	17.05	2.53	0.40	0.77
	RMSE ($\mu g/m^3$)	6.15	8.93	2.75	0.54	0.77
	Bias ($\mu g/m^3$)	-0.57	-2.53	-0.85	0.12	0.07
2000-2009	mean correlation	0.52	0.67	0.40	0.46	0.47
	observed mean ($\mu g/m^3$)	2.16	15.23	1.85	0.38	0.66
	RMSE ($\mu g/m^3$)	2.48	8.09	1.75	0.49	0.70
	Bias ($\mu g/m^3$)	-0.03	-2.87	-0.49	0.10	0.11

(12) Bottom of page 4664 to top of page 4665: Could the underestimation of air concentrations of SO4 be due to underestimating SO2-to-SO4 conversion, or possibly overestimating deposition of SO4? This comment could also apply to the discussion in the second paragraph on page 4674.

- *This could indeed be a reason for the underestimation of SO_4^{2-} concentrations. The following sentence has been added to the corresponding paragraph on page 4665:*
"It is unclear if the underestimation is connected to a lack of SO_2-to-SO_4^{2-} conversion or an overestimation of turbulent mixing leading to too high deposition and vertical mixing."
And the following sentence has been added to the corresponding paragraph on page 4674:
"It has not yet been solved if the underestimation is induced by a lack of SO_2-to-SO_4^{2-} conversion or too high deposition and vertical mixing due to an overestimation of turbulent mixing."

(13) Bottom of page 4665 to top of page 4666: The authors state that the model results compare more favorably in the later years, and that the model predicts N compounds better than S compounds. However, some discussion on statistical significance may be warranted here.

- *This comparison is not done anymore. See response to comment (11).*

(14) Page 4666 lines 4-11 and Figure 4: Does Figure 4 include trends that are not significant? I am not sure both Figure 4 and 5 are needed. I prefer what is shown in Figure 5, as there is an indication of trend significance, as well as magnitude. If the authors remove Figure 4, please modify any subsequent text that refers to this figure.

- *We agree that Figure 5 includes the more distinct information on the observed and modelled trends. Figure 4 has been removed and the corresponding text passages have been adapted*.

(15) Page 4667-4668: I am concerned about reading too much into trends based on only 10 years of data. It only takes a single "unusual" year to confound trend estimates for such a short period. This section could be re-written, it is somewhat confusing. At least there is some mention of significant versus non-significant trends.
- As suggested by the reviewer the Sections 3.2.1 and 3.2.2 have been re-written. In the line with removing Figure 4 the sections on the observed and modelled trends have been re-structured focussing on the former Figure 5 (now Figure 4), Table 6 and Table S2.
- Furthermore, we have added the following paragraph to the text:
"For the time period 2000-2009 Figure 4 shows that the observed trends are non-significant at the majority of stations for all considered components. We would like to stress that this does not necessarily imply that there is no trend present in the data of these stations for 2000-2009 but 10 years may be too short to infer statistically significant trends."

(16) Page 4670-4671: I don't think the first paragraph in section 3.3 is needed, it doesn't add to the paper. Figure 6 is also not needed, since trends in these species are presented in earlier figures.
- Figure 6 and the corresponding paragraph in Section 3.3 have been removed.

(17) Page 4671: The section on SIA formation needs some clarification. What does the 10 kton threshold apply to? See my earlier comment regarding page 4655.
- See reply to comment No. (5) above.

2. To Anonymous Reviewer#2:

General comments:

(1) The first reviewer has two main comments to the paper regarding the statistical significance of trends and the three different time periods used and the possible impacts on the results. I agree with the first reviewer that especially the impact of change in numbers and locations of sites needs to be addressed before the final version of the paper is published.
- See reply to Reviewer#1

(2) I also wonder how it affects the overall conclusions that some of the trend analyses are based on data from e.g. mainly Northern Europe. Shouldn't you then compare it to the trend of the emissions in this region? Even if most of the components have a long life time, the emissions in the specific region will have a significant impact on the level of SIA within the region.

- We agree with the concern of the reviewer. Now that Figure 6 and the associated text have been removed, this comparison is not made anymore. Now the section is totally focussed on the labelling results.

(3) The spatial distribution of the MACC emissions are as far as I understand kept constant (based on the distribution for 2005). It is mentioned only briefly that this could impact the results. However, in order to give a more qualified estimate on how this will influence the simulated trends it would be usefull to include more details on this. This could e.g. be done by comparing the MACC emissions with the EMEP emissions for 1990 (assuming that the EMEP emissions include spatial changes : : :)? This has perhaps already been done during the evaluation of the MACC emissions and a reference to this work and a few lines on the conclusion could be sufficient for the paper.
- We have added the following sentences to the text in section 2.1.2 concerning the choice of using the MACC 2005 emission distribution:
"Although EMEP provides information on changes in the emission distribution from the early 1990s onwards we used the TNO MACC (Denier van der Gon et al., 2010; Pouliot et al., 2012) spatial distribution of emissions for the year 2005 for the entire time period of investigation. We believe that current emission allocation proxies are more reliable than the ones used in the 1990s. Furthermore, the EMEP emission information for the 1990s is only available on a resolution of 150x150 km^2, which is much lower than the resolution of the applied MACC distribution and is therefore not expected to provide an improvement."

(4) On page 4654 it is mentioned that trends are not included in the boundary conditions. Please elaborate a bit on the possible consequences of this in the in the discussion. Maybe other studies have investigated this in more detail? Please also mention where the used background concentrations are taken from.
- The following paragraph was added to the discussion section:
"In the present model set-up, trends in boundary conditions were not considered, although background concentrations are expected to change from 1990 to 2009. We believe that the impact of using time-variant boundary conditions would be most relevant for O_3 levels, which also affect the formation of SIA. Time-variant lateral boundary conditions could be extracted from a global model simulation. Hogrefe et al. (2011) studied the uncertainties associated with chemical boundary conditions from a global model, showing that the representation of the interannual variability of O_3 concentrations was improved when time-variant boundary conditions were used. However, biases in the global simulations significantly affected the O_3 simulations throughout the modelling domain with adverse impact on the simulated O_3 trends."
- Furthermore, the paragraph on the description of the used lateral boundary conditions in Section 2.1.2. has been extended:
"Lateral boundary conditions in LOTOS-EUROS were taken from climatological background concentrations for gases and aerosols. For a number of components we follow the EMEP method (Simpson et al., 2003) based on measured data, in which simple functions were derived to match the observed distributions. Some aerosol species are set to constant at the boundaries. NH3 boundary conditions are

neglected. SO_4^{2-} is assumed to be fully neutralised by ammonium. Nitrate values are assumed to be included in those of HNO_3 (derived following Simpson et al. (2003)) and are zero as well. The climatology fields did not include windblown dust going back to 1990. Hence, dust from e.g., wind erosion, agricultural land management and resuspension by road transport has been neglected, as it does not contribute to the here investigated substances. For ozone (O_3) we have used the climatological dataset by Logan (1998), derived from O_3 sonde data. For the interpretation of the model results we need to keep in mind that there are no trends in boundary conditions considered over the investigated 20 year period."

Specific comments:

(5) P. 4652 line ca 7. It seems like the models has relatively few vertical layers. Please give an indication on the approximate height of the layers.
- *The information on the vertical layers and their approximate height is given in Section 2.1.1 on Page 4652, line 5 and following:*
"The vertical grid is based on terrain following vertical coordinates and extends to 3.5 km above sea level. The model uses a dynamic mixing layer approach to determine the vertical structure, i.e. the vertical layers vary in space and time. The layer on top of a 25 m surface layer follows the mixing layer height, which is obtained from the meteorological input data that is used to force the model. The height of the two reservoir layers is determined by the difference between model top at 3.5 km and mixing layer height."

(6) The source apportionment module needs to be described in more detail.
- *A more detailed description has been added to Section 2.1.1.:*
"LOTOS-EUROS includes a source apportionment module, which enables tracking the source contribution of a set of sources through the model system. The emissions can be categorized in several source categories (e.g. countries or sector) and labelled accordingly before the model is run. The total concentration of each substance for each time step and in each grid cell is modelled as before, but next to this, the fractional contribution of each label to every species is calculated. During each process, the new fractional contribution of each label is defined by calculating a weighted average of the fractions before the process and the concentration change during the process. The labelling routine is only implemented for chemically active tracers containing C, S or N (reduced and oxidized) atoms, as these are conserved and traceable. The source apportionment module is extensively described in Kranenburg et al. (2013)."
- *Furthermore, the set-up of the source apportionment exercise within this study has been better described (see reply to Reviewer#1, under (5))*

Technical corrections:

(7) P. 4652 line 19. Delete "()" around "Erisman et al."
- *Has been corrected.*

(8) There seems to be a general lack of "commas" in the text. Pleas review the manuscript with this in mind.
- *A review of the text has been performed.*

(9) P. 4662, line 10. Make it more clear in the text that Fig 2b is for German stations only, not all of Europe.
- *The text has been adapted.*

(10) P. 4665, line 12. "Other than for : : :" – reformulate, bad English language.
- *This formulation has been removed from the text.*

(11) Fig. 4 – in the current version of the manuscript it is very hard to see the different lines in this plot.
- *Figure 4 has been removed from the manuscript (see reply to Reviewer#1, under (14)).*