We thank the reviewer for the careful reading of the paper. Please find below the point-by-point answers to the reviewer’s comments:

Specific comments:

1) Line 23 (p1): BP and DP are confusing in the abstract.
   We rephrased the sentence in the abstract, by removing the abbreviation BP.

2) Line 25 (p1): It is unclear how the <50%> uncertainty is quantified in the results.
   “As explained in the 2nd paragraph of section 3.2, based on the statistics of the comparison between the model results and the measurements of TP and PO4 aerosol concentrations the NMB is found to be about -67%. This sentence has been rephrased to “…compared with available observations, indicating however an uncertainty of about 70% on current knowledge of the sources that drive P atmospheric cycle’.

3) Line 26-29 (p1): It is better to give the dissolution fluxes, rather than one percentage.
   We have chosen to provide the percentage change in order to show a quantitative comparison with the present day flux for which the absolute number is given earlier in line 21. The absolute fluxes for PAST and FUTURE simulations can be easily derived from these numbers. Therefore, no change has been made to this sentence.

4) Line 30 (p1): "dissolution flux of P" and "P mobilization flux" are confusing in the abstract.
   To avoid confusion we now use only the term ‘P solubilisation flux’.

5) Line 9 (p5): Using a coarse-resolution model to get the horizontal distributions of P concentration and P deposition can lead to substantial biases in the model-observation comparison, which should be considered.
   We re-run the model in its finer horizontal resolution (i.e. 3° x 2° in longitude by latitude) and updated the manuscript and the respective figures.

6) Line 10 (p5): It seems that the model is only run with meteorology for one year (2008). In this case, the question is that there is inconsistency if the observations are derived for other years. It means that the interannual variation of P emissions from mineral dust and sea-spray (related to wind) and the episodic transport of P in the atmosphere (related to wind and wet precipitation) cannot be represented in this study. I expect that the model can be run for more years to get an unbiased estimation of P emissions from mineral dust and sea-spray.
   Our study has to be seen as a ‘climatological’ rather than a year specific study. This is justified by the large uncertainty associated with the sources, the dissolution kinetics and the deposition parameterisation. There is indeed a year-to-year variability in the emissions and atmospheric transport that could be addressed by performing multiple year simulations. However, such simulations are computational intensive and the added value to the main objective of the paper is very small. Furthermore, a model-to-observations comparison for a specific year limits the number of available measurements and thus the statistical significance of the comparison. However, to satisfy the reviewer’s request, we run our model for a longer period that covers a significant number of available measurements to make year specific comparisons. This has been the main reason for the delay in the revision of our manuscript.

7) Equation 1: I am curious to know how the authors get the fraction of P in the emitted dust. The global soil mineralogy datasets are developed for the clay and silt fractions of soils. Therefore, according to my knowledge, the mineralogy of soil is different from the mineralogy of dust. Please explain how the mineralogy of soil is transferred to the mineralogy of emitted dust in the model.
   For the present study, we applied the available P-content distribution of soils on dust emissions making no distinctions between soil and dust particles. We estimated the P-content in dust based on the database by Nickovic et al. (2012) that only provides the mass fraction of P in soils. The emissions of dust were calculated as a function of wind velocities at a height of 10 m over dust source locations and soil types (Tegen et al., 2002) for two modes accumulation and coarse (van Noije et al., 2014) The resulted P emissions were scaled to a global mean of 880 ppm per weight as observed by (Zamora et al., 2013). This
is explained in section 2.1.1. Thus, as in most relevant studies but one (Perlwitz et al. (2015)), airborne dust particle emissions are assumed to have the same mineralogy with that of the soil from which they originate.

We have further added the relevant following discussion in section 2.1.1.

“Although in most relevant modelling studies airborne P-containing dust particle emissions are assumed to have an average P content of 720 ppm (Mahowald et al., 2008; Wang et al., 2014; Brahney et al., 2015), in the atmosphere due to transport, ageing and deposition processes the overall mineralogy may change the chemical composition and size of dust aerosol population. In a recent iron modelling study however (Perlwitz et al., 2015), a significant effort has been made to model the mineral composition of dust considering the differences from the original soil composition. Perlwitz et al. (2015) have found significant overestimate (a factor of 10-30) mainly in the fine aerosol emissions that are the smallest part of dust emissions (e.g. about 7% of the total emissions in our model) and an underestimate in the larger particles emissions both for total dust and for individual minerals when the mineralogy of dust aerosol is assumed to be the same as that of the soil. However for the present study, we did not account for different P content for dust particles in the fine and the coarse mode, since the global soil mineralogy dataset used (Nickovic et al., 2012) does not provide any information of P content in silt and clay soil particles separately. Note also, that recent studies indicate that dust super-coarse particles can be very important for the biogeochemistry over land, since they can represent the dominant fraction of dust close to source regions (Lawrence and Neff, 2009) (Neff et al., 2013). (Brahney et al., 2015) modelling study that focused on the atmospheric phosphorus deposition over global alpine lakes, based on Neff et al. (2013) observations, estimated that only 10% of the mass that travels in the atmosphere is within the <10 μm size fraction. In our study we do not account for super-coarse dust particles because due to their short atmospheric lifetime, they are emitted and deposited in the same model grid box (Brahney et al., 2015). This omission is not expected to have significant impact on our results, since the present work is focused on the P-solubilisation mechanisms occurring via atmospheric long-transport mixing and on the bioavailable P deposition over the marine environment. “

8) Line 29 (p5): I am curious to know what aerosol scheme is used to treat the size evolution of P containing particles in this model. Is it following a modal method or a sectional method? I expect that the size (0.34 um or 1.75 um) is not fixed in the model. If the size was fixed, I would have to say that the authors should take a more advanced scheme before the paper can be published.

The model follows the modal approach to represent aerosol sizes. Specifically, the model configuration used for this study is focused on the chemical interactions of dust particles and the acid P-solubilisation of apatite minerals and, thus, on chemistry and on aerosol mass simulations. The strong point of our model is exactly the representation of atmospheric chemistry in all phases, gas, aqueous and aerosol phase. Overall 32 model P-containing aerosol tracers are used to represent phosphorus in the model of different sizes and solubilities. In TM4-ECPL, different sources emit P-containing aerosols of different sizes represented by lognormal distributions as summarized in the Table below.

Table : P-sources and aerosol size lognormal modes taken into account in TM4-ECPL model.

<table>
<thead>
<tr>
<th>P-sources</th>
<th>Lognormal modes - Dry number median radii in um (sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dust</td>
<td>0.34* (1.59)</td>
</tr>
<tr>
<td>Combustion</td>
<td>0.04 (1.8)</td>
</tr>
<tr>
<td>Primary biological aerosol particles</td>
<td>1.50 (monodisperse)</td>
</tr>
<tr>
<td>Volcanoes</td>
<td>0.04 (1.59)</td>
</tr>
<tr>
<td>Sea-spray</td>
<td>0.09 (1.59)</td>
</tr>
</tbody>
</table>

(*) Mass mean radii

For each aerosol mode and source (Figure 1 of the manuscript) the model accounts for total P, phosphate, insoluble and soluble OP, for the dust source it also accounts for the two P-containing minerals (fluoroapatite and hydroxyapatite) as described in the section 2.1.1. These are individually transported, aged and deposited in the atmosphere. The ‘dry’ aerosol hygroscopic growth in the model is treated as a function of ambient relative humidity and the composition of soluble aerosol components based on experimental work by Gerber (1985; 1988) and this uptake of water on aerosols changes the particle size. In addition, during atmospheric transport there are major changes in the size distribution of aerosols as a
consequence of the removal of larger particles due to gravitational settling. The P-containing aerosols follow the same parameterizations, hygroscopic growth and removal processes are assumed to affect the mass median radius (i.e. size).

Note also that although our model does not include the most sophisticated aerosol scheme (uses the modal lognormal distribution instead of the more compute demanding sectional approach), TM4-ECPL is the first CTM model that takes into account the P-solubilisation due to atmospheric acidity, instead of taking a constant solubility fraction. In addition, TM4-ECPL accounts for all known sources of atmospheric phosphorus including the primary biological particles. These strongly innovative aspects of our study deserve publication.

Relevant text has been added in section 2 and before section 2.1. “To represent phosphorus in the model, overall 32 model P-containing aerosol tracers are used of different sizes and solubilities. In TM4-ECPL, different sources emit P-containing aerosols of different sizes represented by lognormal distributions as outlined in section 2.1. For each aerosol mode and source (Figure 1 of the manuscript) the model accounts for total P, phosphate, insoluble and soluble OP. For the dust source it also accounts for the two P-containing minerals (fluorapatite and hydroxyapatite) as described in the section 2.1.1. These are individually transported, aged and deposited in the atmosphere. The ‘dry’ aerosol hygroscopic growth in the model is treated as a function of ambient relative humidity and the composition of soluble aerosol components based on experimental work by Gerber (1985; 1988) and this uptake of water on aerosols changes the particle size. In addition, during atmospheric transport there are major changes in the size distribution of aerosols as a consequence of the removal of larger particles due to gravitational settling. The P-containing aerosols follow the same parameterizations, hygroscopic growth and removal processes are assumed to affect the mass median radius (i.e. size).”

9) Line 30 (p5): Please explain how the solubility of P (10%) is derived?

The soluble fraction used in our model is based on the measurements of leachable inorganic phosphorus (LIP) for Saharan soil dust, as presented by Nenes et al. (2011). These authors found that LIP represented up to 10 % of total inorganic P in Saharan soil samples and dry fallout collected during Sahara dust storms before acid-treatment. Moreover, Yang et al. (2013) estimated the labile inorganic P in the top soil on the global scale at about 3.6 Pg-P that corresponds to about 10% of the estimates of total soil P on the global scale 30.6-40.6 Pg-P (Smil, 2000; Wang et al., 2010; Yang et al., 2013). To further investigate uncertainties associated with the soluble fraction of P-containing dust aerosol emissions in our model, an additional simulation has been performed neglecting any soluble fraction on initial emissions.

The above discussion has been added in section 2.1.1.

The comparison of model results to the observations shows no significant change in the performance of the model due to this additional soluble P primary source. On the contrary, our results presented in Figure S4 (of the discussion paper) show significant impact of the secondary P-solubilisation source, increasing the soluble P simulated concentrations. Relevant discussion has been added in the revised version of the manuscript in section 3.2.

“Neglecting the P dissolution definitely degrades the comparisons of model results with observations. On the other hand the results show very small sensitivity to the assumption of soluble fraction of the primary emissions of P. This finding supports the importance of the atmospheric processing of dust for the atmospheric DP cycle as well as the potential underestimate of the DP source in all sensitivity simulations. Such underestimate could be associated with an underestimate in the primary source or in the secondary (atmospheric processing) of DP and deserves further studies. “

10) Line 1-5 (p6): Please list the detailed P/BC mass ratios in the emissions (in the Supplementary Materials). I would expect different ratios for emissions from different types of fuel (e.g. coal, oil and biomass, etc).

As clearly stated in section 2.1.2 lines 2 and 3, for the present work, we used the P/BC combustion emissions factors of Mahowald et al. (2008) that are based on BC aerosol emissions for the fine and coarse mode (i.e. P/BC in fine mode of 0.0029 and P/BC in coarse mode of 0.02; independent of fuel type
However, we know that recently new developments on P combustion emissions are published by Wang et al. (2014) following a more comprehensive approach, accounting for different emission factors per fuel type. That estimate is significantly higher than what is derived based on P/BC emission factors and BC emissions. Therefore, as stated at the end of the first paragraph of section 2.1.2, we have performed an additional present-day sensitivity simulation using the P-combustion emissions developed by Wang et al. (2014) (after personal communication with the author). The results of this simulation are discussed in section 3.2 and plotted in the supplementary figures S4-S6.

We comment on the sensitivity simulations in the manuscript as follows (section 3.2):

“In Figures S4 and S5 (supplement) are also presented the results of sensitivity simulations and the base case simulation with the aerosol observations and dry deposition fluxes, respectively, while figures S6 also show the comparison of the annual cycles of the atmospheric concentrations (TP and PO4) and deposition fluxes (dry and wet deposition), against the TM4-ECPL monthly model results. For cruise measurements over the Atlantic and Pacific Oceans (Baker and Croot, 2010; Martino et al., 2014; Powell et al., 2015), and the global compilation of deposition rates (Vet et al., 2014), the observations are also spatially averaged inside the same model grid box. These comparisons show almost similar performance for all sensitivity simulations but one falling in most cases close to the lower edge of observed concentrations and deposition fluxes. However, taking into account the Wang et al. (2014) P-combustion sources, the model performs better over the land (e.g. for TP concentrations at Corsica; Fig. S4g, and for DP concentrations at the Finokalia monitoring station; Fig. 6b,f,i), indicating that the base simulation underestimates either anthropogenic combustion sources or other natural P sources. Neglecting P dissolution definitely degrades the comparisons of model results with observations. On the other hand the results show very small sensitivity to the assumption of soluble fraction of the primary emissions of P. This finding supports the importance of the atmospheric processing of dust for the atmospheric DP cycle as well as the potential underestimate of the DP source in all sensitivity simulations. Such underestimate could be associated with an underestimate in the primary source or in the secondary (atmospheric processing) of DP and deserves further studies.”

Note however that the present study focuses on the natural emissions of phosphorus and how these are affected by human emitted acidic substances. This study does not fully cover all aspects of phosphorus cycle since many questions remain open for future work.

11) Line 8-11 (p6): Please explain how the initial size distributions of P emissions are treated in the two estimations 1) based on P/BC ratios and 2) based on the new estimation by Wang et al.
   See further our reply to point 21

12) Line 14 (p6): What is "TP coarse aerosol emissions"? Please replace this with a more rigorous term.
   This sentence has been rephrased: ‘BC emissions from anthropogenic combustion in the coarse mode are assumed to be 25% of those in the fine mode (Jacobson and Streets, 2009), while biomass burning emissions in the coarse mode are assumed equal to 20% of those of fine aerosols (Mahowald et al., 2008).’

13) Line 20 (p6): I do not think that this assumption is justified. Carbonaceous aerosol (e.g. BC) is mainly contained in fine particles and thus the ageing via coagulation and condensation is very fast. However, P could be more concentrated in coarse particles (as noted by the authors) and the ageing should be slower.
   We thank the reviewer for pointing us this misleading sentence. In our model simulations, we assumed that the insoluble fraction of organic phosphorus (OP) associated with primary emissions of organic aerosol is converted to dissolved OP (DOP) during atmospheric ageing, based on the (Tsigaridis and Kanakidou, 2003) parameterizations for carbonaceous aerosols that were using a mean mass aerosol size of 0.25 µm and a monolayer thickness of 2.5 nm of the particle surface. For the present study, we take into account the different aerosol sizes (based on the lognormal distributions) of hydrophobic OP aerosols.
to compute their conversion rates to hydrophilic aerosols. Note that based on that parameterisation the conversion rate is reduced with increasing particle size. To avoid further misunderstandings, we rephrase this sentence as “The insoluble fraction of OP associated with combustion emissions can be further converted to soluble OP (DOP) during atmospheric ageing, using the ageing parameterization for primary hydrophobic organic aerosols in the model (Tsigaridis and Kanakidou, 2003; Tsigaridis et al., 2006), but for the respective size and lognormal distribution of OP aerosols with the larger particles experiencing the smallest conversion rates.”

14) Line 23 (p7): Please compare the global total sea-spray emissions output from TM4-ECPL to other estimations before using the data. It seems that the super-coarse mode of sea-spray has not been considered in this model, and it might lead to significant underestimation of P concentrations when comparing with cruise measurements over the oceans.

Sea-spray emissions are computed by the model using the parameterization developed and tested by Vignati et al. (2010) and also used in Myriokefalitakis et al. (2010). As in most models, these emissions are distributed into fine and coarse modes, super coarse sea-salt particles are not considered. Emissions are driven by the model’s meteorology. For the year 2008 (base year for this study), TM4-ECPL model calculates a total of 7278 Tg yr⁻¹ of sea-salt emissions from which 36 Tg yr⁻¹ are fine mode particles. These numbers compare well (by 10% lower) with the AEROCOM recommendation of 7925 Tg yr⁻¹ by Dentener et al. (2006) and are within the range of 2272-12462 Tg yr⁻¹ computed by (Tsigaridis et al., 2013) using different parameterisations in the GISS modelE. Relevant discussion is now added in section 2.1.4 right after the first sentence.

“Sea-spray emissions are driven by the model’s meteorology and for the year 2008 the model calculates a total of about 8284 Tg yr⁻¹ of sea-salt emissions (of which 41 Tg yr⁻¹ are in the fine mode). These numbers compare well with the AEROCOM recommendation of 7925 Tg yr⁻¹ by Dentener et al. (2006) and are within the range of 2272-12462 Tg yr⁻¹ computed by Tsigaridis et al. (2013) using several different parameterisations. Note that our sea-salt source estimation is however much lower than the one used in the modelling study by Wang et al. (2014) (i.e. 25300 Tg yr⁻¹), since super coarse sea-salt particles are not considered in the current parameterization.”

However, this omission can explain some discrepancies between model results and observations only when these later concern bulk aerosols in oceanic regions, so they could include super-coarse particles, which is the case for wet or dry deposition samples. In many cases, aerosol samples have been collected with inlet devices that enable collection of specific fractions of aerosols that cut super-coarse particles. When bulk aerosols have been collected, then the presence of super-coarse aerosols might introduce discrepancies between model results and observations. To distinguish such differences, we separated the bulk aerosol observations from the size segregated ones in section 3.2 and in Figures S4, S6.

And at the end of the section 2.1.4 we have also added the following comment: “The omission of the super coarse sea salt aerosol might affect our estimates of P deposition to the ocean. Brahney et al (2015) evaluated this source at 0.0046 Tg-P yr⁻¹, an amount that introduces a 3% underestimate to the here calculated present-day P deposition flux to the oceans.”

15) Line 30-35 (p7): The authors are right to use the surface concentrations of Na and PO4. I am curious to know how deep is defined as the surface water for Na and PO4. It is better to clarify this, although it is not easy to make sure that they are consistent.

We use the consistent dataset for surface concentrations of Na and PO4 from the LEVITUS94 Ocean Climatology database that we have downloaded from the webpage that is provided in the manuscript and is active (e.g. http://iridl.ldeo.columbia.edu/SOURCES/LEVITUS94/ANNUAL/PO4/). Specifically the surface concentrations correspond to the data labelled as for 0m depth, while the next depth with available data is 10m. The same depth has been chosen for the salinity distribution that was used to derive Na concentrations. We now specify this in the manuscript (section 2.1.4).

16) Line 14 (p8): I am not against what is done here, but I would like to make it clearer that we should always be very careful when taking this kind of assumption. For example, the authors took the sulphur emissions from Andres and Kasgnoe (1998) while adopting the size distribution proposed by Dentener et al. (2006). Are they consistent?
Equation 6: I am curious to know what H+ activity is used in this equation. The authors seem to be clear that the H+ activity in aerosol water is different from that in cloud droplet. Which one is used here? or both are used. Please clarify it.

P-dissolution is calculated for aerosol water and cloud droplets separately. In the model we separate these two pathways due to the different properties of the solution. Since the aerosol solution is more condensed than the cloud droplets, we use the term ‘H+ activity’ in contrast to cloud droplets where the H+ activity can be considered equal to the concentrations. For aerosol water, the activity of H+ is calculated on-line in the model by the thermodynamic module ISORROPIA II (Fountoukis and Nenes, 2007). For cloud water, H+ concentration is calculated by the aqueous-phase chemistry module as presented in Myriokefalitakis et al. (2011; 2015). For clarity we have added the last two sentences in the second paragraph of section 2.2.

Section 3.2: Please give maps of the geographic distributions of sites measuring P-containing aerosols concentrations and dry deposition fluxes in the Supplementary Materials to show the spatial coverage of the observational data used in this study.

New figures 4a and 4b have been added to provide the location of the observations.

Line 8-13(p11): Since the model is run with meteorology for only one year (2008), I am curious to know how the authors can compare the modeled P concentrations in 2008 with that measured in other years. If the measurements were also all measured in the year 2008, it should be fine.

Please see answer to reviewer’s comment # 6.

Line 8 (p11): It is unclear if the authors evaluated the dry deposition fluxes, or they have evaluated both the dry and wet deposition. Please make it consistent.

We evaluated separately the dry and wet deposition fluxes, as explained in section 3.2. To limit the number of figures we have used one figure (Figure 4) to show both comparisons but we used different colours to distinguish the dry deposition fluxes (in red) from the wet deposition fluxes (in green). However for clarity we have removed the wet deposition fluxes comparison from this figure and we now present such comparison in the supplement: Figures S5c,d and Figures S6f, h.

Line 20-23(p11): Here, I am not convinced of the conclusion. Accordingly to our knowledge, the emission inventory is important for the modeling of P, but it is not the only factor that matters. For example, the treatment of aerosol scheme and the initial size distribution of P in the emissions can also influence the concentrations and transport of P. Unfortunately the authors do not provide necessary information on these in their methods, making it hard to judge whether the conclusion is right or not. As I know from Wang et al., they have accounted for three size bins of P-containing particles in their model, rather than the two size bins in your model. So, it is at least necessary to repeat the treatment of size distributions by Wang et al. and discuss on the impact.

We thank the reviewer for this remark. In TM4-ECPL model for all our simulations we consider two sizes of combustion aerosols, fine and coarse modes, assuming lognormal distributions for each mode (see also our reply to reviewer’s comment #8). This is described in the beginning of the second paragraph of section 2.1.2 of the discussion paper, i.e.: ‘In the model, a number mode radius of 0.04 μm and a lognormal standard deviation of 1.8 are assumed for fine P emissions, while for coarse P a number mode radius of 0.5 μm and lognormal standard deviation of 2.00 are used as proposed for combustion aerosols by Dentener et al. (2006)’. For the sensitivity study with the database of the P combustion emissions recently developed by Wang et al. (2014), we now consider 3 modes, following the method described in the supplement of that publication. This is now done and clarified at the end of section 2.1.2.

“...To further investigate uncertainties in the P combustion emissions in our model, an additional present-day simulation was performed taking into account the total (bulk) mass of anthropogenic combustion and biomass burning P emissions, as developed by Wang et al. (2014) (R. Wang, personal communication, 2016). According to that database, global anthropogenic emissions from fossil fuels, biofuels and...”
deforestation fires amount to 1.079 Tg-P yr\(^{-1}\) and natural fire emissions equal to 0.808 Tg-P yr\(^{-1}\). For this sensitivity simulation, we apply the size distribution as described in Wang et al. (2014); i.e. by dividing total emissions into 3 modes - one fine (2% of P) and two coarse modes (25% and 73% of P) - with mass mode dry diameters of 0.14 μm, 2.5 μm and 10 μm and lognormal standard deviations of 1.59 and 2.00 for fine and coarse modes, respectively.

22) Line 27 (p11): I would expect that the authors compare the modeled P concentrations with that from cruise measurements for the same days (see Figure 8 in Wang, R. et al. Atmos. Chem. Phys., 15, 6247-6270, 2015).

In the discussion paper we compared simulated and measured concentrations for the same days but not always for the same year since the simulations were done for the year 2008. We now have performed additional 11-year simulations to compare data and model for the specific year and satisfy the reviewer on this point. See also our reply to the comment #6.

23) Line 27 (p13): "SOx, NOx and NHx anthropogenic" -> " anthropogenic SOx, NOx and NHx".
Done.

24) Section 4: This part is interesting and novel. It will be better if the authors can add discussion on what can be improved to get a better understanding of this impact in the future studies or what is the most uncertainty.

We further develop this part in the revised version and in particular we have added some recommendations in section 4.2: “However, large uncertainties are associated with this innovative finding, since the estimates of the global source of PBAPs vary by more than an order of magnitude, their size distributions, their mass density are uncertain and the P-content in these aerosols is also highly variable, spanning 2 orders of magnitude (e.g. (Kanakidou et al., 2012) supplementary material and references therein). All these parameters have to be studied by targeted experiments to improve knowledge of their contribution to the atmospheric P cycle.” and further in the discussion “Our results also show that the P solubilisation from dust aerosol during atmospheric transport and mixing with acidic pollutants is important for DP deposition and deserves further kinetic studies to improve parameterisation of the solubilisation kinetics of various P containing minerals as a function of acidity and temperature.”

References


