We thank the editor and referees for the comments. Here is the point-to-point answer (in roman) to the questions (italic bold).

1) addition of St. Petersburg to the life time / emission study would strengthen the paper very much

We now include the analysis of Stockholm and St. Petersburg similarly to what was done for Helsinki. The results are summarized in Table 1 and discussed in section 3.2. The figures corresponding to Fig. 3 and Fig. 4 are included as supplementary material (Fig S1 and S2). In both cases, the uncertainties in the parameter estimates are larger than in Helsinki because of the larger sensitivity to the selection of the integration range for the linear density calculation. In particular, the effect of other point sources close to the city and the not-perfect symmetry in calm condition reduced the quality of the fit and increased the uncertainty of the results. The results of this sensitivity study are included in the uncertainty estimates for each parameter. In this sense, the Helsinki case study best fulfilled the criteria to successfully apply this methodology for emission and lifetime estimation. Despites these limitations, the lifetime values are consistent with what was found in Helsinki and the emissions are larger in Saint Petersburg than in Helsinki and Stockholm, as expected. Also, the estimated emissions agree within the uncertainties with the EMEP database information.

This text is now added in the manuscript:
L419-437 "The emission and lifetime were similarly calculated for Stockholm and St. Petersburg. The results are summarized in Table 1. The figures corresponding to Fig. 3 and Fig. 4 are shown in the supplementary materials (Fig. S1 and S2, respectively). The wind regimes were similar to Helsinki, with the wind patterns dominated by the westerlies. In both cases the uncertainties on the parameter estimates are larger than in Helsinki because of the larger sensitivity to the selection of the integration range for the linear density calculation. In particular, the effect of other point sources close to the city and the not-perfect symmetry under calm conditions affected the quality of the fit and increased the uncertainty of the results. Despites these limitations, the lifetime values are consistent with the Helsinki case and the emissions are larger in Saint Petersburg than in Helsinki and Stockholm, as expected. Also, the estimated emissions agree within the uncertainties with the EMEP database information (last column in Table 1)."

Abstract, introduction and conclusions are also updated.

2) the error discussion needs to be clarified as requested by the second reviewer and also suggested by the first reviewer in an offline comment

We added the discussion about the emission changes on the AMF in section 3.2 as reported in point 11) at the end of this document.

3) the shipping emissions need some more discussions. In particular with respect to the shipping signals in the Baltic Sea which are

a novel and important aspect for satellite data analysis there are some open points that need to be addressed.

3.1) It is worrying that the largest values are found at the large wind speeds - this is either the result of some sampling artefact or indication of a problem as the opposite would be expected.

Yes, one would expect that the signal would be smaller when strong winds are taken into account. In this case, we think that under strong wind conditions the signal is stronger mainly because of a much larger amount of data are included in the gridding and because of potential transport from land sources from Southwest. Also, we think that, because the shipping lane is directed in SW-NE direction, it can happen that the signal produced in any point along the shipping lane might be transported along the same lane (and so still contribute to the enhanced NO2 signal) by the winds from Southwest, which is one of the dominant wind directions (see Fig.2).

This aspect is now stressed more in the paper. The text in section 3.3 is now rephrased as:
L466-485 "Surprisingly, when only the westerly winds are considered, the NO2 signal is stronger than under calm wind conditions. This can be explained by the fact that a larger number of pixels satisfy the wind-driven condition, thus increasing the signal-to-noise ratio. Furthermore, because the shipping lane is directed in SW-NE direction, the emission produced from any point along the shipping lane might be transported along the same lane (and so still contribute to the enhanced NO2 signal) by the winds from Southwest, which is one of the dominant wind directions (Fig. 2 – lower panel). On the other hand, under strong wind conditions the NO2 patterns over sea can be influenced by the air masses transported from land sources. For example, the red spot in the lower left corner in Fig. 5 (right panel) is most probably caused by the outflow from southern Sweden and Denmark. In order to reduce the effect of this outflow, the time series analysis was performed under calm wind conditions and limited to the central area of the Baltic Sea (the black box in Fig. 6 centred at 57.5° N–19° E), where the minimum mixing between the emissions from land and sea sources is expected."

3.2) The shipping lane in STEAM and in OMI data do not match well in their spatial pattern in Fig. 6. I assume that this is the reason for using different boxes in STEAM and OMI data as shown in the same figure but not discussed in the text. The choice of different regions as well as the high values found close to Gotland need to be discussed in the text.

This aspect will be included in the text as:
L502-508 "The areas including the signal peaks in the central Baltic Sea (black boxes in Fig. 6) were selected from both OMI and STEAM datasets, in order to analyse the temporal evolution. Because the emissions and the OMI data do not exactly match in their spatial patterns and because of the different gridding and resolution, the selected boxes do not perfectly coincides."

For the high values south of Gotland please see the revised text in the answer to point 3.1.

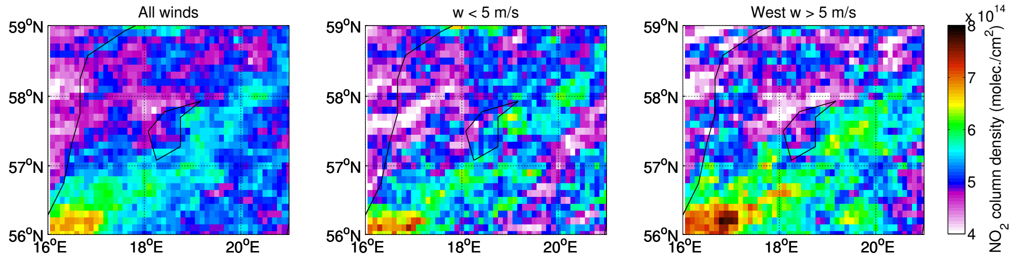
3.3) One of the arguments to support the interpretation of the signals as shipping NOx is the comparison of the temporal evolution with the STEAM inventory. However, there are several problems with that: First, the change in OMI NO2 is (in relative units) larger by a factor of 2 than the change in ship emissions. If anything, I would expect the opposite to be the case. Second, as can be seen in Fig. 1, OMI NO2 was large over the shipping area only in 2007 and 2008, which makes the link to the economic crisis less clear.

- Vinken et al. (2014) derived the sensitivity of the NO2 tropospheric column to the changes in NOx emissions over the same shipping lane in the Baltic Sea, considering the specific in-plume chemistry. They found that the ratio β between the emission change and the NO2 column change is about 0.25 for the Baltic Sea. This ratio β depends on the magnitude of emissions changes and on the local chemical regime (including in-plume chemistry). Thus, a change by 15% as we observe in the emission from 2008 to 2009 would correspond to a change by about 60% in the NO2 tropospheric column, which is even larger than the observed changes (about 30%).
- About the comparison with Fig. 1, we actually expect that the signal would increase from 2005 to 2008 and then decrease in 2009 as seen in the emission temporal evolution. The differences might come from the different time interval of sampling.

This text is now added in the manuscript:
L514-522 "Thus, the NO2 relative change is larger by a factor of 2 than the change in the ship emissions. Vinken et al. (2014) derived the ratio β between the changes in NOx emissions to the changes in NO2 tropospheric columns over several shipping lanes from model calculations. They also found that over the shipping lane the emission changes lead to substantial changes in NO2 columns (e.g.. β=0.25 in the Baltic Sea and β=0.58 in the North Sea), up to 2-to-4 times larger than the emission changes."
Also this text is now added in the caption of Fig.6: "Note that the axes have been selected so that the observed decreases are of the same size in the plot."

3.4) The figures with geometric AMF shown in the response to the reviewers are odd - the values are larger than those using the tropospheric AMF which doesn't make sense. Please check.

Yes, sorry, we found a typo in the code. The result doesn't change much though, because it was a multiplicative factor. Now it should be ok.
The quantity in that figure was calculated as follows:
vcdtrop' = (slant column density – stratospheric slant column density)/AMFgeo (instead of SCD-strat.SCD/AMFtrop)
where AMFgeo=sec(SZA)+sec(Viewing Angle).
The new version of the picture is reported here:



3.5) The absolute values of the shipping signal appear quite large to me. In Fig. 1, the values for 2007 and 2008 are as large as in the centre of Stockholm. This is surprising and I think you should follow the suggestion of the reviewer to check if the magnitude of NO2 columns observed is consistent with what would be expected from the emissions.

First, we should consider that OMI SP retrievals are generally lower in urban regions and higher in remote areas (and overall in agreement with the other measurements within 20%, Lamsal et al., 2014).
The absolute values are anyway similar with what was observed in the Baltic Sea using OMI DOMINO products (see e.g. Fig.6 in Vinken et al. (2014)), with NO2 tropospheric columns up to about $2x10^{15}$ molec/cm2.
If we look at Fig.1- left panel the signal in Stockholm is on average higher than the Baltic Sea area. The NO2 columns might be sometimes comparable (for scarsely polluted city like Stockholm) if we consider that we are looking at summer months, when there is the largest ship activity in the Baltic Sea (Jalkanen et al., 2013) because of the high season of passenger traffic and reduced pollution in the urban sites.

This text is now added to the manuscript:
L158-162 "Comprehensive validation with independent measurements
is presented in Lamsal et al. (2014). They show that OMI retrievals are lower in urban regions and higher in remote areas, but generally in agreement with other measurements within ±20%."
L314-320 "One must note that the largest ship activity in the Baltic Sea is observed in summer (Jalkanen et al., 2013), because of the high season of passenger traffic. On the other hand, the pollution in the cities is generally smaller during summer than in other seasons. Thus, the tropospheric NO2 columns over the shipping lane might be relatively large compared to the signal over the urban sites."
L530-533 "In addition, STEAM data might underestimate the ship emissions because of the effect of the small ships, which are not regularly provided with an AIS system (Jalkanen et al., 2013)."

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., and Knepp, T. N.: Evaluation of OMI operational standard NO_2 column retrievals using in situ and surface-based NO_2 observations, Atmos.

Chem. Phys. Discuss., 14, 14519-14573, doi:10.5194/acpd-14-14519-2014, 2014.

Jalkanen, J.-P., Johansson, L., and Kukkonen, J.: A comprehensive inventory of the ship traffic exhaust emissions in the Baltic Sea from 2006 to 2009, Ambio, http://dx.doi.org/10.1007/s13280-013-0389-3, doi:10.1007/s13280-013-0389-3, 2013.

Referee n.1

4) The authors missed the opportunity to extend their case studies to the Baltic sea region - according to the manuscript title. In particular I do not understand why the authors refuse to estimate emissions and lifetimes for other hot-spots like St. Petersburg and Stockholm. The selection of only one city is not significant for judging on whether the method is applicable to such weak sources. Helsinki might just be a lucky strike.

We now include the analysis of Stockholm and St. Petersburg similarly to what was done for Helsinki, including the discussion of the results. Please see the answer to point 1.

5) An appropriate discussion of errors is still missing. The authors just added some numbers, but it's not clear if they determined these numbers by themselves (if so: how?) or refer to Beirle et al. (if so: are these errors valid for Helsinki as well?)

We now add more details on the uncertainties.
- The uncertainty on the OMI NO_2 columns is provided by the algorithm developers together with the data. We add now the reference.
- The wind effect on the uncertainty is derived from Beirle (2011) and it is consistent with what we observed in Helsinki, looking at the NO_2 patterns obtained using different wind fields at different altitudes or the average over the lowest three levels.
- Already present in the paper is the discussion on the fitting uncertainty, the integration interval choice, the effect of the clear sky bias and the effect of NO_x emissions on AMF calculation (see answer to point 11).
- More considerations about the uncertainty of STEAM data are also updated (see answer to point 3.5).

This text is now added in the manuscript in the manuscript.
L408: We added the reference to the NO_2 algorithm description for the uncertainty on NO_2 tropospheric column. -> Bucsela et al. (2013)
L408-413 "The wind field patterns also affect the NO_2 spatial distribution and, thus, the parameter calculation. Beirle et al. (2011) introduced an uncertainty of 30% due to the wind fields. This is similar to the variability in the NO_2 patterns obtained in Helsinki using the wind fields at different altitudes instead of the average below 950 hPa."

L417-418 “Applying the error propagation rules, the total uncertainty on E’ and τ is larger than 40%.

6) I am still not convinced by the presented ship track. The authors state that it is a weak and noisy signal, but what then is the message? In the reply to the review, the authors presented maps of the ship track with geometric AMFs. It can clearly be seen that the columns south of Gotland are higher for w>5m/s, while I would expect the opposite (as ship emissions are independent on wind speed and higher w leeds to broadening of the plume). In order to investigate how far the NO2 pattern could be caused by ship emissions (at least the order of magnitude), the authors have to relate the STEAM emissions to an expected increase in NO2 column by assuming a NOx lifetime or by asking a CTM.

- Please find the new version of the figure with geometric AMF in the answer to the point 3.4). We found a typo in the previous version. Under strong wind conditions, we attribute the higher NO2 columns south of Gotland to the outflow effect from lower latitudes, mixed with the signal from the ship emissions. For this reason for the timeseries, we select calm conditions and a box at higher latitude in order to reduce the possible effect of outflow from land sources.
We specify now this in the text (see answer to point 3.1 – L478-487)
- As discussed in the answer to point 3.3, the analysis of the local sensitivity of NO2 tropospheric column to the changes in NOx emission was performed by Vinken et al. (2014) for the shipping lane in the Baltic Sea using CTM calculation. They found that the emission changes lead to NO2 columns changes up to 4 times larger than the relative change in the emissions. Our results go in the same direction (see answer to point 3.3).

Further comments:

7) Concerning 1b) (potential outflow from Central Europe) The authors argue that the NOx lifetime is only some hours, so NOx from Central Europe can not reach the Baltic Sea. But obviously (Fig. 5), it does! The authors should be aware that the lifetime of background NOx might be considerably longer (higher altitude, transport in reservoir species like PAN etc.).

We mean here that it is less probable that the signal is reaching the area in the box in the Baltic Sea used then for the timeseries (we use in that case only w<5m/s). It is now stressed again in the text that the plot with strong winds from East can be largely affected by the outflow from the land sources and that we use calm wind conditions and a box at higher latitudes for the temporal evolution analysis. We rephrased this part of the manuscript (please see answer to the point 3.1 and point 6).

8) OMI pixel size: The statement that the OMI pixel size is 13x24 km2 is misleading and only valid for nadir. Still, this statement is made in the revised manuscript three times.

L114 “at nadir” is added
L170-172 This sentence is added: “OMI pixel size ranges between 13x24 km^2 in the center of the swath and about 28x150 km^2 at the edges.
L620: “at nadir “ is added

9) Fig. 2: I agree that the binning of wind data in steps of 30deg is more informative. Consequently, the authors should perform the sorting of OMI observations accordingly. This might offer the possibility to have two independent fits: one for 0-30 and one for 30-60, which would be very helpful in order to judge on the representitativeness and significance of the Helsinki fit results.

This binning would reduce the sample size and increase the noise. One of the reasons why we identified only 4 wind sectors (instead of 8 as in the original method) is that we would otherwise have to few data for every sector, increasing the noisiness of the dataset. So, as the amount of data is critical for this method, we would recommend keeping the wind sector plot as it is, in order to provide more information, and to sample the NO2 data according to 4 different sectors to increase the sample size. On the other hand the results of the emission and lifetime estimates in St. Petersburg and Stockholm are consistent with the results in Helsinki. For example the lifetimes are similar to each other, as one would expect from 3 locations at similar latitudes and the emissions are comparable to the existing database (see Table 1 for details and also the answer to point 1). This, together with the uncertainty estimation, should give a good indication of the representativeness of the results.

This wind sectors choice is now motivated in the manuscript.
L190 “This allows to improve the sample size, thus increasing the signal-to-noise ratio.”

10) Fig. 3: The zoomed-in colorscale makes it hard to "read" the signal. In particular for easterly winds, the map looks very patchy, and any information about spatial patterns below 1.8e15 is lost.

We changed the scale (also in fig. S1) to optimize the information given by the data and at the same time to avoid a plot with a completely blue/purple background which would compromise the clearness of the plot.

Referee n.2

11)The meteorology in GMI may be for 2005-2007 but the emissions are for the late 1990s. Please contact Eric Bucsela for confirmation of this if you remain doubtful. Thus my original comment stands:

"GMI model used in SP retrieval: The emissions used by the GMI model are from
1997 or 1998, and these impact the profile shapes and thus air mass factors and VCDs. How have emissions in the Baltic area changed since then and

discuss how this could bias your emissions numbers."

and needs to be addressed.

According to NCEP database, the NOx emissions in Finland, for example, generally decreased from the late 1990s by about 20%. Higher emission in the a priori information leads to lower AMFs, resulting in higher tropospheric columns (see e.g., Vinken et al., 2014). Thus, the calculated emission factors might be biased high. On the other hand the clear-sky negative bias discussed in the paper could partially compensate this effect.

The following text will be added:
L380-391 "This negative bias could be partially compensated by the effect of emission changes in the AMF (Air Mass Factor) calculation. According to NCEP database, the NOx emissions in Finland, for example, decreased from the late 1990s (when the a-priori information used in OMI algorithm are derived) by about 20%. Higher emission in the a priori information leads to lower AMFs, resulting in higher tropospheric columns. Vinken et al. (2014) found that in Europe about 30% lower emissions produce on average 10% lower tropospheric columns. In our case, the calculated emission factor for Helsinki might be positively biased."