Response to reviewer #2:

General Comments:
This paper examines the impact of shipping emissions in northern Norway on air quality and aerosol radiative effects. Aircraft measurements taken during summer 2012 are combined with modeling using FLEXPART-WRF to derive ship emissions for NOx and SO2, which are then compared to an existing inventory (STEAM2). The STEAM2 inventory is then implemented in WRF-Chem to quantify the impact of these shipping emissions on pollutant concentrations and the impact on aerosol radiative effects. The paper is well written and addresses an important issue for a warming Arctic since shipping emissions are expected to increase in the coming years. There are a few explanations related to the analysis that are not fully developed as outlined in the specific comments below, such as the impact of black carbon from ships on the radiative budget. The scope of the paper is suitable for ACP and the paper should acceptable for final publication provided the following comments are satisfactorily addressed.

COMMENT: 1) Abstract, L9: Consider noting here that the ship emissions implemented in the WRF-Chem simulations are STEAM2.
RESPONSE: This has been changed to read "with and without STEAM2 ship emissions"

COMMENT: 2) Abstract, L22: Could the impact on SO2 also be quantified here?
RESPONSE: Yes, this is a good suggestion. We have calculated the impact of ships on SO_2 (+ 80% along the coast) and SO_4^{2-} (+ 20 % along the coast) and included these numbers in the abstract and the main text. In order to represent these impacts on Figure 9., we have split the Figure in two: Figure 9. (showing the impacts of shipping emissions on surface SO_2, NO_x and O_3 mixing ratios) and a new Figure. 10 (showing the impacts of shipping emissions on surface $PM_{2.5}$, BC and SO_4^{2-} concentrations). The new panels (SO_2 and SO_4^{2-}) are shown below.

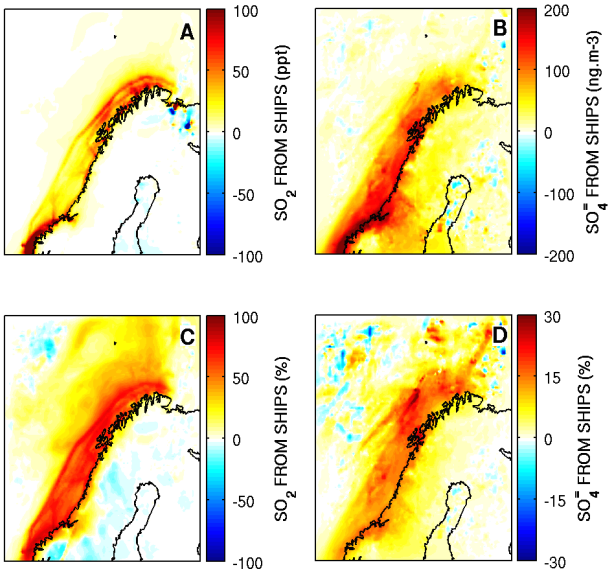

Figure 1. 15 day average (00:00 UTC 11 July 2012 to 00:00 UTC 26 July 2012) of (top) absolute and (bottom) relative surface enhancements (CTRL – NOSHIPS) in (a, c) SO_2, (b, d) SO_4^{2-} due to ship emissions from STEAM2 in northern Norway.

COMMENT: 3) There is the large negative forcing (cooling effect) due to the ship emissions, but a substantial increase in black carbon. Could there be a few more details added about why the warming effect is not that great. How are the radiative effects for

black carbon parameterized in the model? Is there any uncertainty related to these parameterizations for black carbon and how does that influence the analysis?
RESPONSE: In our simulations, the direct radiative effect of BC is calculated by computing at each model time step the aerosol optical properties through a Mie code (Barnard et al., 2010). This calculation assumes that, in each size bin, BC is internally mixed with other components of the aerosol. The average refractive index of the internally mixed aerosol is calculated as the volume average of the refractive indices of individual chemical components. These calculated optical properties are used in the radiative transfer calculations (direct effect), which means that aerosols (including BC) have an influence on the meteorological model. This influence can inhibit or enhance cloud formation (semi-direct effect). BC-containing particles can also be activated in clouds if they are sufficiently hygroscopic, and can influence cloud properties (indirect effect). Our calculations do not include the effect BC deposition on snow, because it is not included in WRF-Chem.

We did not separate the radiative effect of BC from other aerosol types, or the direct, semi-direct and indirect effects because of the large number of additional simulations and code developments needed to do so. However, previous studies investigating the radiative effect of ships globally (Eyring et al., 2010) and in the Arctic (Ødemark et al., 2012) indicate that the magnitude of the warming by shipping BC is much lower than the cooling by sulfate, especially than the cooling associated with the indirect effect. Ødemark et al. (2012) calculated the direct effect of BC from Arctic ships in July and found that it was very low, ~ +0.15 mW m^{-2}, compared to ~ –10 mW m^{-2} for the indirect effect. In our study, shipping increases BC significantly over the Norwegian and Barents seas because other sources of BC in this region are scarce, causing weak background levels. However, in agreement with Ødemark et al. (2012), we found that these enhanced BC concentrations were not sufficient to cause significant warming.

There are indeed large uncertainties in predicting the radiative impact of aerosols, including BC, on warming/cooling. The issue of uncertainty related to black carbon is a very difficult one – the uncertainty is within the detailed description of aerosols and their interaction with clouds, which are non-linear processes. A short discussion on uncertainties has been added to the text for clarity (described in more detail in the response to comment 20 by reviewer 1). However, a more detailed study of the parameterizations themselves, their uncertainties, and how this impacts black carbon is out of the scope of the present study.

COMMENT: 4) P 18411, L21: What is the time frame for the long-term cooling mentioned here due to reduced CO2 emissions
RESPONSE: The time frame of this cooling is 150 years. We also added in the text that this cooling is both due to reduced CO_2 and to reduced O_3 at lower latitudes.

COMMENT: 5) P 18414, L4: The text mentions that 4 flights are used to derive emissions, it is not clear when these 4 flights occurred since the text seems to indicate that only July 11 (2 flights) and July 12 flight were used.
RESPONSE: There were two flights exclusively dedicated to studying shipping emissions (11 July and 12 July) and two other flights that included some portion of the flight that sampled ships. The two other flights that sampled ships occurred on 19 and 25 July. The text has been updated to clarify this:
"Data collected during these flights are used to derive emissions from operating ships and to evaluate regional chemical transport simulations investigating the impacts of shipping in northern Norway."

has been changed to:
"Data collected during the 11 and 12 July 2012 flights are used to derive emissions from operating ships, and data from the 4 flights (11, 12, 19 and 25 July 2012) are used to evaluate regional chemical transport simulations investigating the impacts of shipping in northern Norway."

COMMENT: 6) P 18415, L10: The analysis requires that the environmental conditions are not strongly varying. Was this condition met? Is that why the contributions from weather effects was neglected as mentioned in P 18418, L25?
RESPONSE: Yes, during the flights the meteorological conditions remained consistent. To show this we have added a figure in an electronic supplement (Figure S1, also shown below) showing measured and modeled (MET simulation) meteorology along the flight tracks during the plume sampling flights (11 and 12 July 2012).

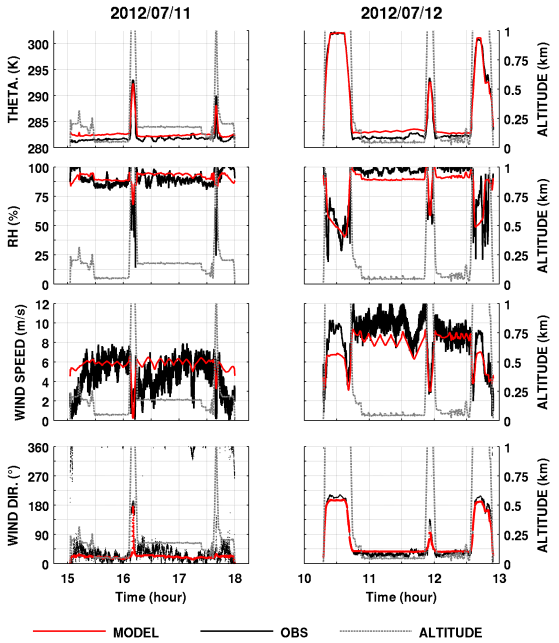

Figure 2. (Figure S1 in the electronic supplement) ACCESS measurements of potential temperature, relative humidity, wind speed and wind direction (black) for the 11 and 12 July flights, compared with WRF-Chem MET simulation results interpolated along the flight track for the same flights (red).

COMMENT: 7) P 18419, L25: Very strong SO2 emissions due to smelting are mentioned in relation to Fig 2d, but this is not evident in the related panel.
RESPONSE: The resolution of the HTAPv2 anthropogenic emissions used in this study is very fine (0.1° x 0.1°), so the localized, very high, SO_2 emission from Russian industrial sources are hard to see on the figure. To make this more clear, the Kola Peninsula is now highlighted in the box on panel 2d. We also note in the text that the highlighted region corresponds to the emission of 11,563 tons of SO_2 during the simulation, 79 % of the total HTAPv2 SO_2 emissions within the domain.

COMMENT: 8) P 18421, L14: Are you able to quantify what you mean by 'Modeled and measured plume locations agree well'? Looking at Fig. 3C and 3D, the observed plume maximum near 14E 68.8N looks to be further to the west than the simulated plume, whereas the agreement seems closer for the other panels (A and B, E and F)

RESPONSE: This displacement of the modeled plume presented in Figures 3C and 3D has been quantified. On average, the plume was displaced by 4.7 km compared to the measurements. This displacement is small considering that, at the end of this flight leg, the plume was being sampled ~80 km away from its source. This displacement is caused by biases in the MET simulation used to drive the plume dispersion model (-16° for wind direction, +14 % for wind speed). In order to clarify this in the paper, we added the following discussion of peak displacements after the presentation of Figure 4 (new text in bold):
"During the second altitude level on 11 July (Fig. 3c and d) the Wilson Leer was farther south and the Costa Deliziosa had moved further north. Therefore, the plumes are farther apart than during the first pass at 49 m. **Modeled and measured plume locations agree well for the first run (z = 49 m). For the second run (z = 165 m), the modeled plume for the Costa Deliziosa is, on average, located 4.7 km to the west of the measured plume. This displacement is small considering that, at the end of this flight leg, the plume was being sampled ~80 km away from its source. This displacement is caused by differences between the simulation (MET) used to drive the plume dispersion model and the observed meteorological conditions (-16° for wind direction, +14 % for wind speed).**"

COMMENT: 9) Why does Fig. 4 not include the July 11, Z= 165m related to Fig. 3 panels C and D? Does this time series differ?
RESPONSE: We have added the whole time series for July 11 including the leg at Z = 165m (Panels A and B) below and to the online electronic supplement (Figure S2) for clarity.

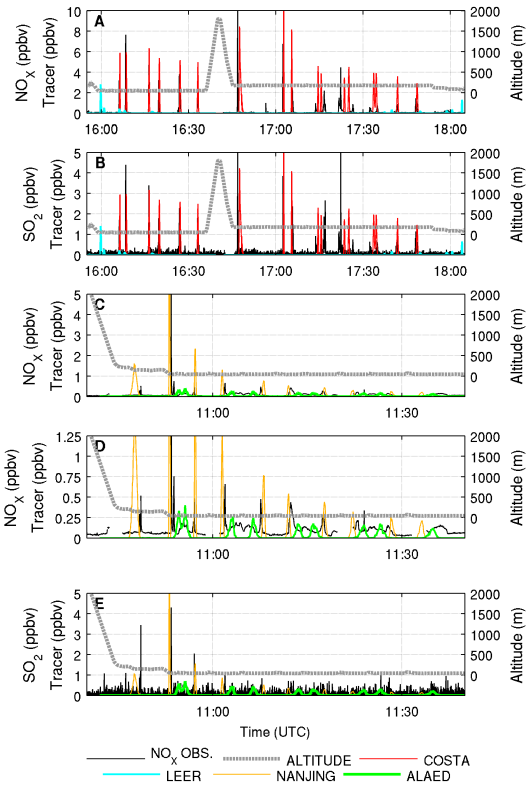

Figure 3. (Figure S2 in the electronic supplement) (a, c, d) NO_x and (b, e) SO_2 aircraft measurements (black) compared to FLEXPART-WRF air tracer mixing ratios interpolated along flight tracks, for the plumes of the (a, b) Costa Deliziosa and Wilson Leer on 11 July 2012 (showing the 2 constant altitude levels at Z ~ 49 m and Z ~165 m) (c, d, e) Wilson Nanjing and Alaed on 12 July 2012. Panel (d) shows the same results as Panel (c), zoomed in. Since model results depend linearly on the emission flux chosen a priori for each ship, model results have been scaled so that peak heights are comparable to the measurements.

This time series does not differ significantly from the other presented time series. There are two points of interest in this second constant altitude level. First, a very intense SO_2 peak was measured before 17:30 (panel B). We already noted p 18423, lines 12-15, that "this large increase in SO_2 in an older, diluted part of the ship plume suggests contamination from another source [than the Costa Deliziosa]", which is why this peak is not reproduced by the plume dispersion model and why it was later excluded from the analysis. Second, this time series illustrates the displacement of the Costa Deliziosa plume during the later part of this second leg, which was pointed out in comment 8 and by reviewer #1 and is now noted in the text.

It is also clear from this updated figure that showing the whole time series makes it difficult to distinguish the modeled and measured plumes in panels A and B, especially during the first flight leg and the beginning and end of the second leg. This is due to the close agreement between modeled and plume locations at these times. However, we also believe that showing the whole time series makes the discussion on plume selection for the analysis in Sect. 4.2 clearer. Because of this, we included this figure as Figure S2 in an electronic supplement.

COMMENT: 10) P 18422, L13: The methodology requires a linear relationship between emission flux and tracer concentration. Can you briefly note why this condition is satisfied?
RESPONSE: The FLEXPART-WRF model results depend linearly on the emission strength used as input. The only source of non-linearity that cannot be taken into account is changes in the emission source strength, which are assumed to be constant in time for the plumes sampled. Given that the ship and meteorological conditions were consistent during sampling, we expect that these effects would be very small.

COMMENT: 11) Eq.(1): How is the background mixing ratio determined?
RESPONSE: This was also brought up by Reviewer 1. We added this discussion in the text:
"The background mixing ratios were determined by applying a 30 second running average to the SO_2 and NO_x measurements. Background values were then determined manually from the filtered time series. For each NO_x peak, an individual background value was identified and used to determine the NO_x enhancement in each plume. For SO_2, a single background value was used for each flight leg (constant altitude)."

COMMENT: 12) There is a large discrepancy for NOx between the derived emissions and the STEAM2 emissions. Since STEAM2 emissions are used in the subsequent calculation of impacts on air quality and radiation, how does this overestimation influence the results and what is the related uncertainty in the presented contribution to the ship emissions to air quality and radiative effects?
RESPONSE: The STEAM2 emission model is based on AIS real-time positioning data, which has a much better coverage than activity datasets used to generate older shipping emission inventories (e.g. ICOADS and AMVER). In addition, these earlier datasets also have known biases for ships of specific sizes or types. In addition, components of the STEAM2 inventory, such as fuel consumption, engine loads, and emission factors have already been studied in detail in the Baltic Sea by Jalkanen et al., (2009) and Jalkanen et al. (2012). Beecken et al. (2015) recently compared STEAM2 emission factors to measurements for ~300 ships in the Baltic Sea. Their results showed that, while important biases were possible for individual ships, STEAM2 performed much better on average for a large fleet. In the

Baltic Sea, STEAM2 NO_x emission factors were found to be biased by +4 % for passenger ships, based on 29 ships, and -11 % for cargo ships, based on 118 ships. For SO_x, the biases were respectively +1 % and +14 % for the same ships. Therefore, we expect that the large discrepancy in NO_x for this individual ship has only a small impact on the quality of the results. We have added a short discussion to the paper.

The mass flow of NO_x in STEAM2 is calculated as a product of instantaneous power (kW) and NO_x emission factor (g/kWh) as a function of time. There are several factors which have an impact on both of these quantities. In STEAM2, engine power prediction is influenced by vessel speed in relation to water, but the value reported by AIS quantifies the speed over ground. If the water itself does not move (no in/outflow from rivers or surface currents) these two are identical, but in reality they can deviate as much as 2-3 knots (3.5-4.6 km/h) if strong currents exist in the study area. There can be also contributions from bad weather (waves, wind), hull fouling (biological growth in ship hulls which increases resistance) and sea ice cover. In this work none of these effects were included and the vessel was treated according to theoretical resistance case during ideal conditions. All aforementioned external effects, except favorable sea currents, tend to increase vessel resistance and power demand. However, if engine power prediction would be the only reason for large discrepancy between modeling results and experimental measurements, both SO_x and NO_x should indicate errors of similar magnitude. In case of Costa Deliziosa this clearly is not the case and engine power prediction alone cannot be blamed for large errors in NO_x emission predictions.

STEAM2 assigns NO_x emission factors for diesel engines based on IMO Tiers and engine rpm. The NO_x emission factor for Costa Deliziosa was assigned with the assumption that the vessel complies with Tier II regulation of the IMO Marpol Annex VI. According to this requirement, NO_x emission factor for Costa Deliziosa must be equal or less to 10.5 g/kWh, based main engine 500 rpm (crankshaft revolutions per minute). Costa Deliziosa was handled as a standard Tier II compliant vessel. Later it was found out that the owner, Costa Crociere, had installed Water in Fuel (WiFE) system on the vessel, which can reduce NO_x up to 50% (IPCO Power, 2015; Woodyard, 2004). However, we could not confirm from the owner whether the system was actually operating during the time of the measurements or not. It is imperative that all necessary details required for accurate ship emission modeling are available, which is clearly not the case for this ship. In order to evaluate the performance of STEAM2, extensive exhaust measurements onboard vessels would be required. This was not possible within this study, but has been done previously and will be published separately (work in progress). We provide three figures for two different vessels as an example of these studies.

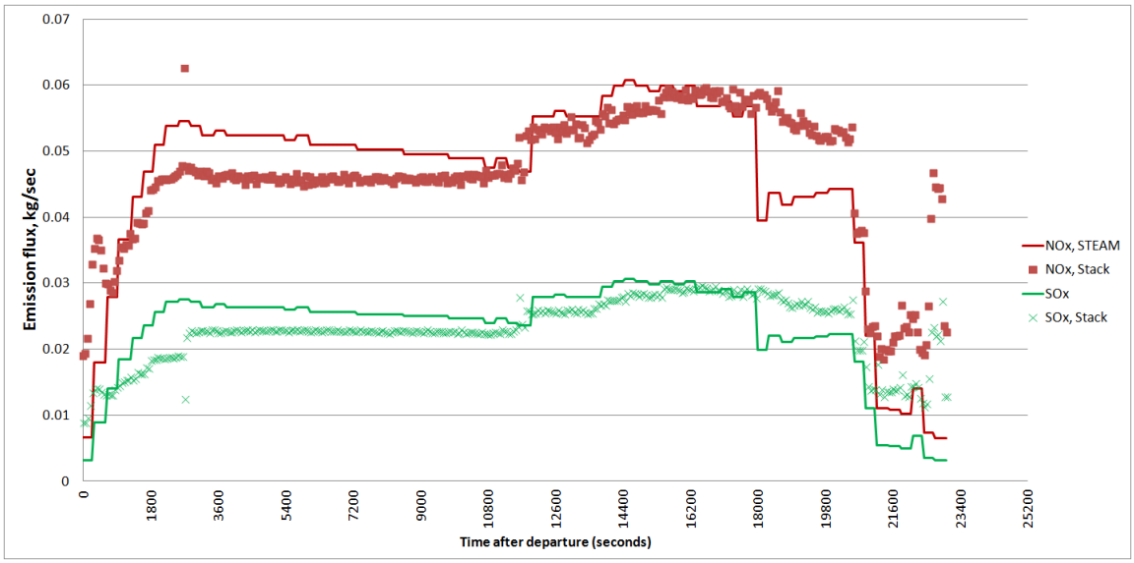

Figure 4. In-stack measurements and STEAM predictions for NO_x and SO_x. Vessel is a 34 000 GT RoPax ship. Measurement data from NUA Umweltanalytik Gmbh.

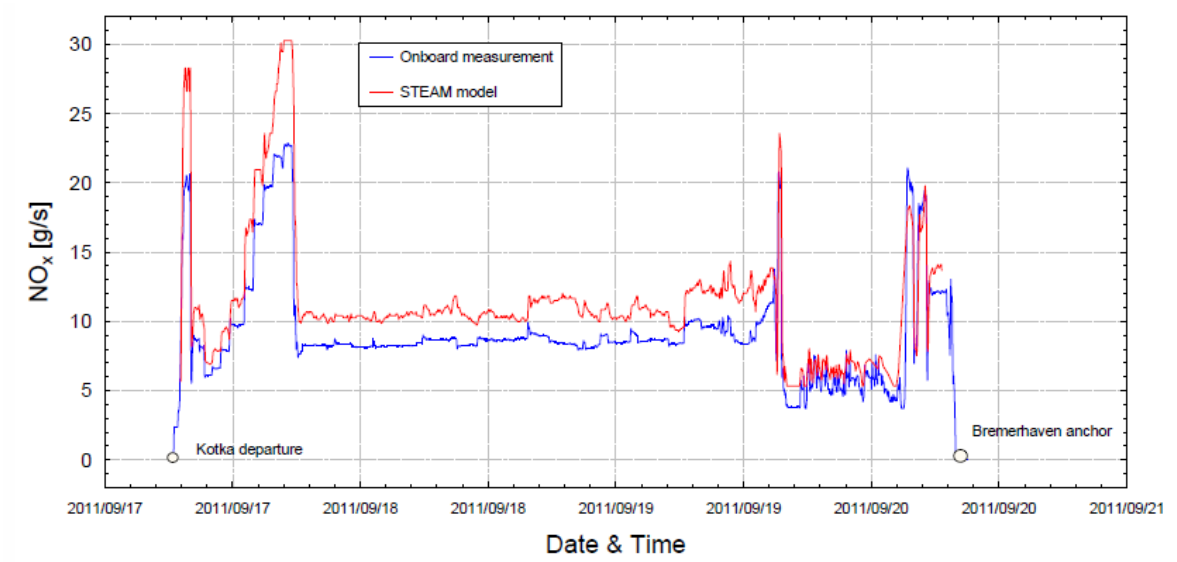

Figure 5. Comparison of measured and STEAM predicted NO_x flux from a 864 TEU container feeder vessel. Data from Maritime Universisty of Szczecin (Borkowski et al, 2012)

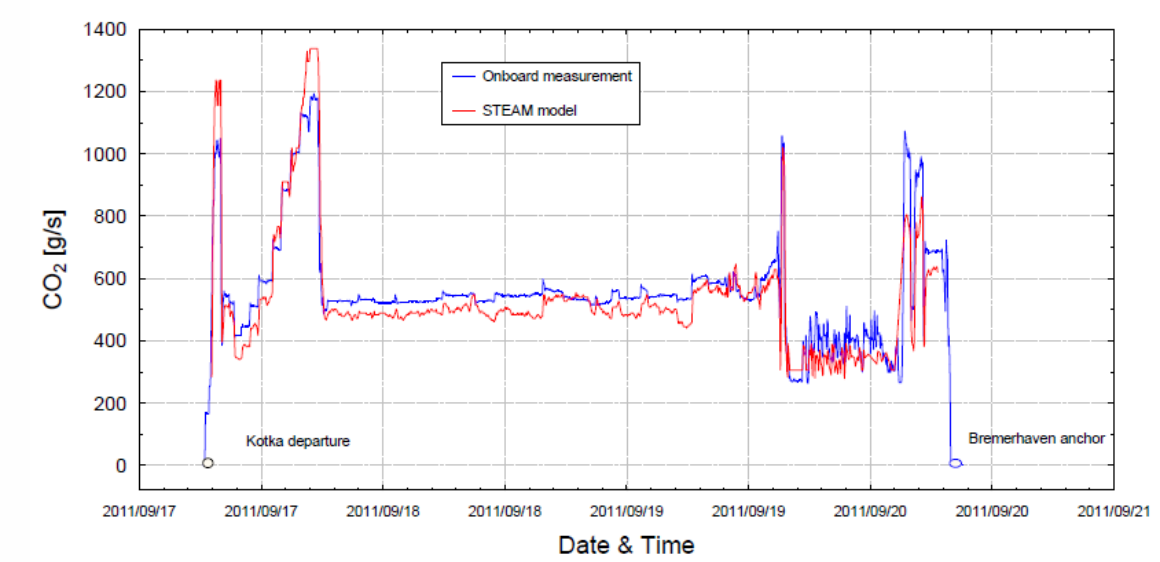

Figure 6. Comparison of measured and STEAM predicted CO_2 flux from a 864 TEU container feeder vessel. Data from Maritime Universisty of Szczecin (Borkowski et al, 2012)

COMMENT: 13) P 18428, L22 (Fig. 6): This comparison is only presented for NOx and O3 on July 11. How did the model and measured NOx and O3 compare for July 12? How does the comparison look for SO2?

RESPONSE: We have added the corresponding results for July 12 to an electronic supplement (Figure S3, also shown below) as well as a short discussion of these results to the paper. SO_2 results for July 11 are shown in the response to comment 16 by reviewer 1.

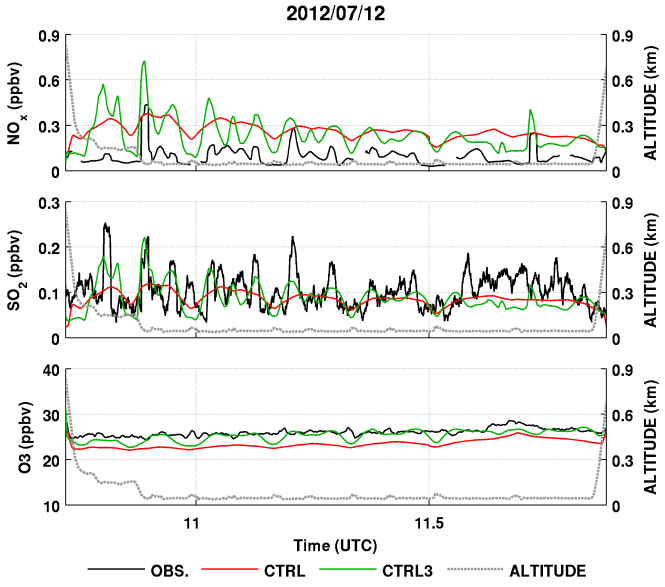

Figure 7. (Figure S3 in the electronic supplement) Time series of measured NO_x, SO_2 and O_3 on 12 July 2012 compared to model results extracted along the flight track for the CTRL and CTRL3 runs. Observations are in black, the CTRL run is in red, and the CTRL3 run is in green. A 56 s averaging window is applied to the measured data for model comparison (approximately the time for the aircraft to travel 2 × 3 km). Flight altitude is given as dashed gray line.

COMMENT: 14) P 18429, L8: Simulation CTL looks to agree more closely with the measurements than CTL3, although both have a negative bias. Is the NOx overestimation the source of this negative bias?
RESPONSE: This is a good point, but the overestimation in NO_x is likely not the only cause of this negative bias for ozone. The overestimation of NO_x discussed in the paper is only for the plume of the Costa Deliziosa, but ozone is also underestimated out of the plume, suggesting that this issue is related to the background O_3 values. This bias for background ozone can be caused by a number of reasons, including biases in the boundary chemical conditions from the MOZART4 model, and biases in photolysis rates, cloud properties and locations, ozone deposition, and/or emissions. A short discussion has been added to the paper to clarify this issue.

COMMENT: 15) P 18430, L5: The text notes that PM1 and SO2 are overestimated by about 25% and 35%, respectively. How does this overestimation influence your subsequent analysis of the aerosol radiative effects.
RESPONSE: These values are only for the Costa Deliziosa ship. We noted above in comment 12 that while STEAM2 can have a large bias for an individual ship, the average values are still very accurate (Beecken et al., 2015). Therefore, we don't anticipate there is a large impact on aerosols or radiative effects from the discrepancy we found for the Costa Deliziosa. A general discussion of uncertainties in calculating the radiative effects of shipping was included in the response to comment 20 by reviewer 1 and in the text.

COMMENT: 16) Fig. 8 shows that the ships have little effect on the vertical profile of PM2.5 – how do we reconcile this with the magnitude of the aerosol radiative effects that are presented?
RESPONSE: The radiative effect of ship emissions is dominated by the indirect effect, which is more sensitive to changes in number concentrations at certain size ranges than to changes in the total $PM_{2.5}$ mass. We added in the supplement a figure (Figure S4, also below) showing the influence of ships along on number concentrations in the second MOSAIC bin (diameters 78 nm to 156 nm), which corresponds to the typical size of aerosols activated as CCN (100 nm; Henning et al., 2002). For the modeled ACCESS vertical profiles, ship emissions increase the number concentrations in the 78 nm to 156 nm size range by + 30 %. We also note that shipping emissions have a large impact on sulfate aerosols (also shown in Figure S4 below), which are efficient cloud condensation nuclei.

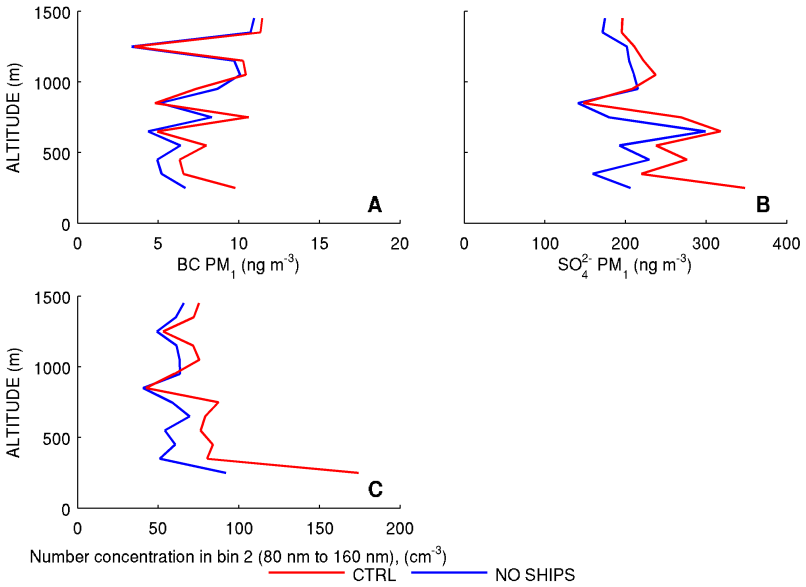

Figure 7. (Figure S4 in the electronic supplement) Average vertical profiles (200 m to 1500 m) of (a) BC PM_1, (b) SO_4^{2-} PM_1, (c) number concentrations in the second bin of MOSAIC (diameters 78 nm to 156 nm), interpolated along the ACCESS flight tracks for the 4 ship flights in the CTRL simulation (red line) and in the NOSHIPS simulation (blue line).

COMMENT: 17) P 18431, L18: The analysis of the regional impacts is based a 15-day period. How does the period chosen influence the results? Would the results for the radiative impacts and concentration enhancements be any different over a different time period?

RESPONSE: This period was chosen due to the availability of the ACCESS aircraft measurements for model evaluation (11 to 25 July 2015), which makes this study unique. Longer simulations were not possible because of the computational resources required to perform these simulations at high resolution. During the ACCESS campaign, sea level pressure anomalies were negative over Northern Norway (Roiger et al., 2015), indicative of more rain and clouds than normal during summer. In our study, we have compared our results with other studies over longer periods. However, it is unclear how the meteorological situation might influence this comparison, as these studies did not include a discussion about weather conditions.

COMMENT: 18) Why does the PM2.5 appear to have little change near the surface in Fig. 8 but greater change for the 15 day average at the surface for Fig. 9?

RESPONSE: The ACCESS measurements included in the comparison shown in Fig. 8 are for specific dates and locations, which are not perfectly representative of the whole domain. In addition, the bottom 200 m of the PBL are also not included in the vertical profiles, since measurements at these heights during ACCESS are scarce, and are only present during the plume sampling portions of the 11 and 12 July flights (already shown in Fig.6).

Compared to the discussion paper, we removed the lowest point of the profiles (100 m – 200 m) in Figure 8, since this range included part of the plume sampling from the Costa Deliziosa (z=165 m). As discussed in comments 8 (technical corrections) and in the paper, these profiles are meant to represent aged ship emissions and should not include fresh plume sampling. We also corrected the bias values, as the numbers given in the text and on the figures were Mean Normalized Bias (MNB) values instead of Normalized Mean Bias (NMB) values as indicated. Normalized mean bias values for the updated 200 m – 1500 m profiles are, for the CTRL simulation, + 14.2 % for NO_x, -6.8 % for SO_2 and -7.0 % for O_3 (+ 6.9 %, -

10.7 % and -7.6 % in the discussion paper). These two updates to Figure 8 do not change the discussion of Figure 8 or the conclusions of the paper.

COMMENT: 19) P 18433, L 5: Can you explain why BC is not efficiently transported away from the source region? Is the lifetime really that short and why?
RESPONSE: In WRF-Chem, the BC lifetime is not prescribed and is calculated within the MOSAIC aerosol scheme based on the included processes. As we mentioned in the answer to reviewer 1 (comment 19), it is not currently possible to obtain an estimate of the average lifetime of BC in WRF-Chem. However, we were able to estimate the average residence time of BC emitted from ships, using the definition of Samset et al. (2014). The BC residence time was calculated as the ratio of the 15-day averaged BC burden (kg m^{-2}) from ships (CTRL – NOSHIPS BC burden) to the average BC emissions from STEAM2 during the simulation (kg m^{-2} day^{-1}). Using this method, we estimate the lifetime of BC from ships during ACCESS is 1.4 days. During the ACCESS campaign, sea level pressure anomalies were negative over Northern Norway (Roiger et al., 2015), indicating poor weather during the campaign. This means that shipping BC was emitted into regions with clouds and rain, increasing the likelihood of removal before BC could be transported long distances. We also note that the behavior seen in Figure 9C is not just due to BC lifetime, but is also caused by dilution of shipping emissions away from the narrow shipping lanes.

COMMENT: 20) Section 5.2.2: Could the discussion and analysis be extended to include a comment of the role of black carbon. The paper attributes a 40% enhancement in black carbon to shipping – how does this influence the radiative effects presented?
RESPONSE: We answered to this comment in the response to comment 3.

COMMENT: 21) P 18435, L20: The PM10 enhancements are 15% higher in the 3km x 3km simulations. How does this influence the results for the effects on the radiative budget calculated at 15 km x 15 km?
RESPONSE: A similar question was asked by reviewer 1 (comment 20). We included in the text a discussion of uncertainties in calculating the radiative effects of shipping. In addition, we would need more aerosol measurements in shipping lanes (including number concentrations and aerosol compositions in ship plumes) in order to better understand and constrain this effect.

Technical Corrections:
COMMENT: 1) P18415, L1: Was the acronym FNL defined?
RESPONSE: FNL is not an acronym but is an abbreviation for "final" (analysis). This is now explained in the paper, when the abbreviation is first encountered.

COMMENT: 2) P 18423, L1: Are the units correct in the text? Please check.
RESPONSE: Yes, the units were wrong, we replaced "kg m^{-3}" by "g mol^{-1}" in the text.

COMMENT: 3) P 18423, L18: I have difficulty to see the 4 SO2 peaks above 1ppb for the Nanjing ship in Fig. 4 please check if this is shown and perhaps consider changing the colors since the pink and red lines are difficult to differentiate. Are not all of the peaks for the various ships mentioned in the text shown in the related figure?
RESPONSE: The colors for this figure have been updated. As we mentioned in the response to comment 9, some of the peaks mentioned in the text were not shown in the Figure, as showing the whole flight made the model and measured peak difficult to distinguish. The rest of the peaks are now shown in the supplement (Figure S2, Figure 3 in this document, also

showing the new colors).

COMMENT: 4) P18427,L 24,25 and elsewhere: Consider removing the word 'see' as a directive to the reader.
RESPONSE: This has been removed.

COMMENT: 5) P 18429, L3: The text comments on an overestimation of NOx in the plumes for simulation CTL. This is not clear in Fig. 6.
RESPONSE: In the CTRL run peaks are not overestimated, because of the dilution in the 15 km grid. This was a mistake and has been removed.

COMMENT: 6) Fig 8: Consider presenting the measurements in black as opposed to color in keeping with standard convention.
RESPONSE: This has been updated such that the measurements are in black, and the text has been updated accordingly.

COMMENT: 7) P 18433, L23: I found the terminology 'averaged over the surface of the Earth' to be somewhat confusing. Consider giving an equation to explain this calculation.
RESPONSE: Shipping studies using global models, such as the Arctic study of Ødemark et al. (2012), quantify the global radiative effect of shipping in W m^{-2}. Since we use a regional model on a limited domain, the average effect in our domain in W m^{-2} is not comparable to this global average, which includes the regions further away from the Arctic where the local impact of Arctic shipping on radiation is very low. In order to compare our own results to those of Ødemark et al. (2012) and other shipping studies, we calculated the total radiative effect in our domain in W, and normalized it with the area of the Earth in order to obtain a comparable global average in W m-2. This is now explained in the text.

COMMENT: 8) P 18435, paragraph 1: Consider quantifying the overestimation of NOx here and consider quantifying that is meant by STEAM2 emissions 'agreeing well with airborne measurements' and reasonably well representing average emissions'.
RESPONSE: We removed "agreeing well" and "reasonably well" from the text since these expressions are too vague, and we replaced them by a quantified discussion of model and emission biases. We mention that the derived emissions of NO_x and SO_2 for individual ships can be very different from STEAM2, especially when the ship specifications are not accurate. We also remind here, that, as we discussed in comment 12, the results of Beecken et al., (2015) indicate that STEAM2 performed much better on average for a large fleet than for individual ships. In the present study, we also showed on Figure 8 that combining WRF-Chem and STEAM2 produced reasonable biases for modeled NO_x and O_3 compared to ACCESS measurements in the lower troposphere (+ 14.2 % and - 7.0 %, using the updated values given in comment 18). The bias on average SO_2 profiles (- 6.8 %) in our study is also significantly better than in a previous model intercomparison by Eyring et al. (2007) based on airborne measurements in July 2004. The average ACCESS profiles presented in Figure 8 represent emissions that had time to mix vertically in the marine boundary layer. Because of this, these plumes are expected to be relatively aged, and are more representative of the regional pollution from shipping in Northern Norway than the 4 specific ships for which we derived emissions. In Figure 8, results from the NOSHIP run also indicate that shipping emissions are significant sources of NO_X and SO_2 at low altitudes along the ACCESS flights. We can thus expect that, if STEAM2 emissions were strongly biased for the total fleet, this would also lead to large biases on the average profiles, except if these biases were

compensated by another deficiency of the model. We included this discussion in the text, and a condensed summary in the conclusion when discussing STEAM2 emissions.

References:

Barnard, J. C., Fast, J. D., Paredes-Miranda, G., Arnott, W. P., and Laskin, A.: Technical Note: Evaluation of the WRF-Chem "Aerosol Chemical to Aerosol Optical Properties" Module using data from the MILAGRO campaign, Atmos. Chem. Phys., 10, 7325-7340, doi:10.5194/acp-10-7325-2010, 2010.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., Jalkanen, J.-P., Johansson, L., Litvinenko, V., Volodin, K., and Frank-Kamenetsky, D. A.: Emission factors of SO2, NOx and particles from ships in Neva Bay from ground-based and helicopter-borne measurements and AIS-based modeling, Atmos. Chem. Phys., 15, 5229-5241, doi:10.5194/acp-15-5229-2015, 2015.

Borkowksi, T., Kowalak, P., Myskow, J., Journal of KONES Powertrain and Transport, Vol 19, No 2, 2012, obtainable http://www.kones.eu/ep/2012/vol19/no2/6.pdf

Eyring, V., Stevenson, D. S., Lauer, A., Dentener, F. J., Butler, T., Collins, W. J., Ellingsen, K., Gauss, M., Hauglustaine, D. A., Isaksen, I. S. A., Lawrence, M. G., Richter, A., Rodriguez, J. M., Sanderson, M., Strahan, S. E., Sudo, K., Szopa, S., van Noije, T. P. C., and Wild, O.: Multi-model simulations of the impact of international shipping on Atmospheric Chemistry and Climate in 2000 and 2030, Atmos. Chem. Phys., 7, 757–780, doi:10.5194/acp-7-757-2007, 2007.

Eyring, V., Isaksen, I. S. A., Berntsen, T., Collins, W. J., Corbett, J. J., Endresen, O., Grainger, R. G., Moldanova, J., Schlager, H., and Stevenson, D. S.: Transport impacts on atmosphere and climate: shipping, Atmos. Environ., 44, 4735–4771, doi:10.1016/j.atmosenv.2009.04.059, 2010.

Henning, S., Weingartner, E., Schmidt, S., Wendisch, M., Gäggeler, H. W., and Baltensperger, U.: Size-dependent aerosol activation at the high-alpine site Jungfraujoch (3580 m a.s.l.), Tellus B, 54, 82–95, 2002.

IPCO, 2015: http://www.ipcopower.com/blog/costa-deliziosa-and-favolosa-equipped-with-wfe-systems-from-ipco-power

Roiger, A., Thomas, J.-L., Schlager, H., Law, K. S., Kim, J., Schäfler, A., Weinzierl, B., Dahlkötter, F., Krisch, I., Marelle, L., Minikin, A., Raut, J.-C., Reiter, A., Rose, M., Scheibe, M., Stock, P., Baumann, R., Bouarar, I., Clerbaux, C., George, M., Onishi, T., and Flemming, J.: Quantifying emerging local anthropogenic emissions in the Arctic region: the ACCESS air- craft campaign experiment, B. Am. Meteorol. Soc., 96, 441–460, doi:10.1175/BAMS-D-13- 00169.1, 2015.

Ødemark, K., Dalsøren, S. B., Samset, B. H., Berntsen, T. K., Fuglestvedt, J. S., and Myhre, G.: Short-lived climate forcers from current shipping and petroleum activities in the Arctic, Atmos. Chem. Phys., 12, 1979–1993, doi:10.5194/acp-12-1979-2012, 2012.

Woodyard, Douglas, 2004, Pounder's Marine Diesel Engines and Gas Turbines, 8th Ed., Elsevier, USA, ISBN 0750658460