This is a combined review of two manuscripts,

1. Aulinger et al " The impact of shipping emissions on air pollution in the Greater North Sea region – Part 1: Current emissions and concentrations"

2. V. Matthias et al, "The impact of shipping emissions on air pollution in the Greater North Sea region – Part 2: Scenarios for 2030",

which together describe the current situation and possible future developments of emissions from ships in the North Sea area. My comments concern both of these manuscripts at the same time, but I will list specific items which require additional attention from the authors. First of the manuscripts describes the methodology of generating ship emissions and reports chemical transport modeling based on the newly developed emission inventory. The second paper concentrates on future scenarios, extending to year 2030.

I cannot avoid having the feeling that a significant piece of work was completed and reported here, but the initial decision was to publish a single paper but later it was found that the story becomes too long. I agree that the decision to split the report in two pieces is a correct one, but the first paper is weak and it should include all relevant information to repeat the work described. Currently it does not and I would suggest the authors do a major revision to answer the points raised in this review.

Manuscript 1, Aulinger et al

General comments:

First, this manuscript describes a newly developed ship emission model, but falls short on details. Emission factors are reported only for BC, but for all other species the reader is referred to MSc thesis of Zeretzke (2013). I could not track down this reference and I would need to see how the authors handle the load dependency of the emission factors in case of other pollutants, since authors use this as justification to current work. I would suggest an addition of a new section, which describes which emission factors have been used and how the engine load and fuel type dependency is applied.

Second, the authors state that emissions from ships in port areas are not included in the current manuscript. I find this a serious shortcoming because it makes the comparisons of consecutive modeling work to air quality measurements difficult. Since the contribution from port areas is included in the measurements, it should also be included on the modeling side.

Third, a new model is introduced but its performance is not evaluated against any kind of experimental data. Or is it done in Zeretzke thesis? This reference seems essential to this manuscript, but I cannot track it down. I found a reference to Vinken et al (2014), which links this work to satellite observations of NOx, but I would have liked to see fuel consumption or stack measurement comparison to judge whether measured emission are correctly reproduced.

Specific comments:

a) Page 11279, L12-15. I would say that EU sulphur directive is also relevant here, because it goes beyond the IMO Marpol regulations.

b) Page 11279, L25: I would be careful with statements like “state-of-the-art” which sound like marketing speech.

c) Page 11281, L1-24: You are missing the third piece of the puzzle. The first is AIS data, the second is the emission model and the third is the technical data of ships. The authors are using vessel specific activity data, but do not mention the tech data here. This should be included in this description.

d) Page 11283, L12-13: “…it was made sure that the interpolated route did not cross solid ground”. I would like to read some more of this part. Linear interpolation between points inevitably leads to vessel trajectories which go over land. What methodology was used to allocate the vessel traffic to water areas?

e) Page 11283, Lines 5-8: Are you saying that you remove data points, which can be reached by the ship (speed required to cover the distance are below the design speed), but do not “fit the speed pattern”? What time interval is used for this? One hour? Changing speed can occur frequently in geographically complicated water areas like the Elbe river and Hamburg harbor entrance.

f) Page 11284, L5-6: The discussion regarding the linking IMO and MMSI together. How do you treat flag changes in this regard? Ships are sold and their flag registry changed frequently. In that case, the same vessel can have at least two MMSI numbers during a year.

g) Page 11284, L7-12: The coverage of auxiliary engine data is quite low in IHS data. GL data is better in this regard, but it does not cover the global fleet in its entirety. What data is used to determine the installed aux engine power? Have you tried other sources than IHS, GL?

h) Page 11284, L24-25: IMO number can be valid, but the commercial IHS databases may have some lag when introducing newly built vessels. Have you tried other sources (yard hull lists, owner data etc)

i) Page 11285, L2-4: The discussion regarding Black Carbon. I noticed that you have listed both BC and POA as emission species. How can you be sure that double counting does not take place? It is usual that organic part of carbonaceous aerosols contain species which are also present in BC absorbance measurements? See for example Andreae and Gelencer, 2006, 6, 3131-3148 or the work of Dan Lack and Andreas Petzold.

j) Page 11285, L8-9: You should not neglect the emissions from ships in port areas. This will inevitably distort the results, especially when you are using the emissions as input to CTM work and then assessing the quality of modeling with AQ measurements. You will never obtain a match with the measurements if you neglect part of the emissions.

k) Page 11285, L16: “For auxiliary engines the load for moving ships was kept constant at 0.3”. How do you come up with this engine load level? It is usual that ships have more than one auxiliary engine and only some of them are used. In these cases individual engines are run with higher loads than 0.3. Also, the need for auxiliary engine power varies a lot according to ship type and power transmission.

l) Page 11285-11286: The discussion of engine load evaluation and the assumptions used. This will lead

to problems, for sure. First, the authors state that engine load is valid only 0.25-1.0 range. To my mind, the cubic relationship with speed will result to 0.25 engine load with relatively high speeds. Let us assume a vessel has design speed of 21 knots. The 0.25 load level will occur already at speeds lower than 13 knots, if you use the cubic dependency. I would say that this methodology only works with ships sailing at or near cruise speeds in high seas. It may not be applicable to the North Sea fleet, because there is a sizeable fleet with four stroke engines which do not follow this dependency. It is more likely that unnecessary engines are switched off to keep the load levels reasonable even at lower speeds.

m) Page 11286, L1-3: The discussion of cases where speeds observed are higher than speeds reported in IHS data. It seems that the authors have not considered the contribution from external effects, like movement of water because of river outflow, sea currents, tide etc. Instead, they modify the speed entry of their ship tech data to match the observed maximum. This is not the way, because you are actually altering the hydrodynamic performance of the vessels, which is not correct. You will change the entire speed/powering curve with this trick and the change becomes severe if vessel travels at high speed because of the cubic speed dependency. The power curve increases sharply near the design speed and large errors in predicted power may occur.

n) Page 11286, L5-11: The part describing the BC emission factor is a refreshing analysis of the work done by Lack & Corbett. However, the rest of the two papers do not discuss BC at all. I would suggest dedicating a separate section for emission factors, because the only reference to emission factors is Zeretzke (2013) which is not available. Also, the load dependency of BC is one thing, but in their paper Lack & Corbett also propose a connection between fuel sulphur content and BC emission factors. I do not see that part in BC emission factor assignment at all. How is this feature taken into account?

o) Page 11287, L20-22: I do not understand the sentence starting “It quantifies also the differences…”. Please clarify.

p) Page 11288, L7-12: There is no description of fuel consumption modeling, or how it is done, even if authors report fuel consumption results in Table 2. What values are used for specific consumption?

q) Page 11288, L19-26: One difference which is directly visible is the neglect of ship emissions in port areas. It is not included in EMEP inventories, either, but the authors conduct CTM runs and compare to AQ observations, which include all emission sources, and then draw conclusions on the improvement made by using the newly developed emission inventories. I disagree with the statement made in L24-26 regarding the differences because a relevant part of the work (emissions from ships in port) was not done.

r) Page 11292, L1-2: I disagree with the statement that SO4 is only from secondary aerosol formation. The work of Petzold et al indicate that 1-5% of fuel sulphur is actually converted to primary SO4. The discussion whether this contribution changes as a function of engine load is still open, though.

s) Pages 11292-11293: I agree that the significance of ship emission contribution to overall air quality is demonstrated with this discussion. However, it goes slightly off the mark with the improvement of

correlation if ship emissions are present or not. In my mind, the real question should be "Is there an improvement when compared with the current ship emissions inventories?" The authors should run a comparison using EMEP emissions and the emissions generated by their ship emission model + do the consecutive CTM part in both cases. The difference in correlation coefficients between these two cases will indicate whether the quality of the inventories has improved.

t) Page 11297, L17: I am guessing here that CCLM is the name of the ship emission model? Or is it Cosmo-CLM referred to in the consecutive paper? This abbreviation is undefined, please clarify.

u) Equations should be numbered

v) Page 11302, Table 1: No difference is made between bulk cargo and container ships? Are you saying that big bulk carriers actually have 13 MW of auxiliary power installed and 0.3 * 13 MW is used during cruising as stated in P11285? I doubt this, even if vessels themselves have deck cargo gear for loading/unloading. The label "cargo ship" is a very wide one and significant differences exist between cargo ships built for different purposes.

w) Page 11316, Figure 7. These are hourly averages, right? If so, then the figure label should say so. Please, clarify.

x) Page 11317, Figure 8: Labels on the bottom image row is incorrect. Currently there are three "winter" and only one "summer".

y) Page 11319, Figure 10. The sulphate concentration in winter case looks odd. It looks like a boundary effect, since there is a very sharp contrast near the East edge of the box. True, Poland uses a lot of coal during winter, but this looks suspicious.

Manuscript 2, Matthias et al

a) Page 11326, L23-24: I would be careful with the transport fuel efficiency statement. True, ship on average, is the most efficient mode of transport, but very large differences exist between various ship types. Large intercontinental container ships or oil tankers are drastically more efficient than RoPax ferries.

b) Page 11327, L5-6 and L8-9: Repetition of 0.1%S fuel requirement. I would also say that the EU sulphur directive is relevant in this context.

c) Page 11327, L15-24. Here the authors shed some light on the NOx emission factors used in part 1, but it is not reported in part 1 at all. Also, IMO Marpol Annex VI Reg 13 Paragraph 7 sets the NOx emission requirements for ships in ECAs and extends the Tier 1 limits to vessels built in 1990's. Was

this considered in the current work?

d) Page 11328, L22-23: Here a reference is made to part 1 of the combined manuscript, but the necessary details are not there unless the authors significantly improve the description of the emissions part of the manuscript.

e) Page 11329, L17: Fleet renewal rate. How is the 2.5% replacement of the fleet done? Do the oldest vessels go first or is this done with random sampling of ships regardless of vessel age?

f) Page 11331: Scenarios for ECA LNG16 and ECA LNG21. I do not see emission factors for engines using LNG fuel (or specific consumption) reported anywhere in part1 or part2. These should be added, as well the discussion regarding methane slip from marine engines.

g) I agree with the statement made by Referee1 that land based emissions change, too. EMEP has made projections for 2030, which could be used for this purpose. The authors are actually quite optimistic with the reduction of land based emissions if they expect no growth between present date and 2030.

h) Page 11335, L17-19: Conversion rates are not the only factors increasing the concentrations, because also ship activity is higher during summer than in winter.

i) Page 11337, L18-21: This “All ships in Tier3 by 2030” is not a very realistic scenario. I understand it sets the game with Maximum Feasible Reduction and the reductions available through drastic measures, but I do not believe for a second that this would happen in reality.

j) Page 11341, Chapter 4.2.5 “PM2.5”: The discussion of PM2.5 feels weak because it is not reported in part1 or part2 how the emission factors for PM are impacted by fuel sulphur content.

k) Page 11354 and Page 11357, Figures 6 and 9: These figures are an excellent idea how to describe the impact of scenarios.