

We thank Anonymous Referee #2 for the helpful suggestions and additional references. Our point-by-point answers to the comments are presented below. Referee comments are in bold and our replies in body text.

Specific comments:

Abstract

Not all people know what the present day situation is. Although the present day situation is described in the introduction, I would suggest to include an explanation of the present day situation of sulfur control also in the abstract, so that the reader is able to compare with the scenario (1) and (2).

We rewrote the description of present-day and geoengineering scenarios into:

“We compared the climate and health effects of a present-day shipping emission scenario (ship fuel sulfur content of 2.7 %) with (1) a simulation with strict emission controls in the coastal waters (ship fuel sulfur content of 0.1 %) and twofold the present-day fuel sulfur content (i.e. 5.4 %) elsewhere

Introduction

Is it possible to explain more clearly your aim in the last paragraph? As it is now, I feel it is a bit too complicated for me.

We think the aim is presented clearly at the beginning of this paragraph (i.e., whether one could simultaneously preserve cooling while reducing mortality) and thus are uncertain what the reviewer refers to.

Methods

“The aerosol model resolves nucleation of new particles (Kazil and Lovejoy, 2007)” Could you add to the text whether the nucleation rate in the planetary boundary layer of this model is dependent on sulfuric acid multiplied with organic carbon concentrations, or if it is only dependent on sulfuric acid? This makes a huge difference to nucleation rates over the oceans, since ships emit mostly SO₂ giving sulfuric acid, but not much organics. This means that with a nucleation rate dependent on both sulfuric acid and organics, we will get very low formation rates over the open oceans even at a heavily trafficked ship lane. On the other hand, if only sulfuric acid is included in the expression for the nucleation rate, then we will have also nucleation along ship routes over the open ocean (Merikanto et al., 2009). Would it be possible to comment on this fact in your paper – in other words whether you have a positive bias of the magnitude of nucleation due to ship traffic (with sulfuric acid only - scheme), or if you have a negative bias of the magnitude of nucleation (sulfuric acid + organics)?

The model considers only nucleation of sulfuric acid and we have now clarified that in the model description. However, we think that one cannot easily speculate whether using the Kazil and Lovejoy nucleation parameterization will lead to positive or negative bias compared to the real atmosphere. Firstly, the exact nucleation mechanisms remain unidentified in many parts of the atmosphere even after the recent CLOUD chamber experiments; for example, Almeida et al. (2013, Nature) conclude that amines can explain only part of the atmospheric nucleation. Therefore we do not know for certain what the “truth” is for the marine boundary layer. Secondly, many aspects of the Kazil and Lovejoy parameterization are based on laboratory measurements, which have been likely contaminated with trace amounts of other compounds than sulfuric acid; therefore, the effect of amines and other organics

may implicitly be included in the parameterization at least to some extent.

“The combination of the model version and the cloud activation parameterization is unpublished and may differ from the official model version to be released with respect to e.g. tuning parameters.” I am not sure what you mean here. Would it be possible to clarify?

We meant that the combination of cloud activation parameterization by Abdul-Razzak and Ghan (2000) with this particular model version was not officially released. As the combination was not officially released, there could have been some differences in the model code compared to the official version to come out. However, now the newest official release of ECHAM-HAMMOZ has also A-R&G parameterization with only little differences to the implementation in our model version. Thus, we have removed this sentence from the manuscript.

“We implemented the model modifications done by Peters et al. (2012) to set all shipping emissions consistently in the first model layer assigning primary sulfate, organic carbon and black carbon emissions from shipping to the soluble Aitken mode with geometric mean radius of 30 nm.”

Two points from my side: 1. What do you think about motivating your choice of emissions centered at 30 nm diameter with the papers by Jonsson et al. (2011) and Petzold et al. (2008)? 2. By examining the references Lieke et al. (2013), Popovicheva et al. (2012), and Xie et al. (2007) the size segregation emissions for the different group of compounds might be different than what you have suggested here. I recognize the difficulty of knowing exactly in what size ranges different chemical species groups are emitted, and that number concentration emissions are dominated at around 30 nm diameter. Hence, I accept the choice of ship emissions in your paper. Please consider though to discuss how the choice of size distribution emission can affect the results. Peters et al. (2012) have discussed some of the problems with the size distribution.

Note that our geometric mean radius, not diameter, of the emission mode was reported to be 30 nm in the original manuscript. During the review phase we found out the radius of primary sulfate particles from shipping had been actually 44 nm in our simulations (see author comment in the discussion). We corrected that number in the manuscript.

Petzold et al. (2008) report that combustion mode is centered at 50 nm diameter being unclear whether that is mode or geometric mean diameter. Jonsson et al. (2011) reported slightly smaller values of up to 40 nm in geometric mean diameter. We have added the following paragraph to the manuscript:

“The chosen mode diameters are smaller than the default sizes in ECHAM-HAM (Stier et al., 2005; Zhang et al., 2012) reflecting recent measurements of ship emissions (e.g., Petzold et al., 2008; Jonsson et al., 2011). Our choice of diameter for carbonaceous aerosols is close to the value reported by Petzold et al. (2008), who measured the number density of the non-volatile combustion mode to be dominated by particles with radius of 40nm. On the other hand, the diameter for primary sulfate emissions is in our study somewhat larger than found for example in the Petzold et al. (2008) study. Since smaller particles mean more cloud condensation nuclei (provided that the particles are still large enough to activate as cloud droplets), the sulfate diameter used in our model version can potentially lead to an underestimation of the aerosol indirect effect. However, different measurements campaigns have yielded highly varying

results for the primary sulfate particle size (Petzold et al., 2008; Jonsson et al., 2011) reflecting the fact that it is difficult to extract the diameter of particles from shipping emissions based on measurements due to e.g. several chemical components involved, plume aging, and variability of engines.”

**“Lacking a precise formulation, we used the original black carbon emissions for all simulations.”
Which is the original study that you are referring to here?**

We refer to the ACCMIP emissions (Riahi et al, 2007,2011). We clarified the sentence to:
“Lacking a precise formulation, we used the unmodified black carbon emissions from the ACCMIP data base for all simulations”

“Due to the model version used, our analysis includes only sulfur, organic carbon, and black carbon aerosol emissions from shipping. Other main aerosol and aerosol precursor compounds in shipping emissions include nitrogen oxides and volatile organic compounds (Eyring et al., 2010), but we expect them to have only minor effects on aerosol-induced premature mortality and radiative forcing.”

I understand that it is hard to know the chemical composition of the ship emissions, since these are poorly characterized, and that the other compounds influence to a minor extent to the total mass emissions. However, wouldn't it be little safer not to make a conclusion that the chemical composition of the other minor compounds has a minor effect on health and radiative forcing? For example, Lieke et al. (2013) show that the particles emitted below 100 nm diameter at a test ship engine contain for example CaSO₄, or other crystalline salts. Even if these compounds do not comprise a significant fraction of the total mass (which is dominated by particles larger than 100 nm diameter), it might still comprise a significant mass fraction of the particles below 100 nm diameter. These particles are possibly influencing cloud activation to a large extent, and hereby the radiative forcing. Please, would you mind considering to change the conclusive remark?

This is a valid point and we have now replaced claim of “only minor effects” with:

“Not including these other compounds leads may lead to an underestimation of aerosol-related climate and health effects of shipping.”

We also added sentence “Lieke et al. (2013) measured also crystalline salts in the ship exhausts.” after the sentence describing major aerosol components from shipping emissions.

Results

Section 3.1. I'm a little surprised by the very large difference between ships-2020 and geo-wide for PM2.5 concentrations over continents, which means that the emissions far away from the coast-line (the difference between geo-wide and ship-2020) can still affect PM2.5 over continents to a significant extent. Ok, I realize that SO2 can be oxidized to sulfuric acid long time after the first ship emissions, and that's why ship emissions far away from the coast can influence continental levels. But, I didn't realize that this could give as high as 0.5 ug/m3 PM2.5 over continents. Maybe you could add one sentence stating that although coast-line emissions are not allowed in geo-wide, the emissions further away from the coast-line still are able to contribute significantly to the continental PM2.5?

We added the following sentence to the end of the paragraph:

“The difference in continental PM2.5 concentration between geo-wide and ships-2020, which have the same coastal emissions, shows that aerosol emissions from the open ocean contributed significantly to continental PM2.5 concentration in geo-wide.”

Section 3.4.1. “We estimated the resulting error by assuming that the ratio between the modeled and bias-corrected PM2.5 values follows a linear fit between modeled and measured PM2.5 concentrations (Fig. 5, red lines).”

Gets complicated. Seems difficult, but is there a way to explain it more clearly?

We rewrote the part into:

“In the first method, we assumed that the model underestimates PM2.5 concentrations in all simulations so that the ratio of the real (or corrected) and modeled PM2.5 concentrations equal the slope of the linear fit between measured and modeled PM2.5 concentrations (Fig. 5, red lines). Using this assumed dependency, we re-calculated the premature mortality due to shipping emissions with total PM2.5 concentrations multiplied with 1.61 (fit to EMEP data) or 1.18 (fit to IMPROVE data).”

We also replaced “Second” with “In the second method” at the beginning of the description of the other way to assess the uncertainty.

Discussion

Section 4.1. Maybe the words “prefer” and “pareto” could be changed to something else?

We have changed ‘others prefer less cooling’ to ‘others from less cooling’. However, we have left the word ‘pareto’ as is since its meaning is explained in the following parenthesis.

“If the sensitivity runs are excluded, the other simulations cannot be put into a preferred order without deciding some conversion method between RFP and mortality rate. For example, geo-narrow offered a stronger cooling (-0.53Wm^{-2} vs. -0.43Wm^{-2}) than geo-wide but had also a greater annual mortality rate ($34\,900\text{ yr}^{-1}$ vs. $15\,400\text{ yr}^{-1}$).” I don't fully understand. Would it be possible to clarify?

The sentence was unnecessary complicated. It has been rewritten into:

“One cannot put simulations geo-wide, geo-narrow, and ships-2020 into a preferred order without deciding some conversion method between ERF and mortality rate.”

References not in the original manuscript

Almeida, J., et al.: Molecular understanding of sulphuric acid-amine particle nucleation in the atmosphere., *Nature*, 502, 7471, 359-363, 10.1038/nature12663, 2013.

Lieke, K. I., Rosenørn, T., Pedersen, J., Larsson, D., Kling J., Fuglsang, K., Bilde, M.: Micro- and Nanostructural Characteristics of Particles Before and After an Exhaust Gas Recirculation System Scrubber, *Aerosol Sci. Tech.*, 47, 1038–1046, 2013.

Jonsson, Å. M., Westerlund, J., and Hallquist, M.: Size-resolved particle emission factors for individual ships, *Geophys. Res. Lett.*, 38, L13809, doi:<http://dx.doi.org/10.1029/2011GL047672>, 2011.

Petzold, A., Hasselbach, J., Lauer, P., Baumann, R., Franke, K., Gurk, C., Schlager, H., and Weingartner, E.: Experimental studies on particle emissions from cruising ship, their characteristic properties, transformation and atmospheric lifetime in the marine boundary layer, *Atmos. Chem. Phys.*, 8, 2387-2403, doi:<http://dx.doi.org/10.5194/acp-8-2387-2008>, 2008.