We thank Prof. Collins for his helpful comments. Our point-by-point answers to his comments and questions are presented below. Referee comments are in bold and our replies in body text.

Abstract: The IPCC AR5 now uses the term "Effective Radiative Forcing" (ERF) rather than RFP to describe fixed SST experiments. The authors may or may not wish to adopt this too.

We changed RFP to ERF in the text, noting that the term RFP has also been used. The original text refered to the definition of forcing in IPCC AR4, and to avoid confusing the reader, we changed the sentence:

"...is comparable to the radiative forcing concept used by the IPCC (Lohmann et al., 2010)." to:

"...and makes it possible to compare total aerosol forcing with forcing from well-mixed greenhouse gases (Lohmann et al., 2010)."

Section 2.1
What is the model resolution?
The model was run with resolution T63 (1.9X1.9 degrees). This information is given at the last paragraph of Section 2.2 and we do not want to repeat it here.

Page 21995, lines 10-12.
Why do the authors have to make assumptions about ACCMIP emissions? This information should be available from Lamarque et al. 2010, (which references Eyring et al. 2010). The authors should understand the origin of this dataset and describe it more fully here.

We were unable to find detailed information on what sulfur content was assumed in each grid cell in the ACCMIP emission files. Lamarque et al. (2010) describes only historical emissions up to year 2000, and does thus not cover the year 2010 emissions used in our work. Lamarque et al. (2013, GMD), which describes the ACCMIP runs, simply states that the 2000 dataset was used for harmonization with the future emissions determined for the four RCPs described in van Vuuren et al. (2011; Climatic Change); however, we were unable to find specific information on the 2010 fuel composition in the van Vuuren paper. A closer look at the ACCMIP data shows that the ship emissions in all grid cells have increased between years 2000 and 2010; this implies that lowering of sulfur content in the emission control areas has not been taken into account. Therefore, we decided to assume that the ship fuel sulfur content is 2.7% in the ACCMIP emissions for the year 2010, which is the current global mean value. We acknowledge that assuming sulfur content of 2.7% everywhere is a simplification but, given the lack of detailed information on the ACCMIP dataset, decided this was an acceptable assumption.

Page 21996, lines 23-25.
Shipping is a large NOx source. The authors should do a rough scaling to estimate the magnitude of the nitrate contribution to health and ERF in order to demonstrate that it has only "minor effects".

We accept that our previous statement about the role of NOx was too strong. We modified that sentence to acknowledge the uncertainty better:

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

Section 3.2
This is the only section where I had some concerns about the methodology. The differences between the health impacts are due to having strips 0, 1 or 2 grid boxes wide. How accurately can the model transport aerosols over 1 grid box? On this scale the amount of pollution reaching the coast might depend on how diffusive or not the advection scheme is, and even on such subtleties as whether the advection is carried out before or after the chemistry and deposition steps. It would be useful for the authors to show detailed maps (eg. of North Sea and Baltic) of the different emissions used so the reader can get a better impression of the size of these coastal strips and their blockiness.

It is true that the transport of aerosols over only one grid-box is difficult for the model to get right. In ECHAM, advection is calculated before aerosol dynamics and deposition in the model. We have now added the following sentence to Section 3.41:

"It should be noted that in our scenarios, the ship-induced PM2.5 concentrations over the continents depend largely on aerosol transport over just one or two grid-cells. This means that the simulated PM2.5 concentrations are sensitive to the accuracy of the advection scheme."

Even though a detailed map to show the blockiness would be good for some readers, we prefer to keep the numbers of figures low and the suggested figure would not help to convey our main message or results. We already have fairly detailed analysis and discussion of the possible sources of errors and uncertainties.

Another way of generating the emission fields would have been to apply the 200 nm strip to the original data (at 0.5 x 0.5 deg resolution) and then downgrade it to the model resolution.

That is true. This method would most probably yield very similar results to our method.

Page 22001, line 25 - page 22002, line 5.
I assume the coastal forcing is calculated by averaging the ERF from each grid box in the coastal zone. This needs to be explained better. I first read this as being the forcing due to coastal emissions, which can only be calculated with a separate experiment. This doesn't really show that emission reductions near the coasts have relatively little effect as you wouldn't expect the effect of coast emissions to be confined to the coastal zone, i.e. coastal emissions would also have an ERF over the land and extend further out to sea.

Emissions from the open sea also contribute to the RFP in the coastal regions, which counteracts the effect that coastal emissions have forcing also outside the coastal zones. However, we agree that the reasoning should be supported by additional experiments to be reliable, and we have removed the analysis of coastal RFP from the paper.

Section 4.1
I am slightly uncomfortable about using words like "desirable" and "optimal" in this section. As discussed in section 4.2 it is not obvious that these words can be defined in a public good sense.

As we state in the paper, it is not straightforward to say that stronger ERF would be better, but replacing those words everywhere would be in conflict with the motivation of the study. Thus, we want to keep the wording as it is and trust that discussion of the limitations is enough to convey the message that global mean ERF is not an adequate metric to express all climate consequences of shipping emissions.

Page 22006, lines 16-18.
Jones et al. found the temperature change was not homogenous, rather the cooling was concentrated in the tropics. This contrasts with the pattern from CO2 where the warming is largest in the Arctic. Therefore geoengineering can't balance the CO2 in each region simultaneously.

We used sloppy wording in the original manuscript. We did not intent to say that temperatures decrease the same amount everywhere, but that they do decrease also outside the geoengineered stratocumulus regions. We have now changed the sentence to:

"Previous studies have shown that regional forcing over the oceans creates a global cooling effect, although the regions with strong local radiative forcing cool the most (Hill and Ming, 2012; Jones et al., 2009; Rasch et al., 2009)."

Page 22007, lines 14-17.
It would be useful to explicitly mention the timescales over which the ships would be needed to be deployed continuously.
We added parenthesis stating the time-scales explicitly: "i.e. centuries or millenia". We also added a reference to Brovkin et al. (2007) who showed the long time scales required for keeping up the geoengineering system to counteract the warming from greenhouse gases.

Page 22007, lines 18-22.
Presumbably there is horizontal mixing between the open ocean and the continental shelf so the acidification would reach the coast. It would be useful to get an order of magnitude estimate of the change in pH compared with that due to doubled CO2.

Estimating the change in oceanic pH is not straightforward and is also out of the scope of this paper. However, a recent study by Hassellöv et al. (2013, GRL) concludes that modeled acidification is greatest along the heavily traffic routes and that some of the highest effects (0.0015-0.002 pH) are observed along the coasts. These effects are of the same order of magnitude as the annual surface water acidification in the open ocean due to increased CO_2 concentration. We have now replaced the paragraph with:

"The increased sulfur emissions over the open oceans in the geoengineering simulations could potentially increase ocean acidification. Hassellöv et al. (2013) concludes that ocean acidification due to SO_x and NO_x from shipping emissions could be in the same order of magnitude as the effect of increased

CO2 concentration near the major shipping routes. However, the coastal areas, which are most vulnerable to acidification (Doney et al., 2007), had either present-day or decreased sulfur emissions in our simulations, although the coastal impact of acidifying compounds transported from the open oceans cannot be totally excluded based on our simulations."

Page 22008, lines 5-8.
This sentence could be deleted. This study hasn't addressed alternative geoengineering techniques or additional greenhouse gas mitigation measures and so doesn't have the information required to recommend exploring them.

We deleted the sentence.

Page 22008, line 12.
Delete "state-of-the-art" as it doesn't add anything.
Deleted.

Page 22009, line 15.
"too precious to lose" is unecessarily emotive. The two sentences could be simply combined -"The cooling effect of aerosol emissions from shipping could be preserved by...".

Done as suggested.

Page 22009, lines 19-20.
The sentence "Therefore, it should..." can be deleted.

We want to keep the sentence to make sure that our study is not interpreted as a proposal to carry out the simulated geoengineering scenarios in the real world.

Figure 6.
x-axis title needed.
The label for the x-axis "Excess mortality (thousands per year)" was left out in the type-setting process. We will put it back again.

References not in the original manuscript

Brovkin, V., Petoukhov, V, Claussen, M., Bauer, E.,Archer, D., and Jaeger, C.: Geoengineering climate by stratospheric sulfur injections: Earth system vulnerability to technological failure, Climatic Change, 92, 243–259, doi:http://dx.doi.org/10.1007/s10584-008-9490-110.1007/s10584-008-9490-1, 2009.

Hassellöv, I.-M., Turner, D., Lauer, A., and Corbett, J. J.: Shipping contributes to ocean acidification, Geophys. Res. Lett., 40, 2731–2736, doi:http://dx.doi.org/doi:10.1002/grl.50521doi:10.1002/grl.50521, 2013.

Lamarque, J.-F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, Geosci.

Model Dev., 6, 179-206, doi:10.5194/gmd-6-179-2013, 2013.

van Vuuren DP, JA Edmonds, M Kainuma, K Riahi, AM Thomson, K Hibbard, GC Hurtt, T Kram, V Krey, J-F Lamarque, T Masui, M Meinshausen, N Nakicenovic, SJ Smith, and S Rose . 2011a. *The representative concentration pathways: an overview* Climatic Change, 109: 5-31. DOI: 10.1007/s10584-011-0148-z.