Response to comments from Reviewer 2

(Reviewer's comments are copied *in italic* here; authors' responses are **in bold**)

General Comments

1. *Although the paper is generally well written and clear, parts of the paper are quite long, and could be written more concisely. In particular, I would encourage the authors to consider whether a summary figure and more concise discussion of the model evaluation against observations could be included, with the minute detail of seasonal vs daily vs hourly comparisons and the information in Table 4 could be moved to supplementary information. Specific aspects of these evaluations of particular relevance for shipping impacts could be summarised in the main text.*

Response: This is a good suggestion. We have revised this section to be more concise. The revised Table 4 presents the hourly statistic scores only, while the extensive statistical scores (previously included in Table 4) are now presented in the supplementary material, and the discussions are more focused. The original manuscript did contain a paragraph at the end of this section summarising the main evaluation results in terms of model's ability of simulating ambient concentrations of criteria pollutants in northern Canada and the Arctic. We have strengthened the discussion in the revised version to emphasise the aspects of particular relevance to assessing shipping emission impacts.

2. *An important aspect of how shipping may impact ambient pollution concentrations in coastal regions is related to the dispersion of pollutant plumes emitted by ships. This is partly related to the vertical boundary layer structure and stability. If possible, and data is available, it would be helpful to include somewhere some assessment of the model vertical BL structure (temperature profile, BL depth), or at least add a comment based on past evaluation of the model.*

Response: Since the meteorological model (GEM, the hosting model of GEM-MACH) is the Environment and Climate Change Canada's operational numerical weather forecast model, the model evaluation in this study has been focused on atmospheric chemistry aspect. However, as the reviewer correctly pointed out, the model's ability to simulate the coastal marine boundary layer would have an important influence on assessing the shipping emission impact on ambient concentrations. Although GEM operational performance has continuously been evaluated against surface and upper air observations and compared against other NWP models of leading Operational Forecasting Centres in the world, the Arctic region alone had not been given significant attention in the past operational evaluation exercise. In order to address the reviewer's comment, we looked at the modelled vertical temperature profiles compared with upper air soundings at a number of coastal sites in the Arctic along the main shipping channels for the month of July in 2010. On average, the modelled vertical temperature profiles compare well with the observations (see Figure R2-1 below). We have also attempted to look into estimation of boundary layer heights. We compared the diagnosed boundary-layer heights from the modelled and observed profiles (determined from potential temperature profiles – the level of maximum

vertical gradient, based on Stull 1988) at the upper air sites. There is an overall negative bias in model diagnosed BL heights compared to those diagnosed from the observation. However, it should be pointed out that under stable conditions (as the case of the Arctic marine BL) there is a large ambiguity in the definition of BL height – the diagnosed BL height can vary significantly depending on the particular method (or parameterization) used (e.g., Aliabadi et al., 2016, Atmosphere-Ocean 54 (1) 2016, 60–74). In the revised manuscript, we have added a brief discussion (at the end of section 4) on the model's ability in simulating the Arctic marine BL structure based on our simple assessment of modelled vertical temperature profiles. A more detailed examination of GEM's forecast capability in the Arctic is being carried out as part of the Year of Polar Prediction (YOPP) activities at Environment and Climate Change Canada.

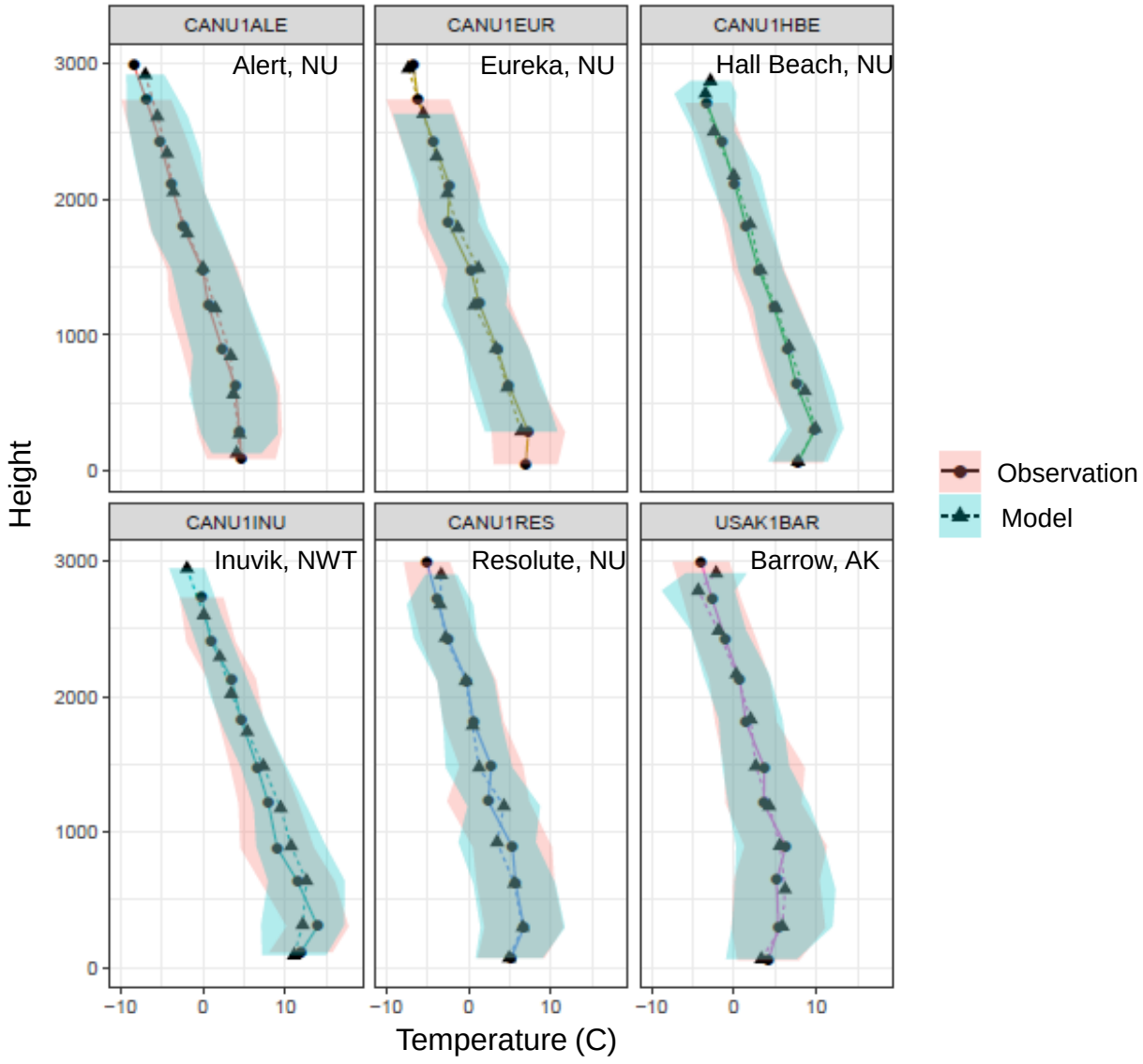

Figure R2-1. Comparison of modelled vertical temperature profiles with observations at 6 Arctic upper air sites for July 2010 (monthly means and standard deviations).

Specific Comments:

Page 3, line 7: Is all of the Arctic pristine? Depending on time of year, "background" PM concentrations may be very different in different areas? This may have implications for the impacts of shipping. How does the Canadian Arctic compare with e.g. N Siberia in terms of background (non-shipping) PM and ozone?

Response: Here we are referring to eastern Arctic (or Canadian Arctic) during summer. It is true that the Arctic may not be as pristine as one would imagine in the wintertime due to long-range transport of pollutants from southern latitudes, and that parts of the Arctic are not as clean due to oil and gas activities (e.g., in northern Siberia and Alaska). We have revised the sentence to qualify this:

"Although Arctic marine shipping currently accounts for a small percentage of global shipping emissions, it makes a proportionally bigger impact on the environment than does shipping at lower latitudes due to the generally pristine Arctic background, particularly in the Canadian Arctic Archipelago."

Page 4, lines 3-4: It would be useful to know more about what was assumed regarding the "limited number of transits of north-west passage". Which sea ice and climate projection scenario are these most consistent with? Which criteria went into this assumption?

Response: The "limited number of transits via the Northwest Passage" was based on restricting the transit to bulk carrier vessels only. This is in consideration for the unpredictable ice conditions in the Canadian Arctic archipelago, even with the projected opening of the NWP (by mid-century based on available climate projections at the time), and economic viability (factoring in extra costs for vessel strengthening, loss of cargo carrying capacity, etc.). The projection for the 2030 NWP transit is based on a gradual (linear) increase from 2020 to a 2050 high-growth (or business-as-usual) scenario assuming bulk carriers would carry the 2050 Northern Europe-Asia bulk trade through the NWP. The 2050 bulk trade between Northern Europe and Asia was projected at an annual rate of increase based on historic trade data between 1975 and 2005. We have added a little more information on the 2030 projection in the revised manuscript but refer readers to the Innovation Maritime Report for more detailed information.

Section 4.1: Discussion of model evaluation against observations. It would be helpful to compare model performance with other model studies focussed on similar regions where possible. e.g. the POLMIP models compared with ARCTAS aircraft data near surface over Canada? (Emmons et al., 2015). Other global modelling studies? For the comparison of SO2 agains observations (but also relevant for other species) -it would be helpful to know how the regional averages and poor model performance are skewed by certain sites. e.g. can the comparisons with the oil sands sites be separated to show how the model compares away from this source (and specifically near to it)?

Response: We were not aware of the POLMIP project and are very glad that the reviewer brought the publication to our attention. Although we recognise that most of the models involved in the POLMIP project are global models at much coarser resolutions and that there may be inherent differences in

comparing model with ground-level observations and aircraft measurements, we have made reference to the POLMIP model comparisons with the ARCTAS-B aircraft observations (O3, NO, NO2, and SO2) in the revised manuscript to provide a general comparison with other existing model studies in our region of interest. For the SO2 evaluation against observations, we have followed the reviewer's suggestion to examine the comparison excluding those sites strongly influenced by oil and gas production/processing activities (e.g., northern BC and Athabasca oil sands area). It clearly demonstrated that those sites were skewing the regional comparison: the large model positive biases were greatly reduced in this case. In fact the comparison shows that the model tends to under-predict SO2 (negative bias) at the more remote sites in the northern region (possibly due to the lack of representation of some local sources in the emission inventory, e.g., diesel generators, garbage burns). We have included this additional analysis in the revised manuscript.

Section 5: The impacts of shipping on ambient pollutant concentrations in future may be closely tied climate system changes - particularly changes in sea ice. It would be interesting to consider how the conditions that may make shipping more favourable (reduced sea ice) may also contribute to a change in the impact of the shipping emissions. While additional model simulations are probably outside the scope of this study, could the authors comment on how reduced sea ice might be expected to impact ozone and PM in the summertime Arctic in context of their results?

Response: Our study focused on assessing the contribution of Arctic shipping emissions to air pollution at current and projected future levels. We are isolating the changes to marine shipping emissions only and have not considered assessing the impact of shipping emissions under a changing climate. Indeed the impacts of shipping on ambient pollutant concentrations in the future will undoubtedly be linked to climate system changes. The reduced Arctic sea ice cover may have an impact on the atmospheric circulation and meteorological systems which will have an impact on the transport and transportation of atmospheric pollutants. The system is highly complex. To comment on how reduced sea ice would impact ozone and PM in future summertime Arctic in isolation without proper study would not be very meaningful in our opinion. We have added a statement in the "summary and conclusions" session to emphasise the focus (and limitation) of our assessment.

Figures 4,5,6,7: How representative are the regional average concentrations, and is there much spread? e.g. it would be helpful to know how variable the concentrations are that make up each average. Can some spread be plotted in shading? This is also relevant to discussion of model evaluation above.

Response: We have addressed the reviewer's question by adding shaded bands showing the lower to upper quartile range, highlighting a measure of spread and how well (or not) the mean represents the distrabution.

Editorial corrections:

Page 3, line 3: "theses" should be "these"

Response: Done. Thank you!