

We thank the reviewer for the constructive comments. This document summarizes our responses and documents the changes made to the manuscript.

All reviewer comments are written in black font

All responses are written in red font

All references to changes in the revised manuscript are written in **bold red font highlighted**

P

The paper's focus is on the validation of the Polair3D chemistry transport model over the geographical domain of Quebec, Canada at different spatial and temporal resolutions. The authors highlight that the model was used for the first time over the Quebec domain and for the first time with a temporal period long enough to evaluate its performance at the seasonal level. The authors present a statistical analysis of the model performance in representing a range of primary and secondary air pollutants comparing model results with ground observation sites. Finally, they conduct a sensitivity test on the primary emissions impact on final concentrations switching off the industrial emissions and evaluating the change in the levels of air pollutants. The manuscript is, on the one hand, highly focused on the evaluation of the model performance in representing a range of pollutants on a large temporal scale. This represents a challenge for any CTM because of the impact that meteorology and chemical mechanisms can have at seasonal levels and at different resolutions. On the other hand, there is a low focus on the expendability of the model (E.g., scenario analysis). The manuscript would benefit from a clearer statement of the use the authors want to do of the model. They mention at the beginning the impact that industrial air pollution has on human health, and they create a scenario to test the model's performance. Nevertheless, there is no health impact quantification analysis of model outputs or scenarios.

A more careful choice of the use of the model would give higher focus to the validation and the choice and description of proposed scenarios. If the choice of the model is motivated by "scenarios-impact" analysis then the validation should focus on high-resolution simulations and on a temporal scale that would allow to evaluate the model against national or international threshold limits at a daily/hourly level. Contrarywise, if the intention is to use the model for monthly regional simulations then the validation could be limited to a 3x3km resolution and evaluate the model representation of seasonal average values analysing the impact of scenarios at the annual/seasonal level.

Major Comments:

43 – 46: *The authors apply the Polair3D model over the domain of Canada highlighting that this particular model has seen only a little use over North America and Canada. It should be made clearer why the use of this particular model should represent a step ahead in air pollution research. Several types of CTMs serve for different purposes from different points of view. These can be related - for example - to the representation (or absence) of particular chemical*

mechanisms to describe the chemical life cycle of some pollutants, or to minor computational costs that make the simulations quicker or smaller in terms of storage space. The authors should make clearer and stronger the motivations that led them to 1) choose the Polair3D model and 2) which use it besides the pure capability of representing concentrations of air pollutants (e.g., scenarios, forecasts, mitigation policies testing, transport/trajectories analysis).

While this paper presents model set-up and validation, the ultimate purpose of this air quality modeling exercise is to support the analysis of sector-specific mitigation policies and their benefits from a population health perspective. This entails the ability to run multiple scenarios and the need for high resolution modeling especially in populated areas. POLAIR3D is part of a suite of air quality models, the POLYPHEMUS platform, and lends itself for plume-in-grid and street-in-grid applications. While this paper does not present either application given that the scope is limited to validation and performance assessment of the CTM, our ultimate goal is to capture the spatial variability of concentrations under mitigation scenarios, to support epidemiologic analyses.

In Canada, two other CTM platforms are typically used: the USEPA's Community Multiscale Air Quality Modeling System (CMAQ) commonly used for studies in North America, and GEM-MACH run by ECCC. CMAQ's major difference with Polair3D lies in its management of the aerosols, while the aerosol module in Polair3D is a sectional model, CMAQ is a modal model. Furthermore, CMAQ has mainly been applied at regional scales, i.e., using horizontal spatial resolutions of 36, 12, or 4 km², with few studies at local scales, i.e., with a resolution of 1 km². In contrast, Polair3D has been widely used with horizontal resolutions of 1 km² with robust performance. The ECCC's GEM-MACH model is currently unavailable for academic use and is not set-up for the type of resolution and scenario analysis required for this project. Several studies have shown that CTMs simulating aerosols for exposure and health assessment analyses over urban areas should be conducted at the finest spatial resolution possible, since higher resolution would reduce exposure misclassification for the population, and enhance the accuracy of health risk estimates.

Text summarizing this was added to the manuscript.

For the benefit of the reviewer, we provide below a table summarizing the sub-modules in Polair3D, CMAQ, and GEM-MACH.

Table 1. Comparison of the sub-modules in Polair3D, CMAQ, and GEM-MACH			
	POLAIR3D	GEM-MACH	CMAQ

Model	Polair3D	Global Environmental Multiscale-Modelling Air-quality and Chemistry (GEM-MACH)	Community Multiscale Air Quality Modeling System (CMAQ)
Type	Eulerian CTM	Online model (meteorology and chemistry are handled within a single model)	Eulerian CTM
Meteorology	WRF, MM5	Includes a “physics and chemistry processor” (GEM physics module)	WRF data
Initial /boundary conditions	MOZART, CAM-Chem, GEMS etc. CAM-Chem recommended for recent dates (CAM-Chem is available for download online)	GEM forecast for boundary conditions	CMAQ Chemistry Transport Model (CCTM), GEMS
Aerosol	Choice of several aerosol modules; SOAP (Secondary Organic Aerosol Processor), SCRAM (new, recommended), SORGAM (requires ISORROPIA) SIREAM is a sectional or size-resolved model (the aerosol size distribution is represented by a number of sections)	Model includes: sedimentation, nucleation, condensation, coagulation, swelling, activation, sea-salt emissions, and inorganic gas-particle partitioning, as well as SOA formation Aerosol capability is based on CAM	(AERO7) is introduced in CMAQv5.3; Aerosol module uses ISORROPIA v2.2 in the reverse mode to calculate the condensation/evaporation of volatile inorganic gases to/from the gas-phase concentrations of known coarse particle surfaces CMAQ is a modal model (log-normal distributions, or modes, are superposed to represent the aerosol size distribution)
Chemistry	CB05 (52 species, 155 reactions), RACM (72 species, 237 reactions),	ISORROPIA (heterogenous chemistry) ADOM/ADOM-II	ISORROPIA, and separate aqueous chemistry module. Aqueous chemistry and scavenging is calculated for resolved clouds as well, using the cell liquid water

	RACM2 (113 species, 349 reactions)	(aqueous and gas phase chemistry)	content and precipitation from the meteorological model.
Species	52 species (with CB05 chemistry module)	Aerosols: PM (2.5 and 10), chemical breakdown (SO4, NO3, NH4, EC, pOC, SOC, CM, SS) Chemical species: O3, NO2	Listed here: https://github.com/USEPA/CMAQ/blob/main/DOCS/Users_Guide/CMAQ_UG_ch06_model_configuration_options.md#6.11_Aerosol_Dynamics
Notes	Polair3D is a Eulerian CTM, and can be coupled with chemistry/aerosol modules e.g., SIREAM Polair3D has been widely used with horizontal resolutions of 1 sq km with robust performance	Developed for AQ forecasting (O3, NO2 and PM); developed by and for Canadian air quality analysis	Relies on the open source Sparse Matrix Operator Kernel Emissions (SMOKE) model to estimate the magnitude and location of pollution sources CMAQ has mainly been applied at regional scales(i.e., using horizontal spatial resolutions of 36, 12, or 4 sq km, with few studies at local scales, i.e., with a resolution of 1 sq km)
References	http://cerea.enpc.fr/polyphebus/doc/Polyphebus-1.11-Guide.pdf https://link.springer.com/article/10.1007/s11869-019-00733-5/tables/1 https://link.springer.com/article/10.1007/s11869-019-00733-5/tables/1 https://doi.org/10.5194/gmd-11-2609-2018 https://slideplayer.com/slide/6540524/	https://github.com/USEPA/CMAQ/tree/main/DOCS/Users_Guide https://doi.org/10.5281/zenodo.5213949	

1. Solazzo, Efisio, et al. "Ensemble modelling of surface-level ozone in Europe and North America for AQMEII." *Air Pollution Modeling and its Application XXII*. Springer, Dordrecht, 2014. 351-356.
2. Silveira, Carlos, Joana Ferreira, and Ana Isabel Miranda. "The challenges of air quality modelling when crossing multiple spatial scales." *Air Quality, Atmosphere & Health* 12.9 (2019): 1003-1017.

61 – 64: The performance of a CTM is highly influenced by the levels of primary emissions and by the representation of the regional meteorology. The authors mention that the meteorology used to drive the model in representing the air pollution was taken by WRF, but they don't mention a validation of this meteorology. Analysing the performance of the model at the seasonal level and focusing on primary and secondary pollutants it would be good to understand the levels of reliability of parameters such as surface temperature, solar radiation, wind speed and temperature, and relative humidity. The evaluation of these parameters should be also analysed in the same temporal dimension (e.g., seasonal, annual) of the air pollution concentrations.

As meteorological data largely affect model results, we examined the meteorological input files from the Weather Research and Forecasting (WRF) model, and validated them against observational temperature and wind data. This validation was done at all spatial resolutions, at hourly temporal resolution. Correlations, root mean square error (RMSE) and biases were calculated for temperature, wind speed, and wind directions. This was done at all meteorological stations in our modeling domain: those operated by Environment and Climate Change Canada (ECCC) and provinces. **We have updated the manuscript and added more information regarding WRF configurations.**

As examples, Figures 1, and 2 present a comparison of WRF outputs and observational data at Montreal Trudeau airport at 1km and 3km resolutions. WRF output for October 1 to 14 2018 were compared against observations. Data were collected by ECCC and were downloaded from their publicly available website.

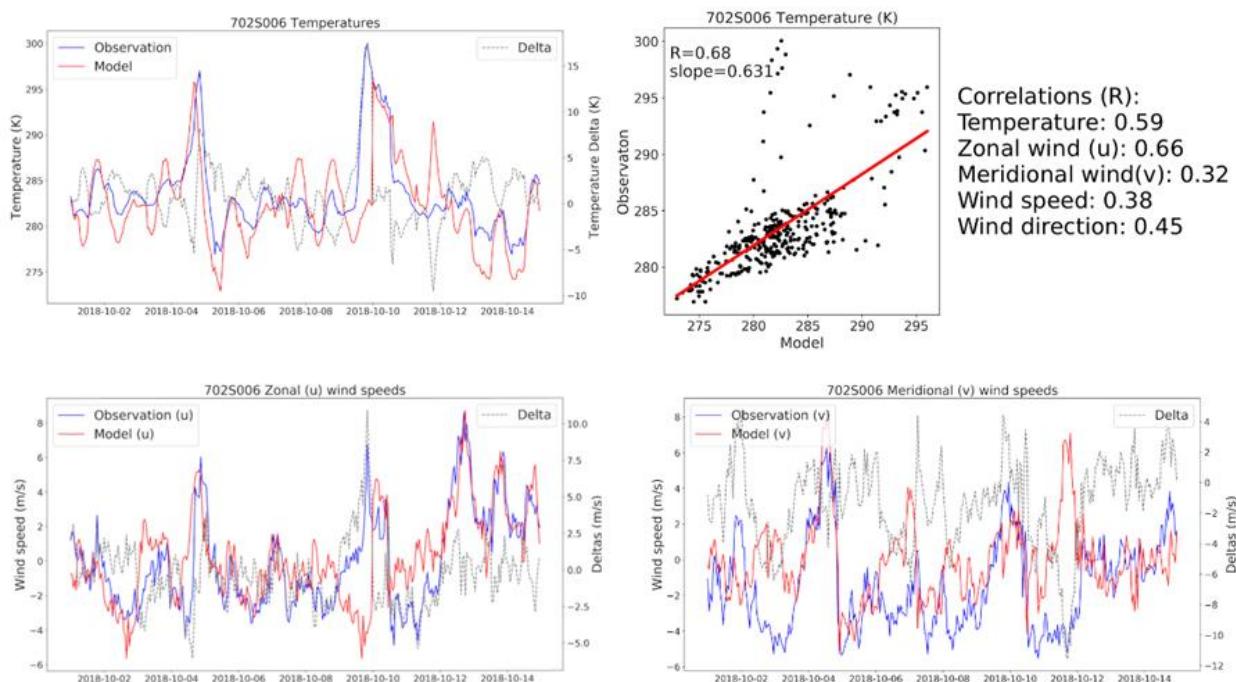


Figure 1. Time series plots of model (at 1km resolution) and in-situ observations of temperature (top left), zonal wind (bottom left), and meridional wind (bottom right) seen at Montreal P.E. Trudeau airport. Temperature correlation plot is also shown (top right). In all the time series plots, observation is indicated by blue, model by red, and the deltas by dashed grey.

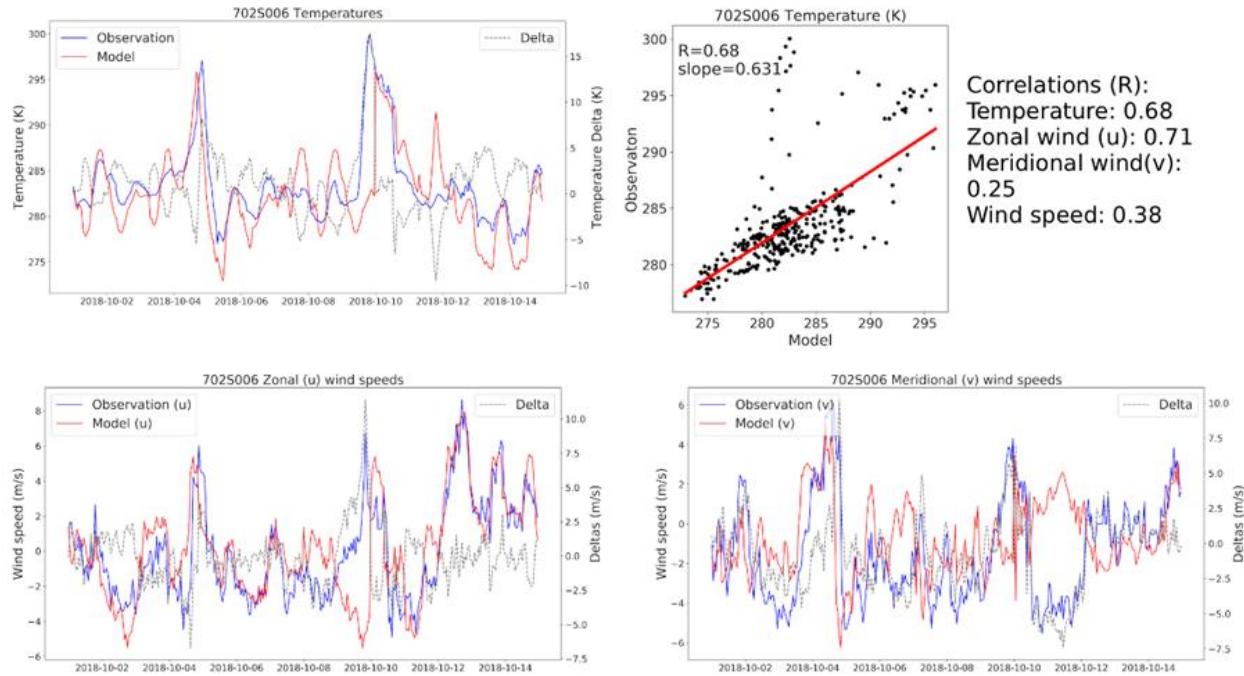


Figure 2. Time series plots of model (at 3km resolution) and in-situ observations of temperature (top left), zonal wind (bottom left), and meridional wind (bottom right) seen at Montreal P.E. Trudeau airport. Temperature correlation plot is also shown (top right). In all the time series plots, observation is indicated by blue, model by red, and the deltas by dashed grey.

92 – 93: Anthropogenic emissions come from NEI 2014 and EPA 2017 inventories. The authors mention using the SMOKE pre-processor and a combination of the “SMOKE-ready” format of these inventories. Are these the most up-to-date inventories to represent the emissions in Quebec? The authors should mention how these two inventories have been processed and speciated before being merged. Is it assumed that – individually – the two inventories have been processed in SMOKE and then merged? If yes, in which way and according to which criteria? Besides the representation of the PM, how the VOCs are represented/speciated in the model? The choice of anthropogenic emissions can be critical in the representation of the final concentrations, and it would be good to have a quantitative analysis (even only in the supplementary material) that shows the annual totals. This is also in light of the scenarios with reduced emissions from Industrial sources. It's good to have maps showing their positions but it would be good to know which percentage of reduction this sector, once deleted, gives on the totals.

Canadian inventory estimates are given in detail in a report by ECCC. We refer to this report in the paper, including a brief excerpt of annual total emissions. We did not include figures in the SI simply because they already exist in the published report, even at a provincial level.
<https://www.canada.ca/en/environment-climate-change/services/air-pollution/publications/emissions-inventory-report-2022.html>

This study uses two emission inventories, the Canadian inventory and the US inventory.

Canada's Air Pollutant Emissions Inventory (APEI) also known as Canadian criteria-air-contaminants (CAC) emissions inventory, prepared and published by Environment and Climate Change Canada (ECCC). The APEI is a comprehensive inventory of anthropogenic emissions of 17 air pollutants such as CO, NH₃, NO_x, PM_{2.5}, PM₁₀, SO₂, and VOC at the national, provincial, and territorial levels and it is compiled from many different data sources. This is the most up to date complete inventory for Canada.

Since the studied domain includes some regions of the United States, the US inventory, also known as the National Emission Inventory (NEI) was processed as well. A combination of SMOKE-ready formats of NEI 2014 and 2017 inventories were taken into account.

The two inventories have been processed using SMOKE individually for each specific source. Through this process, we derived the gridded speciated hourly emission files for each source, then all the area sources were merged into one binary file representing total surface emissions. All industrial sources were merged into one binary file representing total volume emissions. Both surface emissions and volume emissions were used as inputs in the Polair3D model.

For this study, we used the Carbon Bond mechanisms (CB05) for VOC speciation and for the gas-phase mechanism, as well as the AE6 aerosol scheme.

SMOKE-ready format of the Canadian inventory for the year 2015 was used. The APEI is compiled by the Pollutant Inventories and Reporting Division (PIRD) of ECCC. The inventory databases compiled by PIRD are modified by the Air Quality Modelling Applications Section (AQMAS) of ECCC for emissions processing with SMOKE for the Canadian air quality modeling platform. These are the most up to date complete inventories. SMOKE has four individual stages for emission processing including, chemical speciation, temporal allocation, spatial allocation, growth, and control scenario definition. In this project, we consider the three stages of SMOKE emission processing that are shown in Figure 3 to obtain hourly gridded regional emissions for the modeling domain. SMOKE can process different types of emission inventories such as area sources, point sources, and mobile sources. In this study, we have processed all ECCC emission inventories as area sources.

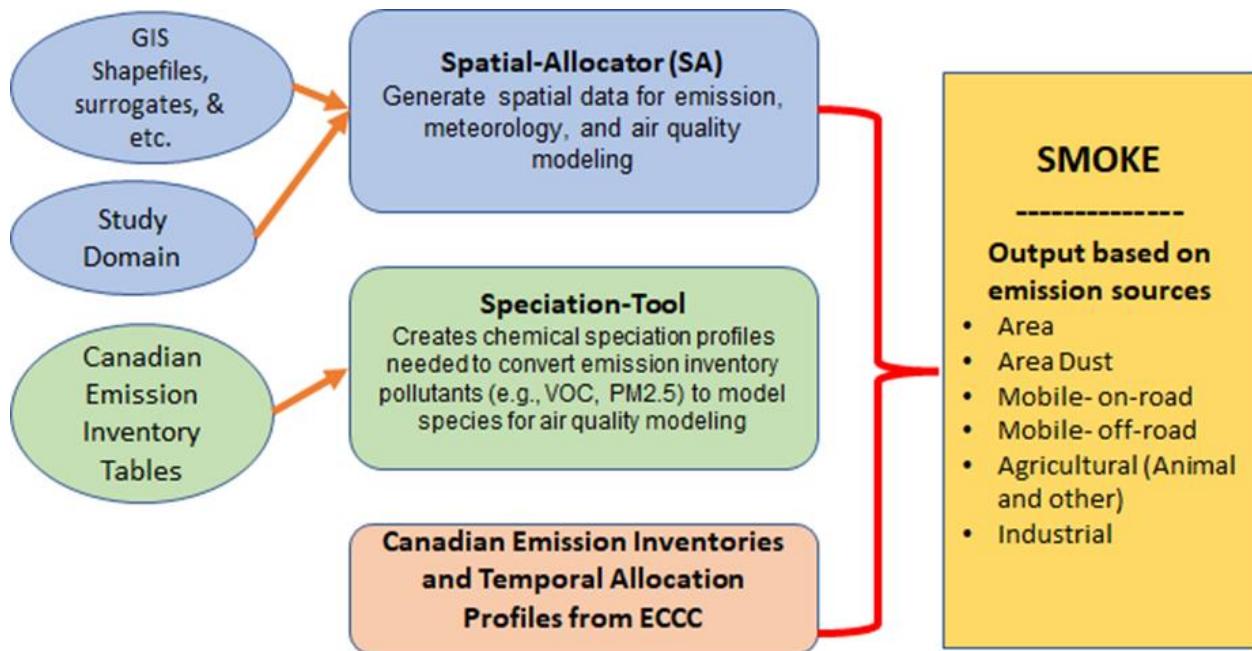


Figure 3. The emissions processing scheme used in this study

113 – 115: The authors mention the use of the NAPS observation sites for the evaluation of the model performance. It would be good to know where these observation points are (maybe in Figure 1) and most of all which type of observation sites these are. Are urban backgrounds and or rural sites? Any of these are road traffic sites? The model performance could sensibly change if different sites are computed together or by type. The suggestion here is to divide the sites by type and perform the statistical analysis again. The evaluation of urban background sites for NO, NO₂ and O₃ could reveal information about how the model represents titration processes in urban environments while in rural areas it could be analysed the impact that biogenic emissions of VOCs have on O₃.

An analysis evaluating model performance at different NAPS site types was added to the revised manuscript. Here, the NAPS designated classifications of regional background (RB), general population exposure (PE), and transportation-influenced (T) were used. Transportation-influenced sites showed the highest correlation for NO₂, NO and O₃, but also had the fewest data points. For O₃, the highest correlation was seen with the RB sites, which is in line with the previous analyses that indicate that the model is able to capture the overall amount of background O₃ that is generated/destroyed. For CO, there were no RB sites available, and both T and PE sites showed similar correlations.

149 – 160: The evaluation of the performance in representing NO, NO₂ and O₃ would benefit from more information about VOCs, and by how the original NOX emissions are partitioned in NO and NO₂. For what concern is the performance of PM2.5 Is there any transport pattern that could influence the seasonal variability in the model performance? The authors mention that the

model performance in summer is lower for all pollutants except for SO₂. Is there any reason related to meteorology or seasonal emissions of SO₂ that could give this?

Adding to the comment above, the site type analyses sheds some light on the model performance of NO_x and O₃. The lack of correlation in PE NO (R=-0.06), along with the PE O₃ R=0.70 and PE NO₂ R=0.48 indicate that the emissions of NO in urban areas may be inaccurate, but NO₂ fares better, and the model is able to still capture the overall urban O₃ trends. For PM_{2.5}, the PE sites had the highest correlation, although the correlation was still relatively low at R=0.38; the model struggles to capture the seasonal and spatial variability of PM_{2.5}.

Across all sources, NO_x emissions were subdivided into nitric oxide (NO) (90%) and NO₂ (10%); NO and NO₂ were speciated within the SMOKE. Please refer to the following documentation: <https://www.cmascenter.org/smoke/documentation/4.5/html/ch02s11.html>

178 – 182: The authors mention that the model performance does not increase with the model resolution for CO and NO₂. This can be seen from the parson coefficient by going to inspect the MB in Table 1 this is always lower in the 1x1km domain. What is not mentioned in the text is that the evaluation of the “3km to 1km” shows higher MB than the “3km”. This could suggest that the agreement in the model performance decreases in urban background sites (that are supposed to be denser in the 1km domain area) and is higher in rural areas. An analysis of the model performance by site type could explain this better.

An analysis at the 3km resolution, but with sites only found in the 1km resolution domain, was done to account for this difference in site selection (see Table 1 in the revised manuscript).

Figure 6: Why the analysis of Ozone is shown only for January when its photochemical activity is lower? Why don't show the same figure for summer and winter?

The figures were swapped for those of July.

Test Scenario: This part of the manuscript would benefit from a deeper understating of the impact of the reduction in industrial emissions on final concentrations and the benefit of the health impact. It's good to know where the emission points are and where the average difference in concentrations is, but it would be also good to know how much change in terms of concentrations at the receptor's sites (observation points). This would require an analysis of the meteorology to understand which sites are downwind and would benefit more from the industrial emission reduction. Additionally, the authors mention the health impact that industrial air pollution has but they don't quantify the impact that the suggested scenario could have on final concentration and health impact.

This run section was run to simply test the model performance under varying emissions, and was not intended as a quantitative sensitivity analysis; sensitivity analysis was not the main focus of this study (that being model validation), it was included as we believed some preliminary analyses of this kind would strengthen the paper.

We have added two more scenarios on top of the “no industry” scenario (turning off emissions from the paper/pulp industry, and turning off emissions from the smelter/refinery industry) to strengthen this part of the paper.

In addition, we have performed additional analyses validating our modeled PM2.5 and NO2 against assimilated monthly ground-level PM2.5 and NO2 dataset products from ACAG (Atmospheric Composition Analysis Group) and CANUE (Canadian Urban Environmental Health Research Consortium) respectively; these datasets are widely used for air quality assessment and exposure studies. The results show fair agreements, indicating that the Polair3D model can be used for health impact analyses. Lastly, please note that we also reran the model for the month of January as we found some aberrations in some of the species in the ECCC offroad emissions (with the emissions themselves, not from our simulations). We reran the month using offroad emissions from April instead (everything else was kept the same), after checking that the offroad values did not vary significantly from month to month (in fact July and April for example were identical).