Reviewer's report on the manuscript by Brett et al. "Investigating processes influencing simulation of local Arctic wintertime anthropogenic pollution in Fairbanks, Alaska during ALPACA-2022", Atmospheric Chemistry and Physics, Manuscript ID: egusphere-2024-1450

The manuscript presents an investigation of the processes influencing dispersion of local anthropogenic pollutants in a sub-Arctic urban setting (Fairbanks, Alaska) during wintertime. Through some detailed analysis of the comparisons between model simulations (using a particle dispersion model) and observations during the ALPACA campaign, the authors explored how the unique local meteorology at Fairbanks in wintertime (dominated by cold, stably stratified conditions) influence the dispersion and transport of pollutants from both surface and elevated sources. Overall, the paper is good and worthy of eventual publication. I do have a few comments (both general and specific) which I hope the authors can address to improve the paper's clarity.

General comments

1. Plume-rise calculation: The use of 12 hourly (03 and 15 local time) radiosonde data at Fairbanks airport for the plume rise (or injection height) calculation is problematic. Why not using the EPA-WRF hourly meteorology at stack locations for this since it is driving FLEXPART?
2. How does FLEXPART-WRF deal with possible plume recirculation (given the topographical feature of Fairbanks) as the model domain (shown in Figure 1) seems to be rather small? I may have missed it – how long is the forward tracer dispersion simulation after particle release?
3. Overprediction of surface SO2: I understand that the authors alluded to vertical mixing being a main factor affecting the model simulation of surface SO2 rather than chemical lifetime (oxidation) in the wintertime at Fairbanks. It might still be interesting to compare modelled SO2 with observed SO2+sulfate (in terms of total sulfur; I believe that sulfate was measured at CTC site as indicated in Simpson et al., 2024). Since the model does not consider sulfur oxidation, this can perhaps provide a check on whether modelled and observed total sulfur is comparable. If they are, this can then lead to questions such as whether oxidation pathways, other than the usual suspects of OH (in clear air) and H2O2, O3 (in aqueous phase), may be at play or whether it could be an emission issue, e.g., a larger portion of sulfur (from space heating, for example) could be emitted as sulfate rather than SO2 perhaps (with reference to Moon et al., ACS EST Air 2024, 1, 139-149). Just a thought.
4. I don't quite see the logic behind the constant emission test being used to illustrate the sensitivity of modelled tracers to meteorology. What the test illustrates is perhaps the impact of emission dependency on meteorology (e.g., dependency of CO emissions from diesel trucks on temperature)? Maybe I misunderstood the CONST-EM run. Were the campaign averaged emissions (or emission rates) applied to surface emissions only or applied to stack emissions also? If applied to stack emissions, were plume-rise calculation done the same way as CTRL run or done differently?

Specific comments

Line 17: Arctic haze occurs primarily during late winter and spring (highest occurrence in April and May) according to Shaw et al. (1995) rather than winter and early spring.

Line 158-160: The problem here is that the buoyancy force is only evaluated at the stack height while the atmospheric stability changes with height as in the case. It is particularly problematic under the strongly stratified situation or with elevated inversion layers. Akingunola et al. (2018) employed a methodology to evaluate the buoyancy force as the plume rises through the vertical column (model levels), which is more computational involved than the capping strategy used here.

Line 168: I don't understand why this is done, by removing negative temperature gradient.

Line 178: The use of +/- 8% of plume height for plume width/thickness – what is the basis for this?

Line 247-251: Were there fog and precipitation events during the ALPACA campaign?

Line 273-277: Are SS and WS regimes determined for the lowest 100 m only?

Figure 5: Are the results shown from the runs conducted with only surface or power plant emissions separately? Also, it may be clearer to denote SS and WS using the notation of filled (solid) and unfilled (open) circles.

Line 368-370: How is wintertime Arctic haze defined here? Are you really considering 0.012 – 0.1 ug/m3 of SO2 and < 0.01 ug/m3 sulfate aerosols as Arctic haze? These are much lower levels compared to the concentration levels (particularly in terms of sulfate) observed during the springtime Arctic haze events at High Arctic sites (e.g., Utqiagvik and Alert).

Line 380-383: The explanation of how the background CO and NOx concentrations are determined is not very clear.

Figure 8(c): Are these supposed to be time series of modelled plume height or the vertical extend of the power plant plumes?

Line 423-426: The difference between the radiosonde wind observation at Fairbanks airport and the LiDAR wind observation close to Aurora power plant would not necessarily affect the modelled dispersion of the Aurora power plant plume. The issue is whether the WRF simulated wind at the Aurora location is different from the LiDAR (close to Aurora) wind observation, since the dispersion calculation depends on the WRF simulated wind not the radiosonde observation at Fairbanks airport (I assume). Of course, the difference in temperature profiles between the airport and the power plant location could affect the plume injection height in this case. If there is a significant vertical wind shear at the power plant location, that would impact the modelled dispersion if the plume were misplaced in the model.

Line 433: I can't see any agreement between the nighttime radio sounding and the Helikite profiling below 270 m from Fig. 7a.

Line 436-437: The statement on the CTRL (with capping) being better than NO-CAP in this case is somewhat subjective. It could be argued that NO-CAP was a bit better than CTRL in terms of plume centerline (or height of maximum concentration) and vertical distribution in Fig.7b (case 2).

Line 444: what about the NOx plume at ~45 m?

Line 451&455: “Fig.D3a” should be Fig.D2a?

Line 459-462: Could perhaps compare the LiDAR aerosol plume with modelled SO2 plumes (given that the aerosols may be dominated by sulfate)?

Line 599: Fig.10 shows that NOx is modestly overestimated in this temperature range (-13C to -23C) without the inclusion of the adjustment for cold temperature.

Line 605-606: Why should the cold temperature effect on NOx emission from diesel vehicles be limited to stably stratified conditions only? In another word, why would the vehicle emissions be dependent on atmospheric stability?

Line 662-668 and Line 727-731: The rather complicated sensitivity to the vertical mixing (via the hmin setting), i.e., conflicting results amongst different pollutants (with enhanced or supressed vertical mixing), may also be an indication of compensating errors in the system.