This work presents a Multi-scale system to predict NO2 at high-spatial resolution. The system combines mesoscale results (CMAQ) with a Gaussian dispersion solver (RLINE). The study deals with the coupling between these two scales. The main innovation is the approach to derive local wind velocity within street canyons. The idea of using CFD simulations to train machine learning models to improve classic Gaussian Dispersion models is interesting and timely. However, some points of the current methodology need to be clarified before publication.

1) As stated in the manuscript, ML methods estimate "the wind vector along X-axis and Y-axis at different heights within the street canyon respectively". However, where is this vertical profile located within the street canyon? Are the edges of the canyon affecting the profile? Are 3D CFD simulations justified?

2) In RLINE, a single wind speed value is associated with each line source. Could you please clarify in the manuscript how this wind speed has been derived for the different presented approaches (CMAQ-RLINE, CMAQ-RLINE-Urban-nc, CMAQ-RLINE-Urban)?

3) How intersections of two different street-canyons are treated?

4) The heat flux condition of the vertical mixing parametrization is different from the one presented in Benavides et al. Also, Benavides's parametrization was adjusted to work with MOST estimation of surface wind speed. How is this parametrization still valid when using a different surface wind speed approach?
Also, to compute the ratio WSsfc/WSbh for the vertical mixing parametrization, how is WSsfc computed along the canyon, given that the ML surrogate provides a single vertical profile?

5) To fairly assess the ML performance, data from a velocity profile for a given set of predictors (Vbgx, Vbgy, z/H, Hl/Hr, H/W,L/W) should not be split between training and test sets. Then for the testing, the ML models will predict the entire velocity profile without seeing any other data point from this profile. This more restrictive data splitting will estimate better the performance in unseen street canyons.

6) How does this approach account for all non-local vehicle emissions?

7) Is the building geometry constraining the Gaussian dispersion?

8) A maximum of 1e4 iterations in the CFD solver are considered to avoid extra CPU time. From the full 1600 CFD simulations, what percentage did not reach the convergence conditions described in the paper?

9) What are the CFD mesh refinements criteria? The CFD mesh should be refined close to walls to capture strong velocity gradients, and it does not seem to be the case.

10) The described boundary conditions are not in accordance with Fig. 2. Also, please report the number of CFD mesh points and simulation time.

11) What is the spatial resolution of CMAQ /WRF?

12) Add ref. for the observation data in Fig. 5. Figure captions should generally be improved to facilitate the reading.

13) Clarify the innovation aspect in the introduction.

14) Section S1 is incomplete. How are Z mix , convective velocity scale w * , surface friction velocity u * , and Monin-Obhukov length recalculated?

15) What are the specific input data used in RLINE?

16) How the facbg is further used in the photochemical scheme? Is [NO2]b = facbg*[NO2]cmaq ? If yes, are O3 and NO also scaled with facbg?