Comment on acp-2022-371
Anonymous Referee #1

Referee comment on "Development and application of a multi-scale modelling framework for urban high-resolution NO2 pollution mapping" by Zhaofeng Lv et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-371-RC1, 2022

The study of Lv et al. presents a multi-scale modelling framework for the simulation of urban scale NO2 and potentially other primary pollutants at high spatial resolution with a focus on traffic-related air pollution. The method combines several different types of models and approaches, a regional chemistry-transport-model (CMAQ), a dispersion model (RLINE), an urban heat island scheme, and machine-learning based simulation of street-canyon flows trained with a CFD model. The overall framework is referred to as CMAQ-RLINE_URBAN.

The overall approach is interesting, but the publication has major deficiencies, is difficult to follow, and leaves many questions unanswered. In my view it cannot be published in the present form but will need substantial improvements.

Major issues:

The individual model components as well as their interplay are very poorly described. Examples:

- The RLINE model is never explained. It remains unclear whether this is Gaussian dispersion or any other type of model. Providing only references without any further details is not sufficient given the fact that this model plays a central role in this study.
- How are emissions released into the model? Is traffic a line source? How are the emissions transported forward and dispersed by the (RLINE) model?
• An UHI scheme is implemented, which "increases atmospheric turbulence intensity around sunset in the afternoon", but it is never explained how this increased turbulence affects the simulation or what is meant by "afternoon". Is the UHI scheme only triggered around sunset? Does it affect the turbulent intensity in RLINE? If so, how exactly? Which localized meteorological parameters are recalculated and how? Is the UHI scheme of Cimorelli et al. (2005) different from the algorithm proposed by Benavides et al., or is the Benavides algorithm based on Cimorelli et al.? The text remains extremely vague despite the fact that, again, this UHI scheme is an essential component of the final model system. Please note that WRF can be run with an urban canopy module (e.g. Barlage et al., 2016; doi:10.1002/2015JD024450), which would alleviate the necessity of implementing an UHI scheme in such a complicated (and unclear) way as done here. Why was this scheme not used to drive CMAQ and to compute the winds and stability above roof level?

• CFD simulations were performed to train a machine-learning based street-canyon flow model (MLSCF) in order to predict airflow in street canyons efficiently. This part of the publication is quite clear, but how the results of the MLSCF are finally applied to compute the dispersion of NO2 is never explained. It should also be noted that the MLSCF model only predicts wind speeds at different locations in the street canyon (in along-canyon and perpendicular direction), but not turbulence, which also varies depending on wind speeds and angle between wind and canyon. The publication mentions the importance of the buildup of a vortex in certain situations, but it remains unclear how the mixing of air pollution induced by this vortex affects the mixing in RLINE. If RLINE is a simple Gaussian dispersion model, how would it be able to represent such a vortex?

• NOx is released by emission sources mainly in the form of NO and then converted to NO2 by reaction of O3. The paper mentions that a "two-reaction scheme" was incorporated into RLINE, but it is not explained which photochemical reactions exactly were considered, how this reaction scheme was implemented in RLINE, or in what form NOx was emitted. Figure 1 suggests that concentrations from vehicles and from background are combined within the "NOx photochemical scheme". Is the scheme applied separately to the two components? How exactly are they combined?

• A "vertical mixing scheme" is mentioned on page 6, which accounts for the "influence of atmospheric turbulence and building geometry on the vertical mixing" and seems to mix background air from roof level into the street canyons. The scheme requires wind speeds at the surface and at roof level, but it is not entirely clear which winds are used here. From the MLSCF scheme? What is the motivation for using the ratio between wind speeds at roof level and street level to compute the contribution of background air? I can only guess, but decisions like this need to be motivated thoroughly.

Why was a resolution of 50 m x 50 m chosen? Note that in Section 2.1 it is suggested that the resolution is only 100 m x 100 m. As mentioned on line 152, the average width of streets in Beijing is about 50 m. Thus, a resolution of 50 m is by far not sufficient to resolve gradients within street canyons.

The machine-learning model is rather simple and little convincing. Complex models are often replaced by artificial intelligence methods using neural networks or Gaussian process...
models, see for example Beddows et al. (2017, doi:10.1021/acs.est.6b05873). A good summary of methods applied in the context of air quality simulations is presented in Conibear et al. (2021, doi: doi.org/10.1029/2021GH000391). Here, a random forest (RF) regression and a MARS approach are used, but these choices are not motivated at all. The RF approach seems to generate quite noisy wind profiles (see Figure 5), but in most cases performs better than MARS. The combination of RF and MARS is referred to as "ensemble learning", but according to page 11, there RF and MARS models have been trained completely independently and there is only a simple switch between the two methods depending on whether the input values are within the range of the predictors used in the training or not. There is a long way from such a simple approach to "ensemble learning".

The introduction section does a fairly poor job in citing relevant literature. Quite many multi-scale air pollution models have been developed recently and also machine learning methods are increasingly used. It is important to place the present study in context and explain where it is different or better than other approaches.

Examples of missing relevant references:
- ADMS-urban is another widely used urban air pollution model not referenced. Note that this model has also been nested into CMAQ (Stocker et al. 2012, doi:). ADMS-urban has been emulated with machine-learning methods by Maillet et al. (2018, doi:10.1016/j.atmosenv.2018.04.009).

Smaller issues:

- I was confused by the usage of the term "receptor". It seems that a receptor can be a grid point but it can also be any other point in the domain, e.g. the location of a measurement station. This needs to be explained much more clearly and earlier in the manuscript. Note that receptor modelling has quite a distinct meaning in air quality modelling and is usually associated with source-apportionment modeling like chemical mass balance or positive matrix factorization.

- The workflow illustrated in Figure 1 is not entirely clear to me: First of all, the arrow between the boxes "receptors in street canyon?" and "receptor information" likely
points in the wrong direction. The most confusing thing is that there is a distinction between "Is a street canyon" and "Receptor in a street canyon". How is it possible that a point can be in a street canyon and at the same time not be inside? Why is there only "road information" needed as input to decide whether we are in a street canyon or not? Shouldn't there also be 3D building data? How the first decision "is a street canyon" is applied is not clear to me at all. Do you choose a road segment and then decide if it is inside a canyon or not? What about points between roads? How do you decide to which road a given point in the city belongs? What about other areas of the city without roads, e.g. parks?

- At many instances in the paper, references to figures, tables and other sections are made in past tense (".. was shown in Figure 1", ".. were discussed in the following section", etc.) but should be in present tense (".. are shown in Figure 1", ".. are discussed in the following section", etc.)

- Data and code availability: Both are only available upon request. Code is only available upon "reasonable request". What is reasonable? Why is the code not made accessible more easily? Advancements in science critically depend on open science and open data.

- Parts of the code seem to be written in Fortran, other parts in R, but it is not clear which. If only CMAQ and WRF are written in Fortran and all other parts in R, then it is not justified to state that a multiscale hybrid model was developed based on Fortran (and R), because there was no development but only application of Fortran code. Whether the model was implemented on Linux (page 5, line 96) or another platform seems irrelevant to me.

- The wind profiles predicted by the MOST scheme presented in Figure 7 look very strange. Apparently, wind speeds reduce to zero at the displacement height, but then jump back to a non-zero value below. Why is this kink in the profile at lower altitude in Figure 7c than in Figures 7a and 7b (despite the higher aspect ratio H/W in case (c) than in (a) and (b)) and why is it not present at all in Figure 7d? Why are the winds at z/H = 1 different between the MOST and the MLSCF schemes? Shouldn't the wind at this level be constrained by the same WRF model output?

- Figure 8 shows differences between simulations with and without the MLSCF scheme. Why are these differences limited to very narrow lines? It is very difficult to see details in this figure. It would be useful to see a zoom into a subregion.