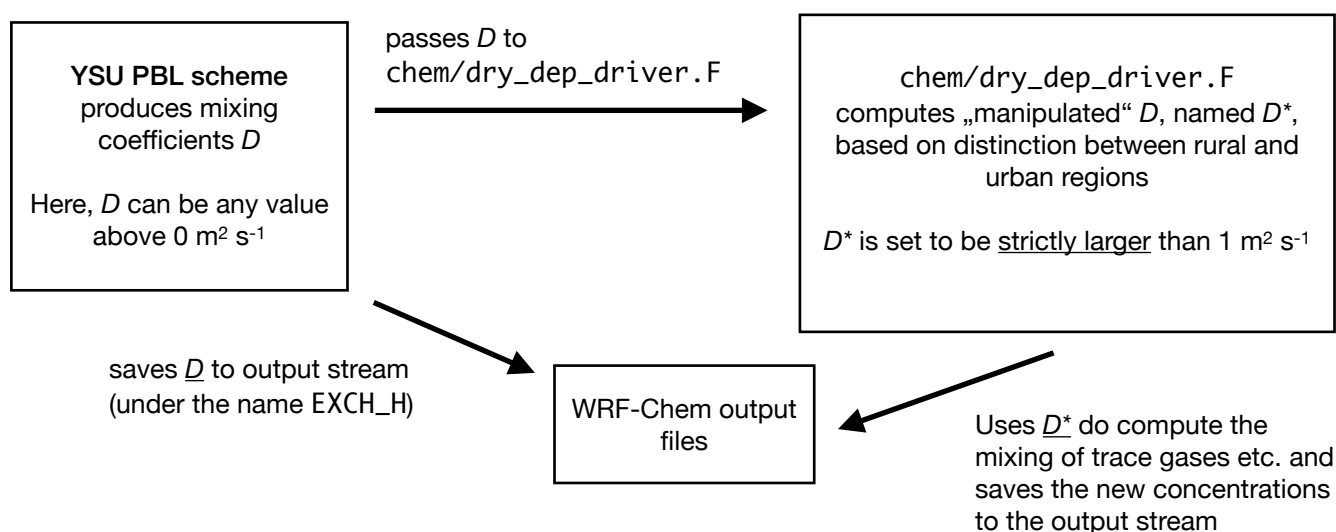


In this paper Yang et al. use the WRF-Chem model to address the connection between horizontal model resolution and mixing strength within the atmospheric boundary layer. Previous literature on this topic had found that the diurnal cycles of modelled trace gases and aerosols show vast discrepancies to measurements, and that this may be caused by faulty parametrizations of mixing. The findings of Yang et al. align well with these previous results.

Because mixing strength appears to be a fundamental issue of the WRF-Chem model, the paper by Yang et al. is a valuable contribution to the field, and suits the scope of ACP. Moreover, the paper is well-written and delivers analyses of the model results. I wish to emphasize that I have truly appreciated the level of depth in these analyses, which reflect the authors' efforts to explain all details of their results based on physical arguments.

That being said, I would generally recommend this paper for publication. However, I do have one central concern of technical nature, which I believe must be addressed before the scientific content of the paper can be evaluated.

Unfortunately, the WRF-Chem source code that handles vertical mixing is far from transparent, and this produces some issues which also affect this article. One of the authors has previously written about the parametrization of vertical mixing (Du et al., 2020), to which the authors refer in line 69. What this paper showed was that PBL schemes, such as the YSU scheme used here, tend to produce unrealistically low mixing coefficients. This aligns well with the overestimation of surface-layer black carbon during nighttime. The problem is, that the associated mixing coefficients are not only obtained via a parametrized PBL scheme, but further manipulated in the mixing routine of the file `chem/dry_dep_driver.F`. Here, the mixing coefficients are clipped to empirically chosen thresholds of $1 \text{ m}^2 \text{ s}^{-1}$ (over rural regions) and $2 \text{ m}^2 \text{ s}^{-1}$ (over urban regions). The distinction between rural and urban regions is made based on the local CO emission strength. The corresponding source code comments indicate, that this clipping is indeed intended to prevent unrealistically low mixing coefficients. I have outlined a sketch of the internal procedure below:



In summary, if the YSU PBL scheme computes a mixing coefficient of e.g. $D = 0.5 \text{ m}^2 \text{ s}^{-1}$, it is set to either $1 \text{ m}^2 \text{ s}^{-1}$ if the corresponding model pixel is considered „rural“ or $2 \text{ m}^2 \text{ s}^{-1}$, if the pixel is considered „urban“, while mixing coefficients above $2 \text{ m}^2 \text{ s}^{-1}$ remain generally unchanged.

This procedure obviously affects the paper by Yang et al. in two regards:

- 1) Because the distinction between rural and urban regions is made based on CO emission strength (not land-use data), the parametrization of mixing implicitly depends e.g. on the resolution of the *emission* data, not the actual model resolution.
- 2) The mixing coefficients that WRF-Chem uses to compute the dispersion of tracers (named D^* above) are not the same as the coefficients written to the output stream (named D above). Figure 7 of the paper shows D , which is why the values are far below the lowest clipping threshold of $1 \text{ m}^2 \text{ s}^{-1}$. This has severe implications, because D goes as low as $0.1 \text{ m}^2 \text{ s}^{-1}$, which is an order of magnitude below D^* .

The authors cannot be blamed for this oversight, because the relevant source code is hard to find. Ironically, I myself have stumbled upon it after reading the paper by Du et al. (2020) and attempting to retrace the full mixing procedure of the WRF-Chem model. Our article on this issue can be found under

Kuhn, L., Beirle, S., Kumar, V., Osipov, S., Pozzer, A., Bösch, T., Kumar, R., and Wagner, T.: On the influence of vertical mixing, boundary layer schemes, and temporal emission profiles on tropospheric NO₂ in WRF-Chem – comparisons to in situ, satellite, and MAX-DOAS observations, *Atmos. Chem. Phys.*, 24, 185–217, <https://doi.org/10.5194/acp-24-185-2024>, 2024.

Please note, that the exact routine described above is subject to further uncertainties, including:

- 1) Slight changes in `chem/dry_dep_driver.F` depending on the version of the WRF-Chem model (these can be traced on the WRF github page). Particularly in commit bda151f, made on 10 January 2022, the routine was changed to exclude the lowest layer and only be effective in the layers above.
- 2) Possible differences to the USTC version of WRF-Chem used in the paper

In order to proceed, I would recommend the following to the authors:

- 1) Add a discussion on the mixing routine, perhaps in sect. 3.2.2 of the paper. I believe a plain-text explanation is sufficient. This should contain a clear outline of the exact model setup the authors have used (e.g. whether they have adapted the adjustments mentioned by Du et al., 2020, in which case the specific numerical values of the adaption should be given).
- 2) Modify Figure 7 to show the values of D^* (see above) in addition to the existing panels. This can be achieved with a simple thresholding applied to the data already shown, but it requires that the authors know exactly which clipping thresholds were in effect during their simulation.

- 3) Modify all corresponding discussions referring to the low values of D (particularly below $1 \text{ m}^2 \text{ s}^{-1}$) in sect. 3.2.2 as to reflect the difference between D and D^* . Please ensure that the line of physical reasoning (which I find to be a very strong point of the paper) remains consistent upon this modification.