Response to review 1

Dear Referee,

Thank you for your detailed and constructive comments aimed at improving our manuscript. We appreciate the time and effort you have taken to provide thoughtful feedback. Below, we provide our detailed responses to your comments and suggestions to improve our manuscript (in blue):

1. **The paper gets into too many very technical details about implementations, while some important characteristics of the simulations are missing. With a minor effort, one can calculate DALES domain is about 1100x1500 cells. How many layers are there, and what are they? How are the boundary conditions for mean wind, for turbulence, and for tracers implemented both in lateral and vertical directions? How much spin-up does the simulation need to become realistic?**

   Response: We thank the reviewer for their thoughtful feedback. We agree that providing more clarity on the simulation characteristics will enhance the understanding of our simulation setup.

   Indeed, the paper lacked information about the vertical grid structure. Therefore, we have added this information to the text (see L407-411): The vertical resolution ranges from approximately 20 meters (within the ABL) to a few hundred meters (in the upper troposphere). This is due to the use of a stretched vertical grid with 128 layers, an initial layer thickness of 25 meters ($dz_0 = 25$), and a stretching factor of 0.017 ($\alpha = 0.017$), which causes the layer thickness to increase geometrically with height.

   Regarding boundary conditions, periodic lateral boundaries are used in the DALES setup in our study. This periodicity applies to mean wind, turbulence, and tracers, meaning that what exits the east boundary re-enters at the west boundary and vice versa (the same applies to the north-south boundaries).

   In the vertical direction, the bottom boundary is a surface flux boundary, where fluxes are computed using the Land Surface Model (LSM) incorporated into DALES. This ensures that surface heterogeneity is properly accounted for in the simulation. The top boundary includes a damping sponge layer, which gradually reduces turbulence and tracers to minimize artificial reflections. These details have been added to the main text (see L199-202).

   Regarding spin-up and initialization periods that should be discarded from the analysis, after further discussion, we decided to adjust the analysis period. The periods to be discarded include not only the DALES initialization phase (i.e., the time required for fields to grow from small turbulent motions to fully developed turbulence), which typically lasts for the first 3 hours (e.g., https://gmd.copernicus.org/articles/12/5177/2019/gmd-12-5177-2019.pdf), but also periods with a stable atmospheric boundary layer (SABL) and HARMONIE forecast initialization.

The SABL (22-6 UTC) time period was excluded because DALES currently becomes less accurate when simulating the stable ABL at this 100 m horizontal resolution. At present, only simulations with horizontal resolutions finer than 10 m provide reasonable representations of the SABL (e.g., https://link.springer.com/article/10.1007/s10546-020-00558-1).

Additionally, the DALES meteorological boundary conditions are derived from the HARMONIE forecast, which resets daily as a new forecast is used. Therefore, when using the HARMONIE forecast for boundary conditions, it is generally advisable to exclude the first 4-6 hours from the analysis (https://gmd.copernicus.org/articles/17/2855/2024/gmd-17-2855-2024.pdf).

Based on this, our evaluation focuses on the diurnal period. We have decided to include June 25 in our analysis (which was previously fully excluded) while excluding the 23-6 UTC period from each simulation day to minimize errors caused by initialization and the well-known limitations of the current simulation approach. This has been described in the main text (see L420-430).

I could not find how surface emissions are handled. Are they just injected into the lowest LES layer?

Response: We thank the reviewer for this question. This information is provided in the main text (see L222-229 for area emissions and Sect. 4.1 for point sources).

We use different approaches to incorporate emission data into DALES. Since plume parameters for area emissions are not available and plume rise cannot be calculated, we compensate for this by following the methodology from Brunner (2019), which provides vertical distribution profiles for certain SNAP categories that are not emitted at the surface. For these categories, we apply emission profiles that redistribute surface emission values between ground level and a height of 150 m, as discussed in Brunner (2019). For other categories, such as SNAP 5, 7, and 10, emissions are injected into the lowest LES layer, as these categories correspond to surface emissions.

For point sources, we utilize available information on stack heights and plume parameters to calculate the vertical distribution (i.e., the plume bottom and top, between which emissions are injected) online using an algorithm based on Gordon et al. (2018) and Akingunola et al. (2018).

There is little information on the reasons why specific technical decisions were taken. It is not clear why one would need to drive an ecosystem model for CO2 fluxes with high-resolution LES fields? Are the features such as second-scale updates of meteorology, clouds, physics of the ABL really needed to simulate CO2 fluxes by ecosystem, or are they just an overhead for such simulations? Same applies to plume-rise calculation.

Response: We appreciate the reviewer for raising this important point. We want to clarify that we are not running an LES to provide details to an ecosystem model that we consider inherently necessary. Rather, we use the output of an ecosystem model in combination with an LES to reproduce CO_2 observations, ensuring consistency in atmospheric CO_2 simulations when the ecosystem model is driven by high-resolution LES.

The decision to drive the ecosystem model for CO_2 fluxes in LES is motivated by the need to resolve fine-scale atmospheric processes that impact CO_2 dispersion and exchange. DALES incorporates a land surface model that simulates ecosystem fluxes (within the A-gs scheme), including CO_2 uptake through photosynthesis and release via soil respiration. Traditional mesoscale or global models often rely on parameterized boundary layer dynamics, which can introduce significant uncertainties in CO_2 flux estimations, particularly under heterogeneous land-cover conditions or complex meteorological scenarios. Additionally, features such as second-scale updates of meteorology and clouds are not merely an overhead but are essential for capturing rapid fluctuations in radiation and turbulence, both of which strongly affect photosynthesis and respiration rates.

These rapid fluctuations (from seconds to minutes) have been observed and lead to significant variations in heat, moisture, and CO_2 fluxes, as demonstrated in the figure below (Vilà-Guerau de Arellano et al., 2020):

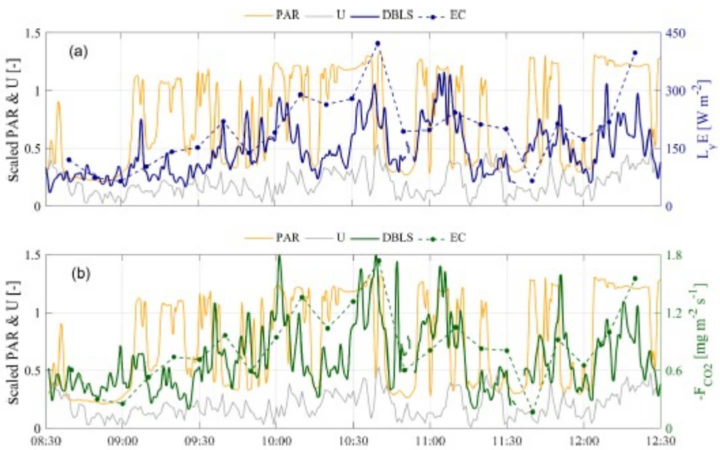


Figure 12. IOP 2 (15 June 2018) time-series of **(a)** latent heat fluxes (LvE) at 1 min intervals with a displaced-beam laser scintillometer (DBLS) and at 10 min intervals with an eddy covariance system (EC) combined with scaled time series of photosynthetically active radiation (PAR, scaled by 1500 μmol m^{-2} s^{-1}) and wind speed (U, scaled by 6 ms^{-1}). Panel **(b)** is the same as **(a)** but for the CO_2 flux (F_{CO_2}).

For example, cloud cover alters the amount of diffuse radiation, which in turn influences canopy light-use efficiency and CO_2 uptake. Similarly, resolving the sub-daily evolution of the boundary layer ensures a more realistic representation of plant stomatal responses and soil respiration fluxes, which are highly sensitive to temperature and moisture conditions. Thanks to new simulations that explicitly resolve clouds at high spatial and temporal resolution, we can improve key drivers of photosynthesis: direct/diffuse radiation, temperature, and water vapor deficit. We have briefly addressed the motivation for using an ecosystem model in LES in the introduction (see L64-70).

Regarding the plume-rise calculation for point sources, we use a parameterization approach because the spatial resolution of used LES grid (100 m) is not fine enough to explicitly resolve narrow plumes. As mentioned in the paper, explicit representation would require prescribing a heat source at the chimney top, allowing LES to compute heating-induced buoyancy effects. However, this approach is only effective if the grid is very fine (<50 m) and the plume remains narrow compared to the grid. This is planned for future development. For our current simulation setup, parameterization remains the more suitable approach for efficiently representing plume rise above stack height when using hectometer-scale horizontal resolution.

Thus, in our case, incorporating an ecosystem model within DALES is not merely an added complexity but a necessary step to ensure a physically consistent representation of CO_2 exchange processes. Further elaboration is required to better understand how the main drivers of photosynthesis change under different conditions.

Reference: Vilà-Guerau de Arellano, J., Ney, P., Hartogensis, O., de Boer, H., van Diepen, K., Emin, D., de Groot, G., Klosterhalfen, A., Langensiepen, M., Matveeva, M., Miranda-García, G., Moene, A. F., Rascher, U., Röckmann, T., Adnew, G., Brüggemann, N., Rothfuss, Y., and Graf, A.: CloudRoots: integration of advanced instrumental techniques and process modelling of sub-hourly and sub-kilometre land–atmosphere interactions, Biogeosciences, 17, 4375–4404, https://doi.org/10.5194/bg-17-4375-2020, 2020.

2. **There is not much info on the computational costs of the DALES CO2 simulations. How many cores and hours needed to simulate one hour of the dispersion? How well it scales with number of cores? How much is the overhead for CO2 with respect to plain DALES? What are cost-benefits of using LES vs using e.g. LOTOS-EUROS at similar resolution for practical applications?**

Response: We thank the reviewer for this insightful question. Indeed, we did not include a detailed discussion of computational costs in the manuscript to not overweight it, and we appreciate the opportunity to clarify this aspect.

The computational cost of simulations with DALES depends, as in many other models, on multiple factors, including domain size, grid resolution, physics parameterizations, and general hardware performance (CPU or GPU using). In our study, we used the following setup:

Domain size: ~172.8 km × 115.2 km
Grid resolution: 100 m in the horizontal (1728 × 1152 grid points) and 128 levels in the vertical stretched grid (~20m resolution within ABL, and corses above, geometrically increasing).
Number of processes: 432 MPI tasks, with a decomposition of 18 × 24 processes
CO2 transport simulated using four scalar tracers (with different CO2 contributions as described in our paper).

Regarding performance, for our simulation setup and domain size, the wall-clock time for one hour of simulation is approximately 1.5-2 times real-time on our specified system and CPU-based simulation (GPU-based free version of DALES is in development). We did not explicitly measure the impact of additional scalar tracers on computational cost, but based on prior experience, adding four passive tracers to a model primarily increases memory usage and I/O operations rather than significantly affecting computational cost. Since these tracers do not involve additional chemistry or deposition processes (only ecosystem fluxes and plume rise for point sources) their impact on the overall model dynamics is relatively modest. However, the increase in I/O and storage usage become a limiting factor, especially with high-frequency output, which is why the applied the limitation on the number of tracers to managing computational resources.

Concerning scalability, DALES is designed for distributed-memory parallelism using MPI. While strong scaling efficiency decreases at very high core counts due to communication overhead, DALES generally scales well up to several hundred cores for domains of this size.

The choice between LES (e.g., DALES) and an Eulerian chemistry-transport model like LOTOS-EUROS should depend on the research question and practical application. In addition, it is worth mentioning that LOTOS-EUROS cannot be run in LES mode. Therefore, the physics and dynamics in the model do not support achieving the same resolution as in DALES. In our case, we focus on the dynamics of CO2 dispersion within the ABL and provide more reliable information of radiation, temperature and water vapor deficit fluctuations for the photosynthesis and soil moisture, making LES the preferred approach. Moreover, with using LSM with LES allows having more detailed soil temperature and moisture in the calculation of surface fluxes.

Some of the used datasets have been declared public, but the procedure of obtaining them has not been clearly described. "The traffic data shapefile is provided by the Dat.mobility company and can be requested from RIVM." Publishing these assets via some open data publishing service or clearly identifying the specific dataset and providing a contact or URL would make it more clear.

Response: We thank the reviewer for pointing this out. We agree that making the NOx traffic data available via Zenodo would improve accessibility. Therefore, we have included the shapefile in the Zenodo archive, along with the 'TVtot_N_sum' variable, which represents the total NOx emissions from all three vehicle types (light, medium, and heavy), and 'Shape_Length'. The dataset can be accessed here:
Doyennel, A. (2025). Annual NOx emission traffic shapefile (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14961517 We have added this link to the main text (see L383-384).

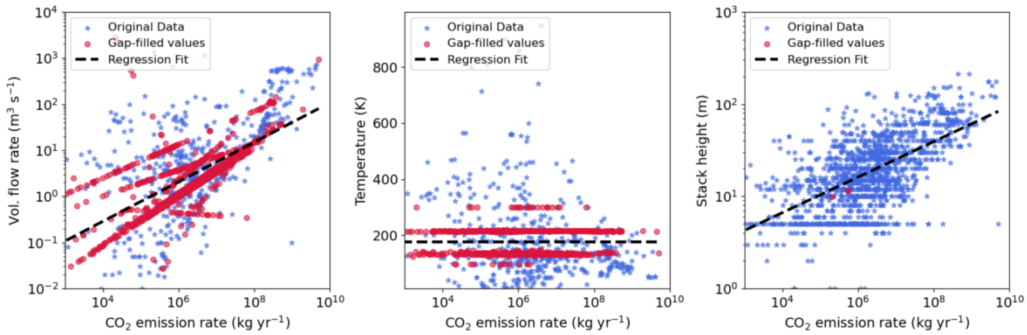
The text is full of vague constructs that do not bring much information: "high-accuracy", "realistic approach", "fine resolution" etc. In a scientific paper each of such constructs if used has to have a very specific meaning.

Response: We appreciate the reviewer's feedback and have carefully revised the text to replace vague or subjective terms such as "high-accuracy," "realistic approach," and "fine resolution" with precise quantitative values, comparative statements, or references to established standards (see e.g., L530-550 and L582-603).

Specific comments:

1. **Table 1 leaves a question on the quality of the fits used. It could be better replaced with regression plots used for get the fit. Then one could judge how good or bad the used approximation is. A simple three-panel figure would do: (a) log(flow rate) vs emission rate (of which species?); (b) Exhaust temperature vs emission rate for specific processes; (c) log height vs emission rate.**

   Response: We thank the reviewer for poining this out. We agree that the figure may be more representative than the table 1 in the paper to show the results of gap filling. Therefore, we replaced Table 1 in the paper with the figure presented below (see L308-309):

   The figure presents scatter plots of three plume parameters: volume flow rate (a), temperature (b), and stack height (c) as a function of emission rate (kg yr^{-1}). Each subplot combines the original data (blue stars) with the gap-filled values (red dots). A general regression fit (black dashed line) is also included in each panel.

   The volume flow rate (a) and stack height (c) show a generally positive relationship with emissions, captured by the regression trend whereas temperature (b) appears more scattered, with the regression fit suggesting a weak dependency on emission rate.

   Note that the resulting fit represents a general trend across the entire dataset, based on the relationship between the emission rate and the corresponding plume characteristics.
   It is important to mention that regression fitting is carried out separately for each subcategory of emission sources, such as chemistry, construction, oil, electricity, etc., with distinct fittings for each subcategory.

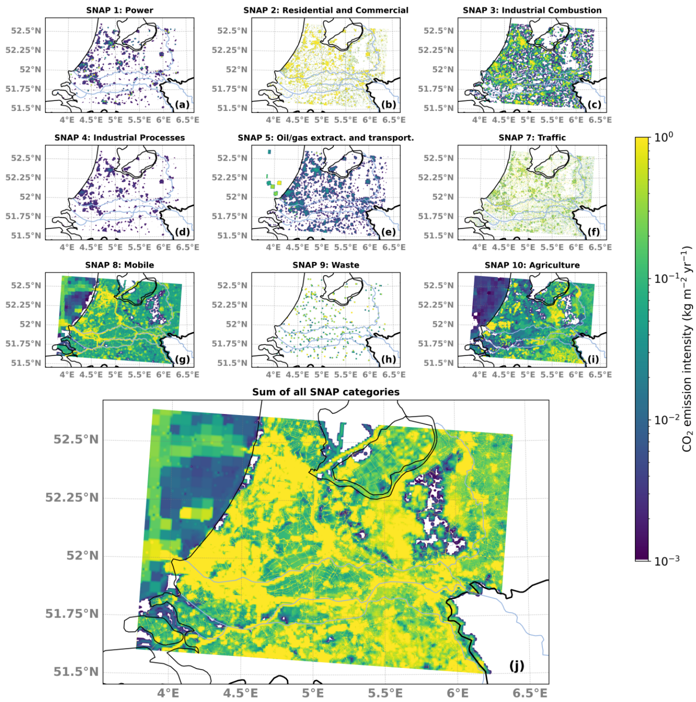
2. **Fig1 It might make sense to mention that SNAP8 includes shipping. SNAP10 has some marine emissions. What is the marine agriculture ?**

   Response: We thank the reviewer for pointing this out. Indeed, we agree that the term 'marine agriculture' does not exist. In our study, we categorized fisheries-related emissions under SNAP10 (Agriculture and Nature) rather than SNAP8 (which includes shipping). This explains the presence of some emissions in SNAP10 over the North Sea. We have updated Table 1 to explicitly mention this classification in SNAPs used in our study (see L118-119). This choice was made because fisheries activities, including fuel use for fishing vessels and fish processing, might be more closely linked to the food production sector rather than transport activities. While these emissions could also be classified under SNAP8 due to the use of vessels (in the more recent emission inventories), the broader link to agriculture and resource extraction justified their inclusion in SNAP10. Some emission inventories (e.g., EMEP/EEA Air Pollutant Emission Inventory Guidebook) include fisheries emissions under agriculture and land use categories, particularly when they are reported together with other land-based food production emissions. Nethertheless, we recognize that classification choices sometimes may vary and appreciate your concern.

3. **Fig1 and Fig2 are labeled to have emissions in kg/year. Normally I would guess that it was meant kg/year/cell, however the figure has at least two distinct cell sizes, which is confusing. Should it be kg/year/ha then? It might make sense to mark the LES domain on these maps, to indicate that missing emissions from Germany (SE corner of the maps) do not affect the simulations.**

   Response: We appreciate the reviewer's comment. The emissions shown in the figures are indeed reprojected onto the target LES grid and are presented in kg/year per grid cell, with each grid cell being 100 × 100 m. The apparent differences in resolution arise because some categories undergo refinement, while others are simply reprojected; since the original resolution is coarse, the coarse grid structure remains visible. Additionally, over the ocean, the original resolution of emission inventory is 5 km instead of 1 km over land.

   We agree that specifying this more explicitly would improve clarity. Therefore, we have converted the emissions to units of kg/year per m^2. Additionally, we have zoomed out the map to show the full extent of the simulation domain and to clarify that emissions from Germany (in the southeastern corner) are absent in the simulation. However, emissions outside the domain should be accounted for in the CAMS EGG4 dataset, which was used for the CO_2 background concentration in DALES. The revised figure is provided below (see L350–351).



4. **l215. The shape file has not been specified properly.**

   Response: We thank the reviewer for pointing this out. For refining, the road transport (SNAP 7) category, we use a NOx annual emission from the Dutch road network provided by RIVM. Specifically, we use a nationally comprehensive road network with attributes such as road segment length and NOx emission intensities from different vehicle categories: light vehicles (LVtot_N), medium vehicles (MVtot_N), and heavy vehicles (ZVtot_N) combined into 'TVtot_N_sum' (sum of NOx emissions from all these vehicle types). The shape file is placed to Zenodo archive and can be accessed here: Doyennel, A. (2025). Annual NOx emission traffic shape file (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14961517. We have added the link to this repository to the main text (see L383-384).

5. **l220 What is the advantage of downscaling 1km traffic emissions instead of just assigning CO2/NOx emission factors per vehicle type and use NOx inventory? At least the latter approach would give more consistent map.**

   Response: Using high-resolution NOx traffic data for downscaling emissions allows for a more detailed and accurate representation of emissions at the local level. This method captures variations in traffic intensity and vehicle types, particularly in areas with heterogeneous traffic flows, which is critical for improving the precision of CO2 emissions in our model. In contrast, relying solely on emission factors and an NOx inventory might provide more uniform emissions map but lacks the granularity needed to represent local variations accurately.
   Worth saying here is that emission factors can also be differentiated further within the modeling setup, but that the focus of our paper is on downscaling to LES resolutions.

6. **l260 What is the benefit of driving A-gs model with a few-second resolution? Is there any benefit in having it online with turbulence-resolving model?**

   Response: We thank the reviewer for this question. The benefit of driving the A-gs model with a few-second resolution lies in capturing rapid fluctuations in environmental conditions that directly affect plant physiological responses. For instance, as shown in Figures 8 and 9 from https://bg.copernicus.org/articles/21/2425/2024/, short-term variations in cloud cover, temperature, and vapor pressure deficit (VPD) lead to dynamic changes in net ecosystem exchange (NEE) and stomatal aperture. These rapid responses are typically averaged out in mesoscale models, which operate at coarser temporal resolutions and do not run in LES mode.

   By coupling the A-gs model online with a turbulence-resolving LES model, we can explicitly resolve sub-minute-scale interactions between the atmosphere and vegetation, rather than relying on parameterizations. This provides a more accurate representation of plant physiological responses under rapidly changing conditions, improving our understanding of carbon flux variability at fine spatial and temporal scales.

7. **l287 Is the model really able to conduct simulations and examine specific aspects of anything? I would say it should be a modeler..**

   Response: Agreed, we reconsidered this part (see L203-206).

8. **l290-310: Sec 3.1 describes unit conversion and linear interpolation in time. Both are quite trivial and, probably not worth the paper space. Same applies to appendix B.**

   Response: Agreed, we omitted this parts from the text (see L210-227).

9. **Sec 3.2 describes quite standard treatment of plume rise. A reference to Gordon (2018) and Akingunola (2018) would be sufficient to describe that, unless something different has been implemented in DALES. In the latter case a brief description with highlighting the differences would be needed.**

   Response: Agreed. We have significantly shortened this discussion, removing all expressions already presented in Gordon (2018) and Akingunola (2018) (see L229-255). However, we kept expressions 2 and 3, as they are not explicitly provided in Gordon (2018) and Akingunola (2018) in the form we presented them.

10. **Table A1 should have a proper reference.**

    Response: Agreed (see L112).