Response to Reviewers

Manuscript: Hoshyaripour et al. (2025), GMD

We sincerely thank both reviewers for their detailed and constructive comments, which helped us improve the manuscript.

Reviewer comments are presented in blue, our replies follow in black, followed by the corresponding changes in the revised manuscript in italic format.

Reviewer 2:

1. L18: "essential for improving predictions related to weather, renewable energy, climate change, air pollution…" consider changing to: "essential for improving predictions and understanding related to weather, renewable energy, climate change, air pollution…"

- Revised accordingly:

Therefore, accurately simulating atmospheric composition is essential for improving predictions and ***understandings*** *related to weather, renewable energy, climate change, air pollution, and associated health impacts.*

2. L59-60: "OEM enables efficient processing of emissions that are constant in time or changing only temporally, but not spatially" This is not clear to me (and may also bother other readers): if different sources are varying differently temporally, it means emissions overall are changing spatially too - please clarify this.

- The reviewer is right that this is not sufficiently clear. The spatial patterns of the emissions of individual categories are constant, but indeed the patterns of the total emissions (sum over all categories) usually change with time because the temporal profiles of the individual categories are different. We changed the text as follows:

OEM enables efficient processing of emissions which can be represented by adding up individual source categories, with the emissions from each category being fixed in space but varying over time.

3. While the paper excels at describing what has been implemented, it could be strengthened by briefly showing why some of the new developments matter in a more quantitative sense. For example: Section 2.4 (Sea Salt): The new Grythe et al. (2014) parameterization is included to better represent SST dependence. A simple zonal-mean plot or a brief statement quantifying the typical change in sea salt emissions or burden in tropical regions compared to the older scheme would be highly illustrative. Section 4.4 (Subpollen Particles): The parameterization for SPP release is described. It would be beneficial to include a sentence stating the typical order-of-magnitude contribution of SPPs to total aerosol number concentration or CCN in relevant regions during pollen season, even if citing another study (Werchner et al., 2022?). Section 3.2 (Detailed

chemistry mechanisms): a brief comment on the typical computational cost increase when moving from a simplified chemistry scheme (like Linoz) to a full mechanism (like MOZART-T1) would provide valuable context for users planning simulations. A percentage increase in runtime, similar to that provided for LINOZ in Section 3.1.3, would be sufficient.

- For the sea-salt emissions, we agree that assessing the impact of the Grythe et al. (2014) parameterization would indeed be valuable. However, a meaningful detailed quantitative comparison requires a comprehensive analysis of the full emission–transport–deposition cycle, rather than an isolated examination of the SST dependence. In addition to the different formulations, the two parameterizations also differ in their particle-size distributions, which further complicates a direct and fair comparison. Such an in-depth evaluation is beyond the scope of the present paper, whose primary aim is to document the model developments and implementation. We therefore prefer to address this analysis in a dedicated follow-up study, which is already underway. We have expanded Section 2.4 with a short qualitative description of the key conceptual differences between the two parameterizations, focusing on their treatment of whitecap coverage, particle-size distribution, and the explicit SST dependence introduced in Grythe et al. (2014).

*MMS and G14 sea-salt emission schemes differ not only in their whitecap formulations but also in their treatment of particle-size distributions and SST-dependent scaling (*Grythe et al., 2014; *Barthel et al., 2019; Li et al.,2024). Barthel et al. (2019) demonstrated that SST corrections can substantially reduce coarse-mode concentrations and may even have a larger impact than switching between source functions. They also found the strongest differences for particles larger than PM2.5, with SST effects further amplifying these differences. These insights highlight that the structural contrasts between MMS and G14 schemes, particularly the inclusion of SST dependence and the size-resolved flux formulation, can significantly influence emitted mass. While a quantitative evaluation is beyond the scope of this study, this context helps to clarify the expected behavior of the new G14 implementation.*

- For the subpollen particle (SPP) parameterization, we have now added a brief statement summarizing the typical magnitude of SPP contributions to total aerosol number and CCN concentrations based on literature values (Werchner et al., 2022):

Werchner et al., (2022) reported for a case study that *SPP concentrations (only used as INP, not as CCN) vary between 10^2 and 10^6 m^{-3}, with a mean value of $4x10^3$ m^{-3} (especially relevant in warmer levels fit for biological ice nucleation), while mean pollen concentration amount to $3.4x10^{-3}$ m^{-3}.*

- In Section 3.2, the following sentence was added to address to question about the typical computational cost increase for full chemistry:

A full-chemistry simulation with MOZART-T1 increases the total runtime by roughly a factor of 10 compared to an ICON simulation without ART (tested on an HPC system with AMD Rome nodes with two AMD Epyc 7742 64-core CPU sockets each), reflecting not only the *computational cost of the chemical mechanism but also the additional overhead from tracer transport, emissions, deposition processes, and model output.*

4. Figure 6: The inner and outer circles are unclear (visually and also in terms of values). For the winter plot (left), the outer rings mostly match the surrounding areas on the contour map but for the summer plot (right) these outer rings are consistently of a lighter shade than the surrounding values in the map which suggests some issue with sampling - please double check. If I disregard the outer rings and only compare the inner circle values with the surrounding values on the simulated map, I see a better model-obs agreement. However, when comparing the inner circles with outer rings, it looks like the model is underestimating surface ozone in both winter and summer. This underestimation doesn't sit well with the broader context of basically all global and regional models overestimating Northern Hemispheric surface ozone (e.g., Young et al., 2013; 2018; Ansari et al., 2025; Nalam et al., 2025, Gao et al., 2025). I suggest that the authors make this figure simpler by only showing one solid circle representing only observed values, and include the overall mean bias, RMSE, and correlation coefficient r for both seasons somewhere in the figure and the text. Accordingly, the text that "the model accurately reproduces…" should be made more nuanced and discussed in the broader context of the aforementioned papers. The authors must also mention which emission inventory was used for these LAM simulations over Europe. The authors should discuss potential reasons for O3 underestimation.

- We adapted the Figure to only show the observations as solid circles including mean bias, RMSE and correlation coefficient. The underestimation noted by the reviewer was caused by missing biogenic emissions of certain compounds in the model. These emissions have now been included, and a new simulation has been performed. The updated results show a slight overestimation, rather than an underestimation, during the summer period.

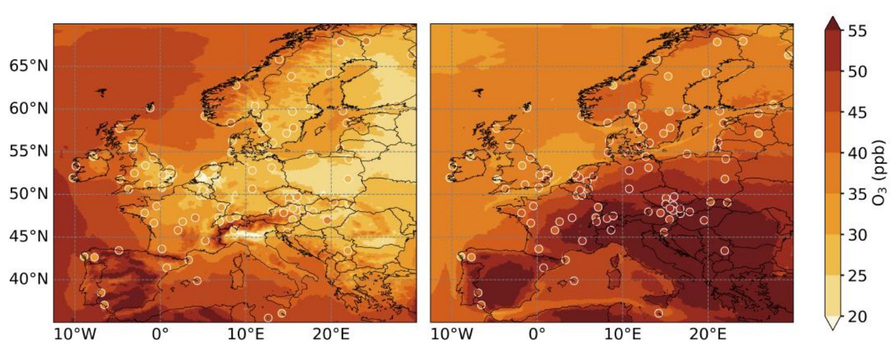

Figure 6. Mean afternoon ground-level O_3 mixing ratios from ICON-ART simulations for winter (JF, left) and summer (JJA, right) 2019. Filled circles indicate observations from EMEP monitoring stations. Elevated sites and stations with less than 75% valid data were excluded.

5. L432: “hats and overbars”?

- We added the following description:

The transport equations in ICON-ART are Hesselberg-averaged (indicated by a hat) meaning a variable $\boldsymbol{\Psi}$ *can be decomposed into a barycentric mean with respect to the air density* ρ_a *and its fluctuations:*

$$\widehat{\Psi} = \Psi - \Psi'' = \frac{\overline{\rho_a \Psi}}{\overline{\rho_a}}$$

The bar over a variable indicates Reynolds-averaging. The prognostic equations for number density ($\widehat{\Psi}_{0,l}$*) and mass mixing ratio (*$\widehat{\Psi}_{3,l}$*) are solved at every fast physics time step and are given by: …*

6. L491: “processes such as removal processes” to “processes such as removal Mechanisms”?

- Changed to mechanisms

7. L493: Describe the key aspects of this alternative method in a couple of sentences here, especially in relation to its computational efficiency.

- The following text is added to the paper:

In contrast to ISORROPIA-2, this approach bypasses full thermodynamic equilibrium calculations and instead uses an empirical hygroscopic growth formulation for sea salt. This greatly reduces computational cost while retaining the key impact of water uptake on particle mass and related aerosol processes.

8. L571: “implemented in other models”: name those models here along with the Citations.

- Done:

(HadCM (Woodward, 2001), LMDz-INCA (Balkanski et al., 2007), WRF-Chem (Zhao et al., 2013), MONARCH (Klose et al., 2021)).

9. L669 (or thereabouts): Also include a couple of sentences on the best practices of using this dusty cirrus parameterization for different (coarser, finer, or variable) grid resolutions. How does it perform across scales? Has this been tested? This could be discussed a bit.

- Following Seifert et al. (2023, ACP), we added a short discussion on the recommended use of the dusty-cirrus parameterization at different model resolutions. Seifert et al. show that the scheme is generally robust across a range of grid spacings, as long as the underlying dust mass and number concentrations are physically consistent. At coarser resolutions, the parameterization captures large-scale cirrus occurrence and radiative effects reasonably well, while finer or convection-resolving grids benefit from the improved representation of vertical motions and aerosol gradients that influence heterogeneous freezing. Although a systematic resolution-sensitivity analysis was not part of the study, available tests indicate consistent behavior of the parameterization across scales. We added the following text to the revised manuscript summarizing these points:

Dusty-cirrus parameterization performs robustly from mesoscale model resolutions (~10–20 km) down to convection-resolving scales (~1–3 km), as long as dust mass and number concentrations are physically reasonable. Coarser grids capture the large-scale cirrus response, while finer or convection-resolving scales benefit from enhanced representation of vertical motions and aerosol gradients.

10. Figure 11: The average OLR value should be shown in enlarged font or ideally printed over the map, or the reader might miss it. The technical name of the simulation experiment is not needed on the figure.

- This figure is revised accordingly (in the revised version as Fig 13)

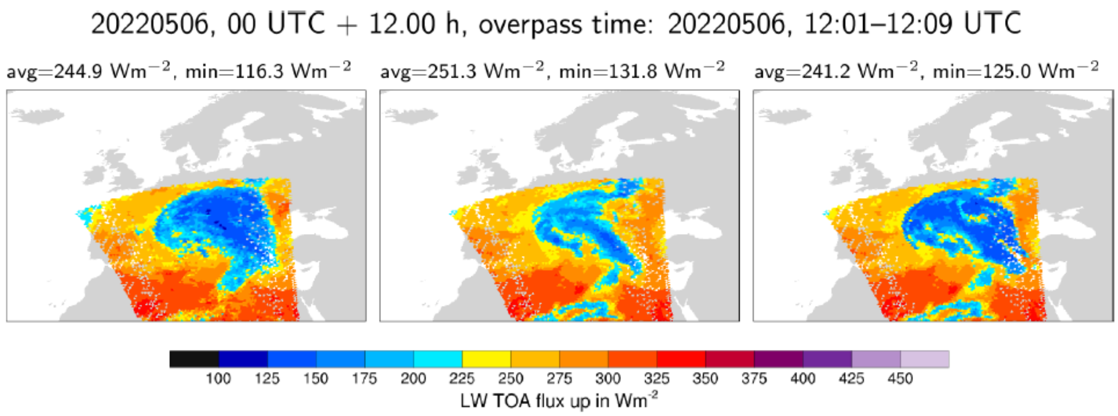

Figure 13. Comparison of global ICON-ART simulation for 12 UTC of 6 May 2022 with CERES Level 2 satellite data of outgoing longwave radiation at the top of atmosphere. Observations (left), ICON-ART without dusty cirrus parametrization (center), and ICON-ART with dusty cirrus parametrization (right).

11. Figure 12: This schematic could be improved: include additional boxes at the top showing input data (for both ICON and ART). Name some typical variables (winds, moisture, pressure; anthro emissions). Similarly, name some typical output variables from ICON and ART; add additional boxes if necessary. Use appropriate arrows along the lines to indicate the direction of control and sequence of execution and data flow. Aim to better depict the loops and subloops within the model time integration workflow. In the caption, consider changing "circles" to "loops".

- This figure is revised accordingly:

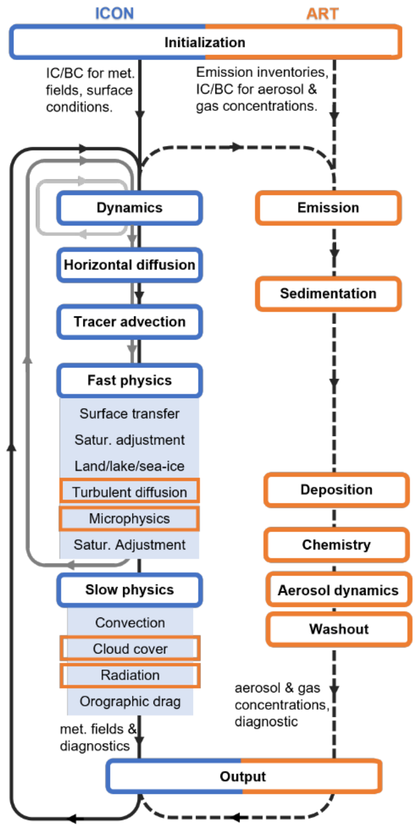

Figure 14. Schematic of the coupling of ICON–ART. The sequence in which processes of ICON are executed is illustrated by the blue boxes. Processes of ART are illustrated by the orange boxes. An orange frame around a blue box indicates, that the according code is part of the ICON tracer framework but ART tracers are treated inside this framework. The gray and black loops indicate the sequences of the time integration. Some examples of input and output fields are also show at the top and bottom of the loop.

12. L698: “since also diagnostic variables can be defined with it” to “since diagnostic variables can also be defined with it”.

- Done