We thank the anonymous reviewers and the editor for the time and attention they dedicated to our paper "Introducing aerosol-cloud interactions in the ECMWF model reveals new constraints on aerosol representation". We welcome the acknowledgment of relevance of the motivation of our research and are happy to address all the concerns raised by the reviewer individually.

Reviewer 1 - Comment 1

1. Model description clarity, effective radius, autoconversion

“The model/simulation description and configuration are not very clear. The study aims to "introduce aerosol-cloud interactions," but it does not mention: a) whether and/or how the computed Nd values will affect cloud droplet effective radius and autoconversion calculations; and b) whether the modified cloud properties/lifetime will provide feedback to radiation or meteorology. For the results discussed, it is often difficult to determine whether they are from offline calculations or online simulations.”

Reply: We acknowledge that the description of the model, especially concerning the IFS cloud scheme, was insufficient. In the updated version of the manuscript we will include a subsection "4.2: Cloud effective radius in the IFS" to better describe that N_d interacts with the IFS radiation scheme but not yet with the cloud physics, with reference to the IFS documentation for extra details on the cloud scheme. We will also make this point clearer in the introduction. In statements throughout the text to make it clearer to the reader whether the results are from offline calculations or online simulations.

4.2 Cloud effective radius in the IFS[1]

The currently-operational IFS cycle CY49R1 has a single-moment cloud microphysics scheme. The scheme’s prognostic variables are: cloud fraction, liquid (LWC) and ice cloud water content (IWC), rain (RWC) and snow water content. Except for cloud fraction, all other variables are represented as mass mixing ratios. Effective radius is a diagnostics produced by the radiation scheme to compute all-sky radiative fluxes and is defined as:

$$r_e = \left(\frac{4(\mathrm{RWC}+\mathrm{LWC})}{3\pi\rho_w\gamma N_d}\right)^{1/3} \quad (5)$$

where $\gamma = \left(\frac{r_v}{r_e}\right)^3$ defines the ratio between mean volumetric (r_v) and effective (r_e) radius, ρ_w is the density of water and N_d the cloud droplet number concentration.

Autoconversion rates of cloud water to rain follow a parametrization derived from (Khairoutdinov and Kogan, 2000) that assumes fixed N_d values over land and sea, and this behavior is not modified within this study. This means that only the first indirect radiative effect of aerosols can be explicitly represented by the model. Although cloud formation and dissipation are eventually impacted by the implied heating rates, our setup forbids any direct impact on lifetime.

2. Lacking evaluation of CAMS aerosol mass fields

“The optimization only considers the impact of aerosol size distribution (or r_{med} changes on N_d, assuming that the aerosol mass fields and thermodynamical fields for activation calculations and the sampling/averaging are "perfect." The authors did not provide any evaluation of these fields or relevant references. For example, do CAMS aerosol

forecasts have known substantial biases and uncertainties over regions with stratiform warm clouds? What is the uncertainty associated with sampling errors? Without considering these aspects, the optimization could compensate for systematic biases in these fields through unrealistic size distribution adjustments."

Reply: It is correct to say that the optimization assumes that the CAMS aerosol mass concentration fields "as perfect". The aerosol model is routinely evaluated by the CAMS community and its description, together with evaluation against a wide range of observations can be found in Rémy (2022), where a good performance of the global fields is assessed (and which we now cite); aerosols are evaluated also in CAMS reports several times a year and their performance (with and without assimilation) can also be assessed with tools like AEROVAL (https://aeroval.met.no/pages/evaluation/?project=cams2-82); moreover we included global evaluation of satellite AOD and AE for the default and optimized size distribution definitions in Figure 5. As discussed in the paper, we are aware that such fields are certainly not perfect and incorporate many unconstrained aspects, concerning speciation, total mass, physical and optical properties, sizes and shapes, and 3D spatial distribution. S are also totally aware that the optimization does tend to compensate for associated systematic errors. However, it is indeed within the scope of the paper to show how those biases whose patterns cannot be described by the few degrees of freedom of the optimization procedure (one global value of r_{med} per each species) are automatically highlighted in the resulting optimized setup. In particular, the biases addressed in this study, such as the Southern Ocean, might be undetected in the standard setups used for validation, due to the lack of reliable observations over open seas and, at the same time, they are insufficiently compensated by the tuning procedure. We will make this clearer in an updated version of the manuscript.
In the case in point, we concluded that these biases reflected issues related to the aerosol population found inside (and directly below) the cloud over the region, which led us to successfully improve the aerosol model in this sense. However, this work is not the last word on the matter - applying the optimization has highlighted areas in both size distribution assumptions and in the CAMS mass concentrations, and the intention is that future work builds on this paper to improve both mass distributions and size distributions simultaneously.

3. Critical model/retrieval sampling mismatch

Critical model/retrieval sampling mismatch. The comparison between model-derived and satellite-observed Nd may suffer from sampling inconsistencies that invalidate the optimization results

The reviewer structured their concerns into the following points (here reported in bold):

a) *different cloud detection methods: ERA5 model diagnostics vs. satellite radiance retrievals define "cloud top" differently.*
Reply: Cloud top properties from MODIS are inferred from brightness temperature measured in some infrared channels. The retrieved temperature is therefore representative of the top layers of the cloud and is typically within 1 cloud optical thickness (τ) in the IR (e.g. 0.72 reported in Wang, 2014). As a rule

of thumb, assuming extinction efficiency Q_{ext} ~ 2 in the visible range, for the IR holds Q_{ext} ~ ∈ [2,4], depending on the dominant droplet size. This says that using visible τ of the order of one is a fair assumption to optically define cloud top consistently with satellite retrievals. In addition to this, the target clouds of this study are mostly warm stratocumulus, that generally present an optically thick cloud top, so the height of cloud top would not differ very much if we were to change the optical-depth threshold from 1 to, say, 10. As a result, we do not expect the exact choice of the τ threshold for cloud top identification to play a significant role. We also incidentally tested this assumption by modulating the threshold on the penetration bias correction formula provided by Grosvenor, 2018 (as mentioned in the Results section) and did not find any significant change in results.

b) *spatial resolution mismatch: 3°×3° model data (aerosol concentrations and meteorological fields) are used to calculate Nd and compared with MODIS data - although MODIS data are also regridded to a 3x3 grid, they are aggregated and averaged from finer resolution, so they represent vastly different sampling volumes.*
Reply: We welcome this concern, but it is unfortunately not clear to us how the reviewer expects such spatial resolution issue to impact the results of this work. However, we concur that an optimization system where local N_d comparison is performed at higher details would be desirable and therefore we have repeated the analysis but at a 1x1 degree resolution; a selection of plots are shown at the end of this document and will be used in the revised version of the manuscript. In future, comparison could potentially be carried out an even higher resolution against level 2 instantaneous overpass data. This could be surely beneficial to verify the ability of the model to simulate sharp changes in aerosol fields, e.g. in the case of strong aerosol emission events like wildfires. The optimization presented in this work is, however, performed against monthly-mean satellite observations, which allow us to catch seasonal aerosol and cloud signals, but is not meant to assess the ability of the model to exactly match the instantaneous N_d values. From this standpoint, the optimization relies on a statistical comparison of simulated and observed N_d values, with an implicit low-pass filtering of frequency below 1 month. We added a remark for the next manuscript version to better clarify this point.

c) *temporal sampling bias: model data (4 times daily, every 5th day) vs. actual satellite overpass times.*
Reply: We are not aware of any specific biases that could be associated to different temporal sampling between model time and satellite overpasses. Potential biases due to diurnal cycles of convection should also be excluded by the cloud top temperature filtering both in observation and simulated data. However, to assess the potential impact of the adopted temporal sampling strategy we also tried excluding model data representing night-time values (that are excluded from MODIS observations used for Nd retrievals) from the optimization procedure. We tried two setups where we defined "dark" model grid-points to be masked out during the tuning procedure:

   1. At each time step we retain only grid points whose local solar zenith angle have values larger than 0.2.

2. At each time step we retain only points whose local solar time is between 9am and 4pm ("Daytime" setup).

In both cases, the impact on the effectiveness of tuning (as relative to the prior) was negligible and the optimal values of aerosol radii were not significantly different. As an example, results for the "Daytime" vs. default "All" setup are reported in the following Table 1, which will be added to the revised version of the paper.

CCN	Prior $r_{\rm med}$ [μm]	Optim. All $r_{\rm med}\{r_{\rm med}^{0.1}, r_{\rm med}^{0.9}\}$ [μm] (num. ratio)	Optim. Daytime $r_{\rm med}\{r_{\rm med}^{0.1}, r_{\rm med}^{0.9}\}$ [μm] (num. ratio)
sulfate	0.1100	0.086 $\{0.082, 0.093\}$ (2.06)	0.082 $\{0.078, 0.086\}$ (2.39)
nitrate	0.0355	0.043 $\{0.038, 0.080\}$ (0.56)	0.047 $\{0.039, 0.131\}$ (0.42)
seasalt1	0.1000	0.063 $\{0.061, 0.066\}$ (3.94)	0.065 $\{0.064, 0.070\}$ (3.66)
organics	0.0900	0.135 $\{0.128, 0.256\}$ (0.29)	0.134 $\{0.117, 0.464\}$ (0.30)
carbons	0.0118	0.007 $\{0.007, 0.015\}$ (4.63)	0.007 $\{0.006, 0.013\}$ (4.76)

Table 1: Optimization result summary for each CCN species. "Optim. All" hours is the default optimizsatione setup, "Optim. Daytime" is the setup where only pixels with local time between 9am and 4pm were used for the optimization. The table reports the Prior and optimized size-distribution median radius $r_{\rm med}$, together with the 0.1 and 0.9 quantiles $r_{\rm med}^{0.1}$ and $r_{\rm med}^{0.9}$ of the optimization procedure results and the implied number ratio $J_{\rm optimized}/J_{\rm prior}$ of the optimized definitions. For seasalt2, $r_{\rm med,seasalt2} = 10\ r_{\rm med,seasalt1}$, therefore the number ratio is identical to seasalt1.

d) *vertical sampling inconsistency: model-diagnosed cloud levels vs. satellite-retrieved cloud properties may sample completely different atmospheric layers. The current study lacks uncertainty estimates related to these issues.*
Reply: Representativity of the aerosol layer constitutes an important challenge to our method but must also be related with the low vertical resolution of used data on pressure levels as pointed out in section 4.3 (Model data selection), which is on average 1.5km throughout the troposphere. However, since we are concerned with aerosol impacts on liquid clouds, we are almost always dealing with boundary-layer clouds, and the vertical difference in sampling between model and satellite is small; in these cases, the aerosol is likely to be well mixed within the boundary layer. Moreover, deep cumuli are typically excluded by the procedure mostly thanks to the cloud-top physical requirements (cloud cover, temperature, phase). We added more details in section 4.3 (Model data selection) and 4.4 (Optimization procedure) for the next manuscript version to inform the reader about the implied vertical smoothing of the aerosol fields for the simulation of Nd.

4. Activation of aerosol population at cloud top - physical inconsistencies
"Physical inconsistency in process representation considered for optimization (Section 4.3). If I understand it correctly, the authors extract aerosol concentrations at "cloud top" and apply a 1 m/s updraft velocity to simulate activation at this level using the lookup table. I assume this approach considers that Nd retrievals are for "cloud top" only, but the method contradicts basic cloud microphysical principles. In reality, CCN activation occurs at cloud base during initial adiabatic ascent where supersaturation develops, and the resulting droplet population is then transported vertically through the

cloud. Nd satellite retrievals often assume adiabatic conditions, where Nd is considered constant throughout the cloud. Additionally, cloud tops typically experience subsiding air masses, entrainment of dry air, and near-zero or negative vertical velocities, so the 1 m/s (upward) vertical velocity assumption at cloud top is very likely unrealistic for most of the time. Aerosol populations at cloud top have been modified by scavenging, entrainment, and chemistry, making them fundamentally different from the original CCN population (consistent with results shown in figure 10). The authors seem to have realized this issue, as indicated by the comparison of InCloud and ClBase3 results in Table 3 and Figures 4 and 10."

Reply: We welcome the call of the reviewer for clarity and consistency on the aerosol activation microphysics. This is indeed a limitation of the presented setup, but it also represents the price we decided to pay for a flexible setup that allowed us to do this study using a relatively long time series of global observations. As reported in the paper, we experimented picking aerosol mixing ratios at different levels within the cloud, the extreme being picking aerosols below the cloud - while this has an impact on the amount of diagnosed sea salt CCN, overall results were quite similar, which might be partially attributable to the intrinsic smoothness of the aerosol fields across the vertical levels, as discussed above. Concerning the choice to pick aerosols at the cloud level to get diagnostic N_d estimates, this is the same approach adopted in the Jones, 2001 scheme of the Met Office climate model, where N_d is, like in our setup, a purely diagnostic quantity as described in West, 2014. The experiment in which we picked aerosols below cloud base was intended to address the specific issue of missing aerosols inside high-latitude clouds, yet besides the modest (but insufficient) impact at high latitudes, we got limited improvement elsewhere and no change in qualitative terms, with a persistence also of the strong bias over Africa as shown in the paper.

Regarding the choice of a fixed updraft velocity of 1 m/s, we agree with the reviewer's concern and have substantially improved the estimate of updraft velocity, as described in our response to Reviewer 2 (comment 3) – see below.

The title of the paper
"If the goal is really to show the introduction of ACI in IFS helps to constrain the aerosol representation, a more comprehensive evaluation of the aerosol properties is needed (e.g., evaluation of aerosol size distribution using in-situ data). In my opinion (and as the authors discussed in the introduction), the value of this work is more on providing a simplified but practical treatment of Nd prediction and ACI representation in the IFS model, which will allow IFS (with CAMS) to consider the impact of aerosols on clouds in the future."

Reply: We agree that it is worth extending the work by bringing in more observations, especially from in-situ measurement sites or specific campaigns. This paper is however quite long already, and we consider investigation in this sense (combining aerosol indirect effects and in-situ aerosol measurements) certainly worth being done and eventually published in the future. This is indeed pointed out in the abstract and in the conclusions, where we provide outlooks and suggestions in this direction stemming from the results of this work. We also believe that the title accurately describes the scope of the paper, which is to introduce, as the reviewer correctly puts it, a simplified but practical treatment of Nd to illustrate that a new set of observations (in this paper N_d and all-sky SW fluxes), can be now brought in to constrain the CAMS aerosol system. Moreover, we present two cases as a proof of concept to illustrate how our approach can reveal regional deficiencies in the aerosol distribution that should be investigated in future.

Punctual comments

We directly adopted most punctual remarks and corrections (well spotted, thanks!) made by the reviewer. Those that required a more articulated response are reported below in bold and followed by our reply.

1) *Page 1, Line 10: "We found that CAMS aerosols allow simulating overall realistic Nd values". Is this conclusion for unoptimized or optimized ASD?*
   Reply: Added clarification in abstract
2) *Page 4, Line 93: The direct aerosol effect should also affect the meteorological fields, not only semi-direct effect.*
   Reply: Rephrased.
3) *Page 4, Line 105: Will hydrophobic aerosols be removed by precipitation?*
   Reply: No, hydrophobic aerosols are not removed by precipitation. Rephrased in text.
4) *Page 4, Line 115: Does hydrophilic BC have hygroscopic growth and is it considered in activation? Or "hydrophilic" BC only applies to wet removal calculation?*
   Reply: Added sentence "All and only hydrophilic species are removed by precipitation" in the texts.
5) *Page 6, Line 144: It seems to me the change in K_{ext} is quite large for certain spices at certain RHs. It would be useful to calculate the difference using the default and optimized r_{med} values.*
   Reply: A possible illustration of this effect is depicted in Figure 5, showing that quite a significant impact is attributed to zonal mean additional 0.01 AOD from Sea Salt, peaking over the Southern Ocean 0.02. Contribution from OM is instead negative and peaks in the tropics over South-East Asia, India, Central Africa.
6) *Page 8, Line 165-170: Please provide the references of Q06, G18, and BR17.*
   Reply: Done.
7) *Page 10, Line 216-217: How large is the uncertainty associated with these assumptions?*
   Reply: We added a discussion of the LUT limitations with references to provide context for the chosen values: "… This means that there is no dependency of the LUT on local atmospheric variables other than aerosol number concentrations. Slightly sub-saturated conditions are generally descriptive of the environment immediately below cloud-base and are in line with previous literature (e.g Leaitch, 1986). The chosen temperature value is up to 5 K higher than the sub-cloud values found in the model and up to 15 K lower than the tropical values (not shown). Similarly, the chosen pressure tends to underestimate by up to 100

hPa high-latitude values and overestimate by up to 200 hPa tropical values (not shown). However, the dependency of the output of the adiabatic parcel activation process on cloud-base pressure has been found to be generally negligible, especially for temperature regimes above 0 K and low CCN concentrations (<200 cm^{-3}) (Leaitch, 1986; Rothenberg and Wang, 2018).

8) *Page 10, Line 224: The assumed updraft/vertical velocity is pretty large, and is associated with large uncertainties.*
   Reply: This point is similar to point 3 of reviewer 2, and is addressed below.
9) *Page 11, Line 238: T255 should be at $~80km$ resolution, instead of $38km$? Please double check.*
   Reply: Well spotted thanks!
10) *Page 11, Line 250: Does the MODIS retrieval apply a similar conditional sampling?*
   Reply: Partly – the ice water ratio IWR = IWC / (IWC+LWC) criterion relaxed due to the very coarse (3x3 degrees) horizontal resolution, where we allowed IWR up to 0.5. In an updated version of this paper, where we upgrade the offline optimization to 1x1 degrees grid, we filter only clouds where the IWR at the top is lower than 0.05.
11) *Page 12, section 4.3: I assume this approach considers that Nd retrievals are for "cloud top" only, but the method contradicts basic cloud microphysical principles. In reality, CCN activation occurs at cloud base during initial adiabatic ascent where supersaturation develops, and the resulting droplet population is then transported vertically through the cloud. Nd satellite retrievals often assume adiabatic conditions, where Nd is considered constant throughout the cloud.*
   Reply: This was addressed in main point above about physical inconsistencies of the activation of aerosol population at cloud top.
12) *Page 12, section 4.3, formula 7: Why only $N_{d,Q06}$ is considered in the numerator?*
   Reply: We added a remark in the Optimization Procedure section clarifying that our choice is because, being Q06 the least restrictive sampling strategy, it contains by design the largest number of valid retrievals.
13) *Page 12, Line 285: a brief description of the "Nelder-Mead" algothrim is necessary. How to simultaneously optimize different r_{med} values for individual aerosol species?*
   Reply: We added a brief paragraph with description and a reference Lagarias, 1998 for the reader

   As an optimization algorithm for the loss function $\mathcal{L}$ we use Nelder-Mead with free bounds. This is a popular minimization method belonging to the class of direct search methods, i.e. that do not require knowledge of the gradients of the function, for functions of the kind $f(\mathbb{R}^n) \to \mathbb{R}$. Nelder-Mead makes use of n-dimensional convex hulls and a series of reflection, expansion and contraction operations to converge on function minima (see e.g. Lagarias et al., 1998). Over the many setups used for the optimization procedure, we never encountered important convergence issues.

14) *Page 12, section 4.3: please also discuss how the temporal co-location and averaging are applied.*
   Reply: We added the following to the paragraph describing the optimization loop flow-chart: "It is worth to stress here that, while aerosol and cloud data are exactly co-located, the comparison with observation is done on monthly mean data, which implicitly makes the loss function L a statistical evaluation of the

simulated N_d. On the one hand, the impact of CCN on results to timescales of variability of one month or larger; on the other, this allows the temporal sampling of model data (four times a day every fifth day) to represent monthly means, with massive reduction of the amount of data to be processed by the optimization procedure.

15) *Page 13, Table 3: Please discuss values in the 4th column (ClBase3).*
Reply: We added this paragraph in the Results section: "Table 3 and Figure 4 also show the effect of diagnosing aerosol fields three model levels below the cloud base (ClBase3). We investigated this setup because of its beneficial impact at high latitudes, which is determined by the sea salt CCN signal and is discussed more in detail in Subsection 5.2. The only other species that is significantly affected by the ClBase3 setup is Sulfate aerosol, for which the optimal PSD implies number concentrations lower by a factor 2.3 compared with the default setup, where aerosols are diagnosed at cloud level (InCloud). In the ClBase3 setup, diagnosing sulfate mass concentrations closer to the surface implies a disproportionately stronger signal in regions with strong anthropogenic emissions such as South-East China, India, Europe and North-Eastern America (compare prior N_d in Figure 4), so that reducing the overall number of sulfate aerosol brings localized benefit. Over remote ocean regions instead, the reduction in N_d is compensated by an increased role of sea salt CCNs. Relevant improvement is also obtained over the southern tropical Atlantic, with a reduction of the large positive N_d bias in the optimized setup. We show in Subsection 5.1 that the positive bias structures in the tropics over Africa and the Atlantic are associated with carbonaceous (OM and BC) aerosols emitted from wildfires. By comparing the OM concentrations in the ClBase3 and InCloud setup (not shown) we assessed that diagnosing aerosol below cloud base implies lower values corresponding to the smoke plumes advected aloft over the Atlantic. This can also be seen by comparing panels (d) and (a) in Figure 4: ClBase3 has stronger positive bias over continental Africa, where carbonaceous aerosols are emitted, but reduced bias over the Atlantic, where the same aerosols are just advected.

16) *Page 19, Figure 8: What is $N_{d,modis}$? Which of Q06, G18, and BR17?*
Reply: Thank you, that was unclear. It is the optimization target Q06, and an updated version of the figure will include this information!

17) *Page 20, Figure 9: How is the IFS "ctrl" simulation configured? Would be useful to compare the simulations with original and modified r_{med} values.*
Reply: We extended the discussion of this figure in the text: "Using the PSD definitions coming from tuning produces a modest shift in the size spectra of the finest peak toward larger values (not shown), which is directly associated to OM aerosol. This results in a degradation with respect to the AERONET spectra, increasing the bias toward lower simulated total aerosol numbers. This adds to the evidence that the optimization procedure is trying to compensate for biases others than PSD assumptions. We need therefore to consider alternative hypotheses to approach the N_d bias, such as errors in the speciation between OM and BC of the emitted carbon in aerosols and in the simulation of ageing processes."

Reviewer 2 - Comment 1

1. *The implementation of aerosol–cloud interactions in weather and climate models has been a longstanding topic of research. Numerous global climate models, seasonal forecasting systems, and modern reanalyses already include such representations (e.g., Seinfeld et al., 2016; Benedetti and Vitart, 2018; Wang et al., 2021; Song et al., 2025). Additionally, aerosol activation parameterizations have undergone extensive development over the past decades, resulting in computationally efficient schemes (e.g., Ghan et al., 2011). Retrievals of CCN concentrations from aerosol reanalyses such as CAMS have also been demonstrated (Block et al., 2024). A more thorough discussion of this existing body of work would have helped guide the authors' methodology and contextualize their contribution.*
   Reply: We welcome the call of the reviewer for a more thorough discussion of the existing literature. While in our manuscript's introduction we already cite five papers (among many available) about the representation of aerosol-cloud interactions in several system, we will expand this by referring to additional literature examining aerosol activation in global circulation models (e.g. Barahona, 2014 and Rothenberg and Wang, 2018). It is however not clear to us how the references reported by the reviewer would imply an insufficient discussion of previous works; one of these (Song et al., 2025) was published well after the submission of this manuscript and another one (Benedetti and Vitart, 2018) reports the inclusion of aerosols in a seasonal forecasting system to represent the direct radiative effect, while explicitly excluding any representation of aerosol-cloud interactions from their work.
   It is also true that the work by Block et al, 2024 provides implied CCN concentrations from CAMS reanalyses, and we will happily reference this in the introduction. However, these are offline calculations designed to be compared with in-situ observations during campaigns, i.e. without assessing which supersaturation values are reached during the activation process, which poses a clear limitation for the modelling of N_d starting from those. Predicting correct supersaturation values is in fact the task of the activation scheme, and it is also the reason why (as the reviewer points out) vertical velocity of activation play a role in the determination of the final number of cloud droplets.

References provided by Reviewer 2 for this question:

- Benedetti, A., & Vitart, F. (2018). Can the direct effect of aerosols improve subseasonal predictability? *Monthly Weather Review*, 146, 3481–3498. https://doi.org/10.1175/MWR-D-17-0282.1
- Block, K., Haghighatnasab, M., Partridge, D. G., Stier, P., & Quaas, J. (2024). Cloud condensation nuclei concentrations derived from the CAMS reanalysis. *Earth System Science Data*, 16, 443–470. https://doi.org/10.5194/essd-16-443-2024
- Ghan, S. J., Abdul-Razzak, H., Nenes, A., Ming, Y., Liu, X., Ovchinnikov, M., Shipway, B., Meskhidze, N., Xu, J., & Shi, X. (2011). Droplet nucleation: Physically-based parameterizations and comparative evaluation. *Journal of Advances in Modeling Earth Systems*, 3(4). https://doi.org/10.1029/2011MS000074
- Seinfeld, J. H., Bretherton, C., Carslaw, K. S., et al. (2016). Improving our fundamental understanding of the role of aerosol–cloud interactions in the climate system. *Proceedings of the National Academy of Sciences*, 113(21), 5781–5790. https://doi.org/10.1073/pnas.1514043113

- Song, C., McCoy, D., Molod, A., Aerenson, T., & Barahona, D. (2025). Signatures of aerosol–cloud interactions in GiOcean: A coupled global reanalysis with two-moment cloud microphysics. *Atmospheric Chemistry and Physics*, 25, 15567–15592. https://doi.org/10.5194/acp-25-15567-2025
- Wang, C., Soden, B. J., Yang, W., & Vecchi, G. A. (2021). Compensation between cloud feedback and aerosol–cloud interaction in CMIP6 models. *Geophysical Research Letters*, 48(4), e2020GL091024. https://doi.org/10.1029/2020GL091024

2. *The manuscript does not clearly define what is meant by the implementation of ACIs. Most of the discussion focuses on adjustments to the aerosol size distribution and the resulting changes in aerosol fields, with minimal reference to cloud properties. It remains unclear whether the computed Nd is actively used to influence or update any cloud-related processes in the model, raising doubts about the actual implementation of ACIs.*
   Reply: We acknowledge that the description of the model was insufficient. We will include a subsection "4.2: Cloud effective radius in the IFS" (see reply to Reviewer 1 Comment 1) to better describe how N_d is used by the IFS radiation scheme and how effects on meteorology and clouds are enabled when the model is run. Within the single-moment cloud scheme of the IFS, N_d is used to simulate the first indirect effect, but no impact on autoconversion rates is allowed, yet. This cautious approach is needed to progressively calibrate the perspective new components of an operational forecasting system.

3. *The assumption of a constant updraft velocity of 1 m/s lacks physical justification and is not supported by observational or theoretical evidence. It is well established that Nd is highly sensitive to updraft velocity, which has profound implications for the aerosol indirect effect (Sullivan et al., 2016). While it is true that updraft is among the most uncertain parameters in ACI modeling, applying a fixed and globally unrealistic value, appropriate only for small marine cumulus clouds, is not defensible.*
   Reply: We agree and have substantially improved the treatment of vertical velocities for activation, as will be described in the updated version of the manuscript. Firstly, the resolution of the tuning procedure has been brought from 3x3 degrees to 1x1 degrees for a sharper identification of cloud regimes (and finer than many climate models). We now include the large-scale vertical wind as mean vertical velocity (as in Van Noije, 2021) and the Deardorff convective scale velocity w* (as defined in Deardorff, 1970). Many models determine vertical velocity scales w' from turbulent kinetic energy (TKE), which is however not viable for us during the offline tuning procedure due to the limited number of diagnostics available in the ERA5 datasets, and no TKE. We are aware that using w* as a diagnostic for TKE would miss non-convectively induced turbulence, especially related to cloud top radiative cooling. However, we need to point out that a satisfactory characterization of vertical updraughts for aerosol activation is still a challenge for global models and these typically need to resort to setting minimum boundaries for the vertical velocity standard deviation σ_w (e.g. 0.1 m/s in Barahona, 2014 and 0.7 m/s in Golaz, 2011), that are reported by Golaz, 2011 to occur more than 90% of the time, and reliable parametrizations for updraft velocities in marine boundary layer clouds are an open area of

research (see e.g. Ahola, 2022 and references therein). Figure 1 of this document (below) shows mean N_d resulting from the optimization setup for the year 2009 in a setup where vertical speed for activation w is a normal distribution N(w_{ls}, 0.4w*), with w_{ls} the resolved large-scale vertical wind of the model, compared with a setup with w=0.5 m/s. This required recomputing the look-up table with two additional dimensions for w_{ls} and w*. Figure 2 of this document shows the difference between the two setups, which is the largest where the highest aerosol loads are, which is consistent with the idea that these are "updraft-limited" regimes for aerosol activation. We ran these optimizations at 1x1 degree resolution, which allowed us to match more strictly the filtering criteria between model data and observations. Figure 3 of this document shows the diagnosed values of the mean and standard deviation of vertical velocity for the same year of simulation. With these changes, we argue that our activation scheme now has a similar level of sophistication to many in the literature.

4. *The treatment of satellite retrieval uncertainties is problematic. The authors appear to treat differences arising from alternate retrieval assumptions as equivalent to experimental error, which is not technically sound. Moreover, model outputs must be sampled in a manner consistent with the assumptions of the satellite retrievals to avoid artificial biases. For instance, retrieval filters based on temperature and cloud fraction tend to exclude clouds with low Nd, particularly in high-latitude regions. If such filters are not consistently applied to the model data, it can lead to systematic biases, potentially causing the authors to erroneously tune scavenging parameters in response to what is essentially a sampling artefact.*

   Reply: Satellite retrievals of Nd from spectrometers are notably affected by substantial uncertainties, largely depending on the cloud-adiabaticity assumption and the accuracy of LWP / r_{eff} (Gryspeerdt, 2022). It is incorrect to say that the datasets produced in the study by Gryspeerdt, 2022 used different retrieval assumptions, because they instead used different sampling strategies aimed at reducing uncertainties "through sampling retrievals with higher confidence" (Gryspeerdt, 2022). While agreeing that this does not imply a precise quantification of the uncertainty of these retrievals, our methodology relies on implicitly assuming that the spread across these datasets is somewhat proportional to the uncertainty of the retrieval. We concede that this assumption constitutes a weakness of our approach, but this is in our opinion nonetheless a step further with respect to previous relevant literature using MODIS Nd data to constrain model representation (see e.g. McCoy, 2018). About the concerns on the sampling of model data, we indeed applied filtering on cloud-top temperature, phase, and total optical thickness as done with the satellite retrievals, to avoid introduction of sampling biases for the comparison as reported in section 4.2 (Model data selection). However, due to much lower resolution of the model data, it was necessary to weaken the filtering criteria when applying them to gridbox-mean quantities, which is a necessary compromise to get enough valid pixels for the comparison. In an updated version of the manuscript, we will increase the resolution of model data from 3x3 to 1x1 degrees, which brings the model data selection criteria closer to the observations and helps improving the reliability of our approach. Finally, concerning the reviewer's remark that such sampling issues eventually could lead to systematic biases to be corrected by changing the wet scavenging parametrization, we are afraid there has been a misunderstanding of the nature of our work. We haven't performed at any point a tuning of the scavenging parameters to improve the comparison between simulated and satellite N_d. The need for an updated wet scavenging parametrization was motivated by the diagnosis shown in figure 11 of the manuscript, where the model cross-section showed virtually no aerosols inside the clouds over the Southern ocean. As described in the manuscript, the algorithm improvements we made to wet scavenging was not based on tuning, but rather on known processes (e.g. Wegener-Bergeron-Findeisen, riming, condensate phase aerosol mass activation rates). Its positive impact on simulated N_d, SW fluxes and satellite AOD over the high-latitudes summer hemisphere could be assessed *a posteriori* as shown in figures 11, 14 and 15. From the reviewer's remark we realize that we should more strongly emphasize this last point, in order to avoid the misunderstanding that there was any tuning of the wet scavenging scheme to compensate from sampling biases in the observations.

5. *The global optimization approach based on satellite retrievals fails to account for their limited validity. Retrieval techniques are typically applicable to vertically homogeneous, adiabatic, low-level clouds, which are the exception rather than the norm. Despite this, the authors extend the use of the retrieval to mixed-phase clouds and regimes strongly influenced by convection, where the assumptions underlying the retrieval no longer hold.*
   Reply: The selection of model data described in section 4.2 is aimed at filtering out clouds in convective-dominated regimes. Thanks to the reviewer's remark, we realized that we inadvertently omitted to specify that we adopted a cloud cover criterion (at least 0.8 on 3x3 degrees grid), in addition to the described criteria on cloud top temperature (>283K) and the presence of cloud-top ice water content (must be less than 50% of total water content on 3x3 degree grid). The combination of these makes it unlikely for clouds strongly dominated by convection to be selected for the optimization procedure. As correctly pointed out by the reviewer, this is a critical aspect of the study, therefore in an updated version of the paper we will use 1x1 degrees data and tighten the ice water content criterion (<0.05 of total water) to make the selection of cloud model data more consistent with the satellite retrievals. This ensures that the analysis of satellite data is dominated by the clouds whose aerosol impacts we are most interested in, namely boundary-layer liquid clouds.
6. *The manuscript does not include a data availability statement. The code availability statement does not refer to the parcel model, nor the lookup table used in this work. Given the nature of the work and its reliance on numerical model development and evaluation, it is essential that both the data and the implementation code be made publicly accessible to support reproducibility and validation.*
   Reply: The parcel model is referenced by citing Rothenberg and Wang (2016) as indicated in the model's documentation (https://pyrcel.readthedocs.io/en/latest/). We will be happy to provide a DOI

pointing to the lookup table as well as the offline optimization code in an updated version of the manuscript.

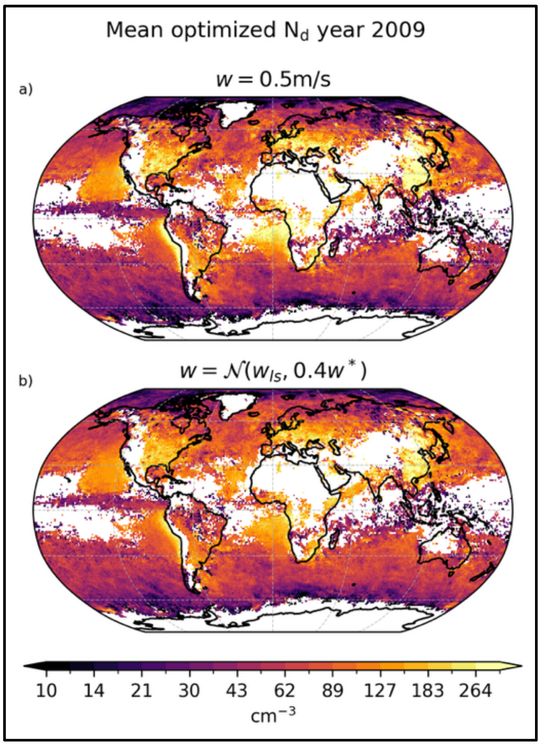

Figure 1: Optimized N_d averages for the year 2009 using 1x1 degrees model data. Top is using fixed vertical speed w = 0.5 m/s, bottom uses the new approach: a normal distribution of vertical speeds with mean equal to large scale vertical velocity and standard deviation w multiplied by 0.4.*

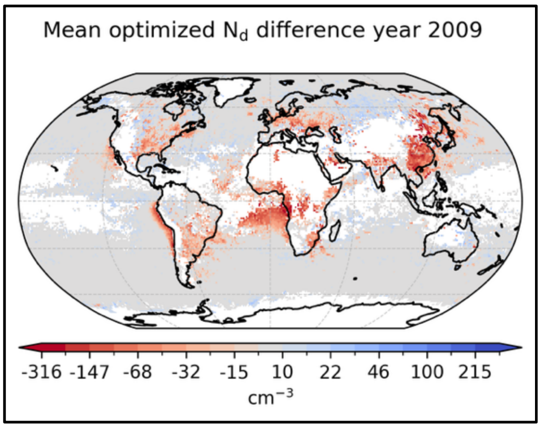

Figure 2: Setup using normal distribution of vertical speeds minus setup using fixed vertical speed w = 0.5 m/s, optimized N_d difference averaged for the year 2009.

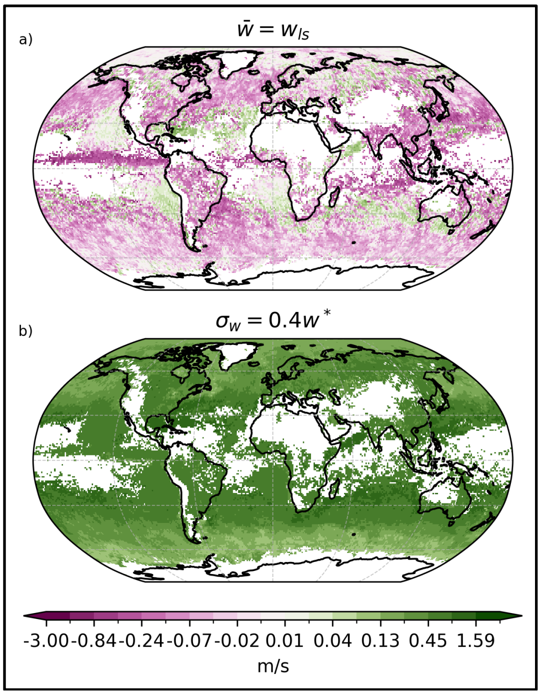

Figure 3: Mean (top) and standard deviation (bottom) of the vertical velocity distribution for the setup using a normal distribution of vertical activation velocities. Values are averages for the year 2009.

References

Ahola, Jaakko et al. (2022). "Technical Note: Parameterising Cloud Base Updraft Velocity of Marine Stratocumuli". In: *Atmospheric Chemistry and Physics* 22.7, pp. 4523–4537. doi: 10.5194/acp-22-4523-2022.

Barahona, D. et al. (2014). "Development of Two-Moment Cloud Microphysics for Liquid and Ice within the NASA Goddard Earth Observing System Model (GEOS-5)". In*: Geoscientific Model Development* 7.4, pp. 1733–1766. doi: 10.5194/gmd-7-1733-2014.

Deardorff, James W. (1970). "Convective Velocity and Temperature Scales for the Unstable Planetary Boundary Layer and for Rayleigh Convection". In: *Journal of the Atmospheric Sciences* 27.8, pp. 1211–1213. doi: 10.1175/1520-469(1970)027<1211:CVATSF>2.0.CO;2.

Golaz, Jean-Christophe et al. (2011). "Sensitivity of the Aerosol Indirect Effect to Subgrid Variability in the Cloud Parameterization of the GFDL Atmosphere General Circulation Model AM3". In: *Journal of Climate* 24.13, pp. 3145–3160. doi: 10.1175/2010JCLI3945.1.

Grosvenor, Daniel P., Odran Sourdeval, and Robert Wood (2018). "Parameterizing Cloud Top Effective Radii from Satellite Retrieved Values, Accounting for Vertical Photon Transport: Quantification and Correction of the Resulting Bias in Droplet Concentration and Liquid Water Path Retrievals". In: *Atmospheric Measurement Techniques* 11.7, pp. 4273–4289. doi: 10.5194/amt-11-4273-2018.

Gryspeerdt, Edward et al. (2022). "The Impact of Sampling Strategy on the Cloud Droplet Number Concentration Estimated from Satellite Data". In: *Atmospheric Measurement Techniques* 15.12, pp. 3875–3892. doi: 10.5194/amt-15-3875-2022.

Jones, Andy et al. (2001). “Indirect Sulphate Aerosol Forcing in a Climate Model with an Interactive Sulphur Cycle”. In: *Journal of Geophysical Research: Atmospheres* 106.D17, pp. 20293–20310. doi: 10.1029/2000JD000089.

Lagarias, Jeffrey C. et al. (1998). “Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions”. In: *SIAM Journal on Optimization* 9.1, pp. 112–147. doi: 10.1137/S1052623496303470.

Leaitch, W. R. et al. (1986). “Cloud Droplet Nucleation and Cloud Scavenging of Aerosol Sulphate in Polluted Atmospheres”. In: *Tellus B* 38B.5, pp. 328–344. doi: 10.1111/j.1600-0889.1986.tb00258.x.

McCoy, Daniel T. et al. (2018). “Predicting Decadal Trends in Cloud Droplet Number Concentration Using Reanalysis and Satellite Data”. In: *Atmospheric Chemistry and Physics* 18.3, pp. 2035–2047. doi: 10 . 5194 / acp -18-2035-2018.

Rémy, Samuel et al. (2022). “Description and Evaluation of the Tropospheric Aerosol Scheme in the Integrated Forecasting System (IFS-AER, Cycle 47R1) of ECMWF”. In: *Geoscientific Model Development* 15.12, pp. 4881–4912. doi: 10.5194/gmd-15-4881-2022.

Rothenberg, Daniel and Chien Wang (2016). “Metamodeling of Droplet Activation for Global Climate Models”. In: *Journal of the Atmospheric Sciences* 73.3, pp. 1255–1272. doi: 10.1175/JAS-D-15-0223.1.

Van Noije, Twan et al. (2021). “EC-Earth3-AerChem: A Global Climate Model with Interactive Aerosols and Atmospheric Chemistry Participating in CMIP6”. In: *Geoscientific Model Development* 14.9, pp. 5637–5668. doi: 10.5194/
gmd-14-5637-2021.

Wang, Chunpeng et al. (2014). “A Physically Based Algorithm for Non-Blackbody Correction of Cloud-Top Temperature and Application to Convection Study”. In: *Journal of Applied Meteorology and Climatology* 53.7, pp. 1844–1857. doi: 10.1175/JAMC-D-13-0331.1.

West, R. E. L. et al. (2014). “The Importance of Vertical Velocity Variability for Estimates of the Indirect Aerosol Effects”. In: *Atmospheric Chemistry and Physics* 14.12, pp. 6369–6393. doi: 10.5194/acp-14-6369-2014