Assessing the potential for simplification in global climate model cloud microphysics

Ulrike Proske[1], Sylvaine Ferrachat[1], David Neubauer[1], Martin Staab[2], and Ulrike Lohmann[1]

[1]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
[2]Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Hannover, Germany

Correspondence: Ulrike Proske (ulrike.proske@env.ethz.ch)

Abstract. Cloud properties and their evolution influence Earth's radiative balance. The cloud microphysical (CMP) processes that shape these properties are therefore important to be represented in global climate models. Historically, parameterizations in these models have grown more detailed and complex. However, a simpler formulation of CMP processes may leave the model results mostly unchanged while enabling an easier interpretation of model results and helping to increase process under-
5 standing. This study employs sensitivity analysis on an emulated perturbed parameter ensemble of the global aerosol-climate model ECHAM-HAM to illuminate the impact of selected CMP cloud ice processes on model output. The response to the perturbation of a process serves as a proxy for the effect of a simplification. Autoconversion of ice crystals is found to be the dominant CMP process in influencing key variables such as the ice water path or cloud radiative effects, while riming of cloud droplets on snow has most influence on the liquid phase. Accretion of ice and snow and self-collection of ice crystals have
10 a negligible influence on model output and are therefore identified as suitable candidates for future simplifications. In turn, the dominating role of autoconversion suggests that this process has the greatest need to be represented correctly. A seasonal and spatially resolved analysis employing a spherical harmonics expansion of the data corroborates the results. This study introduces a new application for the combination of statistical emulation and sensitivity analysis to evaluate the sensitivity of a complex numerical model to a specific parameterized process. It paves the way for simplifications of CMP processes leading
15 to more interpretable climate model results.

1 Introduction

Aerosols and cloud microphysics (CMPs) control cloud properties and thereby exert a large influence on Earth's climate. For example, the cloud water and ice contents determine the cloud albedo and lifetime, and also control precipitation formation
20 (Mülmenstädt et al., 2015). In a changing climate, the correct representation of clouds is especially important to estimate Earth's radiation budget (Sun and Shine, 1995; Tan et al., 2016; Matus and L'Ecuyer, 2017; Lohmann and Neubauer, 2018).

However, clouds and cloud feedbacks are a major source of uncertainty for projections of climate sensitivity in global climate models (Cess et al., 1990; Soden and Held, 2006; Williams and Tselioudis, 2007; Boucher et al., 2013).

Since cloud microphysical processes such as the riming of cloud droplets on snow flakes occur on scales much smaller than the resolution of global climate models (GCMs), they are parameterised, i.e. only their macroscopic effects at the scale of the model grid are described. Responding to the challenge of incorporating these processes in climate models, the community has added more and more processes into GCMs (Knutti and Sedláček, 2013) with increasing detail in their representation (e.g. Archer-Nicholls et al. (2021); Morrison et al. (2020)). As Fisher and Koven (2020) argue for the similar situation in land surface modelling, this may be due on the one hand to scientists' tendency to focus on their own area of expertise. On the other hand, it also reflects the fact that the Earth system is indeed complex and that many processes may matter. However, it is doubtful whether more detail will help us to reduce uncertainty (Knutti and Sedláček, 2013; Carslaw et al., 2018). More complexity also has its downsides: More parameterised processes lead to more parametric uncertainty which in turn scientists investigate and try to reduce with large scientific effort (e.g. Rougier et al. (2009); Lee et al. (2011); Yan et al. (2015); Williamson et al. (2015); Dagon et al. (2020)). In fact, Reddington et al. (2017) argue that "aerosol-climate models are close to becoming an overdetermined system with many interacting sources of uncertainty but a limited range of observations to constrain them", referring to the complexity in the representation of aerosols and their interaction with clouds. This is related to equifinality, meaning that model versions from different regions of the input parameter space may lead to the same results that compare well with observations. These models may simulate a range of aerosol forcings (Lee et al., 2016), which is not possible to constrain with current observations. Morrison et al. (2020) diagnose the same problem for CMP schemes, whose complexity, they say is ""running ahead" of current cloud physics knowledge and the ability to constrain schemes observationally." Climate models have become so complex that they are impossible to comprehend by any one scientist (Fisher and Koven, 2020). More detail means more heterogeneity between climate models, which increases the difficulty of a meaningful comparison of their projections (Fisher and Koven, 2020). But also within a given model, the attention and detail given to some cloud microphysical processes comes at the expense of other less accessible processes. This brings the danger of overinterpreting those processes that are represented in detail while neglecting the impacts of poorly represented ones (Mülmenstädt and Feingold, 2018). Finally, the detail of the aerosol and cloud microphysics increases computational demand and thereby costs (though anticipating the results of Sec. 3.6, the four CMP processes investigated in this study require negligible computing time). It can thereby inhibit other advancements such as the move towards high-resolution simulations (which may themselves also require adaptations of the CMP schemes) or larger ensembles.

In contrast, simple models are easier to interpret and derive understanding from, as long as they represent processes correctly (Koren and Feingold, 2011; Mülmenstädt and Feingold, 2018). Also, assumptions and their consequences are more traceable in simpler or more system-oriented models (Mülmenstädt and Feingold, 2018). For example, conceptual cloud models have been used to investigate the impact of the choice of precipitation particles' attributes on the cloud structure and evolution (Wacker, 1995); or to confirm microphysics findings qualitatively (Wood et al., 2009). Simplifications reduce computational demand, and simplified models yield themselves to other applications, e.g. the use in integrated assessment models (Ghan et al., 2013). At the same time, they may produce similarly good results as more complex models. For example, Ghan et al. (2012) developed

a simple yet physical model for the aerosol indirect effect, whose estimates are comparable to that of complex global aerosol models. Similarly, Liu et al. (2012) compared two aerosol modules with seven and three lognormal modes and find that the simulated aerosol concentrations are remarkably similar.

60   The addition of detail and refinement of a model description is a natural response to the challenge of capturing something as complex as the climate system in a computer model. This is legitimate and beneficial. For example, it may lead to a physically more correct representation and reduce the amount of tuning parameters (e.g. Storelvmo et al. (2008)). And for some applications modelers may need as much detail as possible in one specific module. Hence, scientists tend to call for more detail in process representations (e.g. Gettelman et al. (2013) for warm-rain microphysics; Sotiropoulou et al. (2021) for secondary

65   ice production by break-up from collisions between ice crystals) instead of less. This may in part be because humans are biased towards searching for additive pathways as problem solutions while overlooking subtractive transformations (Adams, 2021). However, due to the reasons mentioned above, a simplified model equifinal to a more complex model may be more useful for gaining understanding of climate models (equifinal meaning that the two model versions lead to similar results). One can therefore question the need for an ever increasing amount of detail, especially in the face of overdetermination (Reddington

70   et al., 2017). In this paper, we propose a new methodology to assess where process parameterisations can be stripped of detail to aid the development of a simplified model as well as to increase understanding of the model.

     The role of CMPs within GCMs has been investigated previously: The influence of CMPs has been shown to dominate over that of aerosol schemes in affecting clouds and precipitation in the Weather Research and Forecast model (White et al., 2017), as well as to dampen the influence of aerosol microphysics on cloud condensation nuclei and ice nucleating particles in

75   a regional model (Glassmeier et al., 2017). For the HadGEM-UKCA global aerosol-climate model, Regayre et al. (2018) have shown that both aerosol and physical atmosphere parameters contribute to uncertainty in aerosol effective radiative forcing. Diving into the importance of single processes on the overall CMPs, Bacer et al. (2021) extracted process rates from the chemistry-climate model EMAC, which is based on the same CMPs as this study's ECHAM-HAM. They found that ice crystal sources in large-scale clouds are controlled by freezing and detrainment from convective clouds, while sinks are dominated

80   by autoconversion and accretion. This approach is somewhat similar to a pathway analysis (e.g. Schutgens and Stier (2014); Dietlicher et al. (2019)) in that it deepens understanding of immediate effects, but is not able to relate the effect of a process on variables further down the process chain.

     A promising method for investigating the effect of model input on output is the use of perturbed parameter ensembles (PPEs) (Murphy et al., 2004; Collins et al., 2011). In a PPE multiple input parameters are perturbed at the same time. In this way, PPEs

85   are expanding upon sensitivity studies that vary one parameter (e.g. Lohmann and Ferrachat (2010); He and Posselt (2015)) or multiple parameters at a time (e.g. Ghan et al. (2013)), allowing to investigate the interaction effects of perturbations within the whole possible parameter space. For example, Sengupta et al. (2021) used a PPE to determine the impact of parameters related to secondary aerosol formation on organic aerosol in a global aerosol microphysics model. In a next step, parameter ranges can be constrained when comparing the PPE to observations (Posselt, 2016; van Lier-Walqui et al., 2014, 2019, note that

90   these studies all used synthetic observations as constraints): Morales et al. (2021) built a PPE of CMP process parameters and

environmental conditions, generated using a Markov Chain Monte Carlo algorithm, in idealized simulations to then constrain the parameters with synthetic observations.

Another benefit is that a PPE does not require any additional changes to model code, in contrast to a pathway analysis that requires additional diagnostics and tracers. The downside is that PPEs require many simulations to sample the whole parameter space which is prohibitive given the cost of global climate model simulations. A remedy is the combination of a PPE with a surrogate model such as an emulator. The emulator is first fitted to the PPE model output and then sampled instead of the GCM which is expensive to run. This technique has been used for example to study the effect of model parameters such as the entrainment rate coefficient on climate sensitivity in a GCM (Rougier et al., 2009); or how model parameters affect forecast model drift (Mulholland et al., 2017).

Global sensitivity analysis is a method to quantify the effect of inputs on model output more formally. It allows to divide the total variation in output into the direct contributions from variations in independent inputs as well as from their interactions. For example, Tan and Storelvmo (2016) used variance-based sensitivity analysis on a generalized model of their PPE to determine that the Wegener-Bergeron-Findeisen time scale is the most influential parameter in determining the cloud phase partitioning in mixed-phase clouds. Bernus et al. (2021) have employed sensitivity analysis on their PPE directly to improve the understanding of their lake model prior to its implementation into a climate model.

When dealing with large models that are expensive to run, a surrogate model that is cheap to run allows for a tight sampling of the whole parameter space which permits for sensitivity analysis on the resulting surface. As such, the combination of a PPE with a surrogate model upon which sensitivity analysis is performed has found wide use in cloud simulation studies (Wellmann et al., 2018; Glassmeier et al., 2019; Wellmann et al., 2020; Hawker et al., 2021a). For example, Lee et al. (2011) emulated a global aerosol model and found that the cloud condensation nuclei concentration in polluted environments is dominated by sulphur emissions, but that in remote regions interactions between different parameters are substantial. In particular, a range of recent studies has employed the methodology to investigate how uncertainty in input parameters (which are often not well constrained within parameterisations) translates to an uncertainty of climate model output: quantifying the effect of aerosol parameters on cloud properties or radiative forcing (Lee et al., 2011, 2012; Carslaw et al., 2013; Lee et al., 2013; Regayre et al., 2014; Johnson et al., 2015; Regayre et al., 2015; Yan et al., 2015; Regayre et al., 2018), but also in various other areas of environmental modelling (e.g. a land model in Dagon et al. (2020)). In a further step, the effect of an observational constraint on the model output can be investigated with the emulator/surrogate models (Tett et al., 2013; Williamson et al., 2013; Lee et al., 2016; McNeall et al., 2016; Johnson et al., 2018), yielding important conclusions about which observations are needed to constrain climate models and on which parameters we need to focus research efforts. The approach also lends itself to an investigation of tuning parameters since these also form a parameter space that needs to be explored and constrained (Williamson et al., 2015; Hourdin et al., 2020; Couvreux et al., 2021).

Here we propose a new application of the combined PPE and sensitivity analysis approach to learn about the needed accuracy in process parameterisations within GCMs. Instead of varying parameters within parameterisations, we perturb the processes themselves, as a whole. By perturbing we mean that we vary the effectiveness of a given process, going from using 50 to 200% of a process's effect in the model. For example, if a process affects the ice crystal number concentration, the change induced
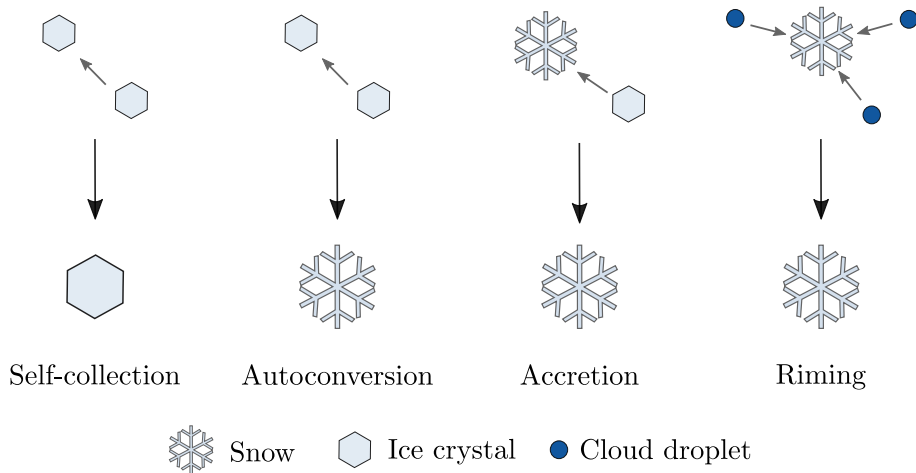
Figure 1. The four cloud microphysical processes investigated in this study, depicted as they are represented in ECHAM-HAM.

on it is multiplied by a perturbation factor between 0.5 and 2 in each timestep. This means that in the extreme cases it would produce half or twice the effect on the ice crystal number concentration that it has in the default model (see Sec. 2.2 for further detail). From the resulting response surface we infer the sensitivity of model output to the CMP processes. The thus generated understanding points to processes whose representation needs to be accurate since they have a large influence, and suggests
130   to simplify those processes that have little influence on model output. Accepting the notion of equifinality, we aim to identify those parts of our current model that do not contribute to variation in output. Thus, we develop a "global sensitivity analysis that can weed out unimportant parameters" as called for by Qian et al. (2016).

To avoid misunderstanding: we are using a surrogate model to learn about sensitivities within the ECHAM-HAM GCM. We are not aiming to replace CMP parameterizations with machine learned substitutes (as e.g. Seifert and Rasp (2020)) or
135   substitute model components (as e.g. Beusch et al. (2020)), because interpretable, physics-based models should be preferred (Rudin, 2019). Instead, in line with Couvreux et al. (2021) we are using emulation and sensitivity analysis as a tool to generate understanding that allows to build a more interpretable model version in a second step. Please note that the potential for simplification is evaluated in the current climate. Thus any derived simplifications would need to be evaluated against a reference model for their suitability in a changed climate state prior to employing it in e.g. climate change projections.
140   In the following Sect. 2 the CMP processes that we investigate, their treatment in the ECHAM-HAM GCM, the generation of the PPE and emulator as well as the sensitivity analysis are described. In Sect. 3 the results from a "one-at-a-time" sensitivity study that explores the axes of the parameter space (Sect. 3.1), the emulated PPE (Sect. 3.2), and of the sensitivity study on the fully sampled parameter space (Sect. 3.3) including a scale dependency (Sect. 3.4) and seasonal analysis (Sect. 3.5) are presented and discussed. Conclusions and an outlook are given in Sect. 4.

145 ## 2 Methods

2.1 Cloud Microphysics in ECHAM-HAM

This study employs the global aerosol-climate model ECHAM6.3-HAM2.3 (Tegen et al., 2019; Neubauer et al., 2019), with a T63 horizontal spectral resolution and 47 vertical levels. The cloud microphysics consist of a 2-moment prognostic scheme for ice crystals and cloud droplets, with additional 1-moment prognostic representation of snow and rain (Lohmann and Roeckner,
150 1996; Lohmann et al., 1999; Lohmann, 2002; Lohmann et al., 2007; Lohmann and Hoose, 2009; Lohmann and Neubauer, 2018). The stratiform cirrus scheme includes homogeneous nucleation of supercooled liquid droplets (Kärcher and Lohmann, 2002a, b; Lohmann, 2003). The stratiform liquid cloud scheme encompasses condensation, aerosol activation, autoconversion of cloud droplets to rain as well as accretion of cloud droplets by rain, evaporation of cloud and rain water, and wet scavenging of aerosol particles (for further details and references see Neubauer et al. (2019)). In stratiform mixed-phase clouds, various
155 CMP processes are included: heterogeneous nucleation via immersion and contact freezing, depositional growth of cloud ice, growth of ice crystals at the expense of cloud droplets via the Wegener-Bergeron-Findeisen process (Wegener, 1911; Bergeron, 1935; Findeisen, 1938), sublimation and melting of ice crystals and snow below clouds. In this study, we are investigating the effect of four different CMP processes involving the ice phase (see Fig. 1): **Self collection** of ice is the process of ice crystals sticking together to form a single ice crystal. **Autoconversion** also has two ice crystals sticking together, albeit forming a
160 snow flake. In **accretion**, a snow flake collects an ice crystal, resulting in a larger snow flake. The fourth process is the only one involving the liquid phase: Cloud droplets are **riming** on a snow flake, again enhancing its size. The implementation of these processes in terms of changes to the ice crystal and cloud droplet mass is detailed in Lohmann and Roeckner (1996), while the implementation of changes to the ice crystal and cloud droplet number concentration is simply in proportion to the mass changes (except for where the mass concentration is unaffected; Lohmann et al. (1999); Lohmann (2002)). The
165 distinction between accretion and autoconversion is necessary due to the separation between ice crystals and snow flakes in their representation as categories of ice in the model. Snowflakes precipitate, while ice crystals are smaller and sediment but do not survive outside of clouds. The four processes were chosen for their comparability as they all represent particle interactions, to represent a range of assumed impacts, as well as for their implementation which is well distinguishable in the code and allowed for easy implementation of the perturbations (see Sec. 2.2). In this study, we do not include any ice
170 multiplication processes. Convective clouds are treated separately from stratiform clouds, except for the interaction through detrained condensate from convective clouds, which is added to stratiform clouds if they exist at the respective model level.

Apart from the perturbations described in the next section, substantial changes that were applied with respect to the published model version ECHAM6.3-HAM2.3 (Neubauer et al., 2019) are:

– Detrained condensate from the convective cloud scheme produced an unrealistically large amount of ice crystals at
175 mixed-phase temperatures, which were then removed with a correction term. The detrained cloud particles are now assumed to be all liquid at mixed-phase temperatures ($0\,°\mathrm{C} < T < -35\,°\mathrm{C}$; Dietlicher et al. (2019); Muench and Lohmann (2020)).

– In line with Muench and Lohmann (2020, Section 3.3.1.2), we now include the immediate, updraft dependent self-collection of detrained ice crystals.

– Previously, a fixed minimal cloud droplet number concentration (CDNC) was applied, which led to unrealistically high CDNCs in high latitude and/or high altitude clouds with low liquid water content (LWC) and hence small droplets. We replace this with a dynamically calculated minimal CDNC, which is calculated from the in-cloud water content and a set maximum volumetric cloud droplet radius (set at $15\,\mu\text{m}$ in the simulations conducted for this study). The resulting minimum CDNC needs to lie between $10^6\,\text{m}^{-3}$ and $4\cdot10^7\,\text{m}^{-3}$. Admittedly, we are replacing the tuning parameter of fixed minimum CDNC with one for a maximum cloud droplet radius. The latter is preferred as it is more physical.

– The model version of Neubauer et al. (2019) contains a mistake in the calculation of the hygroscopicity parameter in the aerosol activation parameterization, leading to an underestimation of the individual aerosol mode solubility. The calculation was updated in Friebel et al. (2019) and subsequently used in Lohmann et al. (2020) and this correction is also applied here.

– In part motivated by the large correction terms highlighted in the process rate study of Bacer et al. (2021) we reduce these if they are unnecessary and/or unphysical. For example, conditions of maximum ice crystal number concentration (ICNC) were enforced after a few CMP processes took place in Bacer et al. (2021). We could reduce the value of that correction term by applying it after each relevant process. Most importantly, the diagnosis of multiple correction terms acting on the same variable led to an artificial increase of corrections. For example, correction terms would enhance ICNC concentrations at model points that later were identified to be outside of a cloud (due to the way the code is structured, the diagnosis of cloud cover happens after e.g. activation/nucleation takes place). In turn, ICNC outside of a cloud were then corrected to be zero, so an unnecessary correction was in fact counted twice. We reduce this artifact by correcting the correction terms themselves. Staying with the example above, the first correction term is now itself set to zero outside of a cloud.

– The sublimation of sedimenting ice crystals appears to be too weak in ECHAM-HAM. This became apparent as in-cloud ICNCs were increasing through sedimentation from above, which indicates that sublimation of ice crystals falling into the cloud-free part of a grid box is too weak. While the underlying problem of a weak sublimation needs to be addressed with future efforts, we introduced a correction of the sedimentation routine: the gain of ice crystal concentrations in the level k into which the ice crystals sediment, $\Delta\text{ICNC}_{\text{sed},k}$ is restricted to the loss of in-cloud ice crystal number concentration in the lowest model level above level j that lost ice crystals by sedimentation. Also, in-cloud ICNC and snow formation rate are now set to 0 outside of clouds inside the ice crystal sedimentation routine where they were previously set to the grid-mean values. This contains the implicit assumption that ice crystals do not survive sedimentation outside of a cloud in ECHAM-HAM.
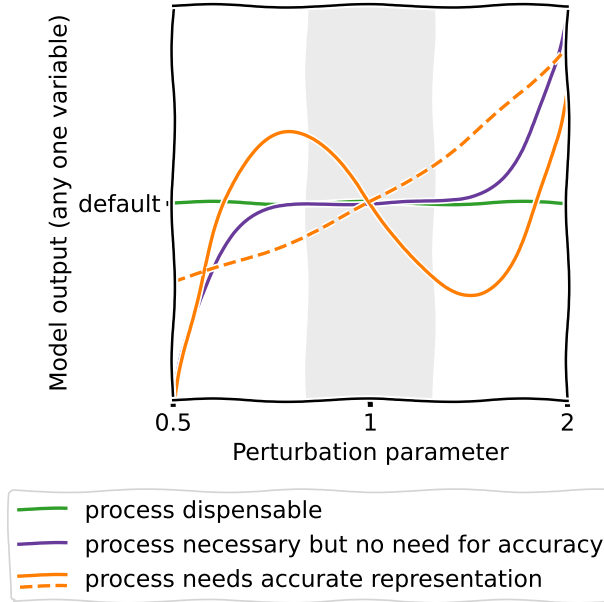
Figure 2. Sketch of the envisioned interpretation. The shading indicates the area that is of most interest to judge the effect of process simplifications on the model output. If the slope in this area is small, this suggests that the process can be simplified (green and purple lines). A large slope indicates that the process needs to be represented accurately (orange lines). If no perturbations of the process in the 0.5 to 2 perturbation parameter range and the suppression of the process (perturbatin parameter of 0, not shown) have a significant influence on the model output, the process may be removed entirely (green line).

With the described changes, the model requires retuning. The tuning procedure follows the one described in Neubauer et al. (2019), with the final tuning parameters given in Table A1 in Appendix A. Model simulations were conducted with the same tuning for all simulations.

2.2 Perturbations as a proxy for complexity

In order to see the effect of whole processes on model output, we can turn processes off in sensitivity studies. In the present study, we achieve this by setting to zero the change that the process induces on prognostic variables. For example, at every model timestep t autoconversion impacts the ICNC:

$$\text{ICNC}_{t+1} = \text{ICNC}_t + \Delta\text{ICNC}_{\text{autc}} \tag{1}$$

We can turn off the effect of autoconversion by multiplying $\Delta\text{ICNC}_{\text{autc}}$, the change in ICNC due to autoconversion in one timestep, by zero when it is added to the affected variables.
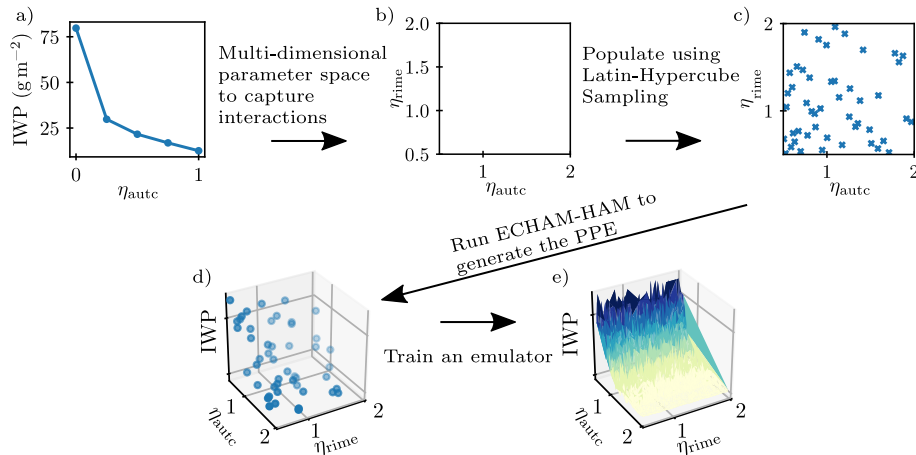
Figure 3. Sketch of the employed methodology: we move from **a)** one-dimensional sensitivity studies where one process is perturbed by varying the parameter η (Sect. 3.1) to **b)** a multi-dimensional parameter space. **c)** The input parameter space is filled with Latin Hypercube sampling and supplied as input to ECHAM-HAM. The simulations form the perturbed parameter ensemble (PPE). The **d)** PPE output is **e)** fitted using a Gaussian Process emulator for each variable of interest to generate a smooth response surface, upon which sensitivity analysis can be applied. Note that this is an illustrative sketch of the method for a PPE with two input dimensions, whereas our PPE has four dimensions, and that the data used to generate it is only illustrative as well. The shading in **d)** illustrates depth only.

More generally, instead of setting to zero the changes induced by a process, we can perturb the process using a newly defined parameter η.

$$\text{ICNC}_{t+1} = \text{ICNC}_t + \eta_{\text{autc}} \cdot \Delta\text{ICNC}_{\text{autc}} \tag{2}$$

This perturbation of whole processes was introduced by van Lier-Walqui et al. (2014) to estimate the uncertainty including errors in the physical assumptions of process formulations. In our case, the parameters aid to understand the sensitivity of the model to each process: From the response of model output to variations in η_i, we can extract information on how accurately a process i needs to be represented in the model (see Fig. 2 for a visualisation). For example, if the model output variable (e.g. ice water path, IWP) as a function of η_i has a slope close to zero at $\eta_i = 1$ (green and purple line in Fig. 2), this suggests that the process i needs to be represented only approximately and that some detail could probably be removed from its parameterisation without much of an effect on the model performance. Note that the perturbations are constant in space and time for each PPE member, serving as a proxy for understanding the effect of possible simplifications, which would likely be variable in time and space. In this study, four cloud microphysical processes, namely self-collection, autoconversion, accretion and riming (see Fig. 1) are perturbed, i.e. $i \in [1,4]$. Combining perturbations of multiple processes allows us to study and take into account possible interaction effects, such as the compensation by one process which is perturbed by another one.

2.3 Generating and emulating the perturbed parameter ensemble (PPE)

In a first scoping study, we perturb each process one by one, by multiplying its effect with $0 < \eta_i < 1$. Multiplicative perturbations between zero and one corresponds to a reduction in the process' effectiveness. However, to take into account interactions, all η_i need to be varied at the same time, thereby creating a multi-dimensional input parameter space in a second step. In addition, the range of η_i is expanded to values up to $\eta_i = 2$ to imitate an overestimation of a given process due to an inaccurate description. As we are most interested in the space around $\eta_i = 1$ and to sample the over- and underestimation equally, we vary η_i from 0.5 to 2 in the multi-dimensional input parameter space. If the process uncertainty were known, it would influence the extent of the perturbation range, which could be different for each process. The perturbations and the procedure described in the following are visualized in Fig. 3. To probe the multi-dimensional input parameter space effectively, the sets of input parameter combinations $(\eta_1, \eta_2, \eta_3, \eta_4)$ to be simulated with the model were generated with Latin Hypercube Sampling (LHS, using the Python library PyDOE (tisimst, 2021)), which maximizes the spacing between inputs and provides good coverage of the parameter space, even when only a few input parameters are important (Morris and Mitchell, 1995). The LHS was applied on the logarithmically scaled input range to account for the multiplicative behaviour of the η_i. Each of the LHS generated input combinations was then used as input for a 1 year ECHAM-HAM model simulation, creating a perturbed parameter ensemble (PPE) with 48 members. This is in line with the suggestion of Loeppky et al. (2009) to use 10 times as many training runs as the number of input parameters for such a computer experiment. To estimate the inter-annual variability, the control simulation with all processes at full effectiveness ($\eta_i = 1 \forall i$) spanned 10 years. This estimate is used to judge whether perturbations observed in the PPE are significantly larger than the inter-annual variability and therefore contain a signal that originates from the perturbation in η_i. As the inter-annual variability exhibited no strong variations throughout the probed phase space in the one-at-a-time sensitivity studies, the 1 year simulations for the PPE members in combination with the control simulation estimate of the variability were deemed sufficient for the analysis. All the simulations were performed with climatological sea surface temperatures and sea ice extents, and aerosol emissions representative for the year 2003. These simulations were not nudged to meteorological data but ran freely so that the full effect of perturbing the processes could be observed. Each simulation included 3 months of spin-up that was not included in the analysis.

Using the PPE output as input for the creation of a surrogate model, we can construct a smooth response surface over the whole parameter space (see Fig. 3e)). As a surrogate model, we choose a Gaussian process emulator (O'Hagan, 2006; Rasmussen and Williams, 2006), which has found wide use in atmospheric and climate science (Lee et al., 2011; Carslaw et al., 2013; Johnson et al., 2015). We prefer the Gaussian process emulator over e.g. a neural network because of its demonstrated suitability and need for fewer input data (see Watson-Parris et al., 2021a, for a more in-depth discussion). Using a recent Python package for emulating Earth System Models (Watson-Parris et al., 2021b, a), the implementation is straight forward. From the PPE, we can construct a surrogate model for every output variable that we are interested in by training a separate emulator for each output variable (ice crystal and cloud droplet number concentration, ice and liquid water path, shortwave and longwave cloud radiative effect, cloud cover, surface precipitation, ice, liquid and mixed-phase cloud cover). For the kernel (or covariance function, Watson-Parris et al., 2021a), an additive combination of the linear, polynomial, bias and exponential

kernels was used as this performed best in preliminary tests (not shown, Duvenaud (2014)). Other model specifics were set as default in Watson-Parris et al. (2021b). As the emulation operates best on standardized data with zero mean and unity variance, the mean was removed from the input data, which was then scaled by dividing it by the standard deviation, prior to emulation.

With the cheap surrogate model a variance-based sensitivity analysis (see Sec. 2.5) becomes feasible (Oakley and O'Hagan, 2004), picking 3000 samples from the emulator as input. This approach is similar to Johnson et al. (2015), except that they perturbed CMP parameters while we vary the effectiveness of whole CMP processes. It allows us to identify the importance of the different η_i for the variables in question and thereby the processes which require a detailed representation.

2.4 Validation

To make sure that the chosen emulators are a fair representation of the model output, we validate them according to Bastos and O'Hagan (2009) except for using leave-one-out validation, as visualized in Fig. 4 for the IWP. In Fig. 4 a) and b), the individual standardized errors, $\frac{Y_{\mathrm{sim}} - Y_{\mathrm{emu}}}{\sqrt{V_{\mathrm{emu}}}}$ (with Y_{sim} and Y_{emu} the output of the ECHAM-HAM simulations and the emulated output, respectively, and V_{emu} the emulator variance), are plotted against the emulated output and input parameters. We observe only few errors larger than 2, which would signal a conflict.

We employ a QQ-plot to determine whether the normality assumption of a Gaussian process is met in the emulator (Bastos and O'Hagan, 2009). The plot compares the quantiles of the standardized errors against those of a Student-t distribution. Fig. 4 c) indicates that the normality assumption holds and that the predictive variability is well estimated by the emulator (Bastos and O'Hagan, 2009). In a direct comparison of emulated and simulated ECHAM-HAM model output (Fig. 4d)), the points should lie close to the line of equality, with the 95% confidence bounds on the emulator predictions crossing it. This should be the case for 95% of the validation points. In our emulations, the number of points with confidence bounds that do not cross the line of equality sometimes is larger (up to 20%), depending on the variable. We attribute this to the disruptive changes that the CMP process perturbations induce as compared e.g. to the aerosol and CMP parameter changes applied by Johnson et al. (2015) (which did not include ice crystal autoconversion and perturbed parameters only within uncertainty bounds instead of whole processes), as well as to the fact that the simulations were not nudged. The difficulty in emulating the response surface for some of the variables was also apparent in computational limitations: some of the leave-one-out validation emulations were not possible to compute because of numerical instabilities in the computations when constructing the emulator. As these were only a few cases (up to six for global means and four for seasonal means in 48 validation emulations), the validation for those variables as a whole is still deemed valid.

The good qualitative agreement with the line of equality and the lack of systematic errors are sufficient for a validation of the emulator, especially considering that we are not aiming for exact quantitative estimates as results of the presented analysis. Rather, we are looking for a conceptual understanding of the need for an accurate description of CMP processes, for which this emulator validation is sufficient.

For the variables which passed the leave-one-out validation, the final emulator used for the sensitivity analysis was trained on all PPE members (note that in a few cases only 47 PPE members were used due to numerical instabilities in the computations when constructing the emulator). Note that the setup of the emulator includes design choices such as the kernel combination to
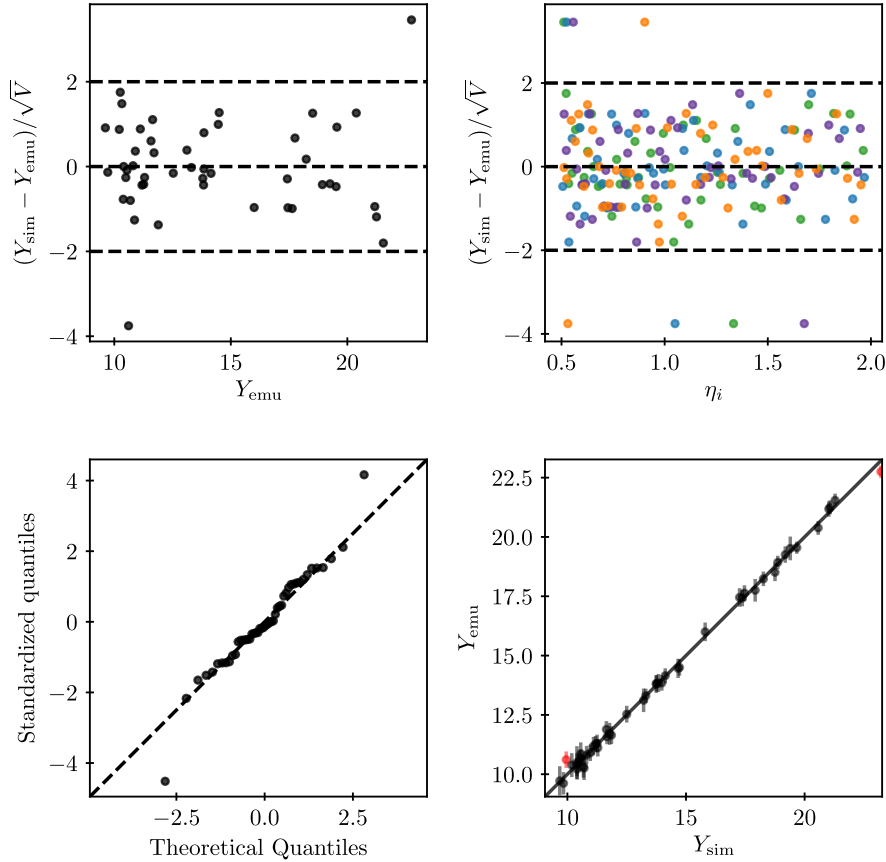
Figure 4. Leave-one-out validation of the emulator for global annual mean IWP. Each point corresponds to the training of the emulator on all points except one and then testing on exactly that point. Individual standardized errors are plotted against **a)** emulator output and **b)** input parameters (colors according to Fig. 5: autoconversion (blue), accretion (purple), riming (green), self-collection (orange)). The dashed lines are drawn at an individual standardized error of zero and 2, which is the threshold discussed in Bastos and O'Hagan (2009). **c)** QQ-plot of the individual standardized errors against a student-T distribution. **d)** Emulator against model output, with the error bars indicating the 95% confidence interval on the emulator predictions. Predictions for which the model result lies outside that interval are marked red.

use. Therefore, the present emulator is only one of multiple possible emulators that could be used to represent the model data. However, as it is shown to validate well, other setups are expected to lead to the same conclusions as this one in the analysis.

2.5 Sensitivity analysis

In our framework, the question of how detailed the representation of a given process i needs to be translates to the question of how sensitive the model output is to a variation of the perturbation parameter η_i. For an answer, we employ variance based sensitivity analysis, following Saltelli (2008). In contrast to derivative-based local methods (Errico, 1997), global variance based sensitivity analysis allows for an investigation of sensitivities within the whole input parameter space. Its main metrics are the first and total order sensitivity indices (S_i and S_{Ti}, respectively). The first-order sensitivity index of η_i measures the contribution of variance in η_i to the variance in an output variable Y. It is constructed as

$$S_i = \frac{V_{\eta_i}\left(E_{\eta_{\sim i}}\left(Y|\eta_i\right)\right)}{V(Y)} \tag{3}$$

E is the average over Y with all η except η_i ($\eta_{\sim i}$) being allowed to vary while η_i is kept fixed at η_i^*. Then V_{η_i} is the variance over that average, for varying η_i^*. S_i is always between 0 and 1, and high values signal an important variable. For additive models all first-order terms add up to one, i.e. $\sum_i S_{\eta_i} = 1$. In non-additive models (e.g. a climate model) interaction terms also have to be taken into account. However, in models with many input parameters the computation of all interaction sensitivities can be cumbersome. The total effect sensitivity index S_{Ti} offers a remedy in that it summarizes all direct and interactive effects a parameter's variance has on the total variance in output (Homma and Saltelli, 1996; Saltelli, 2008). It is defined as

$$S_{Ti} = \frac{V_{\eta_{\sim i}}\left(E_{\eta_i}\left(Y|\eta_{\sim i}\right)\right)}{V(Y)} \tag{4}$$

Here all-but-η_i ($\eta_{\sim i}$) are kept fixed at $\eta_{\sim i}^*$ and only η_i is allowed to vary for the average E_{η_i}. Then the variance of that average over varying $\eta_{\sim i}^*$ is computed and divided by the variance in output Y. Saltelli et al. (1999) argue that the first and total sensitivity index suffice for a meaningful global sensitivity analysis. To compute these indices via the Sobol method, we make use of the Python library SALib (Herman and Usher, 2017).

3 Results and Discussion

3.1 One-at-a-time sensitivity studies

In a first scoping experiment, we perturbed each process separately, which one can imagine as tracing the edges of the cube shown in Fig. 3. The results are presented in Fig. 5. Of the four perturbed processes, turning off autoconversion has the largest effect on model output: the global annual mean ice water path (IWP) is more than doubled, and the increase in cloud cover and decrease in precipitation dwarf the changes induced by turning off the other three processes. In fact, the perturbations induced by perturbing accretion and self-collection are mostly insignificant compared to the inter-annual variability. As autoconversion is a removal process for ice crystals, it is reasonable that its suppression leads to an increase of ice in the atmosphere (note that
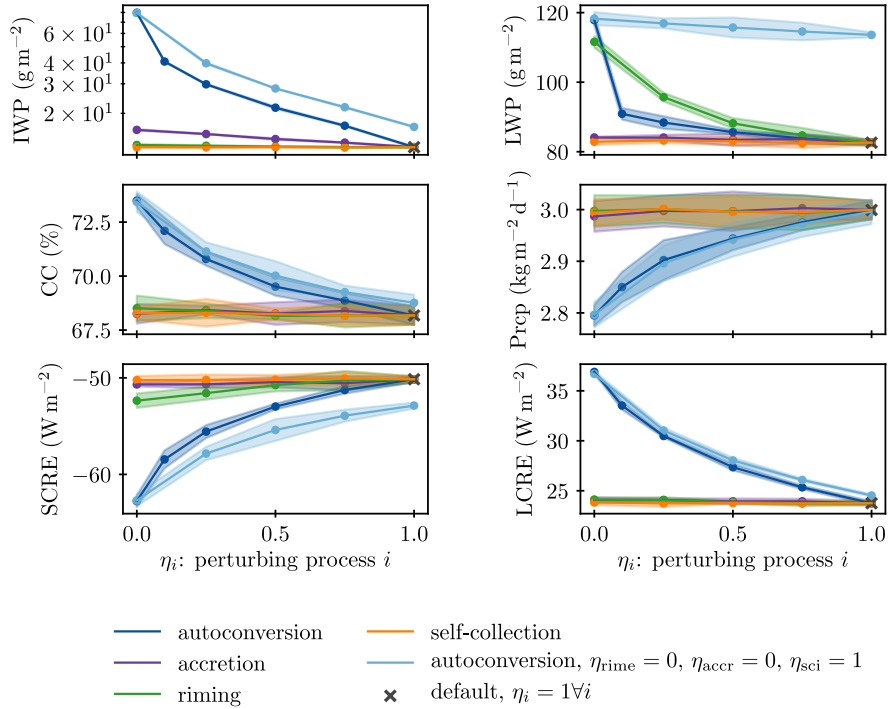
Figure 5. Model response to perturbations of four CMP processes: autoconversion, accretion, riming and self-collection (as illustrated in Fig. 1) in terms of global annual mean IWP, liquid water path (LWP), cloud cover (CC), precipitation (Prcp), shortwave and longwave cloud radiative effect (SCRE, LCRE). An additional experiment was conducted to highlight interactive effects between the perturbation of autoconversion and the suppression of riming and accretion (light blue). The points and line indicate the mean and the shading indicates 2 times the standard deviation of annual mean values of a five-year simulation. Classical sensitivity studies would only show $\eta_i = 0$ and $\eta_i = 1$. Note that we added an extra simulation at $\eta_{\mathrm{autc}} = 0.1$ to better illustrate the threshold behaviour discussed in the text and that for the IWP the shading is hidden behind the lines.

the IWP in ECHAM-HAM only counts ice crystals and not snow). Similarly, riming is a removal process for liquid droplets, so the liquid water path (LWP) increases with its suppression. However, surprisingly the suppression of autoconversion induces a similarly large increase in LWP as that of riming, even though autoconversion includes no direct interaction with liquid droplets. The shape of the model response to the gradual perturbation of the processes holds additional information: while the generated model response is mostly gradual, for low η_{autc} the response is more abrupt. This behaviour, which we call a threshold response, is most striking for the global annual mean LWP, for which the signal for $\eta_{\mathrm{autc}} \geq 0.25$ is not significantly different to that of accretion and self-collection. When autoconversion is completely suppressed, the LWP increases dramatically and the signal becomes stronger than that for riming, which had increased consistently and gradually. This behaviour can be explained by autoconversion acting as a catalytic process for accretion and riming, creating a threshold behaviour when it is turned off. As can be seen from Fig. 1 it is the only process that generates snow flakes. Accretion and riming need the snow flakes to
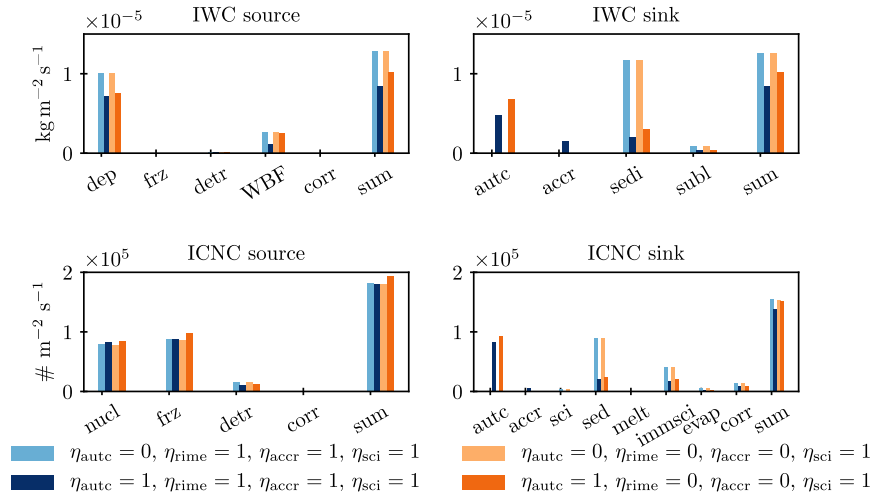
Figure 6. Global annual mean vertically integrated process rates for four experiments that illustrate the suppression of snow formation through turning off autoconversion (mean of a five-year simulation). The rates are diagnosed similar to Bacer et al. (2021), but correction terms were themselves subtracted from process rates where appropriate, i.e. where the correction belongs to the logical entity of the process rate (see Sec. 2.1). The process rates are: deposition (dep), heterogeneous and homogeneous freezing (frz), detrainment (detr), deposition in the Wegener-Bergeron-Findeisen process (WBF), correction terms (corr), autoconversion (autc), accretion(accr), sedimentation (sedi), sublimation (subl), ice nucleation in the cirrus scheme (nucl), melting (melt), immediate self-collection of ice crystals when the ICNC is larger than a maximal threshold (immsci), evaporation (evap).

340   be able to act upon them. Therefore, when autoconversion is turned off, accretion and riming are consequently suppressed as well. In this way, the suppression of autoconversion can strongly influence even the liquid phase. The simulations in which we perturb autoconversion while having riming and accretion turned off confirm this hypothesis (light blue line in Fig. 5): Throughout most of the phase space, turning off accretion and riming reinforces the signal from phasing out autoconversion. However, when autoconversion is turned off, turning off accretion or riming does not change the model output any further. That

345   is because they are both suppressed when autoconversion is turned off and does not generate any snow for them to act upon.

     Fig. 6 further elucidates the reaction of the model to a suppression of autoconversion further: the snow formation rate decreases dramatically, and with increased ice concentrations in the atmosphere, the other removal processes of sedimentation and melting subsequently increase. Again the suppression of riming and accretion only influences the model output when autoconversion is active. When autoconversion is turned off, accretion and riming have no influence.

350   From this one-at-a-time example, one can already see the benefit of the perturbation approach: In classical sensitivity studies, where processes are only turned on and off, only the large signal induced by autoconversion would have been visible. However, here it was the peculiar shape of the model response to the whole perturbations that hinted at the threshold effect of autoconversion. The implications for possible simplifications are different: seeing only the large difference between a simulation with and without autoconversion, one would think that this is an immensely important process. Recognizing it as a threshold process
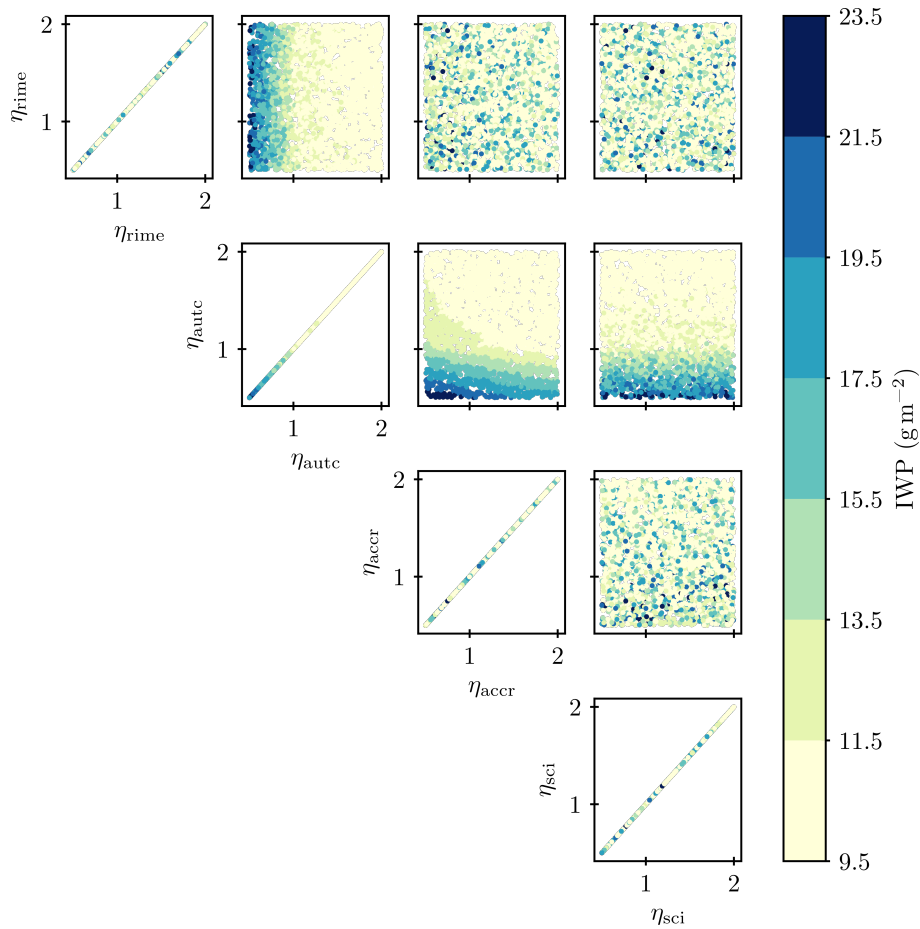
Figure 7. Two dimensional projections of the IWP values sampled from the four dimensional parameter space of the emulated PPE. Each perturbed process is a dimension, and the colorbar denotes the global annual mean ice water path for each input parameter combination.

355 and seeing the gradual response to small deviations from 1.0 in η_{autc} (similar to the purple curve in Fig. 2), it appears that there is potential for a less accurate description of autoconversion in the model. It has also become clear that interaction effects need to be taken into account as well to explain the model behaviour. This is what the PPE expands upon in the next section.

3.2 PPE of global mean variables

Conducting a one year simulation with ECHAM-HAM for each of the 48 input parameter combinations generates the PPE
360 which is then emulated (see Fig. 3). Fig. 7 illustrates the resulting response surface with points sampled from that emulation of the annual global mean IWP. To generate the multi-dimensional response surface 48 one year simulations were needed, compared to the 21 simulations that were needed to investigate the response along only a few of the parameter space edges in Sec. 3.1. This illustrates the value of the chosen approach: the emulated PPE provides more information while needing only
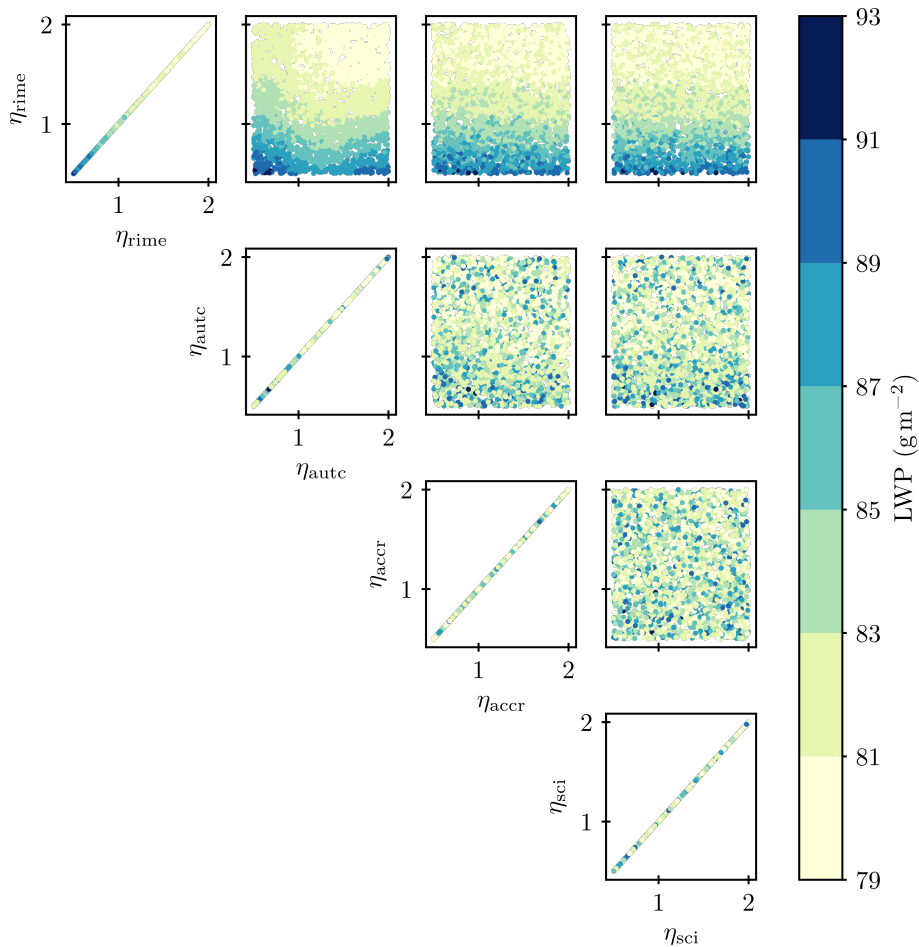
Figure 8. Same as Fig. 7 but for the global annual mean liquid water path. Results for additional variables are presented in Fig. B1 in Appendix B.

roughly twice as many simulations. The surface shows an ordered ascent with decreasing η_{autc}, while the other dimensions

365 exert no control over the value of the IWP. Only for accretion a slight influence is visible from the tilted contours in the phase space shared with autoconversion. Increased accretion depletes the IWP since it converts ice crystals to snow flakes. Fig. 8 shows that the LWP is dominated by η_{rime}, with an additional influence of autoconversion. The LWP decreases with increasing η_{rime} and increasing η_{autc}. This is because riming depletes the atmosphere of cloud droplets and a decrease in autoconversion suppresses riming. The panels in Fig. 7 and 8 exhibiting no order in their parameter space distribution indicate that the processes

370 in question exert no influence on the respective output variable. Similar to the LWP, the CDNC is dominated by riming, and for other cloud variables the dominant influence of autoconversion is confirmed as well (see Fig. B1 in Appendix B).

The ranges in the global annual mean model variables that we observe are mostly larger than what Lohmann and Ferrachat (2010) find for varying uncertain tuning parameters, indicating that whole processes exhibit a larger influence on the model
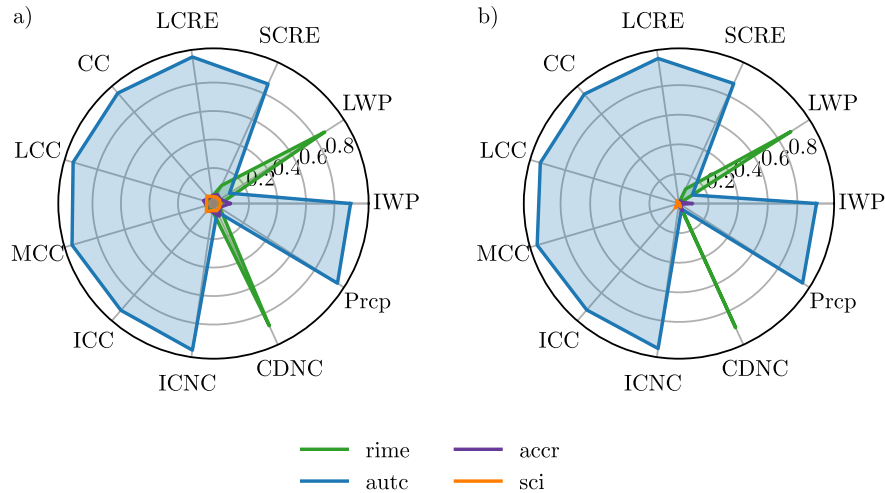
Figure 9. First order (**a**)) and total effect (**b**)) sensitivity indices for the emulated response surface of global annual mean cloud cover (CC), liquid cloud cover (LCC, $T > 0\,°C$), mixed-phase cloud cover (MCC, $0\,°C < T < -35\,°C$) ice cloud cover (ICC, $T < -35\,°C$), longwave cloud radiative effect (LCRE), shortwave cloud radiative effect (SCRE), liquid water path (LWP), ice water path (IWP), and total precipitation. As described in Sec. 2.5, the indices are always between 0 and 1, and high values signal an important variable. Since the climate model is non-additive, the terms do not add up to one as interactions have to be taken into account.

response than those single parameters. Only for LWP do Lohmann and Ferrachat (2010) find a larger range of about $50\,\mathrm{g\,m^{-2}}$
375   when they multiply the autoconversion rate with a factor between 1 and 10. As this warm-rain process is not included in the present analysis, it is reasonable that the observed variation for LWP is smaller.

3.3   Sensitivity analysis

A global variance based sensitivity analysis allows us to quantify the qualitative sensitivities obtained from the graphical representations of the emulated surfaces in the previous section. The results for the first order (S_i) and total effect (S_T)
380   sensitivity indices are presented in Fig. 9. Indeed, the qualitative results are confirmed: the global annual mean LWP and CDNC are dominated by riming while all other variables are dominated by autoconversion, both in first order and total effect.

The observed sensitivities are different from what Bacer et al. (2021) find in their investigation of EMAC ICNC process rates. They find that autoconversion contributes about twice as much as accretion to the ICNCs, while self-collection has a negligible role. In our analysis, the influence of autoconversion dwarfs that of accretion in terms of sensitivity indices as well
385   as for the process rates (see Fig. 6). The sensitivity indices are not directly comparable to Bacer et al. (2021). However, for the default simulation the process rates are diagnosed as in Bacer et al. (2021) and thus comparable. We attribute the observed differences to the slightly different model version used in Bacer et al. (2021), which goes along with a different tuning.

The almost binary results for the sensitivity indices are surprising, as in other studies the sensitivity indices were more evenly distributed (Lee et al., 2011; Wellmann et al., 2018, 2020). However, these studies usually employed a wider suite of input parameters, whereas here only processes from the limited system of ice particle interactions are included. We expect that with additional cloud microphysical processes included, the sensitivities would be more evenly distributed as well. The binary signal is due to the strong dominance of autoconversion throughout the parameter space and not due to the threshold behaviour upon suppression of autoconversion as analysed in Sec. 3.1. This was excluded from the sensitivity analysis as only the input parameter space with $\eta_{\mathrm{autc}} \geq 0.5$ was taken into consideration.

The dominance of autoconversion is hypothesized to originate from the non-linearity in its parameterization. In contrast to the other processes, the conversion rate of autoconversion has a squared dependency on the cloud ice content (see Lohmann and Roeckner (1996)), increasing feedback effects between the two.

Additional reasons for the large role of autoconversion may lie in its role as a tuning parameter in ECHAM-HAM. For tuning, uncertain parameters of the model are used (Neubauer et al., 2019). Historically, the scaling factor for the stratiform snow formation rate by autoconversion, γ_s, has been used as it represents a counterpart to the scaling factor for the stratiform rain formation rate by autoconversion. To reach the tuning goals as detailed in Neubauer et al. (2019), it is brought to unrealistically high values (see Table A1). This enhances the changes induced by perturbing autoconversion in this study using η_{autc}. Additionally, structural problems in the model may enhance the role of autoconversion artificially. For example, by accounting for heterogeneous nucleation in the cirrus scheme, which increased ice crystal sizes, Gasparini et al. (2018) were able to reduce γ_s by an order of magnitude compared to the reference ECHAM-HAM version (personal communication). This in turn would be expected to reduce the importance of autoconversion in the present analysis. Moreover, also the design choices of the CMP scheme, e.g. the order in which processes are called, may influence the results. However, learning about the properties of CMP processes in the ECHAM-HAM model is important, no matter whether they are physically based or artificially introduced through model design.

A caveat to these results is of course that only CMP processes were investigated here. Parameters or processes from other parts of the climate model e.g. the dynamics might exhibit an even larger influence on the investigated model output if they were allowed to be varied. For example, Wellmann et al. (2020), using idealized COSMO simulations, found that environmental conditions are more influential for the diabatic heating rates than microphysical processes. However, for the research question at hand, namely how accurate the representation of these four processes within the CMPs needs to be, the comparison of the processes between each other is sufficient. Indeed, the negligible sensitivity of model output to variations in accretion and self-collection of ice suggests that their representation may be simplified (Lee et al., 2012). Due to the small deviations in the considered variables in response to variations around $\eta_i = 1$ for riming and autoconversion (purple line in Fig. 2), there is potential for simplifications of their formulations. In the grand scheme of CMP parameterization development, however, autoconversion as the most dominant process of the four is a key process to scrutinize given the possibly troubling origin of this dominance in its role as a tuning factor.

3.4 Scale dependency analysis

The analysis of global annual mean values yields clear conclusions, but climate models not only need to simulate global mean values correctly but also their spatial and temporal evolution. Since the emulation and subsequent analysis of grid-point level data is tedious and error-prone due to the small signal and large noise, we compress the information in the data to a space of lower dimensionality. Choosing to reduce the dimensionality but still represent the whole global data rather than picking certain regions allows for a more objectified and unbiased analysis. This is similar to Holden et al. (2015) who also reduce their high-dimensionality output, albeit with singular value decomposition, and Ryan et al. (2018) who use principal component analysis. However, as the model data is complete and on a sphere, a spherical harmonics expansion is our method of choice.

Mathematically, the model data can be represented as a linear combination of the orthogonal spherical harmonics basis functions as follows:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_l^m Y_l^m(\theta, \phi) \tag{5}$$

The data f is then a function of the longitude θ and latitude ϕ, with Y_l^m a spherical harmonics function of degree l and order m (l and m are integers, with $-l \leq m \leq l$). The complex coefficients F_l^m can be computed as:

$$F_l^m = \int_{\Omega} f(\theta, \phi) Y_l^m(\theta, \phi) \mathrm{d}\Omega \tag{6}$$

The coefficients make up the angular power spectrum S_{ff}:

$$S_{ff}(l) = \frac{1}{4\pi} \sum_{m=-l}^{l} |F_l^m|^2 \tag{7}$$

where the sum over the angular power spectrum $\sum_{l=0}^{\infty} S_{ff}(l)$ is the variance of the data. In principle, an expansion up to order 95 would represent the model data at its resolution of 96 latitudinal and 192 longitudinal points perfectly, as they are equidistant in spherical coordinates. In practice, we truncated the expansion at the degree l where it represents 95% of the total data variance $\sum_{l=0}^{\infty} S_{ff}(l)$. Thereby we represent the data with as few as possible but as many as necessary basis functions. Note that in principle, a principle component analysis could yield the same representation with fewer basis functions. However, these functions would depend on the investigated dataset, while the use of spherical harmonics allows for inter-comparability.

Fig. 10 illustrates that a spherical harmonics expansion of the data can serve as an accurate representation, while all the information can be stored in the coefficients up to $l = 20$ instead of on the global grid (see Fig. 10c). Thus confident that the expansion represents the data accurately we can conduct a spatially resolved sensitivity analysis in the spherical harmonics space. For each variable and degree l a separate emulator was trained on the angular amplitude spectrum $\sqrt{S_{ff}}$, from which samples were drawn as input to the sensitivity analysis. The validation procedure was the same as described in Sec. 2.4. Spherical harmonics members of degree l were excluded from the sensitivity analysis when the emulator was found to be defaulting to an equal prediction over the phase space (see Appendix D). This was the case mostly for degrees l for which the coefficients could be seen to have less amplitude in the angular amplitude spectrum already. 4 out of 11 variables had to be excluded because too many members were defaulting or because their variations were too small to be sensibly emulated.
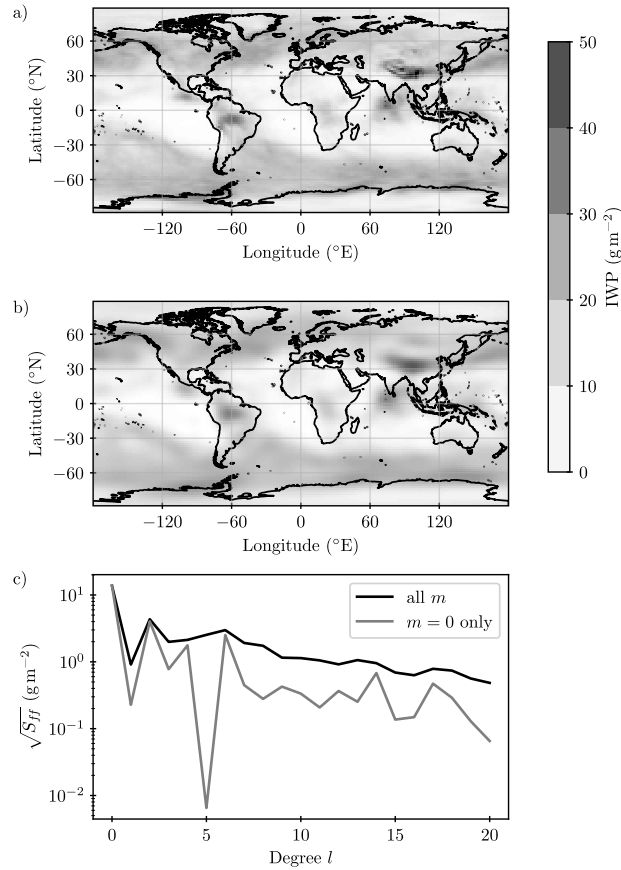
Figure 10. Spherical harmonics expansions for one illustrative PPE member ($\eta_{\mathrm{autc}} \approx 0.87$, $\eta_{\mathrm{accr}} \approx 1.43$, $\eta_{\mathrm{rime}} \approx 0.81$, $\eta_{\mathrm{sci}} \approx 1.89$). **a)** Global annual mean IWP, **b)** expansion of spherical harmonics representing the same data as **a)**, generated from the coefficients of the expansion displayed as an angular amplitude spectrum in **c)** as a function of the degree l (with m independent solutions, where modes of $m = 0$ most strongly resemble rotationally symmetric physical patterns of the Earth system such as a North-South contrast). Note that the variability explained by each degree l in general decreases with increasing l, which allows us to truncate the expansion at the degree l where it represents 95% of the total data variance.
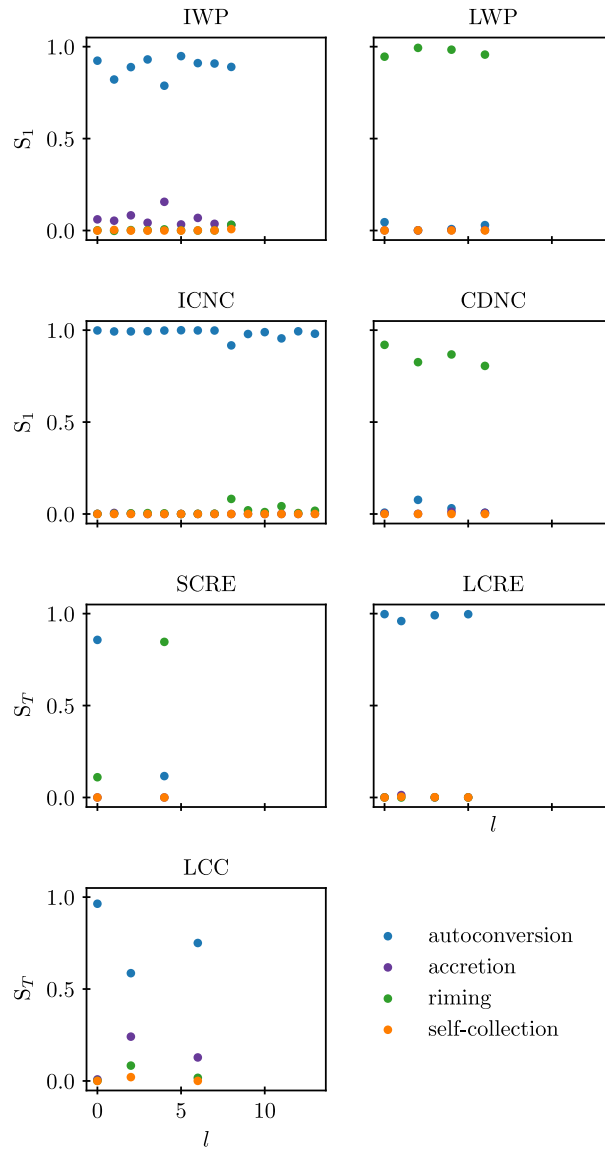
Figure 11. First order sensitivity indices for the emulated angular amplitude spectrum as a function of the spherical harmonics degree l for the variables as described for Fig. 9. As detailed in the text, emulators that were found to be defaulting in the validation procedure were not subjected to the sensitivity analysis so that the results for those l are missing here. The results for the total sensitivity index are shown in Fig. C1 in Appendix C.

The results are displayed in Fig. 11. For those variables that had total sensitivity indices for autoconversion of over 0.9 (IWP, longwave cloud radiative effect, ICNC, and liquid cloud cover) the dominant effect of autoconversion is present on all length scales. Accretion is of secondary importance for the IWP, as indicated by the global sensitivity analysis. The LWP and CDNC are dominated by riming on all regional scales and on the global scale.

The emulated surfaces for the spherical harmonics are more uncertain than those for the global mean values (see Appendix D). This is expected as the training data is more noisy and indicates a less detectable signal on smaller length scales than on the global one. In addition the separate emulation for different degrees l ignores correlations between signals included in multiple degrees l, which may lead to the loss of signals that are small in the different l but correlated and therefore should be addressed in future studies. However, as the results of the sensitivity analysis are clear in that variability is dominated by autoconversion (see Fig. 11), we can conclude that the results of the global sensitivity analysis also hold on regional scales.

Finally, this analysis demonstrates that spherical harmonics expansion is a viable tool to evaluate model output on all length scales in an efficient and objective manner. Future studies may use it to compare results e.g. from different models. As most expansion degrees are physically difficult to interpret, the method may be expanded to use physically meaningful modes such as the land-sea contrast instead.

3.5 Seasonal analysis

Similar to a regional analysis, we use a temporally resolved sensitivity analysis to address the concern that conclusions drawn from annual mean values might not hold on a seasonal scale. Fig. 12 shows the results of the same sensitivity analysis as in Fig. 9, but split by seasons (one emulator per variable was trained and validated for each season; note that in few cases only 47 PPE members were used as with the 48th member the computational constraint was too tight for the emulator). Due to a weaker or less consistent signal in the data on seasonal scales, two variables (mixed-phase cloud cover in DJF and JJA) did not pass the validation procedure as the emulator was found to be defaulting (as described for the spherical harmonics above and in Appendix D). Fig. 12 reveals that indeed the sensitivities to process perturbations are much the same as for the annual mean analysis. This confirms that the conclusions drawn for model simplifications also hold on a seasonal scale. The model is not sensitive to accretion and self-collection of ice and therefore these processes can be simplified, while autoconversion and riming dominate the model response.

3.6 Process costs and implications for simplification

The previous analysis shows that the response of ECHAM-HAM to a suppression of self-collection or accretion is negligible, while for riming and autoconversion a less accurate representation may be appropriate. A potential benefit could lie in the reduction of CPU time per model simulation. Table 1 lists the CPU time spent in the CMP routines of the four processes. The timings represent an estimate of how much time could be gained by removing a process from the model. They show that at most, with naively removing (the most drastic simplification) the whole cold precipitation formation routine, only about 0.2% of total computing time can be saved (since the cold precipitation formation routine makes up 4.8% of the 4.7% of computing
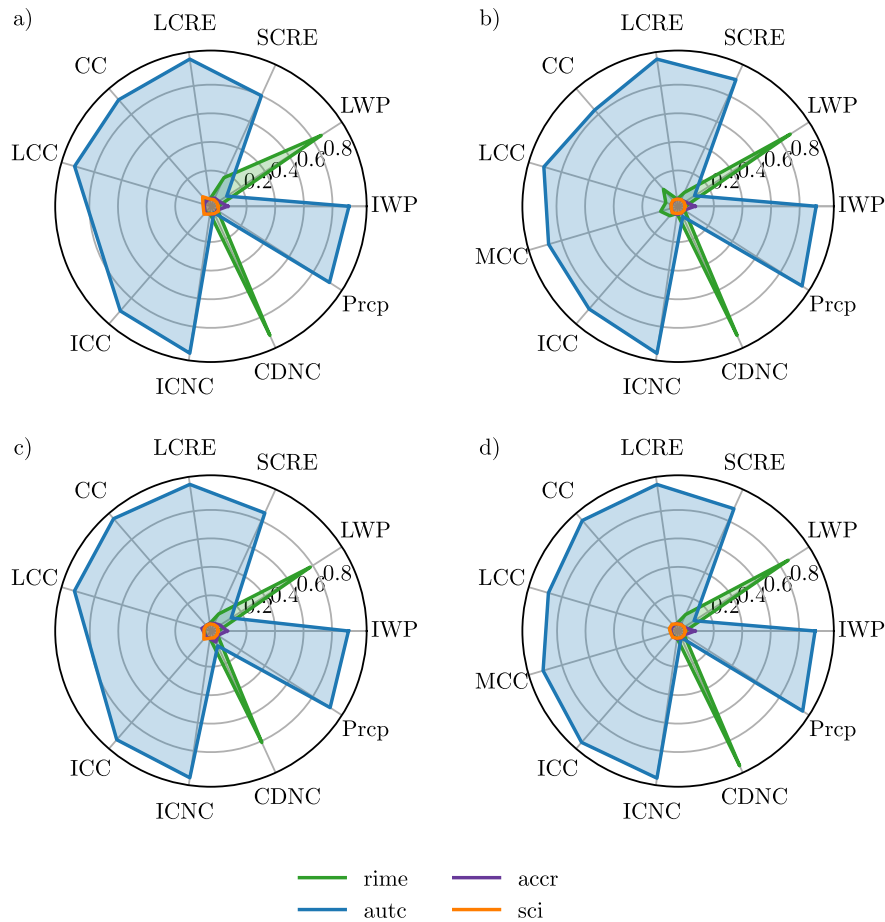
Figure 12. Same as Fig. 9 but with seasonal means (**a**) DJF, **b**) MAM, **c**) JJA, **d**) SON) and only first order sensitivity indices shown. The results for the total sensitivity index are shown in Fig. C2 in Appendix C.

time that the whole CMP take up, see Table 1). In a 10 year simulation this would allow for one additional week of simulation, which is negligible in comparison to the computing needs of e.g. increases in model resolution.

Within the CMP routine there are other physical processes that take up time, but also the calculation of diagnostics and preparatory calculations contribute. Of course, if numerous CMP processes and interactions with aerosols were simplified, this would allow for more drastic steps such as fewer prognostic aerosol variables as those could become redundant. Subsequently, significant reductions in model cost could be achieved. Yet by itself, the isolated removal or simplification of CMP processes provides small leverage for a decrease in computing time. However, as detailed in Sec. 1, there are numerous benefits in simplification that are independent of the associated computing cost, such as a gain in compactness and interpretability.

Table 1. Share of the computing time taken up by the cloud microphysical processes investigated here. In turn, the CMP computing time represents 4.7% of total computing time (excluding diagnostics; all values averaged over a 12-months control simulation with $\eta_i = 1 \forall i$). The time in the subroutine cold precipitation formation that is not attributed to the four processes is used for common initialisations and subsequent processing.

Process	Share of CMP routine cost (%)
Riming	1.8
Autoconversion	0.62
Accretion	0.46
Self-collection	0.046
Subroutine cold precipitation formation	4.8

4 Summary, conclusions and outlook

This study conducted a sensitivity analysis with an emulated PPE to illuminate the impact of selected CMP processes on model output. Different from previous studies (e.g. Wellmann et al. (2020); Hawker et al. (2021b)) we perturb the four CMP processes of autoconversion, riming, accretion and self-collection of ice as a whole. This is achieved by multiplying their process rates with a factor between 0.5 and 2. The resulting response surface of model output and its deviation from results with the default setup serves as a proxy for how accurately a process needs to be represented.

Perturbing only one process at a time reveals that ice crystal autoconversion acts as a threshold process: perturbing it causes the model to deviate, but when it is turned off the deviation is immense. This is because it is the only process that converts ice crystals to snow and as such accretion and riming depend on it. Using only roughly twice as many simulations than in the one-at-a-time perturbations to generate a PPE, we can generate the whole response surface using Gaussian process emulation. A sensitivity analysis of global and seasonal annual means reveals that for cloud cover, ice water path and number concentration as well as shortwave and longwave radiative effect, the perturbation of autoconversion has the most dominant impact by far. Accretion and riming assume a secondary role. As riming is the only investigated process that directly affects the liquid phase, riming has a dominant effect on the liquid water path and cloud droplet number concentration. Self-collection of ice has a negligible impact on the investigated global annual mean variables. Resolving smaller horizontal scales using a spherical harmonics expansion of the output variables corroborates the results of the global annual mean analysis, as does a seasonal analysis. These results as well as the shape of the response surface suggest that the parameterisation of self-collection and accretion can be readily and drastically simplified. While autoconversion and riming have a large impact on the model output considering the whole investigated phase space, the shallow slope of the response surface around the default $\eta_i = 1$ hints that slight modifications of their representations may leave the model output unchanged. The strength of the PPE approach is that interactions are already taken into account, meaning that all four processes could be simplified at the same time. If one wants

to develop the CMP scheme further, autoconversion is the process to scrutinize as it has the largest leverage in the model and therefore most urgent need to be represented correctly.

515   As we find that the processes themselves use a negligible fraction of the overall model computing time, simplifications are proposed as a means to make the model more interpretable, not cheaper (see Sec. 1 and 3.6). Our analysis shows that the representation of the four investigated microphysical processes leaves room for simplification. However, in deciding how drastic these simplifications should be, process uncertainty should also be considered. At the least, when new parameterisations are included in climate models we should question their implementation also regarding the complexity they add, looking for

520   their consistency, interpretability, simplicity and comprehensiveness (Mülmenstädt and Feingold, 2018; Touzé-Pfeiffer et al., 2021). Of course, more drastic simplifications than process reformulations would provide more leverage on interpretability and computing cost. For example, CMP schemes that contain only one category for ice, e.g. the Predicted Particle Properties (P3) ice microphysics scheme (e.g. Morrison and Milbrandt (2015); Eidhammer et al. (2017); Dietlicher et al. (2018, 2019); Tully et al. (2021)) are more physical as well as more interpretable. From this perspective it might seem troubling that in the current

525   CMP scheme the autoconversion process, which is a transfer mechanism between the two artificial classes, is so dominant in its importance. However, while the categories are artificial, the process itself is not: accretion of ice crystals forming larger ice crystals would be the equivalent process with only one ice category. Still, autoconversion is difficult to constrain in observations (Morrison et al., 2020) also because it is not a distinct physical process, and so moving towards a scheme with evolving instead of pre-defined ice categories seems advisable (see e.g. Milbrandt and Morrison (2016); Jensen et al. (2017)).

530   This study introduces the methodological framework to study the sensitivity of a climate model to the representation of CMP processes. To complete it, the analysis needs to be expanded to include other CMP processes in the model: For cold CMP ice formation, regional modeling studies have demonstrated cloud susceptibility to the choice of the ice nucleation parameterisation (Levkov et al., 1995; Hawker et al., 2021b), whereas in ECHAM-HAM heterogeneous immersion freezing in mixed-phase clouds has been shown to be rather inefficient (Villanueva et al., 2021). More generally the heterogeneous ice

535   formation pathway in mixed-phase clouds is small in ECHAM-HAM (Dietlicher et al., 2019; Bacer et al., 2021), hinting at simplification potential. In a sensitivity study of CMP parameters, Tan and Storelvmo (2016) found that the time scale of the Wegener-Bergeron-Findeisen process explains a large variance in supercooled cloud fractions, suggesting that as a whole it may be a dominating process as well. Secondary ice formation (Korolev and Leisner, 2020) may interact with the ice crystal source processes, allowing for interactive sensitivities (Hawker et al., 2021b), and should therefore be included, even though

540   only the Hallet-Mossop process is optionally included in ECHAM-HAM (Neubauer et al., 2019). Moreover, for a complete CMP process investigation, of course the warm rain processes need to be included as well (Wood et al., 2009; Gettelman et al., 2013).

One might argue that our analysis neglects the influence of other factors external to the CMPs on our conclusions. However, as our simulations span the whole globe and a whole year, they cover a range of dynamical situations and the results are

545   therefore robust in the current climate. Whether the conclusions hold e.g. in a future changed climate will have to be evaluated in a future study. It is important to stress that while we propose that simplifications to the CMP representation are possible, care needs to be taken to leave them physically based to ensure that the model can correctly represent differing climates. We

emphasize that our findings are conditional on the design of the ECHAM-HAM model, including the implementation of other processes and parameters that were not varied in the current study. Another factor that has not been investigated here is the model resolution that may affect the CMP behaviour in the model and thereby our conclusions on single processes' importance (Santos et al., 2021). The implementation and design choices of the CMP scheme in ECHAM-HAM may also influence the results, e.g. in the order of processes that are called, the separation between ice and snow, as well as the employed tuning strategy. Thus the results as such are only applicable to this CMP scheme and cannot be transferred to the significance of the investigated processes in other schemes let alone in reality.

Nevertheless, learning about the representation of CMP processes in ECHAM-HAM and how sensitive the model is to their representation helps us to interpret and improve the model, especially when comparing the results to experimental studies. To this end, it will also be fruitful to compare our findings to sensitivities in other models using different CMP schemes.

Code and data availability. The ECHAM-HAMMOZ model is freely available to the scientific community under the HAMMOZ Software License Agreement, which defines the conditions under which the model can be used. The specific version of the code used for this study is archived in the ECHAM-HAMMOZ SVN repository at /root/echam6-hammoz/tags/papers/2022/Proske_et_al_2022_ACP. More information can be found on the HAMMOZ website (https://redmine.hammoz.ethz.ch/projects/hammoz, last access: 17 September 2021). Analysis and plotting scripts are archived at https://doi.org/10.5281/zenodo.5506588 (Proske et al., 2021a). Generated data is archived at https://doi.org/10.5281/zenodo.5506533 (Proske et al., 2021b). The PyDOE library (tisimst, 2021) was used for Latin Hypercube Sampling, ESEm (Watson-Parris et al., 2021b, a) for the construction of the emulator, SALib (Usher et al., 2020) for the sensitivity analysis, and PySphereX (Staab, 2021) for the construction of the spherical harmonics expansion.

Appendix A: Tuning

Table A1. Tuning parameters that differ between this study and the reference of Neubauer et al. (2019). γ_r is the scaling factor for the stratiform rain formation rate by autoconversion. γ_s is a scaling factor for the stratiform snow formation rate by autoconversion. With the changes described in Sec. 2.1 the tuning parameter of the maximum cloud droplet radius, r_{CDNC}, replaces the previous minimum cloud droplet number concentration, $\mathrm{CDNC}_{\min}$. The tuning parameter for immediate autoconversion of detrained ICNC, γ_d, is newly introduced.

Parameter	ECHAM-HAM this study	reference
γ_r	5	10.6
γ_s	600	900
r_{CDNC}	$15 \cdot 10^{-6}\,\mathrm{m}$	–
$\mathrm{CDNC}_{\min}$	–	$40 \cdot 10^{-6}\,\mathrm{m}^{-3}$
γ_d	5	–

Appendix B: PPE results for more variables
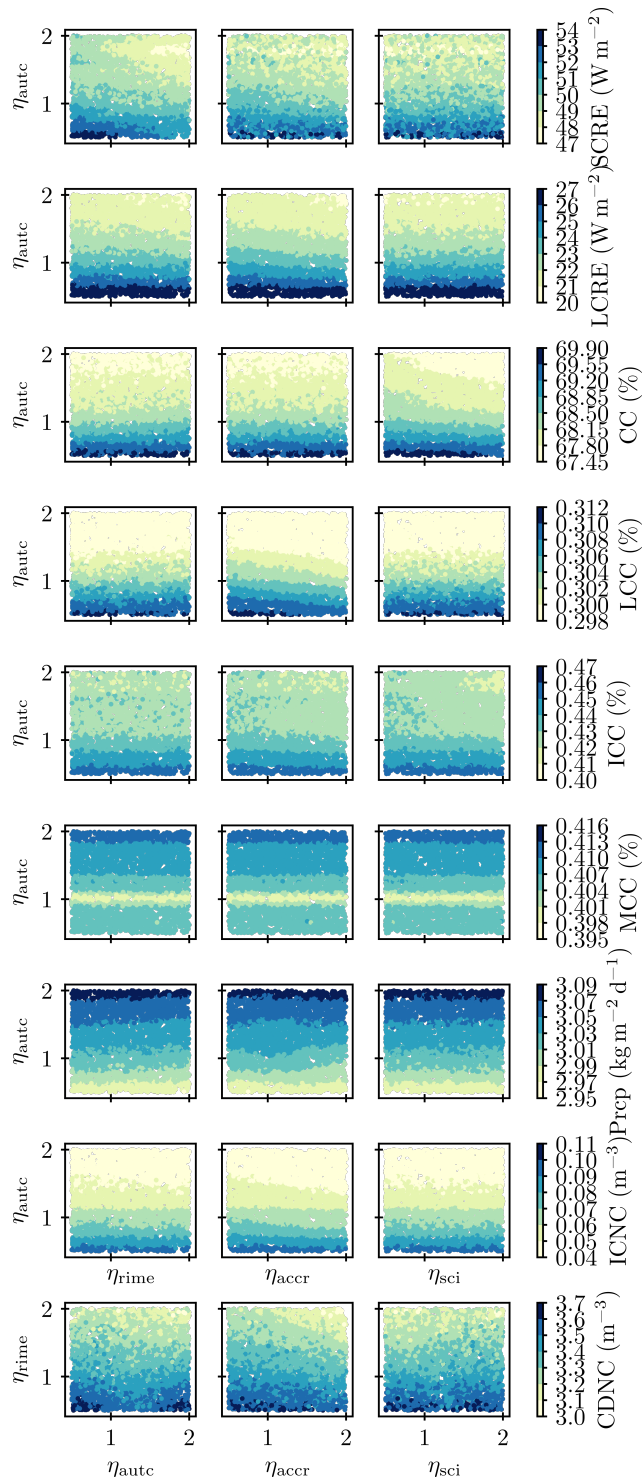
Figure B1. Visualisation of the multi-dimensional response surfaces of the emulated PPEs for multiple variables. Each process is a dimension, and the colorbars denote the global annual mean values. In principle, each surface could be displayed by a full matrix plot as in Fig. 7 and 8, but here only the panels that include the dominating process are shown (autoconversion, except for CDNC in the last row, where riming is the dominant process).

Appendix C: Total sensitivity index
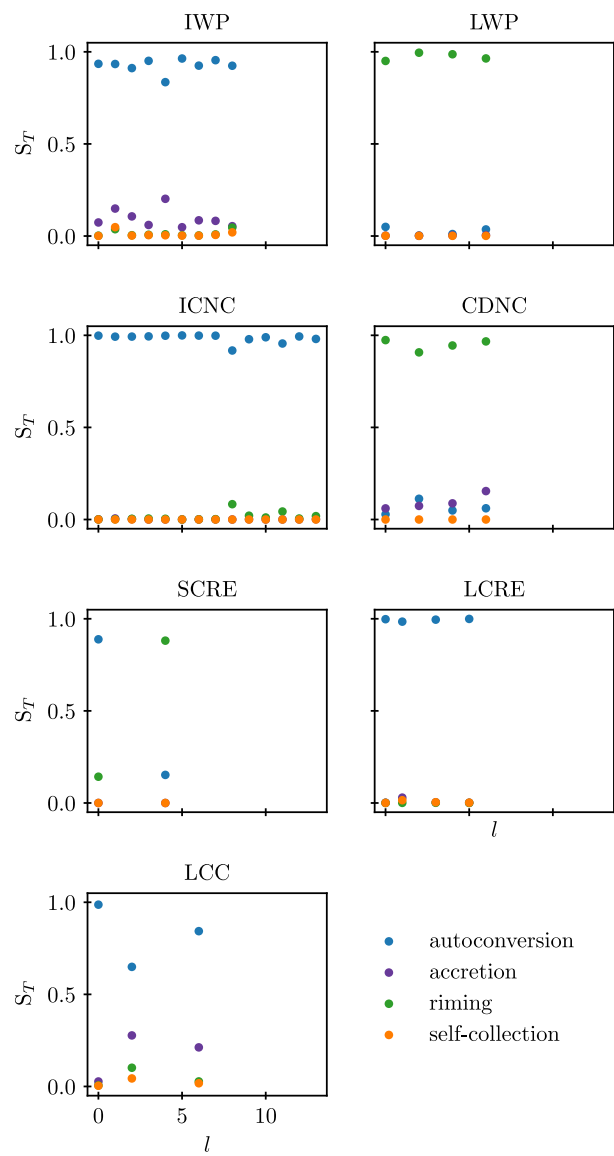
Figure C1. Same as Fig. 11 but for the total sensitivity indices.
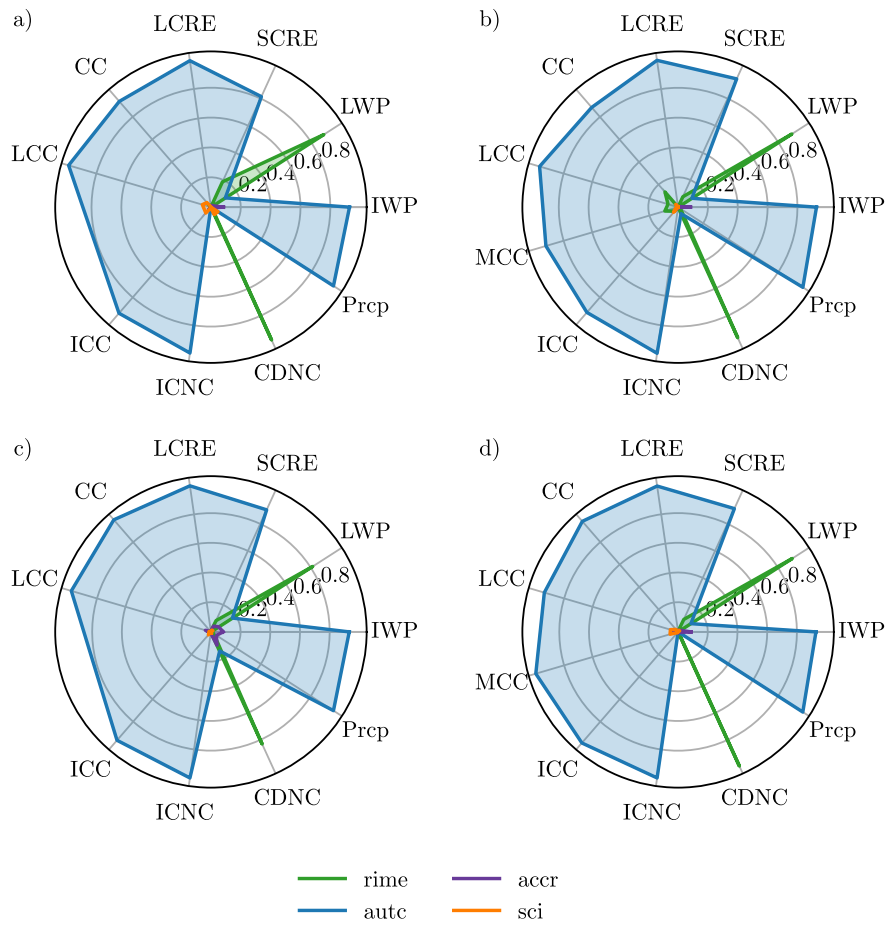
Figure C2. Same as Fig. 12 but for the total sensitivity indices.

Appendix D: Validation of the spherical harmonics sensitivity analysis

570   The validation of the spherical harmonics emulation was carried out as described in Sec. 2.4. Larger uncertainties in the emulation were apparent for almost all variables and degrees l (see Fig. D1 for an example) than for that of the global mean values. However, some emulations were also found to be defaulting, meaning that they predicted a similar output value for the whole phase space (see Fig. D2 and D3 for an example). As this behaviour points to a missing signal in the input, these points were excluded from further analysis, if the following two criteria were not fulfilled (excluding the emulated outliers that are

575   marked red e.g. in Fig. 4):

- The uncertainty in the prediction is smaller than the spread of the variable, i.e. the smallest error bar in Fig. D1d) is smaller than $0.9\Delta Y_{\mathrm{sim}}$.

- The predictions are significantly different from each other, i.e. there is one pair of predictions whose error bars do not overlap.
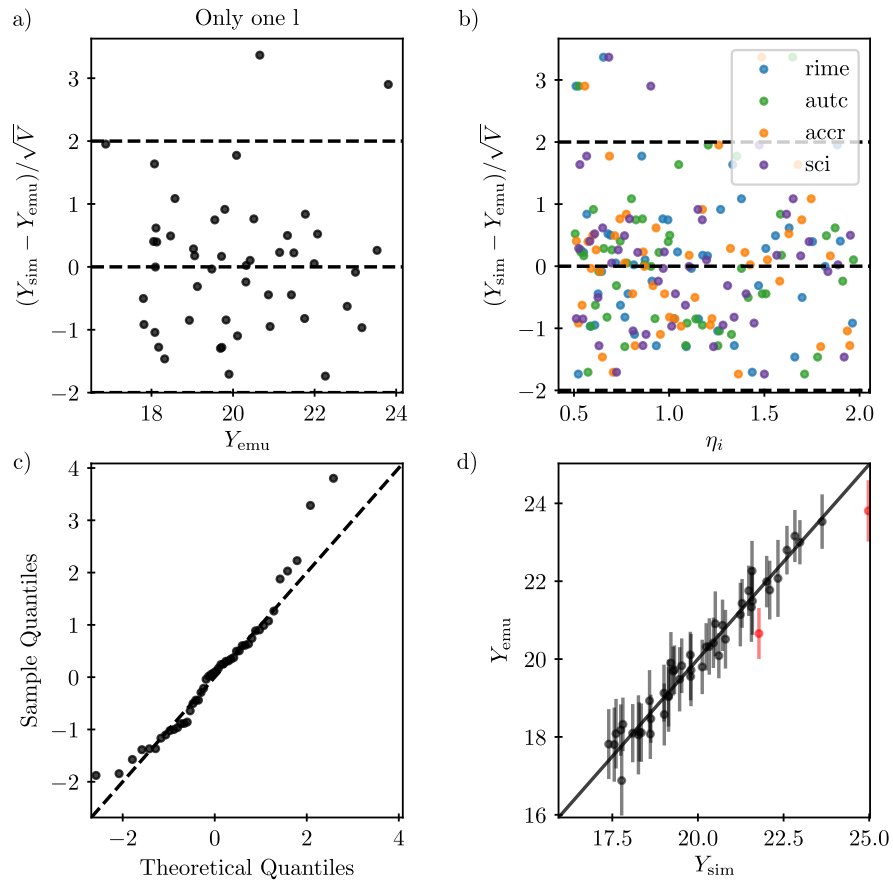
Figure D1. Validation of the emulated angular amplitude spectrum of degree $l = 6$ for the LWP.
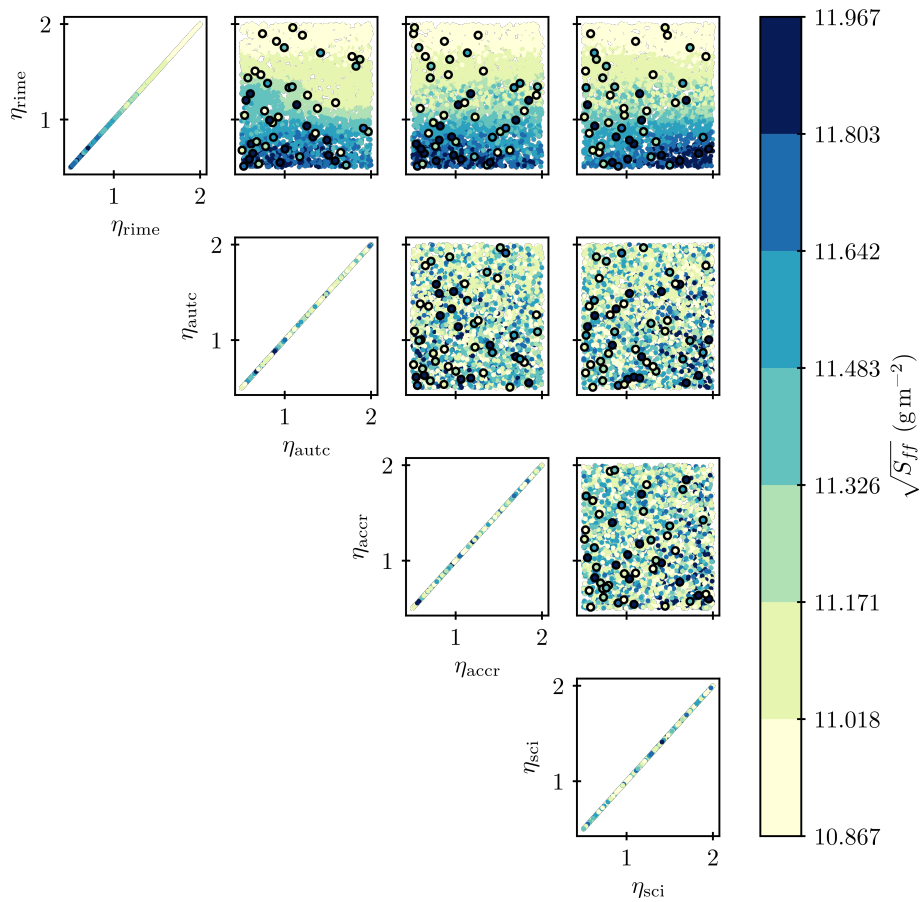
Figure D2. Same as Fig. 7 but for the LWP spherical expansion angular amplitude spectrum of degree $l = 8$. In this case, the emulator was found to be defaulting and therefore failed the validation and was not included in the subsequent sensitivity analysis. The points enclosed by black circles denote the PPE member results used to train the emulator.
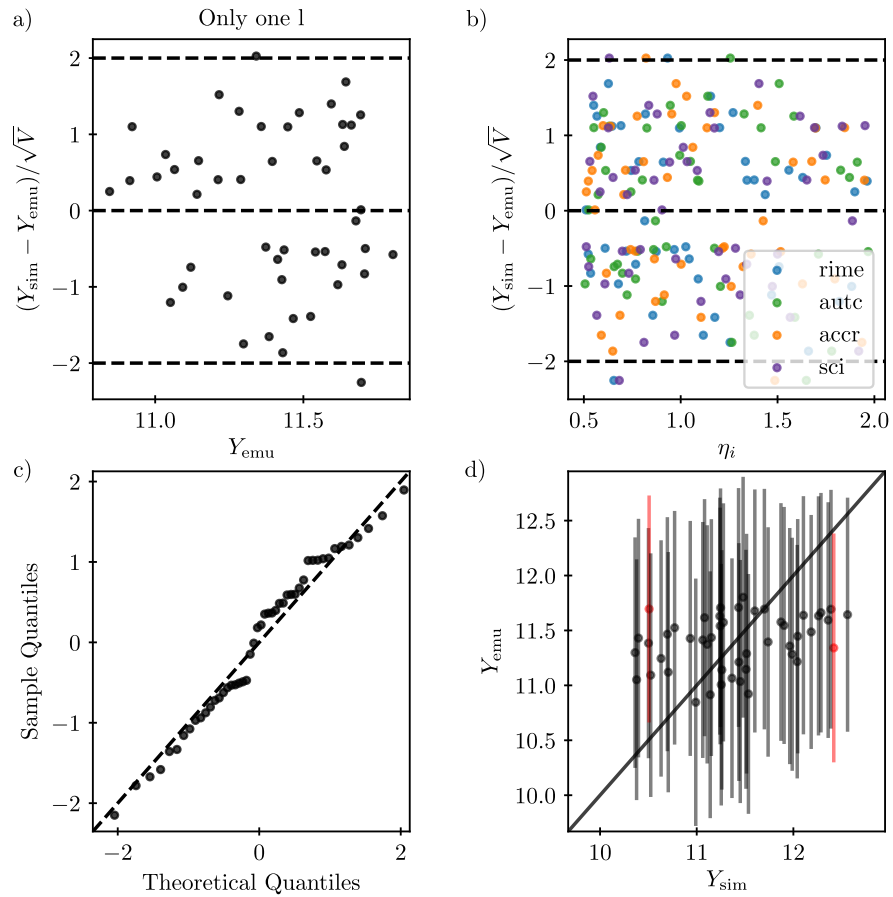
Figure D3. Validation of the emulated angular amplitude spectrum of degree $l = 8$ for the LWP (see Fig. D2), which failed because of diagnosed defaulting.

580 *Author contributions.* UP developed the model code, ran the simulations and emulation, analysed the data and wrote the manuscript. UP, SF, DN and UL developed the study idea and design and analysed the results. MS and UP developed the idea for the spherical harmonics analysis, for which MS developed the code. SF, DN, UL and MS edited the manuscript.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. The authors thank Duncan Watson-Parris for his advice on using ESEm and helpful discussions. They are grateful to
585 Rachel Hawker and Leighton Regayre for their advice on the emulation. The authors would like to thank the two reviewers for their careful and constructive feedback, which has improved this work substantially.

Throughout this study, the programming languages CDO (Schulzweida, 2018) and Python (Python Software Foundation, www.python.org) were used to handle data and analyse it. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 821205 (FORCeS).

590 References

Adams, G. S.: People Systematically Overlook Subtractive Changes, Nature, 592, 17, https://doi.org/10.1038/s41586-021-03380-y, 2021.

Archer-Nicholls, S., Abraham, N. L., Shin, Y. M., Weber, J., Russo, M. R., and ...: The Common Representative Intermediates Mechanism Version 2 in the United Kingdom Chemistry and Aerosols Model, Journal of Advances in Modeling Earth Systems, 13, 50, 2021.

Bacer, S., Sullivan, S. C., Sourdeval, O., Tost, H., Lelieveld, J., and Pozzer, A.: Cold Cloud Microphysical Process Rates in a Global Chemistry–Climate Model, Atmospheric Chemistry and Physics, 21, 1485–1505, https://doi.org/10.5194/acp-21-1485-2021, 2021.

Bastos, L. S. and O'Hagan, A.: Diagnostics for Gaussian Process Emulators, Technometrics, 51, 425–438, https://doi.org/10.1198/TECH.2009.08019, 2009.

Bergeron, T.: On the Physics of Clouds and Precipitation, Proces Verbaux de l'Association de Météorologie, pp. 156–178, 1935.

Bernus, A., Ottlé, C., and Raoult, N.: Variance Based Sensitivity Analysis of FLake Lake Model for Global Land Surface Modeling, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2019JD031928, 2021.

Beusch, L., Gudmundsson, L., and Seneviratne, S. I.: Emulating Earth System Model Temperatures with MESMER: From Global Mean Temperature Trajectories to Grid-Point-Level Realizations on Land, Earth System Dynamics, 11, 139–159, https://doi.org/10.5194/esd-11-139-2020, 2020.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y.: Clouds and Aerosols., in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Carslaw, K., Lee, L., Regayre, L., and Johnson, J.: Climate Models Are Uncertain, but We Can Do Something About It, Eos, 99, https://doi.org/10.1029/2018EO093757, 2018.

Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., Mann, G. W., Spracklen, D. V., Woodhouse, M. T., Regayre, L. A., and Pierce, J. R.: Large Contribution of Natural Aerosols to Uncertainty in Indirect Forcing, Nature, 503, 67–71, https://doi.org/10.1038/nature12674, 2013.

Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Del Genio, A. D., Déqué, M., Dymnikov, V., Galin, V., Gates, W. L., Ghan, S. J., Kiehl, J. T., Lacis, A. A., Le Treut, H., Li, Z.-X., Liang, X.-Z., McAvaney, B. J., Meleshko, V. P., Mitchell, J. F. B., Morcrette, J.-J., Randall, D. A., Rikus, L., Roeckner, E., Royer, J. F., Schlese, U., Sheinin, D. A., Slingo, A., Sokolov, A. P., Taylor, K. E., Washington, W. M., Wetherald, R. T., Yagai, I., and Zhang, M.-H.: Intercomparison and Interpretation of Climate Feedback Processes in 19 Atmospheric General Circulation Models, Journal of Geophysical Research, 95, 16 601–16 615, https://doi.org/10.1029/JD095iD10p16601, 1990.

Collins, M., Booth, B. B. B., Bhaskaran, B., Harris, G. R., Murphy, J. M., Sexton, D. M. H., and Webb, M. J.: Climate Model Errors, Feedbacks and Forcings: A Comparison of Perturbed Physics and Multi-Model Ensembles, Climate Dynamics, 36, 1737–1766, https://doi.org/10.1007/s00382-010-0808-0, 2011.

Couvreux, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., Rio, C., Audouin, O., Salter, J., Bazile, E., Brient, F., Favot, F., Honnert, R., Lefebvre, M.-P., Madeleine, J.-B., Rodier, Q., and Xu, W.: Process-Based Climate Model Development Harnessing Machine Learning: I. A Calibration Tool for Parameterization Improvement, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2020MS002217, 2021.

625 Dagon, K., Sanderson, B. M., Fisher, R. A., and Lawrence, D. M.: A Machine Learning Approach to Emulation and Biophysical Parameter Estimation with the Community Land Model, Version 5, Advances in Statistical Climatology, Meteorology and Oceanography, 6, 223–244, https://doi.org/10.5194/ascmo-6-223-2020, 2020.

Dietlicher, R., Neubauer, D., and Lohmann, U.: Prognostic Parameterization of Cloud Ice with a Single Category in the Aerosol-Climate Model ECHAM(v6.3.0)-HAM(v2.3), Geoscientific Model Development, 11, 1557–1576, https://doi.org/10.5194/gmd-11-1557-2018,
630 2018.

Dietlicher, R., Neubauer, D., and Lohmann, U.: Elucidating Ice Formation Pathways in the Aerosol–Climate Model ECHAM6-HAM2, Atmospheric Chemistry and Physics, 19, 9061–9080, https://doi.org/10.5194/acp-19-9061-2019, 2019.

Duvenaud, D. K.: Automatic Model Construction with Gaussian Processes, Ph.D. thesis, University of Cambridge, 2014.

Eidhammer, T., Morrison, H., Mitchell, D., Gettelman, A., and Erfani, E.: Improvements in Global Climate Model Microphysics Using a
635 Consistent Representation of Ice Particle Properties, Journal of Climate, 30, 609–629, https://doi.org/10.1175/JCLI-D-16-0050.1, 2017.

Errico, R. M.: What Is an Adjoint Model?, Bulletin of the American Meteorological Society, 78, 16, 1997.

Findeisen, W.: Kolloid-Meteorologische Vorgänge Bei Neiderschlagsbildung, Meteorologische Zeitschrift, 55, 121–133, 1938.

Fisher, R. A. and Koven, C. D.: Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2018MS001453, 2020.

640 Friebel, F., Lobo, P., Neubauer, D., Lohmann, U., Drossaart van Dusseldorp, S., Mühlhofer, E., and Mensah, A. A.: Impact of Isolated Atmospheric Aging Processes on the Cloud Condensation Nuclei Activation of Soot Particles, Atmospheric Chemistry and Physics, 19, 15 545–15 567, https://doi.org/10.5194/acp-19-15545-2019, 2019.

Gasparini, B., Meyer, A., Neubauer, D., Münch, S., and Lohmann, U.: Cirrus Cloud Properties as Seen by the *CALIPSO* Satellite and ECHAM-HAM Global Climate Model, Journal of Climate, 31, 1983–2003, https://doi.org/10.1175/JCLI-D-16-0608.1, 2018.

645 Gettelman, A., Morrison, H., Terai, C. R., and Wood, R.: Microphysical Process Rates and Global Aerosol–Cloud Interactions, Atmos. Chem. Phys., p. 13, https://doi.org/10.5194/acp-13-9855-2013, 2013.

Ghan, S. J., Liu, X., Easter, R. C., Zaveri, R., Rasch, P. J., Yoon, J.-H., and Eaton, B.: Toward a Minimal Representation of Aerosols in Climate Models: Comparative Decomposition of Aerosol Direct, Semidirect, and Indirect Radiative Forcing, Journal of Climate, 25, 6461–6476, https://doi.org/10.1175/JCLI-D-11-00650.1, 2012.

650 Ghan, S. J., Smith, S. J., Wang, M., Zhang, K., Pringle, K., Carslaw, K., Pierce, J., Bauer, S., and Adams, P.: A Simple Model of Global Aerosol Indirect Effects, Journal of Geophysical Research: Atmospheres, 118, 6688–6707, https://doi.org/10.1002/jgrd.50567, 2013.

Glassmeier, F., Possner, A., Vogel, B., Vogel, H., and Lohmann, U.: A Comparison of Two Chemistry and Aerosol Schemes on the Regional Scale and the Resulting Impact on Radiative Properties and Liquid- and Ice-Phase Aerosol–Cloud Interactions, Atmospheric Chemistry and Physics, 17, 8651–8680, https://doi.org/10.5194/acp-17-8651-2017, 2017.

655 Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., and Feingold, G.: An Emulator Approach to Stratocumulus Susceptibility, Atmospheric Chemistry and Physics, 19, 10 191–10 203, https://doi.org/10.5194/acp-19-10191-2019, 2019.

Hawker, R. E., Miltenberger, A. K., Johnson, J. S., Wilkinson, J. M., Hill, A. A., Shipway, B. J., Field, P. R., Murray, B. J., and Carslaw, K. S.: Model Emulation to Understand the Joint Effects of Ice-Nucleating Particles and Secondary Ice Production on Deep Convective Anvil Cirrus, Atmospheric Chemistry and Physics, 21, 17 315–17 343, https://doi.org/10.5194/acp-21-17315-2021, 2021a.

660 Hawker, R. E., Miltenberger, A. K., Wilkinson, J. M., Hill, A. A., Shipway, B. J., Cui, Z., Cotton, R. J., Carslaw, K. S., Field, P. R., and Murray, B. J.: The Temperature Dependence of Ice-Nucleating Particle Concentrations Affects the Radiative Properties of Tropical Convective Cloud Systems, Atmospheric Chemistry and Physics, 21, 5439–5461, https://doi.org/10.5194/acp-21-5439-2021, 2021b.

He, F. and Posselt, D. J.: Impact of Parameterized Physical Processes on Simulated Tropical Cyclone Characteristics in the Community Atmosphere Model, Journal of Climate, 28, 9857–9872, https://doi.org/10.1175/JCLI-D-15-0255.1, 2015.

665 Herman, J. and Usher, W.: SALib: An Open-Source Python Library for Sensitivity Analysis, The Journal of Open Source Software, 2, 97, https://doi.org/10.21105/joss.00097, 2017.

Holden, P. B., Edwards, N. R., Garthwaite, P. H., and Wilkinson, R. D.: Emulation and Interpretation of High-Dimensional Climate Model Outputs, Journal of Applied Statistics, 42, 2038–2055, https://doi.org/10.1080/02664763.2015.1016412, 2015.

Homma, T. and Saltelli, A.: Importance Measures in Global Sensitivity Analysis of Nonlinear Models, Reliability Engineering & System 670 Safety, 52, 1–17, https://doi.org/10.1016/0951-8320(96)00002-6, 1996.

Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranque, N., Musat, I., Fairhead, L., Diallo, F. B., and Volodina, V.: Process-based Climate Model Development Harnessing Machine Learning: II. Model Calibration from Single Column to Global, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2020MS002225, 2020.

Jensen, A. A., Harrington, J. Y., Morrison, H., and Milbrandt, J. A.: Predicting Ice Shape Evolution in a Bulk Microphysics Model, Journal 675 of the Atmospheric Sciences, 74, 2081–2104, https://doi.org/10.1175/JAS-D-16-0350.1, 2017.

Johnson, J. S., Cui, Z., Lee, L. A., Gosling, J. P., Blyth, A. M., and Carslaw, K. S.: Evaluating Uncertainty in Convective Cloud Microphysics Using Statistical Emulation, Journal of Advances in Modeling Earth Systems, 7, 162–187, https://doi.org/10.1002/2014MS000383, 2015.

Johnson, J. S., Regayre, L. A., Yoshioka, M., Pringle, K. J., Lee, L. A., Sexton, D. M. H., Rostron, J. W., Booth, B. B. B., and Carslaw, K. S.: The Importance of Comprehensive Parameter Sampling and Multiple Observations for Robust Constraint of Aerosol Radiative Forcing, 680 Atmospheric Chemistry and Physics, 18, 13 031–13 053, https://doi.org/10.5194/acp-18-13031-2018, 2018.

Kärcher, B. and Lohmann, U.: A Parameterization of Cirrus Cloud Formation: Homogeneous Freezing of Supercooled Aerosols, Journal of Geophysical Research, 107, 4010, https://doi.org/10.1029/2001JD000470, 2002a.

Kärcher, B. and Lohmann, U.: A Parameterization of Cirrus Cloud Formation: Homogeneous Freezing Including Effects of Aerosol Size: CIRRUS PARAMETERIZATION, Journal of Geophysical Research: Atmospheres, 107, AAC 9–1–AAC 9–10, 685 https://doi.org/10.1029/2001JD001429, 2002b.

Knutti, R. and Sedláček, J.: Robustness and Uncertainties in the New CMIP5 Climate Model Projections, Nature Climate Change, 3, 369–373, https://doi.org/10.1038/nclimate1716, 2013.

Koren, I. and Feingold, G.: Aerosol–Cloud–Precipitation System as a Predator-Prey Problem, Proceedings of the National Academy of Sciences, 108, 7, https://doi.org/10.1073/pnas.1101777108, 2011.

690 Korolev, A. and Leisner, T.: Review of Experimental Studies of Secondary Ice Production, Atmos. Chem. Phys., p. 31, 2020.

Lee, L. A., Carslaw, K. S., Pringle, K. J., Mann, G. W., and Spracklen, D. V.: Emulation of a Complex Global Aerosol Model to Quantify Sensitivity to Uncertain Parameters, Atmospheric Chemistry and Physics, 11, 12 253–12 273, https://doi.org/10.5194/acp-11-12253-2011, 2011.

Lee, L. A., Carslaw, K. S., Pringle, K. J., and Mann, G. W.: Mapping the Uncertainty in Global CCN Using Emulation, Atmospheric 695 Chemistry and Physics, 12, 9739–9751, https://doi.org/10.5194/acp-12-9739-2012, 2012.

Lee, L. A., Pringle, K. J., Reddington, C. L., Mann, G. W., Stier, P., Spracklen, D. V., Pierce, J. R., and Carslaw, K. S.: The Magnitude and Causes of Uncertainty in Global Model Simulations of Cloud Condensation Nuclei, Atmospheric Chemistry and Physics, 13, 8879–8914, https://doi.org/10.5194/acp-13-8879-2013, 2013.

Lee, L. A., Reddington, C. L., and Carslaw, K. S.: On the Relationship between Aerosol Model Uncertainty and Radiative Forcing Uncer-700 tainty, Proceedings of the National Academy of Sciences, 113, 5820–5827, https://doi.org/10.1073/pnas.1507050113, 2016.

Levkov, L., Boin, M., and Rockel, B.: Impact of Primary Ice Nucleation Parameterizations on the Formation and Maintenance of Cirrus, Atmospheric Research, 38, 147–159, https://doi.org/10.1016/0169-8095(94)00091-Q, 1995.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and Mitchell, D.: Toward a Minimal Representation of Aerosols in Climate Models: Description and Evaluation in the Community Atmosphere Model CAM5, Geoscientific Model Development, 5, 709–739, https://doi.org/10.5194/gmd-5-709-2012, 2012.

Loeppky, J. L., Sacks, J., and Welch, W. J.: Choosing the Sample Size of a Computer Experiment: A Practical Guide, Technometrics, 51, 366–376, https://doi.org/10.1198/TECH.2009.08040, 2009.

Lohmann, U.: Possible Aerosol Effects on Ice Clouds via Contact Nucleation, JOURNAL OF THE ATMOSPHERIC SCIENCES, 59, 10, 2002.

Lohmann, U.: Impact of the Mount Pinatubo Eruption on Cirrus Clouds Formed by Homogeneous Freezing in the ECHAM4 GCM, Journal of Geophysical Research, 108, 4568, https://doi.org/10.1029/2002JD003185, 2003.

Lohmann, U. and Ferrachat, S.: Impact of Parametric Uncertainties on the Present-Day Climate and on the Anthropogenic Aerosol Effect, Atmospheric Chemistry and Physics, 10, 11 373–11 383, https://doi.org/10.5194/acp-10-11373-2010, 2010.

Lohmann, U. and Hoose, C.: Sensitivity Studies of Different Aerosol Indirect Effects in Mixed-Phase Clouds, Atmos. Chem. Phys., p. 18, 2009.

Lohmann, U. and Neubauer, D.: The Importance of Mixed-Phase and Ice Clouds for Climate Sensitivity in the Global Aerosol–Climate Model ECHAM6-HAM2, Atmospheric Chemistry and Physics, 18, 8807–8828, https://doi.org/10.5194/acp-18-8807-2018, 2018.

Lohmann, U. and Roeckner, E.: Design and Performance of a New Cloud Microphysics Scheme Developed for the ECHAM General Circulation Model, Climate Dynamics, 12, 16, 1996.

Lohmann, U., Feichter, J., Chuang, C. C., and Penner, J. E.: Prediction of the Number of Cloud Droplets in the ECHAM GCM, Journal of Geophysical Research: Atmospheres, 104, 9169–9198, https://doi.org/10.1029/1999JD900046, 1999.

Lohmann, U., Stier, P., Hoose, C., Ferrachat, S., Kloster, S., Roeckner, E., and Zhang, J.: Cloud Microphysics and Aerosol Indirect Effects in the Global Climate Model ECHAM5-HAM, Atmos. Chem. Phys., p. 22, 2007.

Lohmann, U., Friebel, F., Kanji, Z. A., Mahrt, F., Mensah, A. A., and Neubauer, D.: Future Warming Exacerbated by Aged-Soot Effect on Cloud Formation, Nature Geoscience, 13, 674–680, https://doi.org/10.1038/s41561-020-0631-0, 2020.

Matus, A. V. and L'Ecuyer, T. S.: The Role of Cloud Phase in Earth's Radiation Budget, Journal of Geophysical Research: Atmospheres, 122, 2559–2578, https://doi.org/10.1002/2016JD025951, 2017.

McNeall, D., Williams, J., Booth, B., Betts, R., Challenor, P., Wiltshire, A., and Sexton, D.: The Impact of Structural Error on Parameter Constraint in a Climate Model, Earth System Dynamics, 7, 917–935, https://doi.org/10.5194/esd-7-917-2016, 2016.

Milbrandt, J. A. and Morrison, H.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part III: Introduction of Multiple Free Categories, Journal of the Atmospheric Sciences, 73, 975–995, https://doi.org/10.1175/JAS-D-15-0204.1, 2016.

Morales, A., Posselt, D. J., and Morrison, H.: Which Combinations of Environmental Conditions and Microphysical Parameter Values Produce a Given Orographic Precipitation Distribution?, Journal of the Atmospheric Sciences, 78, 619–638, https://doi.org/10.1175/JAS-D-20-0142.1, 2021.

Morris, M. D. and Mitchell, T. J.: Exploratory Designs for Computational Experiments, Journal of Statistical Planning and Inference, 43, 22, 1995.

Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, Journal of the Atmospheric Sciences, 72, 287–311, https://doi.org/10.1175/JAS-D-14-0065.1, 2015.

Morrison, H., Lier-Walqui, M., Fridlind, A. M., Grabowski, W. W., Harrington, J. Y., Hoose, C., Korolev, A., Kumjian, M. R., Milbrandt, J. A., Pawlowska, H., Posselt, D. J., Prat, O. P., Reimel, K. J., Shima, S.-I., Diedenhoven, B., and Xue, L.: Confronting the Challenge of Modeling Cloud and Precipitation Microphysics, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001689, 2020.

Muench, S. and Lohmann, U.: Developing a Cloud Scheme With Prognostic Cloud Fraction and Two Moment Microphysics for ECHAM-HAM, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001824, 2020.

Mulholland, D. P., Haines, K., Sparrow, S. N., and Wallom, D.: Climate Model Forecast Biases Assessed with a Perturbed Physics Ensemble, Climate Dynamics, 49, 1729–1746, https://doi.org/10.1007/s00382-016-3407-x, 2017.

Mülmenstädt, J. and Feingold, G.: The Radiative Forcing of Aerosol–Cloud Interactions in Liquid Clouds: Wrestling and Embracing Uncertainty, Current Climate Change Reports, 4, 23–40, https://doi.org/10.1007/s40641-018-0089-y, 2018.

Mülmenstädt, J., Sourdeval, O., Delanoë, J., and Quaas, J.: Frequency of Occurrence of Rain from Liquid-, Mixed-, and Ice-Phase Clouds Derived from A-Train Satellite Retrievals: RAIN FROM LIQUID- AND ICE-PHASE CLOUDS, Geophysical Research Letters, 42, 6502–6509, https://doi.org/10.1002/2015GL064604, 2015.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of Modelling Uncertainties in a Large Ensemble of Climate Change Simulations, Nature, 430, 768–772, https://doi.org/10.1038/nature02771, 2004.

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stier, P., Partridge, D. G., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., and Lohmann, U.: The Global Aerosol–Climate Model ECHAM6.3–HAM2.3 – Part 2: Cloud Evaluation, Aerosol Radiative Forcing, and Climate Sensitivity, Geoscientific Model Development, 12, 3609–3639, https://doi.org/10.5194/gmd-12-3609-2019, 2019.

Oakley, J. E. and O'Hagan, A.: Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66, 751–769, https://doi.org/10.1111/j.1467-9868.2004.05304.x, 2004.

O'Hagan, A.: Bayesian Analysis of Computer Code Outputs: A Tutorial, Reliability Engineering & System Safety, 91, 1290–1300, https://doi.org/10.1016/j.ress.2005.11.025, 2006.

Posselt, D. J.: A Bayesian Examination of Deep Convective Squall-Line Sensitivity to Changes in Cloud Microphysical Parameters, Journal of the Atmospheric Sciences, 73, 637–665, https://doi.org/10.1175/JAS-D-15-0159.1, 2016.

Proske, U., Ferrachat, S., Neubauer, D., Staab, M., and Lohmann, U.: Scripts for the Publication "Assessing the Potential for Simplification in Global Climate Model Cloud Microphysics", Zenodo, 2021a.

Proske, U., Ferrachat, S., Neubauer, David, Staab, Martin, and Lohmann, Ulrike: Data for the Publication "Assessing the Potential for Simplification in Global Climate Model Cloud Microphysics", 2021b.

Qian, Y., Jackson, C., Giorgi, F., Booth, B., Duan, Q., Forest, C., Higdon, D., Hou, Z. J., and Huerta, G.: Uncertainty Quantification in Climate Modeling and Projection, Bulletin of the American Meteorological Society, 97, 821–824, https://doi.org/10.1175/BAMS-D-15-00297.1, 2016.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning, MIT Press, Cambridge, Mass, 2006.

Reddington, C. L., Carslaw, K. S., Stier, P., Schutgens, N., Coe, H., Liu, D., Allan, J., Browse, J., Pringle, K. J., Lee, L. A., Yoshioka, M., Johnson, J. S., Regayre, L. A., Spracklen, D. V., Mann, G. W., Clarke, A., Hermann, M., Henning, S., Wex, H., Kristensen, T. B., Leaitch,

W. R., Pöschl, U., Rose, D., Andreae, M. O., Schmale, J., Kondo, Y., Oshima, N., Schwarz, J. P., Nenes, A., Anderson, B., Roberts, G. C., Snider, J. R., Leck, C., Quinn, P. K., Chi, X., Ding, A., Jimenez, J. L., and Zhang, Q.: The Global Aerosol Synthesis and Science Project (GASSP): Measurements and Modeling to Reduce Uncertainty, Bulletin of the American Meteorological Society, 98, 1857–1877, https://doi.org/10.1175/BAMS-D-15-00317.1, 2017.

Regayre, L. A., Pringle, K. J., Booth, B. B. B., Lee, L. A., Mann, G. W., Browse, J., Woodhouse, M. T., Rap, A., Reddington, C. L., and Carslaw, K. S.: Uncertainty in the Magnitude of Aerosol-Cloud Radiative Forcing over Recent Decades, Geophysical Research Letters, 41, 9040–9049, https://doi.org/10.1002/2014GL062029, 2014.

Regayre, L. A., Pringle, K. J., Lee, L. A., Rap, A., Browse, J., Mann, G. W., Reddington, C. L., Carslaw, K. S., Booth, B. B. B., and Woodhouse, M. T.: The Climatic Importance of Uncertainties in Regional Aerosol–Cloud Radiative Forcings over Recent Decades, Journal of Climate, 28, 6589–6607, https://doi.org/10.1175/JCLI-D-15-0127.1, 2015.

Regayre, L. A., Johnson, J. S., Yoshioka, M., Pringle, K. J., Sexton, D. M. H., Booth, B. B. B., Lee, L. A., Bellouin, N., and Carslaw, K. S.: Aerosol and Physical Atmosphere Model Parameters Are Both Important Sources of Uncertainty in Aerosol ERF, Atmospheric Chemistry and Physics, 18, 9975–10 006, https://doi.org/10.5194/acp-18-9975-2018, 2018.

Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D.: Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments, Journal of Climate, 22, 3540–3557, https://doi.org/10.1175/2008JCLI2533.1, 2009.

Rudin, C.: Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead, Nature Machine Intelligence, 1, 206–215, https://doi.org/10.1038/s42256-019-0048-x, 2019.

Ryan, E., Wild, O., Voulgarakis, A., and Lee, L.: Fast Sensitivity Analysis Methods for Computationally Expensive Models with Multi-Dimensional Output, Geoscientific Model Development, 11, 3131–3146, https://doi.org/10.5194/gmd-11-3131-2018, 2018.

Saltelli, A., ed.: Global Sensitivity Analysis: The Primer, John Wiley, Chichester, England ; Hoboken, NJ, 2008.

Saltelli, A., Tarantola, S., and Chan, K. P.-S.: A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output, Technometrics, 41, 39–56, https://doi.org/10.1080/00401706.1999.10485594, 1999.

Santos, S. P., Caldwell, P. M., and Bretherton, C. S.: Cloud Process Coupling and Time Integration in the E3SM Atmosphere Model, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2020MS002359, 2021.

Schulzweida, U.: CDO User Guide, MPI for Meteorology, 2018.

Schutgens, N. A. J. and Stier, P.: A Pathway Analysis of Global Aerosol Processes, Atmospheric Chemistry and Physics, 14, 11 657–11 686, https://doi.org/10.5194/acp-14-11657-2014, 2014.

Seifert, A. and Rasp, S.: Potential and Limitations of Machine Learning for Modeling Warm-Rain Cloud Microphysical Processes, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2020MS002301, 2020.

Sengupta, K., Pringle, K., Johnson, J. S., Reddington, C., Browse, J., Scott, C. E., and Carslaw, K.: A Global Model Perturbed Parameter Ensemble Study of Secondary Organic Aerosol Formation, Atmospheric Chemistry and Physics, 21, 2693–2723, https://doi.org/10.5194/acp-21-2693-2021, 2021.

Soden, B. J. and Held, I. M.: An Assessment of Climate Feedbacks in Coupled Ocean-Atmosphere Models, Journal of Climate, 19, 3354–3360, https://doi.org/10.1175/JCLI3799.1, 2006.

Sotiropoulou, G., Vignon, É., Young, G., Morrison, H., O'Shea, S. J., Lachlan-Cope, T., Berne, A., and Nenes, A.: Secondary Ice Production in Summer Clouds over the Antarctic Coast: An Underappreciated Process in Atmospheric Models, Atmospheric Chemistry and Physics, 21, 755–771, https://doi.org/10.5194/acp-21-755-2021, 2021.

Staab, M.: PySphereX, Zenodo, https://doi.org/10.5281/ZENODO.5520635, 2021.

Storelvmo, T., Kristjánsson, J. E., Lohmann, U., Iversen, T., Kirkevåg, A., and Seland, Ø.: Modeling of the Wegener–Bergeron–Findeisen Process—Implications for Aerosol Indirect Effects, Environmental Research Letters, 3, 045 001, https://doi.org/10.1088/1748-9326/3/4/045001, 2008.

Sun, Z. and Shine, K. P.: Parameterization of Ice Cloud Radiative Properties and Its Application to the Potential Climatic Importance of Mixed-Phase Clouds, Journal of Climate, 8, 1874–1888, https://doi.org/10.1175/1520-0442(1995)008<1874:POICRP>2.0.CO;2, 1995.

Tan, I. and Storelvmo, T.: Sensitivity Study on the Influence of Cloud Microphysical Parameters on Mixed-Phase Cloud Thermodynamic Phase Partitioning in CAM5, Journal of the Atmospheric Sciences, 73, 709–728, https://doi.org/10.1175/JAS-D-15-0152.1, 2016.

Tan, I., Storelvmo, T., and Zelinka, M. D.: Observational Constraints on Mixed-Phase Clouds Imply Higher Climate Sensitivity, Science, 352, 224–227, https://doi.org/10.1126/science.aad5300, 2016.

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., Rast, S., Kokkola, H., Schultz, M., Schroeder, S., Daskalakis, N., Barthel, S., Heinold, B., and Lohmann, U.: The Global Aerosol–Climate Model ECHAM6.3–HAM2.3 – Part 1: Aerosol Evaluation, Geoscientific Model Development, 12, 1643–1677, https://doi.org/10.5194/gmd-12-1643-2019, 2019.

Tett, S. F. B., Rowlands, D. J., Mineter, M. J., and Cartis, C.: Can Top-of-Atmosphere Radiation Measurements Constrain Climate Predictions? Part II: Climate Sensitivity, Journal of Climate, 26, 9367–9383, https://doi.org/10.1175/JCLI-D-12-00596.1, 2013.

tisimst: PyDOE: The Experimental Design Package for Python, 2021.

Touzé-Pfeiffer, L., Frédéric Hourdin, and Catherine Rio: Parameterization and Tuning of Cloud and Precipitation Overlap in LMDz, in: Improvement and Calibration of Clouds in Models, Conference Presentation, Toulouse, France, 2021.

Tully, C., Neubauer, D., Omanovic, N., and Lohmann, U.: Cirrus Cloud Thinning Using a More Physically-Based Ice Microphysics Scheme in the ECHAM-HAM GCM, Preprint, Clouds and Precipitation/Atmospheric Modelling/Troposphere/Physics (physical properties and processes), https://doi.org/10.5194/acp-2021-685, 2021.

Usher, W., Herman, J., Iwanaga, T., Teixeira, L., CELLIER, N., Whealton, C., Hadka, D., Xantares, Bernardoct, Rios, F., Mutel, C., Cederstrand, E., TobiasKAndersen, Engelen, J. V., ANtlord, Kranas, H., Dixit, V. K., and Bsteubing: SALib/SALib: Public Beta, Zenodo, https://doi.org/10.5281/ZENODO.598306, 2020.

van Lier-Walqui, M., Vukicevic, T., and Posselt, D. J.: Linearization of Microphysical Parameterization Uncertainty Using Multiplicative Process Perturbation Parameters, Monthly Weather Review, 142, 401–413, https://doi.org/10.1175/MWR-D-13-00076.1, 2014.

van Lier-Walqui, M., Morrison, H., Kumjian, M. R., Reimel, K. J., Prat, O. P., Lunderman, S., and Morzfeld, M.: A Bayesian Approach for Statistical–Physical Bulk Parameterization of Rain Microphysics. Part II: Idealized Markov Chain Monte Carlo Experiments, Journal of the Atmospheric Sciences, 77, 1043–1064, https://doi.org/10.1175/JAS-D-19-0071.1, 2019.

Villanueva, D., Neubauer, D., Gasparini, B., Ickes, L., and Tegen, I.: Constraining the Impact of Dust-Driven Droplet Freezing on Climate Using Cloud-Top-Phase Observations, Geophysical Research Letters, 48, https://doi.org/10.1029/2021GL092687, 2021.

Wacker, U.: Competition of Precipitation Particles in a Model with Parameterized Cloud Microphysics, Journal of Atmospheric Sciences, 52, 2577–89, 1995.

Watson-Parris, D., Williams, A., Deaconu, L., and Stier, P.: Model Calibration Using ESEm v1.0.0 – an Open, Scalable Earth System Emulator, Geoscientific Model Development Discussions, p. 24, https://doi.org/10.5194/gmd-2021-267, 2021a.

Watson-Parris, D., Williams, A., and Monticone, P.: Duncanwp/ESEm: V1.1.0, Zenodo, https://doi.org/10.5281/ZENODO.5196631, 2021b.

Wegener, A.: Thermodynamik Der Atmosphäre, J. A. Barth, Leipzig, 1911.

Wellmann, C., Barrett, A. I., Johnson, J. S., Kunz, M., Vogel, B., Carslaw, K. S., and Hoose, C.: Using Emulators to Understand the Sensitivity of Deep Convective Clouds and Hail to Environmental Conditions, Journal of Advances in Modeling Earth Systems, p. 2018MS001465, https://doi.org/10.1029/2018MS001465, 2018.

855 Wellmann, C., Barrett, A. I., Johnson, J. S., Kunz, M., Vogel, B., Carslaw, K. S., and Hoose, C.: Comparing the Impact of Environmental Conditions and Microphysics on the Forecast Uncertainty of Deep Convective Clouds and Hail, Atmospheric Chemistry and Physics, 20, 2201–2219, https://doi.org/10.5194/acp-20-2201-2020, 2020.

White, B., Gryspeerdt, E., Stier, P., Morrison, H., Thompson, G., and Kipling, Z.: Uncertainty from the Choice of Microphysics Scheme in Convection-Permitting Models Significantly Exceeds Aerosol Effects, Atmospheric Chemistry and Physics, 17, 12 145–12 175,
860 https://doi.org/10.5194/acp-17-12145-2017, 2017.

Williams, K. D. and Tselioudis, G.: GCM Intercomparison of Global Cloud Regimes: Present-Day Evaluation and Climate Change Response, Climate Dynamics, 29, 231–250, https://doi.org/10.1007/s00382-007-0232-2, 2007.

Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., and Yamazaki, K.: History Matching for Exploring and Reducing Climate Model Parameter Space Using Observations and a Large Perturbed Physics Ensemble, Climate Dynamics, 41, 1703–
865 1729, https://doi.org/10.1007/s00382-013-1896-4, 2013.

Williamson, D., Blaker, A. T., Hampton, C., and Salter, J.: Identifying and Removing Structural Biases in Climate Models with History Matching, Climate Dynamics, 45, 1299–1324, https://doi.org/10.1007/s00382-014-2378-z, 2015.

Wood, R., Kubar, T. L., and Hartmann, D. L.: Understanding the Importance of Microphysics and Macrophysics for Warm Rain in Marine Low Clouds. Part II: Heuristic Models of Rain Formation, Journal of the Atmospheric Sciences, 66, 2973–2990,
870 https://doi.org/10.1175/2009JAS3072.1, 2009.

Yan, H., Qian, Y., Zhao, C., Wang, H., Wang, M., Yang, B., Liu, X., and Fu, Q.: A New Approach to Modeling Aerosol Effects on East Asian Climate: Parametric Uncertainties Associated with Emissions, Cloud Microphysics, and Their Interactions, Journal of Geophysical Research: Atmospheres, 120, 8905–8924, https://doi.org/10.1002/2015JD023442, 2015.