Tunability of ice supersaturation in the upper troposphere and lower stratosphere in the LMDZ atmospheric model

Sidiki Sanogo[1], Olivier Boucher[1], Audran Borella[1], Frédéric Hourdin[3], Nicolas Bellouin[1,2], Étienne Vignon[3], Ionela Musat[3], Maëlle Coulon-Decorzens[3], Adriana Sima[3], Frédérique Cheruy[3]

[1]Institut Pierre–Simon Laplace, Sorbonne Université / CNRS, Paris, France
[2]Department of Meteorology, University of Reading, Reading, United Kingdom
[3]Laboratoire de Météorologie Dynamique, IPSL, Sorbonne Université / CNRS, Paris, France

Key Points:

- The tunability of ISS parameterization in LMDZ is investigated using a statistical tool based on history matching with iterative refocusing.
- The ISS parameterization can be tuned in LMDZ to document both parametric and structural errors.
- After tuning LMDZ better reproduces the frequency of occurrence of ISS regions at 250 hPa than at 200 hPa because of persistent biases in the tropopause height.

Corresponding author: Sidiki Sanogo, `Sidiki.sanogo@ipsl.fr`

Abstract

Ice supersaturation (ISS) is an atmospheric phenomenon that is necessary for the in situ formation of cirrus clouds and persistent contrails. Including ISS in climate models is essential for more realistic modelling of ice clouds, but this remains difficult because the mechanisms involved are not yet fully understood. In this study, we present a parameterisation of ISS in the LMDZ atmospheric component of the Institut Pierre-Simon Laplace climate model and address the question as to whether this parameterisation can be tuned to fit observational constraints on ISS while maintaining the realism of many other LMDZ features. To this end, temperature biases in the upper troposphere and lower stratosphere were corrected using a nudging approach prior to the tuning phase. Tuning was performed using a statistical tool based on history matching with iterative refocusing and Gaussian processes to adjust 16 free parameters of ISS and large-scale condensation and convection parameterisations in LMDZ. This method identified three LMDZ configurations that best match the observed frequency of occurrence of ISS regions (ISSRs) at 250 hPa in the mid-latitudes of the Northern Hemisphere while also meeting global (top-of-atmosphere radiative budget) and regional (cloud fraction and radiative effects) observational constraints. The main shortcoming of these model configurations is their underestimation of the frequency of occurrence of ISSRs at 200 hPa. This suggests that a realistic representation of tropopause dynamics is an important consideration when developing and choosing climate models to study the radiative forcing of contrails.

Plain Language Summary

Ice supersaturation (ISS) occurs in regions of the atmosphere that are cold and humid enough for water vapor to exceed saturation with respect to ice. These conditions are necessary for the formation of ice clouds, such as cirrus clouds and persistent aircraft condensation trails (contrails). Simulating ISS in climate models requires a dedicated parameterization that also interacts with other cloud-related processes. In this work, we present a parameterization of ISS developed for the LMDZ climate model. We also study whether the parameters of this parameterization can be adjusted to fit a number of observational constraints. The study shows that the tuned version of LMDZ is consistent with observations in the North Atlantic, but has remaining biases in the frequency of occurrence of ISSRs in the USA, Eurasia, and the mid- and high-latitude regions of the Southern Hemisphere.

1 Introduction

Ice supersaturation (ISS) refers to atmospheric situations whereby relative humidity with respect to ice (RHi) is greater than 100%. ISS is essentially a feature of the upper troposphere but can be occasionally observed in the lower stratosphere (Gierens et al., 1999; Petzold et al., 2020; Sanogo et al., 2024) and sometimes near the surface in Antarctica (Genthon et al., 2017; Vignon et al., 2022). The frequency of occurrence of ISS regions (ISSRs) varies according to the season and location (Spichtinger, Gierens, Leiterer, & Dier, 2003; Gierens & Brinkop, 2012; Lamquin et al., 2012; Wolf et al., 2023, 2025; Sanogo et al., 2024; Wolf et al., 2024). The highest frequency of ISSRs is observed in high-latitude regions at pressure levels between 400 and 300 hPa, and in the mid-latitude regions at pressure levels between 300 and 200 hPa (Lamquin et al., 2012). Using RHi measurements carried out every 4 seconds onboard commercial aircraft, Petzold et al. (2020) found an ISSR frequency of the order of 20–25% of the flight time in summer and of the order of 35–40% of the flight time in winter in the troposphere between 350 and 190 hPa in the Northern Hemisphere, over a region extending from North America to Europe.

The frequencies of ISSRs are much lower in the lower stratosphere (below 20% according to the estimation by Petzold et al. (2020) based on in situ measurements), due to the low ambiant relative humidity (Gierens et al., 1999; Petzold et al., 2020; Sanogo

et al., 2024). At 200 hPa, ISSRs are more frequent (6–15%) in tropical regions (Spichtinger, Gierens, & Read, 2003; Sanogo et al., 2024). In terms of size, more than 80% of ice-supersaturated path lengths measured along aircraft flight routes in the North Atlantic are shorter than 10 km (Reutter et al., 2020; Wolf et al., 2024). The occurrence of ISSRs is favored by particular meteorological conditions such as deep convection, mesoscale gravity waves (Spichtinger, Gierens, & Dörnbrack, 2005), the mean divergent flow at the tropopause (Gierens & Brinkop, 2012; Krämer et al., 2009; Wilhelm et al., 2022), or the moist air streams in synoptic disturbances (e.g., warm conveyor belt, cold and warm fronts) (Gierens & Brinkop, 2012; Wolf et al., 2023).

The occurrence of an ISSR is a prerequisite for the in situ formation of natural cirrus clouds (Minnis et al., 2004; Krämer et al., 2016; Kärcher, 2018). These ice clouds play an important role in the Earth energy budget and can form in situ via homogeneous nucleation at temperatures below –38°C and RHi exceeding 140 % (Heymsfield et al., 2017; Kanji et al., 2017) or via heterogeneous freezing at temperatures lower than 0°C and RHi above 100 % in the presence of ice nucleating particles (INP) (Heymsfield et al., 2017). Natural cirrus clouds exhibit a positive climate feedback, estimated at 0.20 ± 0.21 W m^{-2} K^{-1} (Zhou et al., 2014). However, it is worth noting that natural cirrus clouds are not the only ice clouds in the upper troposphere for which the occurrence of an ISSR are a prerequisite. On average, aircraft flying with kerosene fuel at a pressure level, between 325 and 175 hPa can produce contrails for 1–15 % and 6–20 % of their flight time (ratio of flight time in ISSRs to total flight time) in tropical regions and mid- to high-latitude regions, respectively (Sanogo et al., 2024). Contrails can persist and grow for several hours in the ISSRs (Kärcher, 2018). Over time, such persistent contrail can evolve into cirrus clouds and meteorological conditions such as wind shear and turbulence can extend their spatial coverage and affect their ice crystal concentrations (Kärcher, 2018). Contrails can thus enhance the coverage of cirrus clouds (Stubenrauch & Schumann, 2005). Contrails and contrail cirrus together, which is associated with a potentially significant radiative impact, in comparison to the CO_2 effective radiative forcing due to aviation (Boucher et al., 2013; Bock & Burkhardt, 2016; Lee et al., 2021). For instance, for the year 2018, AIC radiative forcing was estimated at 57.4 mW m^{-2}, with a 5%–95% likelihood range of 17 to 98 mW m^{-2}, while that for CO_2 emissions from aviation was estimated to 34.3 mW m^{-2} with a 5%–95% likelihood range of 31 to 38 mW m^{-2} (Lee et al., 2021).

Given the role of ISS in cirrus clouds formation and evolution, it is crucial to represent them in climate models for a more realistic simulation of these clouds. There are various climate models that incorporate the parameterization of ISSRs. Some of these parameterizations represent the formation of cirrus clouds via the homogeneous nucleation pathway (e.g., Tompkins et al., 2007), neglecting heterogeneous nucleation, as the latter is more complex and less well understood (Heymsfield et al., 2017). Despite the lack of representation of heterogeneous nucleation, these parameterizations can reproduce the main features (e.g., spatial distribution) of the ISSRs in a way that is consistent with observations. Other parameterizations take into account heterogeneous nucleation using aerosol models to diagnose the concentration of the INP (e.g., Lohmann et al., 2008; Kärcher & Burkhardt, 2008; Wang & Penner, 2010; Muench & Lohmann, 2020).

In LMDZ, until now, cirrus clouds have been formed by condensing all saturated water, meaning that ISS were not simulated (see Sect. 2.1). ISS parameterization is therefore necessary for more realistic modeling of ice clouds in LMDZ and also to adapt it for the study of contrails: a first version of this new parameterization, which is relatively simple and referred to as first-order parameterization, is described in Section 2.2. Adding such a parametrization requires retuning LMDZ because many of its parameters are only weakly constrained by observations, and the parameterization is expected to strongly change the representation of the upper-atmospheric water budget in the model. This study describes that tuning process.

Parameterizations are a significant source of uncertainty in climate models (Hourdin et al., 2017). For instance, studies by Chen and Gettelman (2013) and Perini et al. (2023) have shown that the radiative effects of AIC are sensitive to the values of free parameters in the climate model (e.g., those that control the contrail lifetime), underscoring the fact that many of these parameters do not have a unique value. To reduce these uncertainties, machine learning methods are more and more used to tune the parametrization of climate models (Jebeile et al., 2023). For example, D. Williamson et al. (2015) used a tuning method based on History Matching and iterative refocusing (see Section 3.3) to explore the parametric uncertainties associated with the global mean temperature and salinity in the NEMO (Nucleus for European Modeling of the Ocean) ocean model. Similarly, Hourdin et al. (2021) applied the same approach to LMDZ, identifying model configurations that were more consistent with observations than the manually fine-tuned base configuration. The use of this method can makes it possible to assess the parametric sensitivity of climate change projections (Hourdin et al., 2023).

In this study, our aim with the first-order parameterization is to explore the feasibility of introducing ISS in LMDZ and to tune it using large-scale observables, while documenting its strengths and limitations. At the same time, a process-based ISSR parameterization has been developed and validated using observations for LMDZ (Borella et al., 2025). This process-based parametrization will be tuned using the insights gained in this study and will be the one used in future work. For the tuning, we used the approach of D. Williamson et al. (2015) and Hourdin et al. (2021, 2023) to also investigate errors in the atmospheric component (LMDZ) of the IPSL climate model, focusing on the spatial and temporal distributions of ISSRs. Furthermore, we aim to identify LMDZ configurations that not only match observed ISSR properties but also align with observations of high-level cloud fractions and global radiative fluxes, to ensure that the model's performance on other critical properties, such as the radiation budget, is preserved.

This manuscript is structured as follows. Section 2.1 describes LMDZ and its parameterizations of ISS and cloud. Sections 3 and 3.3 describe the tuning approach. Section 4 details the experimental setup, while Section 5 presents our results. Finally, Section 6 summarizes our main findings.

2 LMDZ and its parameterizations of clouds

2.1 General description of LMDZ model

LMDZ is the atmospheric component of the Institut Pierre-Simon Laplace climate model known as IPSL-CM which is used for understanding the climate system and its response to various perturbations (Boucher et al., 2020). The version 6A is used in this study. It has 79 vertical levels, that extends up to approximately 80 km above the surface. The resolution is finer in the lower atmosphere to improve the representation of boundary layer transport and associated cloud processes (Hourdin et al., 2020). The first atmospheric layer is centered at 10 m above the surface, and the layer thickness (δz) increases nearly linearly with altitude in the first 3 km, following the relation $\delta z \pm 0.11 z$ (Hourdin et al., 2020). As a result, nearly 40 vertical levels are in the lowest 15 km, including 20 levels within the first 2 km. The version of the model used in this study has a regular horizontal grid of 144 $\times$ 143 points, corresponding to a resolution of 2.5° $\times$ 1.25° in longitude and latitude.

This version of LMDZ incorporates an improved version of the so-called "New Physics", initially introduced in its earlier versions (Hourdin et al., 2020). This physics package notably includes:

- a stochastic triggering scheme for deep convection,
- enhanced representation of stratocumulus and cumulus clouds,

- an improved stable boundary-layer parameterization,
- and a gravity wave drag scheme designed to better simulate the quasi-biennial oscillation.

Radiative transfer is computed using the Rapid Radiative Transfer Model (RRTM). For a comprehensive description of the version of LMDZ used in this study, the reader can refer to Hourdin et al. (2020), Boucher et al. (2020) and Lurton et al. (2020).

The parameterization of clouds in LMDZ is based on the statistical representation of sub-grid total water (q) using a probability density function (PDF, noted $P(q)$) (Hourdin et al., 2020; Madeleine et al., 2020). For stratiform and high-level clouds (i.e., at altitude typically above ∼6.5 km), $P(q)$ is a generalized log-normal distribution (Bony & Emanuel, 2001; Madeleine et al., 2020). It is fully described by its first and second moments, defined here as the mean (q_t) and the standard deviation σ. Due to lack of physical modeling, the standard deviation is simply specified as $\sigma = \xi(p)\ q_t$. The function $\xi(p)$ is used to impose a variation of σ as a function of pressure p. The function $\xi(p)$ increases as pressure decreases to reach an asymptotic value of ξ_{300} at 300 hPa and in the upper troposphere to account for the fact that σ/q_t generally increases. This parameter is an important tuning parameter of LMDZ (Madeleine et al., 2020). In the following, it is renamed ξ_{old} in order to distinguish it from its new formulation in the new parameterization. Given $P(q)$, the gridbox-averaged total water reads:

$$q_t = \int_0^{+\infty} q\, P(q)\, dq \tag{1}$$

Knowing $P(q)$, one can determine the cloud fraction α_{cld} as:

$$\alpha_{cld} = \int_{q_{sat}}^{+\infty} P(q)\, dq \tag{2}$$

where q_{sat} is the total water at saturation. This parameterization is illustrated in Fig. 1a, which shows the saturation adjustment process, where all saturated water condenses into cloud (α_{cld}), thus preventing the representation of ice supersaturated regions (ISSRs) in LMDZ. Note that for the case of shallow convection which is not considered here, the sub-grid water distribution is described by a bi-Gaussian distribution where the thermal plumes and their environment correspond to the small and the main modes of the distribution, respectively (Hourdin et al., 2013, 2020; Madeleine et al., 2020). The parameters required for the computation of this bi-Gaussian distribution are given by a thermal plume scheme (Hourdin et al., 2013, 2020; Madeleine et al., 2020). Cloud phase is determined as a monotonic function of temperature in the mixed-phase temperature regime (Raillard et al., 2024). For a full description of the parameterization of clouds in LMDZ, the reader is referred to Madeleine et al. (2020).

2.2 Ice-supersaturation parameterization: a first approach

2.2.1 The basic framework

The parameterizations of clouds in LMDZ are based on the assumption that water is always at thermodynamic equilibrium between the vapor and liquid phases or between the vapour and solid (ice) phases. However, this assumption is not always met in the atmosphere, as some regions can be supersaturated with respect to ice without being condensed. Our objective is that the model is able to create and maintain such ice supersaturated conditions in a realistic way. Given the coarse resolution of LMDZ (see Sect. 2.1), we consider that a given gridbox can be partitioned between an ice-subsaturated region (a thermodynamic equilibrium state with $q < q_{sat}$), a clear-sky ISS region (a metastable thermodynamic state with $q > q_{sat}$), and a cloud fraction (in thermodynamic equilibrium state with total water $q > q_{sat}$). This new scheme is illustrated in Fig. 1b, where the three fractions are labelled *sub*, *ss* and *cld*, respectively. In the computing process,

this new parameterization needs to retain the cloud fraction from the previous time step (keep the history of cloud fraction, i.e., have memory), to track the fraction of the grid-box where condensation has occurred. The rationale for such a scheme is that an ISS fraction can be converted into a cloudy fraction but the opposite is not true as the cloud can only sublimate in subsaturated conditions. The cloud cover computed at the previous time step, is used as semi-prognostic variable, by considering it as an additional tracer, advected by the dynamical core, along with the specific contents of vapor, liquid and ice. Hereafter, the advected cloud cover and the cloud cover for the current time step are referred to as $\alpha_{cld,adv}$ and α_{cld}, respectively.

We further assume that cirrus clouds form only by homogeneous nucleation. Heterogeneous nucleation is an important process in cirrus cloud formation, but remains a less well-understood physical process (Heymsfield et al., 2017), and accounting for it would require representing INPs in the model, which is beyond the scope of this study but will be the subject of a new study. The tuning method applied in this study could contribute to compensating to some extent for the absence of heterogeneous nucleation processes by making homogeneous nucleation more efficient. The basis of this approach is that ice clouds can form through homogeneous freezing when ice supersaturation exceeds a threshold $\gamma_{ss}\, q_{sat}$ that is approximated as a function of temperature as specified by Koop et al. (2000) and Ren and Mackenzie (2005). The threshold of ice supersaturation ratio, γ_{ss}, is expressed as:

$$\gamma_{ss} = \gamma_0 - T/T_\gamma \tag{3}$$

where T is the temperature (in K). In LMDZ, temperature is assumed to be homogeneously distributed in the gridbox, thus T is the gridbox mean temperature. The values of γ_0 and T_γ were fixed to 2.349 and 259 K (see Table 1), respectively (Ren & Mackenzie, 2005). However, when developing the parametrization, γ_0 and T_γ, are considered as tuning parameters, because the data used to determine Equation 3 may be subject to uncertainty. We defined the allowed intervals for the parameters γ_0 and T_γ by varying the intercept and the slope of the nucleation line within the dispersion range of the data points presented in Figure 1 in Koop et al. (2000). The linear model of Equation 3 is constrained to remain entirely within the limits defined by the dispersion of the observed supersaturation thresholds between 170 K and 240 K, in order to cover the entire range of experimental uncertainties, including variations associated with different crystal sizes. Consequently, the resulting intervals for γ_0 and T_γ represent physically plausible nucleation regimes that encompass the diversity of crystal growth behaviors documented in Koop et al. (2000).

We characterize the three fractions of the gridbox through their respective cover α_k (dimensionless) and their gridbox-average total water q_k ($\mathrm{kg\,kg^{-1}}$). Thus the following conservation equations apply:

$$\alpha_{sub} + \alpha_{ss} + \alpha_{cld} = 1 \tag{4}$$

and

$$q_{sub} + q_{ss} + q_{cld} = q_t \tag{5}$$

where q_{cld} represents the total water content of the cloud, including both water vapor and ice. This parameterization in LMDZ is only applied for temperatures below –30°C. For temperatures warmer than –30°C, the mixed-phase cloud parameterization takes over, with a linear transition applied to avoid discontinuities in α_{cld} and q_{cld}.

2.2.2 Subgrid water vapor distribution

This parameterization fits well the pre-existing one in that is also based on a statistical treatment of water vapor at the subgrid scale (see Sect. 2.1). $P(q)$ is separated into three sub-PDFs P_{sub}, P_{ss} and P_{cld} that describe the total water in the *sub*, *ss*, *cld*

regions, respectively (Fig. 1b). The three PDFs are linked by the following relationship:

$$P(q) = \alpha_{sub}\, P_{sub}(q) + \alpha_{ss}\, P_{ss}(q) + \alpha_{cld}\, P_{cld}(q) \tag{6}$$

The separation of P into distincts parts is a new feature of this parametrization compared to the one introduced in Sect. 2.1. For q values above q_{vc}, P_{cld} is assumed to be proportional to P, otherwise it is equal to 0 (see below). The lower bound of $P_{cld} = q_{vc}$, corresponds to the in-cloud water vapor at the end of the previous physical timestep. The introduction of q_{vc} is needed to diagnose the dissipation of clouds. It should be noted that the evaluation of q_{vc} is an approximation since it relies on the temperature of the gridbox at the previous timestep, while all the other quantities have been advected.

The coefficient of proportionality between P_{cld} and P for $q > q_{vc}$ is denoted N, and derived from the following equation:

$$N = min(1, \alpha_{cld,adv} / \int_{q_{vc}}^{+\infty} P(q)\, dq) \tag{7}$$

As q values can exceed saturation in both the ISS and the cloudy fractions, we have to ensure that the advected cloud fraction is not greater than the diagnosed maximum cloud fraction. Specifically, $N \in]0, 1[$ corresponds to situations where the three fractions coexist in the gridbox, $N = 0$ corresponds to situations where only subsaturated region and ISS fractions coexist, while $N = 1$ corresponds to situations where only subsaturated region and cloudy fractions coexist at the beginning of the physical timestep. Note that if N was limited to 1 because of the min barrier, meaning that the advected cloud fraction is higher than the maximum diagnosed cloud fraction from P, the memory of the cloud fraction is forgotten. In that situation no ISSR is diagnosed, and the cloud scheme falls back to the original cloud scheme of LMDZ. Knowing N and q_{vc}, the normalised PDFs of P_{sub}, P_{ss} and P_{cld} are determined as follows:

$$P_{sub}(q) = \begin{cases} P(q)/\alpha_{sub} & \text{if } q < q_{vc} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$P_{cld}(q) = \begin{cases} 0 & \text{if } q < q_{vc} \\ N\, P(q)/\alpha_{cld,adv} & \text{otherwise} \end{cases} \tag{9}$$

$$P_{ss}(q) = \begin{cases} 0 & \text{if } q < q_{vc} \\ (1-N)\, P(q)/\alpha_{ss} & \text{otherwise} \end{cases} \tag{10}$$

where $\alpha_{sub} = \int_0^{q_{vc}} P(q)\, dq$ and $\alpha_{ss} = 1 - \alpha_{sub} - \alpha_{cld,adv}$ at the beginning of the timestep.

Although the PDFs depend on the in-cloud water vapor of the previous timestep q_{vc}, this is not the case for the ice water content, meaning that the formulation does not guarantee that a formed cloud conserves its total water content. This is particularly challenging for diagnosing the tendencies associated with condensation. The formation of new clouds at each timestep is diagnosed by integrating P_{ss} between $\gamma_{ss} q_{sat}$ and $+\infty$, which corresponds to the fraction of the clear-sky region where humidity is higher than the condensation threshold. When a new cloud is formed, its humidity is by construction higher than $\gamma_{ss} q_{sat}$, and the value of P_{ss} above $\gamma_{ss} q_{sat}$ falls to 0. In the subsequent physical timestep, P_{cld} and P_{ss} are imposed to be proportional to each other. The formed cloud fraction is conserved through $\alpha_{cld,adv}$, but its humidity is redistributed equally between the ISSR and the cloud. More specifically, its humidity decreases to increase that of the ISSR, creating a non-physical flux of humidity from the cloud to the ISSR. Therefore, the humidity in the ISSR is overestimated, as is new cloud formation. To address this limitation, it is assumed that a larger part of the region with $q > \gamma_{ss}\, q_{sat}$ is covered by clouds. For that purpose, the value of the proportionality coefficient N is increased to $N_2 = min(1, \Phi_N\, N)$ for $q > \gamma_{ss}\, q^{vc}$ and to $N_1 = min(\Phi_N, 1)$ for $q^{vc} < q < \gamma_{ss}\, q^{vc}$,

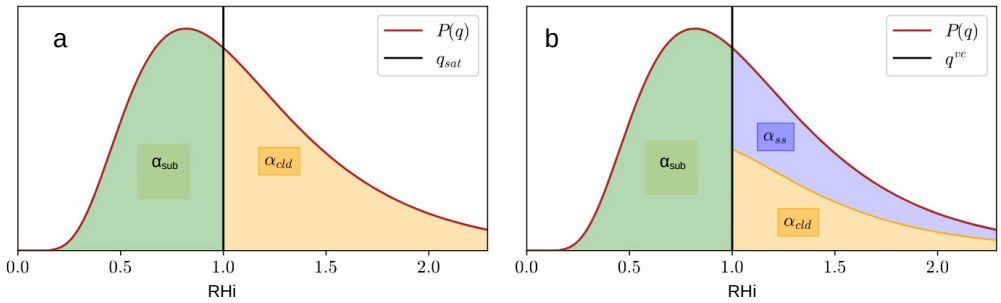

Figure 1. Total water distribution at the subgrid scale versus relative humidity with respect to ice, according to the cloud parameterization in LMDZ6A (a) and the new ice supersaturation parameterization (b). The x-axis shows relative humidity. The vertical black lines in a) and b) define the undersaturated and saturated regions, and correspond to q_{sat} and q^{vc}, respectively. In b) the orange curve represents the $N \cdot P(q) = \alpha_{cld,adv} P_{cld}(q)$ curve. The situation illustrated in (b) is not an equilibrium state, but can be interpreted as the state at the beginning of the timestep before cloud physics is applied.

where Φ_N is a tuning parameter that can take values in the range 1 to 2. In our work, cloud fraction is bounded by a maximum cloud fraction which is determined from P as follows:

$$\alpha_{cld,max} = \int_{q_{sat}}^{+\infty} P(q) dq \tag{11}$$

When $\alpha_{cld,adv} < \alpha_{cld,max}$, cloud fraction is prognostic and the three PDFs of the regions are determined from the prognostic variables. However, when $\alpha_{cld,adv} > \alpha_{cld,max}$, the memory of the cloud properties is lost, and the former diagnostic scheme applies. We refer to this treatment as semi-prognostic. Therefore, cloud fractions are de facto limited by the diagnosis of $P(q)$, and they were shown to be significantly underestimated (Madeleine et al., 2020). This comes from the high value of ξ at cirrus clouds altitudes. To overcome this issue, we reduce the value of ξ when a cirrus cloud was previously present in the gridbox. This treatment allows to simulate high cloud fractions, in line with observations. We assume that the standard deviation of the distribution of total water is inversely proportional to the cloud fraction in a gridbox, and adapt the ξ function accordingly (Eq. 12). This formulation is not directly supported by observations, consequently, we made depending ξ dependent on a tuning parameter (Eq. 12) and test its relevance. The function $\xi(p)$ for the new parameterization is then computed as follows:

$$\xi(p) = \frac{\xi_{oldp}}{1 + \beta \cdot \alpha_{cld,adv}} \tag{12}$$

where β is a tuning parameter whose value is between 0 and 5. Case where $\beta = 0$ corresponds to the condition in which the old parameterization takes over. Tests have shown that the maximum value of 5 is realistic, allowing for scaling that reduces the variance of P compared to the old parameterization. Once the three distributions P_{sub}, P_{ss}, P_{cld}, as well as the total water distribution P in the gridbox, are constructed from q_t, $\alpha_{cld,adv}$ and q_{vc}, the physical processes that modify these quantities are applied. The processes considered are the dissipation of subsaturated clouds, the formation of new clouds, the adjustment of in-cloud water vapor to saturation, and the mixing of clouds with their

environment. When these processes are applied, the shape of the three PDFs are modified, thus Eqs. (8–10) are not valid anymore. Only the distribution of total water in the gridbox is left unchanged throughout the cloud scheme.

2.2.3 Sublimation, condensation, and phase partionning

Sublimation and condensation are diagnosed from the distributions of water in the three regions. First, the region of ice clouds where q_{cld} is lower than q_{sat} is sublimated and transferred to the subsaturated clear-sky region. The corresponding tendency on cloud fraction $(\Delta\alpha_{cld})_{sub}$, is equal to the integral of P_{cld} between q_{vc} and q_{sat}. This process is activated only if $q_{vc} < q_{sat}$, typically because the temperature of the gridbox has increased. Then, the water vapor in the cloud is adjusted to saturation, meaning that the excess vapor deposits onto existing ice crystals if $q_{vc} > q_{sat}$, or the ice crystals sublimate if $q_{vc} < q_{sat}$. Finally, the water vapor in the ISSR that is above the condensation threshold, $\gamma_{ss} q_{sat}$, is condensed and transfered to the cloudy region. The corresponding tendency on cloud fraction, $(\Delta\alpha_{cld})_{cond}$, is equal to the integral of P_{ss} between $\gamma_{ss} \cdot q_{sat}$ and $+\infty$.

The three distributions P_{sub}, P_{cld} and P_{ss}, are modified according to these processes. More specifically, we set $P_{cld} = 0$ and $P_{sub} = P$ between q^{vc} and q_{sat} to account for the sublimation process, and we set $P_{ss} = 0$ and $P_{cld} = P$ above $\gamma_s s q_s a t$ to account for the condensation process. With these new distributions, the total cloudy water content and the humidity in the ISSR can be diagnosed by integrating the corresponding humidity-weighted PDF.

2.2.4 Turbulence

We assumed that cirrus clouds have the shape of a prolate spheroid of volume $\frac{4}{3}\pi a^2 b$, where a and b are its semi-axes (see Fig. 2). The aspect ratio of cirrus clouds is thus defined as $\frac{b}{a}$, and is set to 3 based on sensitivity tests. Small-scale turbulence may mix the cloudy region in the gridbox with the subsaturated and ISS regions. The turbulence is parameterized using a characteristic velocity, v_{turb}, and a characteristic length, l_{turb}. The v_{turb} variable is determined as an approximation of the diagnosed Turbulent Kinetic Energy (TKE) and l_{turb} is a tuning parameter with a typical value of the order of ten to a hundred metres but we used a default value of 50 m. It is also assumed that all the mass that is transferred from the cloud to the environment, or vice versa, occurs within an area around the cloud delimited by a length $L = \min(v_{turb}\Delta t, l_{turb})$. In the environment, this area is that between the boundary (distance R from the centre) and that at distance $R+L$ from the centre (Fig. 2). In the cloud, it is the area between the boundary and that at distance $R - L$ from the centre. It also assumed that these two areas are well mixed by turbulence. The volume of the cloud environment, V_{env}, is the sum of the volumes of the subsaturated region, V_{sub}, and ISS, V_{ss}, which can be expressed through the introduction of a parameter ω, such that:

$$V_{ss} = \omega\, V_{env} \tag{13}$$

and

$$V_{sub} = (1 - \omega)\, V_{env} \tag{14}$$

The impact of turbulence on interactions between clouds and the surrounding air depends on the thermodynamic state of that air and is treated differently for mixing between clouds and ISSRs and mixing between clouds and subsaturated clear air:

2.2.4.1 Turbulent mixing between clouds and ISSRs: In the presence of clouds near an ISSR, turbulence can enhance the diffusion of ice crystals from the cloud into

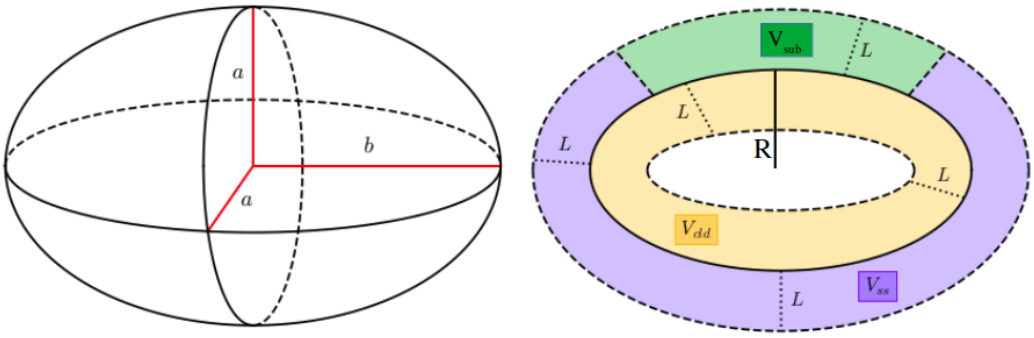

Figure 2. Left panel: Prolate spheroid illustrating our assumption on the shape of clouds. Right panel: cross-section of a cirrus cloud (panel a)). Schematic distribution of the quantities V_{cld}, V_{ss} and V_{sub} in the cloud boundary layer. L is the width of the turbulent mixing zone while R is the distance from the centre of the cloud to the border of the environment around the cloud.

the ISSR, thereby enabling the condensation of supersaturated water vapor, breaking the metastable state. Part of the ISSR becomes cloud, and total water in the cloud increases while that in the ISSR decreases. The cloud fraction tendencies over the time step are computed as follows:

$$(\Delta\alpha_{cld})_{turb,ss} = -\Delta\alpha_{ss} = \frac{V_{ss}}{V_{grid}} \tag{15}$$

$$(\Delta q_{ss})_{turb} = -\frac{q_{ss} \cdot V_{ss}}{V_{grid}} \tag{16}$$

$$(\Delta q_i)_{turb,ss} = -\frac{(q_{ss} - q_{sat}) \cdot V_{ss}}{V_{grid}} \tag{17}$$

where V_{grid} is the total volume of the grid box.

2.2.4.2 Turbulent mixing between clouds and sub-saturated clear sky: For sub-saturated clear-sky regions, two situations are possible, depending on the value of q_{eq} which is the equilibrium total water value after mixing clear and cloudy sky:

$$q_{eq} = \frac{V_{env} \cdot q_{sub} + V_{cld} \cdot q_{cld}}{V_{env} + V_{cld}} \tag{18}$$

- If $q_{eq} > q_{sat}$, meaning that there is enough ice in the cloud region to saturate the clear-sky air and still retain ice crystals, the portion of clear-sky that has been mixed transitions into cloud. The average cloud humidity decreases. These tendencies are computed as follows:

$$(\Delta\alpha_{sub})_{turb,sub} = -\Delta\alpha_{sub} = \frac{V_{sub}}{V_{grid}} \tag{19}$$

$$(\Delta q_{sub})_{turb} = -\frac{q_{sub} \cdot V_{sub}}{V_{grid}} \tag{20}$$

$$(\Delta q_i)_{turb,sub} = -\frac{(q_{sub} - q_{sat}) \cdot V_{sub}}{V_{grid}} \tag{21}$$

Table 1. List of parameters in the ice supersaturation parameterisation. For each parameter, the range of values, the default or First Guess (FG) value and the physical process controlled are indicated. The names of the parameters as used in the tuning process are given in brackets.

Name	Unit	Range	FG value	Controlled processes
T_γ (TG)	K	[249, 269]	259	Saturation rate for homogeneous nucleation
γ_0 (TG0)	–	[2.3, 2.4]	2.349	Saturation rate for homogeneous nucleation
Φ_N (TN)	–	[1, 2]	1.1	Ratio between the P_{cld}, P_{ss} and P
β (RQST)	–	[1, 5]	3.0	Width of P
χ (KHI)	–	[1, 10]	1.1	Boundary size between ISSR, clear sky and cloudy region
l_turb (LTURB)	m	[25, 200]	50.0	Turbulent mixing length around the cloud

- If instead $q_{eq} < q_{sat}$, then all ice crystals sublimate without being able to saturate the clear-sky air. In this case, the cloud transitions to clear-sky, and the mean humidity in the clear-sky region increases. On average, the cloud humidity decreases. The corresponding tendencies are calculated as follows:

$$(\Delta\alpha_{sub})_{turb,sub} = -(1-\omega)\cdot\Delta\alpha_{sub} = \frac{V_{cld}}{V_{grid}} \tag{22}$$

$$(\Delta q_{sub})_{turb} = -\frac{(q_{sat}+q_i)\cdot(1-\omega)\cdot V_{cld}}{V_{grid}} \tag{23}$$

$$(\Delta q_i)_{turb,sub} = -\frac{q_i(1-\omega)\cdot V_{cld}}{V_{grid}} \tag{24}$$

To separate *clr* and *ss* and cloud boundary, we use $\omega = \alpha_{ss}/(\chi\alpha_{sub}+\chi\alpha_{ss})$, where χ is a tunable parameter. This formulation of ω accounts for the fact that there is a higher probability that the cloud and ISS regions are adjacent. In case of a homogeneous distribution in the cloud-free region, $\chi > 1$, otherwise $\chi = 1$.

3 Tuning of LMDZ

The tuning approach adopted is that of Hourdin et al. (2021). It is based on history matching with iterative refocusing and pre-condition the tuning of the 3D version of LMDZ with a tuning of the Single Column version of the Model (SCM). This procedure of Hourdin et al. (2021) is currently the tuning approach of LMDZ and the different metrics are henceforth used as default tuning metrics of LMDZ. The history matching with iterative refocusing procedure and described in several paper is here presented in the 3.3.

3.1 Tuning metrics

3.1.1 Default tuning metrics of LMDZ

The default metrics are those defined by Hourdin et al. (2021) which are metrics of SCM and 3D version of LMDZ. The SCM metrics used are the mixed layer potential temperature and humidity, the average cloud height and the height of the maximum cloud fraction. Large eddy simulation (LES) are used as reference. These metrics and the LES simulations are described in details in Couvreux et al. (2021) and Hourdin et al. (2021). The 3D metrics are the radiative fluxes and precipitation. The radiative metrics include

Table 2. List of radiative and precipitation metrics considered in the LMDZ tuning consisting of radiative fluxes at the top-of-atmosphere averaged over the region shown on the Fig. 3 in Hourdin et al. (2021). The target and σ error chosen for the history matching are shown on the right-hand side of the Table. $\sigma_T = \sqrt{\sigma_r^2 + \sigma_t^2}$, where σ_r represents the uncertainty associated with the target values and σ_t the tolerance to error. The target values are computed from the CERES-EBAF L3b observational data set (Loeb et al., 2009). The regions where these metrics are defined are illustrated in Hourdin et al. (2021, 2023).

Variables	Metric	Unit	Target	σ_T	Products
Net radiation	glob.rt	W m^{-2}	2.5	0.2	LMDZ
Short wave radiation	glob.rsut	W m^{-2}	99.6	5	CERES
Short wave radiation	circAa.rsut	W m^{-2}	24.0	5	CERES
Short wave radiation	subs.rsut	W m^{-2}	84.9	5	CERES
Short wave radiation	weak.rsut	W m^{-2}	81.8	5	CERES
Short wave radiation	conv.rsut	W m^{-2}	103.2	5	CERES
Short wave radiation	etoa.rsut	W m^{-2}	11.0	5	CERES
Long wave radiation	circAa.rlut	W m^{-2}	–48.6	5	CERES
Long wave radiation	subs.rlut	W m^{-2}	274.6	5	CERES
Long wave radiation	weak.rlut	W m^{-2}	264.3	5	CERES
Long wave radiation	conv.rlut	W m^{-2}	235.8	5	CERES
Precipitation	MJO	mm day^{-1}	3.338	0.3	GPCP
Precipitation	PR>50	mm day^{-1}	0.3394	0.17	GPCP
Precipitation	AMMA	mm day^{-1}	1.164	0.63	GPCP

Table 3. List of non-radiative metrics considered in the tuning exercise of LMDZ: Annual average of the high-level cloud (HLC) cover and the Ice Water Path (IWP), the 80^{th} (P1) and 95^{th} (P2) percentile of the distribution of RHi. The metrics of HLC and IWP are computed over NA, EU and the tropical North Atlantic (TNA) and in the Tropical Indo-Pacific (TIP). The metrics P1 and P2 are computed from IAGOS data over the North Atlantic (NA) and Europe (EU) at 250–200 hPa (see Sect. 3.1.2 for the details on the pressure levels). As in Table 2, $\sigma_T = \sqrt{\sigma_r^2 + \sigma_t^2}$.

Variable	Metric	Unit	Target	σ_T	Products
High cloud fraction	NA.clh	%	34.74	11.18	CALIPSO-GOCCP
High cloud fraction	EU.clh	%	33.73	11.18	CALIPSO-GOCCP
High cloud fraction	TNA.clh	%	27.06	11.18	CALIPSO-GOCCP
High cloud fraction	TIP.clh	%	60.43	11.18	CALIPSO-GOCCP
Ice water path	NA.iwp	kg m^{-2}	0.026	0.013	DARDAR
Ice water path	EU.iwp	kg m^{-2}	0.021	0.0105	DARDAR
Ice water path	TNA.iwp	kg m^{-2}	0.018	0.009	DARDAR
Ice water path	TIP.iwp	kg m^{-2}	0.055	0.0275	DARDAR
80^{th} percentile of RHi distribution	NA.250.P1	%	91.98	10	IAGOS
95^{th} percentile of RHi distribution	NA.250.P2	%	113.96	10	IAGOS
80^{th} percentile of RHi distribution	NA.200.P1	%	79.85	10	IAGOS
95^{th} percentile of RHi distribution	NA.200.P2	%	111.87	10	IAGOS
80^{th} percentile of RHi distribution	EU.250.P1	%	96.58	10	IAGOS
95^{th} percentile of RHi distribution	EU.250.P2	%	118.12	10	IAGOS
80^{th} percentile of RHi distribution	EU.200.P1	%	73.15	10	IAGOS
95^{th} percentile of RHi distribution	EU.200.P2	%	104.95	10	IAGOS

metrics of global net radiation (i.e., the imbalance between shortwave (SW) and longwave (LW) and global top-of-atmosphere (TOA) SW upward radiation). As the SW downward radiation is imposed on the model, the global outgoing LW radiation will be constrained automatically by the constraints on the SW and total radiation (Hourdin et al., 2021). Metrics also consider regional averages of the TOA outgoing LW and SW radiation, distinguishing convective, subsidence and intermediate regimes in the tropics and a contrast in latitude between the roaring forties and tropical oceans targeting a circum-Antarctic warm bias in coupled ocean-atmosphere simulations. A specific metric is dedicated to the SW contrast between Eastern Tropical Oceans and the rest of the tropics following (Hourdin et al., 2015). The CERES-EBAF L3b observational dataset (Loeb et al., 2009) is used as reference. The locations where these metrics are computed are indicated in Hourdin et al. (2021) and their values and associated uncertainties are presented in Table 2. In notations such as *glob.rlut*, the first part (*glob*) indicates the spatial domain in which the quantity is computed, and the second part (*rlut*) refers to the specific variable or metric. The acronyms *rt*, *rsut*, and *rlut* refer respectively to the total radiation at the top of the atmosphere, the outgoing shortwave radiation at the top of the atmosphere, and the outgoing longwave radiation at the top of the atmosphere. This naming convention is applied to all the metrics used in this study except those for precipitation (see Fig. S1 and S2 for all cases).

The rainfall metrics used are: (1) the global mean of daily rainfall exceeding 50 mm (PR>50); (2) the mean annual rainfall over a region in West Africa defined for the African Monsoon Multidisciplinary Analysis (AMMA) campaign, where, for simplicity, the acronym AMMA is also used to refer to the metric itself; and (3) an estimate of intra-seasonal variability over the ocean in the Madden–Julian Oscillation (MJO) region, computed from the daily rainfall as the standard deviation of the 20-day running average minus the 120-day running average using the Global Precipitation Climatology Project (GPCP, Huffman et al., 2001) daily precipitation data as reference. The values of these metrics and the associated uncertainties are provided in Table 2. These metrics are now the traditional metrics for tuning LMDZ.

3.1.2 Specific tuning metrics

RHi metrics

In addition to the metrics used in Hourdin et al. (2021), we selected a range of metrics specifically related to our ISS parameterization, aiming to constrain LMDZ on the frequency of ISSR and on the shape of the upper tail of RHi distribution. To this end, we have chosen as metrics of tuning the 80^{th} (P1) and 95^{th} (P2) percentiles of the distribution of RHi, to constraint the range of the upper tail toward the observed one. P1 and P2 are relative to the ISS at least at the sub-grid scale. The advantage of tuning based on these metrics is the ability to adjust the mode of the RHi PDF around 100 %, or even higher values. P1 and P2 are computed at 250 and 200 hPa above the North Atlantic (40–60°N, 65–5°W; NA hereafter) and Europe (40–60°N, 5°W–30°E; EU hereafter). These locations are illustrated in the Fig. S2. Data used as reference are from the Measurement of OZone and water vapour on AIrbus airCraft In-service (MOZAIC) programme (Marenco et al., 1998) over the period 1995-2014 and from the In-service Aircraft for a Global Observing System (IAGOS) programme (Petzold et al., 2015) over the period 2011-2023. For the sake of simplicity, we refer to these two databases as IAGOS data in the following. These data are relevant for the characterization of the statistical properties of RHi (Borella et al., 2024). They data are measured every four seconds with uncertainties ranging from 2% to 8% (5% on average) at the cruising altitude (Smit et al., 2014; Petzold et al., 2020). Due to the inhomogeneous sampling in time (e.g., more measurements in summer) and space (e.g., more measurement at 250 hPa than 200 hPa in the North Atlantic) by the IAGOS and MOZAIC aircraft, some atmospheric conditions are potentially oversampled and this may lead to biases in the long-term RHi distribution

at the regional scale (Sanogo et al., 2024). This temporal and spatial sampling errors are difficult to characterize (Sanogo et al., 2024). To partly take this into account, we consider a total uncertainty of 8%. To make IAGOS data comparable with the LMDZ simulation, a 17-minute moving average (corresponding to approximately 255 km at an aircraft cruising speed of 250 m s^{-1}) is applied to the IAGOS RHi observations. This choice is supported by the findings of Reutter et al. (2020), who showed consistency between ERA5 RHi at 25×25 km resolution and 10-minute averaged IAGOS observations. It should be noted that IAGOS contains a significant amount of missing data and flagged data of poor quality. In order to use robust RHi distributions, we calculated the 17-min data mean only if at least five measurements are flagged as valid. The names of the metrics and their values computed from IAGOS data are provided in Table 3.

Cloud fraction and IWP metrics

Changes in the ISSR fraction impact both upper air cloud cover and the ice water path (IWP), both of which directly affect the radiation balance. Consequently, we used the annual mean cover of cirrus cloud and the ice water path (IWP) as tuning metrics. They are computed over the NA and EU, the tropical deep convection region in the Indo-pacific (10° S–10° N, 60–160° E; IP hereafter) and in the tropical North Atlantic (5–30° N, 20–60° W; TNA hereafter) (Table 3). For the cloud fraction, we used as reference version v2.1 of the GCM-Oriented Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) Cloud Product (GOCCP) noted CALIPSO-GOCCP in what follows and described in Chepfer et al. (2010). This cloud fraction can be subject to uncertainties arising from various sources, such as the choice of horizontal and vertical resolutions, the cloud definition threshold, and the low revisit frequency (16 days) of the CALIPSO satellite. The total uncertainties may be substantial (Chepfer et al., 2010). However, since the cloud fraction from LMDZ is computed in a consistent way to CALIPSO-GOCCP, a total absolute uncertainty of 5 % is deemed to be realistic. It should be noted that clouds are simulated in LMDZ consistently with CALIPSO-GOCCP, by using the CFMIP Observation Simulator Package (COSP). For IWP metrics, a post-processed version of the raDAR/liDAR (DARDAR) dataset in which a dedicated snowfall mask is applied to exclude precipitating snow, is used as a reference (Stengel et al., 2020). The uncertainty for this product is approximately 50 % of the observed value (J. Quaas, personal communication). Consistently with the DARDAR dataset, the LMDZ simulation includes only the contribution from ice crystals and not that from falling snow.

The CALIPSO - GOCCP and DARDAR products have horizontal resolutions of 180×90 and 192×96 grid points in longitude and latitude, respectively. For this study, CALIPSO-GOCCP data from 2006-2009 and DARDAR data from 2007-2009 were used. For comparison with LMDZ, they have been conservatively interpolated to the resolution of LMDZ.

3.2 Tuning parameters

The parameters tuned in this study are those of the ISSR parameterization (Sect. 2.2). Additional parameters that control radiative fluxes, precipitation, cloud cover, and ice water path (IWP) are also tuned. All the parameters are summarised in 4. Some of these additional tuned parameters, such as FALLV and CLTAU, have similar functions in both the convection and large-scale condensation schemes. Their higher values correspond to increased ice crystals fall velocities from cirrus clouds and longer autoconversion timescales, respectively. Higher values of CLC imply a higher rate of conversion of cloud liquid water into rainfall. As described in Section 2.2.2, the standard deviation of P depends on ξ, whose vertical profile is controlled by RQSP0 and RQSDP: The former sets the midpoint, while the latter defines the pressure range over which ξ varies (see Fig. 2 in Madeleine et al. (2020)). The EVAP parameter governs the fraction of precipitation that re-evaporates below the cloud base; higher values enhance rainfall evaporation (see equation 10 in Hour-

Table 4. List of tuning parameters. For each parameter, we give the minimum and maximum values permitted during the tuning process and the way (linear or logarithmic) the interval is sampled. The values of the parameters in LMDZ6A and the First Guess (FG) values of the ISS parameterization (not used in the tuning procedure) are given. Parameters from the ice supersaturation parameterization itself range from TG to LTURB. Other parameters target other aspects of the cloud parameterization.

Code	Unit	Range	6A/FG	Controlled processes
TG	K	[249, 269]	259	Saturation rate for homogeneous nucleation
TG0	–	[2.3, 2.4]	2.349	Saturation rate for homogeneous nucleation
TN	–	[1, 2]	1.1	Ratio between the P_{cld} and P_{ss}
RQST	–	[1, 5]	3.0	Width of P
KHI	–	[1, 10]	1.1	Boundary size between ISSR, clear sky and cloudy region
LTURB	m	[25, 200]	50.0	Turbulent mixing length around the cloud
EVAP	–	[5.10^{-5}, 5.10^{-4}]	$1e^{-4}$	Rainfall evaporation into the atmospheric layers below
EPMAX	–	[0.990 1]	0.999	Precipitation efficiency in deep convection
CLC	kg kg^{-1}	[10^{-4}, 10^{-3}]	6.5e-4	Precipitation conversion in convective clouds
FALLV	–	[0.5, 1]	0.8	Scaling factor for ice crystal fall velocity in convective clouds
CLTAU	s	[800,1900]	900	Autoconversion characteristic time
RQSTOP	–	[0.2, 0.6]	0.4	Width of the large scale high-level cloud P
RQSP0	hPa	[300, 500]	450	Pressure level for the bottom of the layer where RQSTOP varies
RQSDP	hPa	[90, 500]	100	Pressure level for the middle of the layer where RQSTOP varies
REIMAX	μm	[30,100]	61.29	Maximum ice crystal size in cirrus clouds for radiation
REIMIN	μm	[4,20]	16	Minimum ice crystal size in cirrus clouds for radiation

din, 2021). EPMAX regulates precipitation efficiency in convective clouds, controls detrainment and residual cloud water. More specifically, in convective systems, 1-EPMAX refers to the water contained within the portion of the cloud that does not precipitate. It is the main source of moisture in the upper troposphere in LMDZ. The FALLV parameter controls the terminal velocity of ice crystals and acts as a sink. The higher the FALLV value, the faster the ice crystals will fall and the less moisture will remain when the cloud dissipates. This limits the formation of ISSRs. Finally, REIMAX and REIMIN control ice crystal sizes, which is a key variable for cloud-radiative interactions. Unrealistic values (too large or too small) can skew the simulated radiation budget.

3.3 History matching with iterative refocusing

History matching with iterative refocusing is a method used in computer experiment tasks to explore the space of the values of the free parameters of complex models. To apply it, a proxy modeling framework is used. This involves using a computationally cheaper model to efficiently explore the space of values of the parameters of interest, at a lower cost than the full complex model. The goal is to identify regions of the space of the values of the parameters that are not ruled out by comparing the output of the model and observations. In this framework, instead of looking for a single vector of parameter values for which the model performs well, the method rules out implausible values for which the model is unable to reproduce key observed metrics. To avoid overfitting the models, tolerance levels and uncertainty estimates are taken into account

in the procedure. This framework has been described in several studies, notably by D. Williamson et al. (2015); D. B. Williamson et al. (2017) in the context of climate modeling. In previous studies, Couvreux et al. (2021); Hourdin et al. (2021) developed a tool named HighTune Explorer to facilitate the implementation of this framework for the tuning of climate models. This study was conducted using HighTune Explorer.

The implementation of the History Matching with Iterative refocusing using HighTune Explorer consists of six steps: (i) Definition of a set of scalar metrics, for which a reference value and associated uncertainty can be obtained, based on observations (as defined in Table 2 and Table 3) or high-resolution simulations such as LES (as indicated in section 3.1.1). (ii) Identification of the tuning parameters for which acceptable value ranges are specified (as defined in Table 4) to define the space of the values of the parameters. (iii) Perform a set of simulations (Perturbed Physics Ensemble – PPE) with the climate model, using different vectors of free parameter values sampled from the parameter space using a Latin hypercube approach. The different metrics are computed from the PPE to build a train set. (iv) Training of statistical emulators (the proxy model) for each metric. Gaussian Process (GP) emulators are used for their robust probabilistic properties, providing both the expected values and the associated uncertainty estimates across the parameter space. (v) Computation of an Implausibility index I for each vector of parameter values, to quantify the mismatch between the emulator prediction and the observational target. Only vectors of values of parameters with I, below a predefined threshold are retained, and they constitute the Not-Ruled-Out-Yet (NROY) space. In this study, the threshold is set to three, following the rule of Pukelsheim (1994), which states that 95 % of any unimodal Gaussian distribution lies within ± 3 standard deviations of the mean. The index I is defined as:

$$I = \frac{|r - E\left[f(\lambda)\right]|}{\sqrt{\sigma_r^2 + \sigma_t^2 + Var[f(\lambda)]}} \tag{25}$$

where r is the target reference value, σ_r its associated uncertainty, σ_t the tolerance to error. $E[f(\lambda)]$ and $Var[f(\lambda)]$ are, respectively, the expectation of the emulator prediction and the variance of the uncertainty associated with the metric f for each vector λ of parameters. A low implausibility can result either from a good match with observations or from a high emulator uncertainty. (vi) Iterative refocusing process, which aims to reduce emulator uncertainty by performing new simulations within the updated NROY space (iteration over steps iii-iv). Each of these iterations is called a wave. The emulator is re-trained for each wave, and the NROY space is progressively refined. The process is stopped when the NROY space stabilizes, indicating that the emulator uncertainties are sufficiently reduced, and the remaining uncertainties are dominated by those of observations.

In this study, the method was applied to the LMDZ to tune 16 free parameters that are listed in Table 4, using as constraints the metrics provided in Table 2 and Table 3. The default values (considered as First Guess – denoted FG here) for all these parameters are from previous tuning experiences of the LMDZ. Here, we have defined the acceptable value ranges of these parameters around these FG values, taking care to remain within physically plausible limits. This is made possible by the tuning method used, which automatically removes the values of the vectors of parameters incompatible with observations. The NROY space is calculated using a random sample of 10^6 vectors of parameters for which the $I < 3$ criterion is verified by varying the other (16 – 2) parameters. The tuning was carried out in twelve successive waves: nine waves with the 1D version of the model, followed by three waves, in which the metrics from the simulations of the 1D and 3D versions of LMDZ were combined to identify the $NROY$ space. This process is known as preconditioned 1D 3D tuning (Hourdin et al., 2021).

4 Simulations and bias correction

4.1 Experimental protocol and forcing

Simulations with the 3D and 1D (single column) versions of LMDZ are carried out in this study. The 3D simulations are forced by the seasonal cycle of SST and Sea Ice Cover (SIC) following the Atmospheric Model Intercomparison Project (AMIP) protocol. The clouds cover is simulated by using the CFMIP Observation Simulator Package (COSP). The duration of 3D simulations is 2 years.

As a reminder, we wish to tune the 3D version of LMDZ, and for this, a preconditoning tuning of the single column (1D) version of the model in the boundary layer has proved to be a promising method. Couvreux et al. (2021); Hourdin et al. (2021) have shown that this step enables tuning of parameters that matter to represent typical cloud situations, independently of the global state of the model. It also speeds up the 3D tuning, since it eliminates parameter space values that do not produce physically realistic results on typical case studies, at a much lower computational cost. We performed the Single Column Model (SCM or 1D) simulations for four observational case studies for which Large Eddy Simulations (LES) have been developed. The first case is an almost cloud-free convective boundary layer case observed on June 14, 2002 over the Southern Great Plain during the International H_2O Project (IHOP) campaign (Couvreux et al., 2005). The second case deals with the diurnal cycle of shallow convection over land observed on June 21, 1997, with fairly well-developed cumulus at the Atmospheric Radiation Measurement site in Oklahoma (Brown et al., 2002). The third case is a raining cumulus over the ocean (VanZanten et al., 2011). The fourth case is a composite transition case of stratocumulus to cumulus clouds, as described by Sandu and Stevens (2011). These case studies are named IHOP/REF, ARMCU/REF, RICO and SANDU, respectively. The different SCM simulations carried out are not nudged and are described in details with those of the associated LES in Couvreux et al. (2021) and Hourdin et al. (2021).

4.2 Temperature bias correction

Temperature is a key factor in the formation and evolution of ISSRs. However, analysis of the control simulation indicated that the LMDZ model exhibits a temperature bias in the upper troposphere and lower stratosphere (Fig. 3). This cold bias is observed both with the pre-existing and the new parameterization, which means it is a likely a model structural deficiency that is independent of the representation of ISSRs. As our objective is to show the feasibility of tuning the ISS parameterization, reducing this long-standing temperature bias by new model developments would involve complex research beyond the scope of this study, instead we reduce the bias using temperature nudging. To this effect, we first performed a nudged simulation, whereby the LMDZ meridional (u) and zonal (v) wind and the temperature (T) fields are relaxed toward those of the fifth generation of the European Centre for Medium-Range Weather Forecasts atmospheric reanalysis (ERA5) as formalized in the following equation:

$$\frac{\partial X}{\partial t} = F(X) + \frac{X^a - X}{\tau} \tag{26}$$

where F is the operator describing the dynamical and physical processes that determine the unconstrained evolution of X, X^a is the 6-hourly u, v an T fields from ERA5, interpolated to the model grid and model timestep, and τ is a relaxation time constant (Coindreau et al., 2007; Krinner et al., 2020), set to 1 day in this study. The nudging is not applied in the atmospheric boundary layer to allow the model's physics, particularly shallow convection and turbulent mixing, to operate freely. The nudged simulation is carried out for 10 years and the error terms $(X^a - X)/\tau$ are archived. The 10-year (i.e., climatological) monthly-averages of $(X^a - X)/\tau$ are then computed and used to correct the model online in all the remaining simulations of this study. More details

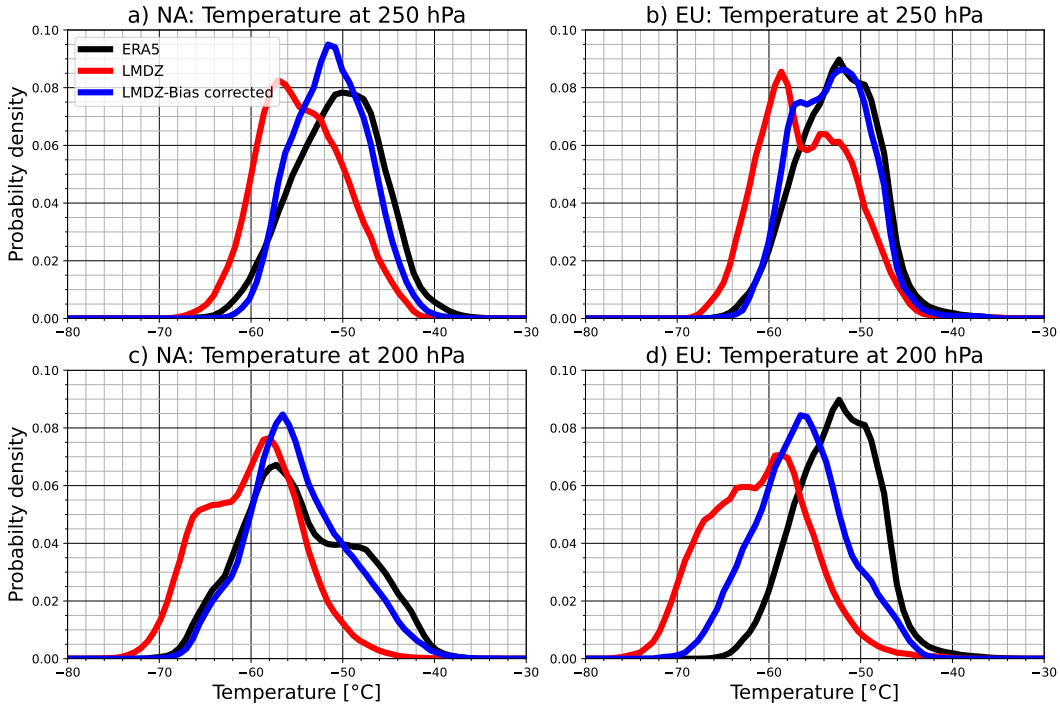

Figure 3. Probability density function of the temperature (°C), binned in 1°C interval, computed from ERA5 (black lines), LMDZ without (red lines) and without (blue lines) online bias correction at 250 and 200 hPa, a,c) over the North Atlantic (NA), and b,d) over Europe (EU). The definition of the NA and EU regions is shown on Supplementary Figure S1.

on this bias correction method, their benefits and their limitations can be found in Krinner et al. (2020). We tested the bias correction method using the averaged relaxation terms of u, v and T and u and v only. The former approach is retained for this study because it allows a better correction of the LMDZ bias in temperature in the UTLS. Fig. 3 illustrates the remaining error for the North Atlantic and Europe at 250 and 200 hPa. The error that persists is greater at 200 hPa in Europe, which can affect the frequency of occurrence of ISSRs since homogeneous nucleation is controlled by temperature (Sections 2.1 and 2.2).

5 Results

5.1 The Implausibility matrix

This section discusses the Implausibility matrix with a focus on wave 11 (Fig. 4). The Implausibility Matrix is a way to illustrate the NROY space, which has 16 dimensions, as 16 parameters are being tuned. In Fig. 4, each panel describes this space in the plane formed by two parameters in the diagonal of the matrix. Each panel is divided into 255 pixels and the density of the NROY space in the remaining 16–2 dimensions is calculated for each pixel. In the submatrix of the lower triangle, the minimum Implausibility value is shown using different values of T ranging from 1 to 3, as represented by the color bar on the left. More information can be found in the figure caption.

In the upper triangular submatrix, the grey areas indicate regions of the parameter space that are incompatible with the imposed constraints. These areas correspond

to the lower triangle in the region defined using $T = 3$ as the Implausibility threshold. The black dot shows the FG values of parameters.

To make the process easier to understand, the Implausibility matrix, from wave 1 ($NROY^1$), is presented in the supplementary material (Fig. S4). The evolution of the density of points highlights the areas of the $NROY^1$ space with the highest probability of finding the relevant vector of parameter values. For wave 1, only the parameter $RQSDP$ is discriminating (Fig. S4). Regions of the parameter space associated with high $RQSDP$ values are implausible. It should, however, be noted that emulator errors are significant at the beginning of the experiment. Consequently, although several parameter vectors in the $NROY^1$ space satisfy the constraints, they do not produce simulations that are consistent with observations. In wave 11, the emulators are sufficiently efficient. The Implausibility index is less dominated by emulator error, making the process more discriminating. However, it remains difficult to clearly identify the most important parameters in the ISS parameterization controlling the RHi distribution, since there is no clear, specific region of the space of parameters that is ruled out (Fig. 4). It is the roles of $EVAP$, $RQSDP$, and $CTAU$ that are clearly evident, and to a lesser extent, so are those of CLC and $SQSN$. These parameters control the spatial variability and the vertical transport of water in LMDZ (Table 4). It is interesting to note that, given the emulator performance at wave 11, the FG parameter vector passed the Implausibility test.

The results mainly discussed in the remainder of this manuscript are those of simulations from wave 12, which were performed using vectors of parameters selected from the $NROY^{11}$ space. The $NROY^{12}$ is also illustrated in the supplementary material (Fig. S5), which indicates that the FG vector of parameters is no longer plausible since at least the points (REIMAX, TG) and (EPMAX, CLC) have been ruled out. It should be noted that the fractured nature of $NROY^{12}$ space and the low density of the remaining points demonstrate the difficulty of satisfying the imposed constraints. Consequently, we did not perform the 13^{th} wave and considered that the waves performed are sufficient for documenting the behavior of LMDZ on the properties of ISSRs.

5.2 Selection of configurations

To select the LMDZ configurations that correspond best to the observation targets, we analysed the 300 3D simulations (see Section 3.3) carried out during waves 10, 11 and 12. For each 3D simulation and tuning metric, we computed the score $S(\lambda)$ as follows:

$$S(\lambda) = \frac{[f(\lambda) - r_f]^2}{\sigma_r^2 + \sigma_t^2} \tag{27}$$

$S(\lambda)$ measures the distance between each simulation and the target for each metric normalised by the total uncertainty. Fig. 5 highlights the metrics well and less well reproduced by the different configurations. The metrics for which the most of the different configurations are consistent with the observations are the metrics NA.clh, EU.clh, TNA.clh, NA.clh and TNA.iwp. The metrics EU.200.P1, EU.200.P2 (metrics of RHi in EU), conv.rlut (longwave outgoing radiation at the top of the atmosphere in the Tropical Deep Convection (TDC) zone of the tropical Pacific) and conv.rsut (shortwave outgoing radiation at the top of the atmosphere in the TDC) are the most challenging metrics. The errors in conv.rsut and conv.rlut reflect the tendency of most configurations to overestimate the high-level cloud fraction and ice content in the TDC region (Fig. 5). The deficiency in EU.200.P1 and EU.200.P2 corresponds to a dry bias of LMDZ at 200 hPa and is discussed in detail in the following section.

In line with the objectives of finding configurations that perform well in terms of ISSR properties, cloud fraction and their radiative effects, we ranked the 300 configurations by averaging S over all metrics. The best configurations, defined here as those

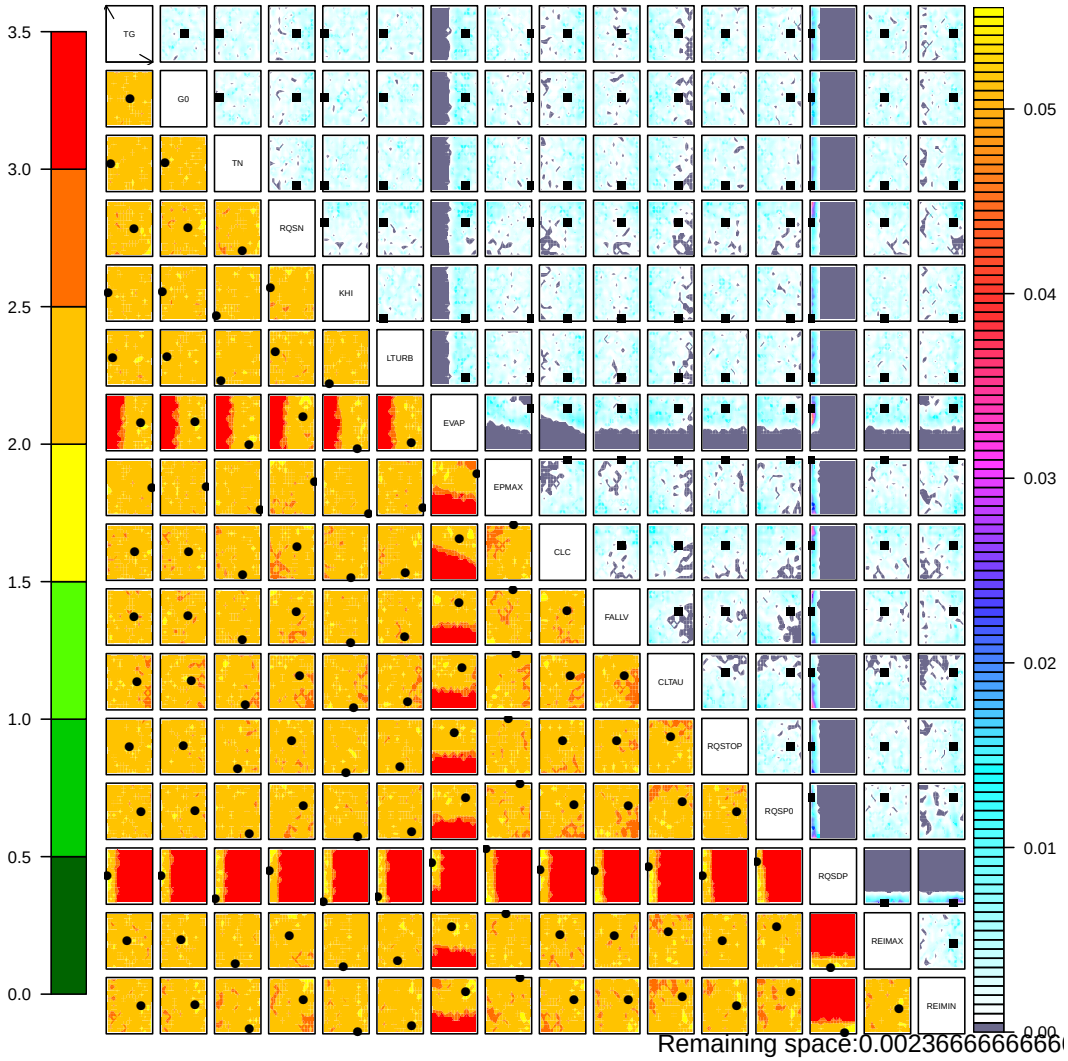

Figure 4. Implausibility matrix for wave 11 of the history matching exploration. The upper-right triangle is made of sub-matrices that display the fraction of points with implausibility lower than the chosen cutoff while the sub-matrices of the lower-left triangle show the minimum value of the implausibility when all the parameters are varied except those used as x- and y-axis. The name of which are given on the diagonal of the main matrix. The squares and black dots correspond to the values of the parameters in the 6A configuration of LMDZ and the first guess values of the parameters in the ISS and cloud parameterizations. The upper right triangle is made up of 2-dimensional sub-matrices, each representing the subspace introduced by the two parameters whose names are given in the diagonal. For example, the x-axis in the top right sub-matrix corresponds to $REMIN$ and the y-axis to TG. x- and y-axis cover the range of values of the corresponding parameters, respectively. The two axes are divided into 15 sub-intervals, so that the sub-matrices are made of 225 pixels. The black square indicates the value used in LMDZ-FG. The density of points with respect to the initial Input space (here $3{\cdot}10^5$ vectors of parameters) with implausibility lower than the cutoff of 3, when varying the 16–2 other parameters, is computed and illustrated by the color bar (from cyan to orange). The sub-matrices of the lower-left triangle display for each pixel the minimum Implausibility obtained when varying the 16–2 other parameters.

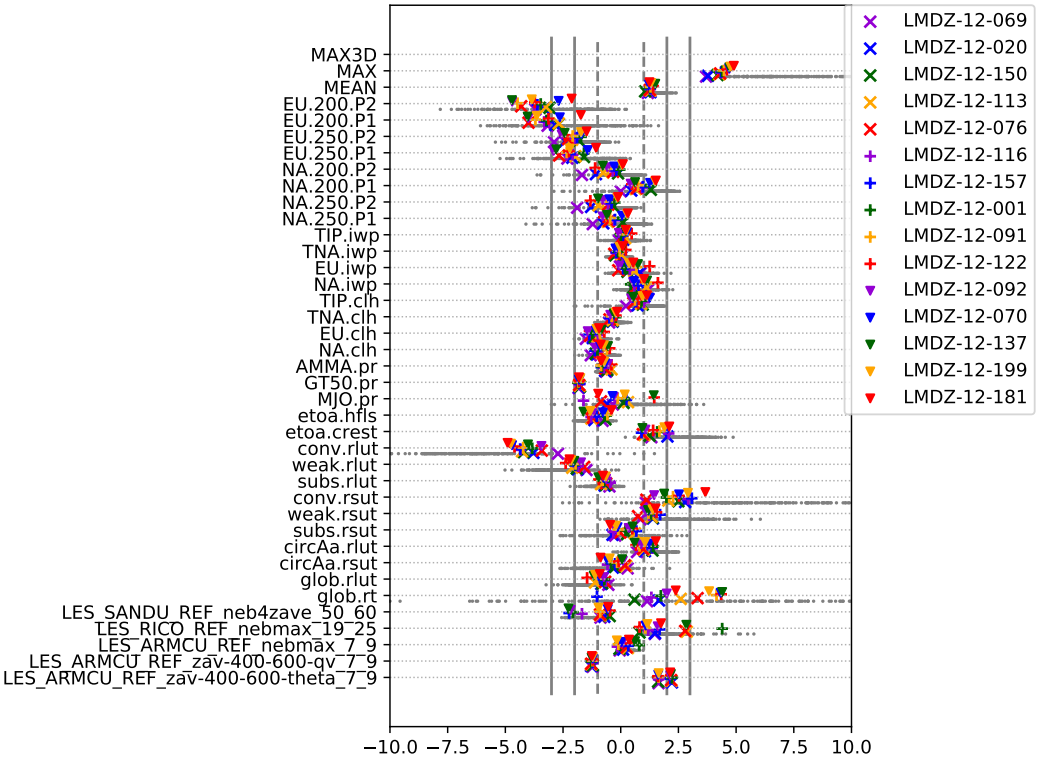

Figure 5. LMDZ configuration ranking scores S, as formalized in Eq 27. For each metric, the sign of the absolute difference is associated to S. The configurations shown with symbols in the legend correspond to those with the lowest average S, ranked in descending order of performance from top to bottom. Metrics listed on the y-axis are, from bottom to top, for processes observed in case studies (SANDU, RICO, ARMCU) of cumulus clouds; radiative flux metrics from glob.rt to etoa.hfls; precipitation metrics from MJO.pr to AMMA.pr; high-level cloud fraction (clh) and ice water path (iwp) metrics from NA.cld to TDC.iwp; RHi metrics NA.P1 to EU.P2 are for the North Atlantic (NA) and Europe (EU). MEAN and MAX correspond to the average and the maximum values of S for each configuration.

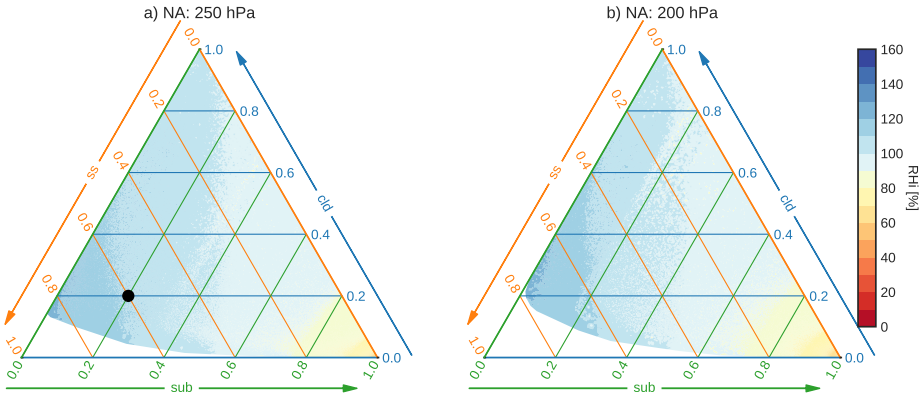

Figure 6. Ternary plots of mean gridbox RHi as a function of the fraction of subsaturated clear sky (sub), ice-supersaturated clear sky (ss) and cloudy sky (cld) in the grid boxes of the North Atlantic (NA) LMDZ at: a) 250 hPa and b) 200 hPa. The black dot in figure (a) is used to illustrate how to read the diagram. It corresponds to a fraction of 0.6, 0.2 and 0.2 of the ISSR (orange sloping line to the left), CLR (green sloping line to the right) and CLD (horizontal blue line) in an LMDZ gridbox, respectively. For each point the sum of the coordinates is equal to 1.

with the lowest maximum of S. However, none of these simulations performed well on all the metrics of interest. In what follows, the 15 best configurations of wave 12 are evaluated on metrics other than those used as tuning constraints. Only the results of wave 12 are shown because the best configuration of these wave performed better than those of wave 10 and 11. It should be noted that, S being an performance score across all the metrics, the 15 configurations selected are not necessarily the best of the 300 configurations on each metric taken individually. In addition, they do not all necessarily perform well on each of the metrics. Therefore, in what follows we also discuss which of these 15 configurations are particularly well performing on the properties of the ISSRs (mean frequency) in different regions, and more generally on the properties of the RHi distribution.

5.3 Performance of the selected configurations

5.3.1 Frequency of occurrence of ISSRs

Comparison with IAGOS and ERA5 over USA, NA and EU

In this section, we evaluate the ability of the selected configurations to reproduce the frequency of occurrence of ISSRs, which is not directly used as a constraint in the tuning process. RHi from ERA5 interpolated on the spatial resolution of LMDZ is also used as a reference, acknowledging that ERA5 may suffer from its own biases. We then defined the ISSRs consistently in IAGOS and ERA5 and in LMDZ, as gridbox average RHi $>$ 100 % collocating with a temperature below –38°C. However, due to the narrowness of ISSRs in IAGOS and ERA5 (Reutter et al., 2020; Wolf et al., 2024), and the fact that ISS in an LMDZ gridbox is often associated with sub-saturated conditions (see Fig. 6), the frequency of occurrence of ISSRs as defined in this study may be underestimated because the spatial mean smooths out the small-scale variability of RHi. A possible solution would be to use a detection threshold lower than 100 %. However, in this study our aim is simply to quantify the degree of agreement between LMDZ and IAGOS, rather

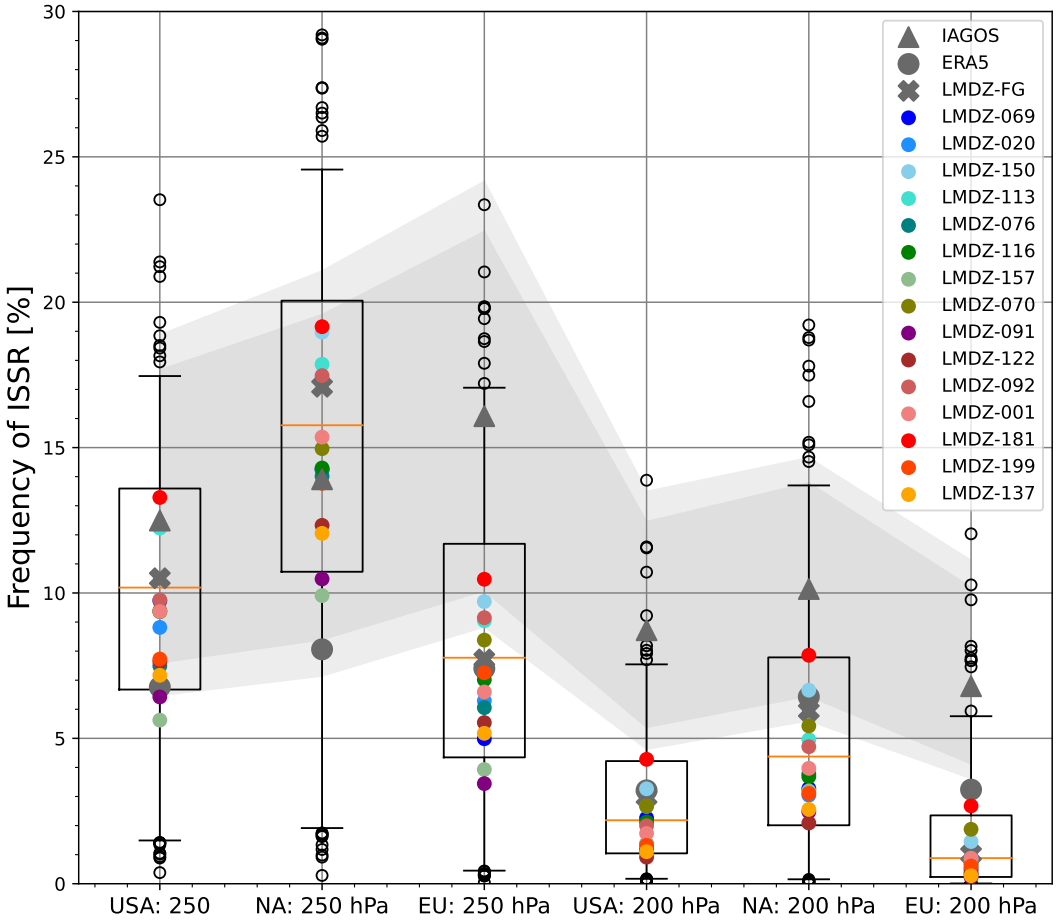

Figure 7. Frequency of occurrence of ISSRs (in %) in the USA (40–60°N, 105°–65°W), the North Atlantic (40–60°N, 65°–5°W; NA) and Europe (40–60°N, 5°W–30°E; EU) at 250 and 200 hPa, from IAGOS (grey triangle), ERA5 (grey dot) and from LMDZ with FG vectors of parameter values (x-shaped markers). The dark and light grey bands correspond to the uncertainties provided with IAGOS data and the total uncertainties used in the tuning of LMDZ (see Sect 3.1.2). The box plots show the distribution of the frequency of occurrence of ISSRs calculated from the 300 simulations of wave 12. Those for the 15 best configurations are shown in the legend. The regions corresponding to USA, NA and EU are shown in Figure S1.

than to document the exact observed frequency of ISSRs. It should be noted that the differences in temperature between IAGOS, ERA5 and LMDZ have little influence on the differences in the frequency of ISSRs in mid- and high-latitude regions discussed below, because in these regions temperatures are rarely greater than –38°C in the UTLS between 350 and 175 hPa (Fig. 3, Reutter et al., 2020).

IAGOS indicates that aircraft flying from North America to Europe will, on average, cross ISSRs more frequently in EU at 250 hPa (14.8 % of the time) and in the NA at 200 hPa (11 % of the time) (Fig. 7). The uncertainties associated with these frequencies in the three regions are of the order of ∼5 % and 3 % 250 and 200 hPa, respectively. At 200 hPa, ERA5 is in agreement with IAGOS on the pattern of the frequency of occurrence of ISSRs (Fig. 7). However, at 250 hPa, in contrast to IAGOS, ERA5 shows that ISSRs occur more frequently in the NA (Fig. 7). Analysis of the spatial distribution of IAGOS measurements showed that the difference between ERA5 and IAGOS at 250 hPa may be due to differences in spatial sampling (Fig. S3). The trajectories of the IAGOS aircraft are more northerly in the NA than over EU (Fig. S3). Consequently, at 250 hPa they sample more dry stratospheric air mass in the NA, resulting in a lower frequency of ISSRs in this pressure layer compared to EU. Lamquin et al. (2012) also found from satellite observations that ISSRs occur more frequently in the NA than in EU at pressure levels between 300 and 100 hPa. This is consistent with the higher occurrence of the warm conveyor belt in the NA (Binder et al., 2020; Joos et al., 2023), which is associated with mid-latitude cyclogenesis and can induce the formation of ISSRs in the upper troposphere (Spichtinger, Gierens, & Wernli, 2005). Interestingly, the selected configurations of the LMDZ and also LMDZ-FG agree with ERA5 on the variation of the ISSR frequency from North America to Europe on average at 250 and 200 hPa, and with IAGOS and ERA5 at 200 hPa (Fig. 7).

All selected and LMDZ-FG configurations are in agreement with IAGOS on the magnitude of the frequency of occurrence of ISSRs at 250 hPa in the NA and in the USA (Fig. 7). However, they are deficient at 200 hPa, with a tendency to underestimate the frequency of occurrence of ISSRs (Fig. 7). It is only in the NA that some configurations (LMDZ-150, LMDZ-070, LMDZ-181, LMDZ-FG) that simulate a frequency of ISSRs that are within the tolerance limit ($\pm\sigma_T$, see Table 3 for the definition of σ_T) at 200 hPa. LMDZ-150, LMDZ-070, LMDZ-181 are configurations performing the best at 200 hPa over EU. They are also the best performing over USA where no tuning metrics are used. It should be noted that the LMDZ-FG is derived from a manual setting in the NA using ISSR frequency as a metric. The negative bias of LMDZ-FG in ISSR frequency in Europe is more pronounced (Fig. 7). Given some of the 300 configurations fall within the IGAOS uncertainties, it can be concluded that the LMDZ-FG bias was partly due to insufficient tuning. It is however, worth noting that in addition to LMDZ-FG most of the configurations are drier than ERA5 at 200 hPa, which is known to have a dry bias in the UTLS against IAGOS (Reutter et al., 2020; Wolf et al., 2024). This confirms the hypothesis that the selected LMDZ configurations are too dry at 200 hPa in the USA and EU. Analysis of the relationship between the frequency of the ISSRs and the height of the tropopause in the 300 configurations of waves 10, 11 (not shown) and 12 (Fig. 8) of LMDZ shows that in the mid-latitude regions of the Northern Hemisphere, the tropopause is lower over the oceans than over the continent. This results in a lower frequency of ISSRs at 250 and 200 hPa in the USA and EU than in the NA, due to a higher presence of dry stratospheric air masses.

For the selected configurations, their dry bias at 200 hPa over Europe and USA is due to a relatively low tropopause, although their cold temperature (Fig. 3), should favour a higher frequency of occurrence of ISSRs. The few configurations that agree with IAGOS on the frequency of occurrence of ISSRs at 200 hPa in the EU and in the USA overestimate the frequency of ISSRs at 250 hPa in the NA and USA. This highlights a structural error in LMDZ in reproducing the frequency of occurrence of ISSRs at 200 hPa

in the mid-latitudes of the Northern Hemisphere. The analyses also show a strong relationship between the frequency of occurrence of ISSRs and the proportion of stratospheric air masses at 250 and 200 hPa (Fig. 8). This suggests that the performance of climate models on the frequency of occurrence of ISSRs depends strongly on their ability to reproduce the spatial and temporal variability of the tropopause height.

Spatial distribution: Comparison with ERA5

As the IAGOS data do not cover the globe, we used ERA5 as a reference to assess the relevance of the selected configurations of the LMDZ on the frequency of occurrence of ISSRs on a global scale. It has been shown that ERA5 reproduces the spatial features of these frequencies of occurrence (Lamquin et al., 2012). This reanalysis product indicates that the frequency of ISSRs is high (4-20 %) everywhere at 250 hPa except in areas of subsidence of dry air masses (0-4 %) located between 40°S and 40°N (Fig. 9a). At 200 hPa, ISSRs occur more frequently in the tropics, and in Antarctica and over the ocean around 40°N/S where the frequency of the warm conveyor belt is higher (Fig. 9b, Joos et al., 2023). LMDZ is relatively consistent with ERA5 in the spatial distribution of these frequencies in the tropics and over the ocean in the mid-latitudes of the Northern Hemisphere (Figs. 9a-d). LMDZ fails to capture the spatial distribution of the frequency of occurrence of ISSRs at mid and high latitudes of the Northern Hemisphere at 200 hPa, and in the high latitude regions of the SH and Eurasia at 250 hPa (Figs. 9a-d and S4). This deficiency is observed in the simulations of waves 10, 11 (not shown) and 12 and could reflect a structural error. This deficiency is not unique to LMDZ, as Irvine and Shine (2015) and Perini et al. (2023) reported similar behaviour for the MIROC5 and ARPEGE-climat climate models, respectively. These errors in the set of 15 configurations is higher in the region where the frequency of occurrence of ISSRs is higher (Fig. 9e-f). In regions of low frequency of occurrence of ISSRs, the parametric uncertainty is more than 30 % of the ensemble mean. The configurations most consistent with ERA5 on the spatial distribution of ISSR are LMDZ-181, LMDZ-150, LMDZ-070 (Figs. S3-S4). It should be noted that these four configurations agree that up to 80 % of gridbox (∼1.2°x2.5° of resolution) can be covered by ISSRs (Fig. 6).

To assess the added value of this tuning work on a global scale, we analyzed the bias with respect to ERA5 for the LMDZ-070, LMDZ-150, LMDZ-181 and LMDZ-FG configurations (Figs. 10, S6, S7). It should be noted that most studies evaluating ERA5 on the statistical properties of ISSRs focus on the Northern Hemisphere. Consequently, the quality of this product is better known in this region, with the dry bias already well documented.

Interestingly, the three configurations LMDZ-070, 150 and 181 show a less pronounced dry bias over Eurasia than LMDZ-FG, at both 250 and 200 hPa (Figs. 10, S6, S7). More generally, LMDZ shows a marked positive bias in areas of high synoptic activity north of 40°N/S, but the extent of areas affected by high bias values remains more pronounced in LMDZ-FG. Moreover, the bias in Antarctic regions is also greater for LMDZ-FG.

In contrast, in the Tropics, LMDZ-070, 150 and 181 appear to have too intense convection compared with LMDZ-FG (Figs. 10, S6, S7). It should be pointed out, however, that the quality of ERA5 on the properties of ISSRs in the tropics is less well known, and that no tuning metrics specific to tropical ISSRs were defined in this exercise.

5.3.2 Properties of the probability density function of RHi

To assess the ability of LMDZ to reproduce the properties of the RHI distribution, for simplicity only the best configuration (LMDZ-181) on the ISSR frequency is first analyzed. We first qualitatively assess the ability of LMDZ-181 to capture the properties of RHi in the UTLS. IAGOS indicates that at 250 and 200 hPa, the PDF decreases and

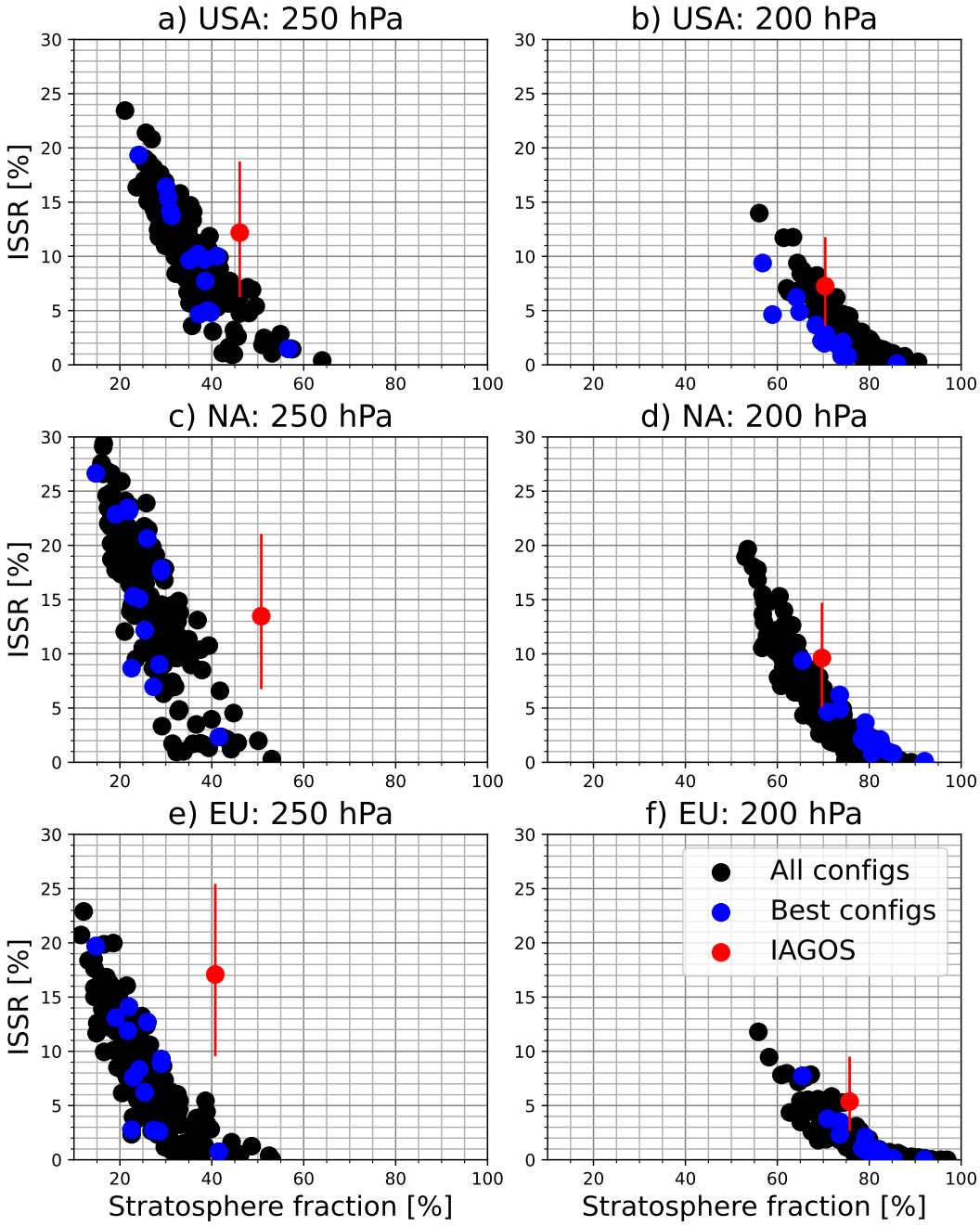

Figure 8. Relationship between the frequency of occurrence of ISSRs (in %) and the fraction (in %) of stratospheric air masses at 250 hPa, a) in the USA, c) in The North Atlantic (NA) and in e) Europe (EU). b,d,f) Same as a,c,e) but for 200 hPa. The fraction of values in the stratosphere are defined as the frequency of measurements collocated with an ozone concentration higher than 100 ppbv. This fraction in LMDZ is computed as the average of a 0-1 flag that identifies stratospheric gridboxes according to a temperature gradient criterion. The vertical red bar corresponds to the uncertainty in the frequency of occurrence of ISSRs in IAGOS. The best 15 selected configurations are shown in the blue-filled dots.

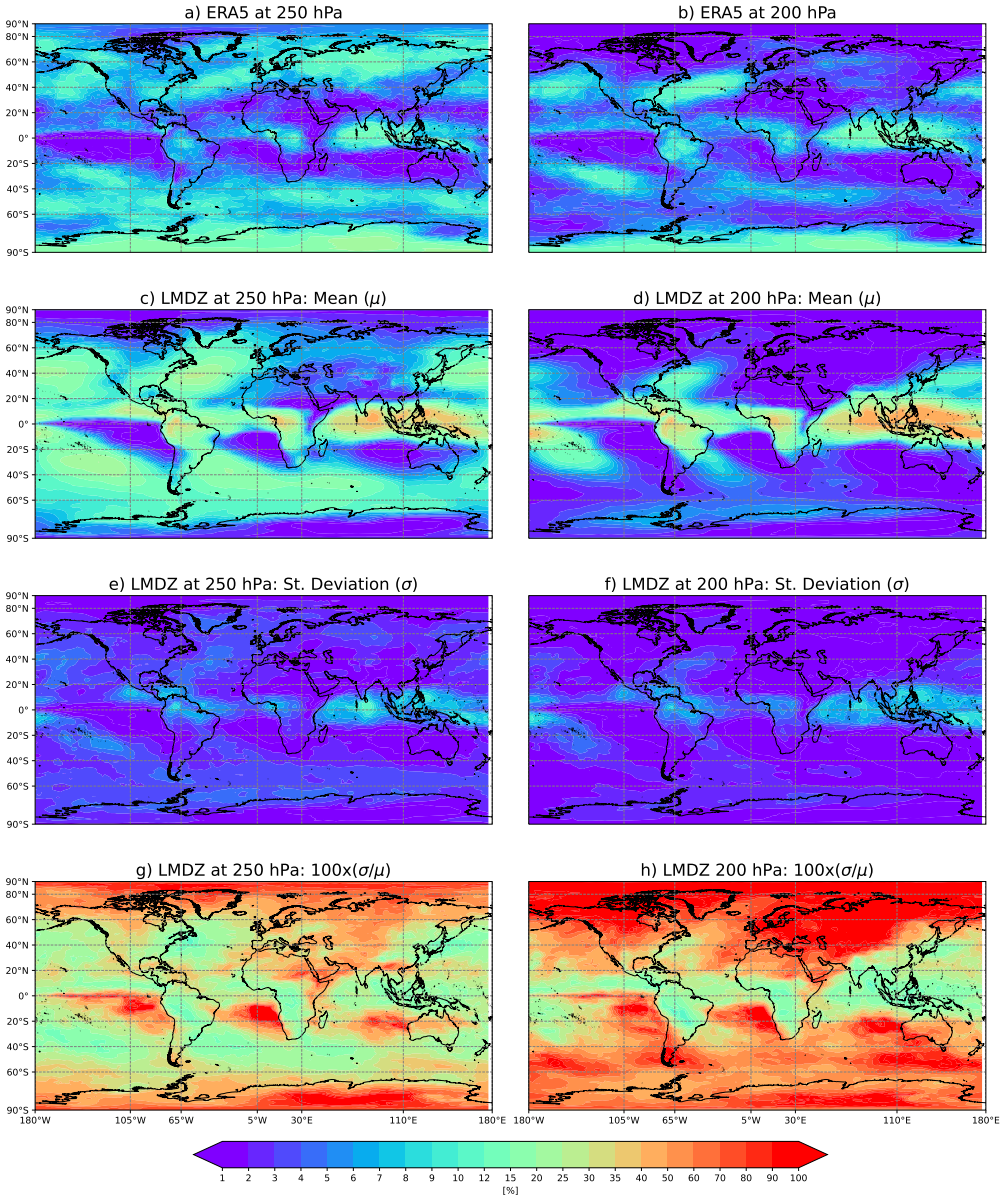

Figure 9. a–b) Frequency of occurrence of ISSRs (in %) in ERA5 averaged over the period 2018-2021. c-d) Same as a-b) but for the ensemble of the 15 selected LMDZ configurations, e-f) standard deviation (in %) of the ensemble, g-h) standard deviation normalised by the ensemble mean (in %).

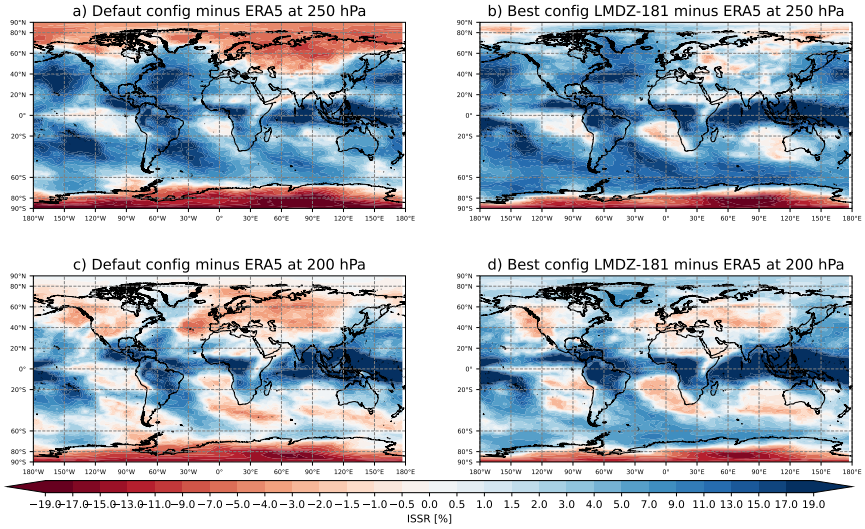

Figure 10. a–c) Bias in ISSR frequency (in %) for the LMDZ simulation using the First Guess (FG) vector of parameter, computed as the difference between LMDZ and ERA5. b–d) Corresponding bias for the best-performing configuration (181).

increases with increasing RHi for supersaturated and sub-saturated conditions in the UT (see the caption of Fig 11 for the definition), respectively (Fig. 11a-f). In the LS (defined in the caption of Fig 11) , the PDF decreases for both subsaturated (for 10 % < RHi < 100 %) and supersaturated conditions with a break in the slope around 100 % (Fig. 11a-f). Interestingly, LMDZ-181 captures these characteristics even over the USA, although no RHi metric is used in the tuning process (Fig. 11a-f). However, there are significant differences with IAGOS. At 250 hPa, the simulated PDFs show a break in the slope at 120 % in the NA in the PDF of RHi under ice-supersaturated conditions in the UT. At 200 hPa, in the UT, the PDF of RHi > 100% is more consistent with IAGOS than at 250 hPa, but the exponentially decreasing trend is overestimated. Overall, the agreement between LMDZ and IAGOS on RHi PDF properties seems to be better in the NA. It should also be noted that stratospheric and tropospheric air masses are not separated by the same criteria in IAGOS and LMDZ, so the comparisons have some limitations as well. But it's worth noting LMDZ's inability to simulate RHi higher than 130% with this first-order parameterization.

Second, we examine how these LMDZ configurations reproduce the total PDF of RHi (Fig. 11g-l). We find that ERA5 overestimates the decreasing slope of the PDF of RHi > 100 %. It also overestimates the PDF for RHi around 100 %. The LMDZ configuration shows a similar bias over the USA and EU, but to a lesser extent than ERA5. Note that the peak around 100 % in LMDZ is more associated with cloudy conditions. This peak is present in IAGOS (at the 4-s resolution) (Petzold et al., 2020) but to a lesser extent than in LMDZ and ERA5, probably due to low sampling of clouds by the IAGOS aircraft (Sanogo et al., 2024). The LMDZ-070 and LMDZ-150 configurations perform like the LMDZ 181 (Figs. 11, S8-S9), while the other selected configurations reproduce the RHi properties only in the NA (not shown).

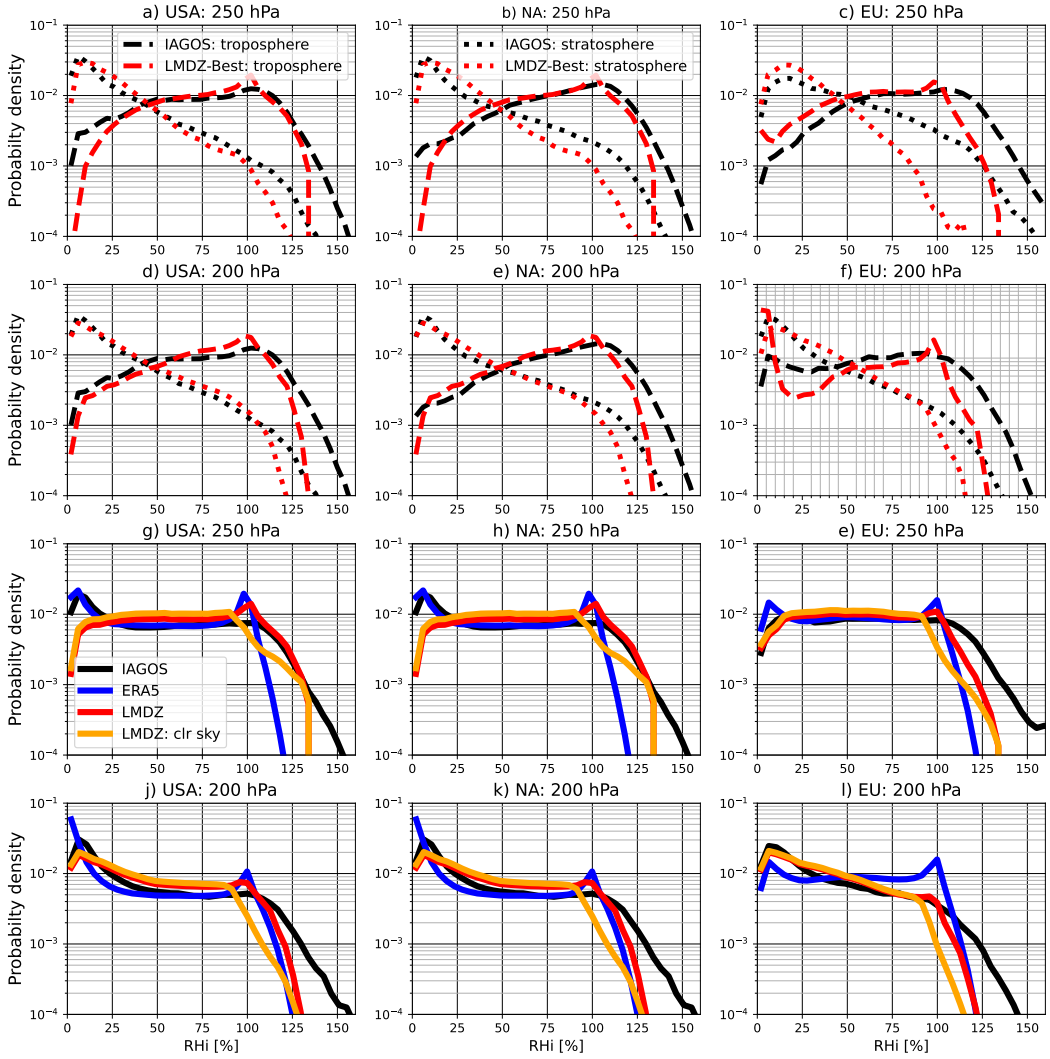

Figure 11. a-f) Probability density function in the lower stratosphere (dotted), the upper troposphere (dashed) of RHi (in %) in IAGOS (black) and LMDZ-181 (red): a-d) in the USA, b-e) in the North Atlantic (NA) and c-f) in Europe (EU). Troposphere and stratospheric measurements are discriminated in IAGOS using the ozone mixing ratio of 100 ppb. The RHi values in LMDZ in the stratosphere are defined as those associated with a probability of presence in the stratosphere higher than 0.6 % and those in the troposphere are defined as those with a probability of presence in the stratosphere of 0.6 %. g-l) Same as a-f) but for the total RHi IAGOS, ERA5 and in LMDZ-181. The PDF in orange is the PDF of RHi in the LMDZ, where the clear sky conditions (ice-subsaturated and -supersaturated conditions) prevail (fraction of grid boxes covered by the cloud, less than 10%).

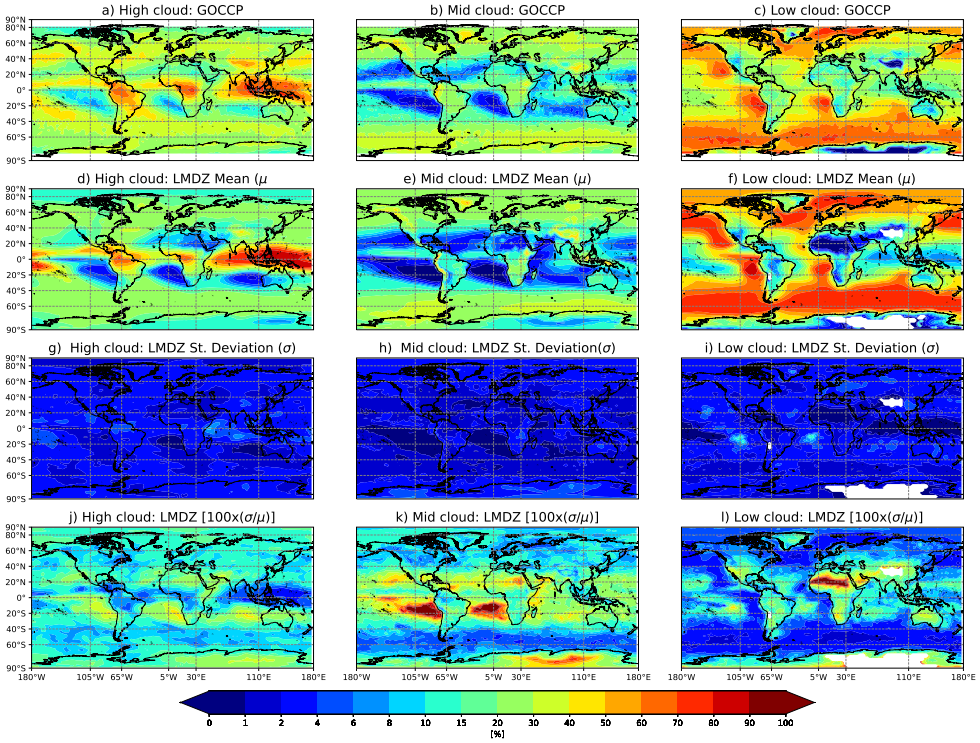

Figure 12. a-c) High, mid and low-level cloud cover (in %) from CALIPSO-GOCCP (averaged over the period 2006-2009). Statistics from the 15 selected configurations of LMDZ: d-f) mean, d-i) standard deviation, j-l) standard deviation normalized by the mean (in %), for high, mid and low cloud cover.

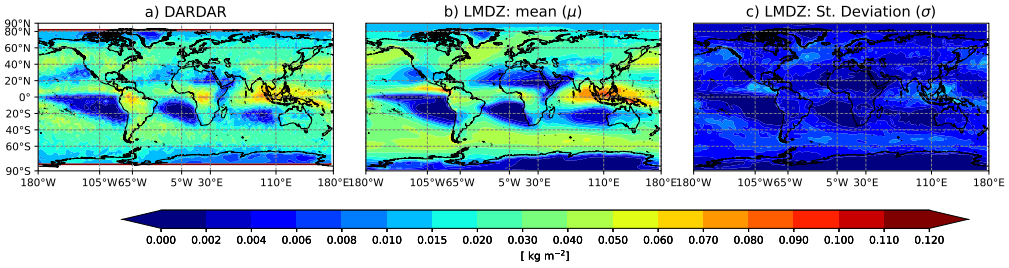

Figure 13. Same as Fig.12a-c) but for the ice water path (in kg m^{-2}).

5.3.3 Cloud cover and the IWP

In this section we explore the strengths and weaknesses of the 15 selected LMDZ configurations in simulating cloud properties (frequency, spatial distribution and ice content) on a global scale, using the CALIPSO-GOCCP dataset as a reference. This observational product shows that high-level cloud cover is larger in the tropical region, where strong updrafts in the intertropical convergence zone transport moisture into the upper troposphere (Fig. 12a). High and mid-level cloud covers are lowest some oceanic regions in 0-40°N/S latitudal bands and over the Sahara (Figs. 12a-b). The low-level cloud cover is higher over the oceans, except in the tropical deep convection region (Fig. 12c). Interestingly, the 15 selected configurations capture all these features of the cloud cover (Figs. 12a-f). The differences varies from 0.01 to 0.08 (i.e., 1 to 8 %), corresponding to 6 to 50 % of the ensemble mean, except in some areas where it is higher for the low and mid-level clouds (Figs. 12c-l). The good performance of these configurations on the spatial distribution of cloud cover is consistent with the results of recent work by Madeleine et al. (2020) who showed that this version of LMDZ is improved in that respect compared to the previous version. However, there are some biases in the magnitude of the cloud cover. However, for all configurations, these discrepancies with CALIPSO-GOCCP are within the uncertainty of the reference product for high and mid-level clouds (±5%), except for some areas in the tropics (Atlantic, Indian and Pacific Oceans, Sahel, South America) and in the mid-latitudes, in particular over China (not shown). The biases are more prominent for low-level cloud. A synthetic view of the analysis is provided in Fig. 14, which illustrates that the discrepancies of the configurations with CALIPSO-GOCCP are higher in the mid-latitudes between 40 and 70°N/S and overall, LMDZ-FG behaves similarly. These low cloud deficiencies may be error-compensating, since only the high cloud fraction is constrained to obtain realistic total radiative effects, and may require adding specific targets on low-level clouds in future work. Taking into account the large

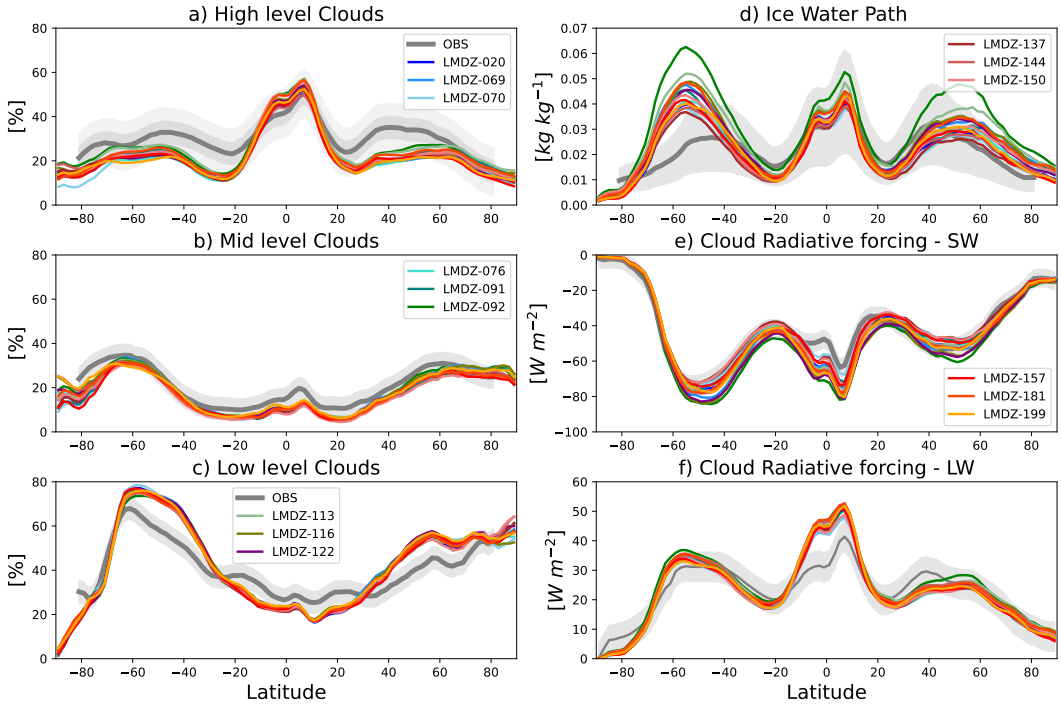

Figure 14. a-c) Zonal mean cloud cover (in %) simulated by the 15 selected configurations of LMDZ compared with the CALIPSO-GOCCP climatology (in grey). d) Same as a-c) but for the IWP (in kg m^{-2}) with the DARDAR observational product in grey. e-f) Same as a-c) but for the top-of-the-atmosphere (TOA) SW and LW cloud radiative effect (CRF, in W m^{-2}) with CERES-EBAF observations in grey. Note that the legend is split across the panels.

differences among cloud observation products (Chepfer et al., 2010), and given the results of Fig.14 and the bias presented in Figs. S10-S12, we conclude that these LMDZ configurations perform well. It should also be the biases of the configurations on the IWP is higher in regions beyond 45-80°N/S of latitudes (Figs. 14-13, S13). It is also worth noting that the configurations LMDZ-070, LMDZ-150 and LMDZ-181 show similar or lesser deviations from observations than the other configurations, illustrating that ISSRs can be tuned in conjunction with cloud cover, IWP and radiation (see Sect. 5.3.4).

5.3.4 Cloud Radiative Effects

The net cloud radiative effect is assessed in the 15 selected configurations of LMDZ (Fig. 15). The CRE is highest in the intertropical convergence zone and over the oceans in the mid-latitudes. In these areas, where cloud cover and IWP are most pronounced, the CRE is of the order of 60 to 80 Wm^{-2} in the shortwave and 20 to 50 Wm^{-2} in the longwave. Overall, the net radiative forcing of clouds on Earth's climate is a cooling effect that can reach 40 Wm^{-2}, except in a few places, the largest of which are Antarctica, Greenland, West and North Africa, and the Middle East, where there is a warming effect (Fig. 15a-c). The 15 configurations, in addition to LMDZ-FG are consistent with each other and with observations of the spatial distribution of the CRE in both the shortwave and longwave parts of the spectrum (Figs. 15). Each configuration simulates the magnitude of the CRE in agreement with the observation. Their biases are within the uncertainty of ±5 Wm^{-2} of the reference product, except in some regions that correspond to higher bias on the high-level cloud in particular in the tropics (Fig. 15). In

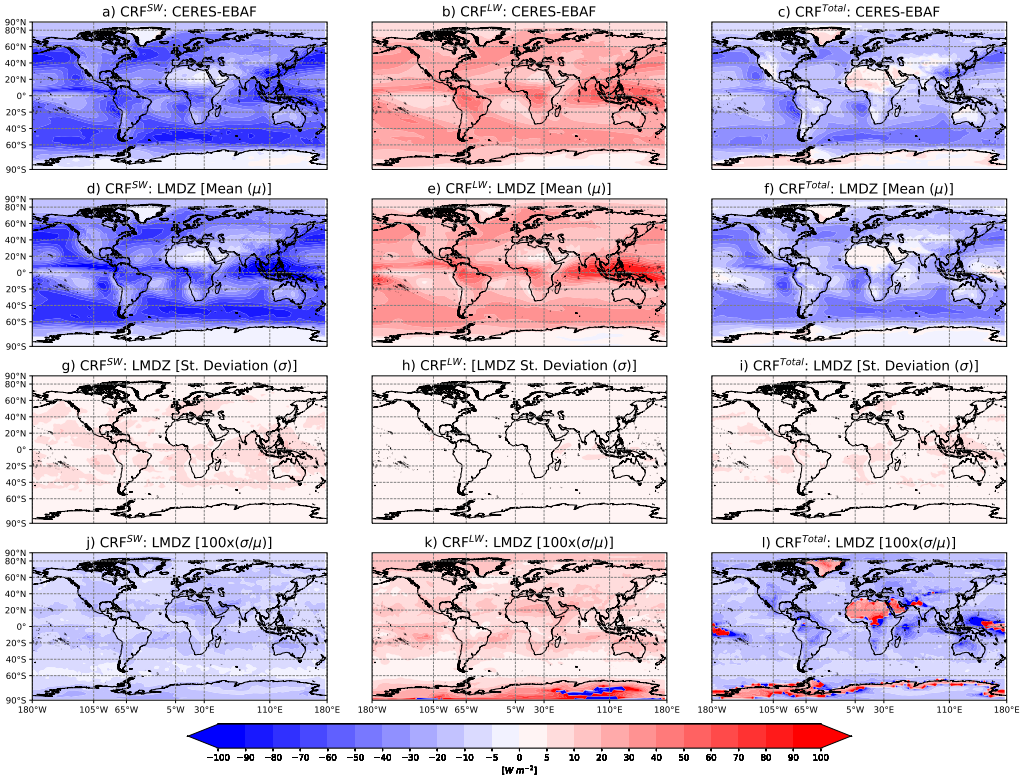

Figure 15. Same as Fig. 12 but for the cloud radiative effect (in W m^{-2}) in the short and longwave radiation. The CERES-EBAF product is used as reference.

the zonal mean, it is only in the tropical region that all configurations are outside the uncertainty range. It should be noted that, as for the cloud fraction, the dispersion of the CRE across the 15 configurations is rather small. The good agreement between the 15 configurations for the cloud fraction and their radiative effects, while there are deficiencies for the frequency of occurrence of ISSRs, especially at 200 hPa, is worth noting. This is due to the fact that the pressure region around 200 hPa is dominated by dry LS air masses (Fig. 8) where only thin cirrus clouds can form (Spang et al., 2024; Zou et al., 2020), some of which are not detected as cirrus clouds (with $\tau < 0.07$) in LMDZ due to their low optical thickness.

6 Conclusions

Ice supersaturation is an essential feature of the upper atmosphere and a prerequisite for the formation of in situ cirrus clouds and contrails, which have important radiative effects and forcing. A parameterization of ice supersaturation has been developed and implemented in the atmospheric component (LMDZ) of the IPSL climate model (Section 2.2). We have used a statistical tool based on Gaussian process and history matching with iterative refocusing to explore the tunability of this ice supersaturation parameterization together with the pre-existing LMDZ large-scale condensation and convection parameterizations, with the aim of obtaining configurations that perform well in terms of ice supersaturation in the Northern Hemisphere but also other large-scale cloud properties without degrading the simulated global climate. For this purpose, we used as tuning constraints a range of regional metrics based on the distribution of relative humidity relative to ice, the cloud cover, the ice water path, radiative fluxes at the top of the atmosphere, and surface precipitation. LMDZ suffers from a temperature bias in the up-

per troposphere that is independent of the ISS parameterization, which we alleviated by applying an online bias correction method. The study identified three configurations, among 15 LMDZ configurations selected based on an aggregated ranking score, that reproduce several aspects of the ISSRs observed by IAGOS and simulated by ERA5.

- The three configurations reproduce the frequency of occurrence of ISSRs at 250 hPa in the mid-latitudes of the Northern Hemisphere, from the United States to Europe consistently with ERA5 and on board aircraft observations IAGOS. They qualitatively reproduce almost all the facets of the PDF of RHi in the upper troposphere and lower stratosphere.
- The three configurations in addition to most of the 15 are also consistent with the CALIPSO-GOCCP high-level cloud cover, except in the tropical Pacific region where they have relatively high biases. They are in agreement with CALIPSO-GOCCP on the mid-level cloud cover. Cloud ice water path is also fairly well reproduced except in 50-80°N/S regions, mainly over oceans where they exhibit relatively high biases. They also simulate the radiative effects of clouds consistent with observations, except in some areas located mainly in the tropics.
- These configurations suggest that ISSRs can have a strong horizontal extension to cover 40 to 80% of a gribox of ∼1.2°x2.5° of resolution.

The study showed that the ice supersaturation parameterization can be tuned in LMDZ. It also highlighted structural errors in LMDZ that may be due to the ISS parameterization or to the host model itself.

- LMDZ fails to reproduce the distribution of the frequency of occurrence of ISSRs in the mid-latitudes of the Northern Hemisphere at 200 hPa: it exhibits large biases over Eurasia and the USA. Configurations consistent with observations at 200 hPa on the frequency of occurrence of ISSRs in these locations overestimate the frequency of occurrence of ISSRs in the North Atlantic at 250 hPa. These shortcomings of LMDZ are partly related to its representation of tropopause, particularly in Europe, where it is too low. Increasing the vertical resolution of the model is among the potential solutions to explore to address this issue, in addition to a more explicit representation of the physical processes of ISSRs in the parameterization.
- LMDZ also fails to reproduce the spatial distribution of the frequency of occurrence of ISSRs in the high-latitudes of the Southern Hemisphere.

Finally, our study shows that reproducing the spatial and temporal variability of tropopause height is a necessary condition for a climate model to accurately represent the occurrence frequency of ISSRs, especially in the mid-latitudes at 250 and 200 hPa, which corresponds to the pressure levels where contrails are most frequent due to the high density of flights. This first version of the ISS parameterization in LMDZ, is based solely on homogeneous nucleation, necessarily has other limitations. The homogeneous nucleation process requires higher levels of ice supersaturation than heterogeneous nucleation to initiate (Heymsfield et al., 2017). Therefore, ignoring heterogeneous nucleation may introduce biases in both the simulated cloud cover (Muench & Lohmann, 2020) and the distribution of the precursor ISSR statistics. In addition, heterogeneous nucleation generally results in larger ice crystals (Heymsfield, 2017). Caution should be exercised when using this parameterisation to study the radiative effects of cirrus clouds or contrails, given that the heterogeneous nucleation was not considered in this study.

A new version of the ISS parameterization is developed (see Borella et al. (2025)), which addressed some of the shortcomings, such as the biases in the frequency of ISSRs over Europe. We consider the tuning methodology presented here to be a necessary prerequisite for evaluating this new version of the parameterization which will be used to quantify the contrail radiative forcing and the associated parametric uncertainty in LMDZ.

Data and Code Availability Statement

The latest version of the LMDZ source code can be freely downloaded from the LMDZ website (`https://lmdz.lmd.jussieu.fr`). The version used for this study is the SVN release from 23 September 2023. The MOZAIC and IAGOS data can be accessed through the IAGOS data portal (`https://doi.org/10.25326/20`). The DARDAR data are available upon request from Stengel et al. (2020). CERES EBAF data were obtained from the NASA Langley Research Center CERES ordering tool (`https://ceres.larc.nasa.gov/`). CALIPSO-GOCCP data were retrieved from the GOCCP website (`http://climserv.polytechnique.fr`). ERA5 data were sourced from the Copernicus Climate Data Store (`https://cds.climate.copernicus.eu/`). GPCP precipitation data were provided by the NOAA/OAR/ESRL PSD in Boulder, Colorado, USA, and can be accessed via their website (`https://www.esrl.noaa.gov/psd/`). All the programs, scripts, and reference LES used to perform the tuning are publicly available via a Subversion through "svn checkout `http://svn.lmd.jussieu.fr/HighTune`". The Python codes used for the analysis are available on the zenodo website (`https://doi.org/10.5281/zenodo.15076522`).

Acknowledgments

This research has been supported by the French Ministère de la Transition Ecologique et Solidaire (grant no. DGAC N2021-39), with support from France's Plan National de Relance et de Résilience (PNRR) and the European Union's NextGenerationEU. The LMDZ experiments were performed using the HPC resources of TGCC under the allocation 2023-A0140107732 (project gencmip6) as provided by GENCI (Grand Equipement National de Calcul Intensif). We would also like to thank Najda Villefranque for the technical support and Hélène Chepfer for the in-depth discussion on uncertainties in the CALIPSO-GOCCP cloud cover dataset.

References

Binder, H., Boettcher, M., Joos, H., Sprenger, M., & Wernli, H. (2020). Vertical cloud structure of warm conveyor belts – a comparison and evaluation of ERA5 reanalysis, CloudSat and CALIPSO data. *Weather and Climate Dynamics*, *1*(2), 577–595. doi: 10.5194/wcd-1-577-2020

Bock, L., & Burkhardt, U. (2016). Reassessing properties and radiative forcing of contrail cirrus using a climate model. *J. Geophys. Res.: Atmos.*, *121*(16), 9717–9736.

Bony, S., & Emanuel, K. A. (2001). A parameterization of the cloudiness associated with cumulus convection; evaluation using TOGA COARE data. *J. Atmos. Sci.*, *58*(21), 3158–3183. doi: 10.1175/1520-0469(2001)058⟨3158:APOTCA⟩2.0.CO;2

Borella, A., Vignon, É., Boucher, O., Meurdesoif, Y., & Fairhead, L. (2025). A new prognostic parameterization of subgrid ice supersaturation and cirrus clouds in the icolmdz agcm. *Journal of Advances in Modeling Earth Systems*, *17*(8), e2024MS004918. doi: 10.1029/2024MS004918

Borella, A., Vignon, É., Boucher, O., & Rohs, S. (2024). An empirical parameterization of the subgrid-scale distribution of water vapor in the utls for atmospheric general circulation models. *Journal of Geophysical Research: Atmospheres*, *129*(20), e2024JD040981. doi: 10.1029/2024JD040981

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... others (2013). 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* (pp. 571–657). Cambridge University Press.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., ... others (2020). Presentation and evaluation of the IPSL-CM6A-

LR climate model. *J. Adv. Mod. Earth Syst.*, *12*(7), e2019MS002010. doi: 10.1029/2019MS002010

Brown, A., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., ... others (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, *128*(582), 1075–1093. doi: 10.1256/003590002320373210

Chen, C.-C., & Gettelman, A. (2013). Simulated radiative forcing from contrails and contrail cirrus. *Atmos. Chem. Phys.*, *13*(24), 12525–12536. doi: 0.5194/acp-13-12525-2013

Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J., Minnis, P., ... Zeng, S. (2010). The GCM-oriented calipso cloud product (CALIPSO-GOCCP). *J. Geophys. Res.: Atmos.*, *115*(D4).

Coindreau, O., Hourdin, F., Haeffelin, M., Mathieu, A., & Rio, C. (2007). Assessment of physical parameterizations using a global climate model with stretchable grid and nudging. *Monthly weather review*, *135*(4), 1474–1489. doi: 10.1175/MWR3338.1

Couvreux, F., Guichard, F., Redelsperger, J.-L., Kiemle, C., Masson, V., Lafore, J.-P., & Flamant, C. (2005). Water-vapour variability within a convective boundary-layer assessed by large-eddy simulations and IHOP_2002 observations. *Quarterly Journal of the Royal Meteorological Society*, *131*(611), 2665–2693. doi: 10.1256/qj.04.167

Couvreux, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., ... others (2021). Process-based climate model development harnessing machine learning: I. A calibration tool for parameterization improvement. *J. Adv. Mod. Earth Syst.*, *13*(3), e2020MS002217. doi: 10.1029/2020MS002217

Genthon, C., Piard, L., Vignon, E., Madeleine, J.-B., Casado, M., & Gallée, H. (2017). Atmospheric moisture supersaturation in the near-surface atmosphere at Dome C, Antarctic Plateau. *Atmos. Chem. Phys.*, *17*(1), 691–704.

Gierens, K., & Brinkop, S. (2012). Dynamical characteristics of ice supersaturated regions. *Atmos. Chem. Phys.*, *12*(24), 11933–11942.

Gierens, K., Schumann, U., Helten, M., Smit, H., & Marenco, A. (1999). A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. *Annales Geophysicae*, *17*(9), 1218–1226.

Heymsfield, A., Krämer, M., Luebke, A., Brown, P., Cziczo, D., Franklin, C., ... Van Tricht, K. (2017). Cirrus clouds. *Meteorological Monographs*, *58*, 2.1-62.26. doi: 10.1175/AMSMONOGRAPHS-D-16-0010.1

Hourdin, F., Ferster, B., Deshayes, J., Mignot, J., Musat, I., & Williamson, D. (2023). Toward machine-assisted tuning avoiding the underestimation of uncertainty in climate change projections. *Science Advances*, *9*(29), eadf2758. doi: 10.1126/sciadv.adf2758

Hourdin, F., Găinusă-Bogdan, A., Braconnot, P., Dufresne, J.-L., Traore, A.-K., & Rio, C. (2015). Air moisture control on ocean surface temperature, hidden key to the warm bias enigma. *Geophysical Research Letters*, *42*(24), 10–885. doi: 10.1002/2015GL066764

Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., ... others (2013). LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, *40*, 2193–2222. doi: 10.1007/s00382-012-1343-y

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., ... others (2017). The art and science of climate model tuning. *Bull. Am. Meteorol. Soc.*, *98*(3), 589–602.

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., ... others (2020). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. *J. Adv. Mod. Earth Syst.*,

12(7), e2019MS001892. doi: 10.1029/2019MS001892

Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranque, N., ... Volodina, V. (2021). Process-based climate model development harnessing machine learning: II. Model calibration from single column to global. *J. Adv. Mod. Earth Syst.*, *13*(6), e2020MS002225. doi: 10.1029/2020MS002225

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., ... Susskind, J. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of hydrometeorology*, *2*(1), 36–50. doi: 10.1175/1525-7541(2001)002⟨0036:GPAODD⟩2.0.CO;2

Irvine, E. A., & Shine, K. P. (2015). Ice supersaturation and the potential for contrail formation in a changing climate. *Earth System Dynamics*, *6*(2), 555–568.

Jebeile, J., Lam, V., Majszak, M., & Räz, T. (2023). Machine learning and the quest for objectivity in climate model parameterization. *Climatic change*, *176*(8), 101. doi: 10.1007/s10584-023-03532-1

Joos, H., Sprenger, M., Binder, H., Beyerle, U., & Wernli, H. (2023). Warm conveyor belts in present-day and future climate simulations – part 1: Climatology and impacts. *Weather and Climate Dynamics*, *4*(1), 133–155. doi: 10.5194/wcd-4-133-2023

Kanji, Z. A., Ladino, L. A., Wex, H., Boose, Y., Burkert-Kohn, M., Cziczo, D. J., & Krämer, M. (2017). Overview of ice nucleating particles. *Meteorological Monographs*, *58*, 1–1.

Kärcher, B. (2018). Formation and radiative forcing of contrail cirrus. *Nat. Commun.*, *9*, 1824.

Kärcher, B., & Burkhardt, U. (2008). A cirrus cloud scheme for general circulation models. *Q. J. R. Meteorol. Soc*, *134*(635), 1439–1461. doi: 10.1002/qj.301

Koop, T., Luo, B., Tsias, A., & Peter, T. (2000). Water activity as the determinant for homogeneous ice nucleation in aqueous solutions. *Nature*, *406*(6796), 611–614. doi: 10.1038/35020537

Krämer, M., Rolf, C., Luebke, A., Afchine, A., Spelten, N., Costa, A., ... others (2016). A microphysics guide to cirrus clouds–part 1: Cirrus types. *Atmos. Chem. Phys.*, *16*(5), 3463–3483. doi: /10.5194/acp-16-3463-2016

Krämer, M., Schiller, C., Afchine, A., Bauer, R., Gensch, I., Mangold, A., ... others (2009). Ice supersaturations and cirrus cloud crystal numbers. *Atmos. Chem. Phys.*, *9*(11), 3505–3522.

Krinner, G., Kharin, V., Roehrig, R., Scinocca, J., & Codron, F. (2020). Historically-based run-time bias corrections substantially improve model projections of 100 years of future climate change. *Communications Earth & Environment*, *1*(1), 29. doi: doi.org/10.1038/s43247-020-00035-0

Lamquin, N., Stubenrauch, C., Gierens, K., Burkhardt, U., & Smit, H. (2012). A global climatology of upper-tropospheric ice supersaturation occurrence inferred from the Atmospheric Infrared Sounder calibrated by MOZAIC. *Atmos. Chem. Phys.*, *12*(1), 381–405.

Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., ... others (2021). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmos. Env.*, *244*, 117834.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., ... Wong, T. (2009). Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *Journal of Climate*, *22*(3), 748–766. doi: 10.1175/2008JCLI2637.1

Lohmann, U., Spichtinger, P., Jess, S., Peter, T., & Smit, H. (2008). Cirrus cloud formation and ice supersaturated regions in a global climate model. *Env. Res. Lett.*, *3*(4), 045022.

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., ... others (2020). Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model. *J. Adv. Mod. Earth Syst.*, *12*(4), e2019MS001940. doi:

10.1029/2019MS001940

Madeleine, J.-B., Hourdin, F., Grandpeix, J.-Y., Rio, C., Dufresne, J.-L., Vignon, E., ... others (2020). Improved representation of clouds in the atmospheric component LMDZ6A of the IPSL-CM6A Earth system model. *J. Adv. Mod. Earth Syst.*, *12*(10), e2020MS002046.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., ... others (1998). Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, An overview. *J. Geophys. Res.: Atmos.*, *103*(D19), 25631–25642. doi: 10.1029/98JD00977

Minnis, P., Ayers, J. K., Palikonda, R., & Phan, D. (2004). Contrails, cirrus trends, and climate. *Journal of Climate*, *17*(8), 1671–1685. doi: 10.1175/1520-0442(2004)017⟨1671:CCTAC⟩2.0.CO;2

Muench, S., & Lohmann, U. (2020). Developing a cloud scheme with prognostic cloud fraction and two moment microphysics for echam-ham. *JAMES*, *12*(8), e2019MS001824. doi: 10.1029/2019MS001824

Perini, M., Terray, L., Cariolle, D., Peatier, S., & Moine, M.-P. (2023). Parametric sensitivity and constraint of contrail cirrus radiative forcing in the atmospheric component of CNRM-CM6-1. *EGUsphere*, *2023*, 1–18.

Petzold, A., Neis, P., Rütimann, M., Rohs, S., Berkes, F., Smit, H. G., ... others (2020). Ice-supersaturated air masses in the northern mid-latitudes from regular in situ observations by passenger aircraft: vertical distribution, seasonality and tropospheric fingerprint. *Atmos. Chem. Phys.*, *20*(13), 8157–8179.

Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A., Gallagher, M., ... others (2015). Global-scale atmosphere monitoring by in-service aircraft–current achievements and future prospects of the european research infrastructure iagos. *Tellus*, *67B*(1), 28452.

Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, *48*(2), 88–91. doi: 10.1080/00031305.1994.10476030

Raillard, L., Vignon, É., Rivière, G., Madeleine, J.-b., Meurdesoif, Y., Delanoë, J., ... others (2024). Leveraging rali-thinice observations to assess how the icolmdz model simulates clouds embedded in arctic cyclones. *Journal of Geophysical Research: Atmospheres*, *129*(16), e2024JD040973. doi: 10.1029/2024JD040973

Ren, C., & Mackenzie, A. (2005). Cirrus parametrization and the role of ice nuclei. *Quarterly Journal of the Royal Meteorological Society*, *131*(608), 1585–1605. doi: doi:10.1256/qj.04.126

Reutter, P., Neis, P., Rohs, S., & Sauvage, B. (2020). Ice supersaturated regions: properties and validation of ERA-Interim reanalysis with IAGOS in situ water vapour measurements. *Atmos. Chem. Phys.*, *20*(2), 787–804.

Sandu, I., & Stevens, B. (2011). On the factors modulating the stratocumulus to cumulus transitions. *J. Atmos. Sci.*, *68*(9), 1865–1881. doi: 10.1175/2011JAS3614.1

Sanogo, S., Boucher, O., Bellouin, N., Borella, A., Wolf, K., & Rohs, S. (2024). Variability in the properties of the distribution of the relative humidity with respect to ice: implications for contrail formation. *Atmos. Chem. Phys*, *24*(9), 5495–5511. doi: 10.5194/acp-24-5495-2024

Smit, H. G., Rohs, S., Neis, P., Boulanger, D., Krämer, M., Wahner, A., & Petzold, A. (2014). Reanalysis of upper troposphere humidity data from the MOZAIC programme for the period 1994 to 2009. *Atmos. Chem. Phys.*, *14*(23), 13241–13255.

Spang, R., Müller, R., & Rap, A. (2024). Radiative effect of thin cirrus clouds in the extratropical lowermost stratosphere and tropopause region. *Atmos. Chem. Phys.*, *24*(2), 1213–1230. doi: 10.5194/acp-24-1213-2024

Spichtinger, P., Gierens, K., & Dörnbrack, A. (2005). Formation of ice supersaturation by mesoscale gravity waves. *Atmos. Chem. Phys.*, *5*(5), 1243–1255. doi:

10.5194/acp-5-1243-2005

Spichtinger, P., Gierens, K., Leiterer, U., & Dier, H. (2003). Ice supersaturation in the tropopause region over Lindenberg, Germany. *Meteorologische Zeitschrift*, *12*(3), 143–156.

Spichtinger, P., Gierens, K., & Read, W. (2003). The global distribution of ice-supersaturated regions as seen by the microwave limb sounder. *Q. J. Roy. Meteorol. Soc.,*, *129*(595), 3391–3410. doi: 10.1256/qj.02.141

Spichtinger, P., Gierens, K., & Wernli, H. (2005). A case study on the formation and evolution of ice supersaturation in the vicinity of a warm conveyor belt's outflow region. *Atmos. Chem. Phys.*, *5*(4), 973–987. doi: 10.5194/acp-5-973-2005

Stengel, M., Stapelberg, S., Sus, O., Finkensieper, S., Würzler, B., Philipp, D., ... McGarragh, G. (2020). Cloud_cci Advanced Very High Resolution Radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties. *Earth System Science Data*, *12*(1), 41–60. doi: 10.5194/essd-12-41-2020

Stubenrauch, C. J., & Schumann, U. (2005). Impact of air traffic on cirrus coverage. *Geophysical Research Letters*, *32*(14). doi: 10.1029/2005GL022707

Tompkins, A. M., Gierens, K., & Rädel, G. (2007). Ice supersaturation in the ecmwf integrated forecast system. *Quarterly Journal of the Royal Meteorological Society*, *133*(622), 53–63. doi: 10.1002/qj.14

VanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A., Burnet, F., ... others (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *J. Adv. Mod. Earth Syst.*, *3*(2). doi: 10.1029/2011MS000056

Vignon, É., Raillard, L., Genthon, C., Del Guasta, M., Heymsfield, A. J., Madeleine, J.-B., & Berne, A. (2022). Ice fog observed at cirrus temperatures at Dome C, Antarctic Plateau. *Atmos. Chem. Phys.*, *22*(19), 12857–12872. doi: 10.5194/acp-22-12857-2022

Wang, M., & Penner, J. E. (2010). Cirrus clouds in a global climate model with a statistical cirrus cloud scheme. *Atmospheric Chemistry and Physics*, *10*(12), 5449–5474. doi: 10.5194/acp-10-5449-2010

Wilhelm, L., Gierens, K., & Rohs, S. (2022). Meteorological conditions that promote persistent contrails. *Applied Sciences*, *12*(9), 4450.

Williamson, D., Blaker, A. T., Hampton, C., & Salter, J. (2015). Identifying and removing structural biases in climate models with history matching. *Climate dynamics*, *45*(5), 1299–1324.

Williamson, D. B., Blaker, A. T., & Sinha, B. (2017). Tuning without over-tuning: parametric uncertainty quantification for the NEMO ocean model. *Geoscientific Model Development*, *10*(4), 1789–1816. doi: 10.5194/gmd-10-1789-2017

Wolf, K., Bellouin, N., & Boucher, O. (2023). Long-term upper-troposphere climatology of potential contrail occurrence over the Paris area derived from radiosonde observations. *Atmos. Chem. Phys.*, *23*(1), 287–309. doi: 10.5194/acp-23-287-2023

Wolf, K., Bellouin, N., & Boucher, O. (2024). Correction of era5 temperature and relative humidity biases by bivariate quantile mapping for contrail formation analysis5. *Atmospheric Chemistry and Physics*, *24*(8), 5009–5024. doi: doi.org/10.5194/acp-24-5009-2024

Wolf, K., Bellouin, N., Boucher, O., Rohs, S., & Li, Y. (2025). Correction of temperature and relative humidity biases in era5 by bivariate quantile mapping: Implications for contrail classification. *EGUsphere*, *25*, 157–181. doi: doi.org/10.5194/acp-25-157-2025

Zhou, C., Dessler, A., Zelinka, M., Yang, P., & Wang, T. (2014). Cirrus feedback on interannual climate fluctuations. *Geophysical Research Letters*, *41*(24), 9166–9173. doi: 10.1002/2014GL062095

Zou, L., Griessbach, S., Hoffmann, L., Gong, B., & Wang, L. (2020). Revisiting

global satellite observations of stratospheric cirrus clouds. *Atmos. Chem. Phys.*, *20*(16), 9939–9959. doi: 10.5194/acp-20-9939-2020