https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Prognostic modeling of total specific humidity variance induced by shallow convective clouds in a GCM

Louis d'Alençon[1], Frédéric Hourdin[1], and Catherine Rio[2]

[1]LMD/IPSL/Sorbonne Université, CNRS, Paris 75005, France
[2]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Correspondence: Louis d'Alençon (louis.dalencon@lmd.ipsl.fr)

Abstract.

Shallow convective cloud cover prediction is a critical element of climate modeling. In most global climate models (GCMs), the cloud scheme relies on a statistical description of the atmospheric water content. While the mean value of specific humidity is a standard output of the climate model, the upper moments, in particular the variance and skewness, must be prescribed by a dedicated model. Observations and large eddy simulations (LES) have shown that the asymmetry of water distribution is linked to the presence of organized convective cells. To capture this asymetry, the boundary layer convective cells are represented in this work by an eddy diffusivity mass flux model, coupled to a bi-Gaussian statistical cloud scheme. Previously, the variance of each component has been prescribed diagnostically from the difference in specific humidity between updrafts and environment. This cloud scheme has demonstrated its capacity to represent dry shallow convective boundary layers, however, it reveals significant inaccuracies in deep convection situations and in the prediction of skewness. We propose here to develop a prognostic model for the total specific humidity variance. We show in particular that the transport of specific humidity and its square in the mass-flux scheme permits to implement the generic prognostic equations of the variance into a GCM. Furthermore, model adjustment is carried out using recently added automatic tuning methods. Based on this adjustment, we show that the prognostic model is consistent with previous work on shallow convective cases while providing significant improvements in the description of variance and third-order moment profiles.

1 Introduction

The importance of cloud representation in the radiative budget of the earth is widely documented (Bony et al., 2006; Stowasser et al., 2006; Zelinka et al., 2017) which has led climate modelers to integrate increasingly accurate cloud schemes in global climate models. Among these cloud schemes, a classical approach consists of building a sub-grid distribution of the (total or relative) humidity or of another thermodynamic quantity such as the saturation deficit (Mellor, 1977; Sommeria and Deardorff, 1977; Jam et al., 2013), which allows to characterize a saturated fraction of the mesh defining the cloud fraction. These distributions were gradually refined in their form, they were first symmetrical (triangle in Smith (1990), rectangle in Le Trent and Li (1991)) before observation and LES studies have shown the importance of asymmetrical distribution especially in the convective boundary layer with cumulus clouds (Lewellen and Yoh, 1993; Larson et al., 2002). Therefore it became essential

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

to capture this asymmetry leading to various PDF forms (log-normal in Bony and Emanuel (2001), beta in Tompkins (2002) or bigaussian in Golaz et al. (2002b) or Jam et al. (2013)) with non-zero skewness. To determine the parameters of these asymmetric PDFs, the modeling of the upper moments of the distribution has also evolved, first a priori (Smith, 1990; Le Trent and Li, 1991) then diagnosed (Bony and Emanuel, 2001; Jam et al., 2013) from thermodynamical variables or prognosed (Tompkins, 2002; Watanabe et al., 2009) with various closure hypothesis. But whether they are established diagnostically or prognostically, it has been necessary to develop a good understanding of the physical sources of these distribution moments, in particular to close the higher-order terms related to turbulent transport.

To study the turbulent source terms of humidity variance and higher moments, the modeling of turbulent transport in GCMs is therefore a central context element. Higher-order turbulent closure models have been proposed in Mellor and Yamada (1982) with assumption of downgradient diffusion for the turbulent fluxes. Though, alongside turbulent diffusion, the importance of convective structures carrying out non-local countergradient transport has long been established (Deardorff, 1966). In fact, this non-local transport is a key element in explaining the asymmetry of the distributions that was highlighted before. Several approaches exist to represent this non-local transport by using for example a joint Bigaussian PDF of velocity, liquid potential temperature and total specific humidity with higher-order closure (Golaz et al., 2002a, b; Larson et al., 2002). In this work we place ourselves within the framework of so-called mass flux schemes to represent non-local transport by organized convective plumes. The association of a mass flux scheme with a classical K-diffusion theory modeling small-scale turbulence (Mellor and Yamada, 1982; Yamada, 1983), has led to the “eddy diffusivity mass flux” (EDMF) models initially developed in Chatfield and Brost (1987) and widely disseminated since (Köhler et al., 2011; Hourdin et al., 2002; Rio and Hourdin, 2008; Pergaud J. et al., 2009; Neggers, 2009).

In LMDZ, the EDMF scheme is associated with a bi-Gaussian statistical model of total specific humidity to determine cloud cover in the lower atmosphere. The humidity variance within each Gaussian component is diagnosed from the specific humidity difference between the ascending plumes and the environment, while the asymmetry of the distribution is related to the relative weight of each Gaussian, given by the surface fraction of the updrafts. This cloud scheme has proven to be very effective in representing dry boundary layers with cloud formation at the tops of ascending plumes so as stratocumulus and transition scenes (Jam et al., 2013; Hourdin et al., 2019, 2021; Madeleine et al., 2020). However it fails to represent some sources of variance, related to subsidence, intensity of detrainment or precipitating downdrafts for example, which may bias the cloud representation, especially in cases of deep convection. The importance of the detrainment terms of convective air masses, in particular, was demonstrated in Klein et al. (2005) (thereafter K05) based on CRM studies, after establishing the corresponding source terms in the prognostic equation of variance. Rather than adding ad-hoc diagnostic terms of variance with more free parameters, our main objective in this work is to present the implementation of K05’s prognostic approach within a GCM, which has not been achieved yet. In fact K05 doubted the feasibility of implementing his developments in large-scale models because of the lack of information on the variance in the thermal plume scheme. Relying on an original methodology to circumvent this difficulties by transporting the specific humidity and its square within the thermal plume scheme, we propose

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

here to integrate into a GCM the complete coupling of the mass flux scheme to the statistical cloud scheme via a prognostic variance equation. This approach allow us to build up a unified vision of the variance evolution and its sources with very limited free parameters. We focus here on the implementation of this prognostic model in the context of shallow convection. This first step is important to lay the theoretical and methodological foundations of the new model and to compare it to the pre-existing model and to large eddy simulations (LES) in cumulus and stratocumulus scenes for which a very solid tuning strategy has been developed (Couvreux et al., 2021; Hourdin et al., 2021). This tuning process employs statistical learning tools based on Bayesian inference models. The principle is to use a set of 1D simulations of the model to produce a gaussian statistical emulator of some specific metrics over the hole space of parameters. These metrics are then compared to LES simulations which makes it possible to constrain a subspace of satisfactory parameters.

The paper is organized as follows : Sect. 2 presents a description of the LMDZ model setup and the methods used to develop and calibrate the prognostic parameterization, Sect. 3 presents the theoretical framework in which the prognostic model is developed, in particular the method leading to its integration into the convective transport scheme. Section 4 documents the results obtained after tuning the free parameters of the prognostic model, especially concerning cloud cover and specific humidity variance and skewness as well as comparisons with the previous model and discussions on the different sources of variability.

2 Model and methods

In this part, we present the LMDZ model as well as the experimental methodologies and the principle of semi-automatic tuning used to calibrate and develop the prognostic physical parameterization.

2.1 The LMDZ model

The parameterization presented in this paper are tested and integrated into the global climate model of Laboratoire de Météorologie Dynamique LMDZ. LMDZ is the atmospheric component of the IPSL coupled atmosphere-ocean model, IPSL-CM, used in particular for the CMIP exercises. Concerning boundary layer subgrid scale vertical transport, LMDZ combines a K-diffusion model based on a prognostic equation of the TKE with a mass flux parameterization (Rio and Hourdin, 2008). This so-called thermal plume model and its coupling to a bi-Gaussian cloud scheme (Jam et al., 2013) is presented in Sect. 3.1, 3.2 and 3.3 as it is an essential element of the proposed parameterization. Deep convection on the other hand is modeled by the Emanuel's scheme in interaction with an original parameterization of cold pools created by reevaporation of convective rainfall (Emanuel, 1991; Emanuel and Živković-Rothman, 1999; Grandpeix and Lafore, 2010; Rio et al., 2009). The LMDZ6A configuration used here, developed for CMIP6, includes developments in many aspects of the model (Madeleine et al., 2020) compared to previous CMIP5 configurations, as for example a modification of the thermal plume detrainment which led to a significant improvement of stratocumulus representation (Hourdin et al., 2019). The LMDZ model is a flexible tool which contains in particular a single column model (SCM) version. The SCM is extensively used for parameterization development, assesment and tuning based on 1D/LES comparisons.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

2.2 The LES cases

The 1D/LES strategy is applied here with LES cases that embrace shallow convection scenes over earth and ocean as well as stratocumulus/cumulus transitions cases. These test cases were widely used in the development of the thermal plume model (Rio and Hourdin, 2008), the bi-Gaussian cloud scheme (Jam et al., 2013) and the modifief formulation of the thermal plume detrainment (Hourdin et al., 2019). They will be refered as IHOP/REF, ARMCU/REF, RICO/REF et SANDU/REF,FAST et SLOW and we will briefly recall the context of these different 1D cases.

IHOP comes from observations carried out during the International H_2O project on June 14, 2002 on the great plains (Couvreux et al., 2005). It represents an almost cloudless boundary layer.

The ARM case is derived from observations collected on 21 June 1997 at the Atmospheric Radiation Measurement site in Oklahoma, USA (Brown et al., 2002). It represents a typical diurnal cycle of the shallow convective boundary layer with appearance of a few cumuli clouds in the middle of the day.

The RICO case (Rain In Cumulus over the Ocean, van Zanten et al. (2011)) is a case of shallow cumuli clouds over the ocean characterized by frequent precipitation.

Finally, the SANDU REF, FAST and SLOW cases are described in Sandu I. and Stevens B. (2011). They were built by compositing the large-scale conditions encountered along a set of individual Lagrangian 3-day trajectories performed for the northeastern Pacific during the summer months of 2006 and 2007. They are meant to represent oceanic boundary layers overhung by stratocumulus clouds which become thinner with a gradual transition to shallow cumulus. REF, FAST and SLOW refer to different configurations involving more or less rapid transitions.

LES reference simulations used in this study have been performed with the MESO-NH non-hydrostatic model (Lac et al., 2018) for the IHOP, ARM and RICO cases. The LES used as a reference for the SANDU case is the one performed with the UCLA LES (Stevens and Seifert, 2008) and described in Sandu I. and Stevens B. (2011).

2.3 Model calibration

As explained in Hourdin et al. (2017), the calibration of the parameterizations of a GCM is a very sensitive step in climate modeling with the multiplication of sub-grid schemes and their associated free parameters. In LMDZ the calibration is done in successive stages from the scale of individual parameterization to the scale of the Earth's climate system as a whole. To facilite this multi step and multi parameter tuning, automatic tuning techniques were developed (Williamson et al., 2013, 2015; Couvreux et al., 2021; Hourdin et al., 2021) allowing to optimize the process and reduce biases. The tool explores the model parameter space in successive waves and build an emulator of given metrics. Once this emulator is built, the parameter space is reduced at each wave according to the gap between the emulated values of the metrics and the target values of the LES. This tuning process on 1D simulations not only makes it possible to significantly reduce the parameter space but it also permits to focus and improve our understanding of the physical processes involved in the parameterizations and thus to adapt the underlying modeling work. More than a simple statistical calibration tool, it is therefore a real working support for the modeler. SCM tuning is then completed by one or several 3D waves.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Based on this approach, and using the previous mentioned 1D cases, Hourdin et al. (2021) were able to finely calibrate the key parameterizations on which we step on in the present work, the thermal plume parameterization and the cloud scheme. Here, we will reinvest the same tuning approach in order to obtain reliable and consistent comparison criteria with previous work.

3 Thermal and cloud parameterization in LMDZ

In this part, we describe the implementation of the variance prognostic model within the statistical cloud scheme. In LMDZ, the asymmetry of the total specific humidity distribution is captured by a bi-Gaussian PDF. The first component of this bi-Gaussian represents the large-scale distribution, while the second component represents the distribution within the thermal plumes. The mass-flux scheme representing the thermal plumes and the statistical cloud scheme are therefore strongly coupled through the parameters of the PDF, and they will be even more so through the establishment of the prognostic equation of the variance. Thus, we will start this section by presenting the mass-flux parameterization of the updrafts before introducing the statistical cloud scheme, in which the specific humidity variance model is integrated.

3.1 The thermal plume model

Let ψ be a conservative variable, $\boldsymbol{v}$ the wind field and ρ the density. Using an Reynolds decomposition of the physical quantities $\psi = \bar{\psi} + \psi'$ and in the case of non-viscous transport, we can write:

$$\frac{\partial \bar{\psi}}{\partial t} = -\boldsymbol{v}.\boldsymbol{grad}(\bar{\psi}) - \frac{\mathrm{div}(\overline{\rho w' \psi'})}{\rho} \tag{1}$$

The first term on the right-hand is the tendency due to large-scale advection, computed by the dynamical core of the model, the second term is the tendency due to turbulent transport which is the sum of a K-diffusive term and a mass flux term :

$$\overline{\rho w' \psi'} = -K\rho \frac{\partial \bar{\psi}}{\partial z} + f(\overline{\psi_{th}} - \bar{\psi}) \tag{2}$$

The first term on the right-hand is the diffusive term, f is the convective mass flux, $\overline{\psi_{th}}$ the value of the variable in the thermal plume and $\bar{\psi}$ its value in the environment assimilated to the average value in the layer. Here the population of thermal plumes is represented by a unique thermal plume (Hourdin et al., 2002; Rio and Hourdin, 2008; Rio et al., 2010), which mass flux is given by $f = \rho \alpha w_{th}$ where α is the surface fraction covered by the thermal and w_{th} its vertical speed.

The vertical variation of the mass flux f is given by $\frac{\partial f}{\partial z} = e - d$ where e is an entrainment term and d a detrainment term. Then we can write the vertical variation of the variable flux as :

$$\frac{\partial f \overline{\psi_{th}}}{\partial z} = e\bar{\psi} - d\overline{\psi_{th}} \tag{3}$$

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Which leads to the following transport equation after some elementary algebra :

$$\frac{\partial \bar{\psi}}{\partial t} = \frac{d}{\rho}(\overline{\psi_{th}} - \bar{\psi}) + \frac{f}{\rho}\frac{\partial \bar{\psi}}{\partial z} \tag{4}$$

This formulation is the one that will make it easiest to derive the variance model. The first term on the right-hand is the tendency due to the thermal detrainment towards the environment, the second term is the tendency associated with the compensatory subsidence in the environment, the impact of which depends on the vertical gradient of the variable in the environment. Here the fact that the tendency of the variable does not depend on entrainment is linked to the assumption that the entrained air is the average air of the environment. This hypothesis simplifies both mean and variance equation.

3.2 The cloud scheme

The LMDZ statistical cloud scheme is based on a Bigaussian PDF of saturation deficit in the presence of thermals (Jam et al., 2013). The saturation deficit is given by $s = a_l(q_t - q_{sat}(T_l)$ where T_l is the liquid temperature, q_t the total specific humidity, q_{sat} the total specific humidity at saturation and a_l is a factor depending on thermodynamic variables defined in (Mellor, 1977; Sommeria and Deardorff, 1977). The Bigaussian PDF of the deficit at saturation reads :

$$PDF(s) = (1-\alpha)\mathcal{G}(s, \overline{s_{env}}, \sigma_{env}) + \alpha\mathcal{G}(s, \overline{s_{th}}, \sigma_{th}) \tag{5}$$

Where $\mathcal{G}$ is a Gaussian function of the variable s with mean value s_{env} and variance σ_{env}^2 in the environment and s_{th} and σ_{th}^2 inside the thermal plume, α being the fraction of the grid covered by the thermal plume. This approach is very powerfull to capture the positive skewness due to ascending plumes but we should also notice that it fails representing the negative asymmetry due to the drying of the atmosphere by subsidence. On the other hand, subsidences are generally less organized than ascendances in local structures which reduces their impact on humidity asymemtry. With this choice of PDF using the thermal plume fraction α to fix the weights of each Gaussian term, the variance and skewness of the saturation deficit are determined by the following equations being given the standard deviations σ_{env} and σ_{th} and the mean values in both environment and thermals (the detailed derivation being given in the appendix).

$$V = \alpha\sigma_{th}^2 + (1-\alpha)\sigma_{env}^2 + \alpha(1-\alpha)(\bar{s}_{th} - \bar{s}_{env})^2 \tag{6}$$

$$S = \frac{3\alpha(1-\alpha)(\bar{s}_{th} - \bar{s}_{env})(\sigma_{th}^2 - \sigma_{env}^2) + \alpha(1 - 3\alpha + 2\alpha^2)(\bar{s}_{th} - \bar{s}_{env})^3}{V^{\frac{3}{2}}} \tag{7}$$

Once the PDF is given, the cloud fraction c_f of the mesh and the cloud water content q_c can be computed as :

$$c_f = \int_0^\infty PDF(s)ds \tag{8}$$

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

$$q_c = \int_0^{\infty} s \cdot PDF(s) ds \tag{9}$$

If the mean values of the saturation deficit in the updrafts and in the environment are explicitly calculated in the GCM, it remains to determine the values of the standard deviations in order to close this cloud scheme. In the next paragraphs, we will present the current diagnostic variance model and then the proposed prognostic model.

3.3 The original diagnostic variance model

In the current diagnostic version, the standard deviations of humidity in thermals and the environment are parameterized as follows :

$$\sigma_{th} = BG2(\alpha + 0.01)^{\gamma_2}(\bar{s}_{th} - \bar{s}_{env}) + b\bar{q}_{th} \tag{10}$$

and

$$\sigma_{env} = BG1\frac{\alpha^{\gamma_1}}{1-\alpha}(\bar{s}_{th} - \bar{s}_{env}) + b\bar{q}_{env} \tag{11}$$

$BG1$, $BG2$, γ_1, γ_2 and b being tunable parameters. This specific parameterization of the variance is based on a heuristic reasoning using the exchange surface and the specific humidity difference between the thermals and the environment (Jam et al., 2013). Although it has enabled significant advances in cloud representations, our aim is to replace this multiparameter heuristic parameterization with a parameterization based on a well-established variance equation.

3.4 The variance prognostic model

Equation 4 cannot apply to the variance of a variable because the variance is not conservative. Analytical equations for the variance evolution in the presence of detraining mass fluxes have been proposed by K05 with one single plume, or by Tan et al. (2018) for a general formalism with N plumes. However their implementation in the climate model requires knowledge of the humidity variance in thermals. Although this is not an insurmountable difficulty in our model where it is predicted, we have chosen to implement a more straightforward, but theoretically equivalent, methodology consisting in transporting q and q^2 within the mass flux scheme. Let us first write the evolution of the variance of the total specific humidity in terms of q et q^2:

$$\frac{\partial \overline{q'^2}}{\partial t} = \frac{\partial \overline{q^2}}{\partial t} - \frac{\partial \bar{q}^2}{\partial t} = \frac{\partial \overline{q^2}}{\partial t} - 2\bar{q}\frac{\partial \bar{q}}{\partial t} \tag{12}$$

To compute the convective variance tendency we only need knowledge of the mean total specific humidity $\bar{q}$, its convective tendency $\frac{\partial \bar{q}}{\partial t}$ and the convective tendency of the total specific humidity square $\frac{\partial \overline{q^2}}{\partial t}$. $\bar{q}$ and $\frac{\partial \bar{q}}{\partial t}$ are already calculated in the

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

thermal plume model, therefore only $\frac{\partial \overline{q^2}}{\partial t}$ remains to be estimated. Noting that the total specific humidity is treated as a conservative variable in the thermal plume transport model, we can initially assume that its square is also conservative. (indeed, the material derivative of a squared quantity is zero if the material derivative of this quantity is zero). Therefore we are justified in transporting the square of the total specific humidity as a conservative tracer, just like the total specific humidity itself. After this initial conservative calculation, we will subsequently add relaxation terms to account for mixing processes in both the environment and the thermals. To compute Eq. 12 in our first conservative approach we only have to call the plume transport routine for both total humidity and its square, instead of just the first one, with no need to implement explicit equations. The consistency of this approach can be checked by deriving the analytical development which joins and provides another perspective on the K05's formalism of variance transport. Applying the Eq. 4 both to q and q^2, we obtain:

$$\frac{\partial \overline{q'^2}}{\partial t} = \frac{d}{\rho}(\overline{q^2_{th}} - \overline{q^2}) + \frac{f}{\rho}\frac{\partial \overline{q^2}}{\partial z} - 2\bar{q}[\frac{d}{\rho}(\overline{q_{th}} - \bar{q}) + \frac{f}{\rho}\frac{\partial \bar{q}}{\partial z}] \tag{13}$$

Subsidence terms combine as follows : $\frac{f}{\rho}\frac{\partial \overline{q^2}}{\partial z} - 2\bar{q}\frac{f}{\rho}\frac{\partial \bar{q}}{\partial z} = \frac{f}{\rho}\frac{\partial \overline{q'^2}}{\partial z}$. Detrainment terms can also be rearranged : $(\overline{q^2_{th}} - \overline{q^2}) - 2\bar{q}(\overline{q_{th}} - \bar{q}) = (\overline{q_{th}} - \bar{q})^2 + (\overline{q'^2_{th}} - \overline{q'^2})$. And the convective variance tendency finally reads :

$$\frac{\partial \overline{q'^2}}{\partial t} = \frac{d}{\rho}[(\overline{q_{th}} - \bar{q})^2 + (\overline{q'^2_{th}} - \overline{q'^2})] + \frac{f}{\rho}\frac{\partial \overline{q'^2}}{\partial z} \tag{14}$$

This approach leads to a similar equation that K05 derives from the mixing of two air masses and its impact on the humidity variance in the environment. This is not surprising because both approaches ultimately rely on the air mass conservation equations. In the present formalism we can underline the absence of the entrainment term, due to the already mentioned fact that we assumed that the entrained air had the average composition of the environment. We can interpret the different terms of the Eq. 14 as follows :

- The first term represents the impact of the mean values difference between the thermal plume and the environment
- The second one expresses the impact of the variances difference between the thermal plume and the environment
- The last one expresses the vertical transport of variance due to compensatory subsidence.

Note that this variance transport equation is limited to organized convective sources which are supposed to be dominant in the studied cases, it could subsequently include the diffusive terms, computed with the same approach by calling the diffusive transport model for both q and q^2. We could also include possible source/sink terms due to evaporation of precipitation for instance, or large-scale advection, this will be the subject of forthcoming publication. Here we place ourselves in a framework where the transport of variance is carried out by the organized convective terms and we only add a classical exponential dissipation term representing the small scale homogenization processes which will be discussed in Sect. 3.5. The present model finally reads :

$$\frac{\partial \overline{q'^2}}{\partial t} = \frac{d}{\rho}[(\overline{q_{th}} - \bar{q})^2 + (\overline{q'^2_{th}} - \overline{q'^2})] + \frac{f}{\rho}\frac{\partial \overline{q'^2}}{\partial z} - \frac{\overline{q'^2}}{\tau} \tag{15}$$

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

where τ represents the relaxation time of the variance in the environment. In practice we can either implement this variance equation term by term or call the thermal transport equation for q and q^2 as described before. However, the two methods lead to very different implementation of the variance model, each of which having its own interest. In the first method we explicitly calculate the different terms of the Eq. 15 which allows us to compare the relative importance of this terms, especially the one associated to the difference in humidity mean and variance. The second way is simpler to implement and can be transposed to other transport processes (diffusive transport or large-scale advection for example). In this case it is no longer necessary to extract all the physical quantities from the thermal model, only the tendencies of $\overline{q^2}$ and $\bar{q}$ are necessary.

As highlighted by K05, an important question underlying this theoretical approach is the determination of humidity variance in updrafts : $\overline{q'^2_{th}}$. In the diagnostic parameterization, it was prescribed using the humidity difference with the environment. However, the transport of q and q^2 within the thermal model now gives us direct access to $\overline{q'^2_{th}} = \overline{q^2_{th}} - \overline{q_{th}}^2$. With the same algebra as above, starting from the Eq. 3 applied to q_{th} and q^2_{th}, we obtain the following expression for the vertical evolution of the variance in the thermal.

$$\frac{\partial \overline{q'^2_{th}}}{\partial z} = \epsilon[(\overline{q_{th}} - \bar{q})^2 + (\overline{q'^2} - \overline{q'^2_{th}})] \tag{16}$$

Where $\epsilon = \frac{e}{f}$. Expression that we can complete using an exponential relaxation representing small-scale diffusion processes in the thermal. The advective-diffusive coupling leads to the introduction of a characteristic diffusion height $z_0 = w_{th}\tau_{th}$ in the thermal plume. Where w_{th} is the vertical advection and τ_{th} the diffusion timescale in the plume.

$$\frac{\partial \overline{q'^2_{th}}}{\partial z} = \epsilon[(\bar{q} - \overline{q_{th}})^2 + (\overline{q'^2} - \overline{q'^2_{th}})] - \frac{\overline{q'^2_{th}}}{w_{th}\tau_{th}} \tag{17}$$

3.5 The relaxation time in the variance model

The Newton relaxation time τ which appears in the dissipation term is an important parameter of the variance calculation. It can be estimated as the relaxation time of the turbulent energy as was proposed in Nieuwstadt and Brost (1986) and Neggers (2009). Tompkins (2002) proposes to add to this small-scale relaxation a second dissipation process associated with very slower large-scale two-dimensional dissipation due to horizontal wind shear instability. This second component would be present even in case of strong temperature stratification. Although it will be a part of our reflection when extending our work to the upper atmosphere, we neglect it here, as we only focus on shallow convective cases where the first component is expected to be dominant. The order of magnitude of the turbulence relaxation time is :

$$\tau = \frac{l}{\sqrt{TKE}} \tag{18}$$

TKE being the turbulent kinetic energy, l a mixing length of roughly 100 m (Blackadar, 1962). We obtain an order of magnitude of 100 s to 1000 s for τ which is a fairly fast relaxation, especially compared to the 10-days estimated 2D relaxation (Tompkins, 2002). The prognostic model depends on a unique tunable parameter. We tested two options which led to small differences in the results either considering τ or l as the tunable parameter in Eq. 18. In the second case, it was necessary, as in Golaz et al. (2002a), to define an upper limit to the relaxation time as our TKE is likely to cancel out even below 3000 m. This prevents the relaxation terms from becoming too small and leading the variance to accumulate too strongly.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

3.6 Injecting the total specific humidity variance into the cloud scheme

To replace the previous diagnostic model with a prognostic model based on humidity variance transport equations, it is first necessary to connect the humidity variance to the variance of the saturation deficit which drives the statistical cloud scheme. This well-known relationship (Tompkins, 2002) can be written as :

$$\sigma(s)^2 = a_l^2(\overline{q'^2} + \alpha_l^2\overline{T_l'^2} - 2\alpha_l\overline{q'T_l'}) \qquad (19)$$

Where $\alpha_l = (\frac{\partial q_{sat}}{\partial T})_{T_l}$, a_l was mentioned in Sect. 3.2 and T_l is the liquid temperature.

We see that a liquid temperature variance term appears in this equation as well as a cross-correlation term whose physical interpretations are described in Tompkins (2003). Several studies have analyzed the relative importance of humidity and temperature contributions to the variance (Price, 2001; Tompkins, 2003). Although the temperature contribution is not completely negligible, it appears from these analyzes that order 1 is mostly dominated by the variability of specific humidity. In the context of this work, it seemed relevant to mainly concentrate on the specific humidity variability. Orders of magnitude of these variabilities can be estimated in LMDZ by simply diagnosing the differences in humidity and liquid temperature between the thermals and the environment, on this basis we obtain an order of magnitude of 10^{-4} for $\overline{q'^2}$ and closer to 10^{-5} for $\alpha_l^2\overline{T_l'^2}$. The comparison in LES of the standard deviations of total specific humidity multiplied by a_l and the saturation deficit also show an excellent first-order agreement. In this work the hypothesis is made that the variance of s is controlled by the variance of the total specific humidity. Therefore the humidity variance computed by the prognostic model is injected into σ_{env} (after multiplication by a_l) instead of being prescribed by the diagnostic parameterization.

4 Results

4.1 Tuning setup

In this part, we sumarize the 1D tuning work carried out on the IHOP, ARMCU, RICO and SANDU cases (Couvreux et al., 2021; Hourdin et al., 2021) when adding the variance model and its relaxation time as a free parameter. We use the same metrics as in paragraph 5.1 of Hourdin et al. (2021) to compare SCM simulations to LES. The free parameters are also kept the same except for the diagnostic model parameter $BG1$ which is replaced by the parameter τ, allowed to vary between 100 s and 2000 s.

We briefly recall the choice of metrics for the tuning process, for more details on the mathematical formalisms, refer to Hourdin et al. (2021). The metrics used to compare 1D GCM simulations to LES are based on three quantities: potential temperature, humidity and cloud cover. These metrics are defined by temporal and spatial integrations of the three quantities over time and altitudes. For cloud cover, three specific metrics are used, the first is linked to the maximum cloud cover on the vertical $\alpha_{cld,max}$, the second represents an average altitude of clouds $z_{cld,ave}$ and the last one defines a typical altitude of the

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Case	IHOP	ARMCU	RICO	SANDU	SANDU	SANDU
Subcase	REF	REF	REF	REF	SLOW	FAST
time	7-9	7-9	19-25	50-60	50-60	50-60
$\theta_{400-600m}$	×	×				
$q_{v,400-600m}$		×				
$\alpha_{cld,max}$		×	×			
$z_{cld,ave}$		×		×		
$z_{cld,max}$		×		×	×	×

Table 1. Metrics for the 1D/LES tuning. Time average is given in hours from the beginning of the simulation. Potential temperature is given in K, humidity in kg kg^{-1} and height in m.

maximum cloud cover $z_{cld,max}$. Table 1 details all the different metrics used in this tuning (the three cloud metrics, and those concerning potential temperature θ and humidity q_v).

The free parameters to be tuned primarily concern the modeling of the entrainment and detrainment rates of the thermal plume model which depend on the buoyancy and the vertical speed in the ascent (these parameterizations are controlled by the parameters $A1$, $B1$, $B2$, CQ). The modification of the detrainment rate proposed in (Hourdin et al., 2019) introduces a characteristic height difference in the calculation of buoyancy which is controlled by the DZ parameter. The $BG2$ and $BG1$ parameters (see paragraph 3.3) are linked to the standard deviation of humidity in the thermal plume and the environment. Both were removed and replaced by the relaxation time of the new model τ and τ_{th}. Finally, the CLC and EVAP parameters are involved in the precipitation and rain re-evaporation model, CLC being associated with the critical incloud water from which precipitation is activated and EVAP being a free parameter of the precipitation flux equation which controls the fraction of precipitation that re-evaporates at a given altitude. The formalism of these parameterizations was introduced by the work of Sundqvist (1978). Table 2 presents a summary of the different parameters to be tuned. More details on the parameterizations and free parameters can be found in Hourdin et al. (2021).

4.2 Simulation scores according to the tuning tool

At each tuning wave, LMDZ SCM-simulations are computed within the range of authorized parameters by the previous wave (NROY : not-ruled out yet). For each of this simulations, a score is assigned per metric, this score being defined from the difference between the target value of the metric (the average value of the LES) and the value for the given simulation, as well as from the accepted tolerance. The closer the score is to 0, the closer the simulation metric is to the target value, a score equal to 1 indicates that the difference between the simulated metric and the target value corresponds exactly to the tolerance which we are willing to accept. The simulations can thus be classified according to the obtained scores, either on an average score

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Name	Min	Max	Ref	Sampling	controls
A1	0.5	1.2	2.	linear	contribution of buyancy to the plume acceleration
A2	$1.5e-3$	$4.e-3$	$2.e-3$	linear	drag term in the plume acceleration
B1	0.	1.	0.95	linear	scaling factor for entrainment and detrainment
CQ	0.	0.02	0.012	linear	influence of humidity contrast on detrainment
DZ	0.05	0.2	0.07	linear	environmental air altitude shift for buyancy computation
EVAP	$5.e-5$	$5.e-4$	$1.e-4$	log	reevaporation of rainfall
CLC	$1.e-4$	$1.e-3$	$6.5e-4$	linear	autoconversion of cloud liquid water to rainfall
τ	100.	2000	700	linear	timescale for variance dissipation in the environment
τ_{th}	100.	2000	700	linear	timescale for variance dissipation in the plume

Table 2. Free parameters of the 1D/LES tuning

criterion or on a maximum criterion. Her we use the maximum criterion to classify the ten most efficient simulations in order to prevent large errors on certain metrics from being compensated by very good results on others.

Figure 1 compares the results obtained on all the metrics between the diagnostic and prognostic models. In particular we see that the ten best simulations achieve almost identical scores in both cases, close to 0.75 which means that the worst metric is at a distance of 0.75 σ from the target value. The distribution of scores in the different metrics is also very similar.

It is important to underline that the tuning process does not aim to select an optimized configuration of the model but rather to highlight a range of parameters leading to acceptable results within the framework we have defined and the typical performance of the model in this range. In this context this tuning showed that the acceptable range of the relaxation time is 100 s to 1300 s which is consistent with the previous qualitative considerations. Or, if using the TKE to define the relaxation time, the acceptable range of the mixing lenght shows up to be between 70 m and 160 m.

4.3 Representation of the moments of the humidity distribution and cloud cover

In this section, we focus on the variance, skewness and cloud cover profiles which are the main observables of this work, being noticed that most physical quantities are little affected by the change in variance parameterization as shown in Sect. 4.2. This is further illustrated in Fig. 2 which shows the time evolution of the vertical profile of cloud cover, as simulated by the LES in ARMCU, RICO and SANDU and its differences with the diagnostic and the prognostic model. This confirms that the simulated cloud cover is quite close between the two models, with a slight decrease of biases with the prognostic model.

Figure 3 shows the vertical profile of cloud cover and the associated standard deviation of specific humidity for specific times of simulations, as well as the time evolution of the standard deviation averaged within the cloud layer. Figure 4 separates

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

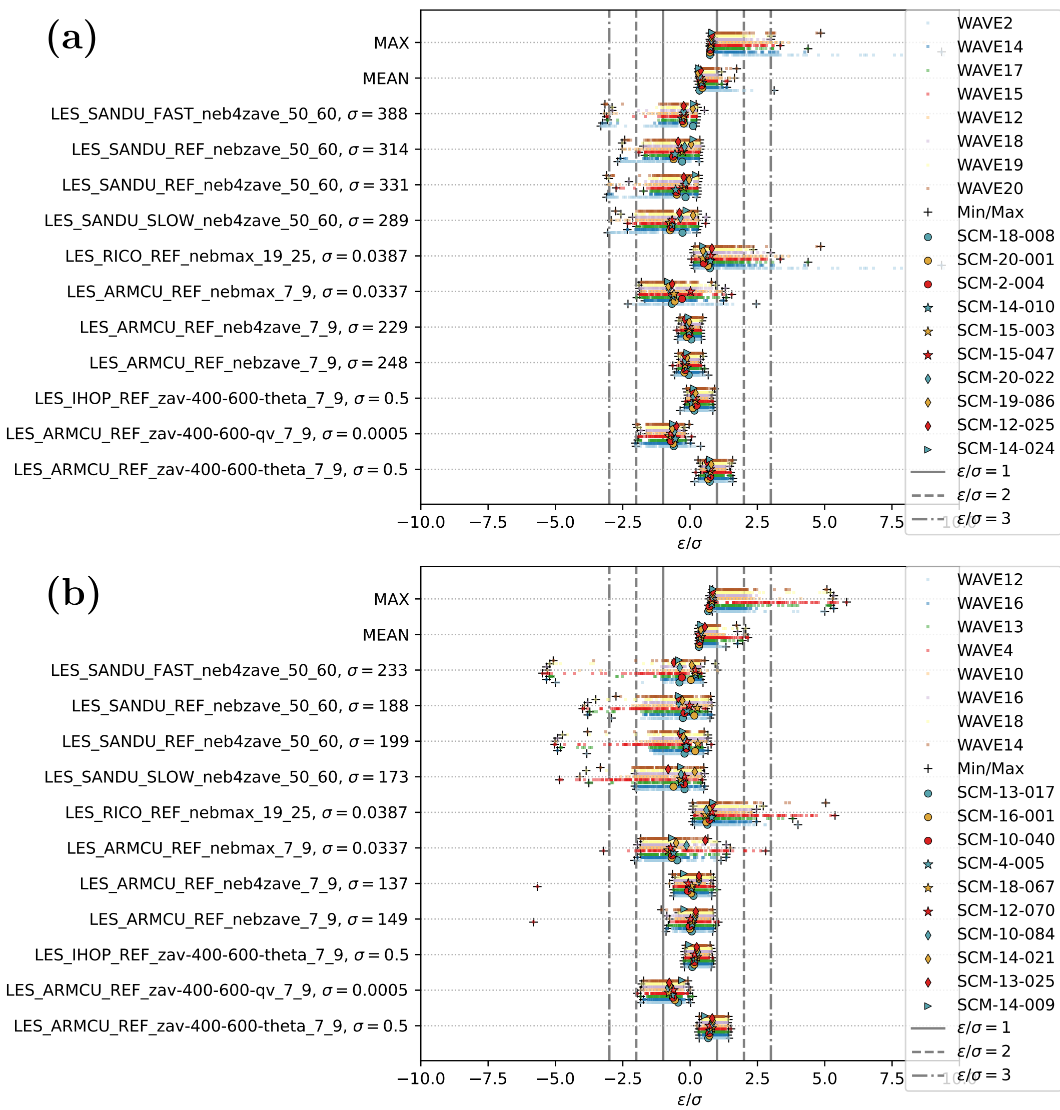

Figure 1. Metric scores (ϵ/σ) over chosen waves including the best simulations. (a) : with the previous diagnostic model. (b) : with the prognostic model. All the scores from the 8 wave simulations are represented with their minimum and maximum scores. The ten best tuning simulations with their scores are highlighted. The 11 metrics are recalled on the left side with their acceptable margins of error σ.

the total standard deviation into the standard deviation within thermals and within their environment, while Fig. 5 evaluates the representation of the third order moment and skewness.

While the cloud cover and the standard deviation of total specific humidity are quite close between the two model versions and the LES (Fig. 3), the prognostic model improves the representation of the standard deviation within thermals (Fig. 4) which was overestimated by one order of magnitude with the diagnostic model. This strongly impacts the third-order moment (Fig.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

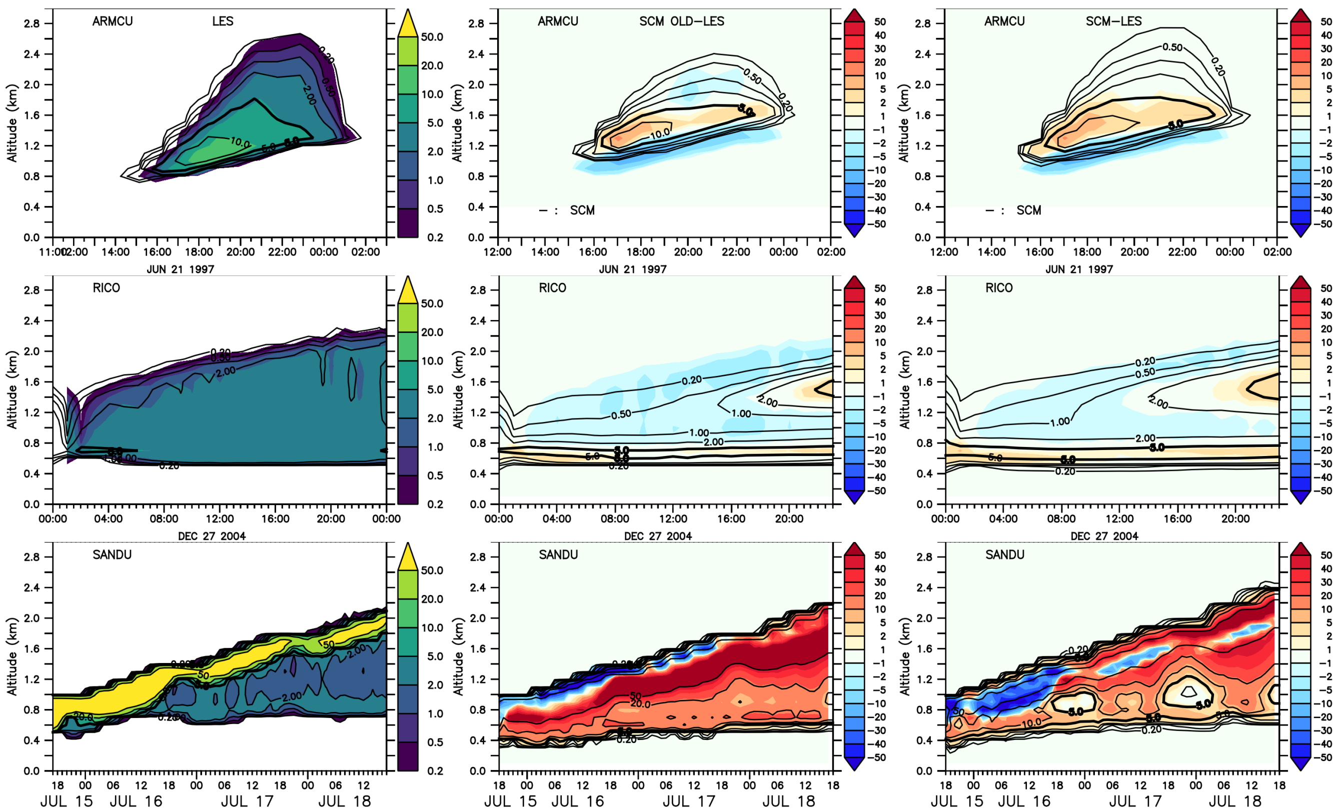

Figure 2. Total cloud cover (%) in the ARMCU/REF (top), RICO/REF (middle) and SANDU/REF (bottom) cases. On the left the LES simulation. In the middle the difference in cloud cover between the best tuning simulation with the old diagnostic model and the LES, black contours indicating the cloud cover in the SCM simulation. On the right the difference in cloud cover between the best tuning simulation with the prognostic model and the LES, black contours indicating the cloud cover in the SCM simulation.

5) through the first term of Eq. 7. This major flaw is largely corrected by the prognostic model, although the standard deviation is there slightly underestimated. The third-order moment, which reflects the asymmetry of the humidity distribution, is now in better agreement with the LES. This demonstrates that an improved representation of the higher moments of the distribution profiles can be obtained by simply exploiting the transport of q and q^2 within the thermals, with a tunable parameter that is easy to estimate using TKE. A possible flaw of the prognostic model is the lack of representation in the PDF of organized subsidizing structures which tend to locally dry out the environment. Although subsidences are taken into account as sources of variance, they cannot negatively impact the third order moment because of the binary structure of our PDF with only thermal

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

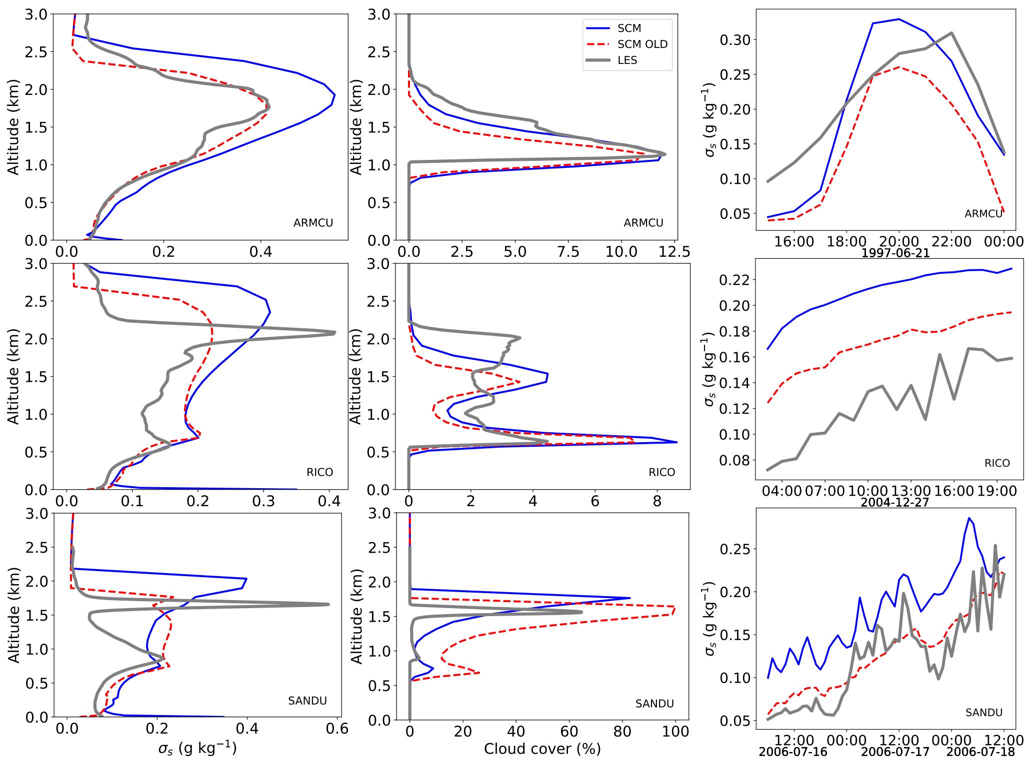

Figure 3. Variance of the saturation deficit, including the contribution of thermals, in the ARMCU/REF (top), RICO/REF (middle) and SANDU/REF (bottom) cases. Thick gray is the LES, blue is the LMDZ simulation with prognostic variance and dashed red with previous diagnostic model. On the left the profiles at 7 p.m. for ARMCU, 8 p.m. for RICO and midnight on the last day for SANDU, in the middle the cloud cover profile at the same dates and on the right side temporal evolutions of the averaged standard deviation between 500 m and 2500 m for ARMCU and RICO and 400 m and 2200 m for SANDU. These data were obtained with the best tuning simulations for each model.

plume and environment. This is particularly visible in rare areas of negative skewness but it should not overshadow the quite accurate estimation of the third-order moments of the model.

Figure 6 further shows the contribution of the different terms of Eq. 14 in the variance transport. It shows the clear predominance of the first term of the variance transport equation, a term which represents the squared difference in humidity between thermals and environment. The two other terms, difference of variance and subsidence, appear to be an order of magnitude smaller so that they could be neglected if one wants to simplify the implementation of the explicit transport equation in the model. This will be very usefull when it comes to deep convection transport, which will be treated in a future publication, and more generally in cases where the model does not predict the humidity variance in the updraft but only its mean humidity. Let us specify that this dominance of the first term doesn't match K05's study where the first two terms are of the same order of magnitude, the last one being smaller. But the analysis of our LES clearly confirms the dominance of the first term with our

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

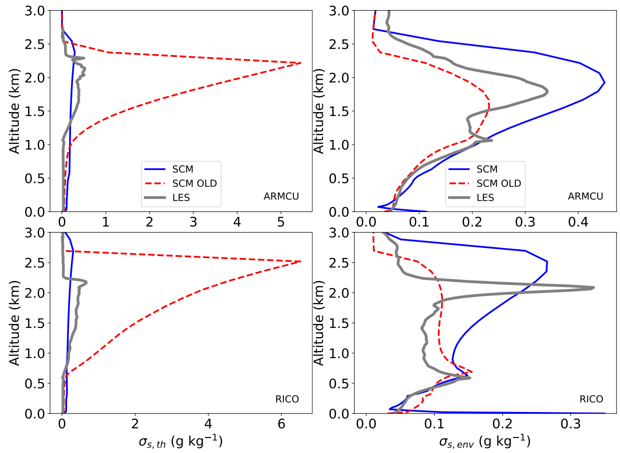

Figure 4. Standard deviation of the saturation deficit in the thermals and in the environment. Top ARMCU/REF, bottom RICO/REF at the same date than previously. Blue : SCM with prognostic variance, dashed red : SCM with diagnostic variance, thick gray : LES.

choice of thermal sampling. The difference between the two analyses may lie in this sampling, but the reference simulations may also have evolved significantly.

The accurate representation of the vertical profile of cloud cover during the transition from stratocumulus to cumulus clouds is particularly challenging (SANDU case). Hourdin et al. (2019) showed how a judicious modification of the detrainment parameterization made it possible to simulate this transition quite accurately. However, Fig. 2 shows that the cloud cover still tends to remain saturated with the diagnostic model while it gradually fades in the LES. Moreover this cloud cover is a little too thick at night in LMDZ especially from the second day of the simulation where it becomes significantly finer in the LES. Figures 2 and 3 show that those two aspects are attenuated with the variance prognostic model. Looking deeper at the variance profile, we can attribute this improvement to a better representation of the variance peak at the top of the clouds. Let's remind, indeed, that an increase in the variance of the statistical cloud scheme, in case of high cloud cover, is likely to reduce the cloud cover by increasing the fraction of air with lower humidity. This better representation of the variance peak can partly be attributed to the thermal plume detrainment peak in this area as we can see in Fig. 6 (in the middle bottom) even if this peak occurs a little too high. In fact, the detrainment is directly involved in the variance transport equation but was not taken into account in the diagnostic model which rather involves a term dependent on the thermal fraction (dashed red curve in the middle of Fig. 6). This term vanishes to zero while the detrainment reaches its peak which explains the difference in behavior between the two models. We can notice, however, that the difference in humidity between thermals and environment also plays a crucial role at these altitudes (in black in Fig. 6 in the middle) but in the diagnostic model this term is multiplied by a factor which tends quickly to zero while in the prognostic model the multiplicative factor, which depends on the detrainment, cancels out higher and can even reach a local maximum at the top of the clouds.

This analyses confirms that the choice of a weighted bivariate Gaussian PDF, based on the fraction of thermals, provides a satisfactory framework for representing the asymmetric distribution of humidity at different atmospheric levels especially

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

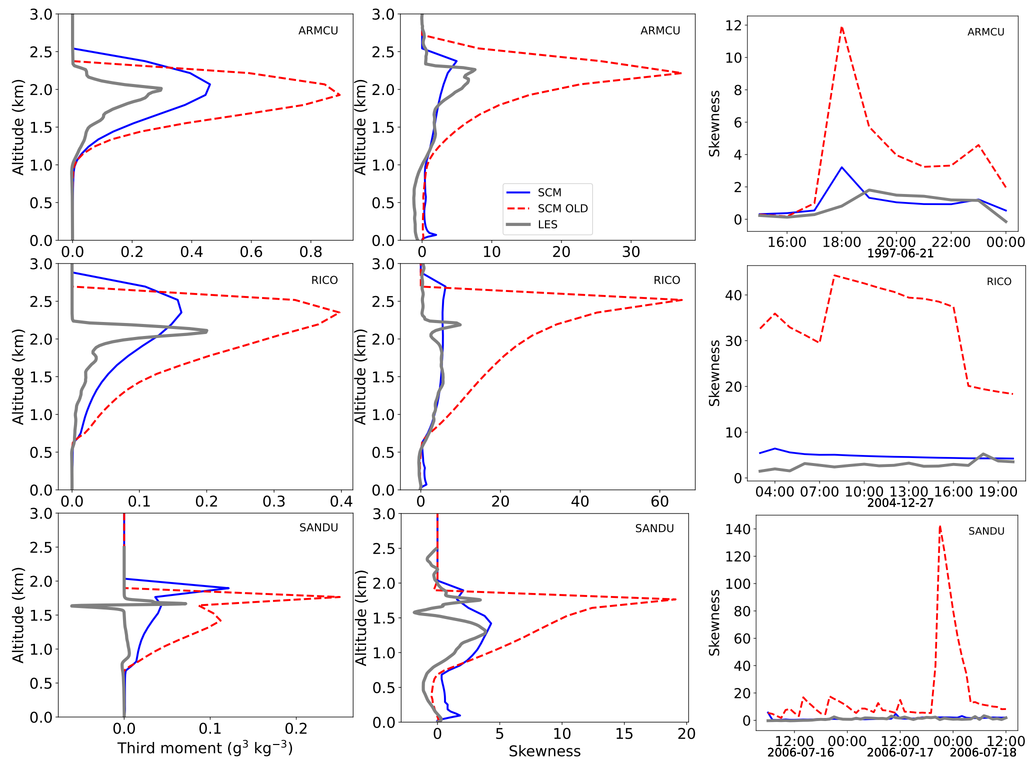

Figure 5. Centered third order moment (numerator of Eq. 7) of the saturation deficit distribution on the left, skewness of the saturation deficit distribution and its evolution in time (averaged over the same vertical interval) on the middle and right side. Top ARMCU/REF, middle RICO/REF and bottom SANDU/REF at the same date than previously for left and middle. Blue : SCM with prognostic variance, dashed red : SCM with diagnostic variance, thick gray : LES.

when associated with the prognostic variance model, which helps avoid inconsistent behaviors of the saturation deficit in the thermals and at the cloud tops through the thermal detrainment term.

5 Conclusions

In the work presented above, we focused on designing and testing in a GCM a prognostic model of the variance of total specific humidity whose physical content is based on the general transport equations. In the context of shallow convection, we made large use of the thermal plume model in which the specific humidity is treated as a conservative variable. This property permits to transport its square and to derive the equation of evolution of the variance by assimilating the diffusive homogenization to a simple Newtonian relaxation involving a time parameter which can be estimated from the tke. Calibration of this free parameter was performed using automatic tuning tools. The implementation of this model coupled to a Bigaussian statistical cloud scheme made it possible to reproduce with great consistency the previous results obtained with the diagnostic model, but also to significantly improve the representation of the higher moments of the specific humidity distribution. These

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

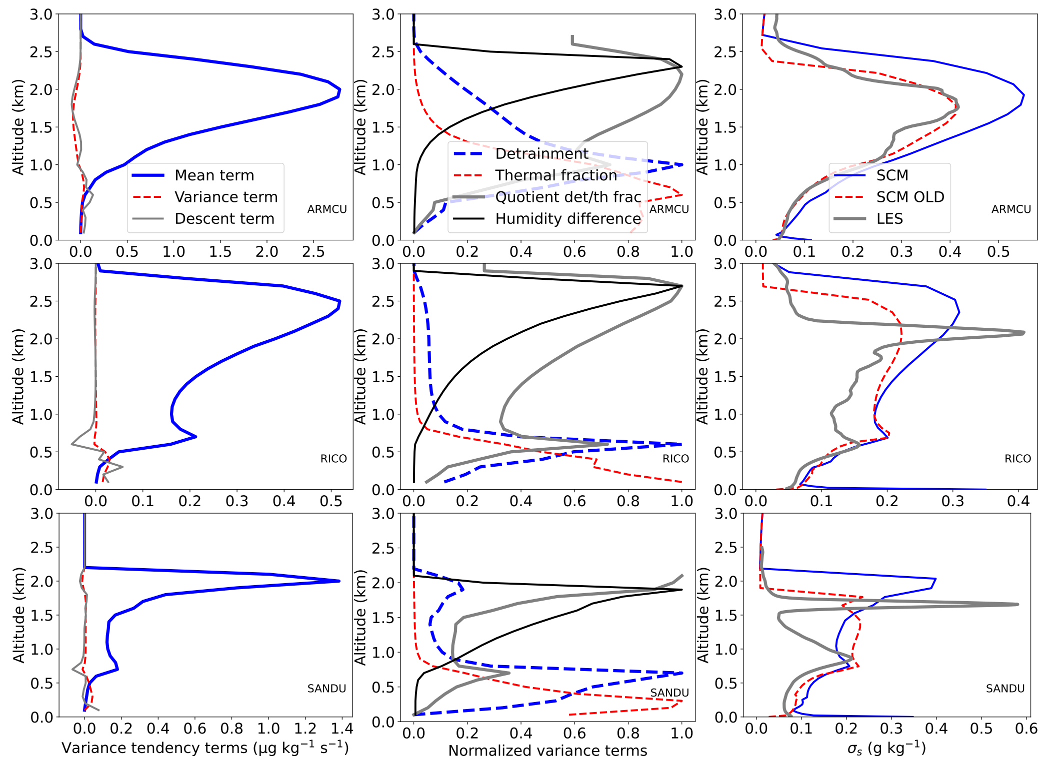

Figure 6. Various profiles in the ARMCU/REF (top) RICO/REF (middle) and SANDU/REF (bottom) cases. On the left : comparison between the 3 source terms of variance in Eq. 14, blue being $\frac{d}{\rho}(\overline{q_{th}}-\bar{q})^2$ the squared mean difference term, dashed red the variance difference term $\frac{d}{\rho}(\overline{q'^2_{th}}-\overline{q'^2})$ and thick gray the subsidence term $\frac{f}{\rho}\frac{\partial\overline{q'^2}}{\partial z}$. In the middle the normalized profiles of some physical quantites involved in the variance model : the detrainment rate d (thick dashed blue), the thermal fraction multiplicative term in the diagnostic model $\frac{\alpha^{\gamma_1}}{1-\alpha}$ (dashed red) and the square of the total humidity difference between thermal and environment $(\bar{q}_{th}-\bar{q}_{env})^2$(black). The quotient of the detrainement term by the termal fraction term is also represented (thick gray). On the right the total humidity variance profiles are remained : blue for the prognostic variance model, dashed red for the diagnostic one and thick gray for the LES. This profiles were computed at the same dates than before.

improvements are closely related to the consideration of detrainment, which plays a key role in the prognostic equations of the variance. Certain inaccuracies of the diagnostic model could therefore be corrected, in particular in the thermal plumes and at the top of the clouds where detrainment is significant. However, negative asymmetries due to downdrafts cannot be accounted for within the framework of the bi-Gaussian scheme. As a natural extension of this work, and it is in progress, it will be relevant to integrate the formalism of the Emanuel's scheme for deep convection into the variance model and further source terms like evaporation of precipitations. The final objective would be to transport the large-scale variance as a state variable in the dynamics of the GCM with the same approach proposed in the present study. This work is therefore an important step to incorporate into a GCM a unified global model of the upper moments of the atmospheric specific humidity relying on the

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

fundamental physical equations and with a minimum of free parameters. It has helped to improve our understanding of the role of thermal updrafts in the variability of specific humidity, particularly through the detrainment process.

Code and data availability. The LMDZ GCM is available at:
http://svn.lmd.jussieu.fr/LMDZ/LMDZ6/trunk
It can be installed automatically on a laptop with the command http://lmdz.lmd.jussieu.fr/pub/install_lmdz.sh The htexplo tool for tuning is available at:
https://svn.lmd.jussieu.fr/HighTune/trunk
The data and scripts used to make the figures of the document will be made available as well on a DOI if the article is accepted for publication. The scripts that will be made available include downloading the correct versions of the LMDZ and htexplo software.

Appendix A: Calculation of second- and third-order moments of saturation deficit

A1 Second order moment of saturation deficit

The second order moment of saturation deficit writes by simply separating the spatial average into a thermal part and an environmental part : :

$$\sigma^2 = \overline{s'^2} = \overline{(s-\bar{s})^2} = \alpha \overline{s'^2_{th}} + (1-\alpha)\overline{s'^2_{env}} \tag{A1}$$

Let us calculate the two terms separately :

$$\begin{aligned} \overline{s'^2_{th}} &= \overline{(s_{th}-\bar{s})^2} = \overline{(s_{th}-\overline{s_{th}}+\overline{s_{th}}-\bar{s})^2} \\ &= \overline{(s_{th}-\overline{s_{th}})^2} + (\overline{s_{th}}-\bar{s})^2 + 2\overline{(s_{th}-\overline{s_{th}})}(\overline{s_{th}}-\bar{s}) \end{aligned} \tag{A2}$$

The last term is zero, the first one is the variance σ^2_{th} of the saturation deficit in the thermal part, and the middle one can be rearranged as follow noting that $\bar{s} = \alpha\overline{s_{th}} + (1-\alpha)\overline{s_{env}}$:

$$(\overline{s_{th}}-\bar{s})^2 = (1-\alpha)^2(\overline{s_{th}}-\overline{s_{env}})^2 \tag{A3}$$

Finally it leads to :

$$\overline{s'^2_{th}} = \sigma^2_{th} + (1-\alpha)^2(\overline{s_{th}}-\overline{s_{env}})^2 \tag{A4}$$

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

In the same way we calculate :

$$\overline{s'^2_{env}} = \overline{(s_{env} - \bar{s})^2} = \overline{(s_{env} - \overline{s_{env}} + \overline{s_{env}} - \bar{s})^2}$$
$$= \overline{(s_{env} - \overline{s_{env}})^2} + (\overline{s_{env}} - \bar{s})^2 + 2\overline{(s_{env} - \overline{s_{env}})}(\overline{s_{env}} - \bar{s}) \tag{A5}$$

Which becomes:

$$\overline{s'^2_{env}} = \sigma^2_{env} + \alpha^2(\overline{s_{th}} - \overline{s_{env}})^2 \tag{A6}$$

By introducing A4 and A6 into A1, we obtain as announced :

$$V = \sigma^2 = \overline{s'^2} = \alpha\sigma^2_{th} + (1-\alpha)\sigma^2_{env} + \alpha(1-\alpha)(\overline{s_{th}} - \overline{s_{env}})^2 \tag{A7}$$

A2 Third order moment of saturation deficit

In the same way, we derive the expression for the third-order moment :

$$M_3 = \overline{s'^3} = \overline{(s-\bar{s})^3} = \alpha\overline{s'^3_{th}} + (1-\alpha)\overline{s'^3_{env}} \tag{A8}$$

Which splits into

$$\overline{s'^3_{th}} = \overline{(s_{th} - \bar{s})^3} = \overline{(s_{th} - \overline{s_{th}} + \overline{s_{th}} - \bar{s})^3}$$
$$= \overline{(s_{th} - \overline{s_{th}})^3} + (\overline{s_{th}} - \bar{s})^3 + 3\overline{(s_{th} - \overline{s_{th}})^2}(\overline{s_{th}} - \bar{s}) \tag{A9}$$

by noting that $3\overline{(s_{th} - \overline{s_{th}})}(\overline{s_{th}} - \bar{s})^2 = 0$. In our framework, the distribution of the saturation deficit is Gaussian in the thermal part. Therefore, the first term of A9, which represents the third-order moment in the thermal part, is zero. Thus, using A3, it remains :

$$\overline{s'^3_{th}} = (1-\alpha)^3(\overline{s_{th}} - \overline{s_{env}})^3 + 3\sigma^2_{th}(1-\alpha)(\overline{s_{th}} - \overline{s_{env}}) \tag{A10}$$

By applying the same reasoning in the environmental part :

$$\overline{s'^3_{env}} = \overline{(s_{env} - \bar{s})^3} = \overline{(s_{env} - \overline{s_{env}} + \overline{s_{env}} - \bar{s})^3}$$
$$= \overline{(s_{env} - \overline{s_{env}})^3} + (\overline{s_{env}} - \bar{s})^3 + 3\overline{(s_{env} - \overline{s_{env}})^2}(\overline{s_{env}} - \bar{s}) \tag{A11}$$

Which leads to :

$$\overline{s'^3_{env}} = \alpha^3(\overline{s_{env}} - \overline{s_{th}})^3 + 3\sigma^2_{env}\alpha(\overline{s_{env}} - \overline{s_{th}}) \tag{A12}$$

And after some simple algebra we obtain the third-order moment by injecting A10 and A12 into A8:

$$M_3 = 3\alpha(1-\alpha)(\overline{s_{th}} - \overline{s_{env}})(\sigma^2_{th} - \sigma^2_{env}) + \alpha(1 - 3\alpha + 2\alpha^2)(\overline{s_{th}} - \overline{s_{env}})^3 \tag{A13}$$

The skewness is then obtained by dividing by $V^{\frac{3}{2}}$ as in the main text (Eq. 7).

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Author contributions. L.A. : work design, diagnostics, analysis, figures and writing
F.H.: work design, diagnostics, analysis and writing
C.R.: analysis and writing

Competing interests. None

Acknowledgements. The authors thank Jean-Yves Grandpeix for helpful discussions. We also acknowledge support from the DEPHY research group of CLIMERI, funded by CNRS/INSU and Météo-France. The work benefited from the development of the htexplo tool within this project and automation of LES/SCM simulations through standardization of IO formats.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

References

Blackadar, A. K.: The vertical distribution of wind and turbulent exchange in a neutral atmosphere, J. Geophys. Res., 67, 3095–3102, https://doi.org/10.1029/JZ067i008p03095, 1962.

Bony, S. and Emanuel, K. A.: A Parameterization of the Cloudiness Associated with Cumulus Convection; Evaluation Using TOGA COARE Data, J. Atmos. Sci., 58, 3158–3183, https://doi.org/10.1175/1520-0469(2001)058<3158:APOTCA>2.0.CO;2, 2001.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G., and Webb, M. J.: How Well Do We Understand and Evaluate Climate Change Feedback Processes?, J. Clim., 19, 3445–3482, https://doi.org/10.1175/JCLI3819.1, 2006.

Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., Lewellen, D. C., Lock, A. P., MacVean, M. K., Moeng, C.-H., Neggers, R. A. J., Siebesma, A. P., and Stevens, B.: Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land, Q. J. R. Meteorol. Soc., 128, 1075–1093, https://doi.org/10.1256/003590002320373210, 2002.

Chatfield, R. B. and Brost, R. A.: A two-stream model of the vertical transport of trace species in the convective boundary layer, J. Geophys. Res.:Atmos., 92, 13 263–13 276, https://doi.org/10.1029/JD092iD11p13263, 1987.

Couvreux, F., Guichard, F., Redelsperger, J. L., Kiemle, C., Masson, V., Lafore, J. P., and Flamant, C.: Water-vapour variability within a convective boundary-layer assessed by large-eddy simulations and IHOP_2002 observations, Q. J. R. Meteorol. Soc., 131, 2665–2693, https://doi.org/10.1256/qj.04.167, 2005.

Couvreux, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., Rio, C., Audouin, O., Salter, J., Bazile, E., Brient, F., Favot, F., Honnert, R., Lefebvre, M.-P., Madeleine, J.-B., Rodier, Q., and Xu, W.: Process-Based Climate Model Development Harnessing Machine Learning: I. A Calibration Tool for Parameterization Improvement, J. Adv. Model. Earth Syst., 13, e2020MS002 217, https://doi.org/10.1029/2020MS002217, 2021.

Deardorff, J. W.: The Counter-Gradient Heat Flux in the Lower Atmosphere and in the Laboratory, J. Atmos. Sci., 23, 503–506, https://doi.org/10.1175/1520-0469(1966)023<0503:TCGHFI>2.0.CO;2, 1966.

Emanuel, K. A.: A Scheme for Representing Cumulus Convection in Large-Scale Models, J. Atmos. Sci., 48, 2313–2329, https://doi.org/10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2, 1991.

Emanuel, K. A. and Živković-Rothman, M.: Development and evaluation of a convection scheme for use in climate models, J. Atmos. Sci., 56, 1766–1782, 1999.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description, J. Atmos. Sci., 59, 3540–3551, https://doi.org/10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002a.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part II: Model Results, J. Atmos. Sci., 59, 3552–3571, https://doi.org/10.1175/1520-0469(2002)059<3552:APBMFB>2.0.CO;2, 2002b.

Grandpeix, J.-Y. and Lafore, J.-P.: A Density Current Parameterization Coupled with Emanuel's Convection Scheme. Part I: The Models, J. Atmos. Sci., 67, 881–897, https://doi.org/10.1175/2009JAS3044.1, 2010.

Hourdin, F., Couvreux, F., and Menut, L.: Parameterization of the Dry Convective Boundary Layer Based on a Mass Flux Representation of Thermals, J. Atmos. Sci., 59, 1105–1123, https://doi.org/10.1175/1520-0469(2002)059<1105:POTDCB>2.0.CO;2, 2002.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D.: The Art and Science of Climate Model Tuning, BAMS, 98, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1, 2017.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M.-P., Brient, F., and Idelkadi, A.: Unified Parameterization of Convective Boundary Layer Transport and Clouds With the Thermal Plume Model, J. Adv. Model. Earth Syst., 11, 2910–2933, https://doi.org/10.1029/2019MS001666, 2019.

Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranque, N., Musat, I., Fairhead, L., Diallo, F. B., and Volodina, V.: Process-Based Climate Model Development Harnessing Machine Learning: II. Model Calibration From Single Column to Global, J. Adv. Model. Earth Syst., 13, e2020MS002 225, https://doi.org/10.1029/2020MS002423, 2021.

Jam, A., Hourdin, F., Rio, C., and Couvreux, F.: Resolved Versus Parametrized Boundary-Layer Plumes. Part III: Derivation of a Statistical Scheme for Cumulus Clouds, Bound.-Lay. Meteorol., 147, 421–441, https://doi.org/10.1007/s10546-012-9789-3, 2013.

Klein, S. A., Pincus, R., Hannay, C., and Xu, K.-M.: How might a statistical cloud scheme be coupled to a mass-flux convection scheme?, J. Geophys. Res.:Atmos., 110, https://doi.org/10.1029/2004JD005017, 2005.

Köhler, M., Ahlgrimm, M., and Beljaars, A.: Unified treatment of dry convective and stratocumulus-topped boundary layer in the ECMWF model, Q. J. R. Meteorol. Soc., 137, 43–57, https://doi.org/10.1002/qj.713, 2011.

Lac, C., Chaboureau, J.-P., Masson, V., Pinty, J.-P., Tulet, P., Escobar, J., Leriche, M., Barthe, C., Aouizerats, B., Augros, C., et al.: Overview of the Meso-NH model version 5.4 and its applications, Geosci. Model Dev., 11, 1929–1969, 2018.

Larson, V. E., Golaz, J.-C., and Cotton, W. R.: Small-Scale and Mesoscale Variability in Cloudy Boundary Layers: Joint Probability Density Functions, J. Atmos. Sci., 59, 3519–3539, https://doi.org/10.1175/1520-0469(2002)059<3519:SSAMVI>2.0.CO;2, 2002.

Le Trent, H. and Li, Z.-X.: Sensitivity of an atmospheric general circulation model to prescribed SST changes: feedback effects associated with the simulation of cloud optical properties, Clim. Dyn., 5, 175–187, https://doi.org/10.1007/BF00251808, 1991.

Lewellen, W. S. and Yoh, S.: Binormal Model of Ensemble Partial Cloudiness, J. Atmos. Sci., 50, 1228–1237, https://doi.org/10.1175/1520-0469(1993)050<1228:BMOEPC>2.0.CO;2, 1993.

Madeleine, J.-B., Hourdin, F., Grandpeix, J.-Y., Rio, C., Dufresne, J.-L., Vignon, E., Boucher, O., Konsta, D., Cheruy, F., Musat, I., Idelkadi, A., Fairhead, L., Millour, E., Lefebvre, M.-P., Mellul, L., Rochetin, N., Lemonnier, F., Touzé-Peiffer, L., and Bonazzola, M.: Improved Representation of Clouds in the Atmospheric Component LMDZ6A of the IPSL-CM6A Earth System Model, J. Adv. Model. Earth Syst., 12, e2020MS002 046, https://doi.org/10.1029/2020MS002046, 2020.

Mellor, G. L.: The Gaussian Cloud Model Relations, J. Atmos. Sci., 34, 356–358, https://doi.org/10.1175/1520-0469(1977)034<0356:TGCMR>2.0.CO;2, 1977.

Mellor, G. L. and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Rev. Geophys., 20, 851–875, https://doi.org/10.1029/RG020i004p00851, 1982.

Neggers, R.: A Dual Mass Flux Framework for Boundary Layer Convection. Part II: Clouds, J. Atmos. Sci., 66, 1489–1506, 2009.

Nieuwstadt, F. T. M. and Brost, R. A.: The Decay of Convective Turbulence, J. Atmos. Sci., 43, 532–546, https://doi.org/10.1175/1520-0469(1986)043<0532:TDOCT>2.0.CO;2, 1986.

Pergaud J., Masson V., Malardel S., and Couvreux F.: A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction, Bound.-Lay. Meteorol., 132, 83–106, https://doi.org/10.1007/s10546-009-9388-0, 2009.

Price, J. D.: A study of probability distributions of boundary-layer humidity and associated errors in parametrized cloud-fraction, Q. J. R. Meteorol. Soc., 127, 739–758, https://doi.org/10.1002/qj.49712757302, 2001.

Rio, C. and Hourdin, F.: A Thermal Plume Model for the Convective Boundary Layer: Representation of Cumulus Clouds, J. Atmos. Sci., 65, 407–425, https://doi.org/10.1175/2007JAS2256.1, 2008.

Rio, C., Hourdin, F., Grandpeix, J.-Y., and Lafore, J.-P.: Shifting the diurnal cycle of parameterized deep convection over land, Geophys. Res. Lett., 36, https://doi.org/10.1029/2008GL036779, 2009.

Rio, C., Hourdin, F., Couvreux, F., and Jam, A.: Resolved Versus Parametrized Boundary-Layer Plumes. Part II: Continuous Formulations of Mixing Rates for Mass-Flux Schemes, Bound.-Lay. Meteorol., 135, 469–483, https://doi.org/10.1007/s10546-010-9478-z, 2010.

Sandu I. and Stevens B.: On the factors modulating the stratocumulus to cumulus transitions, J. Atmos. Sci., 68, 1865–1881, https://doi.org/10.1175/2011JAS3614.1, 2011.

Smith, R. N. B.: A scheme for predicting layer clouds and their water content in a general circulation model, Q. J. R. Meteorol. Soc., 116, 435–460, https://doi.org/10.1002/qj.49711649210, 1990.

Sommeria, G. and Deardorff, J. W.: Subgrid-Scale Condensation in Models of Nonprecipitating Clouds, J. Atmos. Sci., 34, 344–355, https://doi.org/10.1175/1520-0469(1977)034<0344:SSCIMO>2.0.CO;2, 1977.

Stevens, B. and Seifert, A.: Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection, J. Meteor. Soc. Japan, 86, 143–162, 2008.

Stowasser, M., Hamilton, K., and Boer, G. J.: Local and Global Climate Feedbacks in Models with Differing Climate Sensitivities, J. Clim., 19, 193–209, 2006.

Sundqvist, H.: A parameterization scheme for non-convective condensation including prediction of cloud water content, Q. J. R. Meteorol. Soc., 104, 677–690, https://doi.org/10.1002/qj.49710444110, 1978.

Tan, Z., Kaul, C. M., Pressel, K. G., Cohen, Y., Schneider, T., and Teixeira, J.: An Extended Eddy-Diffusivity Mass-Flux Scheme for Unified Representation of Subgrid-Scale Turbulence and Convection, J. Adv. Model. Earth Syst., 10, 770–800, https://doi.org/10.1002/2017MS001162, 2018.

Tompkins, A. M.: A Prognostic Parameterization for the Subgrid-Scale Variability of Water Vapor and Clouds in Large-Scale Models and Its Use to Diagnose Cloud Cover, J. Atmos. Sci., 59, 1917–1942, https://doi.org/10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2, 2002.

Tompkins, A. M.: Impact of temperature and humidity variability on cloud cover assessed using aircraft data, Q. J. R. Meteorol. Soc., 129, 2151–2170, https://doi.org/10.1256/qj.02.190, 2003.

van Zanten, M., Stevens, B., Nuijens, L., Siebesma, A., Ackerman, A., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D., Mechem, D., Nakamura, K., Noda, A., Shipway, B., Slawinska, J., Wang, S., and Wyszogrodzki, A.: Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO, J. Adv. Model. Earth Syst., 3, https://doi.org/10.1029/2011MS000056, 2011.

Watanabe, M., Emori, S., Satoh, M., and Miura, H.: A PDF-based hybrid prognostic cloud scheme for general circulation models, Clim. Dyn., 33, 795–816, https://doi.org/10.1007/s00382-008-0489-0, 2009.

Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., and Yamazaki, K.: History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble, Clim. Dyn., 41, 1703–1729, https://doi.org/10.1007/s00382-013-1896-4, 2013.

Williamson, D., Blaker, A. T., Hampton, C., and Salter, J.: Identifying and removing structural biases in climate models with history matching, Clim. Dyn., 45, 1299–1324, https://doi.org/10.1007/s00382-014-2378-z, 2015.

Yamada, T.: Simulations of Nocturnal Drainage Flows by a q2l Turbulence Closure Model, J. Atmos. Sci., 40, 91–106, https://doi.org/10.1175/1520-0469(1983)040<0091:SONDFB>2.0.CO;2, 1983.

https://doi.org/10.5194/egusphere-2025-5798
Preprint. Discussion started: 17 December 2025

Zelinka, M. D., Randall, D. A., Webb, M. J., and Klein, S. A.: Clearing clouds of uncertainty, Nat. Clim. Change., 7, 674–678, https://doi.org/10.1038/nclimate3402, 2017.