Journal of the Atmospheric Sciences

Deep convection triggering by boundary layer thermals. Part 1: LES analysis and stochastic triggering formulation

--Manuscript Draft--

Manuscript Number:	
Full Title:	Deep convection triggering by boundary layer thermals. Part 1: LES analysis and stochastic triggering formulation
Article Type:	Article
Corresponding Author:	Nicolas Rochetin, Ph.D Columbia University New York, NY UNITED STATES
Corresponding Author's Institution:	Columbia University
First Author:	Nicolas Rochetin, Ph.D
Order of Authors:	Nicolas Rochetin, Ph.D
	Fleur Couvreux
	Jean-Yves Grandpeix
	Catherine Rio
Abstract:	This paper proposes a stochastic formulation of the deep convection triggering by boundary layer thermals in a GCM grid cell. For that, a statistical analysis of a LES cloud field (Couvreux et al. 2012) in a case of transition from shallow to deep convection over a semi-arid land (Niamey, NIGER) is made. Since observations (Lothon et al. 2011) suggest that triggering occurs over the largest cloud base cross-sections, the analysis only focuses on the largest clouds of the study domain. Based on the dynamical and geometrical properties at cloud base, we first propose a new computation of the Available Lifting Energy (ALE) that must exceed the Convective Inhibition (CIN) for triggering. Another triggering condition is then required to make the triggering effective; it is based on the comparison between the distribution law (or PDF) of the maximum cross-sections of the domain, and an arbitrary threshold cross-section. The exceeding of this threshold cross-section is explicitly represented through a random number that has to excess a no-trigger probability, which is computed from the PDF of maximum sizes. Therefore, this new stochastic formulation integrates the whole transition process from the first cloud to the first convective cell, and can be decomposed in 3 steps: (i) the appearance of clouds, (ii) the inhibition layer crossing and (iii) the effective triggering
Suggested Reviewers:	Robert Plant r.s.plant@rdg.ac.uk
	George Craig George.Craig@dlr.de
	Anthony Del Genio anthony.d.delgenio@nasa.gov
	Leo Donner Leo.J.Donner@noaa.gov
	Jon Petch Jon.Petch@metoffice.gov.uk
	Robert Pincus

Powered by Editorial Manager® and Preprint Manager® from Aries Systems Corporation

	robert.pincus@colorado.edu
	Anton Beljaars anton.beljaars@ecmwf.int

Powered by Editorial Manager® and Preprint Manager® from Aries Systems Corporation

Page and Color Charge Estimate Form

Click here to download Page and Color Charge Estimate Form: Page_Charge_Estimation_Form_Part_1.pdf

Manuscript (non-LaTeX)
Click here to download Manuscript (non-LaTeX): ale1_ams.pdf

Deep Convection triggering by boundary layer thermals:

Part I: LES analysis and stochastic triggering formulation

Nicolas Rochetin *

Laboratoire de Météorologie Dynamique, Paris, France

Fleur Couvreux

CNRM-GAME, Météo-France and CNRS, Toulouse, France

Jean-Yves Grandpeix

Laboratoire de Météorologie Dynamique, Paris, France

Catherine Rio

Laboratoire de Météorologie Dynamique, Paris, France

* *Corresponding author address:* Nicolas Rochetin, Laboratoire de Météorologie Dynamique, Boite 99, 4, place Jussieu, F-75252 PARIS CEDEX 05, FRANCE

E-mail: nicolas.rochetin@lmd.jussieu.fr

ABSTRACT

This paper proposes a stochastic formulation of the deep convection triggering by boundary layer thermals in a GCM grid cell. For that, a statistical analysis of a LES cloud field (Couvreux et al. 2012) in a case of transition from shallow to deep convection over a semi-arid land (Niamey, NIGER) is made. Since observations (Lothon et al. 2011) suggest that triggering occurs over the largest cloud base cross-sections, the analysis only focuses on the largest clouds of the study domain. Based on the dynamical and geometrical properties at cloud base, we first propose a new computation of the Available Lifting Energy (ALE) that must exceed the Convective Inhibition (CIN) for triggering. Another triggering condition is then required to make the triggering effective; it is based on the comparison between the distribution law (or PDF) of the maximum cross-sections of the domain, and an arbitrary threshold cross-section. The exceeding of this threshold cross-section is explicitly represented through a random number that has to excess a no-trigger probability, which is computed from the PDF of maximum sizes. Therefore, this new stochastic formulation integrates the whole transition process from the first cloud to the first convective cell, and can be decomposed in 3 steps: (i) the appearance of clouds, (ii) the inhibition layer crossing and (iii) the effective triggering

1. Introduction

Many features of tropical deep convection are accounted for by the quasi-equilibrium hypothesis (QE). According to this hypothesis deep convection responds very rapidly to changes in tropospheric stability due to large scale circulation and radiative forcing, so that the tropical troposphere is permanently close to a state of equilibrium. However, several authors have emphasized that an atmosphere in permanent QE state would exhibit an exceedingly low variability at small scale and at large scale (Neelin et al. (2008), Jones and Randall (2011) , Raymond and Herman (2011)). Using CRM simulations Raymond and Herman (2011) showed that the response of deep convection to a pertubation was very fast (hours) only in the lower half of the troposphere while it was much slower in the upper half. This points to the importance of the depth of moist convection: the QE hypothesis is valid in the region of the troposphere reached by cumulus and congestus clouds, while it is not in the region reached only by deep convection. The basic elementary components of deep convection, the cumulonimbus clouds, are efficient processes warming up the upper troposphere: when present, they bring back the CAPE to very low values in matter of hours. However, they are short-lived (about 30 minutes) and are present only as long as the triggering of new elements goes on. It is then tempting to suppose, following Neelin et al. (2008) and Stechmann and Neelin (2011), that the main reason why deep convection departs from QE is that there are lapses of time where triggering of new convective cells does not occur and where the upper troposphere may wander freely away from QE.

Subcloud lifting processes and convective inhibition (CIN) are known to exert a strong control on deep convection onset and intensity, modulating the entropy flux from the boundary layer to the free troposphere (Emanuel and Bretherton (1994)). Mapes (2000) assumes that deep convective trigger occurs when turbulent kinetic energy in the boundary layer (the ”triggering energy”) is sufficient to overcome CIN. With this picture in mind the question of moist convection occurence and variability in the tropics is strongly dependent on the departure of the troposphere from QE states and thus on action of boundary layer processes on deep convection triggering.

The present series of paper is in the wake of these ideas and adresses the questions of deep convection triggering and of its representation in climate models.

The role of the quasi-equilibrium (QE) hypothesis in the development of deep convection parametrizations in climate models cannot be over-emphasized. It is at the heart of the very concept of parametrizaton: it makes it possible to express the action of deep convection processes as a function of large scale conditions to drive the system towards a state of equilibrium.

However several authors emphasized that departing from QE was a key step to simulate a correct climate variability. Obviously, releasing the constrain of QE yields extra degrees of freedom. The fact that that this yields an increased variability is by no means obvious. However, it seems (Jones and Randall (2011), Neelin et al. (2008)) that adheering strongly to QE leads to an insufficient variability, while attempts to perturb the system away from QE did increase variability.

As described in Jones and Randall (2011) (see also Xu et al. (1992)) two family of methods have been used to drive the local atmospheric system away from QE: (i) in the super-parametrization technique the CRM embedded within each GCM grid cell exhibits internal variability (and sensitivity to initial conditions), thus providing a variability around QE (see also Plant and Craig (2008)), who emphasize the variability provided by CRMs for given large scale conditions); (ii) stochastic parametrizations (Neelin et al. (2008)) may also be used either to perturb the deep convective closure or trigger or to perturb randomly the tendencies; in the first instance (Stechmann and Neelin (2011)) the convective parametrization is still pulling the local tropospheric system towards equilibrium but it does so in a less efficient way, especially it does not when convection is not active (either because trigger is perturbed or because the perturbed closure led to a break in convective activity); in the second instance (Palmer (2012)) the system is no longer driven towards QE but to a target moving randomly around QE (notice that this approach is not respecting conservation laws).

In the present paper we are concerned with ordinary parametrizations, not with super-parametrizations. Moreover, we want to use parametrizations respecting strictly conservation laws. Then, following Neelin et al. (2008), we shall assume that moving away from QE occurs mainly when deep convection is not active. Thus determining the period of activity of deep convection is a key problem for the representation of climate variability.

In observations and in high resolution simulations of moist convection, the onset of deep convection is the time when cumulus clouds reach the upper troposphere, displaying a fast jump from a shallow state to a much deeper state. Prior to this jump the boundary layer enters a transition regime during which cumulus clouds become gradually deeper while remaining in the low troposphere. Thus the onset of deep convection appears as the sudden emergence of a congestus or cumulonimbus cloud in a cumulus field. Chaboureau et al. (2004) show that during the transition phase the updraught vertical velocities at cloud base are large enough for the plumes to overcome the convective inhibition but that entrainment of exceedingly dry air limits their vertical development to the lower free troposphere. It is only when the lower free troposphere is moist enough that the sharp transition to deep convection occurs. Thus they propose a two-step trigger in which stability and moisture are the two critical variables controlling the transition.

We shall follow this idea and attempt to design a two-step trigger applicable to any present boundary layer parametrization using a mass-flux scheme, and coupled with deep convection.

Actually, most of currents GCMs (General Circulation Model) miss this transition phase, and consequently, poorly represent the diurnal cycle of deep convection over land. They simulate a precipitation peak around noon while according to observations it is later in the afternoon (Yang and Slingo (2001), Guichard et al. (2004) and Bechtold et al. (2004)). Guichard et al. (2004) analyse this shift of the diurnal cycle of precipitation as simulated by SCMs relative to high resolution simulations. According to the authors, this is due to the fact that the gradual moistening of the low free troposphere by overshooting cumulus is not represented in GCM parametrizations. Hence, currents GCM cannot capture the succession of dry, shallow and deep convection regimes.

The question of triggering arises when trying to treat separately shallow and deep convective regimes. Especially over lands, in which local processes of shorter time-scales gain influence in controlling the convection life cycle. Furthermore, the continental boundary layer is, on average, more developed than the marine boundary layer (Medeiros et al. (2005)) and capped by a stronger inhibition layer (CIN, Convective INhibition). In such cases, considering the fact that the parcel cannot reach its LFC without some small-scale (i.e subgrid) processes, a subgrid "Trigger function" (Kain and Fritsch (1992)) must be represented. Kain and Fritsch (1992) showed that those "Trigger functions" deeply affect the GCMs and Numerical Weather Prediction (NWP) models ability to forecast the life cycle of deep convection.

Some convection schemes, like Kuo (1974), were designed to initiate convection whenever a threshold value in mass or moisture convergence is exceeded in a grid point. Others, assuming the Quasi-Equilibrium hypothesis (Arakawa and Schubert (1974)) triggers of-the-moment the large scale conditions deviate too far from equilibrium state, in such a way that deep convection adjusts itself quasi-instantaneously to the large scale perturbation. Some other schemes trigger deep convection whenever the buoyancy becomes positive at the vicinity of the cloud base level, and Mapes (2000) introduces in a simplified model a trigger energy depending on the subgrid scale turbulence fluctuations.

In the current version of the atmospheric component of the LMD's GCM (LMDZ5B, Hourdin et al. (2012)), we use the ALE/ALP framework (Grandpeix and Lafore (2010), Rio et al. (2009), Rio et al. (2012)) in which deep convection is coupled with sub-cloud processes thanks to two variables: the Available Lifting Energy (ALE, expressed in J kg^{-1}) and the Available Lifting Power (ALP, expressed in W m^{-2}). Convection triggering and closure are expressed in terms of ALE (Convection is triggered when ALE is larger than CIN) and ALP (cloud base mass flux is proportionnal to ALP) respectively. In the LMDZ5B model, the only lifting processes are the boundary layer thermals and the density currents. The ALE is the maximum of the two lifting energies and ALP is the sum of the two lifting powers. The present paper is only concerned with deep convection triggering hence only with the ALE variable. Moreover, since we are specifically interested in convection intiation, only the lifting energy due to boundary layer thermals has to be considered. In the current versions of the LMDZ5 GCM it is equal to the maximum vertical kinetic energy in the thermal plume. This maximum is generally found near cloud top so that the current implementation of ALE takes somehow into account the size of the cumulus clouds. However, as will turn out in 6, this is quite insufficient to describe the transition regime.

In the present paper we stick to the ALE/ALP framework. Hence our purpose is to modify $\mathrm{ALE_{BL}}$ provided by the boundary layer scheme so that it accounts as well for the lower free troposphere humidity as for the kinetic energy of the boundary layer thermals. The key question investigated is to know what are the critical parameters of the boundary layer which control the transition from a shallow cumulus regime to a deep convective regime.

Several studies using Cloud Resolving Models have been used to characterize this complex transition from shallow to deep convection and gain some insights on what variables control the deep convection triggering. While Chaboureau et al. (2004) proposes that deep convection initiates when a variable called "the Normalized Saturation Deficit" (NSD) at the cloud base reaches its minimum (NSD being strongly linked to the cloud cover, triggering occurs when the cloud cover reaches a critical value), Wu et al. (2009) shows that the virtual temperature profile of the average cloud is a key factor, and Khairoutdinov and Randall (2006) stress the importance of horizontal cloud size. Thus several parameters seem to play a key role in deep convection triggering: (i) at cloud base, the humidity of the troposphere, the cloud cover, the size of individual clouds; (ii) above cloud base, the thermodynamical properties of cumulus clouds.

Given all these questions concerning the transition to deep convection, we tackle the problem of the representation of deep convection triggering in climate models. Thanks to LES (Large Eddy Simulation) datas in a continental case of transition from shallow to deep convection, we extract the statistical properties of the thermal plumes at the cloud base level and propose a new computation of $\mathrm{ALE}_{\mathrm{BL}}$ (Grandpeix and Lafore (2010)). The goal is to propose a simple formulation of the triggering process, easily integrable in a GCM. This new formulation describes the whole transition process and in particular the stochastic nature of the triggering.

Next part describes the theoretical framework and section 3 the method. The cross-section spectrum of the thermal plumes inside the domain is studied in section 4 and the vertical velocity spectrum inside the plumes in section 5. The $\mathrm{ALE}_{\mathrm{BL}}$ computation is given in section 6. The triggering formulation is proposed in section 7 and some final comments are given in section 8.

2. Statistical thermal plumes

a. Single plume versus statistical plumes approaches

The single plume approach is commonly used in present day boundary layer parametrizations with a mass-flux closure. It is justified when considering a quasi-steady regime, as for example shallow cumulus in a subsiding atmosphere. Considering a GCM grid area, the cumulus clouds are numerous enough for neglecting the fluctuations around mean, thus, the "bulk plume" may be a correct predictor of their collective effect. In such cases, the spectrum of plume sizes does not play a significant role in the representation of boundary layer processes, as is the case when computing heat, moisture and buoyancy fluxes. In the single plume approach, there is a bulk plume of cross section S_{tot}, covering a fractional area α_{tot} and a single mean vertical velocity profile inside w'_{u} and outside w'_{e}.

However, this approach is not enough when plume sizes come explicitly into play, for instance when we assume later (see the following sections) that the triggering of deep convection is due to the largest thermal plumes. Indeed, in a transition period, one can expect that fluctuations around the mean become more significant and have to be considered. Therefore, one has then to add a statistical (or spectral) approach to the bulk formulation. Fig 1 illustrates the differences between the single plume and the statistical approaches.

Neggers et al. (2003) and Rodts et al. (2003) shed light on the properties of the cloud field by the use of aircrafts measurements, satellite data and Large Eddy Simulations. It is mentioned that many distribution laws were suspected to fit the cloud size spectrum over the domain, among which the exponential law, the lognormal and some other power laws. Craig and Cohen (2006) proposed an exponential PDF $\mathcal{P}_m$ for representing the individual cloud mass flux spectrum ($\mathcal{P}_m(m) = \frac{1}{\langle m \rangle} \exp\left(\frac{-m}{\langle m \rangle}\right)$, where $\langle m \rangle$ represents the mean mass flux over the plume population). In order to implement this statistical model in a convective parametrization, Plant and Craig (2008) assumed furthermore that, in the vicinity of cloud base, vertical velocities in plumes were independent of the plume size. Then the PDF $\mathcal{P}_s$ of plume sizes is also exponential (since $\mathcal{P}_m(m)\rho\overline{w'_{\mathrm{p}}} = \mathcal{P}_s(s)$):

$$\mathcal{P}_s(s) = \frac{1}{\langle s \rangle} \exp\left(\frac{-s}{\langle s \rangle}\right) \tag{1}$$

The exponential spectrum hypothesis for the mass fluxes has been validated by Craig and Cohen (2006b) thanks to Cloud Resolving Model (CRM) simulations. The uniformity of the mean vertical velocity at cloud base is reported in observational studies such as Warner (1970); it is sometimes assumed in modelling studies (Donner (1993) and, of course, Craig and Cohen (2006b)). Thus the exponential spectrum appears as a likely property of cloud sizes at cloud base. Its relevance will be assessed in subsection 4.a.

The plume's internal fluctuations may also be considered. Emanuel (1991) recalls that pioneer aircraft measurements have shown that in-cloud fluctuations exhibit a typical length-scale of 100m for temperature, liquid water and water vapor. Malkus (1954) and Warner (1970) investigated the properties of in-cloud drafts by means of aircraft measurements. They revealed that vertical velocity fluctuations were, at least, as large as the mean value across the cloud section.

Thereby, those studies suggest that intra-thermal (vertical velocity), as well as inter-thermal (cross-section) fluctuations must be considered. Our aim is now then to propose a corresponding theoretical representation of the thermal plume field.

b. Notations and definitions

In this study we consider as a plume an ensemble of drafts underlying a cloud. This consideration implies that some plumes may have a negative cross-section average velocity, and so cannot be qualified of "thermal plumes" (since a thermal plume is buoyant). Nevertheless, this kind of plumes represent a negligible part of the data which will be exploited in this study.

We consider a domain (a grid cell) of area S_{d} in which N_{tot} plumes are present, covering an area S_{tot} at the lifting condensation level (LCL, or cloud base level). The fractional area covered by the plumes will be denoted α_{tot} ($\alpha_{\mathrm{tot}} = \dfrac{S_{\mathrm{tot}}}{S_{\mathrm{d}}}$). At a given level, the domain is divided into several regions: (i) the individual plumes (p_i, i = 1, N_{tot}), (ii) the plume environment (e). Generally, the overbar ($\overline{x}$) denotes the average over a horizontal region, which may be the whole domain or the plume environment or a given plume (e.g. $\overline{w}$, $\overline{w_{\mathrm{e}}}$, $\overline{w_{\mathrm{p,i}}}$ and $\overline{w_{\mathrm{p}}}$ are the large scale vertical velocity and the mean vertical velocities over the plume environment, over plume i, and over all plumes respectively), while the brackets ($\langle x \rangle$) denote the arithmetic average over the population of thermal plumes.

1) Geometry:

The geometry of plume number i is characterized by the altitudes $z_{\mathrm{lcl,i}}$ and $z_{\mathrm{p,i}}$ of its cloud base and cloud top respectively and by its cross-section s_i at cloud base. As suggested by observational studies, the plume at cloud base may considered as made of independent elementary drafts with typical dimension of a few 100 m. Since the LES horizontal resolution is 200 m, we arbitrary assume, for simplicity, that the cross section of the elementary drafts is š = $4.10^4 \mathrm{m}^2$. We will show in Sec.6 that this arbitrary parameter is of secondary importance, and in Sec c an estimation of a potentially realistic value for š will be suggested.

A plume i is then composed of n_i independent drafts of cross-section š (and underlies a cloud). The number of elementary drafts in a plume i is $n_\mathrm{i} = \frac{s_\mathrm{i}}{\check{s}}$. In the following, this quantity will be named the dimensionless cross-section or the number of drafts per plume.

2) Vertical velocities:

For an air parcel $P(x, z)$ inside plume i two decompositions of the vertical velocity will be used. (i) first the usual decomposition in the domain average $\overline{w}$ and a fluctuation $w'_\mathrm{p,i}$ yields $w_\mathrm{p,i}(x, z) = \overline{w} + w'_\mathrm{p,i}(x, z)$; (ii) then the fluctuation $w'_\mathrm{p,i}(x, z)$ will be further decomposed into a plume average $\overline{w'_\mathrm{p,i}}$ and a second order fluctuation $w''_\mathrm{p,i}$:

$$w_\mathrm{p,i}(x, z) = \overline{w} + \overline{w'_\mathrm{p,i}} + w''_\mathrm{p,i}(x, z) \tag{2}$$

A similar development gives, for any parcel located in the subsiding environment:

$$w_\mathrm{e}(x, z) = \overline{w} + w'_\mathrm{e} + w''_\mathrm{e}(x, z) \tag{3}$$

The present decomposition is illustrated in Fig 2.

3) Main features of the plume population:

- From the individual plume average vertical velocities one may compute the mean vertical velocity of the whole plume field:

$$\overline{w'_\mathrm{p}} = \frac{1}{S_\mathrm{tot}} \sum_{\mathrm{i}=1}^{N_\mathrm{tot}} s_\mathrm{i} \overline{w'_\mathrm{p,i}} \tag{4}$$

  Similarly, the mean second and third order non-centered moments are defined, respectively, by;

$\overline{w_p'^2} = \dfrac{1}{S_{tot}} \sum_{i=1}^{N_{tot}} s_i \overline{w_{p,i}'^2}$ and $\overline{w_p'^3} = \dfrac{1}{S_{tot}} \sum_{i=1}^{N_{tot}} s_i \overline{w_{p,i}'^3}$.

- For each plume i, the vertical velocity standard deviation $\Gamma_{w'_{p,i}}$ and skewness are, respectively:

$$\Gamma_{w'_{p,i}} = \sqrt{\overline{w_{p,i}'^2} - \overline{w'_{p,i}}^2} \tag{5}$$

$$\Phi_{w'_{p,i}} = \frac{\overline{w_{p,i}'^3} - 3\overline{w_{p,i}'^2}\,\overline{w'_{p,i}} + 2\overline{w'_{p,i}}^3}{\Gamma^3_{w'_{p,i}}} \tag{6}$$

- The arithmetical mean cross-section over the plume population gives:

$$\langle s \rangle = \frac{1}{N_{tot}} \sum_{i=1}^{N_{tot}} s_i \tag{7}$$

And similarly the mean cloud base ($\langle z_{lcl} \rangle = \frac{1}{N_{tot}} \sum_{i=1}^{N_{tot}} z_{lcl,i}$) and cloud top altitudes ($\langle z_{top} \rangle = \frac{1}{N_{tot}} \sum_{i=1}^{N_{tot}} z_{p,i}$).

- Finally is defined the arithmetical mean plume average velocity over the plume population:

$$\langle \overline{w'_p} \rangle = \frac{1}{N_{tot}} \sum_{i=1}^{N_{tot}} \overline{w'_{p,i}} \tag{8}$$

And similarly for the second and third order moments.

3. Data and Methodology

a. Case description

The case study investigated here is the AMMA case of 10 July 2006 where a small and short-living convective cell developped over Niamey (Lothon et al. (2011)). The whole transition has been caught by several ground-based instruments (radar, wind profiler and atmospheric soundings) and completed by satellite data. This case study concerns a typical case of transition from shallow to deep convection over semi-arid land with a high Bowen Ratio ($Bo \approx 10$), and associated with an elevated cloud base ($z_{\mathrm{lcl}} \approx 2.5$ km). The structure of the boundary layer clouds is gradually evolving from a "cloud street" organization (morning till noon) to an isotropic structure composed of larger but more heterogenous cells (from noon to mid-afternoon). Around 15:40 LT, deep convective cells develop with associated cold pools. The author noted that the first convective cells developed over the largest horizontal cloud structures; this supports the relevance of the cloud base cross-section in describing the transition process and reinforces the hypothesis made in sec.1.

A modelling set-up has been developed to represent this case and a Large-Eddy simulation is able to represent the main observed features (Couvreux et al. (2012)).

b. The Large-Eddy Simulation

The simulation uses the LES version of the MESO-NH non-hydrostatic model developed by Lafore et al. (1998). The domain is $100 \times 100 \times 20$ km^3 with an horizontal resolution of 200 m, a stretched grid on the vertical (from 50 m to 2000 m) and periodic lateral boundary conditions. The forcing data were collected by the ARM (Atmospheric Radiation Measurement) mobile facility station based at the Niamey Airport. The simulation lasts from 06:00 LT to 18:00 LT, at which time the cold pool generated by deep convection becomes too large relative to the domain. The lower boundary condition consists in imposed homogeneous fluxes of heat and water vapour. However, the observations show a large positive surface temperature anomaly (around 5K), over which develops the first cell (at 15:40 LT). This heterogeneity is suspected to play an important role in the triggering of deep convection (enhancing mesoscale circulation and breeze convergence over the hot spot, see Taylor et al. (2011)). In order to simulate a similar onset of deep convection the model is forced with a low-level moisture convergence in the morning linked to the moonsoon flow, and a low-level ascent of 1.5 cm/s during the afternoon. With these conditions, the LES yields a trigger of deep convection around 16:30LT. This simulation has been evaluated against observations in Couvreux et al. (2012).

The material used in the present study consists in various fields extracted from the simulation every hour from 12:00LT to 18:00LT and for each cloud: (i) cloud top and cloud base altitudes; (ii) cloud base cross-section; (iii) cloud base average of vertical velocity, of its square and of its cube; (iv) cloud base maximum vertical velocity.

c. *Method*

Our final goal is to propose a new formulation of $\mathrm{ALE_{BL}}$, that is to compute a maximum kinetic energy provided by the thermal plumes, which has to be compared with CIN. This may be resumed as compute an estimate of the maximum speed over the domain. Then, the following LES analysis is aimed at finding the maximum value distribution for the plume cross-sections and for the plume vertical velocities, for computing $\mathrm{ALE_{BL}}$.

Our starting hypothesis are (i) a two-step triggering (as suggested by Chaboureau et al. (2004)) and (ii) that the cloud base cross-section plays a crucial role in controlling the deep convection triggering (see Lothon et al. (2011)).

4. LES analysis: maximum cross-section distribution at the cloud base

a. *Cross-section spectrum: $\mathcal{P}(s)$*

Mapes (2000), Khairoutdinov and Randall (2006), Rio et al. (2009), Grandpeix et al. (2010) and Del Genio and Wu (2010) suggested that the subcloud layer processes play a key role in producing the mechanical forcing, which lifts the parcel from the surface layer to its Level of Free Convection (LFC). In a conditionally unstable atmosphere, the Lifting

Condensation Level (LCL) nearly corresponds to the top of the boundary layer and to the bottom of the convective inhibition layer (CIN). We shall consider it as the most relevant level where to represent the couplings between boundary layer processes and deep convection. Consequently the present study focuses on the thermal plume properties at cloud base.

Fig 3 displays with logarithmic coordinates the N-normalized cross-section spectrum ($\mathcal{N}_n = N_{\text{tot}}\mathcal{P}$, where $\mathcal{P}$ is the PDF) at two different times. Since the strong peaks at low cross-sections are incompatible with exponential distributions, the spectra are fitted with double exponential PDFs, also displayed in Fig 3:

$$\mathcal{N}_n(n) = \frac{N_1}{n_1} \exp\left(\frac{-n}{n_1}\right) + \frac{N_2}{n_2} \exp\left(\frac{-n}{n_2}\right) \tag{9}$$

Where $n_1 = \frac{S_1}{\check{s}}$ and $n_2 = \frac{S_2}{\check{s}}$ are the average, dimensionless cross-section of each type of plumes.

If considering the cross-section $s = n\check{s}$ (instead of the dimensionless cross-section n) the N-normalized distribution becomes:

$$\mathcal{N}_s(s) = \frac{N_1}{S_1} \exp\left(\frac{-s}{S_1}\right) + \frac{N_2}{S_2} \exp\left(\frac{-s}{S_2}\right) \tag{10}$$

Where N_1 and N_2 represent the total number of clouds of each type, and S_1 and S_2 their average cross-sections.

The two exponential PDF corresponding to each type of distribution are:

$$\mathcal{P}_1(s) = \frac{1}{S_1} \exp\left(\frac{-s}{S_1}\right) \tag{11}$$

$$\mathcal{P}_2(s) = \frac{1}{S_2} \exp\left(\frac{-s}{S_2}\right) \tag{12}$$

Category 1 gathers a very large population of small cumulus clouds, presumably topping the smallest CBL's thermal plumes. Their cloud base area is ranging from $n = 1$ to 40 drafts (see Fig 3), and their depth fluctuates between 50 m to 500 m (not shown). The corresponding thermals play an important role since they moisten the lower free troposphere and thus favour the growth of future cumulus clouds. However, due to their small size, we do not expect them to contribute to the triggering of deep convection.

Category 2 concerns small and intermediate clouds building the distribution tail (i.e the right branch of the N-PDF plotted in Fig 3). Their cloud base area is ranging from $n = 1$ to 160 drafts (see Fig 3), and their depth fluctuates between 500 m to 2000 m (not shown). Knowing that size is an important proxi for describing the transition phase, we expect that type-2 plumes are the only one category of interest.

The remaining class of clouds (not shown) is not represented by the fitting function given in Eq 10, it concerns deep convective clouds (i.e congestus and cumulonimbus appearing after 16:30 LT in the LES).

b. Cross-section spectrum evolution

Fig 4 a) represents the fitting N-normalized PDF evolution (defined in Eq 10) for the afternoon hours of the simulation (12:00 to 18:00 LT). The slope of the exponential distribution of type-2 clouds decreases with time, while it does not seem to vary appreciably for type-1 clouds.

Fig 4 b) and c) gives further details about the evolution of each cloud population. First, Fig 4 b) shows that N_2 is decreasing all along the transition period. It is less trivial to extract a tendency for population 1, as errorbars are very important at 12:00 LT (only small clouds are present) and 13:00 LT: at those times the populations 1 and 2 are more or less confounded. Therefore, the population N_1 as well as the average cross-section S_1 stays nearly constant from 13:00 LT up to the trigger time 16:30 LT. On the contrary, Fig 4 c) shows that S_2 increases form 12:00 LT up to 18:00 LT. In other words, the transition from shallow to deep convection gives rise to fewer structures but larger ones, suggesting that the gradual drying and deepening of the boundary layer (Lothon et al. (2011) and

Couvreux et al. (2012)) is correlated with fewer but larger plumes feeding deeper cumulus clouds. Since the tendencies of population N_2 and S_2 are of opposite signs, the fractional coverage α_{tot} (as suggested by Chaboureau et al. (2004) through the NSD) is *a priori* not the best proxi for describing the transition process. The present study shows that the average cross-section is a more pertinent predictor. This also reinforces the relevance of considering spectral plumes rather than a bulk plume, and both the plume population and their mean cross-section separately.

The largest plumes are the key elements of the transition. To make this more precise, in the following, we study the statistical properties of the type-2 plumes. However, since the cloud base cross-section is a variable absent from boundary layer parametrizations using the single plume approach, we first turn to establishing empirical relationships between cloud base cross-section and vertical cloud development.

c. *Vertical vs horizontal cloud development*

The cloud base altitude and the cloud depth are largely determined by the thermodynamic profiles of the environment and the air parcel. Hence, this subsection dedicated to study the potential link existing between the vertical characteristics of the type-2 cumulus, and their horizontal lengthscale. The typical horizontal lengthscale of cloud i is $\sqrt{s_{\mathrm{i}}}$, with statistical mean over the population $\left\langle\sqrt{s_{\mathrm{i}}}\right\rangle$ (similar to Eq 7). The vertical lengthscales of the type-2 cloud field are given by the mean cloud base $\langle z_{\mathrm{lcl}}\rangle$ and cloud top $\langle z_{\mathrm{top}}\rangle$ altitudes.

We assume that there is a linear relationship between the root mean square plume diameter $\sqrt{S_2}$, the boundary layer height and the mean cloud thickness. Assuming that the mean cloud base altitude $\langle z_{\mathrm{lcl}}\rangle$ is a good approximation for the boundary layer height, this linear relation reads:

$$\sqrt{S_2} = a(\langle z_{\mathrm{top}}\rangle - \langle z_{\mathrm{lcl}}\rangle) + b\,\langle z_{\mathrm{lcl}}\rangle \tag{13}$$

where a and b are parameters to be tuned The first term corresponds to a simple cloud model, in which the cloud width is proportional to its height. The second term accounts for the fact that the aspect ratio of the PBL coherent structures may be considered fixed (and cloise to 2) so that thicker boundary layers display larger cells which, in turn, allow wider clouds. Coefficients a and b were determined by fitting $\sqrt{S_2}(t)$, $\langle z_{\text{top}} \rangle (t)$ and $\langle z_{\text{lcl}} \rangle (t)$ at times t in the range 12:00 to 16:00 LT (i.e before deep convection triggers) with Eq 13. The results were that coefficients a and b were poormy constrained but highly correlated, so that the results of the fit may be approximated by:

$$a \;=\; 1.5 \;\pm\; 0.8 \quad \text{and} \quad b \;=\; 0.25 \;-\; 0.1\,(a - 1.5)$$

So that we decided, for simplicity, to assume the arbitrary parameters $a \;=\; 1.$ and $b \;=\; 0.3$. Nevertheless, the large uncertainties should require more LES results to better constrain those values.

The quality of the fit with parameters $a = 1$ and $b = 0.3$ is visible in Fig 5, where the time-evolution of S_2 and its approximation $\mathcal{S}_2$ following Eq 13 are displayed. The difference between the two variables is within two standard deviations during the whole transition period (from 12:00 to 17:00). Another important result is that neither a nor b are compatible with zero (at two standard deviations for a and more than three for b): hence both the dependence on the boundary layer height and on the cloud thickness are necessary to determine the cloud base cross section.

The dependence of the typical size of the thermals at cloud base to the cloud base altitude is consistent with the constant aspect ratio of the boundary layer structures. Nevertheless, the fact that cloud base size is also correlated with the cloud depth is more tricky. We briefly discuss it through 2 diabatic processes. Those two processes may be responsible for the gradual widening of the cloud base, associated with the cloud layer deepening. Nevertheless, nothing in the present study could help to confirm and/or to dismiss one of those mechanisms, they are just presentend here as hypothesis.

The first process is the diabatic cooling by rain evaporation. High resolution simulations (Khairoutdinov and Randall (2006), Matheou et al. (2011), Boing et al. (2010)) showed that density currents induced by rain evaporation often appear before deep convection triggers, and play a role in the transition from shallow to deep convection. They suppress convection in their core and favour it on their edges by lifting the environment unstable air, in particular where density currents collide. This leads to the emergence of sparse but strong updrafts, yielding deeper and larger cloud structures. Even though this process is more and more suspected to govern the transition phase, Couvreux et al. (2012) noticed that the absence of evaporative cooling did not affect the deep convection triggering. The second one is the diabatic heating by condensation. Clark et al. (1986) evokes the fact that mid-size cumulus cloud heating can trigger gravity waves. Indeed, the author shows in a 2D framework, that convective heating by shallow convection can enhance the vertical propagation of gravity waves, which reflect on the tropopause and feed back on the low levels, selecting eddies whose horizontal lengthscale is comparable with the gravity wave spacing. Such a mechanism would operate a scale selection on thermal eddies and favour more sparse and larger horizontal structures during the transition.

Since we assume that the triggering occurs over the largest cloud of the domain, we now look at the maximum cross-section distribution of the type-2 plumes.

d. Maximum cross-section distribution: $\mathcal{P}_{max}(S_{max})$

As mentioned earlier, the type-2 plumes contains the largest thermals. At a given time, it is described by the PDF $\mathcal{P}_2(s)$ given in Eq 12. As shown in the Appendix, the cross-section $S_{\max}$ of the largest plume is a random variable with CCDF $\mathcal{F}_{\max}(S_{\max})$ given in Eq A7 from which may be derived a PDF $\mathcal{P}_{\max}(S_{\max})$ which verifies Eq A2, that is:

$$\mathcal{P}_{\max}(S_{\max}) = \frac{-d\mathcal{F}_{\max}(S_{\max})}{dS_{\max}} \tag{14}$$

The median $\mathcal{S}_{\max}$ of the $S_{\max}$ distribution is given by the approximate formula Eq A8 with $\pi_t = \ln(2)$:

$$\mathcal{S}_{\max} = S_2 \ln\left(\frac{N_2}{0.7}\right) \tag{15}$$

The maximum value PDF $\mathcal{P}_{\max}(S_{\max})$ and estimated $\mathcal{S}_{\max}$ at various times are plotted in Fig 6. Fig 6 a), $\mathcal{P}_{\max}(S_{\max})$ confirms that the distribution tail is increasing while transition occurs. Larger structures are appearing in the domain but coexist with still numerous small ones. As a result the cross-section spectrum is widening and prolongs itself to the high values of $S_{\max}$. Indeed, during the early afternoon (i.e 12:00 and 13:00) $\mathcal{P}_{\max}(S_{\max})$ is relatively peaked, and accordingly, $\mathcal{S}_{\max}$ fits well with the simulated values ($S_{\max}$) (see Fig 6 b)). Then later on, the distance $|\mathcal{S}_{\max} - S_{\max}|$ seems to increase with time, as predicted by the spectrum widening $\mathcal{P}_{\max}(S_{\max})$ (except at 15:00 for which, by chance, the estimated value is almost equal to the simulated one). This suggests that the exponential distribution $\mathcal{P}_2(s)$ pertinently describes the tail of the cross section density spectrum. This can be further assessed when looking at the maximum values CCDF ($\mathcal{F}_{\max}(S_{\max})$) histogram (not shown) is also compatible with a flat distribution for the 7 realisations (12:00 till 18:00) of $S_{\max}$ considered here. Fig 6 b) also shows that, from 17h onwards, the estimated values diverge from the simulated ones. This is consistent with Fig 5, indeed once deep convection has triggered Eq 13 is no more valid.

Therefore, from the PDF $\mathcal{P}_{\max}(S_{\max})$, we have extracted an estimator of the maximum cross-section $\mathcal{S}_{\max}$ of the domain, over which the triggering has the largest probability to occur. Since we are trying to compute an $\mathrm{ALE_{BL}}$, our goal is then to estimate the statistical maximum vertical velocity corresponding to that "maximum" plume (for getting a statistical $\mathrm{ALE_{BL,stat}}$). For that, we have first to characterize the vertical velocity spectrum of the plumes, and to find out the maximum vertical velocities distribution of the plumes.

5. LES analysis: maximum vertical velocities distribution of type-2 plumes

a. Method

In the preceeding section, we suggested that only type-2 plumes were involved in the transition process. Hence, we only focused on the dynamical properties of those plumes, and filter out the type-1 plumes. The aim is to extract from the LES data a large enough sample composed of type-2 plumes only. For that, the whole LES simulation (i.e from 12:00 to 18:00 LT) is gathered in a unique dataset of 9500 clouds. Then, only clouds which verifies $n > 40$ drafts (i.e $s > 4.10^4$ m^2, or $D > 1500$ m) are taken into account; the resulting dataset is made of about 900 clouds of type-2 exclusively. Finally, this dataset is divided in ten samples sorted into increasing cross-section.

Table 1 displays, for each sample k, characterized by its n range and composed of $N_{\mathrm{tot,k}} =$ 90 clouds, the arithmetic means $\langle . \rangle_{\mathrm{k}}$ of various fields at cloud base: (i) the average vertical velocity (defined in Eq 4), (ii) the second and third order non-centered moment, (iii) the maximum velocity $w'_{\mathrm{max,i}}$ and (iv) the standard deviation $\Gamma_{w'_{\mathrm{p,i}}}$ and the skewness $\Phi_{w'_{\mathrm{p,i}}}$.

b. Vertical velocity moments

Trying to characterize the vertical velocity distribution inside the plumes, we first look at the sensitivity of the velocity moments to the mean cross-section of each sample.

Fig 7 displays the pairs $\left[\langle n \rangle_{\mathrm{k}} : \left\langle \overline{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}}\right]$, $\left[\langle n \rangle_{\mathrm{k}} : \left\langle \overline{w'^2_{\mathrm{p}}} \right\rangle_{\mathrm{k}}\right]$ and $\left[\langle n \rangle_{\mathrm{k}} : \left\langle \overline{w'^3_{\mathrm{p}}} \right\rangle_{\mathrm{k}}\right]$. From Fig 7, it seems that the sample mean of the cross-section averaged vertical velocities $\left\langle \overline{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}}$ does not depend on the mean dimensionless cross-section of the sample $\langle n \rangle_{\mathrm{k}}$. Hence, whatever the sample k, $\left\langle \overline{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}} = \left\langle \overline{w'_{\mathrm{p}}} \right\rangle$, where $\left\langle \overline{w'_{\mathrm{p}}} \right\rangle$ is the arithmetic average over the whole population (composed of the 10 samples). Fig 7 also shows that it is also true for the second and third order non-centered moments of w'_{p}; that is, whatever the k, $\left\langle \overline{w'^2_{\mathrm{p}}} \right\rangle_{\mathrm{k}} = \left\langle \overline{w'^2_{\mathrm{p}}} \right\rangle$ and

$\left\langle \overline{w_{\mathrm{p}}'^3} \right\rangle_{\mathrm{k}} = \left\langle \overline{w_{\mathrm{p}}'^3} \right\rangle$. Extending this result to the individual plume scale yields that the plume averaged $\overline{w_{\mathrm{p,i}}'}$, second order moments $\overline{w_{\mathrm{p,i}}'^2}$ and third order moments $\overline{w_{\mathrm{p,i}}'^3}$ are not sensitive to the cross-section s_{i} of the plume considered.

Hence, whatever the plume i:

$$\overline{w_{\mathrm{p,i}}'} = \overline{w_{\mathrm{p}}'} \tag{16}$$

$$\overline{w_{\mathrm{p,i}}'^2} = \overline{w_{\mathrm{p}}'^2} \tag{17}$$

$$\overline{w_{\mathrm{p,i}}'^3} = \overline{w_{\mathrm{p}}'^3} \tag{18}$$

This gives, for the standard deviation and skewness:

$$\Gamma_{w_{\mathrm{p,i}}'} = \Gamma_{w_{\mathrm{p}}'}$$

$$\Phi_{w_{\mathrm{p,i}}'} = \Phi_{w_{\mathrm{p}}'}$$

Thus, the vertical velocity spectrum may be considered uniform over the plume field. This means that, at a given time, all the plumes of the domain exhibit the same spectrum for the draft velocities $\mathcal{P}(w_{\mathrm{p}}')$. Since the vertical velocity inside the plume results from a balance between buoyancy, pressure and friction forces (Simpson and Wiggert (1969)), this result suggests that all type-2 plumes (i.e with diameter greater than 1.5 km) experience the same balance of forces. Since lateral entrainment only involves the plume's peripheral, we may expect that, above a certain diameter, it does play a negligible role in the plume's motion (because the area of the external ring is much less than the plume's cross-section).

c. *PDF of vertical velocity* $\mathcal{P}(w_{p,i}')$

In order to increase the statistical significance, we decide temporarily (only in this subsection) to divide the dataset in only 5 samples of 180 plumes each, sorted into increasing cross-sections. Fig 8 displays the five corresponding histograms of mean vertical velocity at cloud base $\overline{w'_{\rm p,i}}$: the spectrum $\mathcal{P}(\overline{w'_{\rm p,i}})$ in the various samples look very similar to gaussians with widths $\Gamma_{\overline{w'_{\rm p,i}}}$ roughly proportionnal to $\frac{1}{\sqrt{n}}$. This is in agreement with the hypothesis that the elementary draught vertical velocities are independent gaussian random variables such as:

$$\mathcal{P}(w'_{\rm p,i}) = \frac{1}{\sqrt{2\pi}\Gamma_{w'_{\rm p,i}}} \exp\left(-\frac{(w'_{\rm p,i} - \overline{w'_{\rm p,i}})^2}{2\Gamma^2_{w'_{\rm p,i}}}\right) \tag{19}$$

Of course Eq 19 is not strictly valid since the skewness of the vertical velocity distribution is non-zero (see Table 1 column 8). However we shall assume that the vertical velocity PDF differs from a gaussian only in the low velocity region and that Eq 19 represents accurately the PDF in the region of the velocities relevant for triggering. This assumption will be justified a posteriori by the results of the next subsection.

Supplementary remarks

1.Reference cross-section š of the drafts:

It has been noticed that the approximated relationship $\Gamma_{\overline{w'_{\rm p,i}}} \approx 4\frac{1}{\sqrt{n}}$ was verified for the 5 samples considered. This suggests that the independent, gaussian drafts may have a reference cross-section nearly equal to $4\check{s} = 1.6 * 10^5$ m^2, that is a typical lenghtscale of $l = 400$ m (instead of the arbitray $l = 200$ m which has been chosen in sec.b).

2.Vertical velocity mean and standard deviation:

From Table 1 columns 4 and 7, the sample averages of the mean and of the standard deviation of the cloud-base vertical velocity $\left\langle\overline{w'_{\rm p}}\right\rangle_{\rm k}$ and $\left\langle\Gamma_{w'_{\rm p}}\right\rangle_{\rm k}$ are nearly equal. Therefore, we shall assume that the cloud-base vertical velocities within each plume display uniform and equal mean and standard deviation:

$$\overline{w'_{\mathrm{p}}} = \Gamma_{w'_{\mathrm{p}}} \tag{20}$$

This assumptation has already been used by Grandpeix and Lafore (2010) and Grandpeix et al. (2010), and will also be used in the stochastic parametrization for deep convection triggering presented in the second paper of this series.

d. *Maximum vertical velocities distribution* $\mathcal{P}_{max}(w'_{max,i})$

Since we have characterized the velocity spectrum ($\mathcal{P}(w'_{\mathrm{p,i}})$), the next step is to look for the maximum-value distribution of $w'_{\mathrm{p,i}}$, given a thermal i made of n_{i} drafts (i.e of cross-section s_{i}). According to Appendix A, from the vertical velocity PDF, it is possible to retrieve a distribution law for the maximum values $\mathcal{P}_{\mathrm{max}}(w'_{\mathrm{max,i}})$ (see Eq A2) and to compute an estimator $W'_{\mathrm{max,i}}$ (see Eq A11) of the maximum velocity at the cloud base. A representative value of this estimator can be the median value, corresponding to $\pi_t \approx 0.7$. Introducing this statement in Eq A11, taking into account the uniformity of $\overline{w'_{\mathrm{p}}}$ and $\Gamma_{w'_{\mathrm{p}}}$ and averaging over each sample k yields the estimated (median) maximum velocity:

$$\left\langle W'_{\mathrm{p,max}} \right\rangle_{\mathrm{k}} = \overline{w'_{\mathrm{p}}} + \Gamma_{w'_{\mathrm{p}}} \sqrt{\ln\left(\frac{\langle n\rangle_{\mathrm{k}}^2}{2\pi\pi_t^2}\right) - \ln\left(\ln\left(\frac{\langle n\rangle_{\mathrm{k}}^2}{2\pi\pi_t^2}\right)\right)} \tag{21}$$

Fig 9 a) displays the sensitivity of the maximum values PDF ($\mathcal{P}_{\mathrm{max}}(w'_{\mathrm{max,i}})$) of a plume i to its cross-section s_{i}. The PDF is relatively peaked and thin in all cases, becoming slightly narrower as the cloud base area increases. Consequently, the estimator $W'_{\mathrm{max,i}}$ is expected to give a good approximation of the simulated $w'_{\mathrm{max,i}}$ at any time (contrary to $\mathcal{S}_{\mathrm{max}}$). Fig 9 b) compares the pairs $\left[\left\langle w'_{\mathrm{p,max}}\right\rangle_{\mathrm{k}} : \langle n\rangle_{\mathrm{k}}\right]$ and $\left[\left\langle W'_{\mathrm{p,max}}\right\rangle_{\mathrm{k}} : \langle n\rangle_{\mathrm{k}}\right]$ for each sample k. The sample-mean estimator $\left\langle W'_{\mathrm{p,max}}\right\rangle_{\mathrm{k}}$ is in line with the sample-mean maximum $\left\langle w'_{\mathrm{p,max}}\right\rangle_{\mathrm{k}}$ extracted from the LES data. This result suggests that the hypothesis considering the mean ascending thermal plume as an ensemble of independent drafts, with a Gaussian velocity spectrum, seems relevant. Moreover, the maximum velocity encountered does not depend in anything else that the sampling effect: the more numerous the updrafts at the cloud base, the more the probability to get a strong one. Indeed, the analytical formulae given in Eq 21 does not take into account of entrainment/detrainment mixing or any other physical process.

In current thermal plume parametrizations using a mass-flux scheme, it is supposed that the entrained air from the subcloud layer is a rest gas. Hence, entrainment is a braking term in the parcel's equation of motion. Since entrainment only affects the peripheral zone of the thermal plume, one can expect that larger plumes are less sensitive to lateral entrainment (see Sec b). Thereby, larger plumes can host stronger updrafts in their core. But here it is shown that the sampling effect alone largely determines the maximum velocity $w'_{\mathrm{max,i}}$ encontered in a plume i of cross-section s_{i}. In other words, if considering a plume i, an increase of the cross-section s_{i} is accompagned by a corresponding increase of the number of random samplings for the vertical velocity $w'_{\mathrm{p,i}}$, finally leading to a statistical increase of the maximum velocity $w'_{\mathrm{max,i}}$. Then, according to this study, the fact to consider that larger clouds host more undiluted parcels in their core is not the best way to explain the velocity maximum, at least at the could base level.

This concordance between simulated and calculated maximums also shows that the tail of the Gaussian distribution of the velocity field ($\mathcal{P}(w'_{\mathrm{p,i}})$) in each plume i is pertinent. Since our concern is the deep convection triggering, we do focus on high velocities. And we try to verify if the independent gaussian draft is relevant, at least for the distribution tail. A way to do that is to plot the histogram of the complementary cumulative distribution function of $w'_{\mathrm{max,i}}$ (CCDF $\mathcal{F}(w'_{\mathrm{max,i}})$, i.e the probability to have a larger value than $w'_{\mathrm{max,i}}$ for each thermal plume i) given in Eq A1 (Appendix A). For that, for each thermal plume i (of type-2) we compute the CCDF $\mathcal{F}_{\mathrm{max}}(w'_{\mathrm{p,i}})$, then we divide them in bins of 0.1, and we plot the $\mathcal{F}_{\mathrm{max}}(w'_{\mathrm{max,i}})$ histogram displayed in Fig 10. The flat distribution shows that the simulated $w'_{\mathrm{max,i}}$ of each plume i is equally distributed on both sides of the PDF of the maximum velocities $\mathcal{P}_{\mathrm{max}}(w'_{\mathrm{max,i}})$. This proves that, at least for the tail of $\mathcal{P}(w'_{\mathrm{p,i}})$, (i) the hypothesis of the independent drafts is relevant, and (ii) the Gaussian PDF is pertinent too.

e. *Sum up*

To sum up, the dynamical properties of the type-2 plumes are uniform over the plume field. Since the cross-section spectrum for type-2 plumes is exponential, this result is somehow consistent with the exponential distribution for individual mass fluxes proposed by Plant and Craig (2008). Moreover, each thermal plume can be considered as composed of independent drafts (i.e with no spatial coherence), following a Gaussian distribution for the vertical velocity, in which the average is quasi equivalent to the standard deviation. Finally, the Gaussian distribution well describes the maximum values statistics, which mostly depend on the cloud base cross-section.

We shall add that a uniform vertical velocity spectrum over the plume field gives some relevance to the single plume approach, at least when considering the mean dynamical properties of the thermal plume ensemble.

6. Statistical Available Lifting Energy $\mathrm{ALE}_{\mathrm{BL,stat}}$

a. *$ALE_{BL,stat}$ computation*

The statistical $\mathrm{ALE}_{\mathrm{BL,stat}}$ corresponds to the maximum kinetic energy found over the plume spectrum. From Sec 4 and Sec 5 we extracted, respectively, a median value for the maximum cross-section $\mathcal{S}_{\max}$ (Eq 15), and a median value for the maximum vertical velocity $\left\langle W'_{\mathrm{p,max}} \right\rangle$ (Eq 21) of a plume sample. Knowing that $\left\langle W'_{\mathrm{p,max}} \right\rangle$ is an increasing function of cross-section, the strongest updraft is hosted by the largest thermal $\mathcal{S}_{\max}$ of the domain. Hence, when combining Eq 15 with Eq 21, and introducing Eq 20, we get a statistical maximum velocity inside the largest thermal:

$$\mathcal{W}'_{\max} = \overline{w'_{\mathrm{p}}} \left[1 + \sqrt{ \ln \left(\frac{ \left(\frac{S_2 \ln(N_2)}{\check{s}} \right)^2 }{2\pi} \right) - \ln \left(\ln \left(\frac{ \left(\frac{S_2 \ln(N_2)}{\check{s}} \right)^2 }{2\pi} \right) \right) } \right] \tag{22}$$

We could noticed that the arbitrary value š had a limited influence on $\mathcal{W}'_{\mathrm{max}}$. Hence, supposing that N_2, S_2 and $\overline{w'_{\mathrm{p}}}$ are known, we can finally compute the statistical maximum kinetic energy at cloud base:

$$\mathrm{ALE}_{\mathrm{BL,stat}} = \frac{1}{2}\mathcal{W}'^{2}_{\mathrm{max}} \tag{23}$$

Fig 11 shows the time evolution of $\mathrm{ALE}_{\mathrm{BL,stat}}$. It is maximum around 13:00 LT and decreases later on. Actually, $\mathcal{W}'_{\mathrm{max}}$ is approximately in phase with $\overline{w'_{\mathrm{p}}}$ (not shown), itself correlated to the sensible heat flux (not shown). Although the maximum cross-section $\mathcal{S}_{\mathrm{max}}$ is around two times larger at 16:00 LT than at 13:00 LT (see Fig 6 b) the surface heating is less, consequently, the mean velocity of the plume population $\overline{w'_{\mathrm{p}}}$ is around 30% less (not shown). This correlation between $\mathcal{W}'_{\mathrm{max}}$ and $\overline{w'_{\mathrm{p}}}$ can be easily understood by a growth-compared analysis applied to the 2 terms of the product in Eq 22: the first term is $\overline{w'_{\mathrm{p}}}$, and the second varies with $\sqrt{\ln(\mathcal{S}^2_{\mathrm{max}})}$ (or $\sqrt{2\ln(\mathcal{S}_{\mathrm{max}})}$). Thus, during the transition phase, the $\overline{w'_{\mathrm{p}}}$ decrease dominates the $\mathcal{S}_{\mathrm{max}}$ increase.

According to the LES, the morning time large-scale inhibition is very high, and $\mathrm{ALE}_{\mathrm{BL,stat}}$ reaches the CIN (not shown) around 13:00 LT. Therefore, since observational (Lothon et al. (2011)) as well as LES (Couvreux et al. (2012)) data shows that deep convection triggers near 16:00 LT, the dynamical threshold $\mathrm{ALE}_{\mathrm{BL,stat}} > |\mathrm{CIN}|$ alone is not sufficient to describe the whole transition process.

b. *Towards a new formulation of triggering*

Lothon et al. (2011) shows that, around 12:00-13:00 LT, the boundary layer moves from a regular, steady cloud-street organization to a more istropic structure consisting of bigger clouds. This period correspond to the beginning of the transition phase. Then, if $\mathrm{ALE}_{\mathrm{BL}} > |\mathrm{CIN}|$ is apparently not a pertinent threshold for the deep-convection triggering, it may be relevant for describing the threshold from a shallow cumulus regime, to an transition regime.

In the shallow cumulus regime, no clouds cross the inhibition layer. In the transition regime, many cumulus clouds have enough kinetic energy to overshoot the CIN, but are still too small for reaching the high troposphere. Then, we shall impose a complementary constraint on the size of the thermal plumes to permit the triggering of deep convection.

7. Deep convection triggering formulation

In the current LMDZ model version, the deep convection triggering by boundary layer thermals is exclusively based on the threshold condition $\mathrm{ALE_{BL}} > |\mathrm{CIN}|$. Since the associated thermal plume representation is deterministic, either not any plume triggers, or all the plumes trigger. But, since a thermal plume spectrum is considered here, we can *a priori* expect to represent, in a given domain, both passive boundary layer cumulus clouds, and overshooting clouds. As already mentioned, the plume size looks of primary importance in the triggering process; Lothon et al. (2011) noticed that first deep convective cells occur over a zone covered by the largest horizontal structures of the observed domain. Chaboureau et al. (2004) also stressed the existence of a two-step triggering, in which a transition phase clearly appears.

Hence, the triggering formulation main idea is that the thermal plume field must require (i) at least one thermal plume whose maximum kinetic energy exceeds the CIN, which means $\mathrm{ALE_{BL,stat}} > |\mathrm{CIN}|$, and (ii) a sufficient number of thermals whose size may potentially exceeds a certain threshold value S_{trig}. This threshold corresponds to an arbitrary limit, from which the cloud base do not anymore correspond to a cumulus, but to a congestus or a cumulonimbus cloud. One might expect that the largest thermal plume size grows gradually up to the time when it reaches this threshold.

Let S_{trig} be the threshold value for deep convection triggering and assume that t_0 corresponds to the instant when $\mathrm{ALE_{BL,stat}} > |CIN|$. The triggering probability P_τ for one plume scene of duration τ, composed of N_2 plumes, is the probability that $S_{\max} > S_{\mathrm{trig}}$; that is the CCDF $\mathcal{F}_{\max}(S_{\max})$ given in Eq A7:

$$P_\tau = \mathcal{F}_{\max}(S_{\text{trig}}) = 1 - (1 - \widehat{\mathcal{F}}(S_{\text{trig}}))^{N_2}$$

The no-trigger probability is then:

$$\widehat{P}_\tau = \widehat{\mathcal{F}}_{\max}(S_{\text{trig}}) = 1 - P_\tau$$

Giving, for every independent cloud scene of duration τ (e.g the average life expectancy of a thermal plume, ≈10 min):

$$\widehat{P}_\tau = \left(1 - \exp\left(\frac{-S_{\text{trig}}}{S_2}\right)\right)^{N_2} \tag{24}$$

The no-trigger probability definition $\widehat{P}_\tau$ can be generalized to every time period $\Delta t = n\tau$, composed of n independent scenes of duration τ:

$$\widehat{P}_{\Delta t} = \prod_{k=1}^{n} \widehat{P}_\tau$$

A continuous formulation (i.e whatever Δt) of the no-trigger probability $\widehat{P}_{\Delta t}$, which verifies $\widehat{P}_{\Delta t} = \widehat{P}_\tau$ if $\Delta t = \tau$ is:

$$\widehat{P}_{\Delta t} = (\widehat{P}_\tau)^{\frac{\Delta t}{\tau}}$$

When combining with Eq 24, this yields:

$$\widehat{P}_{\Delta t} = \left[\left(1 - \exp(\frac{-S_{\text{trig}}}{S_2})\right)^{N_2}\right]^{\dfrac{\Delta t}{\tau}} \tag{25}$$

Thus, during every time period Δt, we can compute a no-trigger probability $\widehat{P}_{\Delta t}$.

Looking back to Fig 6, the distribution of $S_{\max}$ is broad, meaning that $S_{\max}$ may vary a lot around the median value $\mathcal{S}_{\max}$ (Eq 15), and the median value $\mathcal{S}_{\max}$ does not represent the large fluctuations of $S_{\max}$. Therefore we have to consider the triggering process $S_{\max} > S_{\text{trig}}$ is stochastic. Considering a time period Δt, the probability that $S_{\max} > S_{\text{trig}}$ is equal to the probability that a random sample $0 < \mathcal{R} < 1$ exceeds the non triggering probability per unit time $\widehat{P}_{\Delta t}$. By the same token, in a time period Δt, the stochastic triggering happens if $\mathcal{R} > \widehat{P}_{\Delta t}$.

The triggering process not only governs the deep convection beggining, but also its end. Indeed, deep conection happens as long as it is triggered. Consequently, to be coherent this triggering must last a certain time for allowing deep convection to produce significant rain. For that we suggested to double the decorrelation time τ (from 10 min to 20 min) once deep convection has triggered; arguing that the typical timescale for a deep convective updraft is around two times more than for a thermal plume.

Sum up: the three steps of the transition process

1.Preliminary condition

The boundary layer must be cloudy to allow the deep convection triggering.

2.The dynamical threshold

This threshold governs the transition from a regime in which cumulus clouds cannot reach their level of free convection (LFC) (i.e stays under the inhibition layer (CIN)) to a transient regime where at least some cumulus overshoot the CIN but do not reach the high troposphere in a significant number. It is also a deterministic threshold, which uses a PDF approach. It takes place when the statistical maximum kinetic energy produced by the boundary layer thermals $\mathrm{ALE}_{\mathrm{BL,stat}}$ exceeds the CIN:

$$\mathrm{ALE}_{\mathrm{BL,stat}} > |\mathrm{CIN}| \tag{26}$$

3. The geometric threshold

Once the dynamical criterion is reached, the boundary layer enters a transition regime, in which some cumulus overshoot the inhibition but do not reach the high atmosphere. The geometric criterion is stochastic, and governs the abrupt transition from the transient regime to the deep convection regime. It considers the type-2 plumes population spectrum, and states that every cloud scene of duration Δt can potentially trigger at the condition that a random sample $\mathcal{R}$ exceeds the no-trigger probability $\widehat{P}_{\Delta t}$:

$$\mathcal{R} > \widehat{P}_{\Delta t} \tag{27}$$

If deep convection has already triggered, the procedure is the same, but with a τ two times more impotant.

Fig 12 illustrates the conceptual view of this formulation, from the first cloud to the deep convection triggering. From this new formulation, a stochastic triggering parametrization is proposed in a companion paper.

8. Discussion and conclusion

To consider the plume field like a statistical ensemble, with intra-thermal velocity fluctuations and inter-thermal cross-section fluctuations, made it possible to describe the transition process more in detail than a single plume approach. Data from the LES case AMMA gave us many insights on the geometrical and dynamical properties of the cloudy thermal plumes at the cloud base level during the transition from shallow to deep convection.

The thermal plume field is divided into two populations, each one following an exponential distribution law, and from which an sum of exponential distribution $\mathcal{P}(s)$ for the whole population can be deduced. During the transition time, the distribution slope decreases, thermal plumes are less numerous, have a higher mean cross-section, and feed higher and deeper clouds. A simple linear relationship between the cloud horizontal lengthscale (at LCL), the cloud depth and the the altitude of the cloud base has been proposed and verified on the AMMA case. This relationship suggests a link between the cloud thermodynamic properties and the cloud geometry. The cross-section maximum distribution $\mathcal{P}_{\max}(S_{\max})$ is consistent with the LES, but spreads over a large range of values. Indeed, the estimated median value $\mathcal{S}_{\max}$ gradually moves away from the simulated $S_{\max}$ while transition evolves.

Suspecting that the thermal size plays a key role in the triggering process, we filter out the type-1 plumes, and focus on the dynamical properties of type-2 plumes exclusively. Type-2 plumes can be described as a sum of independent drafts whose velocity distribution $\mathcal{P}(w'_{\mathrm{p,i}})$ is nearly a gaussian, and is constant over the plume field. The gaussian mean and standard deviation are similar. Since the maximum value distribution $\mathcal{P}_{\max}(w'_{\mathrm{max,i}})$ is also consistent with the simulated values, the gaussian distribution is pertinent for describing the maximum velocities distribution as well.

Combining an analytical formulae of the median maximum plume size $\mathcal{S}_{\max}$ and the median maximum velocity $\left\langle W'_{\mathrm{p,max}}\right\rangle$ over a plume sample, a statistical maximum velocity $\mathcal{W}'_{\max}$ inside the largest thermal has been computed to get a statistical estimate of $\mathrm{ALE}_{\mathrm{BL,stat}}$. In addition, the new triggering consider a threshold size S_{trig}, which has to be exceeded by the maximum $S_{\max}$ to trigger deep convection. Knowing that the maximum size distribution $\mathcal{P}_{\max}(S_{\max})$ is wide, $S_{\max}$ fluctuations are important. Then it is pertinent to consider triggering (i.e $S_{\max} > S_{\mathrm{trig}}$) as a stochastic process, in which a random sample $\mathcal{R}$ has to exceed a no-trigger probability $\widehat{P}_{\Delta t}$ for triggering.

The present formulation proposes a three-steps transition and consists in two consecutive thresholds; the first one is deterministic and the second one is stochastic. The first threshold is dynamic ; it governs the inhibition crossing by at least one plume of the domain (i.e $\mathrm{ALE}_{\mathrm{BL,stat}} > |\mathrm{CIN}|$). It represents the moment when shallow clouds start to overshoot the inhibition layer and reach their Level of Free Convection (LFC); that is the transition phase. The second one is geometric and rules the deep convection triggering. Since deep convection tends to trigger where the largest horizontal structures are, there is a threshold cross-section which has a certain probability to be exceeded at every independent cloud scene.

This new triggering formulation has the great advantage to allow the existence of a particular stage between shallow and deep convection, during which the inhibition layer is overcome but clouds are still too small for reaching the high troposphere. This transient regime is generally missed in most of GCMs.

However, to integrate such a formulation in a parametrization of deep convection triggering by boundary layer thermals is still a difficult work. The main difficulty is to retrieve a cross-section spectrum from the variables given by the boundary layer parametrization, which is single-plume based in most of the cases. A triggering parametrization for the LMD's model (LMDZ) based on this formulation is proposed in a companion paper.

One may contest that this triggering formulation is inspired from only one case study, and so has few chances to be applicable in other situations. That is why the robustness of this formulation will be futher investigated in the second part of this paper; the corresponding parametrization will be tested over various environmental conditions (continental and oceanic) and also in conditions favourable, and not favourable, for triggering. It will be first tested in a single-column framework on different case studies, and then in the global framework to estimate the added value when compare to the deterministic approach in the full GCM.

9. Figures and tables

a. Figures

b. Tables

Acknowledgments.

The research leading to these resulthas been both supported by the French Department of Teaching and Research, and by the European Union, Seventh Framework Program (FP7/2011-2015) under grant agreement n 282672. The author would like to thank Franoise Guichard and Jean-Philippe Lafore, from Meteo-France (Toulouse), for their very helpful comments, and for the numerous and enlightening discussions we had about deep convection issues.

APPENDIX A

Maximum of a large ($\simeq 100$) number of random variables with identical probability density functions.

We consider a set of N independent random variables $(x_i)_{i=1,N}$ with identical probability density function (PDF) $\mathcal{P}$, cumulative distribution function (CDF) $\widehat{\mathcal{F}}$, complementary cumulative distribution function (CCDF) $\mathcal{F}$. The CDF $\widehat{\mathcal{F}}$ (resp CCDF $\mathcal{F}$) is defined by: $\widehat{\mathcal{F}}(X) = \{\text{probability that } x_i < X \text{ (resp } x_i > X)\}$. The following relations hold:

$$\mathcal{F}(X) = 1 - \widehat{\mathcal{F}}(X) \quad ; \quad \mathcal{P}(x) = \frac{d\widehat{\mathcal{F}}}{dx} = -\frac{d\mathcal{F}}{dx}$$

(i) CCDF of the maximum:

We seek the CCDF $\mathcal{F}_{\max}$ of the maximum of the $(x_i)_{i=1,N}$. The probability that $\max(x_i)$ exceeds a given value X is equal to the probability that at least one of the x_i exceeds X, which is equal to $1 - \{\text{probability that, for all i, } x_i < X\}$. Since the $(x_i)_{i=1,N}$ are independent, the last probability reads $1 - (1 - \mathcal{F}(X))^N$. Thus the CCDF $\mathcal{F}_{\max}$ of the maximum of the $(x_i)_{i=1,N}$ reads:

$$\mathcal{F}_{\max}(X) = 1 - (1 - \mathcal{F}(X))^N \tag{A1}$$

Which gives for the PDF of the maximum values:

$$\mathcal{P}_{\max}(X_{\max}) = \frac{-d\mathcal{F}_{\max}(X_{\max})}{dX_{\max}} \tag{A2}$$

(ii) Inverse formula

Given a probability $P_t < 0.9$ we seek the corresponding threshold value X_t such that the probability that $\max(x_i)_{i=1,N} > X_t$ is equal to P_t:

$$\mathcal{F}_{\max}(X_t) \;=\; P_t \tag{A3}$$

Note that we are interested in large values of the x_i, which implies that some upper bound be imposed upon P_t. As will appear later, an upper bound of 0.9 is sufficient for the oncoming developments.

Substituting the expression of $\mathcal{F}_{\max}$ (Eq. A1) in Eq. (A3) and solving for $\mathcal{F}(X_t)$ one gets:

$$\mathcal{F}(X_t) \;=\; 1 \;-\; (1 \;-\; P_t)^{\frac{1}{N}} \tag{A4}$$

This is an exact formula. We shall rather use an approximate form taking into account the fact that N is large. To that end we rewrite Eq.(A4):

$$\mathcal{F}(X_t) \;=\; 1 \;-\; \exp\left(\frac{1}{N}\ln(1 \;-\; P_t)\right)$$

Introducing the new variable

$$\pi_t \;=\; -\ln(1 \;-\; P_t)$$

which verifies $0. < \pi_t < 2.3$, the equation reads:

$$\mathcal{F}(X_t) \;=\; 1 \;-\; \exp\left(-\frac{\pi_t}{N}\right)$$

Since π_t/N is in the order of 10^{-2}, the exponential may be replaced with a first order expansion :

$$\mathcal{F}(X_t) \;=\; \frac{\pi_t}{N} \tag{A5}$$

Thanks to this equation, finding X_t amounts merely to inverting $\mathcal{F}$. In particular, the median X_{med} of the distribution of the maximum, which corresponds to $P_t = 0.5$ and $\pi_t \approx 0.7$, is given by:

$$\mathcal{F}(X_{\text{med}}) \;=\; \frac{\ln(2)}{N}$$

(iii) The thermal cross-section s case

The CCDF is

$$\mathcal{F}(S) \;=\; \exp\left(-\frac{S}{\langle s\rangle}\right) \tag{A6}$$

and the number of random variables is the number N_{tot} of thermals in the grid-cell. The threshold cross-section $\mathcal{S}_t$ is given by Eq A5 where expression A6 is substituted for $\mathcal{F}(X_t)$, that is:

$$\exp\left(-\frac{\mathcal{S}_t}{\langle s\rangle}\right) = \frac{\pi_t}{N_{\text{tot}}} \tag{A7}$$

thus:

$$\mathcal{S}_t = \langle s\rangle \ln(\frac{N_{\text{tot}}}{\pi_t}) \tag{A8}$$

(iv) The vertical velocity $w'_{p,i}$ case

The CCDF is

$$\mathcal{F}(W'_{\text{p,i}}) = \frac{1}{2}\text{Erfc}\left(\frac{W'_{\text{p,i}} - \overline{w'_{\text{p,i}}}}{\sqrt{2}\Gamma_{w'_{\text{p,i}}}}\right) \tag{A9}$$

and the number of random variables is the number n_{i} of elementary drafts in the thermal indexed i. The threshold vertical velocity $W'_{\text{t,i}}$ verifies Eq A5 where expression A9 is substituted for $\mathcal{F}(X_t)$, that is:

$$\frac{1}{2}\text{Erfc}\left(\frac{W'_{\text{t,i}} - \overline{w'_{\text{p,i}}}}{\sqrt{2}\Gamma_{w'_{\text{p,i}}}}\right) = \frac{\pi_t}{n_{\text{i}}} \tag{A10}$$

Since $\dfrac{\pi_t}{n_{\text{i}}} << 1$, one may use the asymptotic form of $\text{Erfc}^{-1}(x)$ in the limit $x \longrightarrow 0$, $\text{Erfc}^{-1}(x) \;\approx\; \dfrac{1}{\sqrt{2}}\sqrt{\ln\left(\dfrac{2}{\pi x^2}\right) - \ln\left[\ln\left(\dfrac{2}{\pi x^2}\right)\right]}$, which yields:

$$W'_{\text{t,i}} = \overline{w'_{\text{p,i}}} + \Gamma_{w'_{\text{p,i}}}\sqrt{\ln\left(\frac{n_{\text{i}}^2}{2\pi\pi_t^2}\right) - \ln\left(\ln\left(\frac{n_{\text{i}}^2}{2\pi\pi_t^2}\right)\right)} \tag{A11}$$

REFERENCES

Arakawa, A. and W. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, part i. *J. Atmos. Sci*, **31 (3)**, 674–701.

Bechtold, P., J. Chaboureau, A. Beljaars, A. Betts, M. Köhler, M. Miller, and J. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3119–3137.

Boing, S., H. Jonker, A. Siebesma, and W. Grabowski, 2010: Influence of subcloud-layer structures on the transition to deep convection. *19th Symposium on Boundary Layers and Turbulence*.

Chaboureau, J., F. Guichard, J. Redelsperger, and J. Lafore, 2004: The role of stability and moisture in the diurnal cycle of convection over land. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3105–3117.

Clark, T., T. Hauf, and J. Kuettner, 1986: Convectively forced internal gravity waves: Results from two-dimensional numerical experiments. *Quarterly Journal of the Royal Meteorological Society*, **112 (474)**, 899–925.

Couvreux, F., C. Rio, F. Guichard, M. Lothon, G. Canut, D. Bouniol, and A. Gounou, 2012: Initiation of daytime local convection in a semi-arid region analysed with high-resolution simulations and amma observations. *Quarterly Journal of the Royal Meteorological Society*, **138 (4)**, 56–71.

Craig, G. and B. Cohen, 2006: Fluctuations in an equilibrium convective ensemble. part i: Theoretical formulation. *Journal of the atmospheric sciences*, **63 (8)**, 1996–2004.

Del Genio, A. and J. Wu, 2010: The role of entrainment in the diurnal cycle of continental convection. *Journal of Climate*, **23 (10)**, 2722–2738.

Emanuel, J. D. N., K. A. and C. S. Bretherton, 1994: On large scale circulations in convecting atmospheres. *Quart. J. Roy. Meteor. Soc.*, **120**, 1111–1143.

Emanuel, K., 1991: A scheme for representing cumulus convection in large-scale models. *Journal of the atmospheric sciences*, **48 (21)**, 2313–2335.

Grandpeix, J. and J. Lafore, 2010: A density current parameterization coupled with emanuel's convection scheme. part i: The models. *Journal of the Atmospheric Sciences*, **67 (4)**, 881–897.

Grandpeix, J., J. Lafore, and F. Cheruy, 2010: A density current parameterization coupled with emanuel's convection scheme. part ii: 1d simulations. *Journal of the Atmospheric Sciences*, **67 (4)**, 898–922.

Guichard, F., et al., 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3139–3172.

Hourdin, F., et al., 2012: Lmdz5b: the atmospheric component of the ipsl climate model with revisited parameterizations for clouds and convection. *Climate Dynamics, in Press*, 1–30.

Jones, T. and D. Randall, 2011: Quantifying the limits of convective parameterizations. *Journal of Geophysical Research*, **116 (D8)**, D08 210.

Kain, J. and J. Fritsch, 1992: The role of the convective ?trigger function? in numerical forecasts of mesoscale convective systems. *Meteorology and Atmospheric Physics*, **49 (1)**, 93–106.

Khairoutdinov, M. and D. Randall, 2006: High-resolution simulation of shallow-to-deep convection transition over land. *Journal of the atmospheric sciences*, **63 (12)**, 3421–3436.

Kuo, H., 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *Journal of Atmospheric Sciences*, **31**, 1232–1240.

Lafore, J., et al., 1998: The meso-nh atmospheric simulation system. part i: Adiabatic formulation and control simulations. *Annales Geophysicae*, Springer, Vol. 16, 90–109.

Lothon, M., B. Campistron, M. Chong, F. Couvreux, F. Guichard, C. Rio, and E. Williams, 2011: Life cycle of a mesoscale circular gust front observed by a c-band doppler radar in west africa. *Monthly Weather Review*, **139**, 1370–1388.

Malkus, J., 1954: Some results of a trade-cumulus cloud investigation. *Journal of Atmospheric Sciences*, **11**, 220–237.

Mapes, B., 2000: Convective inhibition, subgrid-scale triggering energy, and stratiform instability in a toy tropical wave model. *Journal of the atmospheric sciences*, **57 (10)**, 1515–1535.

Matheou, G., D. Chung, L. Nuijens, B. Stevens, and J. Teixeira, 2011: On the fidelity of large-eddy simulation of shallow precipitating cumulus convection. *Monthly weather review*, **139 (9)**, 2918–2939.

Medeiros, B., A. Hall, and B. Stevens, 2005: What controls the mean depth of the pbl? *Journal of climate*, **18 (16)**, 3157–3172.

Neelin, J., O. Peters, J. Lin, K. Hales, and C. Holloway, 2008: Rethinking convective quasi-equilibrium: observational constraints for stochastic convective schemes in climate models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **366 (1875)**, 2579–2602.

Neggers, R., H. Jonker, and A. Siebesma, 2003: Size statistics of cumulus cloud populations in large-eddy simulations. *Journal of the atmospheric sciences*, **60 (8)**, 1060–1074.

Palmer, T. N., 2012: Towards the probabilistic earth-system simulator: a vision for the future of climate and weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **138 (665)**, 841–861.

Plant, R. and G. Craig, 2008: A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences*, **65 (1)**, 87–105.

Raymond, D. and M. Herman, 2011: Convective quasi-equilibrium reconsidered. *Journal of Advances in Modeling Earth Systems*, **3 (8)**, M08 003.

Rio, C., F. Hourdin, J. Grandpeix, and J. Lafore, 2009: Shifting the diurnal cycle of parameterized deep convection over land. *Geophys. Res. Lett*, **36**, L07 809.

Rio, C., et al., 2012: Control of deep convection by sub-cloud lifting processes: the alp closure in the lmdz5b general circulation model. *Climate Dynamics*, 1–22.

Rodts, S., P. Duynkerke, and H. Jonker, 2003: Size distributions and dynamical properties of shallow cumulus clouds from aircraft observations and satellite data. *Journal of the atmospheric sciences*, **60 (16)**, 1895–1912.

Simpson, J. and V. Wiggert, 1969: Models of precipitating cumulus towers. *Monthly Weather Review*, **97 (7)**, 471–489.

Stechmann, S. N. and J. D. Neelin, 2011: A stochastic model for the transition to strong convection. *J. atmos. Sci*, **68 (665)**, 2955–2970.

Taylor, C., A. Gounou, F. Guichard, P. Harris, R. Ellis, F. Couvreux, and M. De Kauwe, 2011: Frequency of sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, **4 (7)**, 430–433.

Warner, J., 1970: The microstructure of cumulus cloud. part iii. the nature of the updraft. *Journal of Atmospheric Sciences*, **27**, 682–688.

Wu, C., B. Stevens, and A. Arakawa, 2009: What controls the transition from shallow to deep convection? *Journal of the Atmospheric Sciences*, **66 (6)**, 1793–1806.

Xu, K., A. Arakawa, and S. Krueger, 1992: The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *Journal of the atmospheric sciences*, **49 (24)**, 2402–2420.

Yang, G. and J. Slingo, 2001: The diurnal cycle in the tropics. *Monthly Weather Review*, **129 (4)**, 784–801.

List of Tables

Sample k	n range (drafts)	$\langle n \rangle_{\mathrm{k}}$ (drafts)	$\left\langle \overline{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}} (m.s^{-1})$	$\left\langle \overline{w'^2_{\mathrm{p}}} \right\rangle_{\mathrm{k}} (m^2.s{-}1)$	$\left\langle \overline{w'^3_{\mathrm{p}}} \right\rangle_{\mathrm{k}} (m^3.s{-}1)$	$\left\langle \Gamma_{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}} (m.s{-}1)$	$\left\langle \Phi_{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}}$	$\left\langle w'_{\mathrm{p,max}} \right\rangle_{\mathrm{k}} (m.s{-}1)$
1	40:43	41.0	1.04±0.06	2.67±0.21	7.74±0.93	1.08±0.04	0.26±0.05	3.35±0.15
2	43:47	44.7	0.91±0.05	2.23±0.18	5.98±0.78	1.01±0.04	0.23±0.05	3.14±0.13
3	47:51	49.0	1.11±0.06	2.89±0.21	8.53±0.96	1.11±0.04	0.14±0.06	3.66±0.14
4	51:57	53.8	1.04±0.05	2.63±0.17	7.42±0.75	1.12±0.04	0.22±0.04	3.79±0.15
5	57:63	60.1	1.07±0.05	2.60±0.18	7.02±0.75	1.06±0.04	0.13±0.04	3.59±0.14
6	63:71	66.8	1.07±0.04	2.44±0.15	6.25±0.58	1.04±0.03	0.19±0.04	3.53±0.12
7	71:84	77.2	1.07±0.05	2.59±0.16	6.65±0.63	1.08±0.03	0.16±0.04	3.81±0.11
8	85:103	99.5	1.11±0.04	2.80±0.15	7.85±0.70	1.16±0.03	0.24±0.05	4.08±0.13
9	103:137	116.9	1.02±0.04	2.57±0.13	6.55±0.57	1.14±0.03	0.21±0.04	3.97±0.12
10	138:430	215.9	1.08±0.04	2.72±0.12	6.64±0.48	1.17±0.03	0.06±0.05	4.23±0.10

TABLE 1. Mean dynamical characteristics of the 10 thermal plume samples of category 2

List of Figures

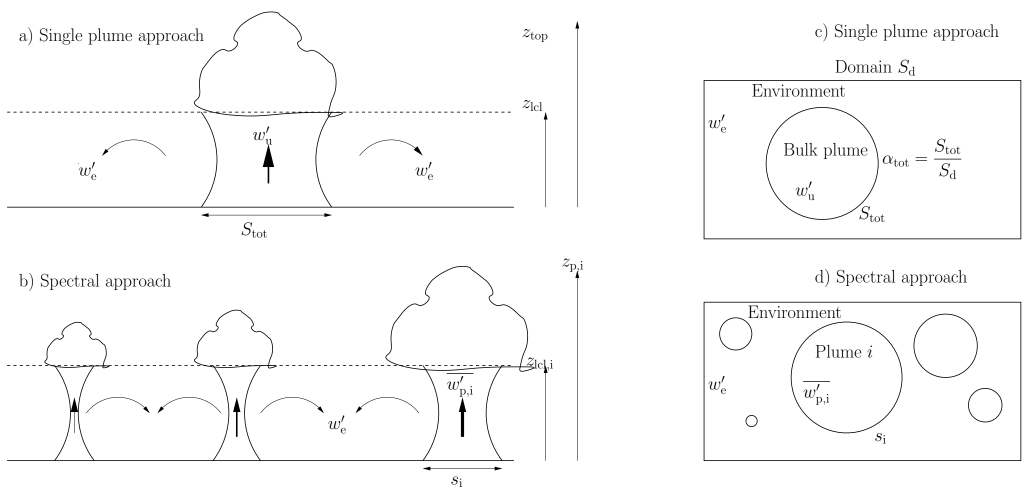

FIG. 1. Side view (left panel) and top view (right panel) of the single plume (a and c) and statistical (b and d) boundary layer coherent structure representation

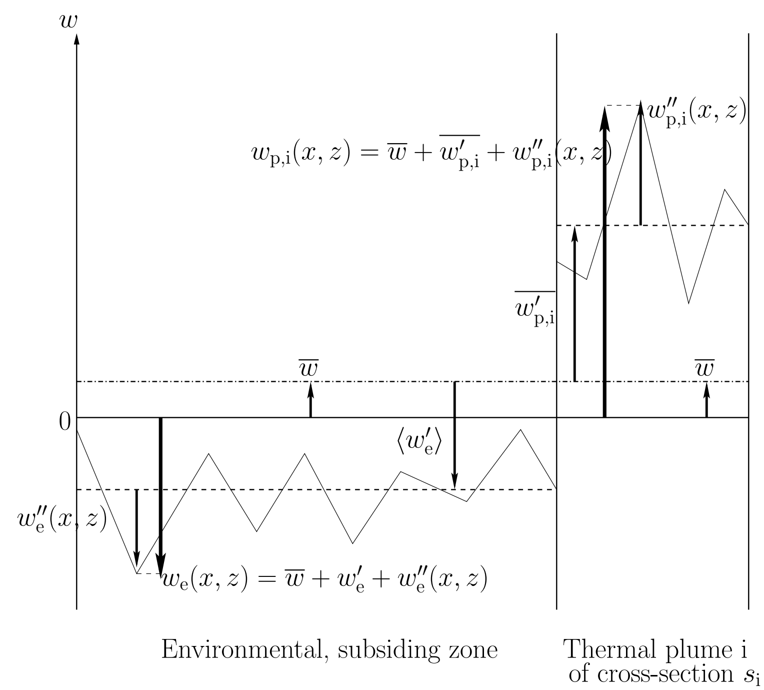

FIG. 2. Vertical cross-section of the thermal plume and its environment

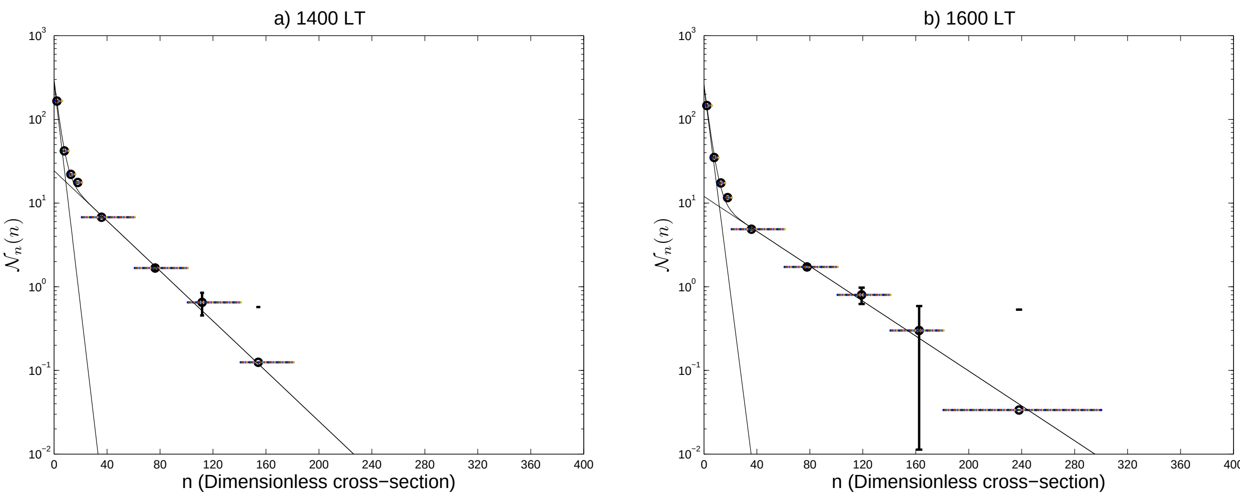

FIG. 3. N-normalized dimensionless cross-section distribution ($\mathcal{N}_n(n)$, see Eq 9) of the thermal plumes at LCL at a) 1400 LT and b) 1600 LT. Horizontal lines display the dimensionless cross-section bins. Vertical lines are errorbars

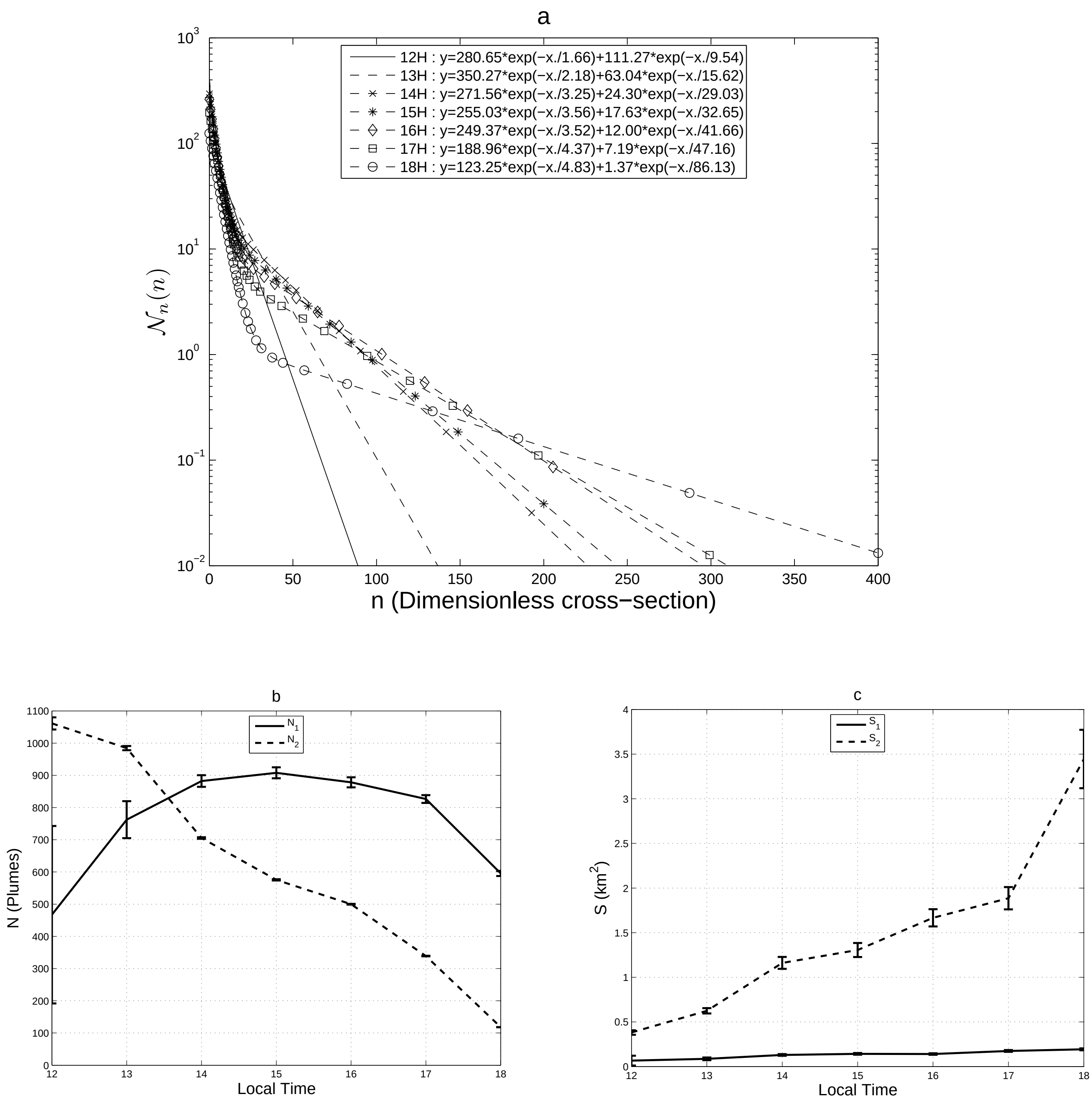

FIG. 4. a) Time evolution of the N-normalized cross-section distribution ($\mathcal{N}_n(n)$) fitting function at LCL. b) N_1 and N_2 time series c) S_1 and S_2 time series

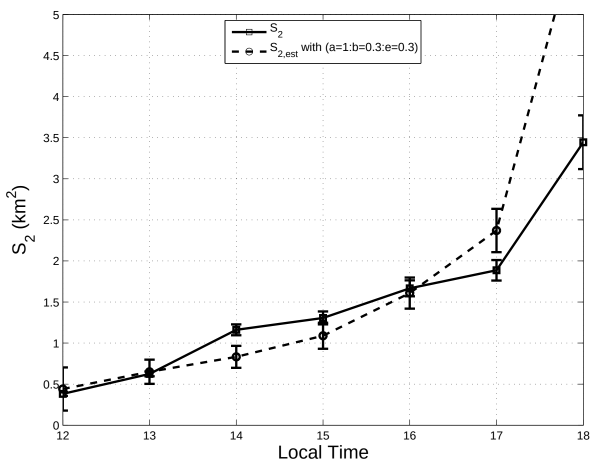

FIG. 5. Average cross-section of clouds 2 at the cloud base S_2 (m^2) for the LES (solid) and $\mathcal{S}_2$ calculus following Eq 13 with parameters $\{a = 1 : b = 0.3\}$ (dashed)

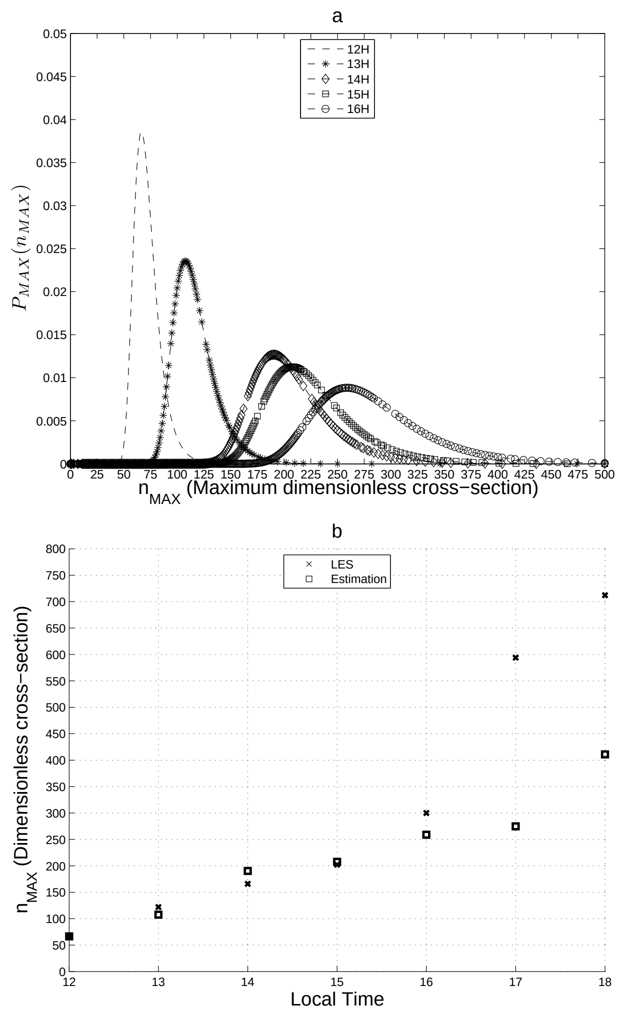

FIG. 6. a) $\mathcal{P}_{\max}(S_{\max})$ time evolution from 1200 to 1600 LT. b) Time-series of the estimated maximum cross-section $\mathcal{S}_{\max}$ (squares) and simulated $S_{\max}$ (crosses) from 1200 to 1800 LT.

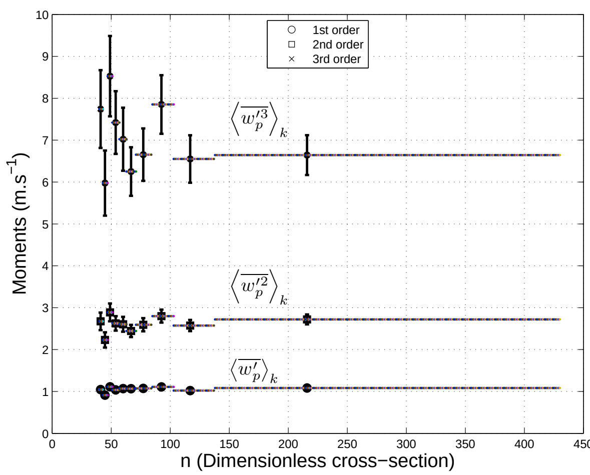

FIG. 7. Scatterplots of sample mean $\left\langle \overline{w'_{\mathrm{p}}} \right\rangle_{\mathrm{k}}$, second order $\left\langle \overline{w'^2_{\mathrm{p}}} \right\rangle_{\mathrm{k}}$ and third order non-centered moments $\left\langle \overline{w'^3_{\mathrm{p}}} \right\rangle_{\mathrm{k}}$ of cloud base vertical velocity, as a function of the dimensionless cross-section (n). Horizontal lines display the n bins and vertical lines display errorbars

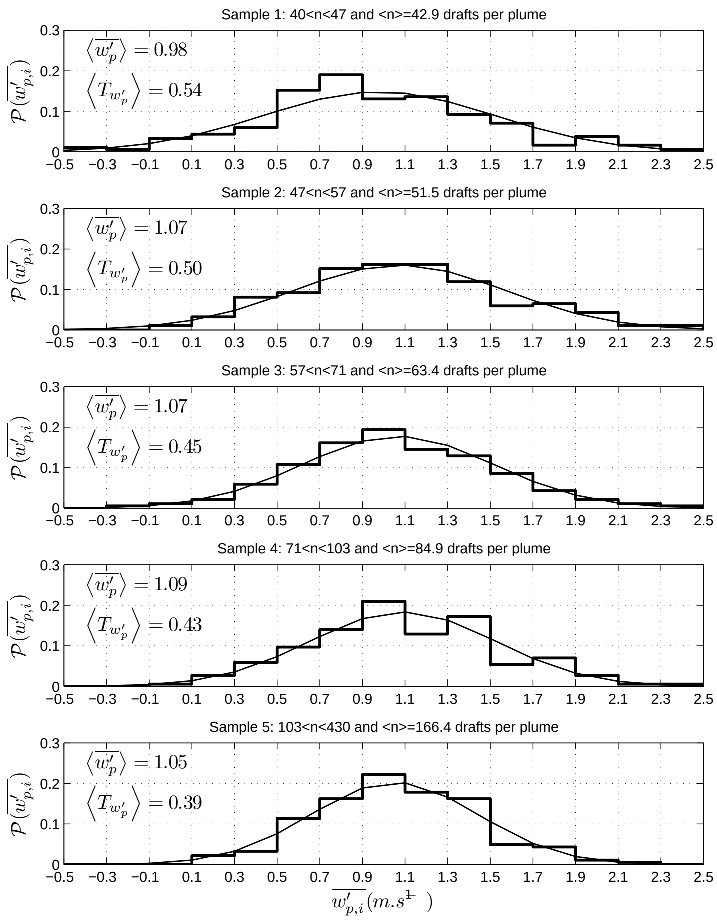

FIG. 8. Normalized histogram of $\overline{w'_{\mathrm{p,i}}}$ and fitting PDF $\mathcal{P}(\overline{w'_{\mathrm{p,i}}})$ for each sample. The sample mean cross-section averaged vertical velocity $\langle \overline{w'_{\mathrm{p}}} \rangle$ and standard deviation $\langle \Gamma_{w'_{\mathrm{p}}} \rangle$ are displayed on the upper left corner

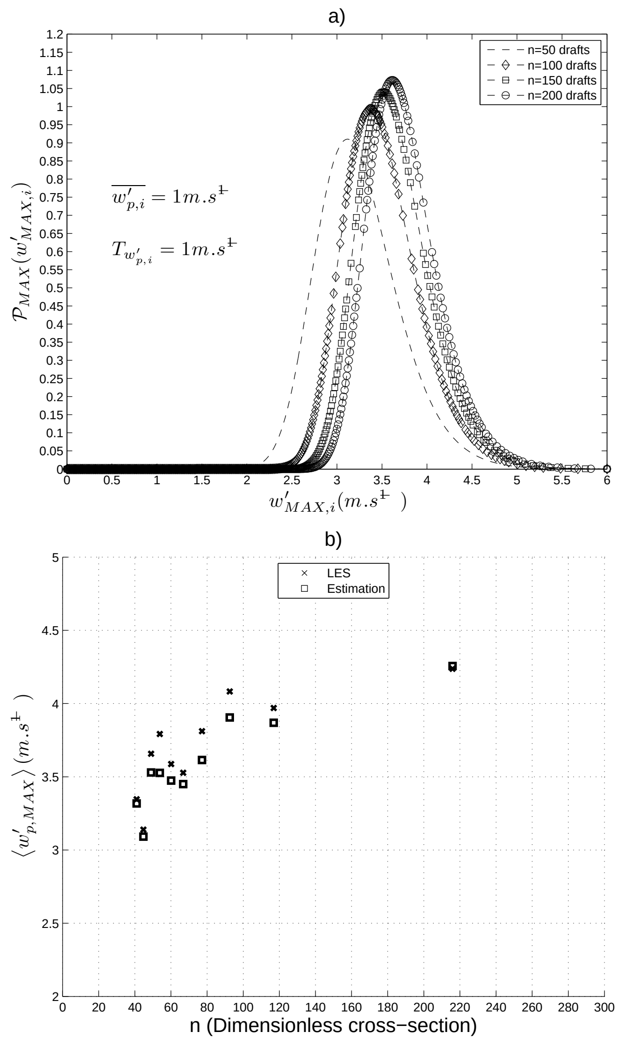

FIG. 9. a) $\mathcal{P}_{\max}(w'_{\mathrm{max,i}})$ sensitivity to the thermal i cross-section s_{i} with $\overline{w'_{\mathrm{p,i}}} = 1$ m.s^{-1} and $\Gamma_{w'_{\mathrm{p,i}}} = 1$ m.s^{-1}. b) Scatterplot of the estimated maximum velocity $\left\langle W'_{\mathrm{p,max}} \right\rangle_{\mathrm{k}}$ (squares) at the cloud base and simulated $\left\langle w'_{\mathrm{p,max}} \right\rangle_{\mathrm{k}}$ (crosses) as a function of cross-section

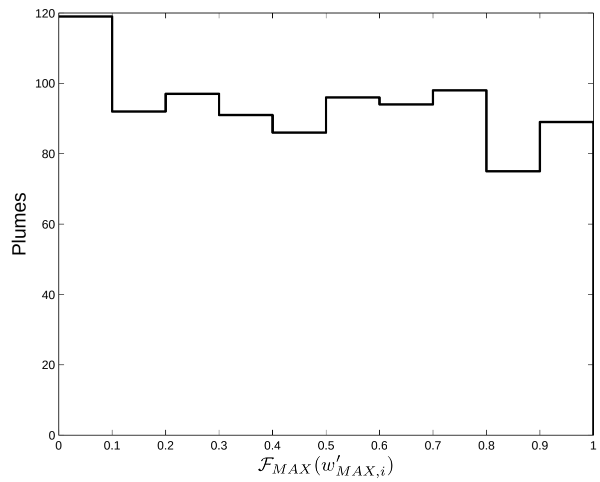

FIG. 10. Histogram of the CCDF of $w'_{\mathrm{max,i}}$ ($\mathcal{F}_{\mathrm{max}}(w'_{\mathrm{max,i}})$) for the cloud base of type-2. The horizontal axis represents the CCDF $\mathcal{F}_{\mathrm{max}}(w'_{\mathrm{max,i}})$ for each plume i and the vertical axis displays the number of plumes in each bin (0.1)

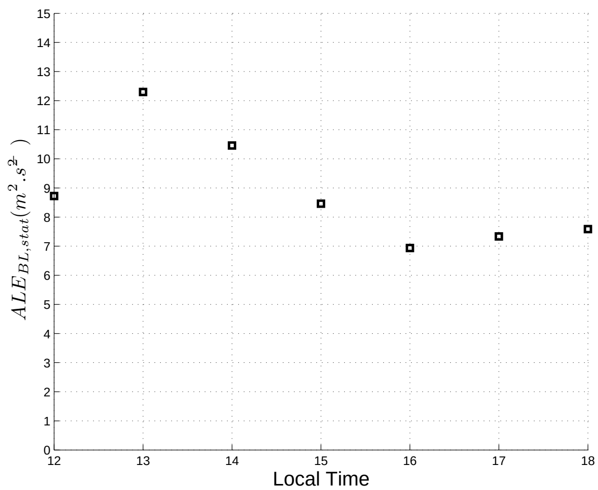

FIG. 11. Time series of $\mathrm{ALE_{BL,stat}}$

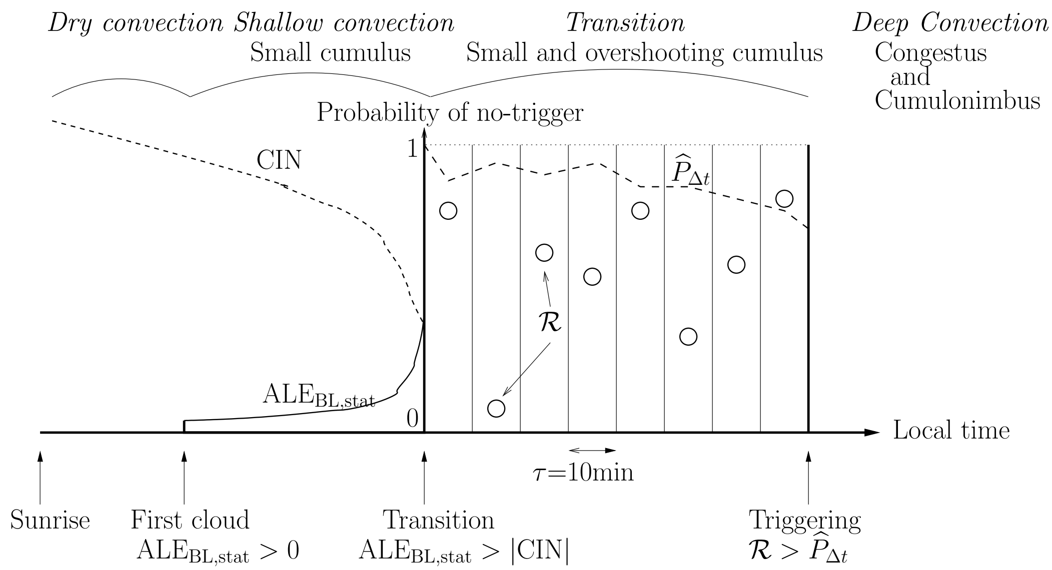

Fig. 12. Sketch of the transition from shallow to deep convection