



Diagnosing Coherent Structures in the Convective Boundary Layer by Optimizing Their Vertical Turbulent Scalar Transfer

Georgios A. Efstathiou¹ · John Thuburn¹ · Robert J. Beare¹

Received: 6 February 2019 / Accepted: 7 September 2019
© The Author(s) 2019

Abstract

A new method is introduced to identify coherent structures in the convective boundary layer, based on optimizing the vertical scalar flux in a two-fluid representation of turbulent motions as simulated by a large-eddy simulation. The new approach partitions the joint frequency distribution (JFD) of the vertical velocity and a transported scalar into coherent structures (fluid 2) and their environment (fluid 1) by maximizing that part of the scalar flux resolved by the mean properties in fluid 2 and fluid 1. The proposed method does not rely on any a priori criteria for the partitioning of the flow nor any pre-assumptions about the shape of the JFD. Different flavours of the optimization approach are examined based on maximizing either the total (fluid 1 + fluid 2) or the fluid-2 resolved scalar flux, and on whether all possible partitions or only a subset are considered. These options can result in different derived area fractions for the coherent structures. The properties of coherent structures diagnosed by the optimization method are compared to the conditional sampling of a surface-emitted decaying tracer, in which coherent structures are defined as having tracer perturbation greater than some height-dependent threshold. Results show that the optimization method is able to smoothly define coherent thermal structures in both the horizontal and the vertical. Moreover, optimizing the turbulent transfer by the fluid-2 resolved flux produces very similar coherent structures to the tracer threshold method, especially in terms of their area fraction and updraft velocities. Nonetheless, further analysis of the partitioning of the JFD reveals that, even though the area fraction of coherent structures might be similar, their definition can occupy different quadrants of the JFD, implying the contribution of different physical mechanisms to the turbulent transfer in the boundary layer. Finally, the kinematic and thermodynamic characteristics of the coherent structures are examined based on their definition criteria.

Keywords Conditional sampling · Thermals · Tracer threshold · Two-fluid approach · Updraft fraction

✉ Georgios A. Efstathiou
g.efstathiou@exeter.ac.uk

¹ Department of Mathematics, University of Exeter, Harrison Building, North Park Road, Exeter, UK

1 Introduction

In the convective boundary layer (CBL), the predominant source of turbulence is buoyancy resulting from the transfer of heat from the surface to the adjacent fluid. Buoyant plumes rise from the surface layer to form large coherent structures (thermals) that are responsible for transferring heat to the whole depth of the CBL by cooling the surface layer and warming the mixed layer. Thermals produce strong mixing in the boundary layer resulting in an almost uniform, well-mixed profile of potential temperature and mixing ratio of advected scalars such as moisture or trace species. The strongest of the thermals are able to maintain their ascent in the weakly-statically-stable upper part of the CBL, reaching the inversion and penetrating into the strongly stable layer above the boundary layer. The overshooting (negatively buoyant) boundary layer thermals and the entrainment of potentially warmer air from the stable layer contribute to a pronounced heat-flux minimum in the vertical sensible-heat-flux profile. Through mass conservation, these coherent structures force a broader, compensating descent in the CBL.

Because of this key role played by coherent structures in the dynamics and transport within the CBL, there is a need to be able to define them objectively and to identify them in large-eddy simulations (LES), both to improve our understanding of their behaviour and to parametrize their effects in numerical weather prediction (NWP) and climate models.

Small-scale turbulent eddies transfer heat gradually through adjacent vertical levels, so this heat flux can be modelled reasonably well as a diffusive flux down the local potential temperature gradient. In contrast, the depth scale of coherent structures is greater than the depth scale over which the mean potential temperature varies; thus the structure of thermals, and the heat flux carried by them at any height z , depend in a non-local way on conditions throughout the entire boundary layer. Because of these properties, it is common to decompose the total vertical turbulent heat flux ($\overline{w'\theta'}$) in the CBL as

$$\overline{w'\theta'} = -K_H \frac{\partial \bar{\theta}}{\partial z} + \overline{w'\theta'^{NL}}, \quad (1)$$

where the overbar denotes a horizontal average over the region of interest (e.g. one grid cell of a climate model), and the primes denote deviations from the mean. In addition, K_H is the thermal turbulent eddy diffusivity coefficient, θ is the potential temperature, and $\overline{w'\theta'^{NL}}$ is the non-local heat flux assumed to be associated with coherent structures. The downgradient term (first term on the r.h.s) expresses the local heat flux as mentioned above.

Such a partitioning of the heat flux into local and non-local components is widely used when parametrizing the CBL in NWP and climate models. Troen and Mahrt (1986) used a countergradient term (γ) to model the non-local heat flux (Deardorff 1972) in the form $\overline{w'\theta'^{NL}} = \gamma K_H$, an approach that has been used for some time in global and mesoscale modelling (Hong and Pan 1996; Lock et al. 2000; Hong et al. 2006). Chatfield and Brost (1987) were one of the first to suggest a decomposition of Eq. 1 based on a mass-flux concept. In addition, Hourdin et al. (2002) proposed a mass-flux representation of boundary-layer thermals, while Siebesma and Teixeira (2000), Soares et al. (2004), and Siebesma et al. (2007), introduced the mixed eddy-diffusivity mass-flux (EDMF) approach for modelling the CBL and shallow cumulus convection. In this approach the non-local heat flux due to the ascending thermals is represented by a mass-flux formulation $\overline{w'\theta'^{NL}} = M(\theta_u - \bar{\theta})$ where M is the mass flux expressed as $M = \sigma(w_u - \bar{w})$ with σ the updraft area fraction, w_u the updraft vertical velocity, and θ_u the updraft potential temperature (see also Rio and Hourdin 2008; Pergaud et al. 2009; Angevine et al. 2010; Sušelj et al. 2013). Some versions of the EDMF

approach explicitly or implicitly specify the vertical profile of σ in the boundary layer. Others (e.g., Angevine 2005) calculate the profile of σ via a vertical integral of the updraft mass budget. More recently, EDMF-like schemes have been introduced in which updraft properties including w , θ , and the area fraction are prognostic variables (Tan et al. 2018; Thuburn et al. 2019). For all of these schemes, information on the profile of σ and other updraft properties from high-resolution reference simulations is essential either as direct input to the scheme or for calibration and validation.

Although non-local transport and transport by coherent structures are conceptually closely related, and are often assumed to be identical, an important caveat is that they are not identical when defined mathematically—see Sect. 2.3. Note, also, that the terms ‘coherent structure’, ‘thermal’, and ‘updraft’ are often used interchangeably, though in what follows we will not assume that they comprise exclusively ascending fluid.

A further reason for interest in identifying coherent structures in the boundary layer is that the definition and diagnosis of entrainment into and detrainment out of coherent structures depends critically on how those coherent structures themselves are defined (e.g., Romps 2010; Yeo and Romps 2013; Dawe and Austin 2013). Quantifying entrainment and detrainment is important both for modelling the boundary layer using EDMF schemes such as those discussed above, and because entraining parcel models for shallow and deep cumulus convection are typically initiated in the boundary layer.

It becomes obvious that there is a broad spectrum of motions in the CBL originating from the surface layer due to strong surface heating. Even though coherent structures and small-scale processes might appear conceptually straightforward to separate, the continuous flow of energy from the turbulence production to the dissipation scales exhibits the strong interconnection between eddies of different size in the CBL. Siebesma et al. (2007) identified the updraft area as the grid points with vertical velocity larger than the p -percentile of the w distribution as derived from LES at each height. They used p values of the order of 1–5% to define updrafts and their properties. This method has the inherent limitation of setting a rather arbitrary cut-off point in the continuous spectrum of vertical velocities while considering the updraft fraction to be constant with height. Couvreux et al. (2010) developed a new method to diagnose the coherent structures based on the release of a “radioactive” passive scalar from the surface. The definition they used was based on the comparison of the turbulent fluctuations of the scalar mixing ratio with the standard deviation of the scalar mixing ratio similar to Berg and Stull (2004), providing a more physical representation of coherent updrafts.

Mahrt and Paumier (1984) analyzed joint frequency distributions (JFD) of w and θ from aircraft measurements to identify the relative contribution of thermals to turbulent transport. Wyngaard and Moeng (1992) followed a statistical approach to address the closure problem in a mass-flux parametrization, fitting a Gaussian distribution to the joint probability density of w and θ . In a similar way, Berg and Stull (2004) examined the JFD of θ and vapour mixing ratio. Chinita et al. (2018) utilised the JFD method to partition the flow in the $\theta - w$ space by fitting a joint Gaussian to the JFD to represent local mixing while the remaining part represents coherent updrafts. Even though this approach is physically-based and has the advantage of not requiring the specification of any sampling criteria, it still requires an assumption about the functional form of the JFDs.

The motivation for the present study is twofold: first is to introduce a new method for separating the flow into coherent structures and small-scale turbulence by adapting a two-fluid representation of the CBL. Optimizing the two-fluid formulation of the vertical scalar flux leads to the partition of the JFD into coherent motions and local turbulence without the need to introduce any arbitrary thresholds or other separation criteria. The new approach is compared with different flavours of the Couvreux et al. (2010) tracer threshold method. The

second aim is to provide a better understanding of the physical mechanisms represented by the partitioning of the turbulent transfer in the CBL into coherent structures and their environment. Using quadrant analysis of the joint θ and w distribution, the physical processes behind the assumptions from the different definitions of coherent motions are revealed, showing that, even though different methods can diagnose similar area fractions, the mechanisms of heat transfer might be different. Also, the kinematic and thermodynamic properties of coherent motions from the different approaches are examined based on the release and tracking of Lagrangian particles in the LES (e.g., Heus et al. 2008) of a quasi-steady CBL.

2 Methodology

2.1 Simulations

The Met Office Large Eddy Model (LEM) version 2.4 is used to simulate a dry CBL driven by a strong sensible heat flux of 240 W m^{-2} and small geostrophic wind speed of 1 m s^{-1} . The set-up follows Sullivan and Patton (2011) with a strong capping inversion above the boundary layer ($z_* \approx 1000 \text{ m}$) to avoid any significant CBL development during the simulation. Horizontal and vertical grid spacings are set to 25 m and 10 m respectively; the domain size is $4.8 \times 4.8 \text{ km}^2$ while the model top is placed at 2000 m . Sensitivity tests doubling the size of the domain in both horizontal directions did not reveal any significant differences in our analysis (not shown). The convective turnover time for the simulation ($t_* = z_*/w_*$) is $\approx 500 \text{ s}$, where w_* is the convective velocity scale and z_* is the CBL depth. The incompressible, Boussinesq equations were integrated for a time $30t_*$ and only data from after $23t_*$ were taken into consideration for the analysis to allow for the spin-up of turbulence (Efstathiou and Beare 2015). In addition, the vertical profiles presented in the following sections have been time averaged for $3t_*$ ($25 - 28t_*$).

2.2 Passive Tracer Threshold Method

The tracer threshold method of Couvreux et al. (2010) is implemented to diagnose the coherent structures in the CBL. A passive tracer is emitted from the surface with a constant surface flux and evolves according to

$$\frac{Dq}{Dt} = -\frac{q}{\tau_0}, \quad (2)$$

where q is the tracer mixing ratio. The tracer undergoes exponential (radioactive) decay on a constant time scale τ_0 . An activity operator I_2 is introduced; I_2 takes the value 1 at those grid points where the sampling criterion for the definition of a coherent structure is met, and takes the value zero elsewhere. (For later use, it is convenient to define $I_1 = 1 - I_2$.) For the method of Couvreux et al. (2010), fluid is defined to be part of a coherent structure at those locations where the tracer perturbation exceeds a height-dependent threshold,

$$I_2 = \begin{cases} 1 & \text{if } q'(x, y, z) > m_q \max[s_q(z), s_{\min}(z)] \quad \& \quad w(x, y, z) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

with $q'(x, y, z) = q(x, y, z) - \bar{q}(z)$ where the overbar now denotes a horizontal average at height z , $s_q(z)$ is the standard deviation of scalar mixing ratio at height z , and $s_{\min}(z)$ is a minimum threshold of the standard deviation (see Couvreux et al. 2010).

Couvreur et al. (2010) also examined the sensitivity to the additional $w > 0$ criterion in their definition of coherent structures in a shallow cumulus case. Here, we revisit this test in the dry CBL case in order to identify and construct the thermals as coherent structures. Strong downdrafts from penetrating convection and the detraining thermals from the CBL-top entrainment zone can penetrate far below the CBL top, down to the middle of the boundary layer (Crum and Stull 1987). Arguably, such eddies could also be considered part of the coherent structures, but would be excluded by the condition $w > 0$. Therefore, the criterion for the definition of coherent structures ($I_2 = 1$) from Eq. 3 is modified to

$$q'(x, y, z) > m_q \max[s_q(z), s_{\min}(z)]. \tag{4}$$

The extent to which Eq. 4 captures coherent downdrafts will depend on the overturning time of coherent structures relative to τ_0 and the time scale for lateral mixing. Nonetheless, the relaxation of the sampling criteria in Eq. 3 can provide insight into the sensitivity of the tracer threshold method and the turbulent transport mechanisms that it captures.

In order to achieve a clean implementation of the method we choose $m_q = 1$ and $\tau_0 = 15$ min as in Couvreur et al. (2010). Setting $\tau_0 = 60$ min led to a slight increase of the tracer mixing ratio at the CBL top; however, it did not significantly affect the area fraction of coherent structures (not shown).

2.3 Optimization of the Vertical Scalar Transfer

Any given distribution of the activity operator $I_2(x, y, z)$ defines a partition of the fluid into two parts: fluid 2 with $I_2 = 1$, and fluid 1 with $I_1 = 1 - I_2 = 1$, corresponding to coherent structures and their environment, respectively. If c is the potential temperature or the mixing ratio of any scalar quantity, then we can define

$$\sigma_i = \overline{I_i}; \quad \sigma_i c_i = \overline{I_i c}; \quad \sigma_i w_i = \overline{I_i w}. \tag{5}$$

Here $\sigma_2(z)$ is the area fraction labelled as coherent structures, $c_2(z)$ is the mean value of c within coherent structures at that height, and $w_2(z)$ is the mean value of w within coherent structures at that height, with analogous definitions for $\sigma_1, c_1,$ and w_1 . The total flux of the scalar is then given by

$$\begin{aligned} \overline{wc} &= \sigma_1 \overline{wc}^1 + \sigma_2 \overline{wc}^2, \\ &= \sigma_1 (w_1 c_1 + \overline{w'c'^1}) + \sigma_2 (w_2 c_2 + \overline{w'c'^2}), \end{aligned} \tag{6}$$

or

$$\begin{aligned} \overline{w'c'} &= \overline{wc} - \overline{w}\overline{c} = \sigma_1 \overline{w'c'^1} + \sigma_2 \overline{w'c'^2} \\ &\quad + \sigma_1 (w_1 - \overline{w})(c_1 - \overline{c}) + \sigma_2 (w_2 - \overline{w})(c_2 - \overline{c}), \end{aligned} \tag{7}$$

where \overline{wc}^i denotes averaging over the area fraction occupied by fluid i , and $\overline{w'c'}^i$ denotes contributions due to departures from the mean over fluid i .

The decomposition of the flow as presented in Eq. 7 is equivalent to the environment-updraft partitioning of the flow described in Siebesma and Cuijpers (1995) and Siebesma et al. (2007). Siebesma et al. (2007), as in much of the literature, then assume $\sigma_2 \ll 1$ and hence approximate $\sigma_1 \overline{w'c'^1} + \sigma_2 \overline{w'c'^2}$ by $\overline{w'c'^1}$, corresponding to the first term on the r.h.s. of Eq. 1, and neglect the term $\sigma_1 (w_1 - \overline{w})(c_1 - \overline{c})$ leaving $\sigma_2 (w_2 - \overline{w})(c_2 - \overline{c})$, corresponding to the second term on the r.h.s. of Eq. 1. Thus there is an approximate mathematical correspondence between the decomposition into coherent structures and environment in Eq. 7 and

the decomposition into local and non-local transport in Eq. 1. Note, however, that σ_2 may take values of around 0.2 in the CBL, so the approximations resulting from the assumption $\sigma_2 \ll 1$ may be of marginal validity. Consequently, some more recent EDMF schemes (e.g., Sušelj et al. 2019; Tan et al. 2018) relax this assumption.

Thuburn et al. (2018) proposed a multi-fluid approach to modelling the atmosphere in which they derived prognostic equations for the mean quantities in each type of fluid σ_i , w_i , c_i , etc. They envisaged applications in which these mean quantities would be predicted, and therefore resolved, by a model dynamical core for $i = 1$ and $i = 2$. For such a multi-fluid scheme the contributions

$$H = H_1 + H_2, \quad H_1 = \sigma_1(w_1 - \bar{w})(c_1 - \bar{c}), \quad H_2 = \sigma_2(w_2 - \bar{w})(c_2 - \bar{c}) \quad (8)$$

would be resolved while the contributions $\sigma_i \overline{w'c'^i}$ would be unresolved and therefore would still need to be parametrized. In this context, it is attractive to consider whether I_2 can be defined in such a way as to maximize the resolved contribution H or H_2 , and thereby minimise the contribution that needs to be parametrized.

Coherent structures are identifiable as relatively uniform objects with different thermodynamic and kinematic characteristics compared to the mean state of the atmosphere, as expressed by Eq. 8. In the mass-flux framework, they are considered responsible for most of the turbulent transport (see Wang and Stevens 2000). Furthermore, observations show that coherent motion, associated with strong thermals, is significantly more efficient in the transfer of scalars in the CBL compared to the contribution of local mixing (Lenschow and Stephens 1980). Therefore, it would be physically plausible to attempt to derive a coherent-structures area fraction that would in turn maximize the contribution of the resolved coherent motion to the total turbulent scalar transport. Additionally, defining I_2 so as to maximize H or H_2 may be attractive even when H or H_2 is to be parametrized, as is the case for mass-flux convection schemes and EDMF type schemes. The idea of defining coherent structures by maximizing their contribution to some scalar flux has the inherent advantage that no specification of arbitrary thresholds or other sampling criteria, nor any assumptions about the shape of the JFD, are necessary for the definition. With the multi-fluid approach in mind, in what follows we refer to H_1 and H_2 as the ‘resolved’ contributions to the flux in fluid 1 and fluid 2, respectively, and to $\sigma_1 \overline{w'c'^1}$ and $\sigma_2 \overline{w'c'^2}$ as the ‘unresolved’ contributions.

Two optimization methods for maximizing the resolved flux are introduced and tested here. In simple terms, both involve gradually incorporating more air in the diagnosed coherent structures, increasing their area fraction, until the resolved scalar flux does not increase any further and the optimum is found. One method (called OPT2) finds the global optimum, i.e. the optimum over all possible partitions, of the scalar flux. The result given in the Appendix shows that to find the global optimum it is sufficient to consider partitions in which I_2 is a function of c and w , i.e. we can work in terms of the JFD in $c - w$ space. The second method (called OPT1) also considers partitions in which I_2 is a function of c and w , but restricts itself to a limited subset of possible partitions. It has the benefit of simplicity and efficiency, and can be implemented easily in JFD space or in grid space, while producing results for many quantities of interest that are close to those from method OPT2, though there are also some interesting differences in detail. Method OPT1 is described first.

Method OPT1 scans the grid points at each level z and labels them as fluid 2 through the operator I_2 when the following condition is satisfied

$$I_2 = \begin{cases} 1 & \text{if } w(x, y, z) > w_p(z) \ \& \ c(x, y, z) > c_p(z) \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where p is the percentile of the distributions; note the same p is used for both the c and w distributions. For any given p , σ_2 is given by

$$\sigma_2 = \frac{1}{N_n N_m} \sum_{n=1}^{N_n} \sum_{m=1}^{N_m} I_2(n, m) \tag{10}$$

with N_n and N_m the number of grid points in the x and y direction respectively as expressed by the n, m indexes at each level. The mean vertical velocity perturbation of fluid 2 is then given by

$$\sigma_2(w_2 - \bar{w}) = \frac{1}{N_n N_m} \sum_{n=1}^{N_n} \sum_{m=1}^{N_m} I_2(n, m) (w(n, m) - \bar{w}), \tag{11}$$

(note that \bar{w} is identically zero for the Boussinesq simulation employed here), and similarly for the scalar mixing ratio perturbation in fluid 2

$$\sigma_2(c_2 - \bar{c}) = \frac{1}{N_n N_m} \sum_{n=1}^{N_n} \sum_{m=1}^{N_m} I_2(n, m) (c(n, m) - \bar{c}). \tag{12}$$

The parameter p is gradually increased, incorporating more grid points into fluid 2 and increasing σ_2 , until the resolved fluid-2 flux H_2 in Eq. 7 given by

$$H_2 = \sigma_2(w_2 - \bar{w})(c_2 - \bar{c}) \tag{13}$$

reaches its maximum value. Figure 1a illustrates three possible partitions during the search process in the OPT1 method.

Equation 13 corresponds to the mass-flux part of the EDMF scheme. As the proposed methodology is not based on any pre-assumptions about the partitioning of the flow into coherent structures and their environment, it is not self-evident that in a two-fluid representation of convection the contribution from the resolved fluid 1 should be by definition negligible (as usually assumed in the EDMF approach). Therefore, in order to further examine the relative contribution of the resolved fluxes in the decomposition of the flow and explore the impact of resolved fluid 1 on the optimization method, the same procedure is applied aiming to maximize this time the total resolved scalar flux

$$H = H_1 + H_2 \tag{14}$$

instead of H_2 , with H_1 the resolved scalar flux of fluid 1 representing a form of mass-flux by the compensating environment,

$$H_1 = \sigma_1(w_1 - \bar{w})(c_1 - \bar{c}). \tag{15}$$

The OPT1 method is relatively inexpensive since it involves only a one-dimensional search along the parameter p , but it considers only a subset of all possible partitions. The alternative OPT2 method optimizes over all possible partitions but is computationally more expensive. Since the optimum partition is known to be given by a straight line in $c - w$ space (see the Appendix), it is convenient to work in $c - w$ space and consider only activity operators of the form $I_2(c, w)$. First, at the height z of interest, the JFD $f(c, w)$ is constructed by assigning each grid point to suitably chosen bins in $c - w$ space. The mean quantities in fluid 2 are then given by sums over bins in $c - w$ space rather than over grid cells in physical space,

$$\sigma_2 = \sum_{n=1}^{N_c} \sum_{m=1}^{N_w} I_2(n, m) f(n, m), \tag{16}$$

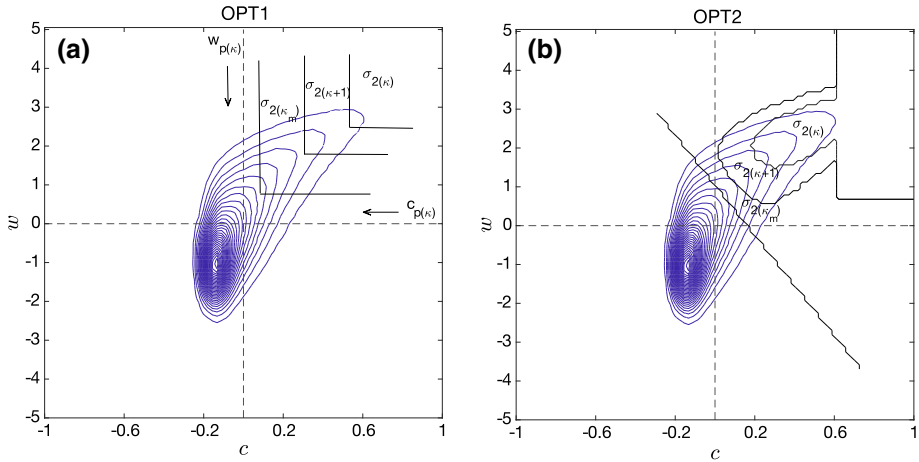


Fig. 1 Schematic description of the two different optimization methods, **a** OPT1, and **b** OPT2 of the $w - c$ JFD. $w_{p(\kappa)}$ and $c_{p(\kappa)}$ denote the percentiles (fractions) of the w and c distributions respectively at each iteration (κ) of the procedure (with κ_m the iteration where the maximum resolved flux has been achieved). σ_2 expresses the fluid-2 area fraction. The straight line in **b** corresponds to $\sigma_{2(\kappa_m)}$ (see also the Appendix). The contour lines represent the JFD of $w - c$

$$\sigma_2(w_2 - \bar{w}) = \sum_{n=1}^{N_c} \sum_{m=1}^{N_w} I_2(n, m) f(n, m) (w(m) - \bar{w}), \tag{17}$$

$$\sigma_2(c_2 - \bar{c}) = \sum_{n=1}^{N_c} \sum_{m=1}^{N_w} I_2(n, m) f(n, m) (c(n) - \bar{c}), \tag{18}$$

with N_c and N_w the number of bins in the c and w directions respectively.

A first guess for $I_2(c, w)$ is defined, for example as the top right-hand corner of $c - w$ space (Fig. 1b). Bins along the boundary of fluid 2 in $c - w$ space are then checked and are added to fluid 2 or removed from fluid 2 if this operation increases the value of H_2 or H (at each iteration the bin that makes the biggest increase is relabelled). When none of the bins checked gives an increase in H_2 or H then the maximum value has been found. Figure 1b illustrates three possible partitions during the search process in the OPT2 method. Having found the optimum I_2 as a function of c and w , it is then straightforward to label grid points at that height in grid space by checking their c and w values. Since the OPT2 method performs a global optimization over all possible partitions, it has no built-in assumptions about which part of the JFD corresponds to the coherent structures. The OPT1 method builds in the minimal physical assumption that the coherent structures are dominated by the first quadrant (see Fig. 1a), rather than one of the other quadrants. As will be shown later, the two optimization methods result in incorporating different physical mechanisms for the scalar flux associated with coherent structures.

In order to better understand the optimization procedure, Fig. 2 shows the different components of the scalar flux (Eq. 7) as a function of the fluid-2 area fraction σ_2 for the OPT1 and OPT2 optimization methods for a typical horizontal cross-section of the flow in the middle of the boundary layer. It can be seen that the flux components do not exhibit significant differences between the optimization methods, especially concerning the position of the maxima.

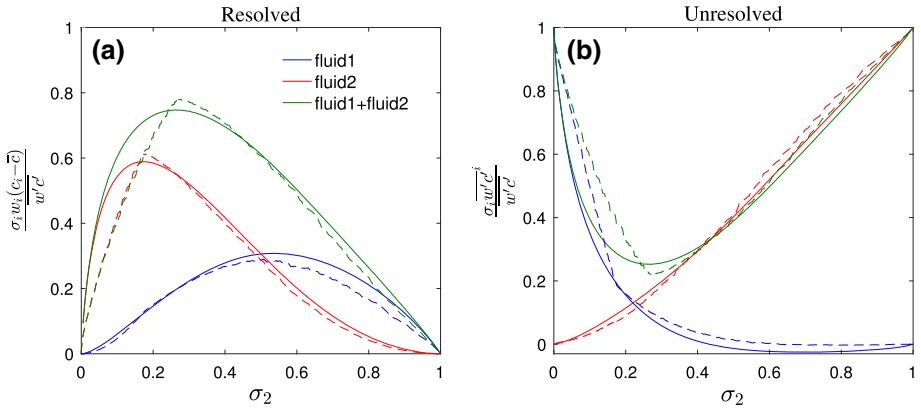
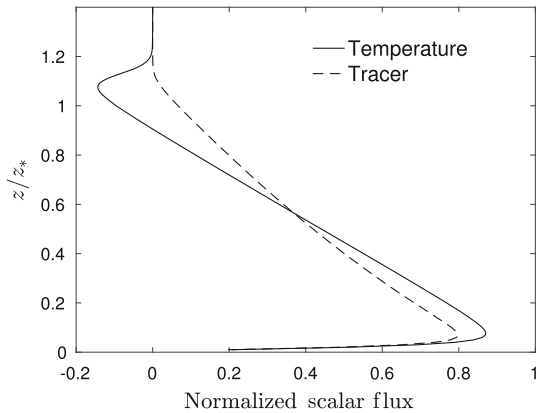


Fig. 2 Normalized, **a** resolved, and **b** unresolved scalar flux from the decomposed total scalar flux as shown in Eq. 7 from fluid $i = 1, 2$ as a function of fluid 2 area fraction (σ_2), at $z/z_* = 0.5$. Solid lines represent the OPT1 method and dashed lines the OPT2 method. Note that for OPT2 the fluxes are not unique functions of σ_2 , but they depend on the first guess for I_2 . Here we combine results for two first guess conditions: (i) only the top right corner of the JFD is included, so that σ_2 increases from approximately zero to its optimum value, and (ii) the whole JFD except the bottom left corner is included, so that σ_2 decreases from approximately one to its optimum value

Fig. 3 Time- and horizontally-averaged potential temperature flux (bold line) and decaying tracer flux (dashed line), normalized by their respective surface fluxes from the LES run



Nonetheless, it becomes obvious from Fig. 2 that H_2 and H_1 each have a clear maximum, for H_2 at $\sigma_2 \approx 0.2$ and H_1 at $\sigma_2 \approx 0.5$, while their sum (H) maximizes at $\sigma_2 \approx 0.3$.

It should be noted that the proposed method can be applied using either a passive tracer or the potential temperature as the scalar c . The choices $c = q$ (governed by Eq. 2) and $c = \theta$ give almost identical results for the lower half of the boundary layer. However, when $z/z_* > 0.5$, the choice $c = \theta$ fails to capture thermals that have lost their buoyancy as heat flux turns negative (see also Couvreur et al. 2010) and the maximization of the resolved flux is achieved at large area fractions σ_2 , as will be discussed later in Fig. 4. This can be seen in Fig. 3 where the vertical profile of the horizontally-averaged sensible heat flux and the “radioactive” tracer flux from the LES run is shown. The heat flux turns negative at about $z/z_* > 0.8$ while the tracer flux retains its positive sign up to the top of the CBL. Note also that the tracer-flux profile exhibits a slight curvature as the tracer-flux convergence balances the radioactive sink.

2.4 Particle Tracking

A particle model was applied in order to obtain some Lagrangian information about the fluid identified as belonging to coherent structures. High frequency LES output (every 10 s) was used to derive particle paths from the simulated flow field following Taylor et al. (2014). A number of particles (36,864) were released from the surface layer at $20t_*$ and the analysis of Lagrangian paths started after the spin-up of the flow, allowing for the particles to become well mixed. The different methods described above and shown in Table 1 were used to define coherent structures, and hence to determine whether each particle belongs to a coherent structure at any given time. For each continuous period during which a particle belongs to a coherent structure, its maximum vertical velocity and maximum buoyancy over that period were diagnosed. It should be noted that the Lagrangian approach is not used here to define coherent motion but to derive these diagnostics that highlight some interesting differences amongst the methods tested for defining coherent structures.

3 Results

Table 1 summarises the different methods examined for identifying coherent structures. TRACER and TRACERNW use the radioactive tracer threshold method of Couvreur et al. (2010), with and without the condition $w > 0$, as given by Eqs. 3 and 4, respectively. All of the optimization methods except for OPT2HF use the same radioactive tracer q for the scalar c , and use either the OPT1 or OPT2 method to maximize either H_2 or H , as indicated by their mnemonics. Finally, OPT2HF uses the OPT2 method to maximize the sensible heat flux ($c = \theta$).

3.1 Area Fraction

Figure 4 shows the vertical profile of the coherent-structure area fraction σ_2 diagnosed using all seven methods in Table 1. The default tracer threshold approach (TRACER) returns an area fraction of about 10 – 20%, very similar to the findings of Couvreur et al. (2010) (and references therein) from simulating an evolving CBL. Moreover, the derived σ_2 in our quasi-steady state CBL is also similar to the vertical profiles of boundary-layer thermal fraction from Rio et al. (2010) for the ARM shallow cumuli case. Excluding the $w > 0$ criterion (TRACERNW - Eq. 4) results in small differences in σ_2 that are more pronounced near the inversion. This will be analyzed further when examining the vertical and horizontal cross-sections of the defined thermals.

The area fraction of coherent structures from the optimization of H_2 , using methods OPT1H₂ and OPT2H₂, closely follows the area fraction profile from the tracer threshold method, although it remains slightly larger for the whole of the CBL compared to the tracer threshold method. As shown in Fig. 2, methods OPT1H₂ and OPT2H₂ produce almost identical area fractions except near the inversion where some marginal differences are present. In contrast, coherent structures defined by methods OPT1H and OPT2H occupy a significantly larger area than all other approaches, reaching about 40% near the surface.

Additionally, Fig. 4 shows the vertical distribution of σ_2 when optimizing the resolved heat flux instead of the tracer flux using the OPT2 optimization method approach (OPT2HF). Results are almost identical to method OPT2H up to $z/z_* \approx 0.5$; however, they start to deviate significantly further upwards as heat flux tends to zero (Fig. 3), with OPT2HF exhibiting two

Fig. 4 Time-averaged vertical profiles of area fraction occupied by coherent structures (fluid 2), derived from the different methodologies shown in Table 1

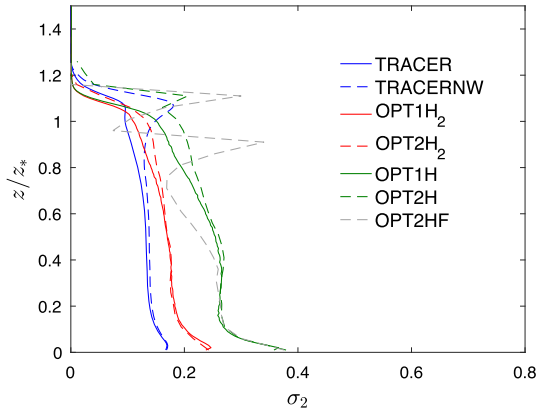


Table 1 Summary of the different methods used in this study for the partitioning of the flow in coherent structures (fluid 2) and local small-scale turbulence (fluid 1)

Method	Partitioning criteria
TRACER	Equation 3
TRACERNW	Equation 4
OPT1H ₂	Optimize over p -percentiles—maximizing H_2 ($c = q$)
OPT2H ₂	Full optimization—maximizing H_2 ($c = q$)
OPT1H	Optimize over p -percentiles—maximizing H ($c = q$)
OPT2H	Full optimization—maximizing H ($c = q$)
OPT2HF	Full optimization—maximizing total resolved heat flux ($c = \theta$)

spikes in σ_2 where the total sensible heat flux reaches zero and then just above where it attains its minimum value (see Fig. 3).

The vertical and horizontal cross-sections of vertical velocity from the LES at 3.2 h after the start of the simulation are shown in Figs. 5 and 6 respectively, together with the outlines of thermals as defined by the different methods. In accordance with Fig. 4, the two flavours of the tracer approach do not exhibit any significant differences up to the middle of the CBL (compare Fig. 6a, b) and produce different results only in the upper part of the CBL (compare Fig. 5a, b). The tracer approach without $w > 0$ (TRACERNW) incorporates some of the returning flow ($w < 0$) from ascending thermals and entrainment flux from the inversion which increases σ_2 in the inversion layer (see Figs. 4 and 5b). The vertical (Fig. 5c, e) and horizontal (Fig. 6c, e) cross-sections from the H_2 optimization produce very similar structures to the tracer threshold method. One notable difference is the merging of low-level plumes into larger ones that leads to larger area fractions in the optimization approaches (e.g. $x \approx -800$ m near the surface in Fig. 5, see also Fig. 4). Coherent structures defined using the H optimization (OPT1H and OPT2H) occupy a significantly larger area compared to the other methods as is especially obvious in the horizontal cross-sections in Fig. 6d, f. One thing to note is that method OPT2H diagnoses slightly larger thermals than method OPT1H, mainly on the upper part of CBL (see Fig. 5d, f), which also can be seen in Fig. 4.

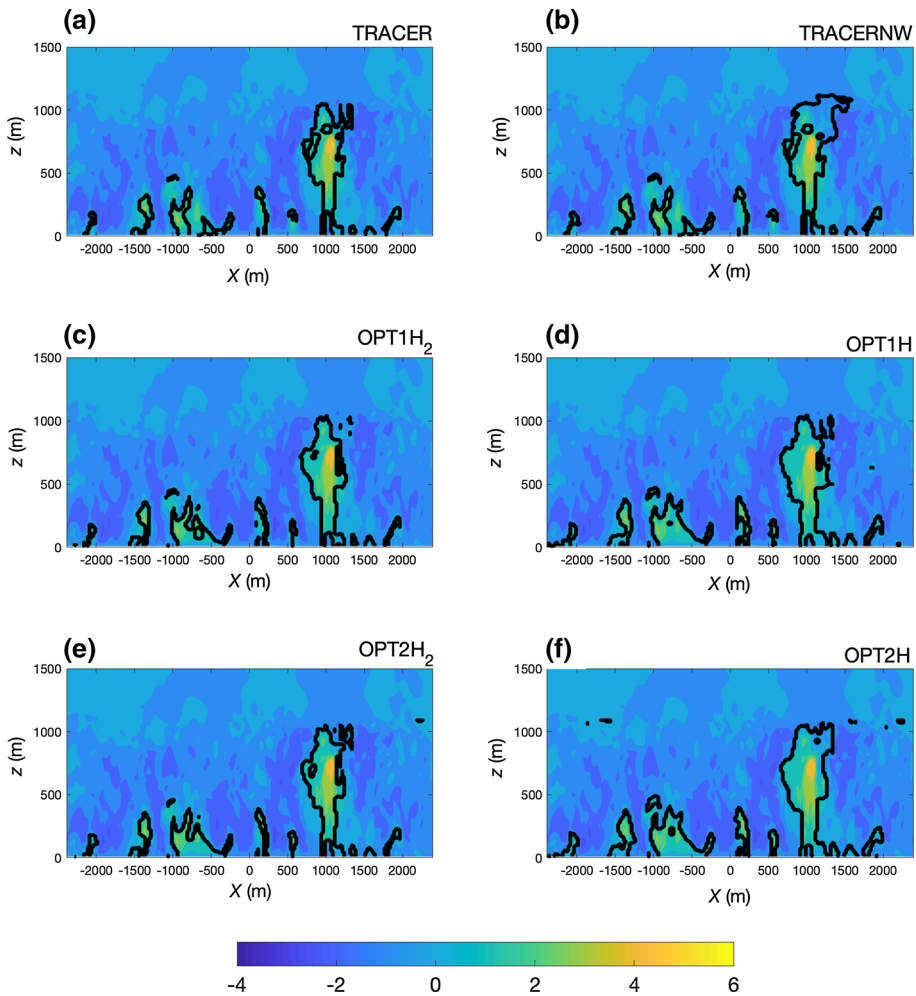


Fig. 5 Vertical cross-section of the w field (colourbar in m s^{-1}) from the LES and the outlines of defined thermals (black bold line) according to the different methods from **a** to **f** (see also Table 1) at $t = 3.2$ h after the start of the simulation

3.2 Sensible-Heat-Flux Partitioning

Based on the partitioning of the flow into coherent structures (fluid 2) and their environment (fluid 1) using the different methods, Fig. 7 shows the partitioning of the sensible heat flux into resolved and unresolved components for both fluids according to Eq. 7 (with $c = \theta$). Using the tracer threshold approach, most of the heat transfer is being done by the coherent structures (Fig. 7d), as captured by the mean, i.e. resolved properties of fluid 2 (see Couvreur et al. 2010), while the next most important mechanism is the unresolved heat flux in fluid 1 (local mixing). This is similar to the partitioning of the heat flux in the EDMF scheme. However, the increase in σ_2 when applying the optimization technique (Fig. 4) leads to a reduction in the heat flux due to unresolved motion in fluid 1 (Fig. 7a) and an increase in the contributions due to resolved motion in fluid 1 (Fig. 7b) and due to unresolved motion

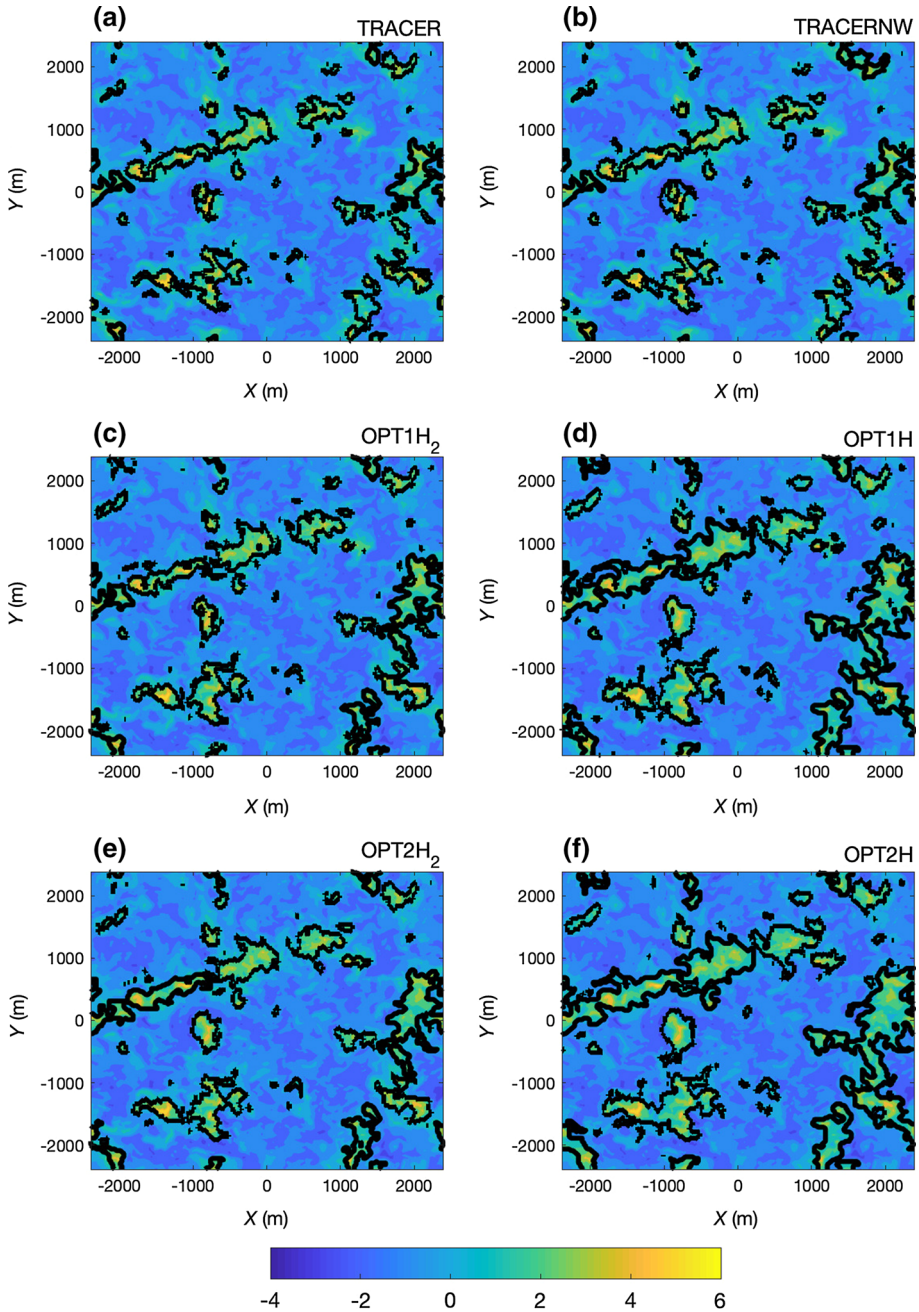


Fig. 6 Same as Fig. 5 but for the horizontal cross-sections at $z/z_* = 0.5$

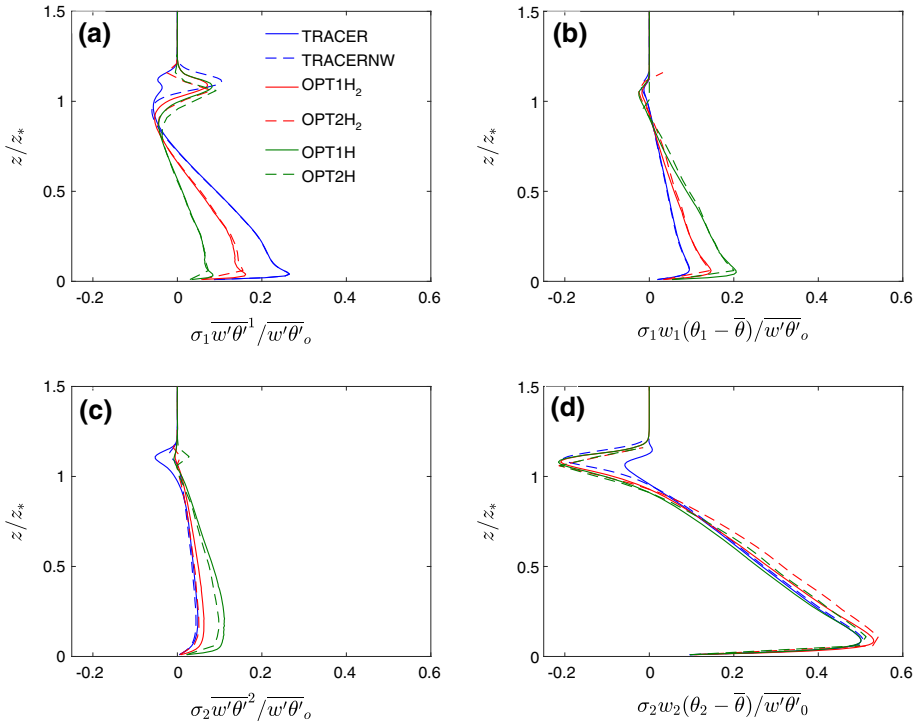


Fig. 7 Time- and horizontally-averaged components of the total heat flux (see also Eq. 7) from the LES according to the different definitions of fluid 2: **a** subfilter fluid-1 flux, **b** resolved fluid-1 flux, **c** subfilter fluid-2 flux, and **d** resolved fluid-2 flux. Fluid 2 represents coherent fluxes responsible for the non-local heat flux. $\overline{w'\theta'_0}$ denotes the surface heat flux

in fluid 2 (Fig. 7c). This effect becomes more pronounced when maximizing H (OPT1H and OPT2H) and the area fraction of coherent structures is substantially larger. Similar to Figs. 4, 5, and 6, differences between the optimization methods OPT1 and OPT2 are very small.

In contrast, the non-local transport due to coherent structures (fluid 2) remains almost unchanged for the bulk of the CBL regardless of the method used for the partitioning of the flow (Fig. 7d), representing almost 50% of the total heat flux in the lower CBL. The most apparent difference among the methods is seen in the strong negative heat flux in the inversion when the $w > 0$ criterion is excluded, which is partly compensated by some positive heat flux due to unresolved motions in fluid 1 in the inversion layer (Fig. 7a). The default TRACER method (the only method to require $w > 0$ in fluid 2) has a much smaller negative heat flux at the inversion (Fig. 7d), and the boundary-layer-top entrainment flux is represented primarily through the unresolved motions in fluid 2 (Fig. 7c).

3.3 Quadrant Analysis

In order to shed some light on the physical mechanisms that are incorporated in the different definitions of coherent motion, the JFD of $\theta - w$ at two levels, $z/z_* = 0.5$ and 0.9 , is presented in Fig. 8a, b respectively. The four quadrants of the JFD represent different

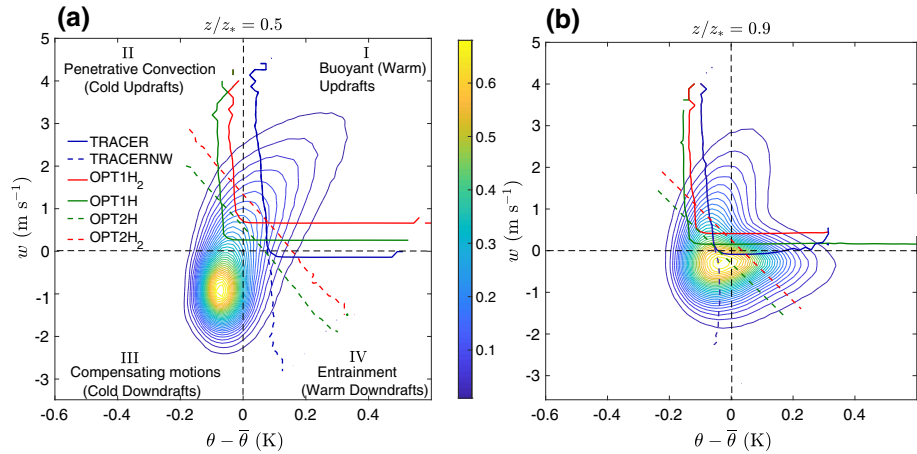


Fig. 8 Normalized JFD of \bar{w} and $\bar{\theta}$ perturbation (coloured contours and colour bar) in **a** the middle of the CBL ($z/z_* = 0.5$) and **b** near the inversion ($z/z_* = 0.9$). The bold borderlines depict the 0.01 contour of the ratio between the JFD from the grid points where $I_2 = 1$ to the total JFD as derived from the different fluid-2 definitions (see Table 1)

physical processes responsible for the turbulent heat transfer in the CBL (see Mahrt and Paumier 1984): The first quadrant (I) includes buoyant updrafts (thermals), (II) includes cold updrafts related to penetrative convection (updrafts that have lost their buoyancy but still have enough momentum to keep rising in the CBL), (III) includes cold downdrafts that express compensating motions in the CBL, and (IV) includes warm downdrafts originating from the entrainment of potentially warmer air from the inversion layer.

For the methods that optimize the tracer flux, the partitioning of the flow into fluid 1 and fluid 2 is described exactly by a curve in $c - w$ space (Fig. 1). However, in $\theta - w$ space, the borderline between fluid 1 and fluid 2 is no longer perfectly sharp because a given bin in $\theta - w$ space may contain some grid points labelled as fluid 1 and some grid points labelled as fluid 2 (because they have different values of c). Similarly, for the TRACER and TRACERNW methods the borderline between fluid 1 and fluid 2 is not perfectly sharp in $\theta - w$ space. Nevertheless, for all the methods, the partitioning may be visualised in $\theta - w$ space as follows. The JFD in $\theta - w$ space is computed first using only those grid points at height z where $I_2 = 1$ and second using all grid points at height z . In the middle of the boundary layer, the ratio of these two JFDs changes fairly abruptly between 1 and 0, and the 0.01 contour of the ratio gives a clear indication of the partition. Near the inversion the ratio changes rather more gradually between 1 and 0, indicating that the partition in $\theta - w$ space becomes rather blurred at these heights. Figure 8 shows this contour for several of the methods listed in Table 1.

In the middle of the CBL (Fig. 8a), the standard tracer threshold method (TRACER) only incorporates upward motion ($w > 0$) in the definition of coherent structures, including a large part of quadrant (I) of buoyant upward motion, and with a minimum buoyancy threshold of about 0.1 K. If the tracer threshold method is not restricted to upward motion only (TRACERNW), it incorporates the entrainment fluxes (Figs. 5b, 6b) at almost the same buoyancy threshold, which results in the increase of σ_2 especially near the inversion (see Fig. 4). On the other hand, optimizing H_2 (OPT1H₂) leads to a definition of coherent motion that includes buoyant updrafts (quadrant (I)), similar to the tracer threshold method, together with a small

part of the cold ‘overshooting’ thermals (quadrant (II)) close to neutral buoyancy. However, method OPT2H₂ shows a different partitioning of the flow with a straight line to define the relative fractions of fluid 1 and 2. Even though OPT2H₂ incorporates the same mechanisms and occupies a similar area fraction as method OPT1H₂ (see Fig. 4), it excludes some low velocity and less buoyant thermals while including only the strongest of the cold updrafts. Optimizing the total resolved scalar flux (H) in method OPT1H increases σ_2 including more cold updrafts in OPT1H while method OPT2H starts to encompass the contribution of the warmest downdrafts (entrainment fluxes) from a small area of quadrant (IV).

Near the top of the boundary layer (Fig. 8b), the shape of the JFD exhibits a double ‘lobe’ structure with most updrafts concentrated in near-neutral or negative buoyancy, and most warm perturbations occurring in descending air from the entrainment zone. All of the methods examined here incorporate most of the warm updrafts (quadrant (I)) and some cold updrafts, i.e. ‘overshooting’ thermals (quadrant (II)), in their diagnosed coherent structures. TRACERNW and OPT2H both include contributions from all quadrants, with method OPT2H incorporating more of the cold compensating downdrafts (see Figs. 4, 5f). Method OPT2H₂ includes the warm updrafts as well as some cold updrafts and some warm downdrafts, but almost none of the cold compensating subsidence. Although the different methods produce similar magnitude of negative heat flux (see Fig. 7d) in the entrainment zone (except for TRACER due to the $w > 0$ criterion), they include different parts of the JFD.

4 Kinematic and Thermodynamic Characteristics of Diagnosed Coherent Structures

4.1 Vertical Velocity Profiles

Figure 9 shows the vertical profile of w_1 and w_2 according to the different definitions of fluid 2. The tracer threshold (TRACER and TRACERNW) and optimization of H_2 (OPT1H₂ and OPT2H₂) methods produce similar vertical velocities, with the maximum w_2 comparable to the convective velocity scale (w_*) in the middle of the CBL. The TRACER method exhibits slightly stronger w_2 compared to the optimization methods, as well as smaller area fraction σ_2 (Fig. 4), implying that it captures less of the more weakly ascending fluid (see also Fig. 8). The TRACERNW method shows a reduced w_2 in the upper half of the CBL, along with a larger σ_2 compared with TRACER, because it captures descending as well as ascending fluid (again see Fig. 8).

At the same time, OPT1H₂ and OPT2H₂ methods have a more pronounced w_1 than the tracer threshold method, which results in the increase of the resolved fluid 1 contribution to the sensible heat fluxes shown in Fig. 7b. Optimizing H (OPT1H) leads to a further reduction of w_2 ($w_2 \approx 0.8w_*$), which in turn increases the magnitude of w_1 (see also Fig. 4). As a result the ‘environment’ mass-flux contribution (resolved fluid 1) reaches its maximum at the expense of the unresolved fluid 1 heat flux (see Fig. 7a, b).

The choice of the optimization method (OPT1 or OPT2) does not seem to have significant impact on the vertical velocity profiles compared to the effect of maximizing H_2 versus H . One notable difference between the OPT1 and OPT2 methods is the slight increase of w_2 at $z/z_* \approx 1.1$ in OPT2, associated with some strong overshooting thermals ($w \gg 0$) that occupy small area fractions (see Figs. 4, 8b). In any case, the differences due to the optimization method between OPT1H₂ and OPT2H₂ are similar to those between methods

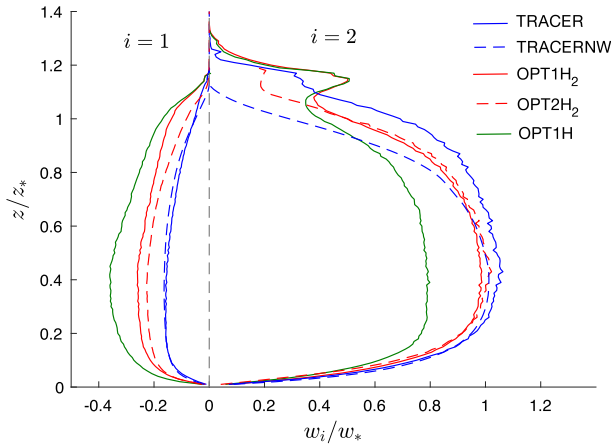


Fig. 9 Profile of the time-averaged vertical velocity (w_i) in fluid 1 ($i = 1$) and fluid 2 ($i = 2$) normalized by the convective velocity scale (w_*), as derived from the different definitions of fluid 2. The grey dashed line depicts the zero w line

OPT1H and OPT2H (not shown). For the rest of the analysis, only differences between the OPT1H₂ and OPT2H₂ methods will be shown.

4.2 Distributions of Maximum Buoyancy and Vertical Velocity

The term ‘coherent structure’ implies that fluid parcels move coherently over some distance and some period of time. However, the definitions of coherent structure discussed here and listed in Table 1 do not directly use any Lagrangian information about the fluid parcel’s history (though some such information is contained implicitly in the distribution of potential temperature and other scalars). It is useful, therefore, to examine some of the Lagrangian properties of the fluid identified as belonging to coherent structures. Some small but interesting differences are found between the different methods of diagnosing coherent structures.

Using high-frequency output from the LES, a large number of particle paths were computed, as described in Sect. 2.4. For each particle and for each time t , the w , θ , and q were interpolated to the particle position, so that any of the methods listed in Table 1 could be used to determine whether the particle was within a coherent structure. For each continuous time period that a particle belonged to a coherent structure, its maximum vertical velocity and maximum buoyancy over that time period were diagnosed. These diagnostics were accumulated over all such time periods for all particles during roughly the last hour of the LES to compile the results presented below. For comparison, the same diagnostics were also computed based on continuous time periods for which a particle had $w > 0$, and based on time periods for which a particle had $w > w_{p_w}$ with $p_w = 5\%$.

The bold black lines in Fig. 10 present the frequency distribution of maximum attained vertical velocity (Fig. 10a) and maximum buoyancy (Fig. 10b) using the $w > 0$ criterion, i.e. over periods of continuous particle ascent. The maximum w distribution exhibits an almost bimodal behaviour comprising a narrow peak, where most of the particles are found around a low maximum major mode of w values, and a second broader distribution including the strongest maximum attained w . Selecting coherent structures defined according to the TRACER method produces a distribution of maximum w that includes all of the strongest

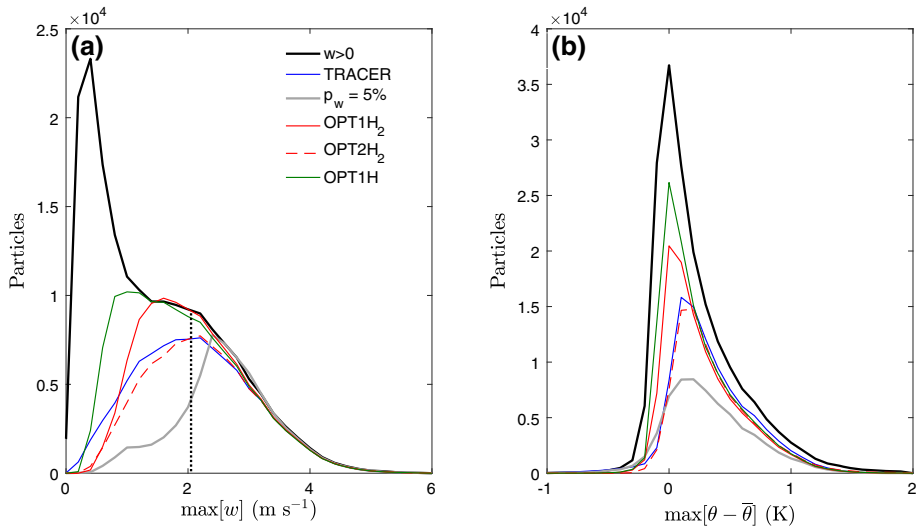


Fig. 10 Frequency distribution of **a** maximum w and **b** maximum buoyancy ($\theta - \bar{\theta}$) attained from fluid-2 labeled particles. The bold black line represents the distribution of all upward moving particles. Different colour lines represent the distributions from the different definitions of coherent motion (see legend). The grey bold line depicts the distribution of coherent motion as the 5% percentile of w . The black dotted line corresponds to the value of w_* .

maximum vertical velocities. The mode of the distribution is comparable to w_* , while the frequency of lower maximum w is significantly reduced. Similar behaviour is exhibited by using the OPT1H₂ and OPT2H₂ methods, with the OPT2H₂ distribution being very similar to TRACER. One thing to note is that there is a minimum value of maximum w greater than zero in both OPT1H₂ and OPT2H₂ as also shown in Fig. 8. The OPT1H method has similar behaviour, however it includes more events of lower maximum w , moving the mode of the distribution to lower values (see also Fig. 9).

Figure 10a also shows the distribution of maximum attained vertical velocity for those particles with w in the top $p_w = 5\%$ of w values at the same height (grey bold line). In contrast to the physical behaviour of the previous methods, the distribution for $p_w = 5\%$ is missing a large part of the strong updrafts close to w_* even though it incorporates the strongest of the maximum updrafts. Additionally, as shown in Fig. 10b, choosing the top 5% of the w distribution omits many of the most buoyant particles from the diagnosed coherent structures. On the contrary, TRACER incorporates most of the buoyant particles similar to the optimization approaches. However OPT1H₂ and especially OPT1H label more non-buoyant particles as fluid 2, moving the mode of their distribution closer to zero. The OPT2H₂ method exhibits a similar distribution to TRACER as also seen in Fig. 10a.

For each time period that a particle spends in a coherent structure, the in-thermal ascent h is calculated as

$$h = z_{\text{end}} - z_{\text{start}}, \quad (19)$$

where z_{start} is the height at which a particle enters fluid 2 and z_{end} is the height at which it leaves fluid 2 (relabelled as fluid 1). Figure 11 shows the relative frequency distribution of maximum w for four different definitions of coherent structure when attention is restricted to certain ranges of h . When all values of h are included (Fig. 11a) the distributions from the different definitions are similar to each other as also shown in (Fig. 10a), with the OPT1H

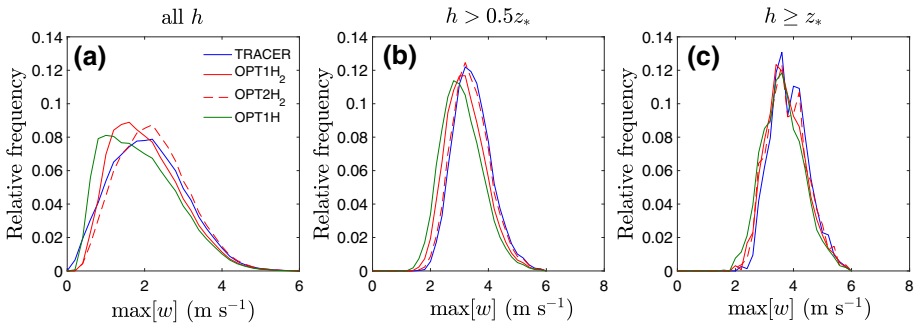


Fig. 11 Relative frequency distribution of maximum attained w from fluid-2 particles defined in coherent structures using different methods (see legend) according to their in-thermal ascent (h): **a** all thermals **b** thermals deeper than $0.5z_*$, and **c** thermals equal or deeper than z_*

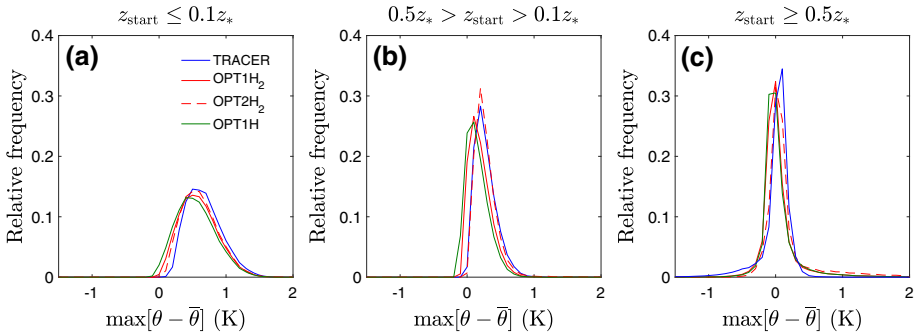


Fig. 12 Relative frequency distribution of maximum attained buoyancy ($\theta - \bar{\theta}$) from fluid-2 particles defined in coherent structures using different methods (see legend) with starting points (z_{start}) **a** the surface layer ($0.1z_*$) **b** in the lower half of the CBL ($0.5z_*$), and **c** in the upper half of the CBL (z_*)

method including in fluid 2 more particles that have low maximum w . When attention is restricted to $h > 0.5z_*$ (Fig. 11b) the mode of all distributions increases, and even though the differences among the methods are not substantial, the OPT1H mode is again slightly shifted to lower maximum w values. As expected, the deepest in-thermal ascents ($h \geq z_*$) attain higher vertical velocities (Fig. 11c) while any differences among the methods used for the definition of coherent structures are not significant in this case.

Figure 12 shows the relative frequency distributions of maximum buoyancy for particles identified as belonging to coherent structures according to four of the criteria examined, for three different ranges of starting height z_{start} . It can be seen that there are only minor differences among the four criteria. Particles starting their ascent from the surface layer ($z_{\text{start}} \leq 0.1z_*$) are positively buoyant and include the most buoyant particles due to the strong surface heating in the superadiabatic layer. Only the OPT1H method incorporates some slightly negatively buoyant particles as it significantly increases the area fraction of coherent structures in the surface layer compared to the other methods (see also Fig. 4). Similarly, particles starting above the surface layer and up to the middle of the boundary layer ($0.5z_* > z_{\text{start}} > 0.1z_*$) are still buoyant even though their distribution is narrower and their mode is closer to neutrally buoyant values. Particles in coherent structures starting their ascent from the upper, slightly stable part of the CBL ($z_{\text{start}} \geq 0.5z_*$) are distributed almost uniformly around the neutral buoyancy values and are probably being entrained into thermals

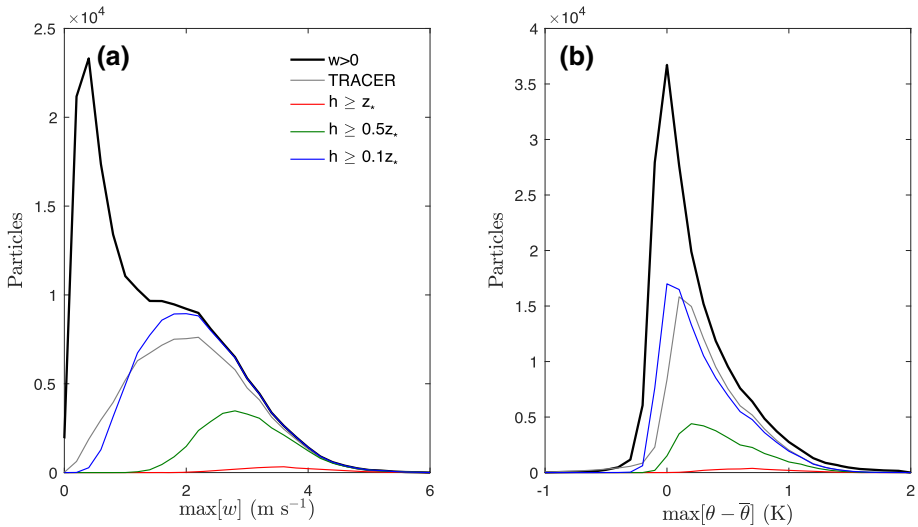


Fig. 13 Frequency distribution of, **a** maximum w , and **b** maximum buoyancy $(\theta - \bar{\theta})$ attained by the upward moving particles. The bold black line represents the distribution of all upward moving particles. Different colour lines represent the distributions from upward moving particles with different in-thermal ascent. For comparison, the grey bold line depicts the distribution of coherent motion as defined by the TRACER method

ascending from lower in the boundary layer. The distribution based on the TRACER method peaks at slightly higher buoyancy values compared to the other methods, exhibiting a longer tail towards negative buoyancy particles.

Finally, Fig. 13a depicts the distribution of maximum w for different h when a coherent structure is defined solely by $w > 0$. As shown in Fig. 11, the deepest thermals attain the highest maximum vertical velocities. However, continuously ascending particles with $h \geq z_*$ account for only a small fraction of strong upward motion as is evident in Fig. 13a. The same is true for the most buoyant particles (Fig. 13b). In fact, we must include all particles with $h \geq 0.1z_*$ in order to capture the total of the maximum vertical velocities (see Fig. 13a) and most of the buoyant particles. The distributions for $h \geq 0.1z_*$ seem to be very similar to those for TRACER, which are also presented in Fig. 13. That said, most particles ascend only about $0.1z_* - 0.2z_*$ during a continuous period within a coherent structure defined using the TRACER approach (not shown).

5 Discussion and Concluding Remarks

A new method for identifying coherent structures in the CBL has been presented. It is based on labelling the fluid in each grid box as either fluid 2 (coherent structure) or fluid 1 (environment) in such a way as to maximize the contribution to some scalar flux that is resolved by the mean properties of the two types of fluid. Even though the optimization method makes no explicit use of the concept of a coherent structure, it is found that fluid 2 picks out the thermals with length scales comparable to the boundary-layer depth (see Fig. 5) that are responsible for the non-local vertical transport in the CBL. The new approach does not depend on any arbitrary parameters or assumptions about the distribution of the scalar flux; it merely scans the JFD of $c - w$ until the maximum resolved flux is achieved. Four variants of the optimization

method have been explored, with two methods for scanning the JFD (OPT1 and OPT2) and two choices for maximizing either the total resolved flux H or the fluid-2 resolved flux H_2 .

The OPT2 method searches for the global maximum of the scalar flux by considering arbitrary partitions of the $c - w$ JFD, adding (or removing) bins until the maximum is achieved. The resulting partition is a straight line in $c - w$ space, as shown in the Appendix. The method OPT1 considers only a subset of the possible partitions of the $c - w$ JFD, as described in Sect. 2.3. Although it does not achieve the global maximum flux, the results for many diagnostics, including fluid 2 area fraction, heat flux, and vertical velocities, are very similar to those from method OPT2. However, OPT1 and OPT2 methods can incorporate different physical heat transfer mechanisms as represented by different quadrants of the $\theta - w$ JFD. At the same time OPT1 is computationally faster than OPT2, and is also more similar to currently used methods for identifying thermals or updrafts.

The main difference amongst the optimization approaches comes from the choice of maximizing either H or H_2 . Using H results in larger area fractions σ_2 and weaker w_2 resulting in broader thermals as the method might incorporate more low velocity and less buoyant updrafts and/or some returning flow from the entrainment zone. In contrast, H_2 produces the strongest resolved upward velocity w_2 , with maximum close to w_* , while the derived optimum σ_2 agrees better with previous studies, compared to maximizing H . This difference is also obvious in the heat-flux decomposition where OPT1H and OPT2H methods exhibit the most pronounced contribution of resolved heat flux from fluid 1 (the mass-flux contribution from the environment) in the total heat flux.

Results from the optimization method for the definition of coherent structures are compared to those from the tracer threshold method (TRACER) proposed by Couvreux et al. (2010). Maximizing H_2 produces similar profiles of σ_2 and almost identical w_2 to TRACER, diagnosing coherent structures with very similar horizontal and vertical structures. The impact of the $w > 0$ criterion in the tracer threshold method was also tested and differences were only found in the upper part of the boundary layer, where removing the $w > 0$ constraint led to the inclusion of entrainment fluxes and returning flow in the definition of coherent structures. This is in agreement with the findings of Couvreux et al. (2010) in the sub-cloud layer, although differences in the cloud layer might be more pronounced. Moreover, the tracer threshold method here returns a coherent structure area fraction very similar to the cloud-free case of Couvreux et al. (2010) (and to that found in references from measurement-focused studies therein) using the natural choice $m_q = 1$ (Eq. 3). Nevertheless, as shown in Chinita et al. (2018) the tracer threshold method exhibits some increased sensitivity to the m_q and τ_0 parameters in the cloud-free CBL.

Optimization of the total sensible heat flux (OPT2HF) gives almost identical results to OPT2H in the lower half of the boundary layer. However, near the inversion, where the total sensible heat flux changes sign, OPT2HF produces unsatisfactory results that do not correspond to our intuitive ideas of coherent structures. The spikes in the diagnosed fluid 2 area fraction near the inversion (Fig. 4) are one symptom of this. It is instructive to consider why optimizing the sensible heat flux is unsatisfactory in these circumstances. Intuitively, maximizing the sensible heat flux seems inappropriate when the coherent structures of interest are known to carry a negative heat flux. This idea is backed up by examination of Eq. 30 with $c = \theta$, which shows that, provided both \tilde{w}_2 and $\tilde{\theta}_2$ are positive, the optimal partition is described by a straight line of negative slope, which assigns predominantly quadrant (I) of the $\theta - w$ JFD to fluid 2. However, near the inversion a large proportion of the thermals become negatively buoyant and move to quadrant (II) (Fig. 8b). Then the optimal partition of the $\theta - w$ JFD is of the wrong form to capture these. Equation 30 also shows that when the total resolved heat flux H is zero the optimal partition line must pass through the origin;

this suggests an area fraction σ_2 close to 0.5 (depending on the exact shape of the JFD), and partly explains the spikes seen in Fig. 4.

Using quadrant analysis reveals that for all of the methods the coherent structures are dominated by quadrant (I): the buoyant updrafts. However, there are some differences in the detail. In the middle of the boundary layer the TRACER method picks out fluid with $w > 0$ and buoyancy greater than a threshold of about 0.1 K, whereas the OPT1H₂ method picks out positively buoyant fluid with vertical velocity greater than a threshold of about 0.5 m s⁻¹ (see Fig. 8a). Further up, near the inversion, both methods include cold updrafts (quadrant (II)) that have enough momentum to overcome the stable stratification. However, the OPT2H₂ method produces a different partitioning of the flow. Even though it produces the same σ_2 as the OPT1H₂ method and also includes quadrants (I) and (II), it excludes some low vertical velocity low buoyancy fluid from fluid 2, and includes more large vertical velocity negatively buoyant fluid.

Decomposing the heat flux according to the different definitions of coherent structure results in four contributing components to the total heat flux, following Eq. 7 and similar to Siebesma and Cuijpers (1995). Regardless of the method used, the resolved fluid 2 explains most of the total heat flux. This can be attributed to the dominant contribution of quadrant (I) in the definition of fluid 2 for all methods. Using TRACER or TRACERNW leads to a further significant heat flux contribution from the unresolved perturbations in fluid 1, which represents small scale processes in the environment. This is in accordance with the simplifications of the updraft-environment decomposition in the EDMF scheme (Siebesma and Cuijpers 1995; Siebesma et al. 2007), which only takes into account the unresolved fluid-1 contribution (local fluxes in the environment) and the resolved fluid-2 contribution (non-local fluxes from coherent structures). In contrast, the increased area fraction in OPT1H₂ and OPT2H₂ results in a greater resolved fluid-1 contribution and a reduced unresolved fluid-1 contribution to the vertical turbulent heat transport (see Fig. 2). The resolved fluid-1 contribution becomes much more pronounced, together with the unresolved fluid-2 fluxes, in OPT1H and OPT2H, as mentioned above. The contribution of the resolved environment was also identified in Chinita et al. (2018) when applying their JFD approach in a shallow cumulus cloud layer. It seems that each method for the partitioning of the flow into coherent structures and environment, or local and non-local components, could be a better fit to a particular parametrization scheme making the objective partitioning of the flow extremely difficult as also discussed in Chinita et al. (2018).

The distribution of maximum w attained by upward moving particles is almost bimodal. All of the methods investigated for defining coherent structures are consistent with the interpretation of the minor mode as representing coherent structures and the major mode as representing their environment. Overall, the differences amongst the distributions from the different definitions of coherent structure are not very pronounced. However, one noticeable difference is the inclusion of more low maximum velocity particles when maximizing H , which leads to a shift of the fluid-2 distribution closer to the major mode, with a peak further from w_* . Moreover, particle analysis has shown that continuous ascent over a large vertical extent cannot be the sole criterion for the identification of coherent structures as there exist strongly buoyant particles that ascend over a relatively small vertical extent. This behaviour is probably related to the entrainment-detrainment processes and the continuous exchange of particles between fluid 2 and fluid 1. Furthermore, Lagrangian particle tracking is known to produce excessive switching of particle labels between the coherent structures and their environment (see Yeo and Roms 2013; Thuburn et al. 2019).

The main advantage of the optimization approach is its ability to define vertically and horizontally coherent structures (as shown in Figs. 5 and 6), with smooth area fractions

(Fig. 4) and heat fluxes (Fig. 7) without the need to specify a priori any sampling criteria or assumptions about the functional form of the JFD. Interestingly, no direct information about the length scales of the structures to be included in fluid 2, or about the Lagrangian history of the parcels in fluid 2, is included in the optimization methods, since such information is not present in the JFD. Nevertheless, the method picks out structures with horizontal and vertical scales comparable to the CBL depth and which remain coherent as they ascend over comparable scales.

One possible drawback of the optimization method can be considered the increased computational cost especially when applying the OPT2 approach. The tracer threshold method on the other hand, employs a rather physical and easy to calculate sampling criterion for the identification of coherent motion. However, the use of a tuneable parameter (m_q), induces some extra sensitivity to the tracer threshold method (see Couvreur et al. 2010; Chinita et al. 2018) which might be application-dependent. All in all, the two approaches produce similar results (setting $m_q = 1$ for the tracer threshold method) with comparable fluid-2 area fractions, distributions of maximum vertical velocities and decomposition of the sensible heat.

The proposed methodology for the identification of coherent structures can in principle be used with any scalar. However, as discussed above, tests using potential temperature produced some noisy spikes in σ_2 near the inversion where the heat flux changes sign. Therefore, in this study as a starting point, we have used the same decaying scalar as in the tracer threshold method to also facilitate a better comparison between the two methods. Moreover, we are currently exploring how the optimization method can be generalized to three-fluid flow (updrafts, downdrafts and environment) to account for cases where downdrafts significantly affect the thermodynamic structure of the atmosphere. Work is underway for the extension of the optimization method to cloudy boundary layers and deep convective clouds.

Acknowledgements This work was funded by the Natural Environment Research Council (NERC) under the ParaCon program, grant number NE/N013123/1 and used Monsoon2, a collaborative High Performance Computing facility funded by the Met Office and the Natural Environment Research Council.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

Define the vertical velocity and scalar mixing ratio perturbations as,

$$\tilde{c} = c - \bar{c}; \quad \tilde{c}_i = c_i - \bar{c}; \tag{20a}$$

$$\tilde{w} = w - \bar{w}; \quad \tilde{w}_i = w_i - \bar{w}. \tag{20b}$$

The resolved contribution to the scalar flux Eq. 7 is given by

$$H = \sigma_1 \tilde{w}_1 \tilde{c}_1 + \sigma_2 \tilde{w}_2 \tilde{c}_2, \tag{21}$$

where $\sigma_1 = 1 - \sigma_2$. It is a straightforward exercise in variational calculus to show that the partition of the fluid that gives the maximum of H at any height z is described by a straight line in $c - w$ space.

First note that

$$\sigma_i = \frac{1}{A} \int I_i dA, \quad \sigma_i \tilde{c}_i = \frac{1}{A} \int \tilde{c}_i I_i dA, \quad \sigma_i \tilde{w}_i = \frac{1}{A} \int \tilde{w}_i I_i dA, \quad (22)$$

where $I_1 = 1 - I_2$ and A is the domain area. Therefore it is convenient to write

$$\sigma_1 \sigma_2 H = \sigma_2 (\sigma_1 \tilde{w}_1) (\sigma_1 \tilde{c}_1) + \sigma_1 (\sigma_2 \tilde{w}_2) (\sigma_2 \tilde{c}_2). \quad (23)$$

Using the fact that $\sigma_1 \tilde{c}_1 = -\sigma_2 \tilde{c}_2$, $\sigma_1 \tilde{w}_1 = -\sigma_2 \tilde{w}_2$, and $\sigma_1 + \sigma_2 = 1$, this simplifies to

$$\sigma_1 \sigma_2 H = (\sigma_2 \tilde{w}_2) (\sigma_2 \tilde{c}_2). \quad (24)$$

Now consider the changes (indicated by δ) that result from a change δI_2 to I_2 ,

$$\sigma_1 \sigma_2 \delta H + \sigma_2 H \delta \sigma_1 + \sigma_1 H \delta \sigma_2 = (\sigma_2 \tilde{w}_2) \delta (\sigma_2 \tilde{c}_2) + (\sigma_2 \tilde{c}_2) \delta (\sigma_2 \tilde{w}_2), \quad (25)$$

where

$$\delta \sigma_2 = \frac{1}{A} \int \delta I_2 dA = -\delta \sigma_1, \quad (26)$$

$$\delta (\sigma_2 \tilde{c}_2) = \frac{1}{A} \int \tilde{c}_2 \delta I_2 dA, \quad (27)$$

$$\delta (\sigma_2 \tilde{w}_2) = \frac{1}{A} \int \tilde{w}_2 \delta I_2 dA, \quad (28)$$

so that

$$\sigma_1 \sigma_2 \delta H = \frac{1}{A} \int \{ \sigma_2 H - \sigma_1 H + (\sigma_2 \tilde{w}_2) \tilde{c}_2 + (\sigma_2 \tilde{c}_2) \tilde{w}_2 \} \delta I_2 dA. \quad (29)$$

In order for H to be a maximum, δH must equal zero for arbitrary changes δI_2 . This will be satisfied only if the expression in curly braces in Eq. 29 vanishes at the locations where I_2 may change, i.e., along the boundary between fluid 2 and fluid 1. Thus, that boundary is given by

$$(\sigma_2 - \sigma_1) H + (\sigma_2 \tilde{w}_2) \tilde{c}_2 + (\sigma_2 \tilde{c}_2) \tilde{w}_2 = 0. \quad (30)$$

This result is useful in several ways. First, it shows that the optimal boundary between fluid 2 and fluid 1 is a straight line in $c - w$ space. Thus, the search for the optimum partition may be carried out in terms of the JFD in $c - w$ space rather than in physical $x - y$ space. Second, it may be used to accelerate the practical computation of the optimum partition. For example, using the OPT2 search method described in Sect. 2.3, a first approximation to the optimum partition may be found quickly, starting from an arbitrary $I_2(c, w)$, using rather coarse data bins in $c - w$ space; that first approximation may then be used to calculate the coefficients in Eq. 30, which may then be used to provide a good first guess for a rerun of the search using much finer bins. Third, this result also provides some understanding of why the optimization method is less useful when applied to the sensible heat flux, whose vertical profile changes sign, as discussed in Sect. 3.1 and Sect. 5.

A slight modification of the above derivation shows that the partition that maximizes H_2 is described by

$$-\sigma_2 H_2 + (\sigma_2 \tilde{w}_2) \tilde{c}_2 + (\sigma_2 \tilde{c}_2) \tilde{w}_2 = 0. \quad (31)$$

The derivation may be adapted straightforwardly to show that the partition that maximizes an arbitrary linear combination of the resolved scalar flux, scalar variance, and w variance (either the total or the contribution only from fluid 2) is also described by a straight line in $c - w$ space.

References

- Angevine WM (2005) An integrated turbulence scheme for boundary layers with shallow cumulus applied to pollutant transport. *J Appl Meteorol Climatol* 44(9):1436–1452
- Angevine WM, Jiang H, Mauritsen T (2010) Performance of an eddy diffusivity-mass flux scheme for shallow cumulus boundary layers. *Mon Weather Rev* 138(7):2895–2912
- Berg LK, Stull RB (2004) Parameterization of joint frequency distributions of potential temperature and water vapor mixing ratio in the daytime convective boundary layer. *J Atmos Sci* 61(7):813–828
- Chatfield RB, Brost RA (1987) A two-stream model of the vertical transport of trace species in the convective boundary layer. *J Geophys Res Atmos* 92(D11):13,263–13,276
- Chinita MJ, Matheou G, Teixeira J (2018) A joint probability density-based decomposition of turbulence in the atmospheric boundary layer. *Mon Weather Rev* 146(2):503–523
- Couvreur F, Hourdin F, Rio C (2010) Resolved versus parametrized boundary-layer plumes. Part I: a parametrization-oriented conditional sampling in large-eddy simulations. *Boundary-Layer Meteorol* 134(3):441–458
- Crum TD, Stull RB (1987) Field measurements of the amount of surface layer air versus height in the entrainment zone. *J Atmos Sci* 44(19):2743–2753
- Dawe JT, Austin PH (2013) Direct entrainment and detrainment rate distributions of individual shallow cumulus clouds in les. *Atmos Chem Phys* 13:7795–7811
- Deardorff JW (1972) Theoretical expression for the countergradient vertical heat flux. *J Geophys Res* 77(30):5900–5904
- Efstathiou GA, Beare RJ (2015) Quantifying and improving sub-grid diffusion in the boundary-layer grey zone. *Q J R Meteorol Soc* 141(693):3006–3017
- Heus T, van Dijk G, Jonker HJJ, Van den Akker HEA (2008) Mixing in shallow cumulus clouds studied by Lagrangian particle tracking. *J Atmos Sci* 65(8):2581–2597
- Hong SY, Pan HL (1996) Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon Weather Rev* 124(10):2322–2339
- Hong SY, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon Weather Rev* 134(9):2318–2341
- Hourdin F, Couvreur F, Menut L (2002) Parameterization of the dry convective boundary layer based on a mass flux representation of thermals. *J Atmos Sci* 59(6):1105–1123
- Lenschow DH, Stephens PL (1980) The role of thermals in the convective boundary layer. *Boundary-Layer Meteorol* 19(4):509–532
- Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB (2000) A new boundary layer mixing scheme. Part I: scheme description and single-column model tests. *Mon Weather Rev* 128(9):3187–3199
- Mahrt J, Paumier J (1984) Heat transport in the atmospheric boundary layer. *J Atmos Sci* 41(21):3061–3075
- Pergaud J, Masson V, Malardel S, Couvreur F (2009) A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction. *Boundary-Layer Meteorol* 132(1):83
- Rio C, Hourdin F (2008) A thermal plume model for the convective boundary layer: representation of cumulus clouds. *J Atmos Sci* 65(2):407–425
- Rio C, Hourdin F, Couvreur F, Jam A (2010) Resolved versus parametrized boundary-layer plumes. Part II: continuous formulations of mixing rates for mass-flux schemes. *Boundary-Layer Meteorol* 135(3):469–483
- Roms DM (2010) A direct measure of entrainment. *J Atmos Sci* 67(6):1908–1927
- Siebesma AP, Cuijpers JWM (1995) Evaluation of parametric assumptions for shallow cumulus convection. *J Atmos Sci* 52(6):650–666
- Siebesma AP, Teixeira J (2000) An advection-diffusion scheme for the convective boundary layer: description and 1d-results. In: *Proceedings of the 14th symposium on boundary layers and turbulence*, American Meteorological Society, Aspen, CO, pp 133–136
- Siebesma AP, Soares PMM, Teixeira J (2007) A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *J Atmos Sci* 64(4):1230–1248
- Soares PMM, Miranda PMA, Siebesma AP, Teixeira J (2004) An eddy-diffusivity/mass-flux parameterization for dry and shallow cumulus convection. *Q J R Meteorol Soc* 130(604):3365–3383
- Sullivan PP, Patton EG (2011) The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulation. *J Atmos Sci* 68(10):2395–2415
- Sušelj K, Teixeira J, Chung D (2013) A unified model for moist convective boundary layers based on a stochastic eddy-diffusivity/mass-flux parameterization. *J Atmos Sci* 70(7):1929–1953
- Sušelj K, Kurowski MJ, Teixeira J (2019) On the factors controlling the development of shallow convection in eddy-diffusivity/mass-flux models. *J Atmos Sci* 76(2):433–456

- Tan Z, Kaul CM, Pressel KG, Cohen Y, Schneider T, Teixeira J (2018) An extended eddy-diffusivity mass-flux scheme for unified representation of subgrid-scale turbulence and convection. *J Adv Model Earth Syst* 10:770–800
- Taylor AC, Beare RJ, Thomson DJ (2014) Simulating dispersion in the evening-transition boundary layer. *Boundary-Layer Meteorol* 153(3):389–407
- Thuburn J, Weller H, Vallis GK, Beare RJ, Whittall M (2018) A framework for convection and boundary layer parameterization derived from conditional filtering. *J Atmos Sci* 75(3):965–981
- Thuburn J, Efstathiou GA, Beare RJ (2019) A two-fluid single-column model of the dry, shear-free, convective boundary layer. *Q J R Meteorol Soc* 145(721):1535–1550
- Troen I, Mahrt L (1986) A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Boundary-Layer Meteorol* 37(1–2):129–148
- Wang S, Stevens B (2000) Top-hat representation of turbulence statistics in cloud-topped boundary layers: a large eddy simulation study. *J Atmos Sci* 57(3):423–441
- Wyngaard JC, Moeng CH (1992) Parameterizing turbulent diffusion through the joint probability density. *Boundary-Layer Meteorol* 60(1):1–13
- Yeo K, Roms DM (2013) Measurement of convective entrainment using lagrangian particles. *J Atmos Sci* 70(1):266–277

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.