# Detecting Cold Pool Family Trees in Convection Resolving Simulations

**Jannik Hoeller[1,2]** 🔸, **Romain Fiévet[2]** 🔸, and **Jan O. Haerter[1,2,3,4]**

[1]Integrated Modeling, Leibniz Centre for Tropical Marine Research, Bremen, Germany, [2]Niels Bohr Institute, Copenhagen University, Copenhagen, Denmark, [3]Physics and Earth Sciences, Constructor University Bremen, Bremen, Germany, [4]Department of Physics and Astronomy, University of Potsdam, Potsdam, Germany

**Abstract** Recent observations and modeling increasingly reveal the key role of cold pools in organizing the convective cloud field. Several methods for detecting cold pools in simulations exist, but are usually based on buoyancy fields and fall short of reliably identifying the active gust front. The current cold pool (CP) detection and tracking algorithm (CoolDeTA), aims to identify cold pools and follow them in time, thereby distinguishing their active gust fronts and the "offspring" rain cells generated nearby. To accomplish these tasks, CoolDeTA utilizes a combination of thermodynamic and dynamical variables and examines the spatial and temporal relationships between cold pools and rain events. We demonstrate that CoolDeTA can reconstruct CP family trees. Using CoolDeTA we can contrast radiative convective equilibrium (RCE) and diurnal cycle CP dynamics, as well as cases with vertical wind shear and without. We show that the results obtained are consistent with a conceptual model where CP triggering of children rain cells follows a simple birth rate, proportional to a CP's gust front length. The proportionality factor depends on the ambient atmospheric stability and is lower for RCE, in line with marginal stability as traditionally ascribed to the moist adiabat. In the diurnal case, where ambient stability is lower, the birth rate thus becomes substantially higher, in line with periodic insolation forcing—resulting in essentially run-away mesoscale excitations generated by a single parent rain cell and its CP.

**Plain Language Summary** Cold pools are cooled air masses below thunderstorm clouds, produced when rain evaporates underneath such clouds. Cold pools are important, as they produce strong gusts and have been associated with clumping of rain cells, whereby heavy rainfall over relatively small areas could be generated—with implications for flooding. The current work describes a method that helps identify such cold pools in computer simulation data. In contrast to earlier methods, we here show that the interaction between a CP and its surroundings can be reconstructed by the method. We show that this identification works under a range of contexts, such as when horizontal wind is applied in the simulations or when the surface temperature is not constant—as might often be the case over a land surface. The identification reveals interesting dynamical effects, such as that in some cases, cold pools can kick-start a form of chain reaction, by which "rain cell children" of it give rise to additional cold pools that again produce children, and so forth. The dynamics revealed is in line with expectations of widespread, so-called mesoscale convective systems over land, whereas over an ocean surface the dynamics is much less explosive.

## 1. Introduction

Convectively-generated cold pools are dense air masses forming beneath precipitating thunderstorm clouds when a fraction of hydrometeors evaporates within the subcloud layer (Droegemeier & Wilhelmson, 1985; Simpson, 1980). Besides the higher density compared to their surroundings, cold pools manifest themselves by a divergent near-surface wind field and vortical structures along their "dynamical edges," that is, the demarcation between the cold pool (CP) and the ambient atmosphere. Many recent studies have considered the characteristics of cold pools (Drager & van den Heever, 2017; Fournier & Haerter, 2019; Gentine et al., 2016; Henneberg et al., 2020; Schlemmer & Hohenegger, 2016; Torri & Kuang, 2019), because they have been implicated in the spatio-temporal organization of the convective cloud field—thus impacting non-random organization of rainfall, with consequences for extreme events, such as flooding (Böing, 2016; Haerter et al., 2019; Jensen et al., 2021; Schlemmer & Hohenegger, 2016; Simpson, 1980; Tompkins, 2001b).

Observationally, measuring cold pools is difficult, as high-resolution spatio-temporal records of the subcloud boundary layer are required to track a given spreading CP. Field campaigns thus often remain restricted to a

particular, often spatially confined area (Feng et al., 2015; Hohenegger et al., 2023), use point-like or lower-dimensional measurements (Engerer et al., 2008; Zuidema et al., 2017), or need to resort to indirect signals, such as cleared areas in radar imagery (de Szoeke et al., 2017) or combinations of point and radar measurements (Kruse et al., 2022). Yet, from existing observational studies, it is evident that CP interactions help structure the cloud field and qualify as a mechanism to induce thunderstorm precipitation cells (Moncrieff & Liu, 1999). Measurements of the ocean temperature after CP events also point to lowered sea surface temperatures after pronounced CP occurrences (Pei et al., 2018), thus possibly giving rise to locally anti-correlated rain activity.

Numerical approaches hardly suffer from a lack of data coverage. Yet, in the past, three-dimensional simulations at cloud-resolving $\mathcal{O}(100m)$ resolution have not been possible over domain sizes of several hundred kilometers and weeks, required for reaching steady state dynamics. In recent years, however, the required scales have become accessible, and numerical simulations aimed at improving our understanding of cold dynamics have become feasible. For instance, process-oriented computer simulations have helped appreciate the role of cold pools in organizing the thermal and momentum structure within the subcloud layer (Böing et al., 2012). Further, numerical studies have helped elucidate the role of surface heat fluxes, induced by cold pools as they move along (Torri & Kuang, 2016), a dynamics which might affect the formation of so-called moisture rings (Chandra et al., 2018). The partitioning of surface heat fluxes into latent and sensible contributions was found to be strongly dependent on soil moisture, a variable which conversely also affects the size and propagation speed of cold pools due to changes in boundary layer humidity and rain shaft areas (Drager et al., 2020; Gentine et al., 2016).

On the larger-scale level of simulated mesoscale organization, cold pools have been suggested as a possible cause for non-random distributions of rain cells in space (Haerter et al., 2019), which may arise when CP gust fronts collide to trigger new cells at the loci of collision (Böing, 2016; Moseley et al., 2016; Torri & Kuang, 2019). Cold pools were also found to affect the moisture distribution in the subcloud layer (Böing et al., 2012; Schlemmer & Hohenegger, 2016) and may thereby aid the transitioning from shallow to deep convection (Kurowski et al., 2018). Even the paradigmatic "convective self-aggregation" (CSA), a system-scale symmetry breaking, was found to be affected by CP activity (Haerter, 2019; Jeevanjee & Romps, 2013; Muller et al., 2022; Nissen & Haerter, 2021). In an effort to study more realistic configurations, recent numerical experiments employ temporally varying surface temperatures to investigate the impact of the diurnal cycle on correlated CP activity, especially regarding the formation of mesoscale convective systems (MCS) (Haerter et al., 2020; Jensen et al., 2021), stimulating conceptual models to help explain the complex processes involved (Niehues et al., 2021).

Interestingly, only few attempts have been made at parameterizing cold pools in larger-scale climate models (Grandpeix & Lafore, 2010; Rio et al., 2009; Rooney et al., 2021). This may be due to an incomplete understanding of the fundamental processes affecting CP interaction, dynamics, and thermodynamic modifications of the boundary layer, as well as the pre-moistening of the lower free troposphere. Process-focused studies that leave out parts of the components involved (Fiévet et al., 2023; Haerter & Schlemmer, 2018; Meyer & Haerter, 2020; Romps & Jeevanjee, 2016), may help improve parameterizations further.

Several recent works have brought forward methods to track cold pools, in particular their gust fronts, in numerical simulations (Drager & van den Heever, 2017; Fournier & Haerter, 2019; Henneberg et al., 2020; Hirt et al., 2020; Rochetin et al., 2021; Schlemmer & Hohenegger, 2016). Yet, the detection of cold pools in numerical simulations is far from trivial: (a) the area affected by rainfall, where the CP is fed by rain evaporation, is often not cleanly delimited and subject to setting a threshold value, for example, on the precipitation rate; (b) the CP density current, as a highly dynamic object, experiences turbulent mixing and heat exchange with the surface and ambient environment, and is subject to spontaneous symmetry breaking, such as under the formation of "lobe-and-cleft" instabilities—again feeding back on the dynamics (Härtel et al., 2000; Markowski & Richardson, 2010; Simpson, 1972; Wakimoto, 2001); (c) the larger-scale pattern of cold pools and the rain cells produced by their collisions, is highly complex, with families formed by CP-rain cell networks making for a challenging tracking problem.

The current work addresses these challenges by building an algorithm which identifies cold pools and their causal chains based on a combination of buoyancy and dynamical measures. We apply our algorithm to cloud resolving simulation data based on radiative convective equilibrium (RCE) and diurnal cycle setups with and without wind shear to examine the impact of the simulation setup on CP dynamics. Based on the identified causal chains of cold pools, we investigate how the total accumulated rainfall of a given CP relates to (a) its capacity to trigger new rain

cells and (b) the number of colliding cold pools which triggered its parent convection. Furthermore, we show examples of derived CP family trees and propose a simple conceptual model that replicates the observed cascade-like dynamics of cold pools within MCS.

## 2. Methods

### 2.1. Cloud Resolving Simulation Data

We employ the System for Atmospheric Modeling (SAM) (Khairoutdinov & Randall, 2006) to simulate the convective cloud field for a set of numerical experiments. The horizontal domain size is chosen to be square $L \times L$, with $L = 240$ km in all simulations, and the horizontal grid spacing $\Delta h$ is chosen as $\Delta h = 200$ m throughout, thus yielding $N = 1,200$ grid boxes in each horizontal dimension. Cyclic boundary conditions are chosen in both horizontal dimensions. We found this horizontal mesh resolution a suitable compromise, where relatively large domain sizes can be simulated for several days, yet, key CP effects, such as the dynamic gust front, can be simulated at satisfactory detail (Bryan et al., 2003; Fiévet et al., 2023; Hirt et al., 2020; Meyer & Haerter, 2020; Straka et al., 1993). The mesh is discretized along the vertical direction $z$ using 64 levels of increasing depth $\Delta z$, ranging from $\Delta z = 50$ m at the first level ($z = 25$ m) to 1,000 m at the domain's top boundary ($z = 26$ km). The model resolves a non-hydrostatic anelastic form of the Euler equations, and uses the liquid water static energy and the total precipitating/nonprecipitating water mixing ratios as prognostic variables. We use the embedded one-moment microphysics scheme (Khairoutdinov & Randall, 2003) along with a first-order turbulent subgrid scheme (Smagorinsky, 1963). The radiative fluxes are calculated using the National Center for Atmospheric Research Community Atmosphere Model version 3 (Collins et al., 2006).

The surface conditions are prescribed and taken as a (saturated) sea surface of a given, horizontally-homogeneous surface temperature $T_s$, that is, $T_s(x, y, t) = T_s(t)$. We distinguish two sensitivities:

1. To mimic a *diurnal cycle*, we allow $T_s(t)$ to vary sinusoidally as

$$T_s(t) = \overline{T}_s - \Delta T \cos(2\pi t/T). \tag{1}$$

here, $\overline{T}_s = 298$ K and $\Delta T$ is either chosen as 0 K, which we term "rce," or as 2 and 4 K, respectively, termed "diu."

2. To simulate *large scale wind shear*, we additionally impose a height dependent wind tendency in the $x$-direction, with a piecewise linear profile given by

$$u_0(z < 1 \text{ km}) = 0,$$

$$u_0(1 \text{ km} < z < 19 \text{ km}) = \delta_u \left( \frac{z - 1 \text{ km}}{18 \text{ km}} \right), \tag{2}$$

$$u_0(z > 19 \text{ km}) = \delta_u.$$

The velocity field is nudged toward this profile using a typical linear-relaxation term in the momentum equation with a timescale of $\tau = 1$ day. This timescale is chosen to preserve the internal subdiurnal variability occurring through convective organization, such as through CP dynamics.

The above settings for $\Delta T$ and wind shear give rise to the six setups used in this paper, namely rce0K, diu2K, and diu4K, which can each have $\delta_u = 0$ or $\delta_u = 16$ m s$^{-1}$. For each setup we run the simulation for 7.5 days.

The integration time step, $\sim \mathcal{O}(s)$, varies accordingly with a maximum admissible Courant-Friedrichs-Lewy (CFL) number of 0.8. The simulation output frequency is of one every $\Delta t = 10$ min. Each $\Delta t$, the following instantaneous fields are output: surface rainfall intensity $I(x, y, t)$, near-surface specific humidity $q(x, y, t) = q(x, y, z = 25 \text{ m}, t)$, virtual temperature $T_v(x, y, t) = T_v(x, y, z = 25 \text{ m}, t)$, the horizontal wind vector $\mathbf{u} \equiv (u, v)$, with $u = u(x, y, z = 25 \text{ m}, t)$ and $v = v(x, y, z = 25 \text{ m}, t)$ as the lowest-level horizontal wind components in the $x$ and $y$ directions, and near-surface vertical wind speed $w = w(x, y, z = 25 \text{ m}, t)$. For simplicity, we refer to $\Delta t$ as the "time step" throughout the paper. The discrete time $t = t_n \equiv n\Delta t$, measured from the beginning of the simulation, is often denoted by the integer time step $n$. Similarly, we use integers $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, N\}$ to

label the two discrete positions $x = x_i \equiv i\Delta h$ and $y = y_j \equiv j\Delta h$. Note that $x_i = x_{i+N}$ and $y_j = y_{j+N}$ in accordance with the cyclic boundary conditions. For a given 2D field $Q$ we often use $Q(i, j, n)$ to refer to the numerical value of $Q$ at horizontal position $(x_i, y_j)$ and time $t_n$.

## 2.2. Cold Pool Detection and Tracking Procedure

The main goal of the proposed algorithm is to detect and track cold pools in data from numerical simulations, while observing and identifying the involved causal relations between cold pools and rain cells: On the one hand, cold pools require evaporating precipitation and are thus inevitably linked to particular rain cells. To accommodate this causal relation, rain cells are taken as the origin of cold pools in the proposed algorithm. Thus, a rain cell is a required condition for the identification of a CP. Each rain cell in the domain is assigned a unique rain identification number (ID). In case a rain cell generates a CP, the CP receives the same ID as the associated rain cell. On the other hand, CP gust fronts can trigger new rain cells, which might again generate new cold pools (Böing, 2016; Haerter et al., 2019; Tompkins, 2001a). Accordingly, the algorithm stores rain cells triggered at a CP gust front, as well as potential cold pools emerging from these rain cells, as children of that particular CP. In this manner, the algorithm not only detects and tracks cold pools, but also their relationships.

The proposed CP detection and tracking algorithm (CoolDeTA) consists of four main steps: (a) identification of rain patches (RPs), (b) segmentation of the domain into "potential CP" and "no CP" regions, (c) derivation of markers, and (d) labeling of cold pools. These steps are performed for every output time step of the respective simulation or simulation time window.

For notational convenience we introduce several definitions:

- The deviation of any given scalar quantity $Q(i, j, n)$ relative to its spatial mean is denoted as

$$Q'(i,j,n) = Q(i,j,n) - \overline{Q}(n), \tag{3}$$

where $\overline{Q}(n)$ denotes the spatial mean at $t_n$.
- Occasionally, we require $Q$ at a given time step $t_n$ to be normalized to the range between zero and one. This is accomplished by the rescaling

$$\hat{Q}(i,j,n) = \frac{Q(i,j,n) - \min(Q(n))}{\max(Q(n)) - \min(Q(n))}, \tag{4}$$

where $\min(Q(n))$ or $\max(Q(n))$ refer to the minimum and maximum of the two-dimensional field $Q(n)$ at time step $t_n$.
- For simpler notation we write the Kronecker delta symbol as $\delta(a, b)$, which equals unity for $a = b$ and zero otherwise.
- Two (spatial) indices $(i, j)$ and $(i', j')$ are considered contiguous, if $dist((i,j) - (i',j')) = 1$, where $dist$ denotes the shortest distance operator, which ensures cyclic boundary conditions on the toroidal square lattice geometry.
- We use $p_r(i, j, n)$ and $p_{cp}(i, j, n)$ to indicate $N \times N$ integer fields for labeling RPs and CP patches at time step $t_n$. Analogously, we also define an auxiliary field $p_{seg}(i, j, n)$. For either, we define a superscript $m$ as labeling the subset of indices $(i, j, n)$ of $p_r$ or $p_{cp}$, where the integer is equal to $m$, for example, $p_r^m \equiv \{(i,j,n) \in [1,N] \times [1,N] \times [1 \wedge p_r(i,j,n) = m]\}$.

(i) *Identifying RPs*. At any discrete time $t_n$ a *RP* is defined as a spatially contiguous region in which the rain intensity $I(x_i, y_j, t_n) \geq I_0$. Generally, we employ $I_0 = 2$ mm hr$^{-1}$, which has been determined as critical rate for downdrafts penetrating into the subcloud layer (Barnes & Garstang, 1982). The algorithm assigns a unique ID to every RP identified by this procedure. In this way, $p_r(i, j, n)$ is populated at each time step $t_n$, where

$$p_r(i,j,n) = \begin{cases} 0 \text{ if } I(i,j,n) < I_0 \\ \text{otherwise: ID of respective rain patch.} \end{cases} \tag{5}$$

For robustness, a given RP, labeled $m$, at time step index $n$ is only retained if its area $A(p_r,n,m)$ $\equiv \Delta h^2 \sum_{i,j=1}^{N} \delta(p_r(i,j,n),m) \geq A_0$, where we use $A_0 = 2\,\text{km}^2$ as a threshold. Otherwise, the corresponding values of $p_r$ are set to zero.

(ii) *Segmentation*. At each time $t_n$, we use an indicator field

$$\mathcal{I}_{seg}(i,j,n) = \hat{s}'(i,j,n) + \left(1 - \hat{T}'_v(i,j,n)\right), \tag{6}$$

where $s'(i, j, n)$ are the horizontal wind speed fluctuations, computed as

$$s'(i,j,n) \equiv \sqrt{u'(i,j,n)^2 + v'(i,j,n)^2}. \tag{7}$$

$\mathcal{I}_{seg}(i,j,n)$ is a measure which we assume to be increased in areas where cold pools are present. Since cold pools are associated with fluctuating horizontal wind speeds but decreased virtual potential temperatures, $\hat{T}'_v(i,j,n)$ in Equation 6 is subtracted from unity, that is, relatively low $T_v$ increases the probability of a CP. The two quantities are combined with equal weights to ensure equal importance is given to both thermodynamic and dynamical features. We choose perturbations rather than absolute quantities in Equation 6 to be able to apply the method to simulations with other boundary conditions, such as non-zero surface winds, and thus increase the method's scope of application.

A k-means algorithm (Pedregosa et al., 2011) now subdivides every grid cell $(i, j)$ into two cluster groups, namely "potential CP" or "no CP" by minimizing the one-dimensional distance metric within both cluster groups. The result is a horizontal $N \times N$ binary segmentation $p_{bin}(i, j, n)$ with $p_{bin}(i, j, n) = 1$ for all "potential CP" grid cells and $p_{bin}(i, j, n) = 0$ for all "no CP" cells. As an example, Figure 1a shows the corresponding histogram of $\mathcal{I}_{seg}$ for diu4K at time step $t_{471}$ representing a morning scene. The lower and upper cluster groups represent regions where cold pools are considered prohibited versus feasible. A "cold pool" grid cell must fall within the feasible cluster group. Yet, additional criteria below are required.

Spreading quickly beyond the boundaries of their RPs, fully developed cold pools should at least encompass the same area as their related RPs. Thus, we apply the minimum area $A_0$ criterion from (i) also to the identified "potential CP" regions. For this reason, spatially contiguous "potential CP" regions in $p_{bin}(i, j, n)$ are identified and labeled with a unique integer number. The result is, analogous to (i), a horizontal $N \times N$ integer field $p_{seg}(i, j, n)$ where a patch, labeled $m$, is only retained if its area $A(p_{seg},n,m) \equiv \Delta h^2 \sum_{i,j=1}^{N} \delta\left(p_{seg}(i,j,n),m\right) \geq A_0$.

(iii) *Deriving markers*. A marker could be defined as a center of mass (COM) of a given RP $m$ at time $t_n$. We however found that the actual (dynamical) downdraft within this patch more precisely marks the CP center. A good choice for the downdraft field $\mathcal{I}_{Tw}(i,j,n)$ was found to be a combination of normalized temperature and vertical velocity, namely

$$\mathcal{I}_{Tw}(i,j,n) = \hat{T}(i,j,n) + \hat{w}(i,j,n), \tag{8}$$

so that for a given patch, labeled $m$,

$$\mathcal{I}_{mar}^m(i,j,n) = 1 - \hat{\mathcal{I}}_{Tw}^m(i,j,n), \tag{9}$$

which is large for locations of minimal temperature and vertical velocity. The corresponding COM, $\left(C_x^m, C_y^m\right)$, follows as

$$C_x^m = \frac{\sum_{i,j=1}^{N} x_i \mathcal{I}_{mar}^m(i,j,n)}{\sum_{i,j=1}^{N} \mathcal{I}_{mar}^m(i,j,n)}, \tag{10}$$
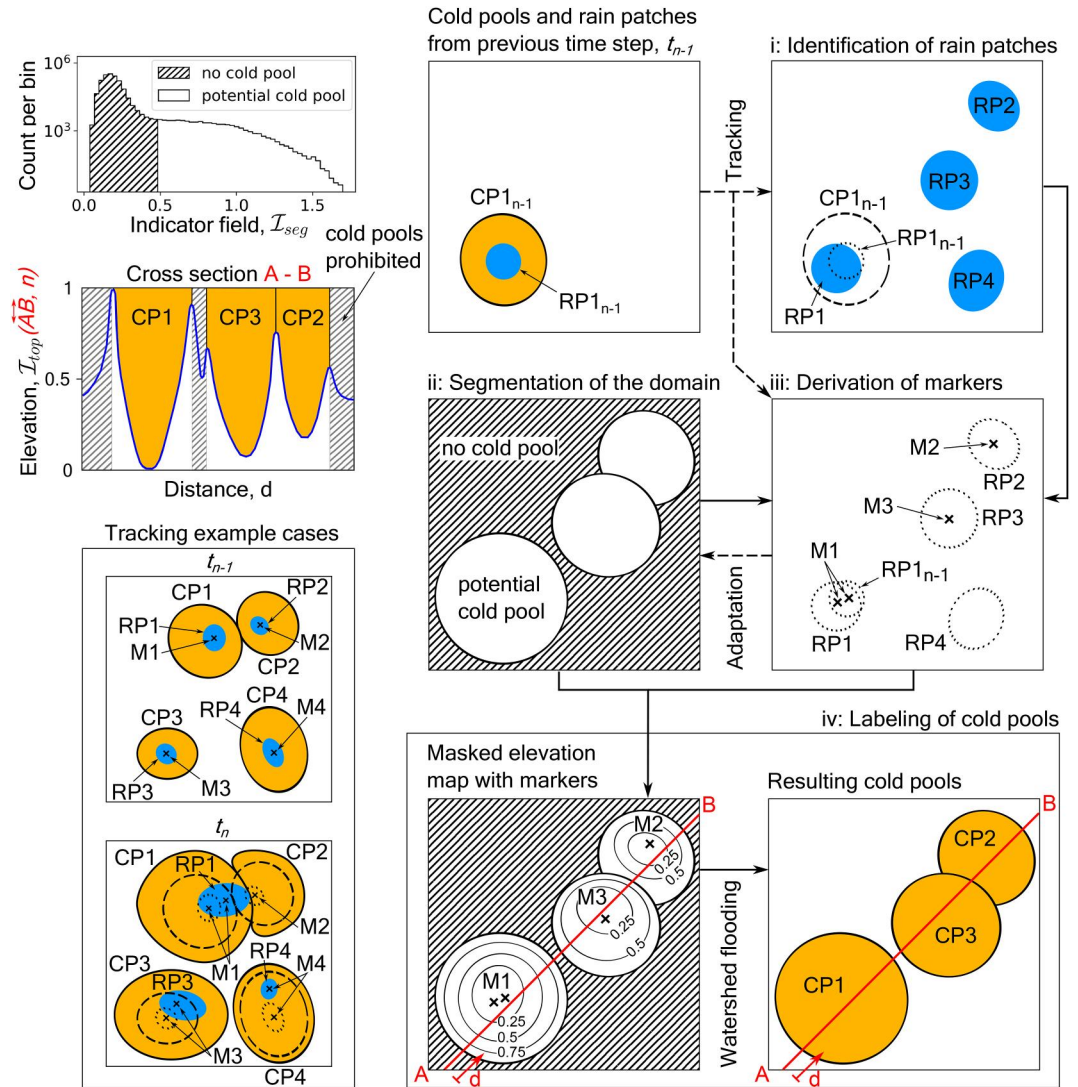
**Figure 1.** *Proposed cold pool (CP) detection and tracking algorithm.* (a) Histogram of the indicator field, $\mathcal{I}_{seg}(i,j,n)$ (Equation 6), for diu4K at $t_{471}$. "no cold pool (CP)" and "potential CP" grid cells, as determined by the k-means algorithm, are shown as hatched and striped curve fillings. (b) Main algorithm steps: Rain patches (RPs) and CPs are shown as blue and orange areas, respectively. Markers (Ms) are denoted as "x"-symbols; Dashed and dotted lines indicate CP and RP contours, respectively. The steps are repeated for every simulation output time step $\Delta t$. (c) Elevation map values (blue curve), $\mathcal{I}_{top}\left(\overleftrightarrow{AB},n\right)$, of grid cells on line $\overleftrightarrow{AB}$ defined in (b), step (iv) for time step $t_n$. Hatched areas are locations classified as "no CP" by the segmentation, and thus not flooded by the watershed algorithm. The elevation of these "no CP" regions can be thought of as infinite. Orange areas represent labeled CPs which result from flooding the elevation map from the markers M1, M2 and M3. (d) Example cases for the tracking of RPs and CPs. In time step $t_n$ (labeled above panels), RP1, RP3 and RP4 are tracked from $t_{n-1}$. Each marker represents an RP center of mass and acts as starting point for the watershed flooding.

and analogously for $C_y^m$. $\left(C_x^m, C_y^m\right)$ maps onto a 2D index $\left(i_{mar}^m, j_{mar}^m\right)$ by determining the nearest integer pixel indices. We collect these marker indices into a field $p_{mar}(i,j,n)$, where $p_{mar}\left(i_{mar}^m, j_{mar}^m, n\right) = m$ and which is zero for all locations without markers.

When the rain associated with a CP labeled $m$ from the previous time step $t_{n-1}$ stops, the CP will have no active RP $p_r^m$ in $p_r(i,j,n)$ anymore, that is, $\sum_{i,j=1}^{N}\delta(p_r(i,j,n),m) = 0$. However, cold pools may still spread and trigger new convection. Therefore, we in this case introduce their latest rain marker to $p_{mar}(i,j,n)$ instead.

Apart from the *rain* marker in the center of the active or latest RP, cold pools which were present in the previous time step $t_{n-1}$ obtain an *origin* marker. An origin marker represents the first rain marker which initially generated the CP, that is, the center of its initial RP. Analogous to the rain markers, origin markers are introduced to $p_{mar}(i, j, n)$. The combination of origin and rain markers increases the robustness of the final CP labeling, which will be illustrated in step (iv).

(iv) *Labeling cold pools*. For a time step $t_n$, the final field with labeled cold pools, $p_{cp}(i, j, n)$, is—analogous to the RP field $p_r(i, j, n)$ (compare Equation 5)—a horizontal $N \times N$ integer ID field where $p_{cp}(i, j, n)$ specifies the ID of a CP or $p_{cp}(i, j, n) = 0$, if no CP is present at $(i, j)$ and time index $n$. The CP field is obtained using a watershed algorithm (van der Walt et al., 2014), which can be thought of as placing different water sources in a topographical map and flooding it up to the highest elevation (see Figure 1c), such that the water table is equal in all basins. The resulting watersheds represent the boundaries between the different water sources and the corresponding basins represent regional minima.

In order to derive the labeled CP field, we provide the watershed algorithm with three ingredients: the locations of the water sources, that is, the marker locations $p_{mar}(i, j, n)$ computed in step (iii), an elevation map $\mathcal{I}_{top}$, and a mask, which acts like a stencil preventing certain regions from being flooded. The elevation map is computed as

$$\mathcal{I}_{top}(i,j) = \hat{T}(i,j) + \hat{q}^2(i,j) + \hat{w}(i,j). \tag{11}$$

Cold pool regions are characterized by cold (low $\hat{T}(i,j)$) and often dry air (low $\hat{q}(i,j)$) which forms basins with regional minima in the elevation map, whereas horizontal convergence with associated vertical updrafts towers up walls at CP boundaries (large $\hat{w}(i,j)$), thus separating the CP basins from each other. The mask, as a final ingredient to the watershed algorithm, is provided by the binary segmentation $p_{bin}(i, j, n)$ from step (ii). Whereas CP1 and CP3 in Figure 1c are separated by the mask's "no CP" region between them, CP2 and CP3 are in contact with each other. The contact location, that is, the point where the two cold pools collide (in 2D, it is a line), depends on the elevation map $\mathcal{I}_{top}$. Since water from different sources, representing different cold pools, does not mix, the contact location remains constant for the rest of the flooding.

Since only its gust front should separate a CP from its environment, we do not allow holes, that is, enclosed grid cells with different ID values, within individual cold pools. Thus, we close potential holes within individual cold pools by assigning each grid cell of the hole the ID of the surrounding CP. By this procedure, we thus merge cold pools if the grid cells of the hole carry the ID of other cold pools.

In later stages, a CP basin in the elevation map can consist of multiple local minima, caused by newly formed RPs within the CP area or an advected initial RP. By providing both rain and origin markers to cold pools which were present in the previous time step $t_{n-1}$, we make sure that cold pools in later stages of their life cycle are not accidently flooded by neighboring cold pools.

Figure 1b summarizes the main steps of CoolDeTA. Note that both tracking and the interactions between steps (ii) and (iii) will be described separately in the following Sections 2.3 and 2.6, respectively.

### 2.3. Cold Pool Tracking

Provided that RPs, $p_r(i, j, n - 1)$, and/or cold pools, $p_{cp}(i, j, n - 1)$, from the previous time step $t_{n-1}$ are available, we track cold pools already during the identification of RPs in step (i). Since a RP, labeled $m$ in $p_r(i, j, n)$, obtains a corresponding marker $p_{mar}(i^m_{mar}, j^m_{mar}, n) = m$, an associated CP which might be identified based on the watershed algorithm in step (iv) would be labeled $m$ in $p_{cp}(i, j, n)$, as well. The label $m$ thus represents the ID of a RP in $p_r(i, j, n)$ and the ID of an associated CP in $p_{cp}(i, j, n)$. Accordingly, cold pools can be tracked by adjusting the label of tracked RPs in step (i). For this purpose, we evaluate the spatial overlap of patches from consecutive time steps. A RP labeled $m'$ at time step $t_n$ is tracked if the overlap with a RP labeled $m$ from time step $t_{n-1}$, $O_r(m, m', n - 1, n) \geq f_r \times A(p_r^{m'})$ with $f_r = 0.01$ and

$$O_r(m, m', n - 1, n) \equiv \sum_{i=1, j=1}^{N,N} \mathbf{1}((i,j) \in S_1), \tag{12}$$

where **1** is the indicator function and the set $S_1 \equiv \{(i, j) \in [1, N] \times [1, N], p_r(i, j, n-1) = m \wedge p_r(i, j, n) = m'\}$, or if the overlap with a CP labeled $m$ from time step $t_{n-1}$, $O_{cp}(m, m', n-1, n) = f_{cp} \times A(p_r^{m'})$ with $f_{cp} = 1$ and

$$O_{cp}(m, m', n-1, n) \equiv \sum_{i=1, j=1}^{N,N} \mathbf{1}((i, j) \in S_2), \tag{13}$$

where the set $S_2 \equiv \{(i, j) \in [1, N] \times [1, N], p_{cp}(i, j, n-1) = m \wedge p_r(i, j, n) = m'\}$. Figure 1d visualizes both cases: On the one hand, RP3 is tracked in time step $t_n$ since the corresponding RP overlaps RP3 from the previous time step $t_{n-1}$. On the other hand, RP4 is tracked in time step $t_n$ since the corresponding RP fully overlaps CP4 from time step $t_{n-1}$.

If a RP labeled $m'$ is overlapped sufficiently by a previous RP labeled $m_1$ and a previous CP labeled $m_2$, that is, $O_r(m_1, m', n-1, n) \geq f_r \times A(p_r^{m'}) \wedge O_{cp}(m_2, m', n-1, n) = f_{cp} \times A(p_r^{m'})$, the previous CP provides its label, that is, the label of RP $m'$ is replaced with the label $m_2$. In case RP $m'$ is overlapped by multiple RPs from the previous time step with deviating labels, the RP with the largest overlap, $O_r^{\max}$, provides its label. For this reason, the new RP in time step $t_n$ of Figure 1d which overlaps RP1 and RP2 from the previous time step $t_{n-1}$ is tracked as RP1.

Apart from the rain markers derived from RPs, all cold pools in $p_{cp}(i, j, n-1)$, which existed in the previous time step, receive origin markers and—if they do not have an active RP anymore—also their latest rain marker (see step (iv)). Accordingly, CP1, CP3 and CP4 in Figure 1d each receive two markers in time step $t_n$, representing the COM of the current RP and that of the initial RP from the previous time step $t_{n-1}$, respectively, whereas CP2 only receives its latest rain marker from time step $t_n$. However, due to its origin marker CP2 is still present at time step $t_n$ although it does not possess an active RP anymore.

### 2.4. Cold Pool Relationships

The association of RPs and cold pools allows the method to consistently label cold pools. Beyond this, through the spatial dependencies, we are also able to capture relationships between individual cold pools. By tracing which CP (gust front) triggered which RP and thus also a potential CP resulting from it, CoolDeTA reveals CP networks and their associated causal chains.

In order to describe the relationships between cold pools, we introduce the roles of "parent" and "child." A RP labeled $m'$ at time step $t_n$ which was not present in time step $t_{n-1}$, is a child of a CP, labeled $m$, from time step $t_{n-1}$, if the overlap $0 < O_{cp}(m, m', n-1, n) < f_{cp} \times A(p_r^{m'})$, that is, if the RP intersects the boundary of the CP, but is not fully located within its area. Conversely, CP $m$ becomes a parent of RP $m'$. If RP $m'$ is a child of CP $m$ and generates a new CP $m'$, the new CP adopts CP $m$ as a parent and becomes its child, as well. Since cold pools can trigger new rain events, they can become parents—both of RPs and the resulting cold pools—as well as children, whereas RPs can only be children.

In case a RP labeled $m'$ at time step $t_n$ which was not present in time step $t_{n-1}$ overlaps multiple cold pools from time step $t_{n-1}$, any CP $m$ with an overlap $0 < O_{cp}(m, m', n-1, n) < f_{cp} \times A(p_r^{m'})$ becomes a parent of RP $m'$ and thus also of potential cold pools resulting from it.

### 2.5. Cold Pool Dissipation

When the rain associated with a CP stops and thus no longer feeds the density current anymore, the surface boundary layer recovers until the CP is eventually dissipated. However, providing origin markers to cold pools from the previous time step $t_{n-1}$ in time step $t_n$, cold pools could potentially exist forever, even after their rain has stopped. To rule out infinite lifetime, we implemented a dissipation mechanism within step (iii) of CoolDeTA. To this end, we differentiate between "dissipating" cold pools that are still in the process of dissipating and "dissipated" cold pools that have completely dissipated. We define a CP, labeled $m$, as dissipating at time step $t_n$ if it has no active rain anymore, that is, $\sum_{i=1, j=1}^{N,N} \mathbf{1}((i, j) \in \{(i, j) \in [1, N] \times [1, N], p_r(i, j, n) = m\}) = 0$, and if parts of its latest RP $p_r^m$ from time step $t_n^{last}$ are classified as "no CP" in $p_{bin}(i, j, n)$, that is, $O_{seg}(m, 1, n^{last}, n) < 1.0 \times A(p_r^m)$ with

$$O_{seg}\left(m,1,n^{last},n\right) \equiv \sum_{i=1,j=1}^{N,N} \mathbf{1}((i,j) \in S_3), \tag{14}$$

where the set $S_3 \equiv \{(i,j) \in [1,N] \times [1,N], p_r(i,j,n^{last}) = m \wedge p_{bin}(i,j,n) = 1\}$. For each CP we count the number of time steps in which it is dissipating, $n_{dis}$. We consider a CP, labeled $m$, as dissipated if $n_{dis}^m \geq 3$ and $O_{seg}(m, 1, n^{last}, n) = 0$, that is, if it is dissipating for $\geq 3$ time steps ($=30$ min) and if its latest RP $p_r^m$ is completely classified as "no CP" in $p_{bin}(i,j,n)$. A CP which is dissipated at time step $t_n$ does not obtain markers anymore in $p_{mar}(i,j,n)$ and will thus not exist in the labeled CP field $p_{cp}(i,j,n)$.

## 2.6. Additional Algorithm Rules

The evolution of the populated fields of both cold pools, $p_{cp}$, and RPs, $p_r$, is highly complex. On the one hand, relatively cool air associated with remnants of dissipated cold pools might lead to the misclassification of grid cells as "potential CP" in the segmentation, $p_{bin}$, in time steps with very weak or suppressed convection and thus only weak down and updrafts. On the other hand, processes such as the merging of RPs or cold pools, as well as the formation of new RPs at CP gust fronts, complicate the tracking. In the following, we introduce several rules implemented in CoolDeTA to increase its robustness with respect to these and other cases.

### 2.6.1. Divergence Criterion

Cold pools are driven by atmospheric density gradients resulting from the evaporation of rain. Evaporative cooling causes the affected air to sink toward the ground and spread outwards, creating wind gusts. Consequently, cold pools are associated with a horizontally divergent flow in their center and a horizontally converging flow at their gust front. Thus, we only keep a "potential CP" region, labeled $m$ in $p_{seg}(i, j, n)$ at time step $t_n$, in the segmentation, $p_{bin}(i,j,n)$, if the mean divergence of the horizontal wind field, $\overline{\nabla \cdot \mathbf{u}(i,j,n)} \geq 0$ for grid cells $(i,j)$ in its patch interior $\text{int}\, p_{seg}^m$ and if $\overline{\nabla \cdot \mathbf{u}(i,j,n)} \leq 0$ for grid cells in its patch boundary $\partial p_{seg}^m$. The divergence is approximated using a central difference scheme. We define a grid cell $(i,j)$ as part of the interior $\text{int}\, p_{seg}^m$ of patch $m$ if all contiguous grid cells are contained in $p_{seg}^m$. A grid cell $(i,j)$ is part of the boundary $\partial p_{seg}^m$ of patch $m$ if it encompasses at least one contiguous grid cell contained in $p_{seg}^m$ and one contiguous grid cell not contained in $p_{seg}^m$. Since the "potential CP" region $m$ in $p_{seg}(i,j,n)$ can contain multiple CP instances and thus colliding gust fronts, we exclude a grid cell $(i,j)$ from $\text{int}\, p_{seg}^m$ if $\nabla \cdot \mathbf{u}(i,j,n) < \overline{\nabla \cdot \mathbf{u}(n)} - z_{95th}\, \sigma_d(n)$ where $\sigma_d$ represents the standard deviation of the horizontal wind field divergence and $z_{95th} = 1.645$.

Furthermore, new rain cells might be triggered at the boundary of the "potential CP" region $m$ compensating its convergent flow. Analogous to $\text{int}\, p_{seg}^m$ we thus exclude a grid cell $(i, j)$ from $\partial p_{seg}^m$ if $\nabla \cdot \mathbf{u}(i,j,n) > \overline{\nabla \cdot \mathbf{u}(n)} + z_{75th}\, \sigma_d(n) \vee I(i,j,n) \geq 1$ mm/h with $z_{75th} = 0.675$. We chose deviating factors $z_{95th}$ and $z_{75th}$ for the exclusion of grid cells from CP interior and boundary, as colliding CP gust fronts will be associated with relatively strong convergence in the patch interior, whereas new rain cells triggered at the boundary will be associated with relatively weak divergence compared to the domain mean. The values of both $z_{95th}$ and $z_{75th}$ were determined by experiment. The exclusion of potentially misleading grid cells, as well as the weak conditions in terms of $\text{int}\, p_{seg}^m$ not being convergent and $\partial p_{seg}^m$ not being divergent ensure that no correctly classified "potential CP" region is dropped.

Due to CP interactions such as collisions and the formation of new cold pools at gust fronts, a CP can loose its occupied area to another CP. To not supersede this process artificially, a CP, labeled $m$, at the previous time step, $t_{n-1}$, only obtains an origin marker $p_{mar}^m\left(i_{mar}^m, j_{mar}^m, n\right) = m$ if it still occupied the marker location $\left(i_{mar}^m, j_{mar}^m\right)$ at $t_{n-1}$, that is, if $p_{cp}\left(i_{mar}^m, j_{mar}^m, n-1\right) = m$.

Apart from CP interactions, where a CP gust front can be pushed toward the CP center, CP gust fronts spread outwards only, due to the horizontally diverging flow in the CP interior. As long as a CP, labeled $m$, is not dissipated and receives at least one marker at time step $t_n$, we thus adapt the segmentation $p_{bin}(i,j,n)$ in step (iii) so that all grid cells occupied by the CP at the previous time step $t_{n-1}$, $p_{cp}(i,j,n-1) = m$, are labeled as "potential CP" in $p_{bin}(i,j,n)$. Whether CP $m$ looses parts of its occupied area to an intersecting CP in time step $n$ depends on the topography of the elevation map, $\mathcal{I}_{top}(i,j,n)$, and is determined during the watershed flooding in step (iv).

**Table 1**
*Parameters and Conditions in the Proposed Algorithm*

| Parameter | Value/condition used | Description |
|---|---|---|
| $I_0$ | 2 mm hr$^{-1}$ | Surface rainfall intensity threshold |
| $A_0$ | 2 km$^2$ | Potential CP and RP area threshold |
| $f_r$ | 0.01 | Tracking factor for RPs overlapping RPs |
| $f_{cp}$ | 1 | Tracking factor for RPs overlapping CPs |
| | $0 < O_{cp} < 1.0 \times A(p_r)$ | Overlap condition for a new RP to become a child of a CP |
| $n_{dis}$ | 3 time steps (30 min) | Dissipation threshold |
| | $\overline{\nabla \cdot \mathbf{u}} \geq 0$ | Divergence criterion for potential CP interior |
| | $\overline{\nabla \cdot \mathbf{u}} \leq 0$ | Divergence criterion for potential CP boundary |
| $f_s$ | 0.75 | RP-potential CP overlap factor |
| | $1 \leq A(p_{cp})/A(p_r) \leq 3$ | New CP-RP area ratio |

*Note.* Parameters handling rain intensity threshold, minimal potential cold pool (CP) and rain patch (RP) areas, overlap conditions, and CP dissipation. Parameters and conditions below the horizontal line represent additional algorithm rules.

### 2.6.2. Additional Overlap Criteria

When new cold pools start to form, grid cells with strong negative temperature perturbations inside the downdraft region might already be classified as "potential CP" in $p_{bin}$. In order to detect fully developed cold pools only, a new RP, labeled $m$, detected at time step $n$, obtains a rain marker at time step $n$ only if $O_{seg}(m,1,n,n) \geq f_s \times A(p_r^m)$ with $f_s = 0.75$.

The final check is applied to the labeled CP field $p_{cp}$ derived based on the watershed flooding in step (iv). By comparing the area, $A_{cp}$, of newly formed cold pools with the area, $A_r$, of their RPs, we get an indication of erroneous labeling. Erroneous labeling can result from cool air remnants being classified as "potential CP" during time steps with very weak or suppressed convection. In this case, markers of isolated rain events could flood areas $A_{cp} \gg A_r$. Another potential source of erroneous labeling is new RPs, triggered at CP gust fronts: Since not each of these RPs generates an independent child CP which is separated from its parent CP by their colliding gust fronts, child cold pools may flood into the area of their parent cold pools and vice versa. Thus, an independent and fully developed CP should, on the one hand, at least feature the size of its RP. On the other, a new CP cannot occupy an area substantially larger than its RP, given the time step of $\Delta t = 10$ min. To intervene only in erroneous cases, we keep a new CP, labeled $m$, if $1 \leq A(p_{cp}^m)/A(p_r^m) \leq 3$. Otherwise, we either drop it, that is, $p_{cp}^m = 0$, if it has no other CP in its neighborhood, or we assign the label of another CP, labeled $m'$, that is, $p_{cp}^m = m'$, otherwise. In the latter case, the CP label $m'$ is determined based on its number of parents, $k_p$: if $k_p = 0$, it is either the label of the CP which overlaps the largest area of RP $m$ (if $A(p_{cp}^m)/A(p_r^m) < 1$), or the label of the CP which lost the largest area with respect to the previous time step $t_{n-1}$ to CP $m$ (if $A(p_{cp}^m)/A(p_r^m) > 3$). If $k_p = 1$, CP $m$ is simply assigned the parent label, $m_p$, that is, $m' = m_p$. The final case, $k_p > 1$, is analogous to the first but based on the parents of CP $m$.

Table 1 summarizes the parameters and conditions used in the proposed algorithm. Recommendations for the application of CoolDeTA to other simulation setups are provided in Text S1 in Supporting Information S1.

### 2.7. Analysis of the Simulation Data

We consider the first 3 days (432 time steps) of the simulations as spin-up period. The spin-up period is chosen so that the total domain rainfall per time step, $R$, has visibly reached a steady state in both the RCE and diurnal cycle setups (Figures 2e and 2f). Further time series are visualized in Figure A1. To be able to track the full life cycle of cold pools and their causal relationships, the onset of convection is an ideal starting point for the analysis with CoolDeTA. Apart from rce0K, where $R$ is essentially constant (Figure 2f), we identify the corresponding starting time step based on the standard deviation of the virtual temperature, $\sigma_{Tv}$, and its time derivative, $\dot{\sigma}_{Tv}$. Based on their 1 hr running average, we find $\sigma_{Tv} \geq 0.15$ K and $\dot{\sigma}_{Tv} \geq 0.02$ K s$^{-1}$ as good indicators for the onset of
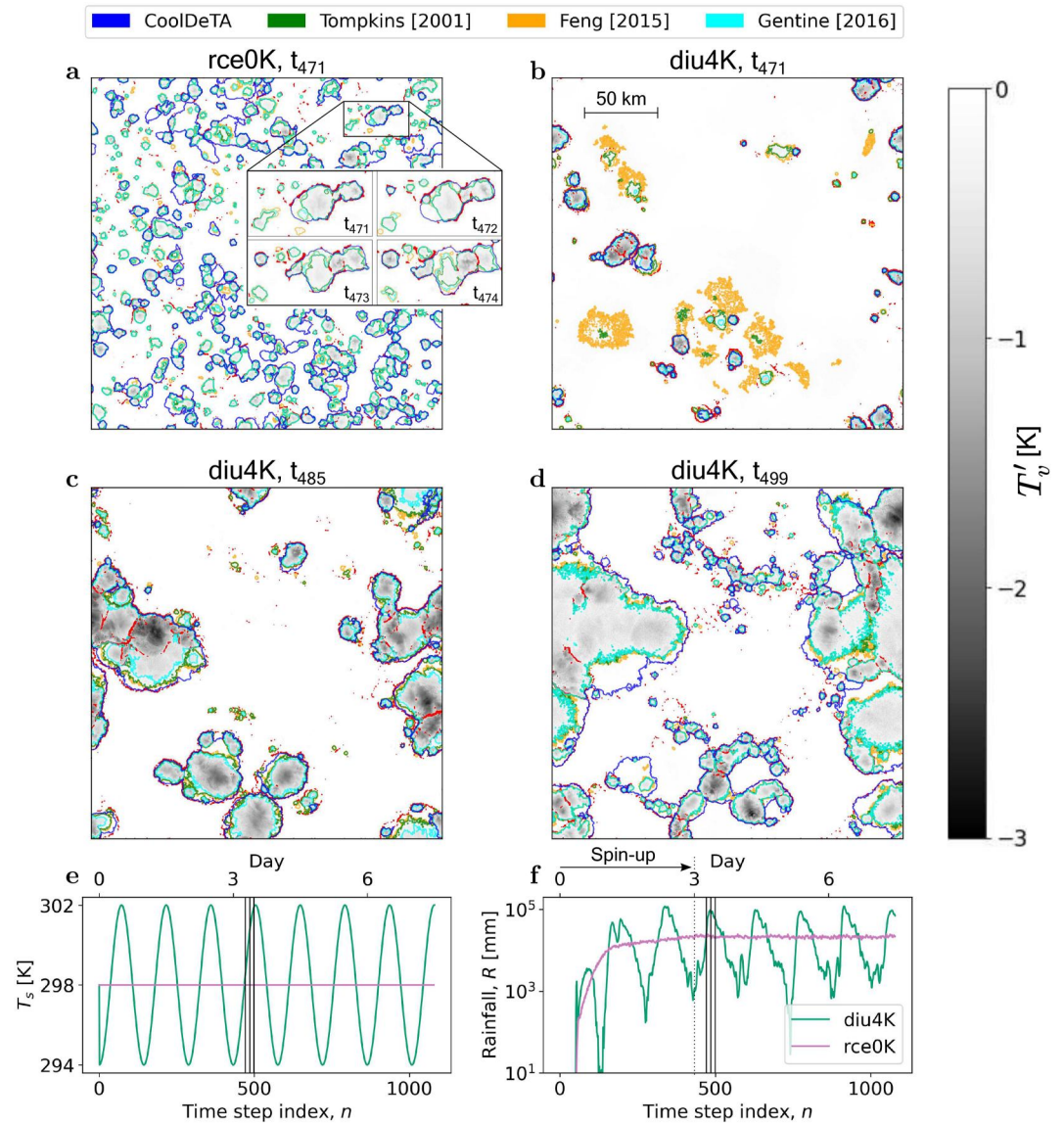
**Figure 2.** *Comparison of cold pool (CP) detection methods.* (a) Time step index 471 of rce0K, showing near-surface virtual temperature perturbations, $T'_v$, with superimposed dynamical gust front, that is, $w > \overline{w} + 2\sigma_w$ (red scatter); The superimposed colors represent contours of CP patches $\geq A_0$ based on different CP detection methods. The enlarged inset shows a magnification for a subregion highlighting a sequence of four time steps. (b) Analogous to (a) but for diu4K. (c) Analogous to (b) but for time step index 485, corresponding to the peak of diurnal rainfall. (d) Analogous to (b) but for time step index 499. (e) Surface temperature, $T_s$, for the two simulation setups visualized in (a)–(d); Solid vertical lines within the plot denote the time steps depicted in (a)–(d). (f) Analogous to (e) but for total domain rainfall per time step, $R$; The chosen spin-up period of 3 days (432 time steps) is indicated by the vertical dotted line and the horizontal arrow above the panel.

convection in our simulations. In the diurnal cycle setups, as well as the rce0K setup with wind shear, we choose the starting time step on day four, $t^{start}$, so that the two conditions are fulfilled at least for the next hour (6 time steps). For rce0K we simply start the analysis with CoolDeTA at the first time step of day four, at $t_{432}$.

In the diurnal cycle setups, $R$ oscillates nearly periodically (Figure 2f). Especially in the two diu4K simulation setups, convective activity is largely suppressed during time steps with low $T_s$. Since the k-means algorithm always clusters the horizontal domain into two groups, namely "potential CP" and "no CP," the segmentation cannot work properly in the absence of cold pools. In such cases, also the additional rules may be insufficient. For this purpose, we stop the CoolDeTA analysis if $R_{px} < 0.001$ mm for more than 30 min (3 time steps) or as soon as $\sigma_{Tv} < 0.15\,\mathrm{K} \wedge R_{px} < 0.0015$ mm, with $R_{px} = \frac{R}{N \times N}$. Analogous to the definition of the start time step, we apply the

conditions to the 1 hr running averages of $\sigma_{T_v}$ and $R_{px}$. The stop conditions affect only the two diu4K setups, which we stop at the end of each day and re-start on the following day based on the defined start conditions.

## 3. Results

In the following, we present results of our CP detection and tracking method. We first compare our method to previous ones (Feng et al., 2015; Gentine et al., 2016; Tompkins, 2001a). We then employ our current method to discuss the effect of the simulation setup, that is, wind shear or diurnal cycles, on CP expansion. Third, we make use of our method to analyze the causal relationships between cold pools and their effect on CP rainfall. Based on our findings we then propose a simple model to capture CP spreading-triggering dynamics. We finish the chapter by visualizing the family trees of two cold pools.

### 3.1. Comparison to Other Methods

Several previous CP detection methods have exploited thermodynamic CP features (Feng et al., 2015; Gentine et al., 2016; Tompkins, 2001a). While Feng et al. (2015) and Tompkins (2001a) use buoyancy thresholds of $-0.005\ \mathrm{m^2\,s^{-1}}$ and $-0.003\ \mathrm{m^2\,s^{-1}}$, respectively, for the detection, Gentine et al. (2016) apply a k-means algorithm to the horizontal virtual temperature field of the first model level to distinguish cold pools from their environment. For an RCE-like setup, rce0K (Figure 2a), as well as a diurnal cycle case, diu4K (Figures 2b–2d), we compare CoolDeTA to such existing methods. The time steps shown are highlighted in the time series (Figures 2e and 2f) as thin vertical lines. Inspecting the different examples (Figures 2a–2d), it is apparent that the CP dynamics in rce0K occurs at relatively small scales $\mathcal{O}(10\,\mathrm{km})$, whereas the dynamics in diu4K ranges from scattered small scale events during the early stages of the diurnal cycle to elaborate mesoscale features $\mathcal{O}(100\,\mathrm{km})$ at the late stages of the diurnal cycle.

A primary goal of CoolDeTA is to detect both thermodynamic and dynamical features of cold pools. To obtain a reference for the latter, we highlight grid cells of pronounced positive vertical velocity in each panel, typically indicating locations of dynamical CP gust fronts (red pixels in Figures 2a–2d). Inspecting the plots, it is apparent that on several occasions, the detection results by all four methods nearly coincide, especially during early stages of CP expansion. However, important differences can also be seen, notably within the inset to panel (a), where CoolDeTA is able to capture the locations of the dynamic gust front (red pixels) reliably, whereas methods based on thermodynamics typically indicate gust fronts that are located further inward, seen for example, by comparing the lines of different colors in the inset. As is visible on the left side of the larger CP patch in $t_{473}$ of the inset, CoolDeTA keeps dissipating CP instances as long as they feature a partly active dynamic gust front.

For the early stage of the diurnal cycle (Figure 2b) the agreement of CoolDeTA, Gentine et al. (2016) and Tompkins (2001a) is arguably good. However, as cold pools spread in the course of the day, the dynamic gust fronts tend to detach from the thermodynamic ones, leading to an increased discrepancy between the methods. It can be visually verified that CoolDeTA is still capable of identifying most regions of strong vertical velocity.

Apart from the common minimum area $A_0$ which we apply in Figures 2a–2d for CP patches of all methods, CoolDeTA applies the threshold $A_0$ also to RPs associated with individual cold pools and conducts further checks to confirm the existence of a CP (see Section 2.6). Consequently, the minimum CP area detected by CoolDeTA tends to be larger than for other methods. As a case in point, some of the smaller patches identified by the other methods either suddenly disappear or do not feature a spreading dynamic gust front (Figure 2a, enlarged area), suggesting that CoolDeTA manages to successfully detect only robust cold pools.

### 3.2. Effect of the Simulation Setup on Cold Pool Expansion

We now employ CoolDeTA to investigate CP temporal expansion and how it depends on the choice of simulation setup. Cold pool expansion is a basic dynamical CP feature, which we here capture as the change in CP area, $A_{cp}$, over time, on the one hand for cold pools without offspring during their lifetime (Figure 3a) and on the other for cold pools with offspring (Figure 3b). To not distort the results with cold pools which shrink due to CP interactions, we only consider a CP as long as its area is expanding and thus $A_{cp}(n) > A_{cp}(n-1)$.

Conditioning on cold pools without children (Figure 3a), it is noticeable that within the log-log representation, all curves start out following a modest increase of CP area with time. To make contact with previous works suggesting power law dependencies of CP radius on time (Meyer & Haerter, 2020; Rooney, 2015), we indicate a
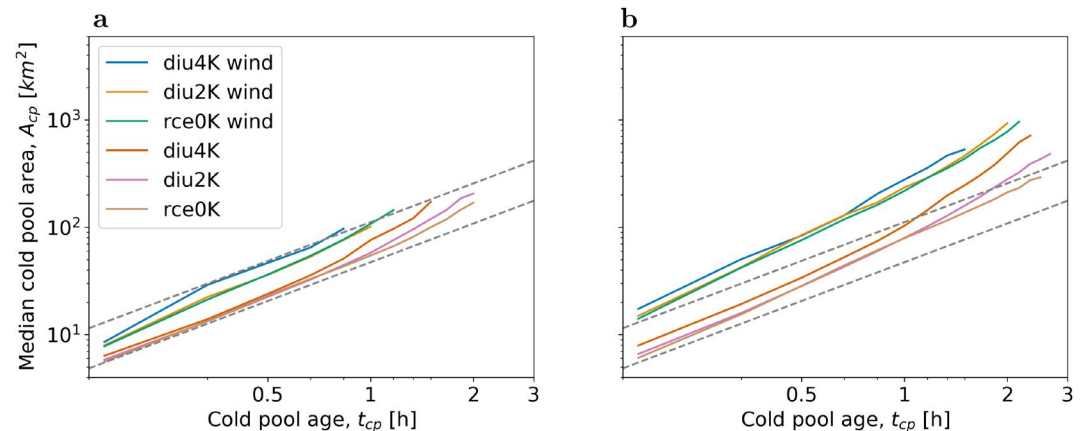
**Figure 3.** *Cold pool (CP) expansion for various simulation setups.* (a) Median CP area, $A_{cp}$, versus CP age, $t_{cp}$ for cold pools without any children during their lifetime. Colors represent different simulation setups; Dashed gray lines indicate areas $\propto t_{cp}^{1.2}$; Only CP ages with $\geq 50$ cold pools are depicted. (b) Analogous to (a) but for cold pools with children during their lifetime. Note the logarithmic vertical axis scaling.

dependency of $A_{cp} \sim t^{\alpha}$ with $\alpha = 1.2$ (dashed gray lines). After approximately 1 hr of expansion, the curves of setups without wind shear somewhat increase in exponent, thus departing from a power-law dependency. We attribute this to the higher density of cold pools for setups without wind shear, which increases the probability for cold pools to merge with others. The departure is even stronger for diu2K and diu4K where cold pools increasingly cluster within thermodynamically favorable subregions.

The curves for cold pools with children exhibit an increased power-law exponent already at earlier stages (Figure 3b). In comparison to the curves of cold pools without children, the exponent seems to be increasing more strongly, especially during the later stages of expansion. We interpret this strong increase as a consequence of enhanced triggering near CP edges for cold pools with children, whereby the original CP can merge with additional offspring excited by it—thus developing into a mesoscale CP. Apart from this, the diurnal cycle again appears to promote rapid CP expansion, especially during the later stages.

In both Figures 3a and 3b, cold pools under wind shear conditions have larger areas from the outset. A possible reason is the overlap that a RP is required to have with a "potential CP" region in order to obtain a marker and thus form an associated CP: when wind shear is present, rain might reach beyond the emerging CP and limit their overlap.

### 3.3. Effect of Cold Pool Relationships on Cold Pool Rainfall

Given these findings on CP dynamics we now investigate: (a) how the total accumulated rainfall corresponding to any given CP relates to its capacity to trigger offspring (Figure 4a); and (b) how this rainfall relates to the number of parents of a given CP (Figure 4b). For this purpose, we define the total accumulated rainfall of a CP labeled $m$ as

$$R_{cp}(m) \equiv \sum_{n} \sum_{i,j=1}^{N} \Delta t \, I(i,j,n) \, \delta\big(p_r(i,j,n),m\big). \tag{15}$$

Overall, the majority of cold pools do not give rise to offspring, with approximately 50% of cold pools without detectable children (Figure 5a). However, as we inspect the probability distributions, it is found that, moving from rce0K to diu4K, the distribution function is close to an exponential for rce0K (dotted line), whereas it is well approximated by a power-law $\propto k^{-1.5}$ in the case of diu4K (dashed line). This qualitative finding is also mirrored by the wind shear cases. A power-law distribution hints at a "rich-gets-richer" feedback, where cold pools that have already grown, may be more likely to grow further.

At the same time, it is clear that for large numbers of CP children, associated RPs tend to have much greater total rainfall rates (Figure 4a). These rates range from a factor of two for rce0K to more than an order of magnitude for
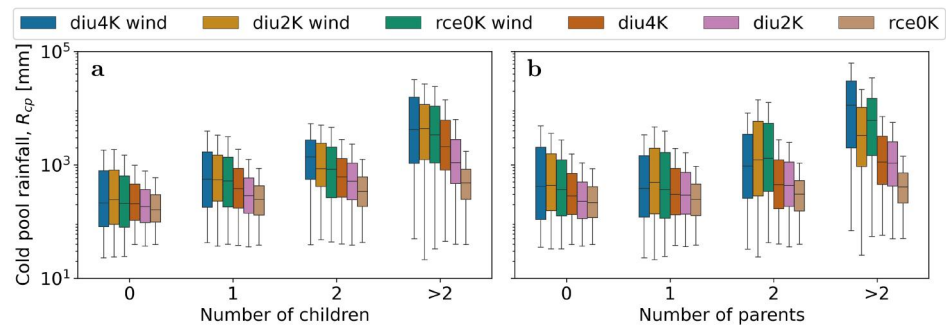
**Figure 4.** *Total cold pool (CP) rainfall*. (a) Total accumulated rainfall associated with cold pools, $R_{cp}$, depending on their number of children for different simulation setups. Colored bars represent the interquartile ranges $IQR = Q3 - Q1$ of the simulation setups, with the first quartile $Q1$ and the third quartile $Q3$, along with the corresponding median (horizontal dash). Whiskers range from $Q1 - 1.5 \times IQR$ (minimum) to $Q3 + 1.5 \times IQR$ (maximum); Outliers w.r.t. this range are not visualized. (b) Analogous to (a) but for the number of parents. Note the logarithmic vertical axis scaling.

diu4K. Notable differences also exist between the different setups: diurnal $T_s$-forcing and wind shear seem to promote larger numbers of children and generally come with greater parent rainfall.

Apart from cold pools with one parent, the relation between parental rainfall and the number of children is almost perfectly mirrored by the number of parents (Figure 4b): larger total rainfall typically implies a larger number of parents. Also, for setups with $T_s$-diurnal cycles or wind shear, the role of parent cold pools contributing to rain cell formation increases. Yet, the rainfall rates of cold pools with one parent hardly differ from those of cold pools without a parent. We attribute this to the higher risk that a single-parent CP has of getting merged into the parent CP: whereas multiple parents are associated with CP collisions, which interrupt the expansion of the parents, single-parents might still spread and thus merge with excited offspring.

### 3.4. Simple Cold Pool Offspring Model

To capture CP spreading-triggering dynamics, we suggest a simple model (Figure 5b) where cold pools primarily grow by peripheral spawning of new rain cells, thus child cold pools, that then help the CP expand further. The model is hence different from those in the recent literature (Böing, 2016; Haerter, 2019; Haerter et al., 2019; Nissen & Haerter, 2021) in that the dynamics described previously was inward-directed, such that existing cold pools could excite new convection by collisions among two or three gust fronts. As a result, a given population of cold pools could topologically only cause new convection within the interior of an ensemble of spreading cold pools. In the current model we describe an outward-directed dynamics, where a given evolving mesoscale convective system, embedded in an unstable environment, can give rise to cascading additional convective activity at its periphery—thus leading to an outward expansion of the population of convective cells.

In our model each CP is initialized with an area $a$ of $a = a_0 \equiv 6 \text{ km}^2$ representing the median area of all cold pools from Figure 3 at $t_{cp} = 10 \text{ min}$, and circumference $l = 2(\pi a_0)^{1/2} \approx 8.7 \text{ km}$. This circumference is subdivided into segments, each of length $l_0 \equiv 1.6 \text{ km}$, taken to represent the approximate diameter of a new rain cell of area $A_0$. Within each segment and time interval $\Delta\tau \sim 20 \text{ min}$, chosen to represent the timescale of a convective rain cell, the model allows for a new rain cell to be spawned with a small probability $0 < p \ll 1$. The constant $l_0 \ll l$ is thereby assumed to be substantially less than the total circumference $l$, such that even small cold pools can have several segments. Any new rain cell will add an area increment $a_0$ to the existing CP area $a$. In addition, as the CP area increases, its total number of segments $n_s = l/l_0$ will also increase due to the larger circumference. Thus, larger cold pools will typically have a larger number of successful spawning events.

As mentioned, the dynamics proceeds in discrete time steps $\Delta\tau$, during which each segment will be able to experience spawning at probability $p$. If none of the segments give rise to spawning during the time step, the CP is considered terminated and no further expansion is possible. If a number $m_s > 0$ of segments give rise to spawning, the CP area is incremented by $\Delta a \equiv m_s a_0$ and the expansion proceeds iteratively (tree diagram in Figure 5b). At each fork of the tree diagram, at probability $1 - P(a_i)$ the CP expansion is terminated. $1 - P(a_i)$ represents the probability that none of the segments spawn a new rain cell, that is, for $n_s = l/l_0$ segments and spawning probability $p$,
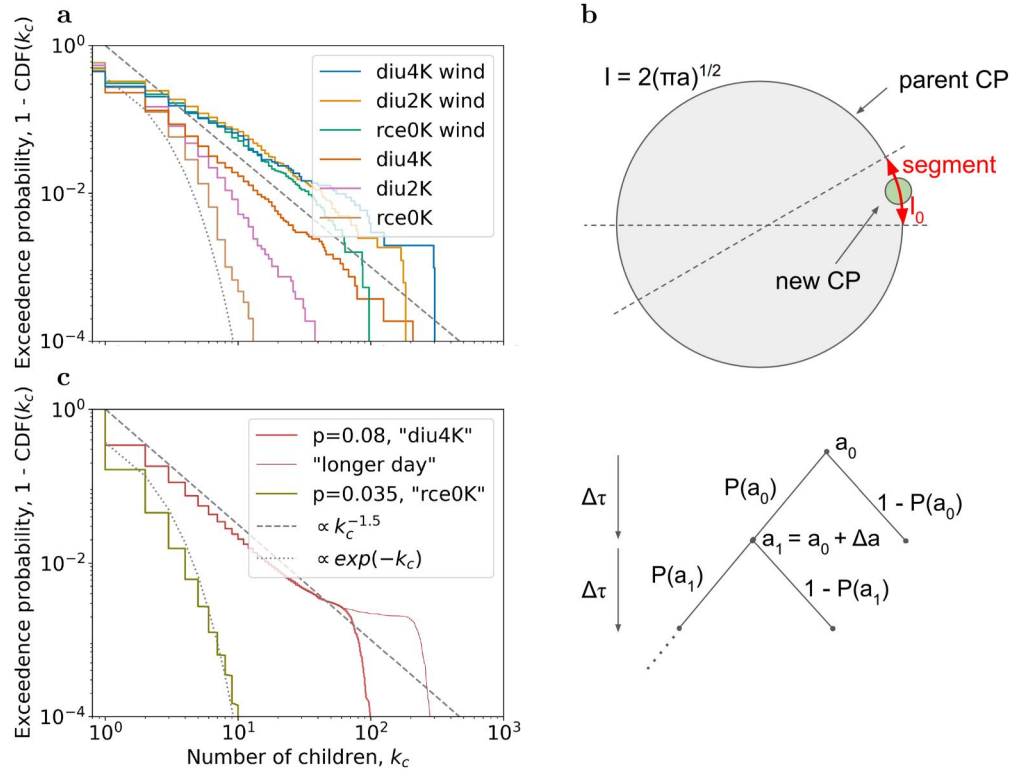
**Figure 5.** *Children probability distribution.* (a) Exceedance probability for a cold pool (CP) to have a certain number of children; Colors indicate different simulation setups. (b) Simple model for CP expansion and spawning of new rain cells. (c) Analogous to (a) but for our simple model.

$$1 - P(a_i) = (1 - p)^{n_s}. \tag{16}$$

Conversely, the probability for continued CP expansion is

$$P(a_i) = 1 - (1 - p)^{n_s}. \tag{17}$$

As cold pools expand, $n_s$ continues to increase and $P(a_i) \to 1$—thus allowing larger cold pools to experience a larger probability of further expansion.

Simulating the model using a total of 30 time steps $\Delta\tau$ and $10^5$ CP realizations allows us to mimic the statistics found for the cloud-resolving simulations (Figure 5c): whereas small values of spawning probability, for example, $p = 0.035$, yield a nearly exponential distribution of the number of children, a larger value of $p = 0.08$ gives rise to approximate power-law decay. Notably, as cold pools grow to large areas, spawning will be essentially guaranteed at subsequent time steps—thus giving rise to run-away expansion. This is verified by allowing for a greater number of 60 time steps (thin brown curve), where a fraction of "successful" cold pools accumulates ever larger numbers of children. Using this model, our interpretation of the CP dynamics is twofold: spawning probability $p$ may be determined by the convective instability surrounding a given parent CP, and different $p$ can yield very different offspring distributions.

For the RCE case, $p$ may be overall modest, as the atmosphere tends to be in a marginally stable state close to the moist adiabat. For the diu cases the situation is quite different, where activity is invigorated by the destabilizing effect of the increased diurnal surface heating—thus allowing for larger values of $p$, yielding potentially long-tailed offspring distributions. A complete run-away effect is likely prevented by the finite duration of the diurnal heating. It may thus be interesting to revisit the statistics (Figure 5a) for an artificially lengthened diurnal cycle, say, corresponding to a 48 hr-day.
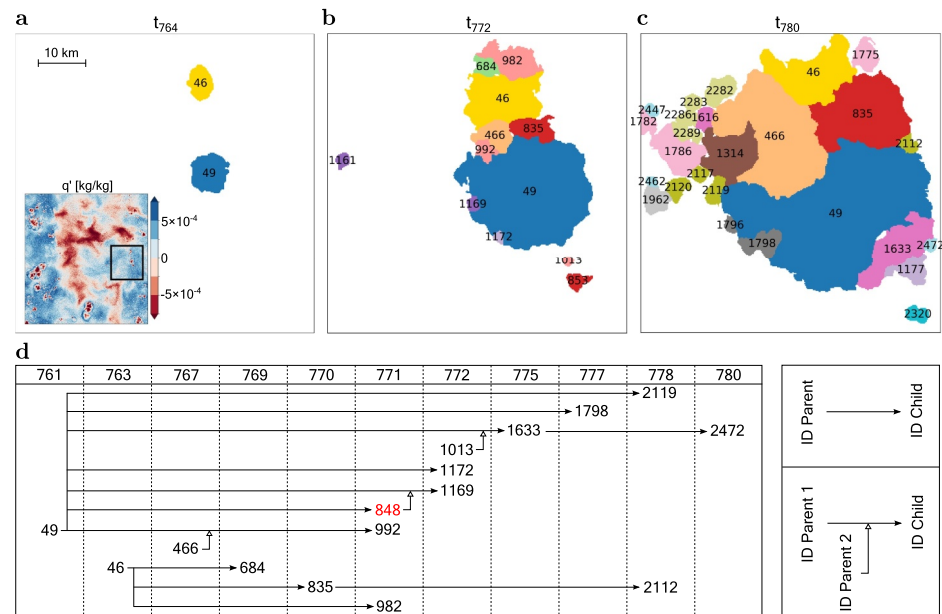
**Figure 6.** *Cascades of tracked cold pools.* (a) Labeled cold pools for a subsection of the full simulation domain at time step index 764 of diu4K. The inset plot depicts full domain specific humidity anomalies at the same time step; The black square indicates the subsection of interest. (b) Analogous to (a) but without inset plot and for time step index 772. (c) Analogous to (b) but for time step index 780. (d) Family tree for CP "46" and CP "49." The tree features the IDs of all related cold pools which are present in (a)–(c); Columns represent the time steps indices of CP formation; Filled arrows point from the IDs of CP parents to the IDs of CP offspring; Empty arrows denote additional parents; Red text indicates the IDs of cold pools that are not present in the depicted time steps of (a)–(c), but which contributed to the related cold pools.

### 3.5. Cold Pool Family Trees

Based on a subsection of the full simulation domain we now visualize an example of the temporal evolution of a labeled CP field from diu4K (Figures 6a–6c), along with the corresponding family trees of the two initial cold pools (Figure 6d). The subsection of interest is indicated as black square in the inset which shows the specific humidity anomalies, $q'$, for the full simulation domain at $t_{764}$. Like most of the other regions in which cold pools start forming (dark red blobs with large gradients), the subsection of interest features a moisture-rich surface layer. Here, the two initial cold pools depicted in Figure 6a, CP46 (yellow patch) and CP49 (blue patch), formed at $t_{763}$ and $t_{761}$, respectively (Figure 6d).

In Figure 6b, 80 min later, the two cold pools have expanded and several offspring have been generated. However, CP848 (red ID in Figure 6d), one of the first children of CP49 which was triggered at the freely expanding gust front, could only exist for one time step until it was merged by its parent. Although CP848 contributed to the formation of the new CP CP1169 (see Figure 6d), it is thus not present anymore at $t_{772}$. Comparing Figures 6b and 6c, it is apparent that also the other cold pools which were triggered at freely expanding gust fronts, namely CP684, CP982, CP1169 and CP1172, shared the same fate. Only the two cold pools CP466 and CP835 in the collision zone of CP46 and CP49 are still present at $t_{780}$ and managed to expand over the former area of CP46 and CP49. Since it is not evident from Figure 6b we note that CP466 and CP835 formed before the collision of CP46 and CP49, otherwise both CP46 and CP49 would be their parents.

## 4. Discussion

Cold pools have been implicated in a large range of cloud field properties, including the formation of MCS (Haerter et al., 2020; Houze Jr, 2004; Jensen et al., 2021; Kain & Fritsch, 1992; Schumacher & Rasmussen, 2020), the evolution of squall lines (M. D. Parker, 2008; D. J. Parker & Diop-Kane, 2017; Rotunno et al., 1988; Weisman et al., 1988), the organization of trade-wind cumuli (Dauhut et al., 2022; Seifert & Heus, 2013; Vogel et al., 2021; Zuidema et al., 2012) and stratocumuli (Glassmeier & Feingold, 2017), and idealized studies on convective self-aggregation (Haerter, 2019; Haerter et al., 2019; Jeevanjee & Romps, 2013; Muller & Bony, 2015; Nissen &

Haerter, 2021) and the convective diurnal cycle (Böing, 2016; Böing et al., 2012; Haerter et al., 2019; Haerter & Schlemmer, 2018; Schlemmer & Hohenegger, 2016). In all of these systems, the exact role of cold pools in giving structure to the cloud and precipitation fields is still not well-understood. Yet, it is becoming more and more obvious, especially as higher model grid resolutions are approached, that cold pools do play a potentially crucial role in providing an interaction mechanism between rain cells in space and time.

A number of CP detection methods exist, which often make use of a buoyancy threshold (Feng et al., 2015; Gentine et al., 2016; Tompkins, 2001a) or a detection of the gust front through a criterion on the dynamics (Fournier & Haerter, 2019; Henneberg et al., 2020). We have presented a method, CoolDeTA, that combines thermodynamic and dynamic CP signatures. CoolDeTA detects and tracks cold pools as well as the location of their dynamical gust front. On this basis, CoolDeTA is able to identify relationships between cold pools and subsequent rain cells and thus CP "offspring." CoolDeTA therefore can be used to analyze interactions between one or more cold pools and the ones emerging in the surroundings.

Like most other methods (Drager & van den Heever, 2017; Feng et al., 2015; Fournier & Haerter, 2019; Gentine et al., 2016; Henneberg et al., 2020), CoolDeTA cannot fully refrain from setting some threshold values. In the present work we apply a surface rainfall intensity threshold, $I_0 = 2$ mm hr$^{-1}$, to delineate spatially contiguous RPs, and a corresponding area threshold, $A_0 = 2$ km$^2$, for RPs and potential CP regions. Thresholds are generally case dependent and thus limit a method's scope of application. Although particularly the CP labeling of RCE frameworks would occasionally benefit from an even smaller $A_0$, we however claim that both of our thresholds serve only as a preselection of robust rain and CP patches and do not impact the final CP contours: Since we provide each new RP above $A_0$ one marker in its downdraft COM, all of these RPs have the same chance of forming a CP. The CP contour, and whether a RP forms a CP at all, depends on the result of the k-means algorithm. A sufficiently large RP and an associated marker are accordingly necessary, rather than sufficient, conditions with respect to the formation of a CP. Also the additional conditions summarized in Table 1 serve solely the purpose of optimizing the preselection in order to increase the robustness of the CP labeling. To verify this, we deactivated the dissipation threshold, $n_{dis}$, as well as all additional conditions and reproduced Figures 2 and 6 based on this parameter setup in Figures S1 and S2 in Supporting Information S1, respectively: whereas the deactivation causes essentially no change in the diu4K plots (Figures S1b–S1d in Supporting Information S1), several of the smaller patches in the rce0K plot which were identified as cold pools by other methods but not by CoolDeTA before, get classified as cold pools without the additional conditions (Figure S1a in Supporting Information S1). The same applies to the CP relationships (Figure S2 in Supporting Information S1), which are identical under deactivated conditions, except for a few additional CPs that disappear shortly after being detected. With respect to dissipating cold pools, the omission of $n_{dis}$ can even positively affect the results in the rce0K plot since some of the weak cold pools dissipate in less than 30 min.

We see the benefit of CoolDeTA in allowing for a systematic disentanglement of the processes leading up to organized convective cloud and rainfall fields in the systems mentioned above. We have shown that the method can build "cold pool family trees" in RCE and diurnal cycle frameworks, with and without wind shear. These family trees lay the groundwork for a more elaborate analysis of the evolution of convective organization throughout longer simulated periods. Moreover, as CoolDeTA considers dynamic CP signatures, the identified CP boundaries align with the cloud patterns associated with cold pools in satellite images or corresponding simulation output. CoolDeTA thus offers a systematic and objective ground truth labeling for artificial intelligence methods that detect cold pools from simulated cloud fields and that potentially pave the way for future satellite-based CP observations (Hoeller et al., 2024).

As an extension of the current method, it could be interesting to devise a moisture tracking for simulated convective cloud fields. Especially for hysteretic phenomena, such as convective self-aggregation, the long-term memory of mesoscale moisture anomalies could play a critical role in maintaining a persistent atmospheric circulation (Bretherton et al., 2005; Jensen et al., 2021; Muller & Bony, 2015). In RCE, horizontal moisture variations were suggested to give rise to imbalances in thermal radiative emission, which would then drive a persistent circulation and resultant upstream moisture transport—reinforcing the initial moisture imbalance (Emanuel et al., 2014). Recent numerical work extends this notion to diurnal cycle simulations, where MCS emerge and cause abrupt transitions to such a moisture segregated state (Jensen et al., 2021). Importantly, the mesoscale convective system dynamics was found to occur only at sufficiently high horizontal model grid

resolution, where CP interactions could be resolved sufficiently. Deciphering the complex moisture and precipitation dynamics requires a tool, such as CoolDeTA, where causal relations can be objectively mapped.

## 5. Conclusion

We have devised a CP detection and tracking method, CoolDeTA, which uses a combination of thermodynamic and dynamical variables to track cold pools from their initial forcing by rainfall evaporation until their dissipation stage. The method is shown to function well in a range of contexts, involving RCE and diurnal cycle experiments, as well as simulations with or without wind shear. In comparison to existing buoyancy-focused CP detection methods, CoolDeTA offers the additional benefit of determining the actual dynamical gust front, which is often displaced several kilometers from the thermodynamic boundary. In this way, the method is particularly well suited for the tracking of CP families, where the entire causal cascade of cold pools and their "offspring" can be identified. Such offspring tracking and attribution to parent cold pools would be hampered if only the thermo-dynamic gust front were available—many new rain cells are triggered at the dynamic gust front.

## Appendix A: Supplementary Data

The following appendix presents a figure with supplementary data for the assessment of the spin-up periods in the conducted simulations (Figure A1).
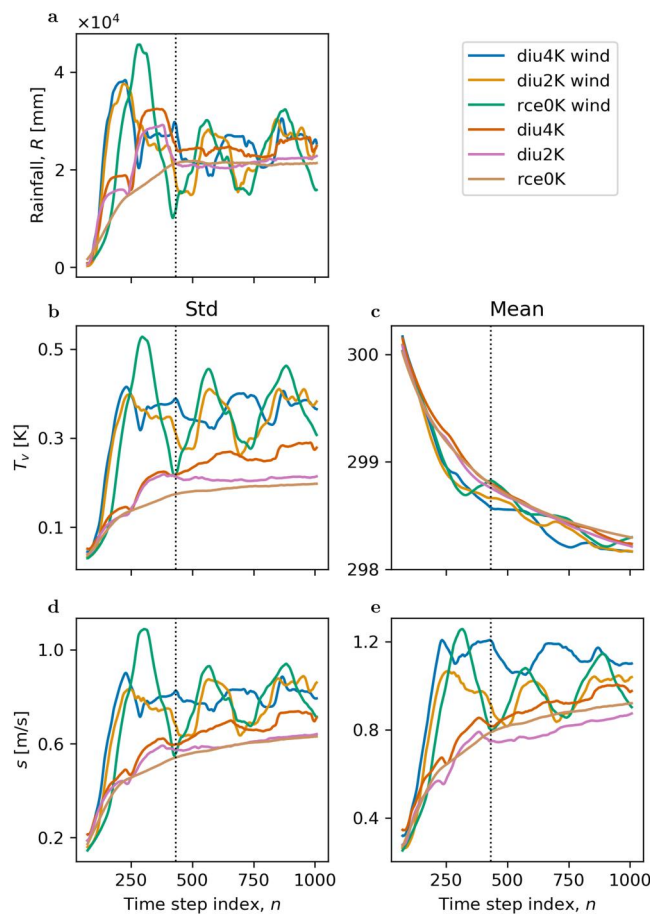


**Figure A1.** *Definition of spin-up time steps*. (a) Total domain rainfall per time step as running average of 1 day (144 time steps) for different simulation setups; The dotted black line indicates the chosen spin-up period of 3 days (432 time steps). (b) Analogous to (a) but for domain-standard deviation of virtual temperature, $T_v$. (c) Analogous to (b) but for domain-mean. (d) Analogous to (b) but for horizontal wind speed, $s$. (e) Analogous to (d) but for domain-mean.

## Data Availability Statement

Figures were made with Matplotlib version 3.5.2 (Caswell et al., 2022; Hunter, 2007) and seaborn version 0.12.2 (Waskom, 2021). The code for the simple CP offspring model is licensed under MIT and published on GitHub https://github.com/Shakiro7/coldPool-detection-and-tracking. The CP detection and tracking algorithm (Cool-DeTA) was used in version 1.0 and is licensed under Creative Commons Attribution 4.0 International (Hoeller, 2023). CoolDeTA makes use of a watershed algorithm (van der Walt et al., 2014) and a k-means algorithm (Pedregosa et al., 2011). The simulations are run with the cloud-resolving three-dimensional atmosphere simulator System for Atmospheric Modeling (SAM) (Khairoutdinov & Randall, 2003), version 6.11.

## References

Barnes, G. M., & Garstang, M. (1982). Subcloud layer energetics of precipitating convection. *Monthly Weather Review*, *110*(2), 102–117. https://doi.org/10.1175/1520-0493(1982)110<0102:sleopc>2.0.co;2

Böing, S. J. (2016). An object-based model for convective cold pool dynamics. *Mathematics of Climate and Weather Forecasting*, *2*(1), 43–60. https://doi.org/10.1515/mcwf-2016-0003

Böing, S. J., Jonker, H. J., Siebesma, A. P., & Grabowski, W. W. (2012). Influence of the subcloud layer on the development of a deep convective ensemble. *Journal of the Atmospheric Sciences*, *69*(9), 2682–2698. https://doi.org/10.1175/JAS-D-11-0317.1

Bretherton, C. S., Blossey, P. N., & Khairoutdinov, M. (2005). An energy-balance analysis of deep convective self-aggregation above uniform SST. *Journal of the Atmospheric Sciences*, *62*(12), 4273–4292. https://doi.org/10.1175/jas3614.1

Bryan, G. H., Wyngaard, J. C., & Fritsch, J. M. (2003). Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, *131*(10), 2394–2416. https://doi.org/10.1175/1520-0493(2003)131<2394:rrftso>2.0.co;2

Caswell, T. A., Droettboom, M., Lee, A., de Andrade, E. S., Hoffmann, T., Klymak, J., et al. (2022). matplotlib/matplotlib: Release (version 3.5.2) [Software]. Zenodo. https://doi.org/10.5281/zenodo.6513224

Chandra, A. S., Zuidema, P., Krueger, S., Kochanski, A., de Szoeke, S. P., & Zhang, J. (2018). Moisture distributions in tropical cold pools from equatorial Indian Ocean observations and cloud-resolving simulations. *Journal of Geophysical Research: Atmospheres*, *123*(20), 11–445. https://doi.org/10.1029/2018jd028634

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., et al. (2006). The formulation and atmospheric simulation of the Community Atmosphere Model version 3 (CAM3). *Journal of Climate*, *19*(11), 2144–2161. https://doi.org/10.1175/JCLI3760.1

Dauhut, T., Couvreux, F., Bouniol, D., Beucher, F., Volkmer, L., Pörtge, V., et al. (2022). Flower trade-wind clouds are shallow mesoscale convective systems. *Quarterly Journal of the Royal Meteorological Society*, *149*(750), 325–347. https://doi.org/10.1002/qj.4409

de Szoeke, S. P., Skyllingstad, E. D., Zuidema, P., & Chandra, A. S. (2017). Cold pools and their influence on the tropical marine boundary layer. *Journal of the Atmospheric Sciences*, *74*(4), 1149–1168. https://doi.org/10.1175/jas-d-16-0264.1

Drager, A. J., Grant, L. D., & van den Heever, S. C. (2020). Cold pool responses to changes in soil moisture. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS001922. https://doi.org/10.1029/2019ms001922

Drager, A. J., & van den Heever, S. C. (2017). Characterizing convective cold pools. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1091–1115. https://doi.org/10.1002/2016ms000788

Droegemeier, K., & Wilhelmson, R. (1985). Three-dimensional numerical modeling of convection produced by interacting thunderstorm outflows. Part I: Control simulation and low-level moisture variations. *Journal of the Atmospheric Sciences*, *42*(22), 2381–2403. https://doi.org/10.1175/1520-0469(1985)042<2381:TDNMOC>2.0.CO;2

Emanuel, K., Wing, A. A., & Vincent, E. M. (2014). Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, *6*(1), 75–90. https://doi.org/10.1002/2013ms000270

Engerer, N. A., Stensrud, D. J., & Coniglio, M. C. (2008). Surface characteristics of observed cold pools. *Monthly Weather Review*, *136*(12), 4839–4849. https://doi.org/10.1175/2008mwr2528.1

Feng, Z., Hagos, S., Rowe, A. K., Burleyson, C. D., Martini, M. N., & de Szoeke, S. P. (2015). Mechanisms of convective cloud organization by cold pools over tropical warm ocean during the AMIE/DYNAMO field campaign. *Journal of Advances in Modeling Earth Systems*, *7*(2), 357–381. https://doi.org/10.1002/2014MS000384

Fiévet, R., Meyer, B., & Haerter, J. O. (2023). On the sensitivity of convective cold pools to mesh resolution. *Journal of Advances in Modeling Earth Systems*, *15*(8), e2022MS003382. https://doi.org/10.1029/2022ms003382

Fournier, M. B., & Haerter, J. O. (2019). Tracking the gust fronts of convective cold pools. *Journal of Geophysical Research: Atmospheres*, *124*(21), 11103–11117. https://doi.org/10.1029/2019jd030980

Gentine, P., Garelli, A., Park, S.-B., Nie, J., Torri, G., & Kuang, Z. (2016). Role of surface heat fluxes underneath cold pools. *Geophysical Research Letters*, *43*(2), 874–883. https://doi.org/10.1002/2015gl067262

Glassmeier, F., & Feingold, G. (2017). Network approach to patterns in stratocumulus clouds. *Proceedings of the National Academy of Sciences*, *114*, 201706495.

Grandpeix, J.-Y., & Lafore, J.-P. (2010). A density current parameterization coupled with Emanuel's convection scheme. Part I: The models. *Journal of the Atmospheric Sciences*, *67*(4), 881–897. https://doi.org/10.1175/2009jas3044.1

Haerter, J. O. (2019). Convective self-aggregation as a cold pool-driven critical phenomenon. *Geophysical Research Letters*, *46*(7), 4017–4028. https://doi.org/10.1029/2018GL081817

Haerter, J. O., Böing, S. J., Henneberg, O., & Nissen, S. B. (2019). Circling in on convective organization. *Geophysical Research Letters*, *46*(12), 7024–7034. https://doi.org/10.1029/2019GL082092

Haerter, J. O., Meyer, B., & Nissen, S. B. (2020). Diurnal self-aggregation. *npj Climate and Atmospheric Science*, *3*(1), 30. https://doi.org/10.1038/s41612-020-00132-z

Haerter, J. O., & Schlemmer, L. (2018). Intensified cold pool dynamics under stronger surface heating. *Geophysical Research Letters*, *45*(12), 6299–6310. https://doi.org/10.1029/2017GL076874

Härtel, C., Carlsson, F., & Thunblom, M. (2000). Analysis and direct numerical simulation of the flow at a gravity-current head. Part 2. The lobe-and-cleft instability. *Journal of Fluid Mechanics*, *418*, 213–229. https://doi.org/10.1017/S0022112000001270

Henneberg, O., Meyer, B., & Haerter, J. O. (2020). Particle-based tracking of cold pool gust fronts. *Journal of Advances in Modeling Earth Systems*, *12*(5), e2019MS001910. https://doi.org/10.1029/2019ms001910

Hirt, M., Craig, G. C., Schäfer, S. A., Savre, J., & Heinze, R. (2020). Cold-pool-driven convective initiation: Using causal graph analysis to determine what convection-permitting models are missing. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 2205–2227. https://doi.org/10.1002/qj.3788

Hoeller, J. (2023). CoolDeTa: Detection and tracking of convective cold pools and their causal chains in cloud-resolving simulation data release (version 1.0) [Software & Dataset]. Zenodo. https://doi.org/10.5281/zenodo.10115957

Hoeller, J., Fiévet, R., Engelbrecht, E., & Haerter, J. O. (2024). U-net segmentation for the detection of convective cold pools from cloud and rainfall fields. *Journal of Geophysical Research: Atmospheres*, *129*, e2023JD040126. https://doi.org/10.1029/2023JD040126

Hohenegger, C., Ament, F., Beyrich, F., Löhnert, U., Rust, H., Bange, J., et al. (2023). Fesstval: The field experiment on submesoscale spatio-temporal variability in lindenberg. *Bulletin of the American Meteorological Society*, *104*(10), E1875–E1892. https://doi.org/10.1175/bams-d-21-0330.1

Houze, R. A., Jr. (2004). Mesoscale convective systems. *Reviews of Geophysics*, *42*, RG4003. https://doi.org/10.1029/2004RG000150

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment [Software]. Computing in Science & Engineering, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jeevanjee, N., & Romps, D. M. (2013). Convective self-aggregation, cold pools, and domain size. *Geophysical Research Letters*, *40*(5), 994–998. https://doi.org/10.1002/grl.50204

Jensen, G. G., Fiévet, R., & Haerter, J. O. (2021). The diurnal path to persistent convective self-aggregation. arXiv preprint arXiv:2104.01132.

Kain, J., & Fritsch, J. (1992). The role of the convective "trigger function" in numerical forecasts of mesoscale convective systems. *Meteorology and Atmospheric Physics*, *49*(1–4), 93–106. https://doi.org/10.1007/bf01025402

Khairoutdinov, M. F., & Randall, D. (2006). High-resolution simulation of shallow-to-deep convection transition over land. *Journal of the Atmospheric Sciences*, *63*(12), 3421–3436. https://doi.org/10.1175/jas3810.1

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the arm summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities [Software]. Journal of the Atmospheric Sciences, *60*(4), 607–625. https://doi.org/10.1175/1520-0469(2003)060<0607:crmota>2.0.co;2

Kruse, I. L., Haerter, J. O., & Meyer, B. (2022). Cold pools over The Netherlands: A statistical study from tower and radar observations. *Quarterly Journal of the Royal Meteorological Society*, *148*(743), 711–726. https://doi.org/10.1002/qj.4223

Kurowski, M. J., Suselj, K., Grabowski, W. W., & Teixeira, J. (2018). Shallow-to-deep transition of continental moist convection: Cold pools, surface fluxes, and mesoscale organization. *Journal of the Atmospheric Sciences*, *75*(12), 4071–4090. https://doi.org/10.1175/jas-d-18-0031.1

Markowski, P., & Richardson, Y. (2010). *Mesoscale meteorology in midlatitudes*. John Wiley and Sons. https://doi.org/10.1002/9780470682104

Meyer, B., & Haerter, J. O. (2020). Mechanical forcing of convection by cold pools: Collisions and energy scaling. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002281. https://doi.org/10.1029/2020ms002281

Moncrieff, M. W., & Liu, C. (1999). Convection initiation by density currents: Role of convergence, shear, and dynamical organization. *Monthly Weather Review*, *127*(10), 2455–2464. https://doi.org/10.1175/1520-0493(1999)127<2455:cibdcr>2.0.co;2

Moseley, C., Hohenegger, C., Berg, P., & Haerter, J. O. (2016). Intensification of convective extremes driven by cloud–cloud interaction. *Nature Geoscience*, *9*(10), 748–752. https://doi.org/10.1038/ngeo2789

Muller, C., & Bony, S. (2015). What favors convective aggregation and why? *Geophysical Research Letters*, *42*(13), 5626–5634. https://doi.org/10.1002/2015GL064260

Muller, C., Yang, D., Craig, G., Cronin, T., Fildier, B., Haerter, J. O., et al. (2022). Spontaneous aggregation of convective storms. *Annual Review of Fluid Mechanics*, *54*(1), 133–157. https://doi.org/10.1146/annurev-fluid-022421-011319

Niehues, J., Jensen, G. G., & Haerter, J. O. (2021). Self-organized quantization and oscillations on continuous fixed-energy sandpiles. arXiv preprint arXiv:2111.04470.

Nissen, S. B., & Haerter, J. O. (2021). Circling in on convective self-aggregation. *Journal of Geophysical Research: Atmospheres*, *126*(20), e2021JD035331. https://doi.org/10.1029/2021jd035331

Parker, D. J., & Diop-Kane, M. (2017). *Meteorology of tropical West Africa: The forecasters' handbook*. John Wiley & Sons.

Parker, M. D. (2008). Response of simulated squall lines to low-level cooling. *Journal of the Atmospheric Sciences*, *65*(4), 1323–1341. https://doi.org/10.1175/2007jas2507.1

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python [Software]. *Journal of Machine Learning Research*, *12*, 2825–2830. Retrieved from https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

Pei, S., Shinoda, T., Soloviev, A., & Lien, R.-C. (2018). Upper ocean response to the atmospheric cold pools associated with the Madden-Julian Oscillation. *Geophysical Research Letters*, *45*(10), 5020–5029. https://doi.org/10.1029/2018gl077825

Rio, C., Hourdin, F., Grandpeix, J.-Y., & Lafore, J.-P. (2009). Shifting the diurnal cycle of parameterized deep convection over land. *Geophysical Research Letters*, *36*, L07809. https://doi.org/10.1029/2008gl036779

Rochetin, N., Hohenegger, C., Touzé-Peiffer, L., & Villefranque, N. (2021). A physically based definition of convectively generated density currents: Detection and characterization in convection-permitting simulations. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2020MS002402. https://doi.org/10.1029/2020ms002402

Romps, D. M., & Jeevanjee, N. (2016). On the sizes and lifetimes of cold pools. *Quarterly Journal of the Royal Meteorological Society*, *142*(696), 1517–1527. https://doi.org/10.1002/qj.2754

Rooney, G. G. (2015). Descent and spread of negatively buoyant thermals. *Journal of Fluid Mechanics*, *780*, 457–479. https://doi.org/10.1017/jfm.2015.484

Rooney, G. G., Stirling, A. J., Stratton, R. A., & Whitall, M. (2021). C-Pool: A scheme for modelling convective cold pools in the Met Office unified model. *Quarterly Journal of the Royal Meteorological Society*, *148*(743), 962–980.

Rotunno, R., Klemp, J. B., & Weisman, M. L. (1988). A theory for strong, long-lived squall lines. *Journal of the Atmospheric Sciences*, *45*(3), 463–485. https://doi.org/10.1175/1520-0469(1988)045<0463:atfsll>2.0.co;2

Schlemmer, L., & Hohenegger, C. (2016). Modifications of the atmospheric moisture field as a result of cold-pool dynamics. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 30–42. https://doi.org/10.1002/qj.2625

Schumacher, R. S., & Rasmussen, K. L. (2020). The formation, character and changing nature of mesoscale convective systems. *Nature Reviews Earth & Environment*, *1*(6), 300–314. https://doi.org/10.1038/s43017-020-0057-7

Seifert, A., & Heus, T. (2013). Large-eddy simulation of organized precipitating trade wind cumulus clouds. *Atmospheric Chemistry and Physics*, *13*(11), 5631–5645. https://doi.org/10.5194/acp-13-5631-2013

Simpson, J. E. (1972). Effects of the lower boundary on the head of a gravity current. *Journal of Fluid Mechanics*, *53*(4), 759–768. https://doi.org/10.1017/S0022112072000461

Simpson, J. E. (1980). Downdrafts as linkages in dynamic cumulus seeding effects. *Journal of Applied Meteorology*, *19*(4), 477–487. https://doi. org/10.1175/1520-0450(1980)019<0477:dalidc>2.0.co;2

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, *91*(3), 99–164. https://doi.org/10.1175/1520-0493(1963)091<0099:gcewtp>2.3.co;2

Straka, J. M., Wilhelmson, R. B., Wicker, L. J., Anderson, J. R., & Droegemeier, K. K. (1993). Numerical solutions of a non-linear density current: A benchmark solution and comparisons. *International Journal for Numerical Methods in Fluids*, *17*(1), 1–22. https://doi.org/10.1002/fld. 1650170103

Tompkins, A. M. (2001a). Organization of tropical convection in low vertical wind shears: The role of cold pools. *Journal of the Atmospheric Sciences*, *58*(13), 1650–1672. https://doi.org/10.1175/1520-0469(2001)058<1650:ootcil>2.0.co;2

Tompkins, A. M. (2001b). Organization of tropical convection in low vertical wind shears: The role of water vapor. *Journal of the Atmospheric Sciences*, *58*(6), 529–545. https://doi.org/10.1175/1520-0469(2001)058<0529:ootcil>2.0.co;2

Torri, G., & Kuang, Z. (2016). Rain evaporation and moist patches in tropical boundary layers. *Geophysical Research Letters*, *43*(18), 9895–9902. https://doi.org/10.1002/2016gl070893

Torri, G., & Kuang, Z. (2019). On cold pool collisions in tropical boundary layers Giuseppe. *Geophysical Research Letters*, *46*(1), 399–407. https://doi.org/10.1029/2018GL080501

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). scikit-image: Image processing in Python [Software]. PeerJ, *2*, e453. https://doi.org/10.7717/peerj.453

Vogel, R., Konow, H., Schulz, H., & Zuidema, P. (2021). A climatology of trade-wind cumulus cold pools and their link to mesoscale cloud organization. *Atmospheric Chemistry and Physics*, *21*(21), 16609–16630. https://doi.org/10.5194/acp-21-16609-2021

Wakimoto, R. (2001). Convectively driven high wind events (pp. 255–298). https://doi.org/10.1007/978-1-935704-06-5_7

Waskom, M. L. (2021). Seaborn: Statistical data visualization [Software]. Journal of Open Source Software, *6*(60), 3021. https://doi.org/10.21105/ joss.03021

Weisman, M. L., Klemp, J. B., & Rotunno, R. (1988). Structure and evolution of numerically simulated squall lines. *Journal of the Atmospheric Sciences*, *45*(14), 1990–2013. https://doi.org/10.1175/1520-0469(1988)045<1990:saeons>2.0.co;2

Zuidema, P., Li, Z., Hill, R. J., Bariteau, L., Rilling, B., Fairall, C., et al. (2012). On trade wind cumulus cold pools. *Journal of the Atmospheric Sciences*, *69*(1), 258–280. https://doi.org/10.1175/jas-d-11-0143.1

Zuidema, P., Torri, G., Muller, C., & Chandra, A. (2017). A survey of precipitation-induced atmospheric cold pools over oceans and their interactions with the larger-scale environment. *Surveys in Geophysics*, *38*(6), 1–23. https://doi.org/10.1007/s10712-017-9447-x