Detecting Cold Pool Family Trees in Convection Resolving Simulations

Jannik Hoeller[1,2], Romain Fiévet[2], Jan O. Haerter[1,2,3]

[1]Integrative Modeling, Leibniz Centre for Tropical Marine Research, Fahrenheitstr. 6, 28359 Bremen, Germany
[2]Niels Bohr Institute, Copenhagen University, Blegdamsvej 17, 2100 Copenhagen, Denmark
[3]Physics and Earth Sciences, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

Key Points:

- Our CoolDeTA algorithm reliably detects and tracks cold pools.
- The algorithm identifies the causal chain of cold pools and the rain cells triggered.
- CoolDeTA opens for new studies into the dynamics of convective self-organization through cold pools.

Corresponding author: Jannik Hoeller, `jannik.hoeller@leibniz-zmt.de`

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 in reliably identifying the active gust front. The current algorithm, termed CoolDeTA, aims to detect and track cold pools along with their active gust fronts and the "offspring" rain cells generated nearby. We show how CoolDeTA can reconstruct cold pool family trees. Using it allows us to contrast RCE and diurnal cycle cold pool dynamics, as well as cases with vertical wind shear and without. The results suggest a conceptual model where cold pool triggering of children rain cells follows a simple birth rate, which is proportional to a cold pool'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 cold pool.

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 cold pool 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 expectation on 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 (J. Simpson, 1980; Droegemeier & Wilhelmson, 1985). 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," i.e., the demarcation between the cold pool and the ambient atmosphere. Many recent studies have considered the characteristics of cold pools, because they have been implicated in the spatio-temporal organization of the convective cloud field — thus impacting on non-random organization of rainfall, with consequences for extreme events, such as flooding.

Observationally, measuring cold pools is difficult, as high-resolution spatio-temporal records of the subcloud boundary layer are required to track a given spreading cold pool. Field campaigns thus often remain restricted to a particular, often spatially confined area (Feng et al., 2015), 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 cold pool 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 cold pool events also point to lowered sea surface temperatures (SST) after pronounced cold pool 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. With key features such as rain evaporation simulated to a reasonable degree of realism, 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 (Gentine et al., 2016; Drager et al., 2020).

On the larger-scale level of simulated mesoscale organization, cold pools have suggested as a possible cause for non-random distributions of rain cells in space (Haerter et al., 2019), which may arise when cold pool 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 cold pool activity (Jeevanjee & Romps, 2013; Haerter, 2019; Nissen & Haerter, 2021; Muller et al., 2022). 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 cold pool activity, especially regarding the formation of mescoscale 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 (Rio et al., 2009; Grandpeix & Lafore, 2010; Rooney et al., 2021). This may be due to an incomplete understanding regarding the fundamental processes affecting cold pool 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 (Romps & Jeevanjee, 2016; Haerter & Schlemmer, 2018; Meyer & Haerter, 2020; Fiévet et al., 2022), may help improve parameterizations further.

Several recent works have brought forward methods to track cold pools, in particular their gust fronts, in numerical simulations (Schlemmer & Hohenegger, 2016; Drager & van den Heever, 2017; Fournier & Haerter, 2019; Henneberg et al., 2020; Hirt et al., 2020). Yet, the detection of cold pools in numerical simulations is far from trivial: (i) the area affected by rainfall, where the cold pool is fed by rain evaporation, is often not cleanly delimited and subject to setting a threshold value, e.g., on the precipitation rate; (ii) the cold pool 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 (J. E. Simpson, 1972; Härtel et al., 2000; Wakimoto, 2001; Markowski & Richardson, 2010); (iii) the larger-scale pattern of cold pools

and the rain cells produced by the their collisions, is highly complex, with families formed by cold pool-rain cell networks making for a challenging tracking problem.

The current work addresses these challenges by building a detection and tracking procedure based on a combination of buoyancy and dynamical measures.

2 Methods

2.1 Cloud Resolving Simulation Data

We employ the System for Atmospheric Modeling (SAM) (M. 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 Δh is chosen as $\Delta h = 200\,m$ throughout, thus yielding $N = 1200$ 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 cold pool effects, such as the dynamic gust front, can be simulated at satisfactory detail (Straka et al., 1993; Bryan et al., 2003; Meyer & Haerter, 2020; Hirt et al., 2020; Fiévet et al., 2022). The mesh is discretized along the vertical direction z using 64 levels of increasing depth Δz, ranging from $\Delta z = 50\,m$ at the first level ($z = 25\,m$) to $1000\,m$ 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 (M. F. Khairoutdinov & Randall, 2003) along with a first-order turbulent subgrid scheme (Smagorinsky, 1963). The radiative fluxes are calculated using the the National Center for Atmospheric Research (NCAR) 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
$$T_s(t) = \overline{T}_s - \Delta T \cos\left(2\pi\, t/T\right) . \tag{1}$$
Here, $\overline{T}_s = 298\,K$ and ΔT is either chosen as $0K$, which we term "rce", or as $2K$ and $4K$, 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
$$\begin{aligned} u_0(z < 1\,km) &= 0\,, \\ u_0(1\,km < z < 19\,km) &= \delta_u \left(\frac{z - 1\,km}{18\,km} \right), \\ u_0(z > 19\,km) &= \delta_u\,. \end{aligned} \tag{2}$$

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

The above settings for ΔT 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 $\delta_u = 16\,ms^{-1}$. 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 CFL number of 0.8. The simulation output frequency is of one every $\Delta t = 10\,min$.

Each Δt, the following instantaneous fields are output: surface rainfall intensity $I(x, y, t)$, near-surface specific humidity $q(x, y, t) = q(x, y, z = 25\,m, t)$, virtual temperature $T_v(x, y, t) = T_v(x, y, z = 25\,m, t)$ and the horizontal wind vector $\mathbf{u} \equiv (u, v)$, with $u = u(x, y, z = 25\,m, t)$ and $v = v(x, y, z = 25\,m, t)$ as the lowest-level horizontal wind components in the x and y directions. For simplicity, we refer to Δt 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 cold pool. Each rain cell in the domain is assigned a unique rain identification number (ID). In case a rain cell generates a cold pool, the cold pool receives the same ID as the associated rain cell. On the other hand, cold pool gust fronts can trigger new rain cells, which might again generate new cold pools (Tompkins, 2001; Böing, 2016; Haerter et al., 2019). Accordingly, the algorithm stores rain cells triggered at a cold pool gust front, as well as potential cold pools emerging from these rain cells, as children of that particular cold pool. In this manner, the algorithm not only detects and tracks cold pools, but also their relationships.

The proposed cold pool detection and tracking algorithm (CoolDeTA) consists of four main steps: (i) identification of rain patches, (ii) segmentation of the domain into "potential cold pool" and "no cold pool" regions, (iii) derivation of markers, and (iv) 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 rain patches and cold pool patches at time step t_n. Analogously, we also define an aux-

iliary 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, e.g., $p_r^m \equiv \{(i,j,n) \in [1,N] \times [1,N] \times [1 \wedge p_r(i,j,n) = m]\}$.

(i) Identifying rain patches (RPs). At any discrete time t_n a *rain patch* 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 $I_0 = 2\ mm\,h^{-1}$, 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 rain patch 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 rain patch, 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 $A_0 = 2\,km^2$ 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

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

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 Eq. 6 is subtracted from unity, that is, relatively low T_v increases the probability of a cold pool. We choose perturbations rather than absolute quantities in Eq. 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 cold pool" or "no cold pool" 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 cold pool" grid cells and $p_{bin}(i,j,n) = 0$ for all "no cold pool" cells. As an example, Fig. 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 vs. 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 rain patches, fully developed cold pools should at least encompass the same area as their related rain patches. Thus, we apply the minimum area A_0 criterion from (i) also to the identified "potential cold pool" regions. For this reason, spatially contiguous "potential cold pool" 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 $A(p_{seg},n,m) \equiv \Delta h^2 \sum_{i,j=1}^{N} \delta(p_{seg}(i,j,n),m) \geq A_0$.

(iii) Deriving markers. A marker could be defined as a center of mass (COM) of a given rain patch m at time t_n. We however found that the actual (dynamical) downdraft within this patch more precisely marks the cold pool 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

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

which is large for locations of minimal temperature and vertical velocity. The corresponding center of mass, (C^m_x, C^m_y), follows as

$$C^m_x = \frac{\sum_{i,j=1}^{N} x_i \mathcal{I}^m_{mar}(i,j,n)}{\sum_{i,j=1}^{N} \mathcal{I}^m_{mar}(i,j,n)}\,, \tag{10}$$

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

When the rain associated with a cold pool labeled m from the previous time step t_{n-1} stops, the cold pool will have no active rain patch p^m_r in $p_r(i,j,n)$ anymore, i.e., $\sum_{i,j=1}^{N} \delta(p_r(i,j,n), m) = 0$. 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 rain patch, 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 cold pool, i.e., the center of its initial rain patch. 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 cold pool 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 rain patch field $p_r(i,j,n)$ (*compare*: Eq. 5) - a horizontal $N \times N$ integer ID field where $p_{cp}(i,j,n)$ specifies the ID of a cold pool or $p_{cp}(i,j,n) = 0$, if no cold pool is present at (i,j) and time index n. The cold pool 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* Fig. 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 cold pool field, we provide the watershed algorithm with three ingredients: the locations of the water sources, i.e., 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 typically characterized by cold (low $\hat{T}(i,j)$) and 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 cold pool boundaries (large $\hat{w}(i,j)$), thus separating the cold pool 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 Fig. 1c are separated by the mask's "no cold pool" region between them, CP2 and CP3 are in contact with each other. The contact location, i.e., the point where the two cold pools collide (in 2D it is a line), depends on the elevation map $\mathcal{I}_{top}$ and remains constant for the rest of the flooding.

Since only its gust front should separate a cold pool from its environment, we do not allow holes, i.e., 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 cold pool. 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 cold pool basin in the elevation map can consist of multiple local minima, caused by newly formed rain patches within the cold pool area or an advected initial rain patch. 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 accidentally flooded from neighbouring cold pools.

Fig. 1b summarizes the main steps of CoolDeTA. Note that both tracking and the interactions between the 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 rain patches, $p_r(i,j,n-1)$, 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 rain patches in step (i). Since a rain patch, labeled m in $p_r(i,j,n)$, obtains a corresponding marker $p_{mar}(i^m_{mar}, j^m_{mar}, n) = m$, an associated cold pool 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 rain patch in $p_r(i,j,n)$ and the ID of an associated cold pool in $p_{cp}(i,j,n)$. Accordingly, cold pools can be tracked by adjusting the label of tracked rain patches in step (i). A rain patch labeled m' at time step t_n is tracked if the overlap with a rain patch 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 $\mathbf{1}$ 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 cold pool 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 $S_2 \equiv \{(i,j) \in [1,N] \times [1,N], p_{cp}(i,j,n-1) = m \wedge p_r(i,j,n) = m'\}$. Fig. 1d visualizes both cases: On the one hand, RP3 is tracked in time step t_n since the corresponding rain patch 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 rain patch fully overlaps CP4 from time step t_{n-1}.

If a rain patch labeled m' is overlapped sufficiently by a previous rain patch labeled m_1 and a previous cold pool labeled m_2, i.e., $O_r(m_1,m',n-1,n) \geq f_r \times A(p_r^{m'}) \wedge O_{cp}(m_2,m',n-1,n) = f_{cp} \times A(p_r^{m'})$, the previous cold pool provides its label, i.e., the label of rain patch m' is replaced with the label m_2. In case rain patch m' is overlapped by multiple rain patches from the previous time step with deviating labels, the rain patch with the largest overlap, O_r^{max}, provides its label. For this reason, the new rain patch in time step t_n of Fig. 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 rain patches, 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 rain patch anymore - also their latest rain marker (*see* step (iv)). Accordingly, CP1, CP3 and CP4 in Fig. 1d each receive two markers in time step t_n, representing the center of mass of the current rain patch and that of the initial rain patch from

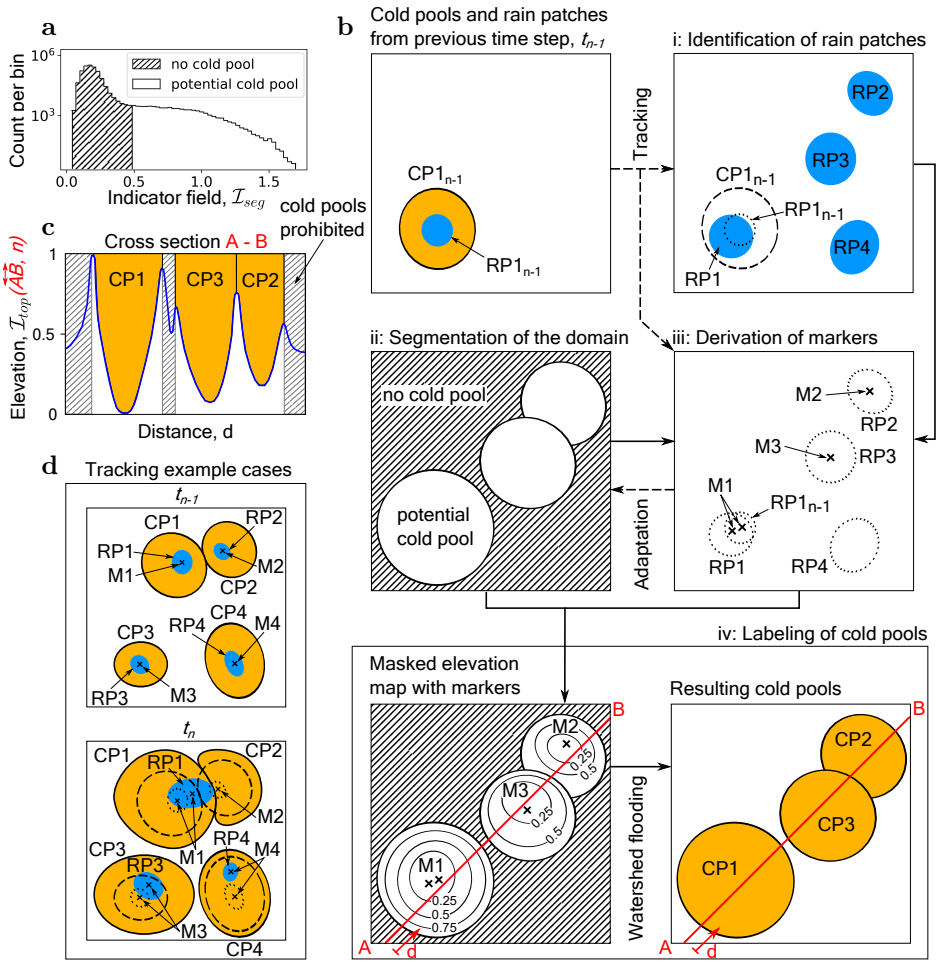

Figure 1: **Proposed cold pool detection and tracking algorithm. a**, Histogram of the indicator field, $\mathcal{I}_{seg}(i, j, n)$ (Eq. 6), for diu4K at t_{471}. "no cold pool (CP)" and "potential CP" grid cells, as determined by the k-means algorithm, are shown as hatched and striped curve fillings. **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 Δt. **c**, Elevation map values (blue curve), $\mathcal{I}_{top}(\overleftrightarrow{AB}, n)$, 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.

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 rain patch anymore.

2.4 Cold Pool Relationships

The association of rain patches 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 cold pool (gust front) triggered which rain patch and thus also a potential cold pool resulting from it, CoolDeTA reveals cold pool networks and their associated causal chains.

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

In case a rain patch 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 cold pool m with an overlap $0 < O_{cp}(m, m', n-1, n) < f_{cp} \times A(p_r^{m'})$ becomes parent of rain patch m' and thus also of potential cold pools resulting from it.

2.5 Cold Pool Dissipation

When the rain associated with a cold pool stops and thus no longer feeds the density current anymore, the surface boundary layer recovers until the cold pool 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 infinite lifetime, we implemented a dissipation mechanism within step (iii) of CoolDeTA. To this end, we distinguish "dissipating" and "dissipated" cold pools. We define a cold pool, labeled m, as dissipating at time step t_n if it has no active rain anymore, i.e., $\sum_{i=1,j=1}^{N,N} \mathbf{1}((i,j) \in \{(i,j) \in [1,N] \times [1,N], p_r(i,j,n) = m)\} = 0$, and if parts of its latest rain patch p_r^m from time step t_n^{last} are classified as "no cold pool" in $p_{bin}(i,j,n)$, i.e., $O_{seg}(m, 1, n^{last}, n) < 1.0 \times A(p_r^m)$ with

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

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 cold pool we count the number of time steps in which it is dissipating, n_{dis}. We consider a cold pool, labeled m, as dissipated if $n_{dis}^m \geq 3$ and $O_{seg}(m, 1, n^{last}, n) = 0$, i.e., if it is dissipating for ≥ 3 time steps ($= 30\,min$) and if its latest rain patch p_r^m is completely classified as "no cold pool" in $p_{bin}(i,j,n)$. A cold pool 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 cold pool field $p_{cp}(i,j,n)$.

2.6 Additional Algorithm Rules

The evolution of the populated fields of both cold pools, p_{cp}, and rain patches, p_r, is highly complex. On the one hand, relatively cool air associated with remnants of dis-

sipated cold pools might lead to the misclassification of grid cells as "potential cold pool" 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 rain patches or cold pools, as well as the formation of new rain patches at cold pool 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.

Divergence criterion. Cold pools are driven by atmospheric density gradients resulting from the evaporation of rain. Evaporative cooling causes the affected air to sink towards 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 cold pool" 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 $\operatorname{int} p^m_{seg}$ and if $\overline{\nabla \cdot \mathbf{u}(i,j,n)} \leq 0$ for grid cells in its patch boundary ∂p^m_{seg}. The divergence is approximated using a central difference scheme. We define a grid cell (i,j) as part of the interior $\operatorname{int} p^m_{seg}$ of patch m if all contiguous grid cells are contained in p^m_{seg}. A grid cell (i,j) is part of the boundary ∂p^m_{seg} of patch m if it encompasses at least one contiguous grid cell contained in p^m_{seg} and one contiguous grid cell not contained in p^m_{seg}. Since the "potential cold pool" region m in $p_{seg}(i,j,n)$ can contain multiple cold pool instances and thus colliding gust fronts, we exclude a grid cell (i,j) from $\operatorname{int} p^m_{seg}$ if $\nabla \cdot \mathbf{u}(i,j,n) < \overline{\nabla \cdot \mathbf{u}(n)} - z_{95th}\,\sigma_d(n)$ where σ_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 cold pool" region m compensating its convergent flow. Analogous to $\operatorname{int} p^m_{seg}$ we thus exclude a grid cell (i,j) from ∂p^m_{seg} 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 cold pool interior and boundary, as colliding cold pool 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 exclusion of potentially misleading grid cells, as well as the weak conditions in terms of $\operatorname{int} p^m_{seg}$ not being convergent and ∂p^m_{seg} not being divergent ensure that no correctly classified "potential cold pool" region is dropped.

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

Apart from cold pool interactions, where a cold pool gust front can be pushed towards the cold pool center, cold pool gust fronts spread outwards only, due to the horizontally diverging flow in the cold pool interior. As long as a cold pool, 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 cold pool at the previous time step t_{n-1}, $p_{cp}(i,j,n-1) = m$, are labeled as "potential cold pool" in $p_{bin}(i,j,n)$. Whether cold pool m looses parts of its occupied area to an intersecting cold pool 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).

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 cold pool" in p_{bin}. In order to detect fully developed cold pools only, a new rain patch, labeled m, detected at time step n, obtains a rain marker at time step n only if $O_{seg}(m,1,n,n) \geq f_s \times A(p^m_r)$ with $f_s = 0.75$.

Parameter	Value/condition used	Description
I_0	2 $mm\,h^{-1}$	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

Table 1: **Parameters and conditions in the proposed algorithm.** 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.

The final check is applied to the labeled cold pool 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 rain patches, we get an indication of erroneous labeling. Erroneous labeling can result from cool air remnants being classified as "potential cold pool" 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 are new rain patches, triggered at cold pool gust fronts: Since not each of these rain patches generates an independent child cold pool which is separated from its parent cold pool 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 cold pool should, on the one hand, at least feature the size of its rain patch. On the other, a new cold pool cannot occupy an area substantially larger than its rain patch, given the time step of $\Delta t = 10\,min$. To intervene only in erroneous cases, we keep a new cold pool, labeled m, if $1 \leq A(p_{cp}^m)/A(p_r^m) \leq 3$. Otherwise, we either drop it, i.e., $p_{cp}^m = 0$, if it has no other cold pool in its neighbourhood, or we assign the label of another cold pool, labeled m', i.e., $p_{cp}^m = m'$, otherwise. In the latter case, the cold pool label m' is determined based on its number of parents, k_p: if $k_p = 0$, it is either the label of the cold pool which overlaps the largest area of rain patch m (if $A(p_{cp}^m)/A(p_r^m) < 1$), or the label of the cold pool which lost the largest area with respect to the previous time step t_{n-1} to cold pool m (if $A(p_{cp}^m)/A(p_r^m) > 3$). If $k_p = 1$, cold pool m is simply assigned the parent label, m_p, i.e., $m' = m_p$. The final case, $k_p > 1$, is analogous to the first but based on the parents of cold pool m.

2.7 Analysis of the Simulation Data

We consider the first three 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 dynamics in both the RCE and diurnal cycle setups (Fig. 2e,f). Further time series are visualized in Fig. 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 (Fig. 2f), we identify the corresponding starting time step based on the standard

deviation of the virtual temperature, σ_{Tv}, and its time derivative, $\dot{\sigma}_{Tv}$. Based on their one hour running average, we find $\sigma_{Tv} \geq 0.15\,K$ and $\dot{\sigma}_{Tv} \geq 0.02\,Ks^{-1}$ as good indicators for the onset of 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 (Fig. 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 cold pool" and "no cold pool", 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 $\sigma_{Tv} < 0.15\,K \wedge R_{px} < 0.0015\,mm$, with $R_{px} = \frac{R}{N \times N}$. Analogous to the definition of the start time step, we apply the conditions to the one hour running averages of σ_{Tv} 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 cold pool detection and tracking method. We first compare our method to previous ones (Tompkins, 2001; Feng et al., 2015; Gentine et al., 2016). We then employ our current method to discuss the effect of the simulation setup, i.e., wind shear or diurnal cycles, on cold pool expansion. Third, we make use of our method to analyze the causal relationships between cold pools and their effect on cold pool rainfall. Based on our findings we then propose a simple model to capture cold pool spreading-triggering dynamics. We finish the chapter by visualizing the family trees of two cold pools.

3.1 Comparison to Other Methods

Several previous cold pool detection methods have exploited thermodynamic cold pool features (Tompkins, 2001; Feng et al., 2015; Gentine et al., 2016). For an RCE-like setup, rce0K (Fig. 2a), as well as a diurnal cycle case, diu4K (Fig. 2b—d), we compare CoolDeTA to such existing methods. The time steps shown are highlighted in the time series (Fig. 2e,f) as thin vertical lines. Inspecting the different examples (Fig. 2a—d), it is apparent that the cold pool dynamics in rce0K occurs at relatively small scales $\mathcal{O}(10\,km)$, whereas the dynamics in diu4K ranges from scattered small scale events during the early stages of the diurnal cycle to elaborate mesoscale features $\mathcal{O}(100\,km)$ 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 cold pool gust fronts (red pixels in Fig. 2a—d). Inspecting the plots, it is apparent that in several occasions the detection results by all four methods nearly coincide, especially during early stages of cold pool 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 e.g., by comparing the lines of different colors in the inset. As is visible on the left side of the larger cold pool patch in t_{473} of the inset, CoolDeTA keeps dissipating cold pool instances as long as they feature a partly active dynamic gust front.

For the early stage of the diurnal cycle (Fig. 2b) the agreement of CoolDeTA, Tompkins (2001) and Gentine et al. (2016) 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 Fig. 2a—d for cold pool patches of all methods, CoolDeTA applies the threshold A_0 also to rain patches associated with individual cold pools and conducts further checks to confirm the existence of a cold pool (*see* Sec. 2.6). Consequently, the minimum cold pool 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 (Fig. 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 cold pool temporal expansion and how it depends on the choice of simulation setup. Cold pool expansion is a basic dynamical cold pool feature, which we here capture as the change in cold pool area, A_{cp}, over time, on the one hand for cold pools without offspring during their lifetime (Fig. 3a) and on the other for cold pools with offspring (Fig. 3b). To not distort the results with cold pools which shrink due to cold pool interactions, we only consider a cold pool as long as its area is expanding and thus $A_{cp}(n) > A_{cp}(n-1)$.

Conditioning on cold pools without children (Fig. 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 cold pool radius on time (Rooney, 2015; Meyer & Haerter, 2020), we indicate a dependency of $A_{cp} \sim t^{\alpha}$ with $\alpha = 1.2$ (dashed gray lines). After approximately one hour 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 favourable subregions.

The curves for cold pools with children exhibit an increased power-law exponent already at earlier stages (Fig. 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 a consequence of enhanced triggering near cold pool edges for cold pools with children, whereby the original cold pool can merge with additional offspring excited by it — thus developing into a mesoscale cold pool. Apart from this, the diurnal cycle again appears to promote rapid cold pool expansion, especially during the later stages.

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

3.3 Effect of Cold Pool Relationships on Cold Pool Rainfall

Given these findings on cold pool dynamics we now investigate: (i) how the total accumulated rainfall corresponding to any given cold pool relates to its capacity to trigger offspring (Fig. 4a); and (ii) how this rainfall relates to the number of parents of a given

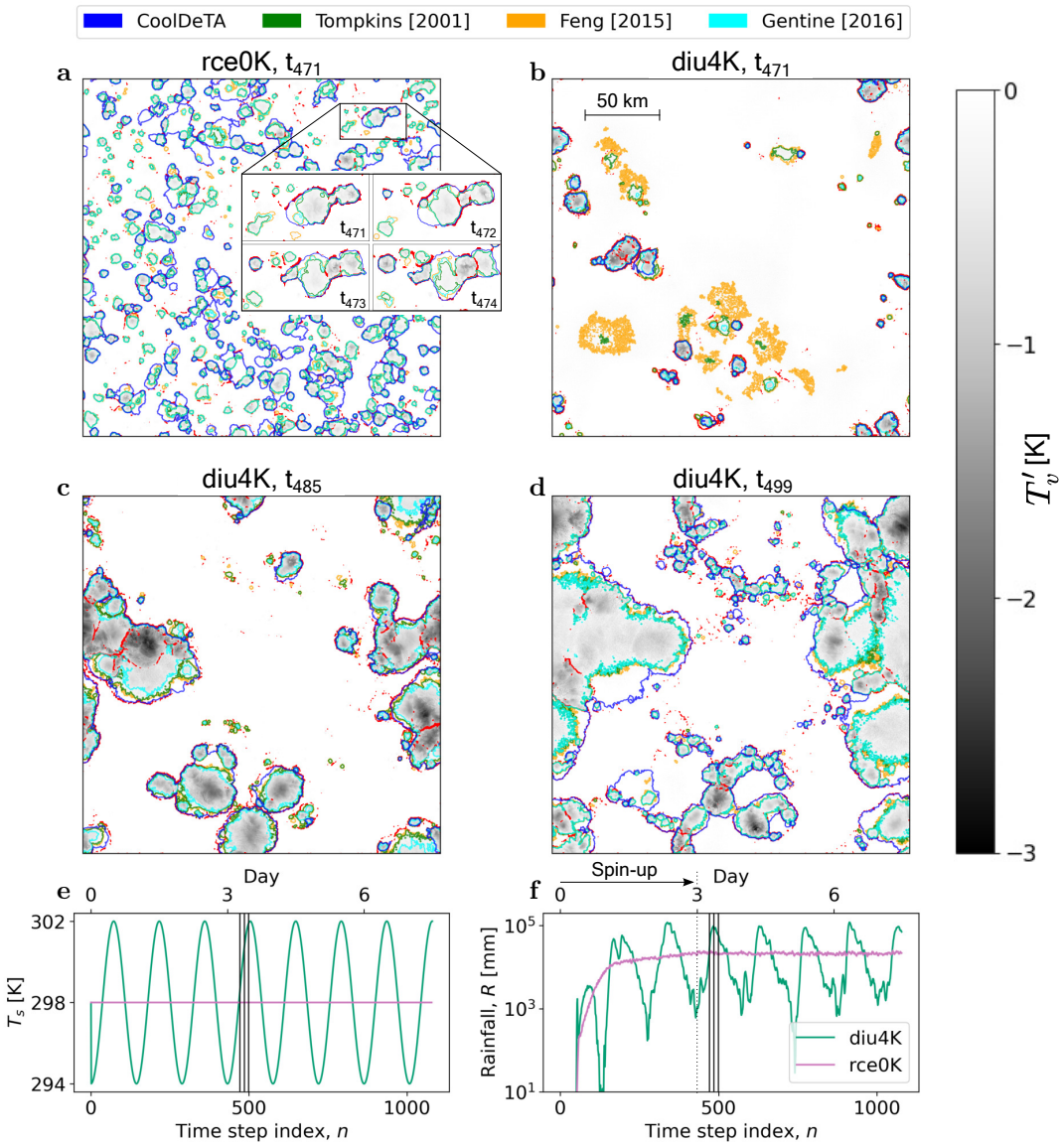

Figure 2: **Comparison of cold pool detection methods. a**, Time step index 471 of rce0K, showing near-surface virtual temperature perturbations, T_v', with superimposed dynamical gust front, i.e., $w > \overline{w} + 2\sigma_w$ (red scatter); The superimposed colors represent contours of cold pool patches $\geq A_0$ based on different cold pool 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 three days (432 time steps) is indicated by the vertical dotted line and the horizontal arrow above the panel.

cold pool (Fig. 4b). For this purpose, we define the total accumulated rainfall of a cold pool labeled m as

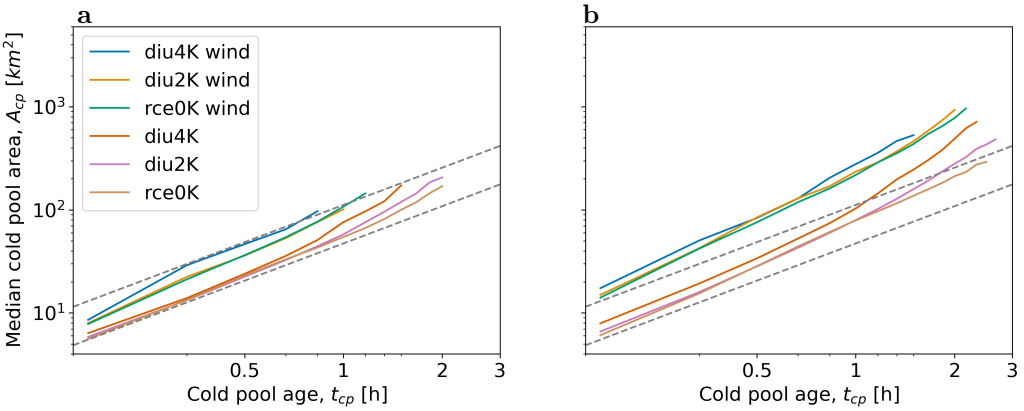

Figure 3: **Cold pool expansion for various simulation setups. a**, Median cold pool area, A_{cp}, vs. cold pool age, t_{cp} for cold pools without any children during their lifetime. Colors represent different simulation setups; Dashed grey lines indicate areas $\propto$ ${t_{cp}}^{1.2}$; Only cold pool 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.

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

Overall, the majority of cold pools do not give rise to offspring, with approximately 50 percent of cold pools without detectable children (Fig. 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 cold pool children, associated rain patches tend to have much greater total rainfall rates (Fig. 4a). These rates range from a factor of two for rce0K to more than an order of magnitude for 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 (Fig. 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 parent. We attribute this to the higher risk that a single-parent cold pool has of getting merged into the parent cold pool: whereas multiple parents are associated with cold pool 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 cold pool spreading-triggering dynamics, we suggest a simple model (Fig. 5b) where cold pools primarily grow by peripheral spawning of new rain cells, thus child cold pools, that then help the cold pool expand further. In our model each cold pool is ini-

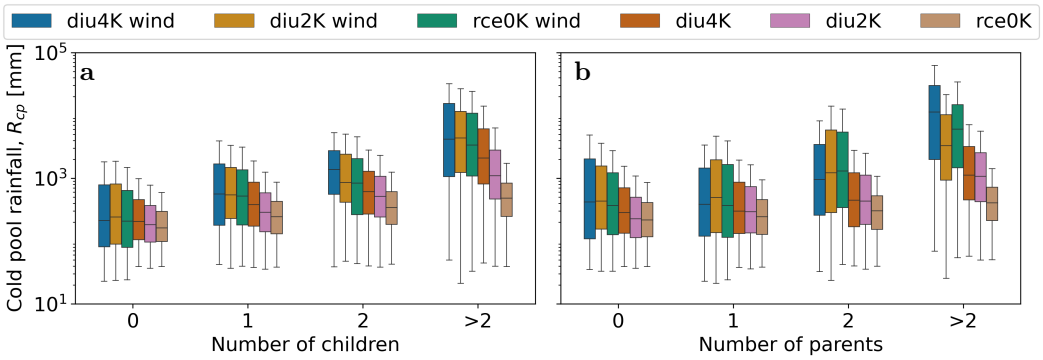

Figure 4: **Total cold pool rainfall. a**, Total accumulated rainfall associated with cold pools, R_{cp}, depending on their number of children for different simulation setups. Colored bars represent the interquartile ranges $IQR = Q3 - Q1$ of the simulation setups, with the first quartile $Q1$ and the 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.

tialized with an area a of $a = a_0 \equiv 6\,km^2$ representing the median area of all cold pools from Fig. 3 at $t_{cp} = 10\,min$, and circumference $l = 2(\pi a_0)^{1/2} \approx 8.7\,km$. This circumference is subdivided into segments, each of length $l_0 \equiv 1.6\,km$, taken to represent the approximate diameter of a new rain cell of area A_0. Within each segment and time interval $\Delta\tau \sim 20\,min$, 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 cold pool area a. In addition, as the cold pool 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 cold pool is considered terminated and no further expansion is possible. If a number $m_s > 0$ of segments give rise to spawning, the cold pool area is incremented by $\Delta a \equiv m_s\, a_0$ and the expansion proceeds iteratively (tree diagram in Fig. 5b). At each fork of the tree diagram, at probability $1 - P(a_i)$ the cold pool expansion is terminated. $1 - P(a_i)$ represents the probability that none of the segments spawn a new rain cell, i.e., for $n_s = l/l_0$ segments and spawning probability p, $1 - P(a_i) = (1-p)^{n_s}$. Conversely, the probability for continued cold pool expansion is $P(a_i) = 1 - (1-p)^{n_s}$. 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 cold pool realizations allows us to mimic the statistics found for the cloud-resolving simulations (Fig. 5c): whereas small values of spawning probability, e.g., $p = 0.035$, 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 cold pool dynamics is twofold: spawning probability p may be determined by the convective instability surrounding a given parent cold pool, 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 (Fig. 5a) for an artificially lengthened diurnal cycle, say, corresponding to a 48h-day.

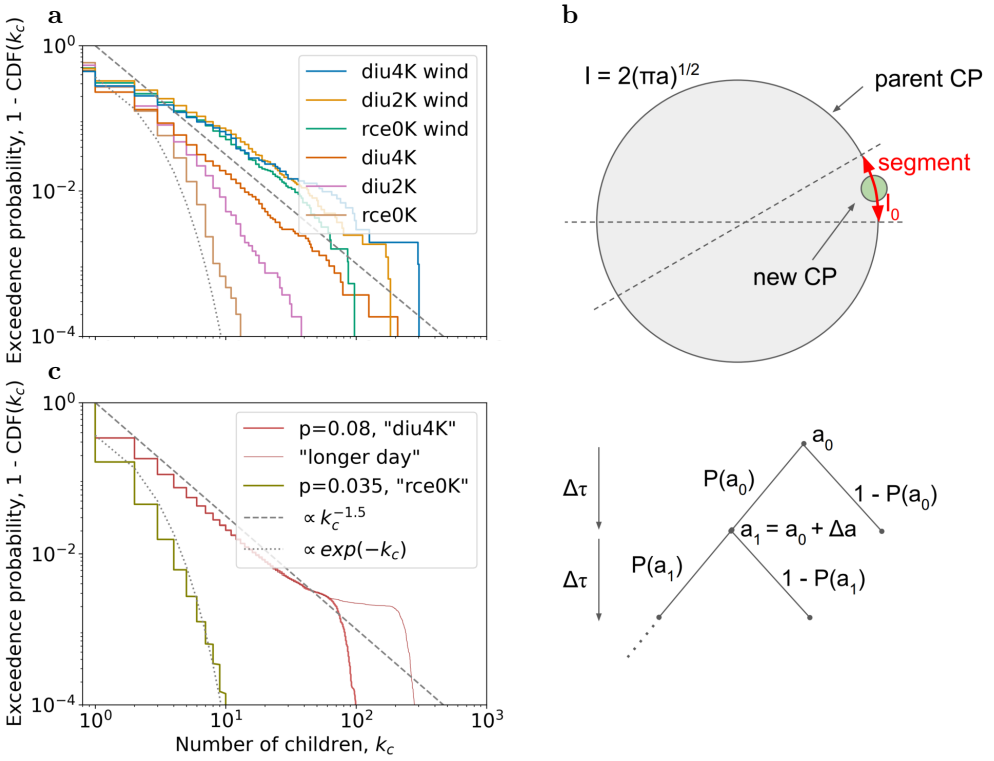

Figure 5: **Children probability distribution.** **a**, Exceedence probability for a cold pool to have a certain number of children; Colors indicate different simulation setups. **b**, Simple model for cold pool expansion and spawning of new rain cells. **c**, Analogous to (a) but for our simple model.

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 cold pool field from diu4K (Fig. 6a—c), along with the corresponding family trees of the two initial cold pools (Fig. 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 Fig. 6a, CP46 (yellow patch) and CP49 (blue patch), formed at t_{763} and t_{761}, respectively (Fig. 6d).

In Fig. 6b, 80min later, the two cold pools have expanded and several offspring have been generated. However, CP848 (red ID in Fig. 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 cold pool CP1169 (*see* Fig. 6d), it is thus not present anymore at t_{772}. Comparing Fig. 6b and c, 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 Fig. 6b we note that CP466 and CP835 formed before the collision of CP46 and CP49, otherwise both CP46 and CP49 would be their parents.

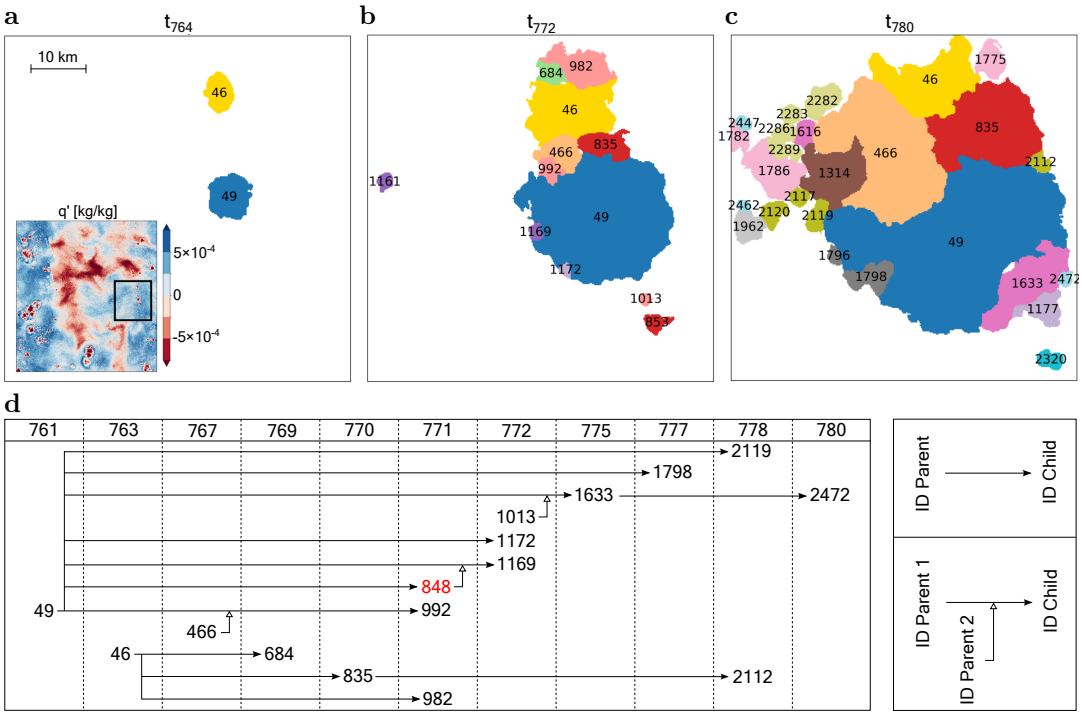

Figure 6: **Cascades of tracked cold pools. a**, Labeled cold pools for a subsection of the full simulation domain at time step index 764 of diu4K. The inset plot depicts full domain specific humidity anomalies at the same time step; The black square indicates the subsection of interest. **b**, Analogous 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 cold pool "46" and cold pool "49". The tree features the IDs of all related cold pools which are present in (a), (b) and (c); Columns represent the time steps indices of cold pool formation; Filled arrows point from the IDs of cold pool parents to the IDs of cold pool 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), (b) and (c), but which contributed to the related cold pools.

4 Discussion

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

& Bony, 2015; Haerter, 2019; Haerter et al., 2019; Nissen & Haerter, 2021) and the convective diurnal cycle (Schlemmer & Hohenegger, 2016; Böing et al., 2012; Böing, 2016; Haerter et al., 2019; Haerter & Schlemmer, 2018). 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 cold pool detection methods exist, which often make use of a buoyancy threshold (Tompkins, 2001; Feng et al., 2015; Gentine et al., 2016) 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 cold pool 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 cold pool "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 (Feng et al., 2015; Gentine et al., 2016; Fournier & Haerter, 2019; Henneberg et al., 2020; Drager & van den Heever, 2017), CoolDeTA cannot fully refrain from setting some threshold values. In the present work we apply a surface rainfall intensity threshold, $I_0 = 2\,mmh^{-1}$, to delineate spatially contiguous rain patches, and a corresponding area threshold, $A_0 = 2\,km^2$, for rain patches and potential cold pool regions. Thresholds are generally case dependent and thus limit a method's scope of application. Although particularly the cold pool 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 cold pool patches and do not impact the final cold pool contours: Since we provide each new rain patch above A_0 one marker in its downdraft center of mass, all of these rain patches have the same chance of forming a cold pool. The cold pool contour, and whether a rain patch forms a cold pool at all, depends on the result of the k-means algorithm. A sufficiently large rain patch and an associated marker are accordingly necessary, rather than sufficient, conditions with respect to the formation of a cold pool. Also the additional conditions summarized in Table 1 serve solely the purpose of optimizing the preselection in order to increase the robustness of the cold pool labeling. To verify this, we deactivated the dissipation threshold, n_{dis}, as well as all additional conditions and reproduced Fig. 2 based on this parameter setup in Fig. A2: whereas the deactivation causes essentially no change in the diu4K plots, 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 (Fig. A2). 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 cold pool signatures, the determined cold pool contours are consistent with observable cloud patterns. CoolDeTA thus further provides a systematic and objective so-called ground truth labeling for artificial intelligence methods that could detect cold pools from satellite-based observations (Hoeller et al., 2022).

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; Muller & Bony, 2015; Jensen et al., 2021). 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 mesoscale convective systems 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 cold pool 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 cold pool 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 cold pool detection methods, CoolDeTA offers the additional benefit of determining the actual dynamical gust front, which is often displaced by several kilometers from the thermodynamic boundary. In this way, the method is particularly well suited for the tracking of cold pool 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 thermodynamic gust front were available — many new rain cells are triggered at the dynamic gust front.

Data Availability Statement

Figures were made with Matplotlib version 3.5.2 (Hunter, 2007; Caswell et al., 2022) and seaborn version 0.12.2 (Waskom, 2021). The code for simple cold pool offspring model, as well as the cold pool detection and tracking algorithm (CoolDeTA) is licensed under MIT and published on GitHub `https://github.com/Shakiro7/coldPool-detection-and-tracking`. 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) (M. F. Khairoutdinov & Randall, 2003), version 6.11.

Acknowledgments
The authors gratefully acknowledge funding by a grant from the VILLUM Foundation (grant number: 13168) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant number: 771859) and the Novo Nordisk Foundation Interdisciplinary Synergy Program (grant no. NNF19OC0057374). This work used resources of the Deutsches Klimarechenzentrum (DKRZ), granted by its Scientific Steering Committee (WLA) under project ID bb1166.

References

Barnes, G. M., & Garstang, M. (1982). Subcloud layer energetics of precipitating convection. *Monthly Weather Review*, *110*(2), 102–117.

Böing, S. J. (2016). An object-based model for convective cold pool dynamics. *Mathematics of Climate and Weather Forecasting*, *2*(1).

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 en-

semble. *Journal of the Atmospheric Sciences*, *69*(9), 2682–2698. doi: 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.

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.

Caswell, T. A., Droettboom, M., Lee, A., de Andrade, E. S., Hoffmann, T., Klymak, J., ... Ivanov, P. (2022, May). *matplotlib/matplotlib: Rel: v3.5.2.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.6513224` doi: 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.

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., ... Zhang, M. (2006). The formulation and atmospheric simulation of the community atmosphere model version 3 (cam3). *Journal of Climate*, *19*(11), 2144 - 2161. Retrieved from `https://journals.ametsoc.org/view/journals/clim/19/11/jcli3760.1.xml` doi: 10.1175/JCLI3760.1

Dauhut, T., Couvreux, F., Bouniol, D., Beucher, F., Volkmer, L., Pörtge, V., ... others (2022). Flower trade-wind clouds are shallow mesoscale convective systems. *Quarterly Journal of the Royal Meteorological Society*.

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.

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.

Drager, A. J., & van den Heever, S. C. (2017). Characterizing convective cold pools. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1091–1115.

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. doi: 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.

Engerer, N. A., Stensrud, D. J., & Coniglio, M. C. (2008). Surface characteristics of observed cold pools. *Monthly Weather Review*, *136*(12), 4839–4849.

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. doi: 10.1002/2014MS000384

Feng, Z., Hagos, S., Rowe, A. K., Burleyson, C. D., Martini, M. N., & 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. doi: https://doi.org/10.1002/2014MS000384

Fiévet, R., Meyer, B., & Haerter, J. O. (2022). On the sensitivity of convective cold pools to mesh resolution. *Authorea Preprints*.

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

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.

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

Haerter, J. O. (2019). Convective self-aggregation as a cold pool-driven critical phenomenon. *Geophysical Research Letters*, *46*(7), 4017–4028. doi: 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. doi: https://doi.org/10.1029/2019GL082092

Haerter, J. O., Meyer, B., & Nissen, S. B. (2020). Diurnal self-aggregation. *npj Climate and Atmospheric Science*, *3*, 30.

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

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.

Hoeller, J., Fiévet, R., & Haerter, J. O. (2022). U-net segmentation for the detection of convective cold pools from cloud and rainfall fields. *Authorea Preprints*.

Houze Jr, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, *42*(4). doi: https://doi.org/10.1029/2004RG000150

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. doi: 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. doi: 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.

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

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the arm summer 1997 iop: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, *60*(4), 607–625.

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.

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.

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

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.

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

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

Muller, C., Yang, D., Craig, G., Cronin, T., Fildier, B., Haerter, J. O., ... others (2022). Spontaneous aggregation of convective storms. *Annual Review of Fluid Mechanics*, *54*, 133–157.

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.

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.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

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.

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*(7).

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.

Rooney, G. G. (2015). Descent and spread of negatively buoyant thermals. *J. Fluid Mech.*, *780*, 457–479. doi: 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*.

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.

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.

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.

Seifert, A., & Heus, T. (2013). Large-eddy simulation of organized precipitating trade wind cumulus clouds. *Atmospheric Chemistry and Physics*, *13*, 5631–5645.

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

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

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. the basic experiment. *Monthly Weather Review*, *91*(3), 99–164.

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.

Tompkins, A. M. (2001). Organization of tropical convection in low vertical wind shears: The role of cold pools. *Journal of the Atmospheric Sciences*, *58*(13), 1650–1672.

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

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

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., ... the scikit-image contributors (2014, 6). scikit-image: image processing in Python. *PeerJ*, *2*, e453. Retrieved from `https://doi.org/10.7717/peerj.453` doi: 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.

Wakimoto, R. (2001, 01). Convectively driven high wind events. In (p. 255-298). doi: 10.1007/978-1-935704-06-5_7

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. Retrieved from `https://doi.org/10.21105/joss.03021` doi: 10.21105/joss.03021

Weisman, M. L., Klemp, J. B., & Rotunno, R. (1988). Structure and evolution of numerically simulated squall lines. *Journal of Atmospheric Sciences*, *45*(14), 1990–2013.

Zuidema, P., Li, Z., Hill, R. J., Bariteau, L., Rilling, B., Fairall, C., ... Hare, J. (2012). On trade wind cumulus cold pools. *Journal of the Atmospheric Sciences*, *69*(1), 258–280.

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*, 1–23.

Appendix A Supplementary data

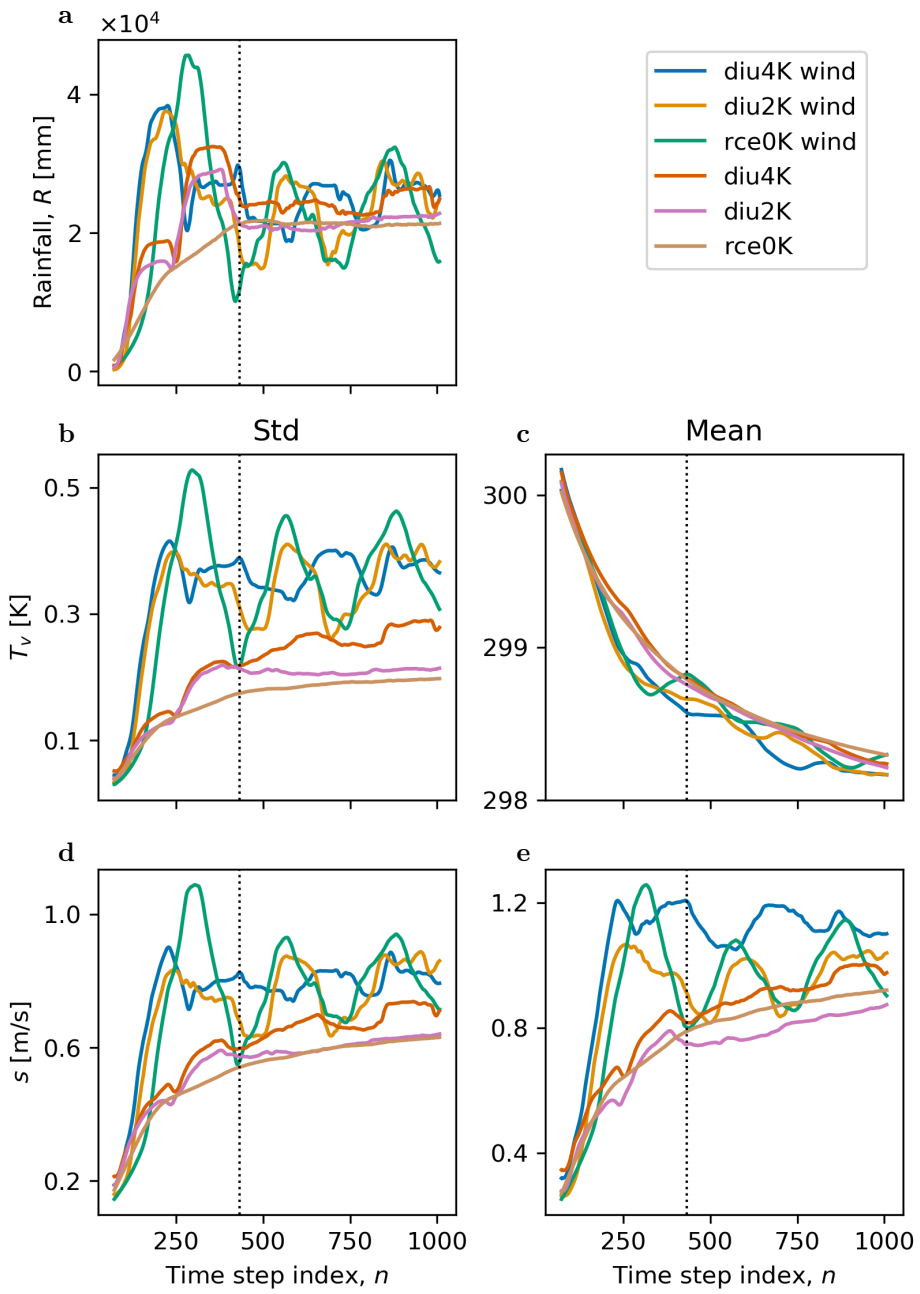

Figure A1: **Definition of spin-up time steps. a**, Total domain rainfall per time step as running average of one day (144 time steps) for different simulation setups; The dotted black line indicates the chosen spin-up period of three days (432 time steps). **b**, Analogous to (a) but for domain-standard 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.

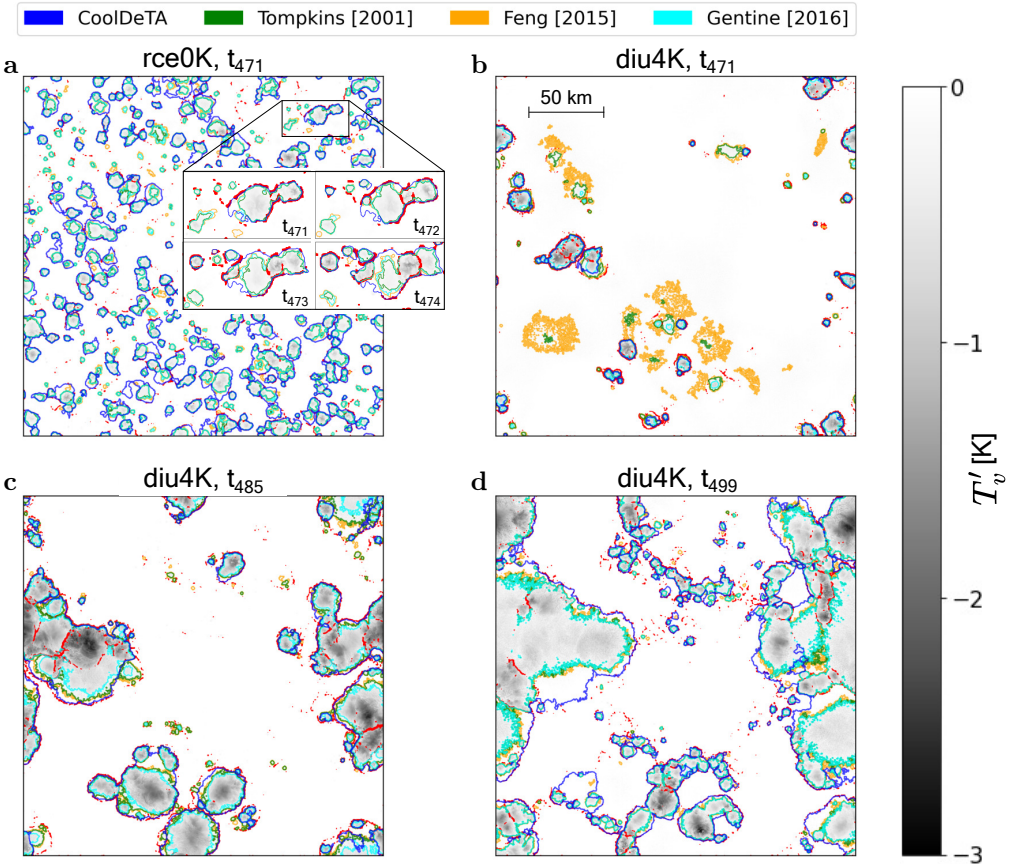

Figure A2: **Comparison of cold pool detection methods.** Analogous to 2a—d but with CoolDeTA contours based on a setup with $n_{dis} = 0$, i.e., without minimum dissipation time for cold pools, and without the additional rules listed in Table 1. Omitted rules comprise the checks of interior and boundary of "potential cold pool" regions with respect to their mean divergence, the required overlap between rain patches and "potential cold pool" regions to obtain a marker, and the required ratio between a newly formed cold pool and its rain patch.