

Master Thesis in Physics

Merge or die: Rain cell dynamics in the absence of cold pools

Author

Ronja Gronemeyer
Date of birth: 27.01.1995
ronja_gronemeyer@yahoo.de
Matriculation number: 4243654

Examiner

1. Prof. Dr. Stefan Bornholdt
2. Prof. Dr. Jan O. Härter

Date of submission: 20.07.2021

Diese Erklärungen sind in jedes Exemplar der Abschlussarbeit mit einzubinden.

Name: _____

Matrikel-Nr.: _____

Urheberrechtliche Erklärung

Erklärung gem. § 10 (10) Allgemeiner Teil der MPO vom 27.10.2010

Hiermit versichere ich, dass ich meine Masterarbeit ohne fremde Hilfe angefertigt habe, und dass ich keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich unter Angabe der Quellen als solche kenntlich gemacht.

Die Masterarbeit darf nach Abgabe nicht mehr verändert werden.

Datum: _____

Unterschrift: _____

Erklärung zur Veröffentlichung von Abschlussarbeiten

Bitte auswählen und ankreuzen:

- Ich bin damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.
- Ich bin damit einverstanden, dass meine Abschlussarbeit nach 30 Jahren (gem. §7 Abs.2 BremArchivG) im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.
- Ich bin nicht damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.

Datum: _____

Unterschrift: _____

Acknowledgements

Foremost, I want to thank Professor Jan O. Härter. His enthusiasm immediately caught me, when we first talked about complex systems and clouds, sitting outside in the sun. I am grateful for the guidance, inspiration, support and opportunities I received since then.

In addition, I want to thank the Atmospheric Complexity group at the Niels Bohr Institute for a warm (virtual) welcome, which allowed to work in a team despite the pandemic home office situation (and ~ 377 km distance). I am looking forward to coffee breaks in person!

Particular thanks to Gorm G. Jensen, volunteering to support me, getting started with Python and developing great ideas for the toy model presented in this thesis. I thank Professor Stefan Bornholdt for volunteering as my 'Bremen-examiner'.

Finally, I want to thank Arne Stefes, my best home office partner for support in so many regards.

1. Abstract

Cold pools (CPs) are associated with strong downdrafts that shut down rain events and are known to inhibit convective self aggregation (CSA) [JR13]. We run a mechanism denial experiment to investigate the role of CPs in CSA by suppressing the re-evaporation of precipitation (NoEvap) in radiative convective equilibrium (RCE) simulations of cloud resolving models (CRM) within a double-periodic square domains of linear horizontal dimension 480 km and 96 km with 1 km and 200 m resolution, respectively. Investigating the NoEvap dynamics is a meaningful idealisation for the formation of tropical cyclones, such as hurricanes. Tropical cyclones are linked to high boundary layer humidity thus weak sub-cloud re-evaporation and tropical cyclones are suggested to originate above the tropical sea [Pal48]. CSA is hampered when CPs are present (FullEvap) and in the NoEvap case can be tracked as the monotonous decrease of convective cells, which decay or pairwise merge. We determine convective cells and their inflow areas, termed 'basins of attraction', that partition the domain, and track how the cells and basins dynamically evolve.

To derive a compressed and discrete representation, each convective updraught is characterised by its horizontal location and the area of its basin. With these three numbers we initialise our model that captures updraughts as point-like objects and derives the resulting horizontal velocity field via a Poisson equation, which we solve analytically based on a Fourier-expansion. To develop a dynamical model, we introduce a gravity-like approach to predict how updraughts move horizontally and successfully calculate positive correlations to our data. Possible physical explanations for the decaying-merging dynamic and observed oscillations in rainfall are discussed, and will be further investigated.

2. Glossary

Acronym	Definition
LES	Large Eddy Simulations
CRM	Cloud Resolving Model
CP	Cold Pool
NoEvap	Simulations with fully suppressed re-evaporation of rainfall. Consequently no CPs form.
FullEvap	Simulations with full realistic re-evaporation of rainfall, which is used as a control simulation where CPs are present.
Evap01	Simulations where the corresponding ventilation coefficient F_v in the Seifert scheme [SB06] is 0.1 times the realistic factor chosen for the FullEvap simulation. Simulations with Evap02 and Evap06 are defined analogously with 0.2 and 0.6 times the realistic factor.
CSA	Convective Self Aggregation
OLR	Outgoing Longwave Radiation in a general context, without concrete reference to its temporal development or its specific height. Consistent notation is used for other variables, where they are used in a more general context.
$\overline{\text{OLR}(t)}_z$	Domain averaged OLR, for a specific layer height z : $\frac{1}{A} \int dx \int dy \text{OLR}(x,y,t)_z$. Consistent notation is also used for other domain averaged physical variables, for example if their temporal development is discussed.
CRH	Column Relative Humidity, associated with the parameter r in models of CSA (introduced in Sec. 4.3)
RCE	Radiative Convective Equilibrium
CAPE	Convective Available Potential Energy
Test parcel	Concept used to calculate CAPE and CIN. A hypothetical lifted idealised air parcel, which does not exchange heat or gas with its environment. A common simplification used in this thesis is, that condensed liquid water can be removed from the test parcel, which describes an irreversible moist adiabatic ascent.
Convective region	Region with positive $\text{CAPE} > 0$, where convection takes place.

CIN	Convective Inhibition
LFC	Level of Free Convection
LNB	Level of Neutral Buoyancy
SST	Sea Surface Temperature
Mesoscale	Horizontal scales in the atmosphere, from a few to several hundred kilometres
Deep convection	We refer to deep convection as convection where a large fraction of the moisture condenses in the test parcel, so the latent heat release provides additional buoyancy and convection reaches high altitudes. In literature the term can be used ambiguously [Ste05].
Final state	Stage of the aggregated (NoEvap) simulation, when only one convective cell is left, with the entire domain as its basin.

Table 2.1.: Acronyms in the context of atmospheric convection and simulations used in this thesis.

2. Glossary

Term	Definition
Active pixel	Any simulation pixel (x, y) where $I(x, y, t) > I_0$, that is, where the rain intensity at the lowest model level (<i>rint</i>) exceeds the threshold I_0 . The threshold depends on the simulation under consideration and is attuned to the applied filtering.
Rain patch	Group of spatially linked active pixels at a specific time t_k (temporal simulation 'snapshot'), that fulfill a condition: (e.g. $I(x_i, y_j, t_k) > I_0 \wedge I(x_{i+1}, y_j, t_k) > I_0$). We consider a 4-pixel neighbourhood for each pixel (x_i, y_j) : $(x_{i+1}, y_j), (x_{i-1}, y_j), (x_i, y_{j+1}), (x_i, y_{j-1})$. Rain patches are later linked over time to a track of temporal persistent rain clusters.
Rain cell	Loosely defined as the joint dynamical object consisting of low level convergence, the associated convective updraft, as well as cloud and precipitation formation. We define the 'live span' of a rain cell as the time period when rain is falling and track rain cells with our cluster algorithm. Notice, that one rain cell corresponds to one convective event, or one updraft detected with the basin approach.
Cluster	Technical term for rain cell, used in the context of the rain tracking algorithm (Sec. 5.3). A rain cluster is defined as consecutive rain patches with spatial overlap: $\exists(x_i, y_j) : I(x_i, y_j, t_k) > I_0 \wedge I(x_i, y_j, t_{k+1}) > I_0$. Clusters have a temporal start and an end point, as the detected rain cell emerges and dies/merges. A cluster object has a unique label m , a start and end time (t_{start}^m and t_{end}^m) and a list of corresponding pixels (the patches) for each $t_{start}^m < t_k < t_{end}^m$.
Decay	Fading of a rain cell, without directly noticeable interaction with other rain cells. Dying is observed as a slow decrease of rain in the context of rain clusters ($I(\text{cluster}, t) \rightarrow 0$) and decreasing convergence towards the corresponding updraft location in the context of the basin approach.
Merging	Two rain cells, fuse to one bigger event. Merging is mainly present in the NoEvap simulations. Here, two rain cells usually move towards each other before they merge, a dynamic well tracked with both algorithms.
Clusterstore	All clusters are organised in a Clusterstore: An object class, that stores 'global' information for each simulated time-step t_i , such as all active clusters and the number of merges or decays.
Final stage	Fully aggregated state, when only one convective cell/rain cell/updraught is left.
Transient stage	The temporal approach to the fully-aggregated final stage.

Convective cell	Similar to rain cell, but more focused on the convective updraft with low level convergence towards it than on the associated cloud and precipitation formation. Convective cell is the preferred term in the context of the basin approach which does not consider rainfall, but inflow areas given by convergence of horizontal wind speed and thus includes larger areas. For the NoEvap case rain usually stops several hours before the convergence towards the corresponding updraft does, thus results differ between tracking rain cells and tracking convective cells.
Basin	Basin of attraction, defined as the inflow area of one convective cell or updraught.

Table 2.2.: Acronyms and terms defined in the context of the two tracking algorithms developed in this thesis (chapter 5)

2. Glossary

Rain variable	Definition
$A(\text{cluster},t)$	Area belonging to a cluster at a given time t , corresponding to the number of pixels detected 'active' in that cluster.
$I(x,y,t)$	LES/CRM output variable 'r_int' of rain intensity at the lowest model level. $I(x,y,t)$ roughly represents surface rainfall and assigns an instantaneous rain value every 5 h of the simulation to each pixel in the two dimensional horizontal field. The value for $I(x,y,t)$ is given in mm/h, since the instantaneous rain volume is projected to one hour under the hypothetical assumption that 'r_int' at the pixel under consideration would stay constant.
$\overline{I(t)}$	Spatial domain average of $I(x,y,t)$, given by the integral over the domain, divided by the area: $\frac{1}{A} \int dx \int dy I(x,y,t)$. $\overline{I(t)}$ can be thought of as a spatial rain density.
$I_{\text{total}}(t)$	$I(x,y,t)$ integrated over the entire domain: $\int dx \int dy I(x,y,t)$ $I_{\text{total}}(t)$ can be thought of as a summed domain summed rain intensity, limited by the global radiation constraint (Sec. 4.1.5).
$I(\text{cluster},t)$	Rain falling inside a rain cluster for a given time in the simulation. $I(\text{cluster},t)$ corresponds to 'r_int' integrated over the active pixels belonging to this cluster and can be thought of as the current total rain intensity of a cluster: $\int_{A_{\text{cluster}}} I(x,y,z) dA_{\text{cluster}}$.
$\overline{I(\text{cluster},t)}$	Mean rain intensity averaged over all pixels of a cluster at a given time t .

Table 2.3.: Acronyms and terms defined to describe precipitation and rain clusters (chapter 6). Notice that time and space dimension are actually discrete since numerical simulations are used. Nevertheless, the integral form allows to give a very specific definition of what is meant by the different rain variables.

Term	Definition
Lowest layer	Lowest horizontal output level (LEV2) of our simulations, with height $z \approx 50m$.
W	Idealised vertical velocities, source function in the Poisson equation.
Updraught	Convective updraught, spatially extended area with clear positive vertical velocity $w_{z=50m} > 0$ in the lowest layer.
Sink	Due to mass conservation, updraughts are sinks in the horizontal velocity field of the lowest layer ($z \approx 50m$). These sinks' locations are inserted in our model reconstruction.
Updraught-sink	To prevent misconceptions and take the dual role of convective updraughts into account, often the term updraught-sink is used in the model context.
Model speed	Absolute value of reconstructed horizontal velocity vector. Although its magnitude is not scaled to correspond to the data, it captures the spatial changes in wind speed well.
$N(t)_{cluster}$	Current number of detected rain cluster.
$N(t)_{basin}$	Current number of detected basins.
$A(t)_{basin}$	Area of an individual basin at a given time t .
$\overline{A(t)_{basin}}$	Mean basin area $A(t)_{basin}$ at a given time t , calculated $1/N(t)_{basin}$.
Coarsening	Coarsening processes typically appear in systems with two distinct phases, in our case moist convective regions and dry, subsidence dominated regions. In our case the separation increases over time, towards the fully aggregated equilibrium state (finale stage). The average 'size' or 'area' of a component reflects the current state and typically the length scale increases as a simple function of time. Often coarsening rates can be described as power laws and typical examples are water droplets on a thin precursor layer or the demixing binary viscous liquids [BOS11].
Increase in scale	Increase of the scale under consideration, for example increasing length scale $\frac{dLS}{dt} > 0$ or area scale $\frac{dAS}{dt} > 0$.
LS	Length scale, in the context of dynamical coarsening.
AS	Area scale, in the context of dynamical coarsening.

Table 2.4.: Acronyms and terms defined in the context of the model reconstruction (chapter 7), the related 'basin approach' (Sec. 5.5, 6.4) and the coarsening process (Sec. 6.5)

Contents

1. Abstract	1
2. Glossary	2
3. Introduction	11
4. Background	14
4.1. Theoretical background	14
4.1.1. Atmospheric convection	14
4.1.2. CAPE in the tropical atmosphere	19
4.1.3. Moist Static Energy (MSE) and Frozen Moist Static Energy (FMSE)	20
4.1.4. Large Eddy Resolving Simulations (LES) and Cloud Resolving Models (CRM)	20
4.1.5. Radiative Convective Equilibrium (RCE)	21
4.1.6. Weak temperature gradient approximation	22
4.1.7. Convective cold pools	22
4.2. Convective Self Aggregation (CSA)	22
4.2.1. CSA and domain size	23
4.2.2. Quantifying CSA	23
4.2.3. Drivers of CSA	24
4.2.4. CSA and CPs	25
4.3. Models of convective self aggregation	26
4.3.1. Bretherton, Blossey, Khairoutdinov, 2005	26
4.3.2. Wing and Emanuel, 2014 and Emanuel, Wing, and Vincent, 2014	29
4.3.3. Craig and Mack, 2013	32
4.3.4. Bi-stability and spatial evolution of CSA	34
5. Methods to detect convective events	37
5.1. Large-eddy (LES) and cloud-resolving model (CRM) simulations used	37
5.2. Detection of convective events	38
5.2.1. Background on cloud tracking	38
5.2.2. Tracking individual rain clusters	39
5.2.3. Motivation of tracking algorithm, based on horizontal velocity field	40

5.3.	Implementation of a rain cluster tracking algorithm	40
5.3.1.	Preprocessing	40
5.3.2.	Structuring	42
5.3.3.	Discussion of the rain tracking algorithm	44
5.4.	Calculating (CAPE) to detect regions where convection can possibly take place	48
5.5.	Tracking of convective events based on the velocity field	50
5.5.1.	Basins of attraction	50
5.5.2.	Determine basins of attraction as inflow areas of convective events	52
5.5.3.	Discussion of the basin approach	54
6.	Data analysis	56
6.0.1.	Reducing the re-evaporation of rain	56
6.1.	Analysis of aggregation dynamics based on rain tracking	57
6.1.1.	Dynamics with and without cold pools (FullEvap versus No- Evap)	59
6.1.2.	Merging	61
6.1.3.	Resolution	61
6.2.	Oscillations	61
6.3.	CAPE	64
6.4.	The low layer velocity field	67
6.4.1.	Analysis of the stationary flow field	67
6.4.2.	Development of basin areas over time	70
6.4.3.	Discussing possible advantages of larger updraughts	70
6.4.4.	Lateral updraughts dynamics	71
6.5.	Dynamical scaling	71
7.	Model	75
7.1.	Theoretical basis	75
7.1.1.	Motivation of the model	75
7.1.2.	Define Poisson-equation which determines the horizontal flow field	77
7.2.	Solving the Poisson equation	79
7.2.1.	Deducing the solution of the Poisson equation	80
7.2.2.	Realisation of the model	81
7.2.3.	Proofing the scaling of basins areas with updraught strength	83
7.3.	Implementation of the model	84
7.3.1.	Choosing parameters of the toymodel	84
7.3.2.	Performance of the basin reconstruction	84
7.3.3.	Discussing the model	86
7.4.	Towards a dynamical model	89
7.4.1.	Performance of the update dynamic	90

Contents

7.5. Contribution of our model to the state of the art	93
8. Discussion	94
8.1. Discussion of possible physical mechanism	94
8.1.1. Transient stage	94
8.1.2. Cloud effects	96
8.1.3. Synchronised oscillations in the final aggregated stage	99
8.1.4. Concluding hypotheses on physical feedbacks possibly contributing to CSA and observed oscillations	99
8.2. Interpretations as convective oscillators	100
8.3. Two coupled oscillators	102
8.4. Conclusion	104
8.5. Outlook	105
A. Appendix	115
A.1. Methods supplements	115
A.1.1. TOBAC tracking algorithm, applicable to a wide variety of situations	115
A.2. Results supplements	116
A.3. Model supplements	116
A.3.1. Scale Analysis, estimating the Reynolds number	116
A.3.2. Fourier expansion of the source function	117
A.3.3. Implementation: Deriving the gradient velocity field from a given updraught-sink distribution	120
A.3.4. Proof: scaling of $A_{basin_i}^{model}$ with strength parameter S_i	123
A.4. Discussion supplements	123
A.4.1. Multiple coupled oscillators	123

3. Introduction

Tropical thunderstorms and tropical cyclones can cause devastating winds, flooding and extreme precipitation, strongly impacting local human communities. Rainfall and storminess are primarily associated with convective clouds, which, under certain conditions, organise into larger clusters that are associated with extreme weather¹. Clustering of clouds is also observed in simulations², for example when the simulated atmosphere spontaneously divides into persistently cloudy, that is convecting, and cloud-free, subsiding, subregions. This phenomenon, known as convective self-aggregation (CSA), is characterised by dynamics of a gradually decreasing number of moist convective regions to a state where rain is finally limited to a single moist patch in the aggregated state.

Apart from spatially re-organising convection, CSA influences the mean state of the atmosphere, such as the radiative budget. Such effects make CSA relevant for the larger scale dynamics and, since it is supposed to play a role in the real atmosphere [Win+17], crucial to consider in the context of climate change. CSA is found to lead to a drastic decrease of domain mean humidity (net drying), accompanied with a warming of the troposphere and a decrease of the total area covered by clouds. Some studies suppose this could lead to a dangerous positive feedback loop, especially since higher temperatures can facilitate CSA [Ema+94], but other studies suggest the rather opposite [Win+18]. Therefore, the radiation feedback associated with CSA is an ongoing research topic, to explore how CSA would influence global warming and how warming, e.g. increased sea surface temperatures, in turn could feed back on CSA, with possible implications for the frequency and intensity of extreme events in the tropics.

In conclusion, understanding CSA is important but multi-faceted, since convective dynamics are complex and clouds organise on a wide range of interacting scales. But reducing complexity can open up for understanding: The comparably simple framework of radiative convective equilibrium (RCE), combined with Large Eddy Simulations (LES) and Cloud Resolving Models (CRM) allows us to investigate the development of CSA at the level of individual rain cells. We even simplify these simulations further, when we remove the inhibiting effect of cold pools (CPs) on CSA. CPs are strong downdrafts that form underneath precipitating clouds and

¹https://isccp.giss.nasa.gov/role.html#SYSTEM_FEEDBACK

²Referring to a variety of simulations ranging from fine resolution Large-Eddy-Simulations (LES), to cloud resolving models (CRM) and Global Circulation Models (GCMs)

3. Introduction

shut down rainfall. In order to remove CPs, the re-evaporation of precipitation is disabled, which enables very persistent rain cells.

We track these rain cells with a new rain tracking algorithm. Many variables describe the state of the atmosphere, but precipitation is the quantity that most directly affects the living environment, which motivated our choice. Moreover, surface rain has practical advantages, for example it can be delimited to discrete rain patches by applying a threshold. Investigating the NoEvap dynamics could yield further insight into the formation of tropical cyclones, as hurricanes (Atlantic) and typhoons (Pacific). These often devastating storms are assumed to originate above the tropical sea [Pal48], where the boundary layer is often very moist and therefore the re-evaporation of precipitation is weak [Gra+20]. Since humidity is high inside tropical storms, investigating convective dynamics with weak or without (NoEvap) re-evaporation of precipitation could be crucial to understanding their formation and development.

Our mechanism denial experiment 'NoEvap' further allows to investigate the role of CPs in the formation of CSA by contrasting the generally artificial NoEvap simulations to simulations with re-evaporation and thus CPs (FullEvap). Whereas convection does not change significantly over our simulations with CPs, strong aggregation appears in their absence. This shows as a monotonous decrease of the number of rain cells by two processes, which we term 'decaying' and 'merging'. 'Decaying' is a gradual decrease in surface rain intensity of rain cells, whereas merging means two rain cells spatially connect and become a larger one. This dynamics remind of a competition, where some rain cells grow and merge and others decay until a single 'winner takes it all.'

To capture this dynamics, we aim to reduce the complexity further with a second approach. We claim that simplified models often are very universal and can capture similar behaviour in different fields. The concept we develop is based on a Poisson equation, commonly used in electrostatics or gravitational physics to derive a field from a given source distribution. We use the Poisson equation to derive the low-level horizontal velocity field from a given distribution of 'point-like' convective updraughts, which we extract from the simulation data. Each convective updraught object is described by the horizontal centre-of-mass coordinates and a 'strength factor,' which scales with its inflow area. To extract these two pieces of information from the data we develop a second tracking algorithm based on the horizontal velocity field. Since updraughts are sinks in the horizontal velocity field, we can follow streamlines of instantaneous fields to detect centres of horizontal convergence, hence updraught locations. We refer to the area from where the flow converges towards such an updraught-sink, as its 'basin of attraction.' Next, we find a compressed reconstruction of the instantaneous flow field, describing each updraught sink by three numbers only. Further we explore how these reconstructions could be expanded to a dynamical model that captures the merging dynamics, dominating the aggregation process in our No-Evap simulations.

Recently, the hypothesis caught attention that the real atmosphere too could be intrinsically drawn towards a stage of aggregation [WH19]. One possible explanation, why this aggregated equilibrium might never be reached in nature could be mechanisms working against CSA, as for example CPs. Taking all these into account, this thesis is motivated to investigate mechanisms working towards and against CSA and finding new condensed representations of key features.

4. Background

This chapter provides useful background to follow the argument of the thesis and to embed this work into the state of the art. After summarising basic theoretical concepts of atmospheric science (Sec. 4.1) we introduce 'Convective Self-Aggregation' (CSA) (Sec. 4.2). Subsequently, established models of CSA are discussed and linked to our work (Sec. 4.3).

4.1. Theoretical background

We revisit concepts of atmospheric convection: Sec. 4.1.1, captures the specifics of Convective Available Potential Energy (CAPE) in 6.3 and explains Moist Static Energy (MSE) and Frozen Moist Static Energy (FMSE) (Sec. 4.1.3) as concepts often used to measure CSA [WE14]. The framework used in this work is outlined by introducing Large Eddy Resolving simulations (LES) and Cloud Resolving Models (CRM) (Sec. 4.1.4), the Radiative Convective Equilibrium (RCE) and the weak temperature gradient approximation (Sec. 4.1.6). Finally, cold pools (CPs) are introduced (Sec. 4.1.7) as the phenomena whose role in CSA we aim to investigate.

4.1.1. Atmospheric convection

In order to investigate convectonal organisation it is important to review its basic dynamics.

In fluid dynamics, convection is the principle of energy transfer where mass transport enables heat exchange. A simple example is a fluid, captured by a plate of higher temperature at the bottom T_{bottom} and a plate of lower temperature T_{top} at the top, causing a temperature gradient ∇T aligned with the gravity vector in the fluid. If $\nabla T \approx T_{\text{top}} - T_{\text{bottom}}$ inside the fluid exceeds a critical value, the fixed point of motionless fluid becomes unstable and heat is no longer transported by diffusion only (heat conduction) but also by (thermal) convection. This convective motion strongly accelerates the heat transport from the warmer bottom plate to the colder upper plate. The characteristics of a fluid's convective flows are associated with the Rayleigh number (Ra). Below a critical value, motionless heat conduction dominates the flow rather than convection. Ra is defined as the ratio between the timescales for thermal transport via diffusion $t_{\text{therm. diffusion}}$ and via convection at a certain speed w , $t_{\text{therm. convection}}^w$. We assume a cubic fluid volume with side lengths L between

4.1. Theoretical background

the plates¹ and with mass density difference $\Delta\rho = \rho_{\text{top}} - \rho_{\text{bottom}} > 0$ caused by the temperature difference. With an acceleration g due to gravity, the force acting on such a cube can be approximated by $\Delta\rho l^3 g$. The viscous drag inhibiting the motion is in the order $\nu L w$, where ν is the fluid's kinematic viscosity. Assuming these forces equal $\Delta\rho l^3 g = \nu L w$, allows to approximate $w \approx \Delta\rho L^2 g / \nu$. With this the ratio of timescales of convective motion $t_{\text{therm. convection}}^w = L/w$ can be approximated as $L/w \approx \nu / \Delta\rho L g$. For thermal diffusion along the distance L the timescale is L^2/α , with α thermal diffusivity. Together, this combines to the Rayleigh number:

$$Ra = \frac{t_{\text{therm. diffusion}}}{t_{\text{therm. convection}}^w} \approx \frac{L^2/\alpha}{\nu/\Delta\rho L g} = \frac{L^3/\alpha}{\nu/\Delta\rho L g}. \quad (4.1)$$

Whereas the Rayleigh number estimates whether a fluid's heat transfer is dominated by convection or by diffusive heat conduction, the Reynolds number estimates the turbulence in a flow. The Reynolds number divides the product of typical fluid velocity U and a typical length scale \tilde{L} by the dynamic viscosity ν :

$$Re = \frac{U \cdot \tilde{L}}{\nu}. \quad (4.2)$$

We use the Reynolds number to estimate the diffusive scale in the horizontal flow, when introducing our model (Sec. 7.1.2).

In atmospheric science, the term 'convection' usually refers to free convection, a vertical upward movement caused by density differences, hence excluding forced convection [Ema+94]. 'Free convection' refers to an unstable vertical mass distribution which causes buoyant air to rise upward against the gravitational field. It can be further divided into two distinct dynamics: Dry convection (Sec. 4.1.1) and moist convection (Sec. 4.1.1). Dry convection is present in the sub-saturated atmospheric boundary layer (BL), the bottom of the troposphere, which is in contact with the earth's surface. It is often turbulent, e.g. due to surface drag, and therefore well-mixed. Moist convection is usually associated with deep convection, which reaches higher levels in the atmosphere.

Stable and unstable conditions

To test whether idealised² stratification conditions are stable or not, one hypothetically lifts a test parcel for an infinitesimal distance δz in an irreversible moist-adiabatically process. Such a process is irreversible, since liquid water that would condense when the test parcel rises, is assumed to be removed by precipitation as soon as it is formed³.

¹with L the distance between the plates

²neglecting friction

³https://glossary.ametsoc.org/wiki/Pseudoadiabatic_process

4. Background

Assuming adiabatic cooling over the ascent, it has to be evaluated whether the parcel's density would be lower than that of the surrounding atmosphere after the hypothetical lifting. If so, the parcel would continue to rise and the stratification is unstable. Otherwise, stratification is stable and the cooler parcel would be accelerated downwards and due to an overshoot of its hydrostatic equilibrium start to oscillate with the Brunt-Väisälä frequency N [SA95].

$$N = \sqrt{\frac{-gd\rho}{dz}}$$

N depends on the acceleration caused by gravity g and the change of density ρ with height. For this reason, it is also referred to as 'buoyancy frequency'. This frequency is the upper frequency limit for gravity waves, internal atmospheric waves, in which buoyancy acts as the restoring force. Gravity waves converge between potential and kinetic energy and can transfer energy between different convective events with speeds up to several m/s [Tsu14].

Dry convection

We refer to dry convection as buoyancy-driven transport where condensation can be neglected, which assumes an adiabatic ascent in an atmosphere totally determined by its temperature field $T(x, y, z, t)$ [Ste05]. Buoyancy b can be described using the potential temperature θ : $b \approx g \frac{\delta\theta}{\theta_0}$. Potential temperature (Eq. 4.3) is a conserved quantity in dry convection and allows to compare air masses at different heights, since it is the temperature a parcel would have if brought adiabatically, or isentropically, to a standard pressure p_0 . With R as the universal gas constant for dry air and c_p as the heat capacity at constant pressure, potential temperature is defined as:

$$\theta = T \left(\frac{p_0}{p} \right)^{R/c_p} \quad (4.3)$$

For a given stratification, the vertical gradient of θ defines whether conditions are:

- stable: $\frac{\partial\theta}{\partial z} > 0$,
- neutral: $\frac{\partial\theta}{\partial z} = 0$,
- unstable: $\frac{\partial\theta}{\partial z} < 0$.

Usually, convection drives towards a neutrally stratified state, where the dry entropy is at its maximum (compare Sec. 4.1.2).

The dry adiabatic lapse rate is defined as the vertical temperature gradient assuming a dry adiabatic lifting [Ran09]:

$$\Gamma_d = \left(\frac{\partial T}{\partial z} \right)_{\text{dry adiabatic}} \quad (4.4)$$

Moist convection

Moist convection is associated with deep convection, when moist air rises high enough to fall below saturation temperature. Then the specific humidity q_v , the mass of water vapour in a mass unit of moist air, reaches the saturation point q_S and water starts to condensate. Condensation releases latent heat which provides extra buoyancy to the water vapour test parcel. This can accelerate a saturated air mass further upward. Since the ongoing latent heat release partly compensates the cooling, the parcel can be lifted a finite distance.

The case of moist adiabatic lifting is often referred to as conditionally unstable, since stability is lost when saturation is reached.

One can imagine a parcel to be trapped in the basin of an unstable fixed point. It only escapes, when enough energy is provided to overcome an energy barrier called convective inhibition (CIN). To overcome CIN, a parcel needs to be raised to a height where it is warmer than its environment, so it gains positive buoyancy. This height is defined as the level of free convection (LFC, Fig. 4.1). Below the LFC, a small perturbation fades and the parcel is pulled back to the local potential minimum. At the LFC, the fixed point becomes unstable and an infinitesimal lifting accelerates the parcel further upwards, until it reaches a stable fixed point: the level of neutral buoyancy (LNB, Fig. 4.1). The LNB is the maximum height the test parcel can reach by convection, since all water is condensed there so the air within the test parcel becomes very dry.

The difference between the potential at LFC and the potential at LNB, is called the convective available potential energy (CAPE, Fig. 4.1). When a parcel overcomes the threshold, given by CIN, CAPE can be converted to kinetic energy and induces the upward movement of a parcel by free convection. A high CAPE indicates regions where deep convection is possible. We refer to deep convection as convection where a large fraction of the moisture condenses in the test parcel, so the latent heat release provides additional buoyancy and convection reaches high altitudes. CAPE will be used to contrast convective dynamics in Sec. 6.3.

Liquid potential temperature

To be precise, the temperature output used in the LES is liquid potential temperature θ_l , which is the temperature a test parcel would have if brought adiabatically to a standard reference pressure⁴, with all water converted to the liquid state [Hol73]. It is linked to temperature T and potential temperature θ as follows:

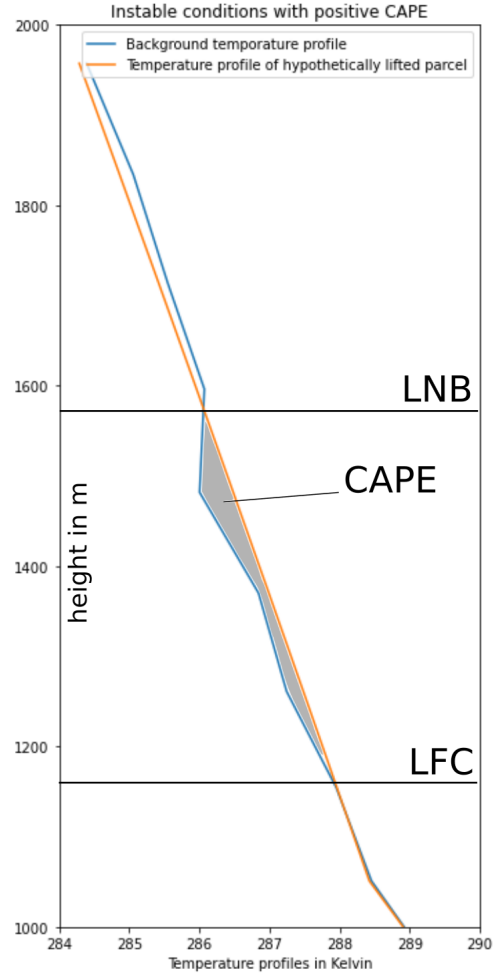
$$\theta_e \approx \theta \cdot e^{\frac{q_l L_v}{c_p^d T}} \quad (4.5)$$

With q_l as liquid water mixing ratio, L_v the latent heat of evaporation and c_p^d the specific heat of dry air. When calculating CAPE and CIN this conversion has to be

⁴usually atmospheric pressure at sea, roughly 100 000 Pa

4. Background

Figure 4.1: CAPE for unstable conditions. Unstable conditions in the atmosphere correspond to positive CAPE (grey shading) and facilitate convection. CAPE is calculated as the difference between the virtual temperature profiles of an adiabatically-lifted test parcel (orange) and the background profile of the environment (blue), vertically integrated from the LFC to the LNB.



taken into account.

Calculating CAPE and CIN

Buoyancy depends on the composition of an air mass, because for a given temperature, moist air is less dense than dry air⁵. The virtual temperature T_v (4.6) is defined as the temperature dry air (H_2O replaced by other gas molecules) would have at the density of a specific composition. It is proportional to the temperature T and the ratio of the gas constant of this composition R , which comprises the gas constant of water vapour R_v and dry air R_d , multiplied with the corresponding mass fractions q_v and q_d :

$$T_v = T \frac{R}{R_d}, \quad R = R_d \cdot q_d + R_v \cdot q_v. \quad (4.6)$$

⁵Moist air is less dense, because the molar mass of H_2O is lower compared the other air components as N_2 and O_2 , but due to equipartition of energy the pressure each molecule exerts is the same

For moist air the composition changes over a hypothetical lifting (Sec. 4.1.1) and buoyancy can only be linearly approximated piece-wise: for the saturated and unsaturated case.

Due to latent heat release, temperature decrease with height is reduced in the moist adiabatic case. Therefore, the moist adiabatic lapse rate $\Gamma_{\text{moist adiabatic}}$, which we consider in our CAPE calculation (Sec. 5.4) is smaller than $\Gamma_{\text{dry adiabatic}}$. $\Gamma_{\text{dry}} > \Gamma_{\text{moist}}$ corresponds to the changing slope in the test parcel's profile (Fig. 4.1) when condensation starts.

CAPE is the difference between the virtual temperature profiles of the test parcel T_v^{parcel} and the background profile of the environment $T_v^{\text{env.}}$ integrated from the LFC to the LNB :

$$\text{CAPE} = \int_{\text{LFC}}^{\text{LNB}} g \left(\frac{T_v^{\text{parcel}} - T_v^{\text{env.}}}{T_v^{\text{env.}}} \right) dz . \quad (4.7)$$

Since positive CAPE is the buoyant energy available to lift an moist air mass adiabatically, CAPE provides information where convection can possibly take place.

4.1.2. CAPE in the tropical atmosphere

In most of the tropics, the population of cumulus clouds is high and convection often develops quite spontaneously, linked to low LFCs. This limits the amount of CAPE building up (Fig. 6.7) until being released as kinetic energy. Despite that, regions of large CAPE do develop, too. Seeking physical explanations, Xu and Emanuel observe buoyancy profiles of cumulus clouds, which are known to be extremely in-homogeneous.

Before, Betts (1982) observed that the tropical atmosphere in deep convective regions is nearly neutral to the moist adiabatic ascent of an undiluted subcloud layer parcel. Xu and Emanuel systematically examined the stability of the western pacific atmosphere and suggest that the atmosphere organises along a moist adiabatic lapse rate.

The buoyancy of an air parcel depends on Θ_e , which is the temperature the parcel would reach, if brought adiabatically to a standard reference pressure, whereby all contained water vapour would condense and release its latent heat. Xu and Emanuel suggest that the 'well mixed' boundary layer is not really mixed in terms of equivalent potential temperature Θ_e , which makes buoyancy crucially depending on the height a parcel originates from and on the cumulus cloud properties. They argue, that in such a conditionally neutral atmosphere, it is problematic to assume vertical motion to be generated by latent heat release and suggest large scale disturbances (as for tropical cyclones, [Ema86], [RE87]) could rather be linked to feedback mechanisms in the subcloud layer.

Concludingly, Xu and Emanuel state that the tropical atmosphere is neutral to non-precipitating clouds, but do not answer the question whether or why the tropical

4. Background

atmosphere is conditionally stable. We will refer to their hypothesis when discussing CAPE in our NoEvap simulations (Sec. 6.3, Fig. 6.7).

4.1.3. Moist Static Energy (MSE) and Frozen Moist Static Energy (FMSE)

MSE is calculated analogously to equivalent potential temperature Θ_e (Eq. 4.5), by assuming a hypothetical adiabatic lifting of an air parcel, but assumes that any kinetic energy is locally dissipated into heat⁶. MSE is approximately conserved in adiabatic ascent or descent and combines the enthalpy $c_p T$ ⁷, potential energy $g \cdot z$ and the energy of potential latent heat release $L_v \cdot q$ [KS79]:

$$MSE = c_p T + gz + L_v q. \quad (4.8)$$

Modelling studies agree that in the CSA process moist static energy variance is dominated by humidity variances above the boundary layer, a feature resulting from the maintenance of weak temperature gradients in the tropical atmosphere [Win+17].

FMSE is an a thermodynamic variable similar to MSE, including an additional term taking the melting and freezing of water into account so FMSE is conserved in deep convective processes. The vertical column integral of FMSE can only be changed by three processes: radiation, surface fluxes and horizontal transport. Thus, it is a suitable quantity to analyse organisation of moist convection of the atmosphere (Sec. 4.15).

4.1.4. Large Eddy Resolving Simulations (LES) and Cloud Resolving Models (CRM)

LES were proposed in the 1960s to simulate atmospheric flows with very high Reynolds numbers [FR02]. Subgrid scale processes, which are not resolved in LES, are assumed to behave as isotropic turbulence, which is strictly applicable for fine horizontal and vertical resolutions of up to 100 m. In addition, horizontal and vertical discretisation should be of similar magnitude. Both requirements are sometimes compromised due to high computational effort, thus LES are used to simulate meteorologic phenomena at coarser scales up to 1 km⁸. Regarding computational cost and accuracy, LES are between direct numeric simulations (DNS), which resolve all fluid properties, and solutions of Reynolds-averaged Navier-Stokes equations (RANS).

CRM is a term that emerged in the 1980s (brief review [GC17]), referring to simulations with resolutions up to a few kilometres, which still allow to study the statistical

⁶though this amount is often negligible

⁷equal to internal energy, if no volume work is done

⁸In this work, we will refer to 200 m resolution as LES and 1 km resolution as CRM

properties of convective cloud systems. Thanks to increasing computational power, today there are even Global Cloud-Resolving Models (GCRMs) that simulate global convection dynamics and still resolve statistical properties of clouds [Sat+19].

4.1.5. Radiative Convective Equilibrium (RCE)

When heat is exchanged by radiative transfer, the thermal equilibrium is called Radiative Convective Equilibrium (RCE). RCE describes the statistical equilibrium of the earth’s atmosphere for a given (constant) solar forcing and aerosol concentration, assuming horizontal energy transport can be neglected (Sec. 4.1.6).

More concretely, at sufficiently large horizontal and temporal scales, typically $10^3 km$ and weeks, the incoming solar radiation balances the outgoing thermal radiation in a dynamic equilibrium. Within this equilibrium, radiative cooling, compactly described by the Stefan-Boltzmann law σT_{eff}^4 , with $\sigma \equiv 5.67 \cdot 10^{-8} W m^{-2} K^{-4}$ the Stefan-Boltzmann constant and T_{eff} the effective emission temperature, balances convective heating. Convective heating warms the atmosphere by upwards transportation of warm and moist air, that for example rises from an (externally heated) warm sea surface (4.1.1).

This global radiation constraint is often found to approximately limit the possible amount of global rainfall: OLR is effectively emitted in the free troposphere, around cloud height. This radiation layer warms by convective heating, which decreases the vertical temperature gradient and prevents convection and thus rainfall, unless the level is cooled enough by OLR.

Studies on such radiation constrains suggest, that domain mean precipitation increases with warming [Tak09] [BR89]. Jeevanjee and Romps recently suggested an explanation for the precipitation increase (2-3%/K) suggested by global climate models [JR18] . They claim water vapour density being a function of temperature only⁹ (in RCE) and clear-sky radiation dominating their RCE simulations, due to a low cloud fraction. This could offer explanations for the global precipitation constrains and is especially interesting, since the cloud fraction is known to decrease with aggregation.

In RCE, this total supply of rainfall within a certain horizontal area has to be partitioned between different ‘microscopic’ rain cells, with typical diameters of $\mathcal{O}(1 km)$. Between these rain cells, typically spaced at $\mathcal{O}(10 km)$ distances, energy can be exchanged via gravity waves, which are generated when buoyancy forces aim to restore an equilibrium. Nevertheless, recent work [EWV14; WH19] suggests that more than one stable equilibrium could be possible, for example one with and one without CSA.

⁹An analytical model for tropical relative humidity, which is linked to the water vapour density had been suggested by Romps[Rom14] before.

4. Background

To investigate such complex phenomena as CSA, it is meaningful to consider RCE as the simplest possible framework [Win+17], as we do in our simulations (Sec. 5.1).

4.1.6. Weak temperature gradient approximation

Additional constraints for diabatic and large-scale fluid dynamics result from the weak temperature gradient approximation [SNP01]: horizontal temperature differences are assumed to be quickly removed by compensating vertical motion. The assumption has been shown to be approximately valid for the tropical atmosphere, where horizontal temperature gradients are observed to be small and the Coriolis force can be neglected.

4.1.7. Convective cold pools

Convective cold pools (CPs) are volumes of relatively cold air surrounded by warmer air. When raindrops evaporate underneath precipitating clouds, evaporative cooling creates a pocket of cold air [Hol73], which accelerates downwards due to its relatively high density. When hitting the ground, the cold air creates a laterally propagating gust front, spreading radially along the earth's surface. The gust front is associated with increasing wind speeds and spreads with typical velocities of $5 \text{ km h}^{-1} - 10 \text{ km h}^{-1}$ [Zui+17]. Due to lower humidity in higher levels of the atmosphere, the gust front is linked to a rapidly dropping relative humidity at the location where the transported air originates [Tom01] [Hae19]. The temperature of the downdraught depends on the potential temperature at the origin of the transported air and CPs interact both dynamically and thermo-dynamically with their environment.

Typically, CPs last on the order of a few hours [Zui+17] and reach maximal radii of $10 \text{ km} - 100 \text{ km}$. Collisions of gust fronts can trigger new convective events with resulting CPs, which influences the temporal and spatial organisation of convection [Hae+19]. It is a matter of current research how exactly CPs interact with the large scale flow and influence convective organisation [Zui+17].

4.2. Convective Self Aggregation (CSA)

CSA is the spontaneous spatial organisation and clustering of convection despite a homogeneous environment [Win+17], first described 1993 by Held, Hemler, and Ramaswamy [HHR93]. After convection is initially distributed approximately randomly, CSA gives rise to separated moist convecting, and dry non-convecting regions. These regions grow on timescales of days to weeks, significantly longer timescales than the lifetime of a typical mesoscale convective system [Win17].

Often, in the final aggregated state rain is limited to a single convective updraught area, surrounded by dry subsidence [WH19]. The size of this region seems to scale

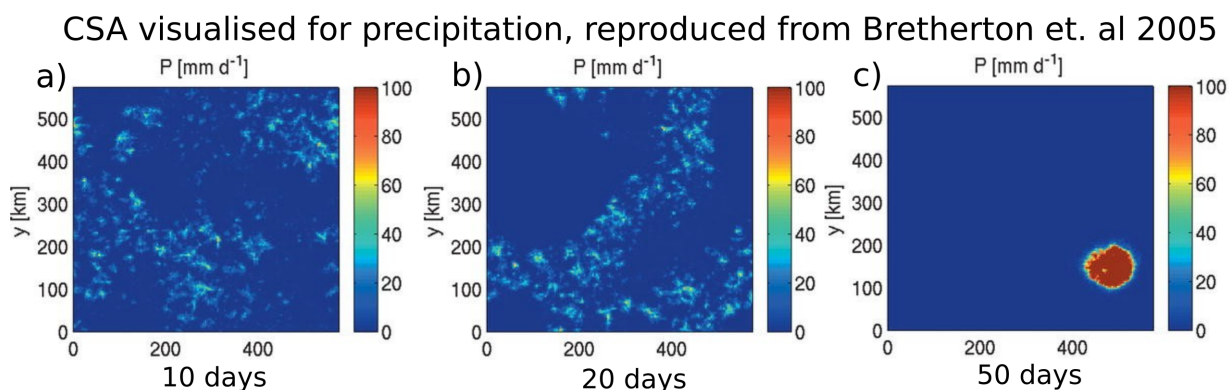


Figure 4.2.: Precipitation aggregates over a CMR simulation (reproduction based on Fig. 1-3 from [BBK05a]) Precipitation P distribution over a $576\text{ km} \times 576\text{ km}$ doubly periodic domain with 3 km horizontal grid spacing. P develops patterns of precipitating and non-precipitating regions after 10 d (a), which become more pronounced as aggregation advances (b) until only one precipitating region is left (c). Notice that the precipitation density is more than three times higher in the fully aggregated state. © American Meteorological Society. Used with permission.

with the domain size, but the final position of the convective region is random, thus only depends on initial noise [WE14].

4.2.1. CSA and domain size

Many studies showed and investigated how domain size and resolution influence the occurrence of CSA, e.g [PR19], [Win+17]. Generally, simulations with larger domains and coarser resolutions are more likely to aggregate than smaller domains and finer resolutions [MH12]. Muller and Held suggest the domain size dependency could be explained by the sensitivity of the low cloud distribution to domain size, since they suggest low clouds play an important role in organising convection [MH12]. Patrizio and Randall suggest the domain size sensitivities to be related to an increase in boundary layer wind speed as the circulation width increases for larger domains [PR19]. Jeevanjee and Romps conclude from previous studies, that CSA emerges when domains exceed a critical width of typically 200 km - 300 km and claim CP effects to impact on this transition (Sec. 4.2.4).

4.2.2. Quantifying CSA

The aggregated state is characterised by positive spatial and temporal correlation between deep convective updraughts. Such positive correlations can result from growing of dry regions or merging of individual convective updraughts. The onset of aggregation can e.g. be measured by an increase of outgoing longwave radiation

4. Background

(OLR) (Sec. 4.3.2, [WE14]). Other explicit metrics used to quantify CSA include the fraction of the domain covered with subsiding air or the number and nearest neighbour distance between convective cells. An inter-comparison project [Win+18] evaluated different cloud resolving models (CRMs) and general circulation models (GCMs) over a wide range of simulations and revealed significant differences between different metrics of CSA. For example some metrics of both model groups suggest an increase of CSA with global warming, whereas others do not. Consequently, quantifying CSA is an ongoing research topic.

CSA is not considered a numerical artefact observed in GCMs and CRMs, but suspected to play a role in the real atmosphere too, especially for tropical regions. Examples include the development of tropical cyclones and planetary-scale organisation, such as the Madden-Julian oscillation [AR15].

4.2.3. Drivers of CSA

Studies agree that CSA is driven by moisture and radiative feedbacks. Windmiller describes that in regions with higher humidity, surface horizontal wind speed and surface fluxes seem to increase [Win17]. The upward transport of moisture is defined as the moisture flux $E_{\text{moisture flux}}$. Friehe and Schmitt describe $E_{\text{moisture flux}}$ by a bulk formula (Eq. 4.9) [FS76], combining the moisture transfer coefficient c_E , the mean water vapour density near the surface $\bar{\rho}_s$, the mean water vapour density at a reference height $\bar{\rho}_a$ and the mean wind speed \bar{V} :

$$E_{\text{moisture flux}} = c_E \bar{V} (\bar{\rho}_s - \bar{\rho}_a) \quad (4.9)$$

Consequently, $E_{\text{moisture flux}}$ is proportional to the wind speed \bar{V} , varies proportional to the specific humidity of air infinitesimally close to the (sea) surface ($\bar{\rho}_s$) and decreases if the atmosphere is more saturated (larger $\bar{\rho}_a$).

When the rising air is moist, more latent heat is released (Sec. 4.1.1) and accelerates the air further upwards, causing vertical divergence. By mass conservation, vertical divergence creates a pull and increases horizontal convergence. Larger updraughts create a stronger pull, increasing horizontal convergence and wind speed, which causes a larger moisture flux and thereby provide an extra energy supply for convection, resulting in a 'rich gets richer' dynamics, discussed in Sec. 6.4.3.

Additional drivers of CSA are water vapour and cloud feedback [MH12]. For example, OLR often increases in dry, non-cloudy regions with less water vapour in the atmosphere ('clear sky effect'), which further cools these regions and inhibits convection there. Since OLR mainly depends on temperature (Stefan-Boltzmann law, $T \approx \sigma T^4$), the effective height of radiation is crucial: assuming a clear sky, the effective temperature determining OLR is close to the warm surface, causing strong radiative cooling. If clouds are present, their height is crucial. Low clouds are usually warmer and often denser than high clouds, cause stronger OLR and often

4.2. Convective Self Aggregation (CSA)

inhibit convection, whereas high clouds have a smaller cooling effect (Sec. 8.1). A detailed investigation of the physical mechanisms driving CSA was made by Wing and Emanuel (Sec. 4.3.2).

4.2.4. CSA and CPs

There seems to be consensus, that positive (Sec. 4.2.3) and negative feedback mechanisms compete on the way to CSA [Win+17]. One mechanism supposed to counteract CSA by redistributing moisture from moist convective to dry regions is the influence of CPs. The dynamical effects of CPs are limited to the radii they reach, which are typically less than 10 km. Therefore fine grids of approximately less than 3 km are needed to resolve the dynamical effect of CPs [Hae19] [Hir+20] [Mos+19]. The consideration of CPs at such fine resolutions could explain why CSA occurs less at higher resolution.

Interestingly, Boye Nissen and Haerter found that for larger domains, spatial organisation of (weak) CPs could actually trigger CSA [BH19] (Sec. 4.3.4).

Nevertheless, permitting CPs in simulations is known to open up for CSA at all domain sizes and resolutions [Kur+18]. Jeevanjee and Romps claim that CPs are responsible for the domain size dependence of CSA (Sec. 4.2.1)

In the absence of CPs, aggregation strength is suggested to smoothly decrease when reducing domain size [JR13].

Muller and Bony show that weak downdraughts (weak CPs) facilitate CSA, since the re-distribution of moisture is inhibited and link this result to the domain and resolution sensitivity of CSA. Weak downdrafts are stated to be sufficient to trigger CSA, at least in the absence of OLR feedbacks [MB15], which seems compatible with our weak CP investigation in Sec. 6.0.1. For these weak CP studies Muller and Bony observe stationary and localised convection (in line with our results, sec. 6.1.1). They further claim that OLR feedback is not responsible for CSA, since a horizontally homogeneous radiative cooling did not prevent CSA and suggest a moisture memory as modelled by Craig and Mack (Sec. 4.3.3).

This project runs a mechanism denial experiment by suppressing the re-evaporation of rainfall (NoEvap) to investigate the influence of CPs on CSA and contrasts NoEvap simulations to corresponding simulations with CPs (FullEvap) (Fig. 6.3).

The absence of CPs in NoEvap simulations leads to very persistent dynamics, more similar to coarser simulations which might not fully resolve CPs (Sec. 4.3.2, Sec. 4.3.3, Sec. 7.3.3). Therefore, it is meaningful to introduce established models for CSA (Sec. 4.3), so this study can be affiliated to their work.

4. Background

4.3. Models of convective self aggregation

Multiple studies of idealised numerical simulations have linked CSA to increasing spatial humidity variance, indicating systematically dry and systematically moist columns. Established conceptual models [EWV14; CM13] and the pioneering approach of Bretherton, Blossey, and Khairoutdinov (2005) are based on humidity feedback mechanisms. All of these studies describe how a horizontal perturbation in humidity evolves over time, but each focuses on different feedback mechanisms and even consider the humidity content at different vertical levels. [WH19] discusses bi-stability in all three models (Sec. 4.3.4), suggesting a bifurcation if humidity reaches a critical value q_{RCE} , a value stated to depend on the specific RCE simulation.

Since key insight regarding CSA was already offered by Bretherton, Blossey, and Khairoutdinov (2005), we briefly summarise their findings (Sec. 4.3.1) before explaining, where the newer models of Craig and Mack (Sec. 4.3.3) and Emanuel, Wing, and Vincent add detail or differ (Sec. 4.3.2). The latter will be discussed together with its accompanied paper [WE14], which investigates the physical mechanisms the model is built on.

4.3.1. Bretherton, Blossey, Khairoutdinov, 2005

In their seminal work, Bretherton et al. studied the organisation of deep moist convection in RCE using prescribed sea surface temperature (SST). Using the weak temperature gradient approximation (Sec. 4.1.6), they analysed CSA as an instability of a homogeneous atmosphere and suggested convection–water vapour–radiation feedbacks that systematically dry the drier columns and moisten the moister ones. Many of their results have been confirmed in a range of setups since then.

Bretherton, Blossey, and Khairoutdinov use the System for Atmospheric Modelling (SAM¹⁰) to run LES/CRMs (Sec. 4.1.4) to simulate a 576 kmx576 km doubly periodic domain with 3 km horizontal grid spacing and pre-described sea-surface-temperature (SST) of 301 K. White noise is used to initiate convection in the otherwise horizontally homogeneous field of liquid-ice static energy s_{li} . This s_{li} is calculated analogously to liquid water potential temperature θ_l (Sec. 4.1.1), by calculating the thermodynamic energy an air mass would have, if all liquid water would evaporate before it is brought adiabatically to a given reference pressure.

$$s_{li} = c_p T + gz - L(q_{li} + q_{ice}) - L_f q_{ice} \quad (4.10)$$

With the constants c_p for isobaric specific heat of dry air, g for gravitational acceleration, L and L_f for the latent heat of vaporisation. The mixing ratios q_{li} and q_{ice} of liquid water and ice combine to the total precipitation mixing ratio $q = q_{ice} + q_{li}$, which together with the total non-precipitating water q mixing ratio describe the

¹⁰Version 6.1 of SAM, updated version of the model introduced by Khairoutdinov and Randall[KR03]

4.3. Models of convective self aggregation

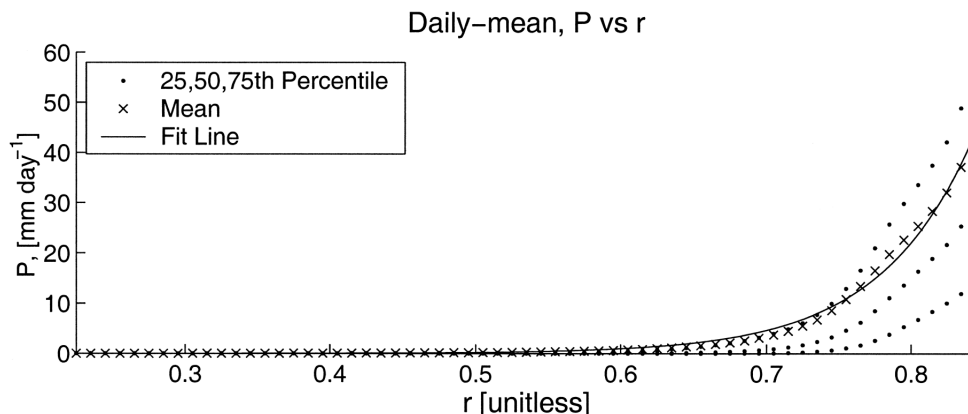


Figure 4.3.: Daily precipitation P depending on column relative humidity r . (Reproduction from [BPB04], Fig. 4) The data was averaged over 2500 grid points and measured monthly in the tropics (between -20° and 20°) in 1998-2001. Dots show the 75%, 50% and 25% percentiles and crosses show bin-mean values of P . The fitted relation is $P(r) = e^{a_d(r-r_d)}$ (with $a_d = 15.6$, $r_d = 0.63$). © American Meteorological Society. Used with permission.

total water $q_t = q + q_p$ mixing ratio.

The simulation output is averaged horizontally over $72 \text{ km} \times 72 \text{ km}$ blocks and temporally over an 24 h interval, in order to focus in mesoscale¹¹ organisation.

Bretherton et. al introduce a semi-empirical model, which approximates a budget for column relative humidity based on column-integrated frozen moist static energy (FSME, Sec. 4.1.3). Since CSA is defined by the process of segregation into dry and moist regions, FMSE is a compact (2D) way to quantify this moisture segregation process¹². FMSE variance increases with the degree of CSA, which can either result from an anomalous source of FMSE in a region of already high FMSE, or an anomalous sink of FMSE in a region of low FMSE. FMSE is used to derive an ordinary differential equation (ODE) for the relative humidity (CRH) of each grid box, assuming a constant temperature profile in each box. This is justified by the focus on earlier stages (5 d-15 d), before the domain-mean warming associated with CSA is present (Sec. 4.2.3).

An earlier observational study of Bretherton, Peters, and Back in 2004 [BPB04] collected data in the tropics (between -20° and 20°) on a monthly basis between 1998 and 2001. This comprehensive study showed, that relating precipitation P to CRH r provides clearer results and lower regional differences than relating P to

¹¹horizontal scales in the atmosphere, from a few to several hundred kilometres

¹²FMSE is exactly conserved for adiabatic displacements, in the governing equations their CRM implements. In general, other similar variables, like low-level moisture or column relative humidity typically give the same qualitative behaviour

4. Background

the water vapour path W . Calculating r was done by dividing W by its saturation value $r = W/W^*$. Afterwards the data was sorted into 100 bins of width 0.01. Precipitation shows an exponential dependency $P(r) = e^{a_d(r-r_d)}$ (with $a_d = 15.6$, $r_d = 0.63$) fitted with a nonlinear least square method (Fig. 4.3).

A fit of comparable form is derived based on the simulation data for the model introduced here (originally introduced in 2005, [BBK05a]), leading to an exponential dependency of precipitation P on CRH r :

$$P(r) = P_{RCE}e^{a_m(r-r_{RCE})}, \text{ with: } a_m = 16.6. \quad (4.11)$$

Where $P_{RCE} = 3.5 \text{ mm d}^{-1}$ is the horizontal mean precipitation rate in the radiative-convective equilibrium of the simulations, r short for CHR and $r_{RCE} = 0.72$ is the corresponding CHR $P(r_{RCE}) = P_{RCE}$.

CRH is assumed to change mainly through the combination of cloud-radiation and surface flux forcing (diabatic forcing), and through moisture convergence. Diabatic forcing linearly depends on precipitation¹³:

$$THF + \Delta R = (c_s + c_R) \cdot L(P - P_{RCE}), \quad (4.12)$$

with the surface total heat flux THF and the column-integrated radiative flux convergence ΔR . Fitting THF and ΔR on precipitation P shows linear relations, which are fitted with slopes of $c_s = 0.12$ and radiation $c_R = 0.17$ respectively.

Finally, the convergence of FMSE in the onset of CSA (days 6-10) also is linearly dependent on precipitation and modelled as:

$$C_h = a_h L(P - P_{RCE})(r_h - r), \text{ with: } r_h = 0.62, a_h = 1.8. \quad (4.13)$$

This C_h depends on the vertical motion profile of a column, which is determined by its radiative cooling profile. Therefore advective and diabatic forcing are closely intertwined. Both forcings are not only determined locally, but depend on the mean circulation. With increasing mean circulations, the surface heat fluxes in the transitions of moist and dry regions increase, an effect which depends on the circulation scale.

Finally, ODE describing the temporal change of CRH is:

$$\frac{dr}{dt} = G(r) = \frac{(c_s + c_R - a_h(r - r_h)) \cdot (P(r) - P_{RCE})}{W/r} \quad (4.14)$$

Where W is the water vapour path and all other variables are defined as introduced above. Bretherton, Peters, and Back stress that their parametrisation is only valid for the onset of aggregation at a certain domain size, since the circulation scale is linked to the domain size and domain mean variables (eg. OLR, temperature)

¹³rather than the water vapour path, which anticipates that diabatic forcing essentially depends on the amount of convection

4.3. Models of convective self aggregation

change with aggregation. For their parameter choice, $r = R_{RCE}$ is an unstable equilibrium, which physically means that anomalously moist columns gain more FMSE (by increased surface fluxes and radiative forcing) than they lose to lateral FMSE divergence. The second (aggregated) equilibrium at $r = r_{max}$, accompanied by a much stronger precipitation rate, is stable. Consequently, columns with initial CRH $r > r_{RCE}$ will moisten to $r = r_{max}$, while columns with initial CRH $r < r_{RCE}$ will continue to dry (unlimited).

In conclusion, the semi-empirical equations focus on local processes increasing humidity perturbations and do not consider humidity exchange between neighbouring regions. Further Bretherton, Blossey, and Khairoutdinov (2005) showed that CSA is suppressed when radiative cooling or surface fluxes are horizontally homogenised.

4.3.2. Wing and Emanuel, 2014 and Emanuel, Wing, and Vincent, 2014

The comprehensive analysis of physical mechanisms in Emanuel, Wing, and Vincent (2014) aims to quantify the feedback contributing to CSA, again by developing a budget method for FMSE. The framework consists of rotation free, cloud resolving SAM-simulations within RCE (Sec. 4.1.5), similar to the ones in Bretherton, Blossey, and Khairoutdinov (2005). SST is fixed to values in the interval 297 K-312 K and includes interactive surface fluxes. We consider the doubly-periodic domain of $768 \text{ km}^2 \times 768 \text{ km}^2$ horizontally and 3 km horizontal grid resolution too coarse to resolve the detailed dynamical aspects of CPs, but elements of their thermodynamic influence, especially local suppression effects on precipitation, are likely still present. Differential equations are solved for the prognostic thermodynamic variables non-precipitating water, total precipitation water and liquid-ice static energy (similar to Eq. 4.10, Sec. 4.3.1). The statistically-homogeneous initial conditions are derived from the domain average of a small $96 \text{ km} \times 96 \text{ km}$ RCE simulation and motion is induced by adding white noise perturbations to the liquid water-ice static energy.

The first aspects analysed are the dependence of CSA on SST and why CSA is reflected in a large decrease of the domain-average column relative humidity (CRH). This decrease is reported to be a consequence of very low water vapour in most columns. Together with overall low cloud fraction, that is, clouds localised within a small sub-region, the overall low CRH allows for a dramatic increase in domain-averaged OLR as simulations aggregate. High SST is known to favour CSA, so a minimum temperature of 300 K is reasonable. Surprisingly, simulations with SST equal or above 310 K do not aggregate, but the authors suggest this would not be valid for larger domains¹⁴ [EWV14].

¹⁴They hypothesise that at higher SST, larger dry static stability occurs, which means the subsidence compensating convective updraughts is weaker and thus requires a larger area.

4. Background

The derived budget for the variance of mass-weighted vertical column integral of FMSE h' is:

$$\frac{\partial \hat{h}'}{\partial t} = SEF' + NetSW' + NetOLR' - \nabla_h \cdot \hat{\mathbf{u}}h \quad (4.15)$$

SEF' is the surface enthalpy flux, equal to the sum of the surface latent heat flux (LHF) and sensible heat flux (SHF): Sensible heat exchange influences the temperature and is contrasted to latent heat exchange, which is the 'hidden' heat exchange not influencing temperature. NetSW is the column short-wave (SW) radiation flux convergence, which is positive if more SW radiative flux enters a column than leaves it. NetOLR is the column longwave (OLR) radiative flux convergence, with $OLR \approx T^4$ according to the Stefan-Boltzmann law. The final term is the advective transport term $\nabla_h \cdot \hat{\mathbf{u}}h$ leaving a given column, which is calculated as a residual from the rest of the budget (diagnostic equation). Assuming a divergence-free flux, the advective transport term has zero domain mean. The resulting approach can be viewed as a spatial model, where an increase in h' indicates a higher degree of aggregation.

Within this framework, Wing and Emanuel (2014) investigated the relevance of a number of positive feedbacks during the progression of aggregation, and summarised the results as follows: Outgoing longwave radiation (OLR) and surface heat fluxes were found to be the dominant drivers for the onset of CSA during the first 20 days of the simulations. Intermediate stages 20 d-60 d were dominated by shortwave radiation feedback, but horizontal convergence seemed equally important in 30 d-60 d. Finally OLR was reported to be the main feedback responsible for the maintenance of CSA (60 d-100 d).

An accompanied paper investigated how the dominant OLR feedback depends on SST [EWV14] and introduced a conceptual two layer model to derive a positive feedback towards CSA based on moisture and temperature.

Two layer model

Emanuel, Wing, and Vincent (2014) model the OLR feedback in two horizontal layers, one for the lower (layer 1) and one for the upper troposphere (layer 2). Radiative emission can be approximated as black body radiation, which is proportional to T^4 , according to the Stefan-Boltzmann law. Each layer is characterised by a corresponding temperature T_1 , respectively T_2 , and respective emission or absorption coefficients η_1 and η_2 , which each depend on the specific humidities, q_1 or q_2 , of the corresponding layer. Convection is an upward mass flux M_U and the downdraught is assigned to M_d . M_d is assumed to be identical for both layers. T_1 , T_2 and the surface temperature T_S are constant in time. In line with their model, SW feedback

4.3. Models of convective self aggregation

is positive, as a dryer atmosphere absorbs less incoming sunlight. For low temperatures this significantly contributes to the perturbations in net radiative heating, but is offset by the stronger negative OLR feedback. For higher temperature it is a small contribution compared to the positive OLR feedback. Thus the temperature dependence of the feedbacks governing CSA is dominated by the OLR effects. According to Emanuel, Wing, and Vincent (2014), an initial dry spot, indicating a negative moisture perturbation, is reinforced in the following feedback loop:

For a sufficient hot and dry upper troposphere (dry spot) the emission decreases, whereas the radiative cooling of the lower troposphere increases. The upward mass flux M_u is decreased and a downward motion through the troposphere establishes. This large scale descent, combined with a decrease of deep convection further dries the troposphere of dry areas, which reinforces the initial negative moisture perturbation. For a positive moisture perturbation, the opposite mechanism holds.

Concluding, Wing and Emanuel claim the SST dependency of CSA to be caused by positive longwave radiation feedback, which is only present for large SST and a resulting moist lower troposphere. If this feedback is present, negative humidity perturbations induce a stronger radiative cooling, resulting in a mean descent of cooled air, since the weak temperature gradient approximation holds (Sec. 4.1.6). Conversely, positive humidity perturbations induce weaker radiative cooling, thus a mean ascent of the warmer air.

From an abstract perspective, aggregation can be considered a linear instability of the state of horizontally-homogeneous moisture within RCE, which occurs when water vapour in lower troposphere becomes so large, that variations of its radiative cooling are mainly governed by variations in upper tropospheric water vapour. At this point, the conditions in the lower troposphere itself can be neglected and a bifurcation occurs, separating the non-aggregated moist regime with large scale ascent from the aggregated dry-regime with large scale descent.

The aggregated dry-regime state is assumed to be supercritical since it appears spontaneously. Therefore, under certain conditions an internal drive towards CSA seems present in the atmospheric simulations. Emanuel, Wing, and Vincent further argue, that CSA being not robust in simulations could be an artefact of running CRMs close to their critical temperature.

Interestingly, Wing and Emanuel emphasise that radiative feedbacks are important drivers of CSA, but question convection-moisture feedbacks to be sufficient to cause CSA. Such feedbacks between convection and moisture are suggested by Craig and Mack and will be discussed in the following. (Sec. 4.3.3)

4. Background

4.3.3. Craig and Mack, 2013

It is known that local processes that are linked via diffusive coupling lead to pattern formations [Mur82] similar to the spatial organisation in CSA. Craig and Mack¹⁵ present a coarsening model for CSA, based on a positive feedback between convection and humidity in the free troposphere [CM13]: First, convection will increase the humidity in the free troposphere by transporting humidity out of the boundary layer. Second, deep convection is more likely to occur in moist regions, since the entrainment of dry air surrounding a given convective updraft reduces buoyancy and this effect will be stronger in the drier sub-regions. Boundary layer processes are not considered, similar to the model introduced above (Sec. 4.3.2).

Coarsening is observed as dynamical scaling, with a power-law increase of correlation length w.r.t. time, that is, the growth of spatial structure is self-similar. The coarsening model is based on an Allen-Cahn equation (Eq. 4.16, [Lif62]) with the troposphere's total humidity I as the 'order parameter', which describes the phase:

$$\frac{\partial I}{\partial t} = -\frac{\delta V(I)}{\delta I} + \nu \nabla^2 I. \quad (4.16)$$

The first term on the RHS is the functional derivative of the potential¹⁶ $V(I)$. Two minima in $V(I)$ correspond to a bi-stable system with two equilibrium values for I . Mixing, through entrainment and detrainment, is described as a diffusive process (second term, RHS), with ν an effective diffusion coefficient. With such a diffusive mixing and a non-conserved order parameter, this equation implies a length scale (LS) increase proportional to $t^{1/2}$.

Craig and Mack physically model coarsening, assuming that subsidence is spatially homogeneous, but regions that already convect are re-moistened by the re-evaporation of rainfall. The evolution of the column humidity is explained by local processes and assumes relative humidity to be exponentially depending on local rainfall.

A global coupling is given through the domain-mean precipitation rate, implying a constrain on total precipitation as required for RCE simulations (Sec. 4.1.5), which then provides a constraint on the local reaction-diffusion-like dynamics.

The coarsening model centrally assumes that the moisture content in the lower troposphere determines where convection takes place, but lacks small scale variability and smooths out the effects of individual convective cells to the moisture content¹⁷. Assuming a uniform horizontal temperature, feedback between cloud or water vapour

¹⁵Mack later changed to Windmiller

¹⁶sometimes associated with the Landau free energy

¹⁷It would be interesting to investigate, whether the 'diffusion' could be more physically modelled as the influence of CPs on the small scale

4.3. Models of convective self aggregation

and radiation is neglected. Instead, the model focuses on a positive precipitation-humidity feedback and argues that the boundary layer and thus surface fluxes and horizontal transport can be neglected, too.

The sensible heat flux is set to zero, thus to keep the domain-averaged precipitation rate constant, a constant radiative heat loss is proposed. This heat loss amounts to constant OLR, which is balanced by latent heat, released in the convective process. The model divides the domain into discrete columns, and models the moisture content of the free troposphere I_v for each column. I_v is given by the water vapour density ρ_v , which is integrated vertically over the troposphere:

$$I_v = \int_{z_B}^{z_T} \rho_v dz . \quad (4.17)$$

The time evolution $\partial I_v / \partial t$ is given by this schematic budget:

$$\frac{\partial I_v}{\partial t} = S + C + T , \quad (4.18)$$

where S is the column integrated subsidence drying, a negative contribution creating a moisture sink at subcloud height. Subsidence is driven by radiative cooling, which causes cooled and thus denser air to descent. T is the horizontal transport of moisture, thus radiative cooling and horizontal transport are actually taken into account. The major moisture source is the convective moistening C , which is mainly detrainment from convective clouds. Lacking a relation for the dependence from C on I_v , Craig and Mack use the precipitation rate $P(I_v)$, which is approximated to be exponentially increasing with I_v :

$$P = a \left(e^{\frac{b \cdot I_v}{I_v^*}} - 1 \right) . \quad (4.19)$$

I_v^* is the saturation integrated water value $I_v^* = 57 \text{ kg/m}^2$, the scaling parameter b is chosen with $b = 11.4$. The form of P is based on Bretherton, Peters, and Back (Sec. 4.3.1), but modified such that no precipitation will occur in a completely dry atmosphere ($I_v \approx 0$). The global radiation constrain is provided by a , which is determined for each iteration step t :

$$a(t) = \frac{P_{av}}{\frac{1}{A_{domain}} \left(e^{\frac{b \cdot I_v}{I_v^*}} - 1 \right) dA_{domain}} \quad (4.20)$$

The constant P_{av} in RCE is estimated with $8 \text{ kg/m}^2/\text{d}$ (based on Tompkins and Craig, [TC98]). Precipitation is linked to convective moistening via a precipitation efficiency η , chosen to increase with I_v as:

$$\eta = \beta \frac{I_v}{I_v^*} . \quad (4.21)$$

4. Background

With β slightly above one, η varies from zero for very dry air ($I_v \approx 0$), which causes strong evaporation of cloud water and thus little rainfall, to β for saturated air¹⁸. All these assumptions combine to:

$$C(I_v) = a \left(e^{\frac{b \cdot I_v}{I_v^*}} - 1 \right) \left(\beta \frac{I_v}{I_v^*} \right) \quad (4.22)$$

A necessary condition for a coarsening process is, that drying dominates over moistening for small I_v , whereas moistening dominates for large I_v . Here, the spatial autocorrelation length exhibits dynamical scaling with $t^{0.5}$, a factor shown to be universal for different values of b . Nevertheless, the coarsening process takes longer for smaller values of b , corresponding to a weaker positive feedback between precipitation P and column humidity content I_v .

Assuming subsidence S to be linearly depending on I_v and the advective transport T to be dominated by down-gradient diffusion, the final budget equation is:

$$\frac{\partial I_v}{\partial t} = -\alpha I_v + a \left(e^{\frac{b \cdot I_v}{I_v^*}} - 1 \right) \left(\beta \frac{I_v}{I_v^*} \right) + K \nabla^2 I_v . \quad (4.23)$$

K is the eddy diffusivity $K \sim v_o L$. Typical values associated with convective motions are horizontal velocity $v_o = 10 \text{ m s}^{-1}$ and length scale $L = 10 \text{ km}$, leading to $K = 10^5 \text{ m}^2 \text{ s}^{-1}$. Craig and Mack acknowledge that a diffusive transport is not justified if the moisture field significantly influences the velocity.

In the present work we adopt the assumption of homogeneous subsidence but explicitly consider the horizontal boundary layer flow, which has been neglected in the two models presented above (Chapter 7). We further re-examine the choice of a re-moistening feedback based on the evaporation of precipitation: In our main simulations, re-evaporation is explicitly suppressed, nonetheless, these simulations do show self-aggregation. Therefore, we claim additional positive feedback mechanisms to be present (Sec. 8.1). We make a reference to the dynamical scaling (Sec. 6.5) and compare it to the LE increase approximated from our NoEvap simulations.

4.3.4. Bi-stability and spatial evolution of CSA

CSA as a general organizing mechanism?

It has not yet been understood how the spatial scale of CSA develops temporally and which mechanisms control this time evolution. After introducing the coarsening model (Sec. 4.3.3), Windmiller and Hohenegger suggested a rather mathematical approach, proposing that in RCE the humidity distribution of the atmosphere can be described as an abstract bi-stable system, an idea that had been discussed before

¹⁸This is mentioned to reflect a net drying tendency for very moist air, observed by Raymond et al., [Ray+09]

4.3. Models of convective self aggregation

[RZ00].

Again, in this framework, CSA means that systems transition from the dis-aggregated to the aggregated equilibrium. A bifurcation occurs when humidity reaches a critical value q_{RCE} , which depends on the respective RCE simulation. If local humidity is perturbed, such that $q > q_{RCE}$, the region develops towards a moist equilibrium, whereas local values $q < q_{RCE}$ cause an entirely dry region.

Windmiller and Hohenegger claim a universal spatial evolution, between the three established models of [BBK05a; CM13; EWW14] introduced above (sec. 4.3.1, 4.3.2, 4.3.3).

Despite varying in detail, all three models base on local mechanisms and in all models homogeneous equilibrium is unstable against local humidity perturbations. Comparing the three respective time evolution equations allows Windmiller and Hohenegger to derive a stochastic model for the (universal) spatial evolution of CSA. The respective equations for the budget evolution are regarded as different possible source terms R in a reaction diffusion equation, which describes the evolution of the humidity content:

$$\frac{\partial q}{\partial t} = R(q) + D\nabla^2 q \quad (4.24)$$

The form of this equation is similar to the Allen Cahn (Eq. 4.16) equation in the previous coarsening model (Sec. 4.3.3, [CM13]) introduced by Craig and Mack. R describes the local processes for each column and differs depending on the model under consideration. It is claimed that the horizontal mixing term $D\nabla^2 q$, which provides a diffusive spatial coupling¹⁹, can result from a stochastic representation of convection²⁰.

Despite that the precise time evolution differs between models (each implying a different R), the key property of coarsening is shown to be universal: Intervals with typical LE increasing proportional to $t^{0.5}$ can be found for all three models. Nevertheless, when moist (dry) regions start to expand at the expense of the dry (moist) regions and moist (dry) regions cover more than half of the domain, the scaling diverges from the $t^{0.5}$ scaling.

The general idea of bi-stability and the supercritical bifurcation proposed by Emanuel, Wing, and Vincent, raises the question whether the atmosphere is actually driven towards an aggregated state under a rather broad range of model assumptions. Thinking this further, if true for simulations, what does this drive imply for the real atmosphere? Aggregation-like effects can be observed in the context of extreme events as tropical cyclones, but synoptic-scale disturbances and homogenising influences, such as CPs, may generally prevent persistent and complete aggregation in

¹⁹which lacks physical motivation in the original coarsening model (Sec. 4.3.3)

²⁰We suggest, that maybe modelling CPs could even be a physically more plausible approach to describe mixing in CSA models, than assuming diffusive processes.

4. Background

nature²¹. The following paragraph provides an overview on studies investigating the role of CPs in organising convection.

CPs (dis)organise convection

Cold pools (CPs) are known to inhibit CSA [JR13], by redistributing moisture to dry regions (Sec. 4.2.4). On the other hand they seem to organise convection in their own way. CP interaction can be used to model the energy transfer in the boundary layer [MH20] and a model proposed by Fuglestedt and Haerter (2020) shows that CPs organise subsequent updraughts in CP driven convective organisation: When colliding, CP outflow boundaries become immobile and force moisture into narrow convergence zones. This effect results in localised pre-moistening, which favours subsequent convection and seems to dominant over surface fluxes. Thus, CP collisions reset the circulation in the boundary layer.

Boye Nissen and Haerter (2019), Haerter (2019) suggest that despite CPs generally working against CSA, the organisation of weak CPs could actually trigger CSA. Many studies agree, that CSA starts from a small dry spot, but do not answer the question how an initial dry spot emerges. Under certain conditions, the dry downdraught of weak CPs could facilitate the emergence of dry spots that are large enough to not be re-moistened and therefore trigger CSA. The emergence of such a dry spot is discussed in a simple geometrical model [BH19]: The model centrally assumes that collisions of expanding CPs, modelled as simple circles of growing radii, initiate new CPs. The expanding CPs can trigger new convection when each of the colliding CPs has a radius between a minimum radius r_{min} , caused by the strongly-negatively buoyant core of the CPs, and r_{max} , the maximum radius they can reach. The 'circle model' shows geometrically that for weak CPs with small r_{max} , dry spots emerge where no new convection can occur. Once a dry spot is large enough, the weakened CPs can not 'bridge across' such a dry spot and no moisture can thus be transported into this region. This phenomenon is referred to as the 'initial symmetry breaking,' which triggers the positive feedback loop of CSA described above (Sec. 4.3.2).

Consequently, weak CPs can open for CSA on very small domain sizes of $96 \text{ km} \times 96 \text{ km}$ at fine horizontal resolutions (200 m).

This thesis focuses on the extreme case of zero re-evaporation (NoEvap) where no CPs form ($r_{max} = 0$) and which always aggregates (Sec. 4.2.4). We propose that, without CPs redistributing moisture, the initial noise on the homogeneous conditions is enough to break the spatial symmetry, and in some regions convection never takes place (Fig. 6.7). This work starts its investigation by building on the same simulations as used in [BH19] (Sec. 5.1) but later decreases horizontal model resolution to 1 km, which allows to simulate a larger $480 \text{ km} \times 480 \text{ km}$ domain at comparable computational effort.

²¹Nevertheless, an internal drive towards aggregated states can not be excluded

5. Methods to detect convective events

Because convective clouds are highly localised, investigation on the scale of individual convective cells is needed in order to understand their physics. Therefore it is crucial to track their temporal development. This chapter provides details on the simulations used (Sec. 5.1).

Sec. 5.2 briefly outlines the tracking methods for convective events and is further divided into a short background on cloud tracking (Sec. 5.2.1) and a motivation of an alternative approach based on the velocity field (Sec. 5.2.3). Afterwards an algorithm is introduced in Sec. 5.4, that calculates the convective available potential energy (CAPE). The last introduced approach tracks convective events based on the lowest level velocity field (Sec. 5.5).

5.1. Large-eddy (LES) and cloud-resolving model (CRM) simulations used

This project uses the University of California, Los Angeles, Large Eddy Simulator (UCLA-LES) to run under RCE-like conditions at horizontal model resolution of 200 m and 1 km. Sub-grid scale turbulence is parameterised after Smagorinsky [Sma63]. Further, a delta four-stream radiation [PS09] and a two-moment cloud micro-physics scheme [Ste+05] are used. Rain re-evaporation was accounted for via a scheme introduced by Seifert and Beheng 2006. Suppressing rain re-evaporation means setting the ventilation factor $F_v = 0$ [SB06], so no condensed water (including ice) converts to water vapour.

Energy is supplied externally by heating a mimicked sea surface to 300 K, hence an infinite heat capacity and a perfect source of water vapour, assuming saturation specific humidity at the atmospheric interface. Solar insolation is kept constant at 655 W/m^2 , which corresponds to an equatorial zenith angle of 50 deg [BBK05b]. We assume interactive surface heat fluxes, which are proportional to the vertical temperature and humidity gradients and to horizontal wind speed (bulk formula, Eq. 4.9, Sec. 4.2.3). The horizontal wind speed is approximated using the Monin-Obukhov similarity theory, which idealises the mean flow profile for a neutral boundary layer based on Prandtl's mixing length theory [Hol73]. Prandtl's theory assumes a logarithmic dependence of the horizontal mean flow to height [PPP01]. To initialise

5. Methods to detect convective events

temperature and humidity a prior 3-day spin-up simulation with 400 m horizontal resolution is used (details [BH19], Fig. S1).

The Arakawa C grid [MO54] used is regular spaced in the horizontal with double periodic boundary conditions¹. The resolution between the 75 vertical levels varies from 100 m in the lowest 1 km, to 400 m in the upper layers. Assuming an equatorial region, the Coriolis force and the mean wind are set to zero. Weak noise generates random initial perturbations to break complete spatial symmetry. Instantaneous surface precipitation intensity, specific humidity, temperature, liquid water mixing ratio, outgoing long-wave radiation, and 3D velocities are calculated every 10 min for different vertical levels. To keep continuity to previous work done in [BH19], we first simulate small domains of 96 km×96 km up to eight days. In addition, we run coarser simulations of 1 km for a larger domain of 480 km×480 km.

5.2. Detection of convective events

Most work concerned with the onset of CSA uses field-like descriptions and models and divides the domain into regular columns ([WE14][CM13][EWV14][MH12]). In contrast, this thesis presents two ideas how to study CSA by describing convective events as discrete objects:

- (i) The 'rain-cluster' approach, in line with the rain tracking introduced in Sec. 5.2.2.
- (ii) The 'basin' approach, for the lowest-layer flow field ($z \approx 50m$)

5.2.1. Background on cloud tracking

Precipitation tracking for data from digital weather radars have existed for more than 40 years [Cra79], with the purpose of early warning for extreme rain and storm events. Later, satellite-based tracking was developed, with an increased interest in long-term analysis of convective mechanisms [Men02].

In more recent years, simple object-based algorithms were suggested [DBB06], which then became more sophisticated to detect complex entities of numerous individual clouds and capture their splitting and merging [Hei+16]. The comprehensive long term analysis by [WBS17] evaluated all tropical and subtropical rain data between 1980 and 2015 and focused on event characteristics, like size and total rainfall.

Another approach, the "forward-in-time" (FiT) algorithm [SBT13], starts by defining multiple thresholds in the precipitation rate that qualify rain events. In agreement with our method (Sec. 5.3.2), unique rain objects are identified for all time steps. Second, objects, linked across multiple time steps, are determined: Two objects are defined as a single event, if any pixel of an object at time t overlaps with any pixel of an object at $t + 1$. 'Forward-in-time' refers to allowing an event to split

¹<https://mitgcm.readthedocs.io/en/latest/algorithm/c-grid.html>

into multiple objects, which are still considered as a single identity. In contrast to FiT, our method allows two objects that are initially unique, to merge at any later time step.

Generally, most algorithms are developed for very specific applications so no consistent rain tracking approach, applicable to all situations, exists ².

5.2.2. Tracking individual rain clusters

To investigate individual rain clusters on their way to CSA, we developed a specialised threshold-based rain cluster detection algorithm. At each time t we consider four neighbours for each domain pixel (x_i, y_j) : $(x_{i+1}, y_j), (x_{i-1}, y_j), (x_i, y_{j+1}), (x_i, y_{j-1})$.

A rain patch is defined as spatially neighboured active pixels:
e.g $I(x_i, y_j, t_k) > I_0 \wedge I(x_{i+1}, y_j, t_k) > I_0$.

A rain cluster is defined as two consecutive rain patches with spatial overlap:
 $\exists (x_i, y_j) : I(x_i, y_j, t_k) > I_0 \wedge I(x_i, y_j, t_{k+1}) > I_0$.

We start by detecting rain patches for each simulated time step and later link them over time. The horizontal distances which individual rain clusters' centres of mass travel in the simulated 10 min interval are small relative to their typical diameters. Thus, it is reasonable to base the connection over time on pixel overlap. Visualising the data (eg. in animations³) showed, that merging of rain clusters is an important part of the dynamics. In contrast, splitting of rain events does not seem to be part of the actual dynamic and can be considered an artefact, when the spatial connection between pixels of rain clusters is (usually only temporary) lost. To 'suppress' splitting, all raining pixels that have been connected once, are considered as the same rain cluster for the rest of their duration. In contrast to standard percolation algorithms (e.g [HK76]), we link rain patches 'forward in time' only, because we want a cluster to be a causally correlated entity: If two well-separated clusters eventually merge much later, there is no reason to believe that they were initially correlated. If two former separated clusters merge, a new rain cluster is created, which stores the information of its history. Tree-like merging dynamics are characteristic for the NoEvap simulations (Fig. 5.4).

²Recently, approaches are made to address this inconsistency in rain tracking methods. For example the TOBAC algorithm (More: Sec. A.1.1) is designed to be applicable to multiple situations

³Visualised in the supplemental material.

5.2.3. Motivation of tracking algorithm, based on horizontal velocity field

The LES data shows highly localised updraughts and a quite homogeneous subsidence for the lower layers (Fig. 6.9). After smoothing the horizontal velocity field using a Gaussian filter, it appears to be vortex-free; clearly separated areas of horizontal convergence and divergence can be observed in the streamlines (Fig. 6.9, left). Each updraught is a sink in the horizontal velocity field with an area of convergence surrounding it. These areas of convergence can be viewed as the corresponding "basins of attraction".

These basins are separated from one another by boundaries of pure horizontal divergence. This 'basin approach' pairs the centre of a given updraught and the corresponding wind-field-derived basins from the LES data. After detecting the basins for several simulation steps, they are linked in time to track their time dependence. This procedure further allows to employ the updraught's horizontal coordinate and the corresponding basin area within a future toy model, revealing the underlying dynamics.

5.3. Implementation of a rain cluster tracking algorithm

We choose an object orientated approach to implement the cluster detection. The workflow is visualised in Fig. 5.1 and the details of the different process stages are explained in the following.

5.3.1. Preprocessing

We base our analysis on precipitation $I(x,y,t)$, which roughly represents the surface rainfall (glossary). Simulating a domain with lateral dimensions L_x, L_y , the $I(x,y,t)$ output is stored in a netcdf file containing an $[L_x \times L_y] \times [t]$ array. To smooth fluctuations, a Gaussian filter is applied. The filter size is chosen to fit the simulation under consideration. For example, for the NoEvap simulation (at 1 km resolution), we use a spatial kernel of 10 km on the horizontal dimensions and a 5 h temporal kernel⁴.

Second, the filtered data is binarised: Active pixels (glossary) are marked '1'. It is important to do **thresholding** in interplay with the choice of filter parameters. Afterwards a specialised **segmentation** algorithm ('cluster creation') detects the rain patches at each time step t with the 'scipy.ndimage.label' function: A $L_x \times L_y$ 'label array' inscribes active pixels of each rain patch with a uniform label. Not active

⁴The 5 h interval roughly corresponds to half of the oscillation period for the last remaining rain cluster.

5.3. Implementation of a rain cluster tracking algorithm

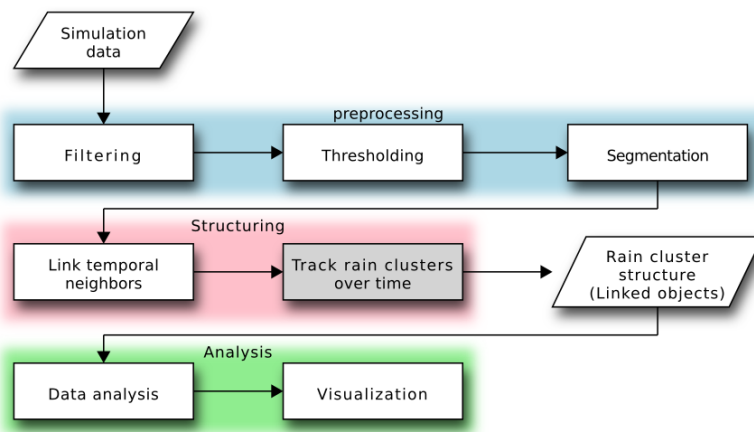


Figure 5.1.: Abstract structure of the rain cluster detection algorithm
The input $I(x,y,t)$ is pre-processed: Applying a Gaussian filter smooths fluctuations and applying a rain threshold I_0 generates binary data. This binary data is then segmented into discrete rain patches. These rain patches are then structured by linking patches with pixel overlap in subsequent time steps and track rain clusters over time, which involves a complex merging routine, introduced in 5.3.2. The outcome are rain clusters that evolve over time and can be further analysed.

5. Methods to detect convective events

pixels are inscribed with zero. To consider the double periodic boundary conditions, we manually adjust the labels of patches reaching across a domain border: If $I(x = 0, y_j, t_k) > I_0 \wedge I(x = L_x, y_j, t_k) > I_0$ or $I(x_i, y = 0, t_k) > I_0 \wedge I(x_i, y = L_y, t_k) > I_0$, the corresponding rain patches are considered to be the same (see Fig. 5.3, yellow cluster). The segmentation algorithm returns the $t \times L_x \times L_y$ label array, and a list of pixels belonging to each patch.

5.3.2. Structuring

Patches

Patch objects are defined for a single time step t_k and stores its corresponding active pixels and the inscribed label. As explained above (5.2.2), we **Link temporal neighbours** of pixels by searching for pixel overlap with patches occurring at t_k+1 . Eventual overlaps with patches defined at t_k+1 or t_k-1 are stored in the patch object, too. Optionally, additional features like the patch centre⁵ or the mean precipitation ($\bar{I}(\text{patch}(t_k))$) can be calculated.

Track rain clusters over time

As explained in Fig.5.2, we link the patches to clusters that persist over time. The algorithm **Loops through time-steps** t_k and for each t_k starts the processing by **grouping patches by their overlap to t_k-1** , distinguishing three cases:

1. **no overlap t_k to t_k-1** : A new cluster object is created and assigned with the start time $t_{start}^{cluster} = t_k$. Such clusters without predecessors are referred to as 'root clusters'.
2. **>1 overlap t_k to t_k-1** : A patch(t_k) overlapping with several patches(t_k-1) can indicate a merge of several clusters (further distinguished in Sec. 5.3.2). If a merge is detected (Fig. 5.3, b→c), the previous clusters are closed and a new cluster is created, storing the 'history' of its predecessors.
3. **1 overlap t_k to t_k-1** : The patch(t_k) is added to the cluster to which the overlapping patch(t_k-1) belongs and the corresponding cluster continues for one more time step (Fig. 5.3, a→b)

Closing clusters and progress to next time step

After the processes described above are conducted, the algorithm collects patches existing at t_k-1 which do not have an successor in t_k . The associated clusters are

⁵The centre calculation is quite elaborate, since double periodic boundary conditions have to be taken into account. Calculating the centre is relevant for tracking the lateral movement of rain clusters or calculating the nearest neighbours distance (NN distance) between rain patches. The NN distance is a measure used to quantify CSA, but the analysis of this implemented feature is out of scope for this thesis.

5.3. Implementation of a rain cluster tracking algorithm

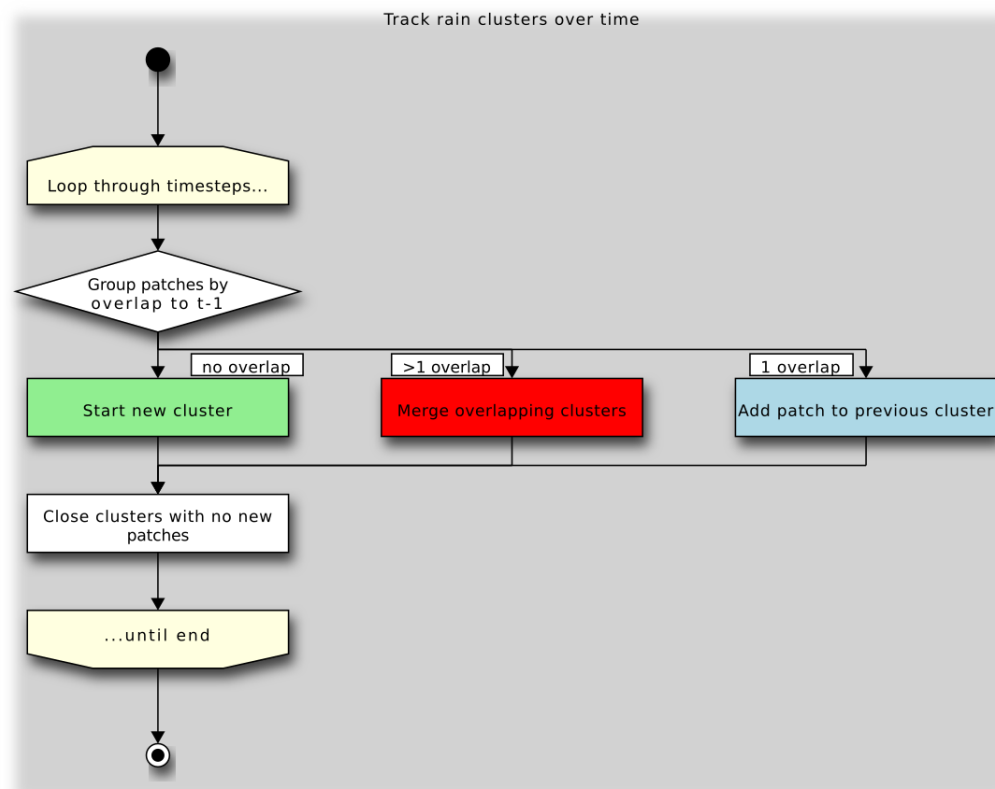


Figure 5.2.: Details of decisive parts of the algorithm: **Linking rain patches(t_k) to rain clusters that persist over time.** We loop through the simulation time t_k and group the detected rain patches(t_k) by the number of pixel overlap to patches(t_{k-1}). Three different cases are separated, depending whether there are no predecessors (green), several predecessors (red) or exactly one predecessor (blue). Based on this either a new cluster is started, several clusters are merged or an existing cluster is simply continued. Afterwards any cluster existing in t_{k-1} without a successor at t_k is closed and removed from the track of active clusters from the current time.

5. Methods to detect convective events

then 'decaying' and closed, by setting $t_{end}^{cluster} = t_k$. In addition, they are removed from the list of 'living' clusters for $t \geq t_k$.

Merging

The most complex part of the algorithm is the merging (Sec. 5.2, red) routine. Again, splitting seems not relevant for the dynamic and removing *splitting* artefacts makes the process much more demanding: If at any time t_k-1 multiple patches belonging to different clusters overlap with a mutual successor patch at t_k , they are merged to one new cluster. From then on ($t > t_k$) all their subsequent patches are considered as one cluster, even if they loose spatial connectivity, until either all 'cluster fractions' die or all overlap again, which also prevents 're-merging' of patches with a common ancestor in an elegant way.

Multi-merges of several patches merging in one step are covered too⁶ and merges are tracked in a 'merging list' in the Clusterstore, which organises all clusters over the simulation.

To test our algorithm we ran it with the largest data set (480 km \times 480 km) unfiltered, and found unexpected corner cases, as the rare case that one patch(t_k) that lost connectivity to other patches(t_k) of its cluster $_m$ (corresponding to a 'suppressed' split) merged to a patch(t_k) of a different cluster $_n$. This case is now covered by merging all patches belonging to cluster $_m$ and cluster $_n$, although they have no direct pixel overlap.

5.3.3. Discussion of the rain tracking algorithm

Challenges of detecting rain clusters

Like all threshold-depending algorithms, our cluster detection algorithm is very filter- and parameter-sensitive. As common (compare Sec. A.1.1, [Hei+19]), we apply filtering before applying a threshold. Changing the filter kernel demands to adjust the threshold, since larger filters need a lower threshold I_0 .

Optimising these parameters is further complicated by the oscillations in clusters' rain intensity $I(\text{cluster},t)$ (glossary). We claim $I(\text{cluster},t)$ results from the same convective cell throughout the oscillation. Therefore, the challenge is not to loose track of a cluster, when rain intensity in all pixels of the cluster temporarily falls below the threshold ($I(x_i, y_i, t_k) > I_0 \notin I(\text{cluster}, t_k)$), but $I(\text{cluster},t)$ increases again shortly after.

We also experimented with a temporal sliding window mean⁷, but in order to keep continuity in raining clusters a temporal Gaussian filter showed to perform better.

⁶here the challenge is to avoid duplicates

⁷which nicely smooths curves of $I(\text{cluster},t)$

5.3. Implementation of a rain cluster tracking algorithm

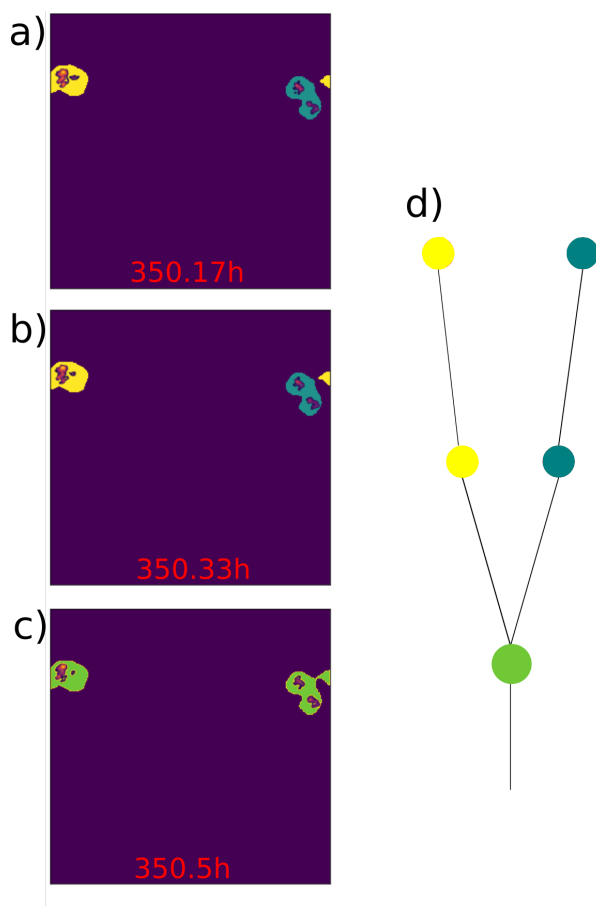


Figure 5.3.: Three snapshots (a-c) visualising the merge of the last two clusters and the corresponding cluster representation provided by our algorithm (d). (a-c) active rain pixels ($I(x,y,t) > I_0$, red contours) are detected in clusters. The different clusters are indicated with yellow, turquoise or green background colouring respectively and the colours correspond to the illustration of the abstract cluster representation (d). Note that despite reaching across the double periodic domain borders, the yellow clusters are detected as one. The yellow and the turquoise cluster continue from (a) to (b) and merge to a new (light green) cluster (c), as soon as their pixels overlap.

5. Methods to detect convective events

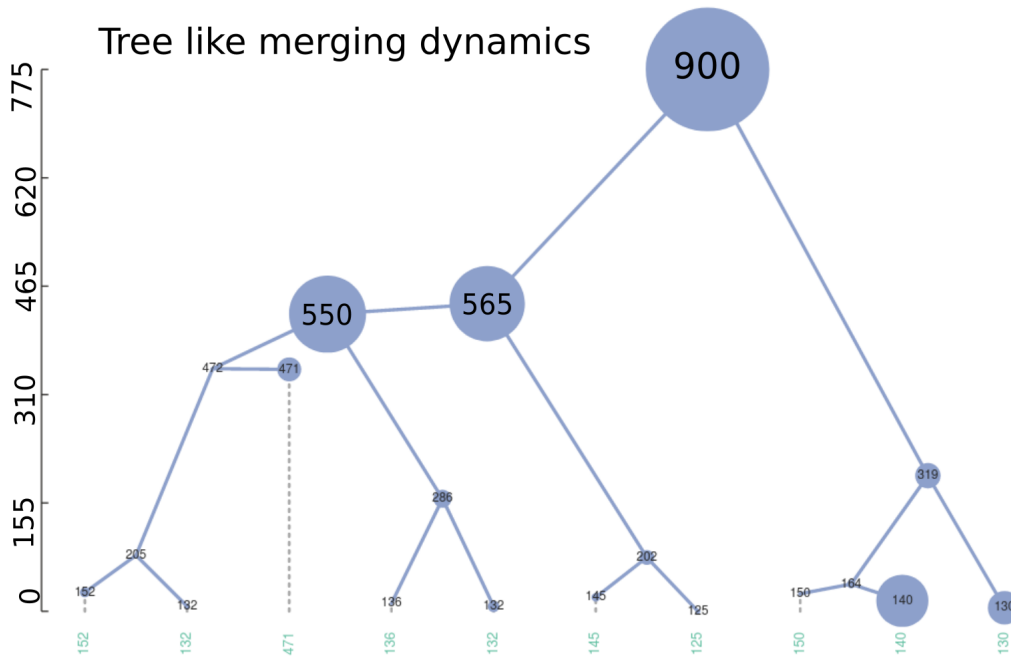


Figure 5.4.: Visualisation of a cluster’s tree like merging history (toytree library) It is possible to inherit properties from the newick library for our cluster objects, which enables various visualisations. The node size scales with the mean cluster area and the branch length scales with the time since the first detection of a cluster, which also corresponds to the node label. Notice, that time (y-axis) is given in simulation steps of 10 minutes and that cluster areas usually increase after a merge, but exceptions occur due to tempo-spatial oscillations in $I(\text{cluster}, t)$.

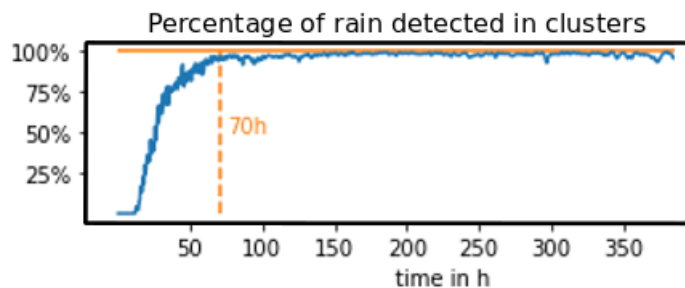


Figure 5.5.: Performance of the rain tracking, measured as the percentage of domain integrated precipitation, that is detected inside all clusters: $100 \cdot \sum_{\text{cluster}} I(\text{cluster}, t) / \sum_{x,y} I(x, y, t)$. After 30 h more than 70% and after 70 h more than 90% of rain is captured in our clusters. We suppose during this ‘spin-up time’ dynamics differ from later stages of the simulation. Since we decided to focus our analysis on later stages (after 100 h), we chose the parameters to capture those well, accepting a weaker performance over the first three days of the simulation.

5.3. Implementation of a rain cluster tracking algorithm

An additional challenge is the increase of spatial scale (glossary): Clusters become larger and their distance increases over time. Thus, compromises have to be made, when adjusting parameters: For our largest simulation (480 km×480 km) with more than 18 days, we focused on filters that represent the stages after three days of the simulation very accurately (Fig. 5.5). To scoop out the algorithm’s full potential, the segmentation procedure should be improved to be less parameter-depending to derive reliable scientific findings. For example additional physical variables could be considered, for which our algorithm could be combined with functionalities provided by existing rain tracking libraries as the TOBAC algorithm introduced in Sec. A.1.1 [Hei+19]. Nevertheless, a main advantage of our threshold-based algorithm is a potential expansion to handle different types of 2D data. Theoretically, it could be enhanced to work for example on back scattering data from satellite observations.

Refactoring the algorithm, we were able to reduce the computational time for the rain tracking of the 480 km×480 km simulation to less than two minutes. In return, some properties as the centres of rain patches and rain clusters are not calculated as a standard feature anymore. Unfortunately, centre calculation demands high computational effort: To take the double periodic boundary conditions into account, we calculate the centre of mass (COM) for all pixels belonging to a rain patch, based on a mathematical method introduced by Bai and Breen[BB08]. Their idea is to calculate the COM for the x- and y-dimension separately, by forming the double periodic domain into two orthogonal tubes. Now, the COM is calculated in 3D for each tube and then is projected back to 2D, which provides the COM of all pixels belonging to a rain patch. This tool does not use approximations and is mathematically elegant, but expensive in computational time. Concluding, we will consider alternatives to use the ‘centre’ feature for a systematical data analysis.

Concluding, our rain cluster algorithm is specialised to capture multiple corner cases of the tree like merging dynamics, occurring in the NoEvap simulations. As a main advantage, it can be easily reviewed visually that the detected number of clusters conforms with the expected count (Fig. 5.3). The implemented algorithm allows to track rain clusters over time, derive basic statistics and offers good evaluation of its performance.

Nevertheless, since rain $I(x,y,t)$ is strongly influenced by tempo-spatial oscillations, we explore additional approaches, which rather focus on the underlying convective dynamics.

5.4. Calculating (CAPE) to detect regions where convection can possibly take place

CAPE describes the buoyant energy of an moist adiabatically rising test parcel (Sec. 4.1.1) and convection is generally limited to regions of positive CAPE. To calculate CAPE we compare the temperature profile of a hypothetical lifted test parcel with the actual temperature profile of the background environment. Whereas the actual temperature profile can be derived from the LES output via prognostic equations calculating the profile of an irreversible moist-adiabatically lifted test parcel is more complicated. To derive the test parcel's profile, we follow the dry adiabatic lapse rate (Sec. 4.1.1) until the water vapour pressure p_v reaches the saturation value p_v^* , which is defined by the Claudius Clapeyron relation. From there on we follow the irreversible moist adiabatic lapse rate, assuming all excess water instantly condenses, releasing its latent heat and leaving the parcel falling down as rain due to gravity. The main challenge is to determine how the fractions of the water phases in moist air change, as the parcel is further lifted. We solve this by using the simplification that condensed water is immediately removed from the test parcel (Sec. 4.1.1) and follow a moist adiabatic lapse rate approximated by Randall [Ran09] with:

$$\Gamma_{moist} = \Gamma_{dry} \left[\frac{1 + \frac{Lq^*(T,p)}{R_d T}}{1 + \frac{L^2 q^*(T,p)}{c_p R_v T^2}} \right] \quad (5.1)$$

L is the latent heat of vaporisation, q^* the saturation specific humidity, R_d , R_v are the gas constants of dry air and water vapour and c_p is the specific heat capacity at constant pressure.

After deriving the two temperature profiles, we compare them starting from maximum height level. The first intersection of profiles is the level of neutral buoyancy (LNB, Sec. 4.1.1). From there the algorithm vertically integrates the positive virtual temperature differences $T_{parcel} - T_{environments}$ until the second intersection is reached with the level of free convection (LFC, Sec. 4.1.1). The common unit of CAPE is energy per mass J/kg, which can also be expressed as m^2/s^2 . Results on how CAPE develops in the absence of CPs are analysed with section 6.3 and compared to the CP case visualised in figure 6.6.

Details of algorithm for calculating CAPE

The following outlines how the calculation of CAPE is implemented⁸. The LES outputs the total water mixing ratio r_{total} and liquid water potential temperature θ_l and the pressure p . From these three variables we derive the background temperature profile and the profile of an irreversible adiabatic lifted test parcel (4.1.1).

⁸Which is based on code provided by the atmospheric complexity group

5.4. Calculating (CAPE) to detect regions where convection can possibly take place

We start by approximating θ_l with equation 5.2:

$$\theta_l \approx \left(\frac{p_0}{p}\right)^\kappa \left(T - \frac{L_v}{c_{pd}} q_l\right), \quad \kappa \approx \frac{R_d}{c_{pd}} \quad (5.2)$$

Where R_d is the gas and c_{pd} the specific heat capacity both for dry air, L_v the latent heat release of vaporisation, and p_0 the surface pressure. This relation (Eq. 5.2) allows to calculate the water vapour mixing ratio r_l . The water vapour mixing ratio $r_v = \frac{p_v R_d}{p_d R_v} = \frac{p_v R_d}{p - p_v R_v}$ and the liquid water mixing ratio r_l combine to the total mixing ratio $r_{total} = r_v + r_l$. This $r_{total} = \frac{\rho_v + \rho_l}{\rho_d}$ is the joint mass fraction $\rho_v + \rho_l$ of water phases relative to the dry components ρ_d .

Combining r_l with r_{total} , which is part of the LES output we can calculate the r_v and the corresponding partial pressure p_v .

To do so, we start with the law of partial pressures p_i (Eq. 5.4), where each p_i is defined by the ideal gas law (Eq. 5.3).

$$p_i = \rho_i R_i T \quad (5.3)$$

$$p = \sum_i p_i, \quad \rho = \sum_i \rho_i, \quad \text{here: } p = p_v + p_d + p_l \quad (5.4)$$

With the p_v , p_d , p_l as partial pressures of water vapour, the dry components of the air and the condensed liquid water.

Assuming that condensed water has no volume in the test parcel, thus $p_l \stackrel{!}{=} 0$ the water vapour pressure $p_v = p - p_d$ can be defined:

$$p_v = p \frac{r_v}{r_v + R_d/R_v} \quad (5.5)$$

Further assuming that p_v can not exceed the saturation pressure $p_v^*(T)$, a critical value $p_v^*(T)$ given by the empirical Claudius Clapeyron relation⁹, which we approximate with Eq. 5.6.

$$p_v^*(T) \approx 611.2 \exp\left(\frac{17.67(T - 273.15)}{T - 273.15 + 243.5}\right) \quad (5.6)$$

We check whether this critical value is reached: $p_v = \min\left(\frac{r_{total}}{r_{total} - R_d/R_v}, p_v^*(T)\right)$. For $p_v < p_v^*(T)$ we derive the absolute temperature T profile based on the dry adiabatic lapse rate Γ_{dry} . Once the critical value is exceeded, we approximate the temperature based on the moist adiabatic lapse rate given with Eq. 5.1.

⁹from https://glossary.ametsoc.org/wiki/Clausius-clapeyron_equation

5. Methods to detect convective events

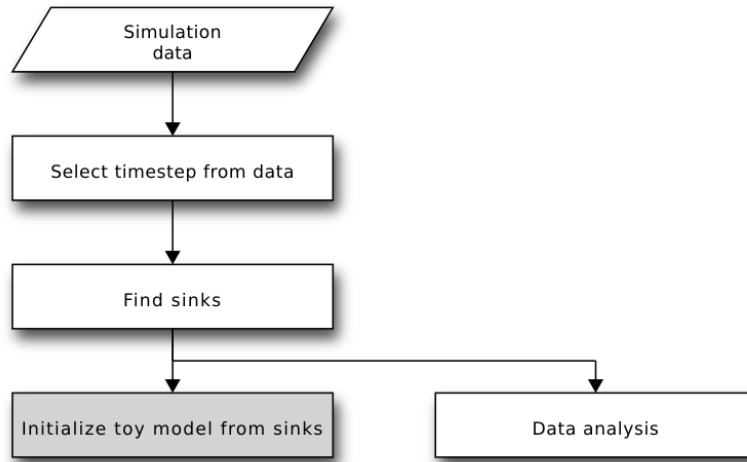


Figure 5.6.: Architecture of the algorithm, extracting sinks and basin areas in the horizontal flow of the data and initialising the model. Details of the simulation data are presented in 5.1. The central 'find sinks' routine that determines sinks and corresponding areas is introduced in 5.5.2. Results of the data analysis are presented in 6 and the model is introduced in 7

Lifting along the dry adiabatic lapse rate Γ_{dry} is straight-forward, since we lift along the potential temperature Θ , which is by definition constant for this case. Lifting along the moist adiabatic lapse rate Γ_{moist} is done step by step for each model-layer: First, the critical water vapour pressure $p_v * (T_{z=i})$ is calculated, so p_v can be set to $p_v * (T_{z=i})$. Second, the lapse rate $\Gamma_{moist}(T_{z=i}, p_v, p)$ is derived. Finally the test parcel is hypothetically lifted to the next level, by updating the temperature $T_{z=i} \rightarrow T_{z=i+1} = T_{z=i} + \Delta z \cdot \Gamma_{moist}$.

5.5. Tracking of convective events based on the velocity field

We introduce an algorithm to locate convective events in the horizontal velocity field of the lowest layer, and to detect their inflow areas as 'basins of attractions', which provides a spatial measure for their influence.

5.5.1. Basins of attraction

This section aims to substantiate the idea of basins of attraction in this context, by discussing different fix-points in the flow (visualised in Fig. 5.7):

1. **Stable fixed points:** updraught locations, which are sinks and centres of convergence in the horizontal velocity field

5.5. Tracking of convective events based on the velocity field

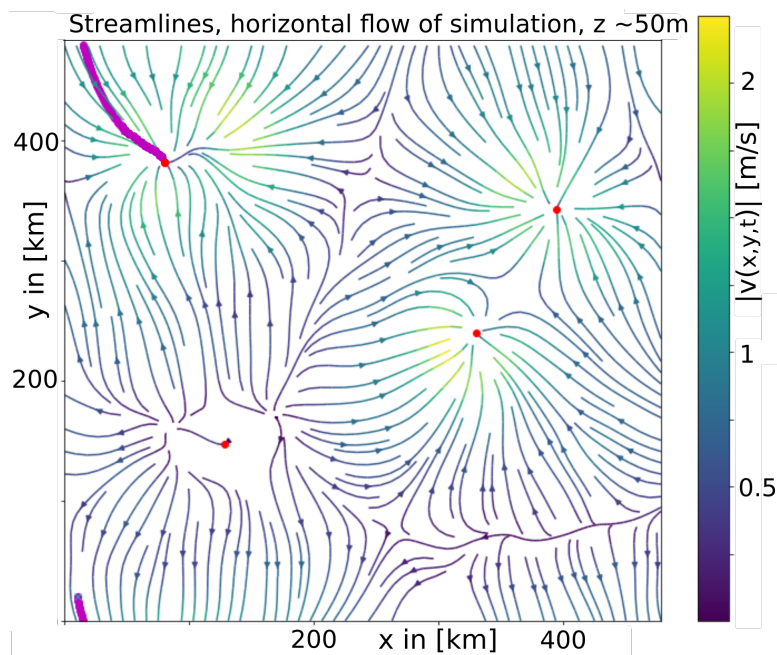


Figure 5.7.: A tracer's trajectory (purple dots) along the lowest level streamlines ($z \approx 50$ m) towards an updraught-sink (red dot). The streamline colouring indicates the local speed $|v(x_i, y_i, t_k)|$, which generally increases towards the updraught-sinks. Notice the nearly faded convective cell in the bottom left which hardly has any inflow left and how the algorithm considers the double periodic boundary conditions.

5. Methods to detect convective events

2. **Unstable fixed points:** Areas of pure divergence sometimes emerging at borders between different basins of attraction, when the velocity is zero and the flow changes direction
3. **Saddle-points:** fixed points with a stable (inflow) and an unstable (outflow) direction, which are also located at the border between different basins of attraction

Following the streamlines from any arbitrary pixel will always end in one of the stable fix points. This is true, as long as step width and convergence time are chosen large enough: The path can not end in an unstable fixed points then, since an infinitesimal small perturbation is enough to leave those.

5.5.2. Determine basins of attraction as inflow areas of convective events

To determine the basins, we use hypothetical tracers¹⁰, place them on a square grid with 10 km lattice spacing and follow the streamlines of a simulation snapshot. After smoothing the low-level wind field with a Gaussian filter (spatial kernel 10 km, no temporal filter) these streamlines are vortex free and trajectories of hypothetical tracers follow these streamlines unambiguously. Therefore, we can iterate through all grid point (x_i, y_i) and follow the flow:

$$x_{i+1} = x_i + dt \cdot u(x_i, y_i), \quad y_{i+1} = y_i + dt \cdot v(x_i, y_i) \quad (5.7)$$

$$\text{with } di = \frac{a}{\sqrt{u(x_i, y_i)^2 + v(x_i, y_i)^2}} \quad (5.8)$$

We use a bi-linear interpolation to derive the flow at locations between the grid-points provided by the 1 km resolution and take the local velocity into account to determine the step width di , which by design is smaller at higher velocities. Step width scales with a factor a , but $a = 1$ finds a good balance between high accuracy and high calculation effort, as empirical investigation shows. After a given number of iterations (chosen sufficiently high in the magnitude of the domain size) the tracer's trajectory converges to a stable fixed point corresponding to an updraught-sink. Then, we check which trajectories end within a 5 km range around an updraught-sink and then sort the start points of these trajectories to the corresponding updraught-sink-object. With this procedure all horizontal grid-points are sorted to their associated sinks. The choice of grid size means balancing computational costs with the accuracy at which the basin areas are detected. The resulting framework allows to divide the domain into 'influence areas' of convective updraughts. For further analyses, properties known to be linked to CSA, such as radiation, humidity or cloud

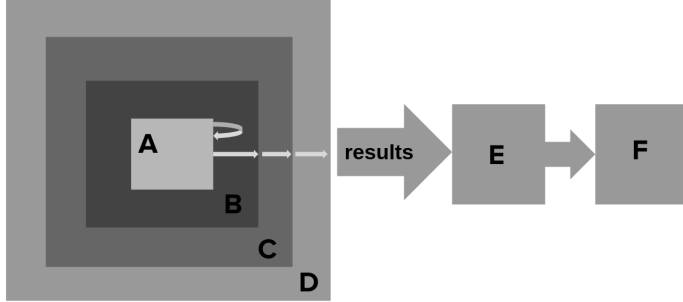
¹⁰Which is originally a common approach to detect the movement of convective events or rain cells e.g [CHR10]

5.5. Tracking of convective events based on the velocity field

cover could be systematically evaluated (e.g with regard to relative basin area) and their development could be tracked as aggregation proceeds over time.

Find sinks by following the flow

Structure of the Algorithm



The algorithm used to determine the basin of attraction belonging to each point-sink is structured as follows:

- A: The core is a function 'interpolate flow', that interpolates the flow in x $u(x_{in}, y_{in})$ and $v(x_{in}, y_{in})$ for a given floating number coordinate-tuple (x_{in}, y_{in}) . The periodic boundary conditions are taken into account, to ensure that the returned new coordinate-tuple 5.8 lies inside the domain.
- B: The function 'Follow-flow' calls function A for a given number of steps t_{max} , starting from a given input point (x_0, y_0) . The return of each interpolation step (each call of A) is iteratively used as input for the next call. The trajectory visited, following the flow from a starting point, can be stored in a list (x_{vis}) .
- C: The streamline-trajectory is determined for starting points on a regular grid of the $L_x \times L_y$ domain. The output result is stored in a list, with entry containing the start and end coordinates for each of the $i = L_x \cdot L_y$ pixels in the domain: $result = [(x_0^i, y_0^i), (x_{final}^i, y_{final}^i)]$ with $i = 0, \dots, L_x \cdot L_y$
- E: The endpoints of the trajectories are then sorted into groups¹¹. The first endpoint opens a new endpoint group. If x_{final}^2, y_{final}^2 is close to x_{final}^1, y_{final}^1 it converges to the same sink. Otherwise a new sink group is started. We then iterate over all endpoints and sort x_{final}^i, y_{final}^i , whereby we start a new sink group whenever an endpoint is not close to the first element of an existing group. Since the spreading of the endpoints is small compared to the area of later simulation stages, this assignment works well when choosing the 'is close' range sufficiently large. After sorting all trajectories, the centre of the

¹¹We use the simple numpy function 'isclose'. To increase the performance and reduce computational costs, this step could be improved by a suitable clustering algorithm, which does not require information on the number of clusters that should be detected.

5. Methods to detect convective events

endpoint group is determined: these coordinates are referred to as the sink locations.

F: Finally, a new class 'sink' is built, to organise the data and simplify post-processing. We create a sink object for each sink we set, when initialising the field Q (see A.3.3). Each sink object contains the information of all starting points, that 'flow' into this sink and thus build it's area of attraction. The number of these points is proportional to the area of the basin. The area later determines the strength factor, which together with the x- and y- coordinate of the sink is enough to reconstruct the flow in our model (introduced with chapter 7).

5.5.3. Discussion of the basin approach

The basin approach allows to divide the domain in areas of influences corresponding to individual convective updrafts and to track the development of these basins over time. The basin algorithm offers a wide range of possibilities for data analysis, which has not been exhausted yet. In order to analyse large data in an automated way the algorithm should be improved, to make the tracking more robust and to save computational cost.

As discussed for the cluster algorithm (Sec. 5.3.3), a main difficulty is choosing parameters that are appropriate over a long period of the simulation. To address this issue, it could be reasonable to determine parameters dynamically, for example depending on the number of basins detected in the previous simulation time step ($N_{basin}(t-1)$). To enhance the determination of updraught-sink locations from the end points of our tracer-trajectories, cluster algorithms from a machine-learning context could be used. Also, it could be reasonable to consider additional variables as the vertical velocity w to determine the sinks (implementation step (E)) more unambiguously. Locating the updraught-sinks independently from the tracers would allow to split step (E) in: First, determine the updraught-sinks locations (E.1) and second sort all grid points to these updraught-sinks, depending on their tracers' trajectories (E.2). In this case, we would not have to place the tracers on a regular grid, but could save computational cost by using approximations. An example would be to start with a coarser grid and determine the trajectories of the tracers. This coarse grid is then only divided further, if not all vertices' tracers converge to the same grid.

One parameter that is particular difficult to determine is the radius r , which determines whether two updraught-sink at t_k and t_{k+1} are considered the same. We try to choose r to the distance $D_i^{t_k \rightarrow t_{k+1}}$ an updraught-sink is typically displaced from $t_k \rightarrow t_{k+1}$. $D_v < r$ provides a lower bound for r , but typical values of $D_i^{t_k \rightarrow t_{k+1}}$ seem to increase over the simulation.

Earlier in the simulations, updrafts are located closer to each other with small

5.5. Tracking of convective events based on the velocity field

distances $D_{i \leftrightarrow j}^{t_k}$ between their centres (x_i, y_i) , (x_k, y_j) . $D_{i \leftrightarrow j}^{t_k} > r$ provides an upper bound on r , such that: $D_{i \leftrightarrow j}^{t_k} < r < D_i^{t_k \rightarrow t_{k+1}}$.

One possible solution would be to chose a dynamical parameter $r(t)$, which increases as $N_{basin}(t-1)$ decreases. Alternatively, larger simulations could be analysed period-wise, assuming that the dynamic changes when aggregation advances.

6. Data analysis

The following chapter presents the results of the Large Eddy Simulations (LES) and Cloud Resolving Model (CRM) data analysis, using the tracking algorithms presented in Sec. 5. First, effects of reduced re-evaporation of rainfall are presented (Sec. 6.0.1). Then NoEvap and FullEvap data is analysed based on rain tracking (Sec. 6.1), aiming to contrast the dynamics with and without CPs and to characterise the transient approach to CSA in the absence of CPs. Sec. 6.2 briefly takes up oscillations, observed in rain intensity $I(t)$ and Sec. 6.3 investigates the convective potential in NoEvap and FullEvap simulations (implementation Sec. 5.4). An analysis of convective dynamics, based on the 'basin approach' (see Sec. 5.5.2) is presented with Sec. 6.4. A change in typical spatial length scales, characteristic for aggregation, is analysed in Sec. 6.5.

6.0.1. Reducing the re-evaporation of rain

In order to align this analysis to previous work, we start with the same, relatively fine-resolution, simulations as used in Boye Nissen and Haerter [BH19]¹. We visualise the effect of reducing rain re-evaporation² in LES simulations (simulation details: Sec. 5.1): Strongly decreasing the re-evaporation opens up for convective self-aggregation (Fig.6.1).

Aggregation³ appears at re-evaporation rates between 20% (Evap02) and 60% (Evap06) of the realistic rate used as control simulation (FullEvap), which can be seen qualitatively by observing the persistently inactive sub-regions opening up in the respective panels, over the first three days (Fig. A.1).

In the Evap02 simulations a more 'band like' aggregation state forms, where rain cells align in a chain (Fig. 6.1, day 4). In the NoEvap a reaction-diffusion-like coarsening dynamic seems to emerge [WH19][CM13], forming small dry and moist patches that gradually grow and merge to larger structures (Fig. 6.1, E). Comparing mean rain intensities $\overline{I(t)}$ shows, that $\overline{I(t)}$ does neither continuously increase or continuously decrease with the re-evaporation rate (not shown)⁴. Simulation with 60% of the re-evaporation show higher $s\overline{I(t)}$ (see glossary, 2) than the control simulation with

¹Later we show, that the qualitative dynamics are similar for coarser resolution (Sec. 6.1.3), which allows to investigate larger domains and numbers of convective events with similar computational effort.

²again, by reducing the ventilation coefficients in the Seifert scheme

³How aggregation can be measured is introduced in 4.2

⁴Investigating this observation further is out of scope for this project and would need additional, longer simulations.

6.1. Analysis of aggregation dynamics based on rain tracking

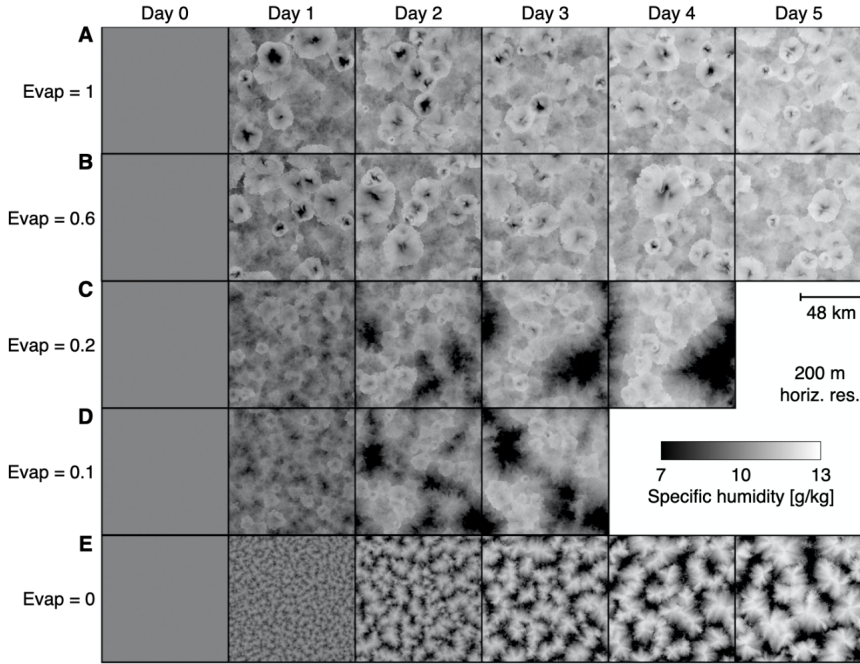


Figure 6.1.: Humidity pattern depending on the re-evaporation of rain. This reproduction of figure one in [BH19] shows the specific humidity within the lowest model level at 50 m height, which shows different patterns, depending on the fraction of re-evaporation ('Evap') compared to realistic evaporation rates (FullEvap, with 'Evap' = 1). Typical aggregation patterns occur for 'Evap' = 0.2 or less.

realistic re-evaporation. The 20% re-evaporation case has significantly reduced $\overline{I(t)}$ and the 10% re-evaporation case even slightly lower values (not shown). In contrast, $\overline{I(t)}$ is highest for NoEvap (always considering the interval of the shortest simulation, Evap01). We observe, that dry and moist patches for Evap02, Evap01 (Fig. 6.1, C, D) are larger than in NoEvap⁵, which could be linked to the reduced domain mean rain intensity.

6.1. Analysis of aggregation dynamics based on rain tracking

This section analyses the dynamics of rain clusters, detected with the rain tracking algorithm (details: Sec. 5.3). First of all, the qualitative difference in the rain field between simulations with and without CPs is presented (Sec. 6.1.1). We subsequently describe observed merging (Sec. 6.1.2) and discuss the effect caused by changes in model resolution (Sec. 6.1.3).

⁵It would be interesting to explore, whether there is any (self)-similarity between the pattern occurring in Fig. 6.1, C, D and 6.1 E.

6. Data analysis

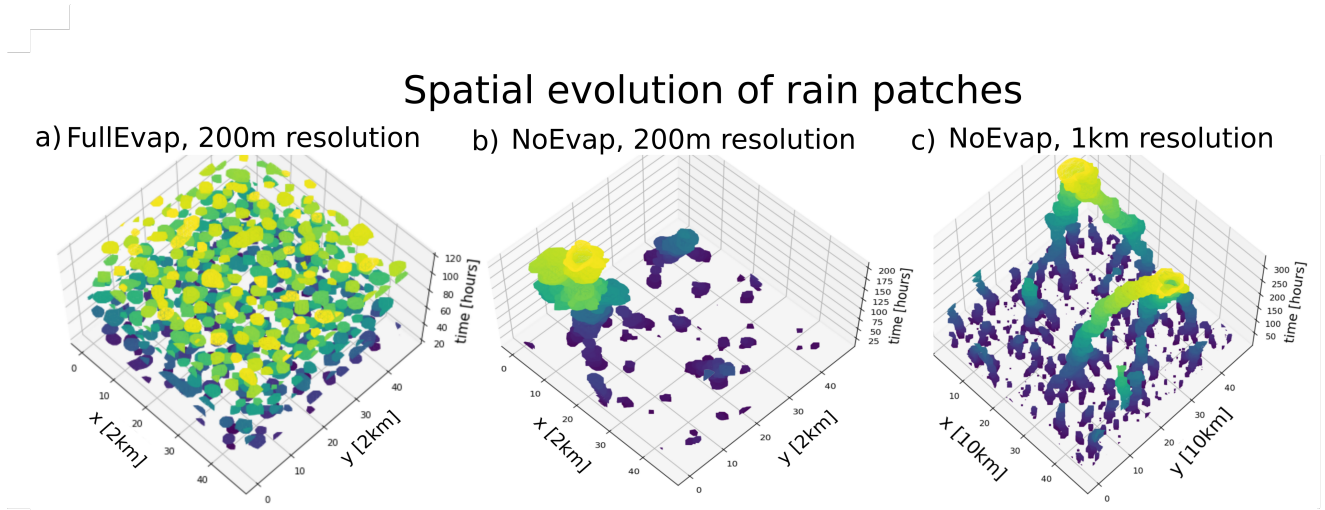


Figure 6.2.: Rainfall pattern $I(x,y,t)$ contrasted for FullEvap (a) and NoEvap (b) cases with 200 m resolution and NoEvap case with 1 km resolution (c). Active pixels (glossary) above a 0.1 mm h^{-1} rain threshold are plotted every three hours, for the FullEvap (a) and NoEvap (b) case. Colour shades indicate the relative time within the simulation, ranging from the early stage (blue) to the final state (yellow). With realistic re-evaporation, no aggregation occurred in the simulated period of 15 days on the small $96 \text{ km} \times 96 \text{ km}$ domain. In contrast, in the absence of CPs the simulation aggregated in less than eight days. At that time rain is restricted to one single contiguous area. Within the transient approach to aggregation, spatially separated rain patches move towards one another, a dynamic that leads to tree-like merging. Other rain patches decay as their area decreases. The NoEvap simulation with 1 km horizontal resolution (c) shows aggregation dynamics similar to the smaller $96 \text{ km} \times 96 \text{ km}$ domain (b), but takes longer to fully aggregate (nearly two weeks instead of about eight days). Notice that due to the double-periodic lateral boundary conditions, 'branches' can leave and re-enter the domain at the respective boundaries of each lateral dimension, as it is the case when the final two 'branches' merge into a single rainy area.

6.1.1. Dynamics with and without cold pools (FullEvap versus NoEvap)

In FullEvap simulations the number of rain cells, detected as $N_{cluster}$ (see glossary) rapidly increases with the onset of convection ($t \approx 10$ h, until it quickly settles at $N_{cluster} \approx 100$ (Fig. 6.3, left). Subsequently, $N_{cluster}$ somewhat fluctuates around this constant value⁶ and so does the number of new cells emerging in a ten-minute interval ('birth rate', spread around ten), yielding an average lifetime of about 100 min per rain cell (Fig. 6.3, a). This 'birth rate' is approximately balanced by merging and 'decay' of rain cells ($N_{cluster} \approx \text{constant}$, Fig. 6.3,b). We use the term 'decaying' to describe the slow fading of a rain cell, which is indicated by a decrease of the cluster's area ($A(\text{cluster},t)$) and rain intensity ($I(\text{cluster},t)$). Merging describes a process where two rain clusters move towards each other, spatially overlap and result in one bigger cell.

In FullEvap rain cells terminate by decay rather than merging (Fig. 6.2, [MHH19]): A CP forms through rain re-evaporation and the resulting dry downdraught locally suppresses further convective updraught. Merging processes do not seem to be a crucial contributor to the dynamics and might simply result stochastically from events joining spatially when they are close to each other, without any far-reaching physical interpretation.

In contrast, the NoEvap dynamic is fundamentally different: After a spin-up stage of roughly 30 hours, $N_{cluster}$ monotonously decreases (Fig. 6.3, c) until only one rain cell remains (final stage, after ~ 350 h). Interestingly, the maximum number of rain cells ever reached ($N_{cluster} \approx 100$) is comparable to that in the FullEvap case. This congruence should be expected, because initial events appear almost simultaneously when convection starts, so the effect of CPs is hardly present. What determines the number of initial events should be the global energy constraint of the RCE simulations (Sec. 4.1.5). In the NoEvap case it takes less than three days, until hardly any new rain cells emerge. Again, two processes reduce their number: decaying and merging.

In the NoEvap simulations, merging is the dominant dynamic after 150 h (Fig. 6.2). Stages before 150 h are generally more noisy and decays are detected more often than merges, leading to a fraction $\frac{n_{merges_{NoEvap}}}{n_{decays_{NoEvap}}} \approx \frac{2}{3}$ over the entire NoEvap simulation. As mentioned, merging hardly plays a role for FullEvap, leading to $\frac{n_{merges_{FullEvap}}}{n_{decays_{FullEvap}}} \approx \frac{1}{40}$, which is more than 25 times smaller than the merging fraction in the absence of CPs (NoEvap case). Interestingly, Patrizio and Randall found pairwise merging when analysing CSA on much larger domains up to 6.144 km with 3 km resolution [PR19], where the influence of CPs should be minor (more Sec. 7.3.3).

⁶Therefore we did stop the simulation after 150 h to save computational cost.

6. Data analysis

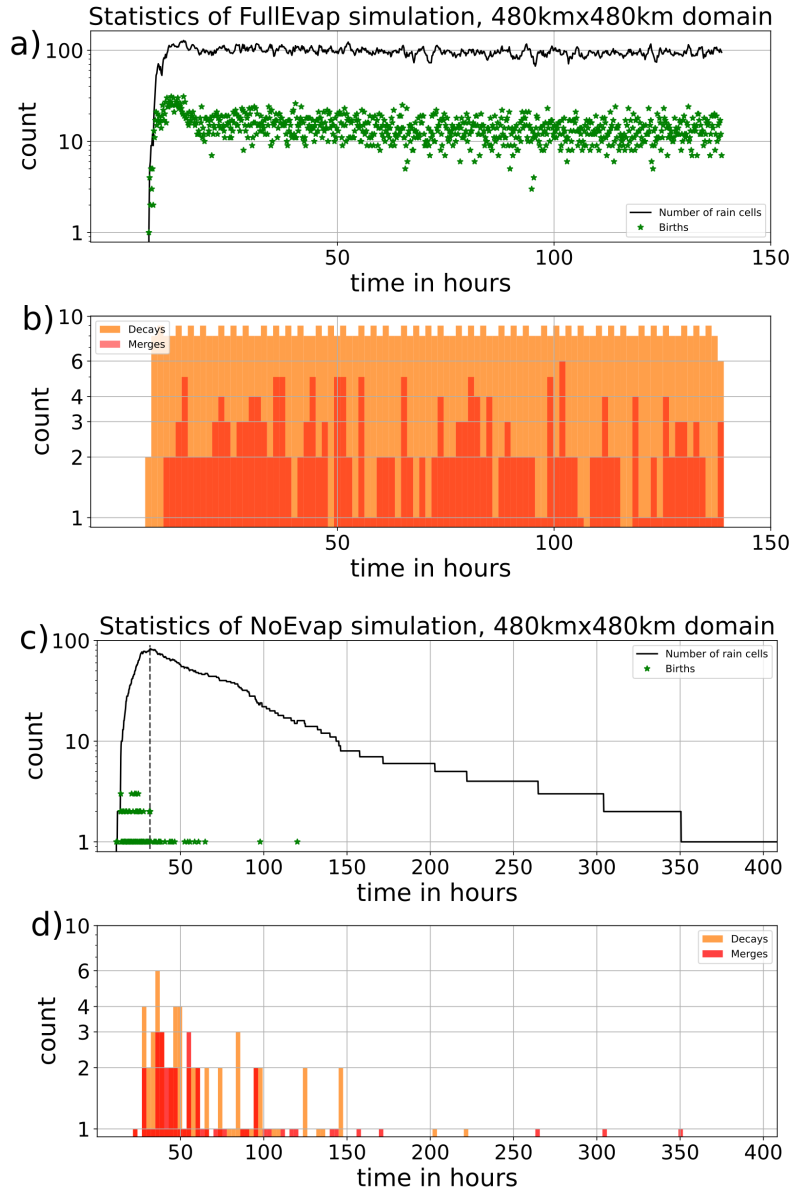


Figure 6.3.: Comparing statistics of FullEvap (a,b) and NoEvap (c,d) simulations. Panel a), c) compare the number of rain cells and the emergence of new ones (green). Panels b), d) show the count of the two processes, that reduce the number of rain events: Decaying and merging. We suppose merging does not play a crucial role in FullEvap, but is characteristic for the NoEvap, where merging is the dominate dynamic from roughly one week on.

6.1.2. Merging

For NoEvap simulations merging seems to be non-random and crucial within the transient approach to the fully self-aggregated state: When long-living rain cells merge, they move towards one another, spatially connect and become one bigger rain cell. The merging dynamic is visualised in the animation of rain clusters (supplement material) and in Fig. 6.4.

Generally, area $A(\text{cluster}, t)$ and rain intensity $I(\text{cluster}, t)$ tend to increase after a merge (Fig. 5.4). Moseley, Henneberg, and Haerter suggested before, that big rain events resulting from several merges show higher rain intensities. The investigation of these merging effects is very relevant, since it could be linked to extreme precipitation events [MHH19].

However, the physical mechanism driving this merging is not fully understood yet. One aspect could be the moisture field, which builds 'bridges' between two moist patches, thus areas corresponding to different convective cells become connected (Fig. 6.9). Further, we propose that cloud-radiative feedback, known to drive CSA, could also facilitate the merging (Sec. 8.1). Possible mechanisms driving the merging based on the surface wind field are discussed with Sec. 7.3.3. With future work, we aim to capturing this merging dynamic in our model (Sec. 7.4) and find data-based physical explanations.

6.1.3. Resolution

Coarser resolutions are known to favour CSA [Win+17]. One aspect might be, that coarse resolutions insufficiently resolve CPs (Sec. 4.1.7) which inhibit the aggregation (Sec. 4.2.4). Irrespective of the 1 km or 200 m resolution, our data shows self-aggregation in the NoEvap, but not in FullEvap. The dynamics are qualitatively similar between 1 km and 200 m resolution, but full aggregation occurs more quickly on the smaller domain with finer resolution (compare Fig. 6.2 to Fig. 6.2). Since coarser resolution allows to investigate larger domains at equal computational cost, thus more convective events at the same computational cost, we focused most of our analysis on the 1 km simulations. Former studies suggest that in the absence of CPs, aggregation strength decreases with smaller domain size [JR13]. Nevertheless, we observe very fast and pronounced aggregation for a very small NoEvap 240 km domain at 1 km resolution, but did not compare the aggregation degree between NoEvap simulations quantitatively.

6.2. Oscillations

We detect strong temporal oscillations in domain $\overline{I(t)}$ and cluster integrated rain $I(\text{cluster}, t)$ for NoEvap simulations (Fig. 6.5). On one hand, these oscillations complicate the detection of rain clusters, on the other hand the oscillations are an interesting phenomenon in themselves.

6. Data analysis

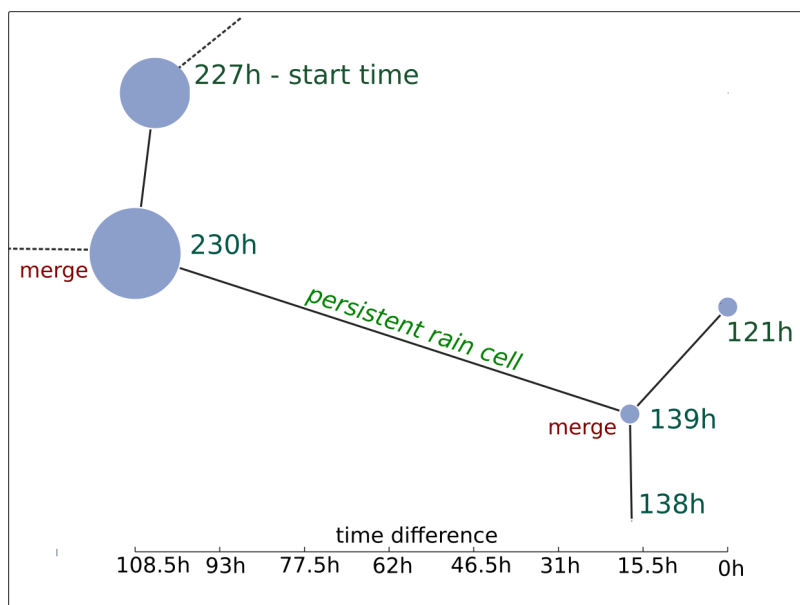


Figure 6.4.: Pairwise merging of rain cells The rain cluster tracking (Sec 5.3.2) allows to visualise the merge history of rain clusters as trees. Here, the size of a tree node scales with the mean area of the cluster $\overline{A(\text{cluster})}$ and the branch length scales with the time between the onset of clusters (written in green). Note that time increases from right to left. The horizontal axis below shows the time difference since the start of the earliest rain cell, merging into this tree. The vertical axis has no interpretation and simply aligns the nodes in a way that prevents the intersection of 'links'. This example illustrates rain cells persist for many days until they merge to one subsequent cluster, which is usually larger than its two predecessors.

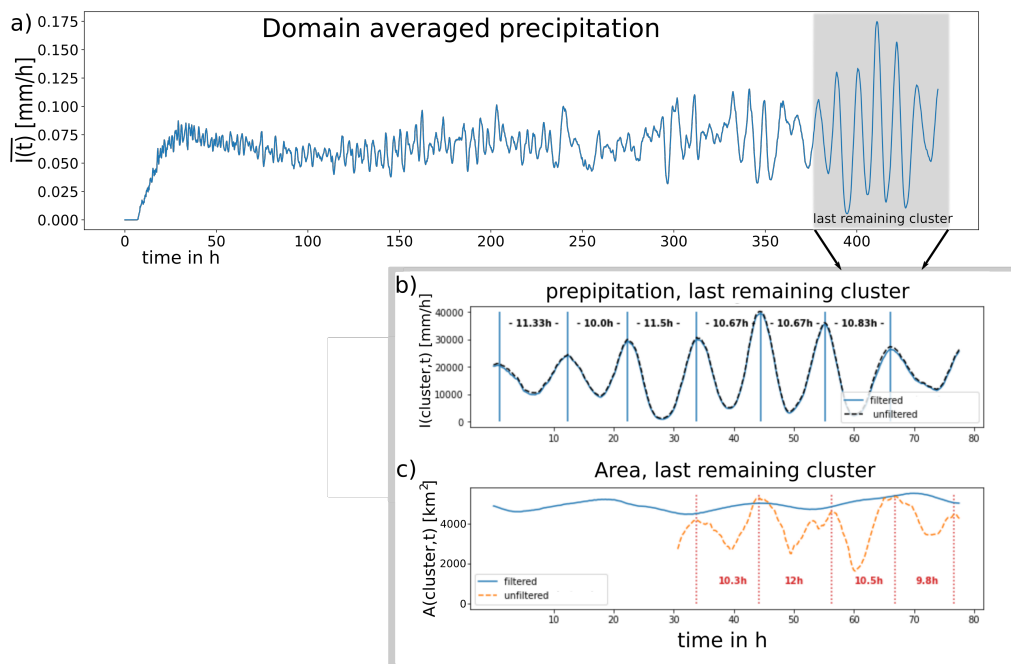


Figure 6.5.: Temporal oscillations in rain intensity $\overline{I}(t)$ (a) and detailed visualisation of $I(\text{cluster},t)$ (b) and $A(\text{cluster},t)$ (c) for the final stage. Once only one rain cell remains (350 h), a significant increase in oscillation amplitude occurs, possibly linked to radiative feedback (Sec. 8.1). This final stage of the simulation shows strikingly regular oscillations in cluster averaged precipitation $I(\text{cluster},t)$ (b) with periods between 10 h-11.5 h. Filtering prevents oscillations in $A(\text{cluster},t)$, which otherwise would oscillate with comparable periods (c) and allows us to focus on the underlying dynamics.

6. Data analysis

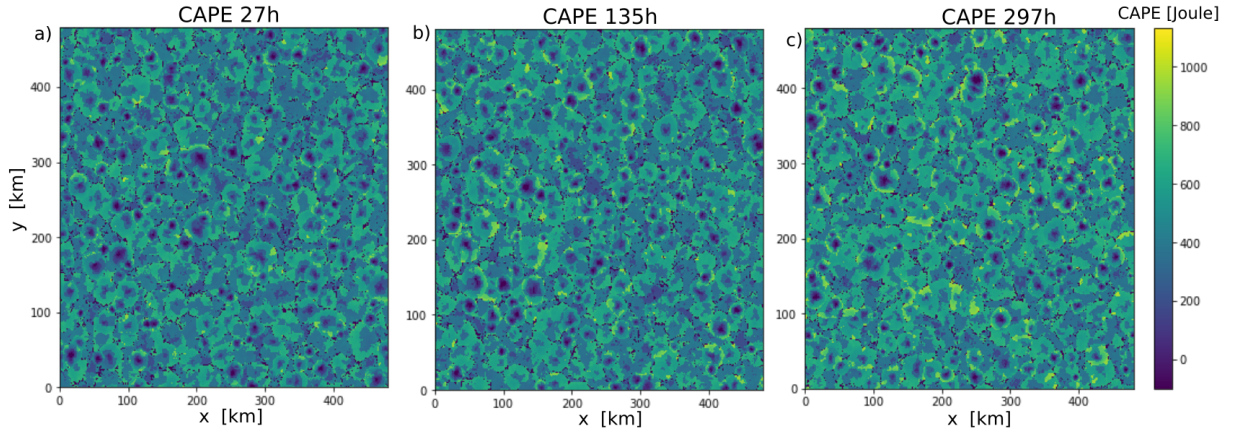


Figure 6.6.: Distribution and magnitude of CAPE are qualitatively similar over the FullEvap simulation. Circles of low convective potential indicate a dry CP inhibiting new convection after rainfall. Sometimes these circles are framed by rings of high CAPE, indicating areas where new convection is favoured.

We use filtering that smooths the temporal oscillation in $A(\text{cluster}, t)$ oscillations (Fig. 6.5, c). This filtering allows an analysis of the slow dynamics and of how cluster area increases with merging (Fig. 6.4). Nevertheless, filtering hardly influences the rain volume (Fig. 6.5, b). $\bar{I}(t)$ oscillates irregularly (Fig. 6.5, a) since the oscillations of individual rain cell are usually neither in phase nor do they have regular periods⁷. In the final aggregated state, oscillations in $\bar{I}(t) \approx I(\text{cluster}_{\text{final}}, t)$ are very regular with periods about 10 h. These oscillations are discussed further in chapter 8.

6.3. CAPE

As visualised before (Fig. 6.2), the convective dynamics is fundamentally different when the re-evaporation of rain is suppressed. 'Normal' convection is accompanied the strong downdrafts of CPs, that create circular patterns of low CAPE between a certain minimal radius R_{min} and a maximum radius R_{max} around the centre of convective cells (Fig. 6.6). This limits the persistence of individual convective cells, so the spatial distribution of convective cells changes. The qualitative dynamic and magnitude of CAPE stays constant over the simulation: besides their inhibiting effect, CPs promote convection at specific radii R_{max} away from convective centres, corresponding to high CAPE regions.

When CPs are absent (NoEvap), persistent areas of positive CAPE form (Fig. 6.7). Here, CAPE decreases monotonically with distance from the centre of convective regions and rainfall $I(x, y, t)$ is limited to areas of positive CAPE. As aggregation

⁷It seems that multiple factors influence and interrupt the periodicity of these oscillations, thus a Fourier analysis of the oscillations did not provide convincing results.

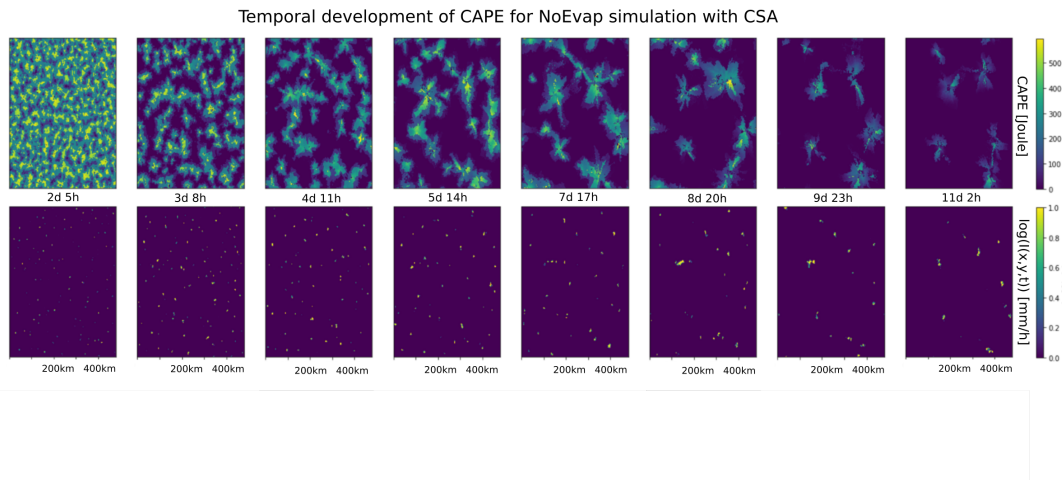


Figure 6.7.: NoEvap aggregation dynamics of CAPE and surface rain intensity $I(\mathbf{x},\mathbf{y},t)$ Convective dynamics are qualitatively different from FullEvap (compare Fig. 6.6): Rain (bottom) is limited to very isolated areas of positive CAPE (top). These areas grow over time, and their number is reduced by decaying and merging of convective cells. Qualitatively we observe that domain-averaged CAPE reduces over the simulation, which might be linked to increased static stability in the upper layers. CAPE was calculated with the algorithm introduced in Sec. 5.4.

proceeds, the number of regions with $CAPE > 0$ decreases. Considering a late stage after 330 h of the simulation (Fig. 6.8): only two rain events are left and located in the only two regions of positive CAPE. Positive CAPE is spatially correlated to a low energy barrier CIN (Fig. 6.8,b) and high relative humidity in low layers, necessary to 'feed' the convection (Fig. 6.8,e). The rest of the domain can be seen as a 'desert' area where no convection is possible, since the LFC is never reached and thus no CAPE can be generated.

In this concrete implementation (Sec. 5.4) this means that the level of free convection (LFC), where temperature profiles of a moist adiabatically lifted test parcel and the temperature profile of the background environment intersect (Fig. 6.8, g), lies above the highest simulated level.

In regions where it is present, CAPE seems to reduce over the course of the simulation. Such an increasing static stability of the atmosphere might be linked to increased static stability in the upper layers, resulting from very persistent convection. Studies state (Sec. 4.1.2, [XE89] [Bet82]) that in RCE the tropical atmosphere tends to organise towards a moist adiabatic vertical temperature profile (Sec. 4.1.5). This implies very low CAPE as the difference between the moist adiabatic temperature profile of a test parcel and the background temperature profile of the tropical atmosphere would vanish. We suggest that an infinitesimal departure $\delta CAPE$ from the moist adiabatic profile could determine where convection occurs: $\delta CAPE > 0$

6. Data analysis

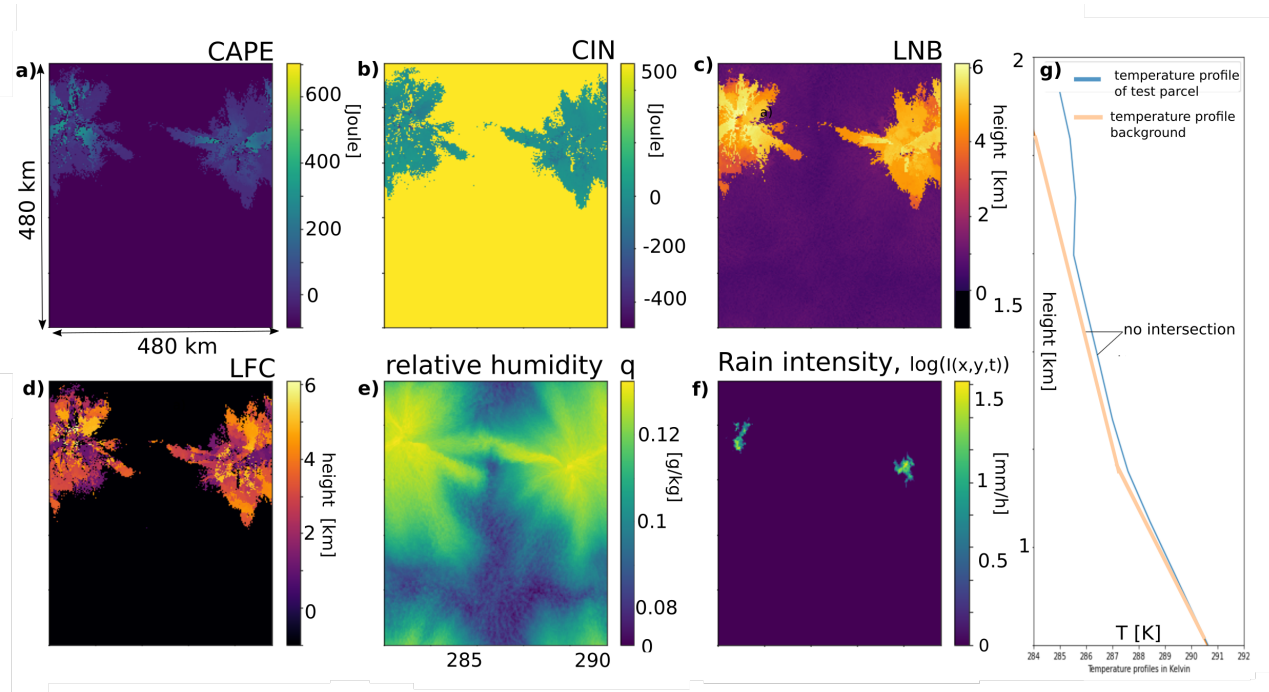


Figure 6.8.: Convective parameters in nearly fully-aggregated NoEvap CAPE (a), CIN (b), LNB (c), LFC (d), relative humidity q (e) and rain intensity (f, logarithmic plot of $I(x,y,t)$) after 330 h and temperature profiles for the special case when the LFC is never reached (g). Analysing these parameters shows, that convection seems restricted to two discrete moist areas only (e, humidity at $z \approx 50$ m), where the last two remaining rain cells are located (f). In the rest of the domain the LFC (d) is never reached, thus the temperature profiles of a moist adiabatically lifted test parcel and the temperature profile of the background environment never intersect (g). No rain falls in these 'desert' areas, and CIN (b) is set to 500 J (yellow). For these regions a continuous energy supply would be needed to lift a test parcel through the height of the atmosphere. Here, the LNB (c) has no validity, since free convection never occurs. The raining areas show CIN close to zero and the LNB (c) lies above 3 km.

could enable an updraught, whereas $\delta CAPE < 0$ could coincide with the 'desert' regions that are dominated by subsidence (Fig. 6.9). Nevertheless, if CAPE is small the question arises what actually does drive the convection. We suppose dynamical effects of persistent circulation could be an important driver (Sec. 6.4).

The combination of an increasing static stability and (partly) dynamically driven convection could possibly explain, why smaller and isolated convective areas seem more likely to fade (Sec. 6.4), which is discussed in Sec. 6.4.3. In conclusion, the convective areas with $\delta CAPE > 0$ are more consistent and informative than the rainfall. This suggests to take additional physical properties into account, in order to investigate the dynamic towards aggregation. The following approach focuses on the dynamics and analyses the flow of the low layer velocity field.

6.4. The low layer velocity field

We investigate the instantaneous simulated velocity field to define a 'basin of attraction' for each convective cell (Sec. 6.4.1.). We devise a method to numerically identify these basins and to track them from one time step to the next (Sec. 6.4.2) and discusses a possible advantage of larger basins to persist over time (Sec. 6.4.3). Finally, we discuss the spatial dynamics of updraughts, corresponding to each basin (Sec. 6.4.4).

6.4.1. Analysis of the stationary flow field

As shown above (Sec. 6.3), rain is limited to a discrete number of convective regions with $\delta CAPE > 0$ (Fig. 6.8). These convective regions correspond to strongly localised updraughts (Fig. 6.9, c). The remainder of the domain shows gentle and homogeneous subsidence, which due to continuity (Eq. 7.2) must exactly balance the updraughts. Dry regions show horizontal divergence and low wind speeds close to zero (Fig. 6.9, b). Within these dry and non-convective regions a strong horizontal velocity gradient points towards the moist updraught regions. Generally, wind speed determines the low layer moisture flux above a homogeneous moisture reservoir (see Sec. 5.1), which implies a moisture transport out of the dry regions towards the already moist regions of convective updraughts. The updraughts' dynamics are quite stationary. We can locate updraughts by following the streamlines of the low-level horizontal flow (Fig. 6.10). These updraughts' locations are the basis for our model that compresses the information of the velocity field to a distinct number of objects (chapter 7).

6. Data analysis

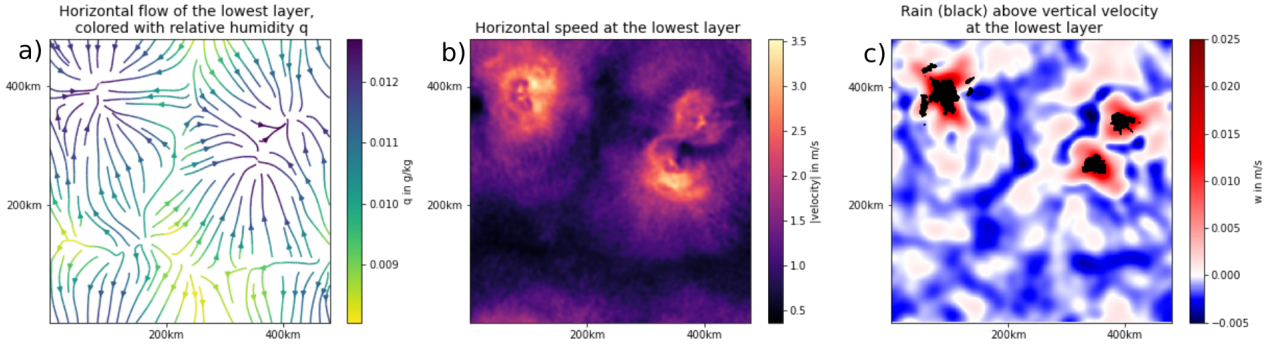


Figure 6.9.: Instantaneous field of horizontal velocity $v(x, y)|_{z \approx 50 \text{ m}}$ (a), horizontal wind speed $|v|_{z \approx 50 \text{ m}}$ (b) and vertical velocity $w(x, y)|_{z \approx 50 \text{ m}}$ (c) at 300 h. (a) Streamlines are coloured with relative humidity $q(x, y)|_{z \approx 50 \text{ m}}$, start from dry areas (yellow) of divergence and converge towards moist areas (dark blue). The dry regions (a), corresponding to minimal wind speeds $|v(x, y)|_{z \approx 50 \text{ m}}$ (b). This $|v(x, y)|_{z \approx 50 \text{ m}}$ increase with radial distance from the centres of dry regions towards the moist centres of horizontal convergence. These convergence centres correspond to areas of updraughts (c), that is, vertical divergence. Surface rain ($I(x, y, t) > I_0 = 0.5 \text{ g kg}^{-1}$, plotted in black, c) is limited to these convective updraught regions. Outside these regions mostly subsidence ($w(x, y)|_{z \approx 50 \text{ m}} < 0$) dominates.

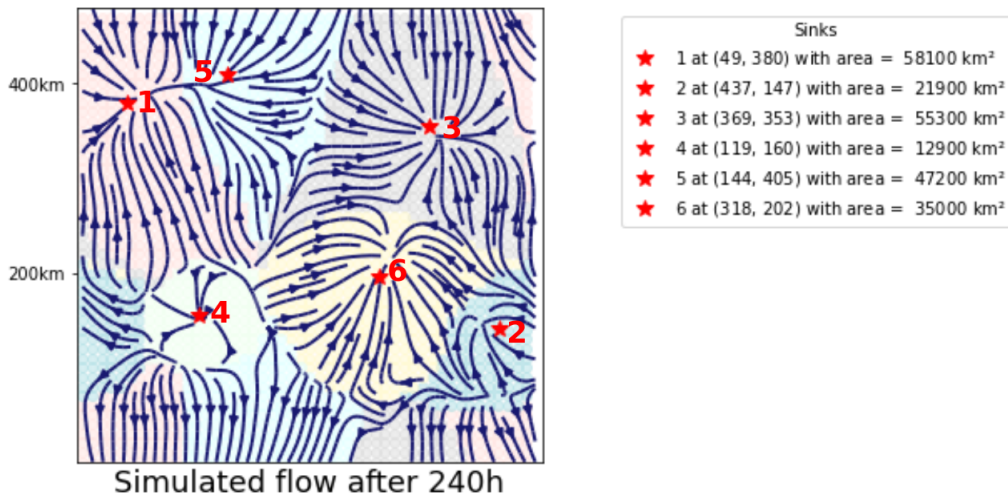


Figure 6.10.: Streamlines of the horizontal velocity above basins of attraction. Red stars mark the updraughts, which are sinks and centres of convergences in the horizontal flow of their corresponding basins. All properties are detected with the algorithm introduced in Sec. 5.5.2.

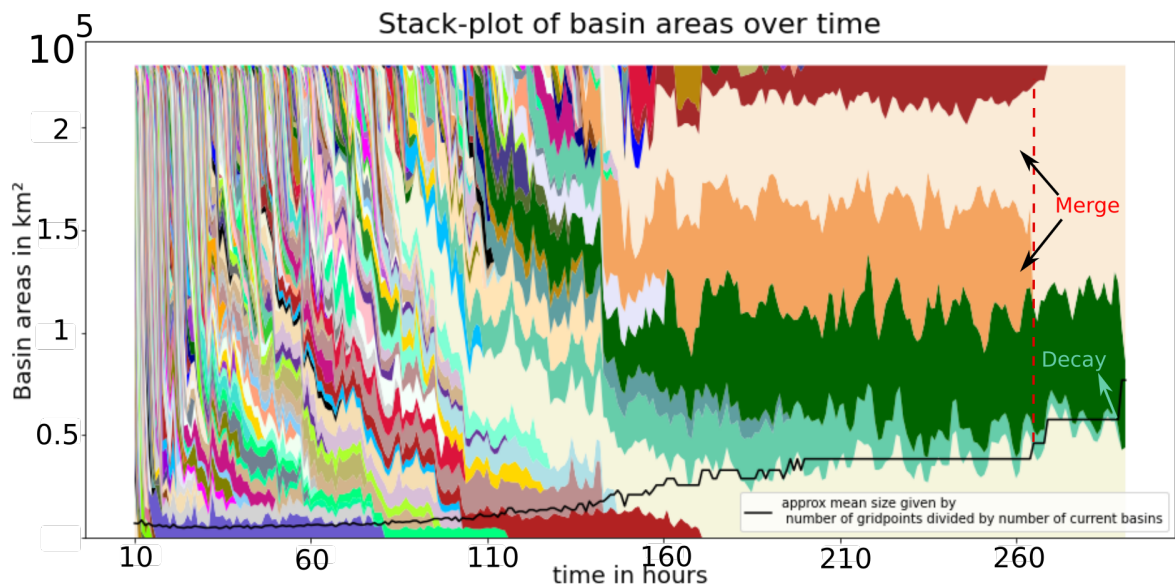


Figure 6.11.: Temporal development of basin areas $A(t)_{basin}$. The area of a basin corresponds to the height of its colour coded area, which are stacked above each other. Notice that the basins are domain filling.

Thus, $\overline{A(t)_{basin}}$ increases as the number of basins $N(t)_{basin}$ decreases over time. The first 150 h show an very unclear picture of many, often short lasting convective cells. Smaller basins gradually decay, causing larger basins to gradually grow. Significant increase in area is observed in discrete steps, when two basins merge.

6.4.2. Development of basin areas over time

The following describes how the velocity field changes over time, as aggregation proceeds. Change can be measured by a decreasing number of updraughts, by different positions of their centres and changing area of the corresponding basins.

Because the number of currently detected basins ($N(t)_{basin}$) decreases over time (Fig. 6.11), on average the area of individual basins ($A(t)_{basin}$) has to increase, since the basins are space filling: $\sum_i A(t)_{basin_i} = A_{domain}$.

Concluding, domain area $A_{domain} = 480 \text{ km} \times 480 \text{ km} = 230\,400 \text{ km}^2$ divided by $N(t)_{basin}$ represents an average basin area: $\frac{A_{domain}}{N(t)_{basin}} = \overline{A(t)_{basin}}$, which is plotted black in Fig. 6.11.

Despite a continuous increase of $\overline{A(t)_{basin}}$, the individual areas $A(t)_{basin_i}$ differ and so does the development of the corresponding basins:

Smaller basins seem to gradually decrease, causing larger basins to gradually increase. Significant increase in area does not result from such 'growing', but takes place in discrete steps, when two basins merge to one (Fig. 6.11).

6.4.3. Discussing possible advantages of larger updraughts

Larger convective cells show wider circulation and faster wind speeds in the inflow area (Sec. 7.6), corresponding to larger $A(t)_{basin_i}$. We suppose a general advantage for the convection of larger updraughts persist over time and suggest an explanation in the following:

When moist air rises, it is further accelerated by latent heat release, creating a low pressure and a pull towards convective regions in low layers, resulting in stronger horizontal wind speed which generally⁸ can collect more moisture (bulk formula, Eq. 4.9, Sec. 4.2.3). Such a positive feedback loop could increase the separation of dry and moist areas: More moisture is transported towards stronger updraughts, which then further accelerates the updraught and promotes already larger updraughts. Meanwhile, these larger updraughts have a stronger heating effect on the free troposphere: Due to the weak temperature gradient approximation (Sec. 4.1.6) this heating energy will be spread laterally by gravity waves and thereby reduce the vertical temperature gradient and thus convective potential (Sec. 6.3, Sec. 4.1.2) of the domain, including regions of weaker updraughts.

Larger updraughts with larger and faster inflow could provide the energy needed to overcome CIN dynamically and keep their convection going. Smaller convective cells with weaker circulation could be more likely to decay as the static stability in the atmosphere increases (Sec. 4.2.3).

Such a decay is indicated by a decreasing inflow accompanied with decreasing basin

⁸Once humidity is close to saturation, the moisture content does not increase with wind speed anymore.

areas (Fig. 6.11).

In addition, we observe that rainfall usually stops several hours before the circulation decays. By design, a global radiation constraint (Sec. 4.1.5) limits the total rainfall $I_{\text{total}}(t)$ on the domain. Therefore, this total supply of rainfall has to be partitioned among the different convective cells⁹, which can interact via gravity waves (as suggested above). Consequently, increasing rainfall $I(\text{cluster}, t)$ and convection in larger basins affect smaller basins adversely.

6.4.4. Lateral updraughts dynamics

We further observe that the lateral positions of updraught centres are not stationary over time. Updraught centres move around with typical speeds up to 1.4 ms^{-1} (5 km h^{-1}) for stages after 150 h (not shown). These later stages have fewer basins, which are more persistent and the convergence at the lowest level is less ambiguous to analyse. Again, we observe a tendency of updraughts to move towards each other and fuse to one convective cell. Based on these observations the dynamics of a phenomenological model can be developed.

6.5. Dynamical scaling

Craig and Mack present a coarsening model (Sec. 4.3.3, [CM13]), based on the re-evaporation of rain as the crucial re-moistening effect in the positive feedback loop towards CSA. Interestingly we observe strong aggregation in NoEvap simulations, where the re-evaporation feedback is explicitly suppressed.

When comparing their results to CRM data, Craig and Mack refer to a study provided by Posselt et al. [Pos+12] which simulates large channel-like¹⁰ domains of 180 km expand in the meridional and up to 10.000 km in the zonal direction (details in [PVS08]). These channel domains are simulated in an RCE framework, aiming to investigate the response of the tropical hydrological cycle to surface warming. Based on a Hovmöller plot (Fig. 2, [Pos+12], OLR for different SSTs) Craig and Mack suggest a qualitative measure to estimate a dynamical scaling factor: An average patch area A_p is obtained by dividing total domain area by the count of dry and moist patches, so that in the course of aggregation A_p will increase. Craig and Mack obtain a power law dependence of the A_p on time t ($A_p \sim t^{\tilde{\alpha}}$, area scaling (AS)).

⁹Therefore significantly higher $I(\text{cluster}, t)$ values are reached for the last remaining convective cell

¹⁰Channel-like domains are a common compromise, since a 3D domain needs to be large enough to not constrain the behaviour of the convective ensemble, but usually available computational resources are limited. Nevertheless, it is not known exactly how such a change in geometry influences the results. We experimented with a channel-like NoEvap simulation, which aggregated very fast within a few days.

6. Data analysis

Assuming that the AS is roughly proportional to length scaling (LS, $A \sim L^2$) implies a power-law LS too: $L_p \sim t^{\tilde{\alpha}}$. For this LS, Craig and Mack state the scaling factor of the simulation with highest SST (302 K) is similar to the scaling obtained from their model: $\alpha_{302\text{ K}} = 0.44 \approx 0.5 = \alpha_{model}$. Despite higher SST was found to promote aggregation speed in some models (Sec. 4.3.2, [EWV14]) the cooler simulations seem to yield larger scaling factors, e.g. $\alpha_{330\text{ K}} = 0.74$ and $\alpha_{298\text{ K}} = 0.81$. Posselt et al. argue that cooler simulations show a different divergent mid-layer circulation¹¹, which transports moisture from moist to dry regions more systematically. Craig and Mack speculate that since replacing diffusive mixing by hydrodynamic streaming flow leads to larger scaling exponents in theoretical coarsening models [Bra03], this change in divergent circulation could lead to larger α despite lower SST.

With respect to the suggested importance of the mixing term, it seems reasonable that our NoEvap simulations, which have little mixing dynamic, show a very different LS. To derive a comparable measure to the counting method suggested by Craig and Mack, we start by calculating the inverse of the number of convective cells ($\frac{1}{N}$). We derive $\frac{1}{N}$ for both our methods and calculate $N(t)_{cluster}$ and $N(t)_{basin}$ for the interval $30\text{ h} < t < 310\text{ h}$, which starts after the spin-up and ends before less than three rain cells are left.

Notice, that $\frac{1}{N(t)_{basin}}$ is a normalised version of dividing the domain size by N , providing a temporal area scaling (AS) (compare to $\overline{A(t)}_{basin}$, Sec. 6.4.2). To obtain a LS comparable to the results of Craig and Mack, we suggest to approximate basin areas a circular shape. Furthermore, we derive a measure for the AS by calculating $\frac{1}{N(t)_{basin}}$ and derive a similar measure $\frac{1}{N(t)_{cluster}}$, although we are aware that the rain clusters are not domain filling. We find that over the first 130 h, $\frac{1}{N(t)_{basin}}$ follows an exponential fit (Fig. 6.12, black), which indicates an exponential AS and thus an exponential LS, again assuming roughly circular basin shapes ($\exp(\text{LS}) \approx \exp(\text{AS}^{1/2})$). We suppose the exponential time dependence could be linked to dominating stochastic processes in this early coarsening stage.

The curve $\frac{1}{N(t)_{cluster}}$ leaves the exponential fit earlier (100 h), for which we suggest multiple explanations.

First, for $30\text{ h} < t < 100\text{ h}$ we detect $N(t)_{cluster} > N(t)_{basin}$. We suggest that in this time interval, several detected raining patches could result from the same convective cell or that the updraught-sinks are too close to each other be detected as distinct basins. After about 200 h $N(t)_{basin} > N(t)_{cluster}$, which relates to the fact that updraughts (basins) are usually still present after the corresponding rainfall (rain cluster) already decayed (Fig. 6.12).

Concluding, we observe a temporal shift between the two measures. In addition, rain clusters not only increase in area, but also the rain density inside the cluster area increases (not shown), which makes it more complex to obtain a scaling

¹¹A divergent mid-layer circulation is discussed to be necessary for CSA to developed [MH12]

for the actual change in convective activity. Nevertheless, results are qualitatively comparable between methods and we unambiguously do not observe any power-law scaling. Thus, an LS comparable to $t^{1/2}$ ($AS \approx t$, Fig. 6.12, green), as Craig and Mack propose for the early coarsening stage, is not valid here. Whether or not later coarsening stages (after 150 h) could show similarity to such a power-law can not be answered at the moment, due to the limited amount of data and resulting small N .

Interestingly, Windmiller and Hohenegger repeat their claim of an universal LS in the evolution of CSA ([WH19], 2019). They investigate scaling parameters for established CSA models (Sec. 4.3.4) and RCE simulations performed by Hohenegger and Stevens[HS16]. After a spin up of roughly one day the average autocorrelation length of these simulation are claimed to follow the mentioned $t^{0.5}$ scaling parameter, assumed for the coarsening process. Longer simulations show departure from $t^{0.5}$ for later stages. This departure suggests coarsening becomes less important, which is in line with our hypothesis, that later stages are dominated by the interaction between individual convective cells.

Notably, Windmiller and Hohenegger observe a more rapid increase in LS for early stages of different simulations and models, which limits the significance of $t^{0.5}$ to specific intervals of coarsening processes¹². In general, we suppose observing a 'faster' LS increase than Craig and Mack in our NoEvap simulations, could coincide to the absence of CPs, since CPs are known to inhibit CSA.

Future work will improve our analysis of the spatial scaling with larger data sets and systematic parameter optimisation for the simulation interval, considered for the dynamical scaling (Sec. 5.5.3).

¹²If this $t^{0.5}$ intervals significantly change between simulations and models, the claim of universality could be questioned.

6. Data analysis

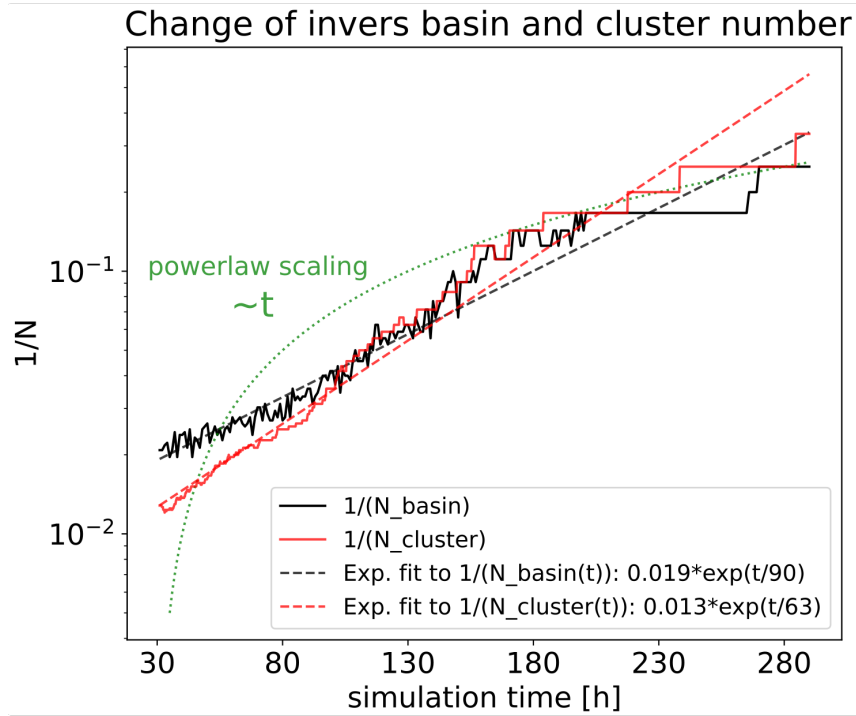


Figure 6.12.: Dynamical scaling of $\frac{1}{N(t)_{basin}}$ and $\frac{1}{N(t)_{cluster}}$ typical for coarsening processes (NoEvap simulations). We offer exponential fit for $\frac{1}{N(t)_{basin}}$ (black) and $\frac{1}{N(t)_{cluster}}$ (red) with $30 \text{ h} < t < 100 \text{ h}$. The green dotted line presents a linear area scaling, as suggested by Craig and Mack. We state that any power-law dependence is not valid for the early stage of coarsening we observe in our NoEvap simulations.

7. Model

Our goal is to define a discrete and compressed representation for the velocity field analysed in Sec. 6.4. Therefore, we will assume very localised updraughts, balanced by homogeneous subsidence. The theoretical basis is introduced in Sec. 7.1. We show that our approach is in line with the Navier-Stokes-equation and introduce a Poisson equation, which derives the velocity field from a given updraught distribution. A solution of this Poisson equation is derived analytically in Sec. 7.2. Section 7.3 sketches how the model is implemented. We continue with a suggestion on how this static model could be expanded into a dynamical model that captures the movement and merging of updraughts (Sec. 7.4). Finally, we briefly evaluate our model and align it to the state of the art (Sec. 7.5). The data referred to in this chapter is always the 480 km × 480 km NoEvap simulation with 1 km resolution. The essence of the model is visualised in Fig. 7.1 already, to give an idea where this chapter leads.

7.1. Theoretical basis

7.1.1. Motivation of the model

The data analysis (Sec. 6.1) showed that precipitation is very localised and limited to few convective regions. To derive a measure that is not influenced by the strong oscillations seen in rainfall intensity, we track the convective updraughts instead of rain cells (Sec. 6.4). We aim to develop a compressed description of the instantaneous horizontal velocity field (always $\vec{v}(x,y)|_{z \approx 50m}$ in this chapter, Fig. 6.9). The approach is to reconstruct the low-level horizontal wind field from a discrete set of localised sinks which correspond to the convective updraughts. Location and strength of these sinks are extracted from the data.

$$\frac{\partial \vec{v}}{\partial t} + \frac{1}{2} \nabla (\vec{v} \cdot \vec{v}) - \vec{v} \times (\nabla \times \vec{v}) = -\frac{1}{\rho} \nabla p \quad (7.1)$$

To do so, we start with the Euler-equation above (Eq. 7.1), which is a formulation of the Navier-Stokes-Equation (NSE). The Euler-formulation describes inviscid fluids and is commonly used to approximate a flow of the atmosphere [Hol73]. It is the momentum equation for a frictionless flow, which combines with the thermal energy equation and the continuity equation to a system of partial differential equations.

7. Model

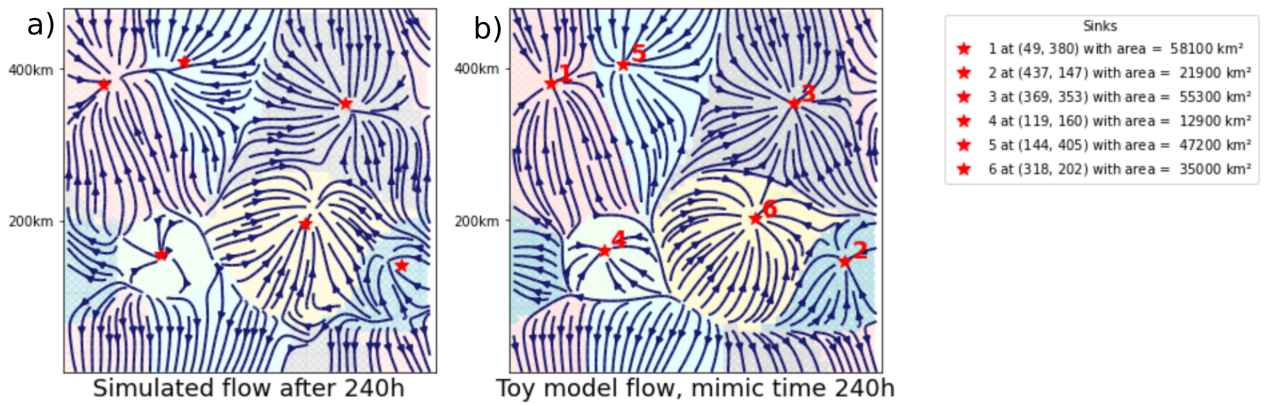


Figure 7.1.: Striking similarity between original simulation data streamlines (a) and our model reconstruction of the flow field (b). The following summarises the essence of our model: First, updraught locations (x_i, y_i) , which are sinks in the horizontal velocity field $\vec{v}(x,y)|_{z \approx 50m}$ (red stars) are extracted from the data (algorithm: 5.5.2). The second information needed is the area $A_{basin_i}^{data}$ of the corresponding basin of attraction ($i = 1 \dots 6$, see legend), which is indicated by the background colouring. Together, these numbers $(x_i, y_i, A_{basin_i}^{data}, i = 1 \dots 6)$ define six discrete objects, from which the $\vec{v}(x,y)|_{z \approx 50m}$ can be reconstructed. First we place a point-like updraught at each sinks location, whose strength is determined by a factor $s \propto A_{basin_i}^{data}$. Then we detect the inflow areas (background colouring) of the sinks in the reconstructed flow and show that these coincide well with the data. Furthermore the absolute value of velocity is captured well too (Sec. 7.6)

7.1.2. Define Poisson-equation which determines the horizontal flow field

Our model will allow to reconstruct the essence of the convective dynamics in the absence of CPs, starting from a discrete distribution of updraught locations and strengths. We show that our approach does not cut across the Navier stokes equation (Eq. 7.3), if we make the initial assumption of a stationary constant density flow. It is reasonable to assume a stationary field and neglect the partial time derivative in Eq. 7.3, since we aim to model the flow at a single time step, corresponding to a 'snapshot' of horizontal velocities. Claiming a constant density flow is in line with the commonly-used Boussinesq-approximation for the well-mixed lower boundary layer: This approximation implies constant density everywhere except for the buoyancy term and therefore implies a constant density flow in the horizontal layer. The density change of a parcel following the flow is given with the total differential:

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\vec{v}),$$

Here, ρ is the fluid density at a given position and time, $\vec{v} = (u,v,w)^T$ is the three-dimensional velocity vector. The two terms on the RHS represent local and advective density changes. The density along a trajectory changes with the divergence of the flow. Since we assume density to be constant along a trajectory and a stationary flow, the divergence of the flow vanishes. The resulting continuity equation links mass density to the velocity field:

$$\nabla \cdot \vec{v} = \frac{-\partial\rho}{\partial t} = 0 \quad (7.2)$$

Navier stokes equation (NSE)

$$\frac{D\vec{v}}{Dt} = \frac{\partial\vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\vec{v} \quad (7.3)$$

The total differential $\frac{d\vec{v}}{dt}$ of the flow vector \vec{v} combines the temporal derivative $\frac{\partial\vec{v}}{\partial t}$ with the advective term $(\vec{v} \cdot \nabla)\vec{v}$. The diffusive term $\nu\nabla^2\vec{v}$ multiplies the dynamic viscosity constant ν with the Laplacian of the flow ($\nabla^2\vec{v} = \Delta\vec{v}$). We assume that density is constant and thus pressure independent, so $\frac{1}{\rho}$ is only a pre-factor to the pressure gradient.

The Reynolds number of the flow (Sec. 4.1.1, Eq. 4.2) can be approximated with a scale analysis to $Re \approx 10^{13}$ (deduction in A.3.1, [YB09]). $Re \gg 1$ implies rather small scale turbulence which might be resolved in the 1 km grid of the data, but is certainly smoothed when we apply the spatial Gaussian filter with a kernel of 10 km (Fig. 6.9). Filtering allows to focus on the bulk movement, rather than local fluctuations or eddies and results in smooth curl free streamlines corresponding to

7. Model

the mean flow (see Fig. 7.1, a). In these streamlines, turbulence is not resolved and viscosity can be neglected, that is, $\nu \stackrel{!}{=} 0$.

Our assumptions allow us to simplify the NSE (Eq. 7.3) by dropping the partial time derivative (stationarity) and setting ν to zero (neglectable turbulence), yielding the simplified equation:

$$(\vec{u} \cdot \nabla) \vec{u} = -\frac{1}{\rho} \nabla p. \quad (7.4)$$

In general, taking the curl on both sides of the NSE yields the vorticity equation for the flow described. In this case both sides become zero:

$$\nabla \times (\vec{u} \nabla \vec{u}) = \nabla \times \left(\frac{-1}{\rho} \nabla p \right) \quad (7.5)$$

$$\rightarrow 0 = 0 \text{ with } \text{rot}(\text{div}) \stackrel{!}{=} 0 \quad (7.6)$$

Consequently the NSE holds no contradiction towards supposing a irrotational velocity field, which is in line with our observation of a purely updraught and subsidence driven flow.

Each rotation free field can be written as the gradient of a scalar field Φ . Therefore we can write the two dimensional velocity field as

$$\nabla \times \vec{v} = 0 \Leftrightarrow \vec{v} = \nabla \Phi \quad (7.7)$$

The boundary conditions of the simulation are double periodic in the lateral dimensions, but the vertical velocity vanishes at the lower boundary:

$$w|_{z=0} \stackrel{!}{=} 0.$$

Thus at the surface layer, horizontal divergence of air parcels has to be balanced locally by vertical subsidence:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = -\frac{\partial w}{\partial z} \approx \frac{w_0 - w_1}{z_1 - z_0} \approx -\frac{w_1}{\delta z}, \quad (7.8)$$

where $\delta z \equiv z_1 - z_0$ is the height of the lowest numerical model level. To be exact, $w = \frac{w_1}{\delta z}$ has to be evaluated at the first level and the horizontal velocity components have to be averaged over the layer thickness δz to be suitably weighted by the z -dependent density.

We assign the divergence of the two dimensional horizontal velocity $v_h \equiv (u, v)$ field to the vertical velocity:

$$\nabla \cdot \vec{v}_h \approx -w$$

Combined with the assumption in Eq. 7.7 this yields

$$\nabla \times \vec{v}_h = \nabla \cdot (\nabla \Phi) = \nabla^2 \Phi = -w. \quad (7.9)$$

7.2. Solving the Poisson equation

This is a Poisson equation (Sec. 7.1.2) that relates a given vertical velocity field w to the corresponding horizontal velocity field.

We now assume N point-like updraughts located at horizontal positions (x_i, y_i) , with $i \in \{1, 2, \dots, N\}$, and assign a 'strength'-factor of S_i to each updraught i .

To represent an updraught distribution that is S_i for all N locations (x_i, y_i) and zero elsewhere, we use the two dimensional Dirac delta distribution with:

$\delta(x - x_i, y - y_i) = 1$ for $x = x_i$ and $y = y_i$ and zero everywhere else. To ensure mass conservation and imply the domain total subsidence cancel, we further subtract one from each term in the summation. The resulting updraught superposition is:

$$w(x, y) \approx \sum_{i=1}^N s_i (\delta(x - x_i, y - y_i) - 1). \quad (7.10)$$

Poisson equation

Mathematically the Poisson equation generalises Laplace's equation¹ and is obtained by adding a constant source term to the RHS. Both equations are sometimes referred to as potential equation, offering a coupling between a scalar potential and a 'sink' or 'source' distribution [Sel00]. The Poisson equation is commonly used in electrostatics to derive an electric potential and further the induced field from a given charge distribution [Nol16]. In gravitational physics, the gravitational potential and further the gravitational force field is derived from a given mass density distribution. In our case, we derive the horizontal velocity field from a sink (updraught) distribution extracted from the simulation data. Solutions of the Laplace equation are always harmonic functions. Thus the solutions of the Poisson equation are harmonic functions, if the 'source' is harmonic too.

7.2. Solving the Poisson equation

This section solves the Poisson equation:

$$\Delta\Phi = w \quad (7.11)$$

for a square domain with doubly-periodic boundary conditions. We approximate the delta function with a single narrow peak centred at a given location (x_0^i, y_0^i) in the form: $W(x, y) = \cos(k(x - x_0^i)/(2L))^2 N \cos(l(y - y_0^i)/(2L))^2 N$. This functional form allows to solve the equation analytically by first expanding $w(x, y)$ as a Fourier series and solving the Poisson equation in Fourier space and transforming the result back to real space.

Since the Poisson equation is straightforward to solve in Fourier space, the mathematical effort lies in finding the Fourier expansion (deduced in A.3.2). The result of

¹An example of a Laplace equation is the continuity equation 7.2

7. Model

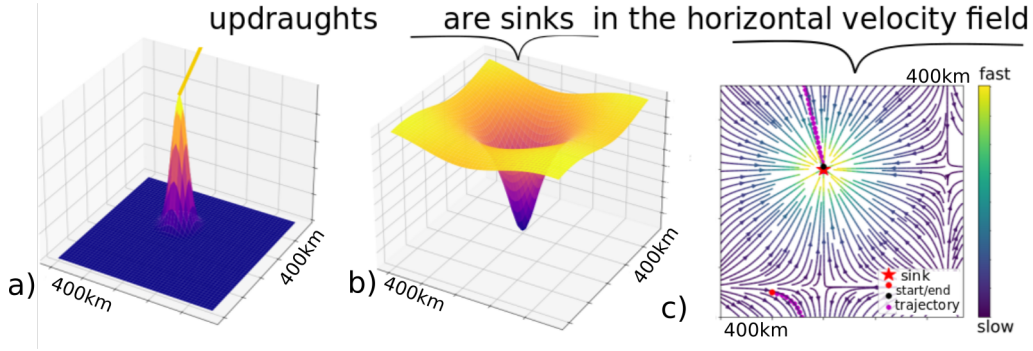


Figure 7.2.: Visualisation of a single modelled updraught. The updraught (a), which is a sink in the pressure potential (b). Furthermore it is an area of convergence in the gradient field, derived from this potential (c). This is illustrated by the path of a random starting point, following the streamlines towards the sink's location.

the Fourier expansion is:

$$W = \sum_{k,l} A \cos\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) + A \cos\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right) \quad (7.12)$$

$$+ A \sin\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) + A \sin\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right), \quad (7.13)$$

$$\text{with } A \equiv \frac{1}{L} \frac{\binom{2n}{n-k} \binom{2n}{n-l}}{\binom{2n}{n}^2} \frac{4}{(1 + \delta_{k,0})(1 + \delta_{l,0})}. \quad (7.14)$$

Multiple sinks at different locations (x_i, y_i) are taken into account by adding their Fourier components, which creates a superimposed potential, where the presence of one sink does influence the depth of another sink (Fig. 7.4).

7.2.1. Deducing the solution of the Poisson equation

To ensure the periodic boundary conditions, we claim the RHS to be L -periodic with L being the length of the quadratic domain. We start with one dimension and choose an L -periodic Ansatz to evaluate how a source at x_0 affects location x :

$$W = \left(1 + \cos\left(2\pi \frac{x - x_0}{L}\right)\right)^N \quad (7.15)$$

$$= \left(2\cos\left(\pi \frac{x - x_0}{L}\right)\right)^{2N} \quad \text{with : } \cos(x - x_0)^2 = \frac{1 + \cos(2(x - x_0))}{2} \quad (7.16)$$

The prefactor of two influences the amplitude only, whereas the width of $\cos(x)^N$ scales inversely with N . The total updraught strength is a constant that decreases with N , but is not relevant since the magnitude of interest will later be the gradient velocity field. In the limit of $N \rightarrow \infty$, W would correspond to a point sink in

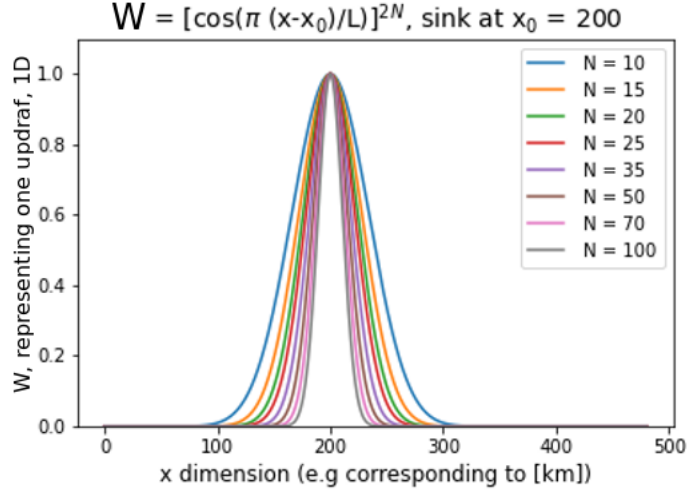


Figure 7.3.: Choice of the function $W(x)$, which represents an updraught in one dimension. The location of this updraught-sink is determined by x_0 and the peak width scales with the exponent N . The limit $N \rightarrow \infty$ corresponds to a perfectly point-like updraught, but is not physically meaningful, because updraughts do have a spatial extent (as visualised in Fig. 6.9, right).

our doubly- L -periodic domain. We take the more realistic finite spatial extent of updraughts into account by empirically choosing N in a way that the derived flow visually reconstructs the data well (discussed further in Sec. 7.3.1). The following outlines how it is reasonable to expand the function W in a Fourier series.

We assume the Poisson equation (Eq. 7.11) can be solved by the general Ansatz: $\Phi_{kl} = \phi_{kl} e^{ikx} e^{ily}$, such that $\Phi = \sum_{k,l} \phi_{kl} e^{ikx} e^{ily}$ and $\Delta\Phi = -(k^2 + l^2) \Phi_{kl}$. Also, any function W can be expanded as a Fourier series, which we do as $W = \sum w_{kl} e^{ikx} e^{ily}$. Since these Fourier components are Eigenvectors of the Laplacian operator $\Delta\Phi$, we can state that the Fourier expansion of $\Phi = \sum \phi_{kl} e^{ikx} e^{ily}$ will be a solution to the equation if and only if:

$-(k^2 + l^2) \phi_{kl} = w_{kl}$ and $\phi_{kl} = \frac{-w_{kl}}{(k^2 + l^2)}$ (and $w_{00} \stackrel{!}{=} 0$) is valid for all modes. If this is true, we can expand the periodic source term W in a Fourier series (solution deduced in App. A.3.2), which later allows to add different source terms in a simple way.

7.2.2. Realisation of the model

To initialise our model we start by finding all stable fixed points, which are sinks of the instantaneous horizontal velocity field, using the algorithm introduced in Sec. 5.5.2. We then record the coordinates (x_i, y_i) and the area of corresponding basins of attraction A_{basin_i} . To reconstruct the original field from this discrete information, we start by finding the Fourier transform (Sec. A.3.2) of the idealised vertical velocity

7. Model

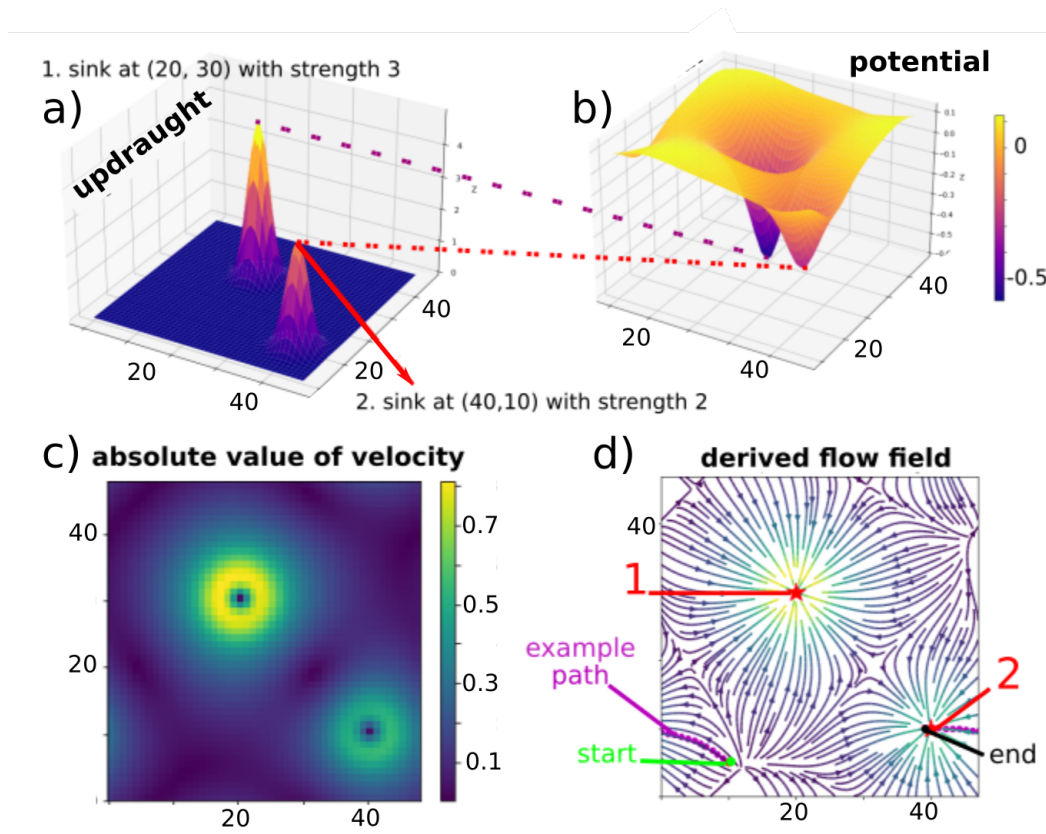


Figure 7.4.: Illustration of the toy model output, ($N = 25$). (a) Updraught-sink distribution (b) Potential resulting from (a). Notice how 'deformations' in the potential minima occur, since the sinks influence each other. (c) The absolute value of the modelled horizontal velocity vector, associated with 'model speed'. This model speed is close to zero in the sinks themselves, but forms a ring with a local maximum in speed around them. This results from the strong slope in the potential, which drives the flow fast. From these rings the speed radially decreases until it reaches zero at the local potential maxima, separating the basins of attraction of both sinks. The stronger the updraught (a), the deeper the sink (b), the higher speed is reached around the sink (c). (d) Stream plot of the derived gradient field corresponding to the modelled velocity vector field and coloured with model speed. The pink dots visualise an example path, following the stationary streamlines from an arbitrary start towards the second sink.

field W , which consists of general subsidence and point-like updraughts at (x_i, y_i) , again with a corresponding strength factor $S_i \propto A_{basin_i}$. We then solve the Poisson equation (Sec. 7.2) in Fourier space and take the gradient of the computed potential field. These steps are the core routine of the model implementation, elucidated further in Sec. A.3.3. Finally, this field is transformed back into real space, so the velocity field can be visualised. The resulting streamlines show good accordance with the original flow (Fig. 6.10).

7.2.3. Proofing the scaling of basins areas with updraught strength

For each updraught we define the basin of attraction (short: basin) as the set of points, from where the flow goes towards the fixed point located at the updraught's centre. In the limit of $N \rightarrow \infty$ it can be shown, that the reconstructed area of a basin ($A_{basin_i}^{model}$) scales with the 'strengths' of the corresponding updraught:

- Step 1: \vec{v}_h is the two dimensional horizontal velocity vector field. The flux of \vec{v}_h through the boundary of the basin of attraction is zero. This results from the definition of the basins and is illustrated in Fig. A.2.
- Step 2: In a conservative vector field \vec{v}_h , the Gaussian theorem holds [Bro+12]: the integral of the flux through the boundary B of a closed surface A equals the integral of the divergence over its area.

$$\oint_B \vec{v}_h \cdot \vec{n} dB = \int_A div(\vec{v}_h) dA \quad (7.17)$$

The line integral is sketched in A.2 and visualises that no flow crosses the border, thus the net-horizontal divergence is zero, too:
 $div(\vec{v}_h) \stackrel{!}{=} 0$.

- Step 3: Since the three dimensional velocity divergence must vanish to satisfy mass conversation and we make use of the lowest level boundary condition, in each basin subsidence has to balance the corresponding updraught:

$$-div(\vec{v}_h) = w_{up} - w_{sub} \stackrel{!}{=} 0$$

- Step 4: Each updraught strength is determined by a factor S_i . The corresponding homogeneous subsidence is subsidence speed c times subsidence area: $A_{sub_i} = A_{basin_i}^{model}$. Thus, every basin area obtained in the model has to be proportional to the corresponding updraught strength S_i :

$$S_i \stackrel{!}{=} c \cdot A_{basin_i}^{model}, \rightarrow S_i' \propto A_{basin_i}^{model}$$

Usage: We initialise the updraught strength parameter S_i' in our model proportional to the basin area A_{basin_i} extracted of the LES data.

7. Model

We are aware that with finite N , the peak width can cause deviations between the relative strength ($S_i/\sum_i S_i$) of an updraught-sink and the corresponding relative basin size ($A_{basin_i}^{model}/\sum_i A_{basin_i}^{model}$). Supposing these deviation could favour larger basins with deeper potential minima, we consider this is not only an artefact, but could yield an interesting dynamic itself: The data analysis showed that larger basins seem to have an advantage and a lower likelihood to decay.

7.3. Implementation of the model

The algorithm implements the analytical solution derived (Eq. 7.15) for given updraught-sink locations. It calculates the velocity vector gradient field \vec{v}_h and the potential Φ from an idealised field W . The algorithm is structured in three functions, where each function's input is based on the output of the primer function. First the joint Fourier components for all updraught-sinks are calculated, second the Fourier coefficients for the variables of interest are derived and finally the solution is transformed back to 'real space'. Details of the implementation are given in A.3.3.

7.3.1. Choosing parameters of the toymodel

When simulating a 480x480 domain with 1 km resolution, a Gaussian filter kernel and basin-detection grid size of ten pixels were used. Accordingly, domain size and grid size were set in the model and c (Sec. 7.2.3) was set to 0.1. We further scaled the peak width with $N = 25$, which derives a flow with close visual resemblance to the original data (see Fig. 7.1).

7.3.2. Performance of the basin reconstruction

To quantify the accuracy of the basin reconstruction, we compare the basin area at reconstructed from a simulation sample at time t ($A(t)_{basin}^{model}$) to the corresponding basin area of the original simulation data $A(t)_{basin}$ (Fig. 7.5). Again, we calculate $A(t)_{basin_i}$ for every basin $i = 1 \dots N_{basin}(t)$ (details Sec. 5.5.2) and initialise the model by placing sinks with strength $S_i \sim A(t)_{basin_i}$ at all extracted updraught locations (x_i, y_i) . Then, we reconstruct the flow field (Sec. 5.6) and determine the actual convergence towards each updraught-sink location (x_i, y_i) , which is the reconstruction basin area $A(t)_{basin}^{model}$.

When an updraught-sink is placed at (x_i, y_i) , but the resulting potential minimum is too 'shallow' for (x_i, y_i) being a stable fix-point, no flow converges to (x_i, y_i) and $A(t)_{basin}^{model} = 0$ (Fig. 7.5). We observe $A(t)_{basin_i}^{model} = 0$ in 11% of the samples (56/520). $A(t)_{basin_i}^{model} = 0$ can occur if a weak updraught with strength S_i is located close to a much stronger updraught (S_j), such that $S_j \gg S_i$ and consequently concerns

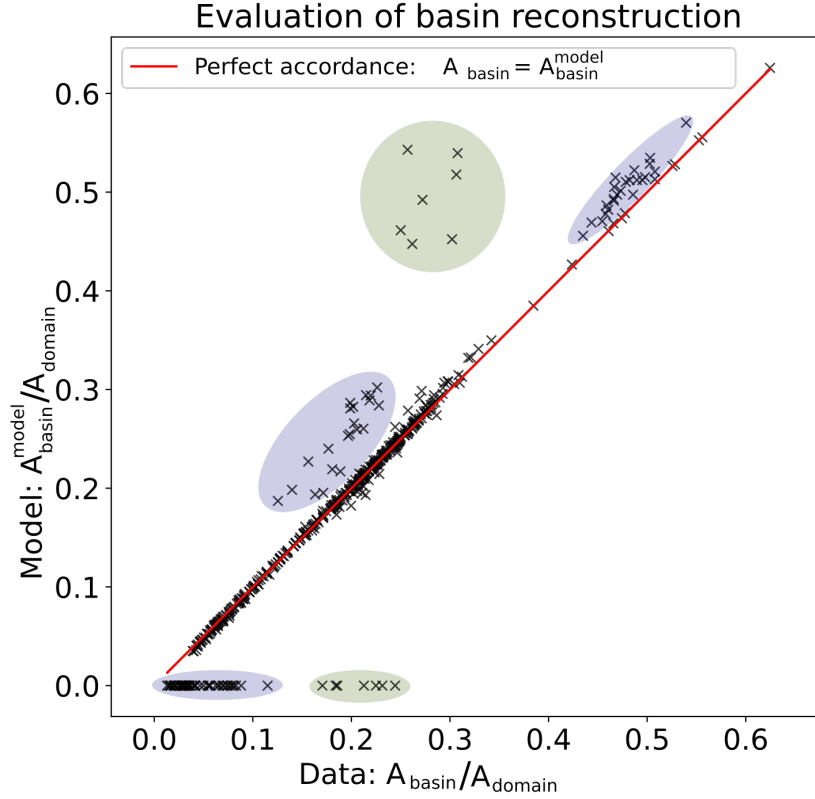


Figure 7.5.: Quantitative evaluation of the reconstructed basin area as domain fraction When plotting $A(t)_{basin}^{model}$ over $A(t)_{basin}$ (for $t \in \{200 \text{ h} \dots 300 \text{ h}\}$) most of the 100 samples are close to line $A(t)_{basin}^{model} = A(t)_{basin}$ (red). Each basin is detected hourly and corresponds to several black crosses. Crosses at $A(t)_{basin} = 0$ indicate that a basin was not reconstructed for a sample in t_i , such that simultaneously at t_i other basins are detected too large (crosses above red line). Purple/green shading highlight samples with a smaller/larger derivation $\delta [A(t)_{basin}^{model}/A_{domain}] < 0.15 [A(t)_{basin}/A_{domain}]$ or $\delta [A(t)_{basin}^{model}/A_{domain}] > 0.15 [A(t)_{basin}/A_{domain}]$. Notice that only the last two remaining clusters can reach the magnitude $0.5A_{domain}$ and that merging can step-wise increase $A(t)_{basin}$, which explains the 'gap' around $0.4A_{domain}$.

7. Model

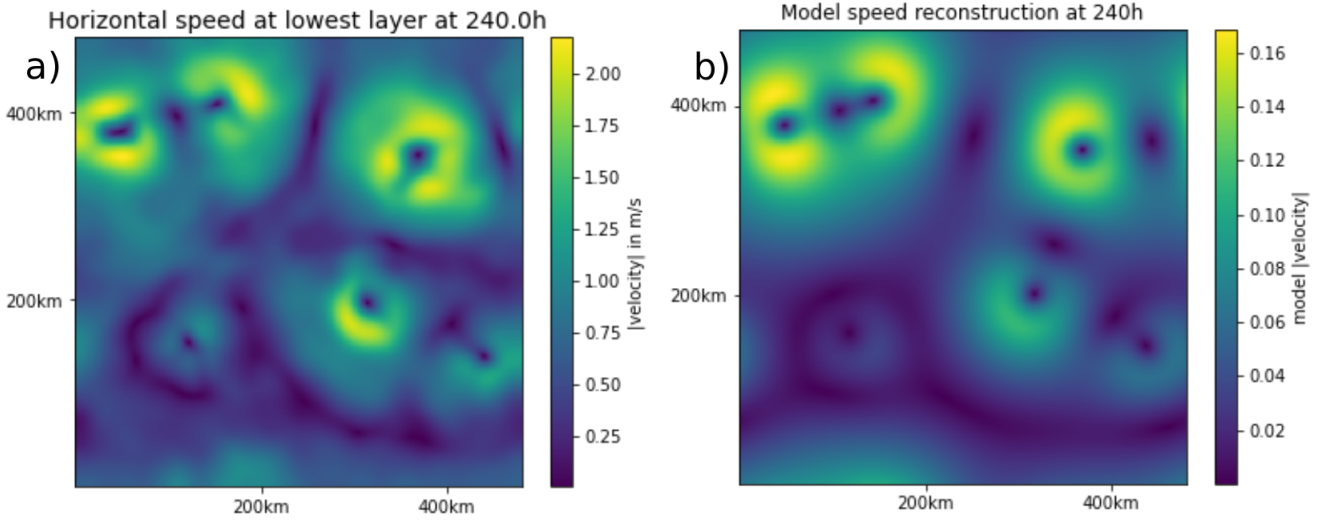


Figure 7.6.: Comparison of horizontal wind speed speed $|v_h|_{z \approx 50m}$ in simulation data (a) and corresponding model reconstruction (b). The instantaneous field $v_h(x,y,t=240\text{ h})$ is smoothed by a Gaussian filter with spatial kernel of 10 km (a). Since the model parameters are not optimised on representing the speed $|v|$ in m/s, the magnitudes are not comparable. But relative speed is reconstructed well: $|v_h|$ increases from basins' borders towards a horseshoe shaped ring of maximum $|v_h|$ around the updraught-sinks.

relatively small basins (usually $A(t)_{basin_i}^{model} < \overline{A(t)_{basin_i}^{model}}$). Since $A(t)_{basin_i}^{model} = 0$ is usually observed a few simulation time-steps before the corresponding basin i decays or merges anyway, we do not think this artefact is too concerning².

7.3.3. Discussing the model

We base our model on the assumption, that turbulence can be neglected to characterise the mean flow (Sec. 7.1.2), which we assume to dominate the long term development of convection. This is implemented for the data by smoothing turbulence with a spatial³ Gaussian filter (10 km) and dropping the turbulence term in Eq. 7.3 for the model. We argue that due to a high Reynolds number (Eq. A.1), turbulence is on a much smaller scale than the 10 km grid we use to estimate the basin areas and therefore does not have to be taken into account. Nevertheless, we are critically aware that the high Reynolds number is not infinite, and it would be meaningful to explore the limitations of the assumption $\nu \stackrel{!}{=} 0$.

²It should be possible to prevent this artefact by parameter optimisation. On the other hand, it could be an interesting 'hint' for the 'fate' (merging/decaying) of basin i when $A(t)_{basin_i}^{model} = 0$.

³Since we want to expand the approach to a dynamical model later, it is important not to apply temporal Gaussian filtering, which would include information of future states in a simulation 'snapshot' we reconstruct and thus distort our prognosis of future developments (Sec. 7.4).

7.3. Implementation of the model

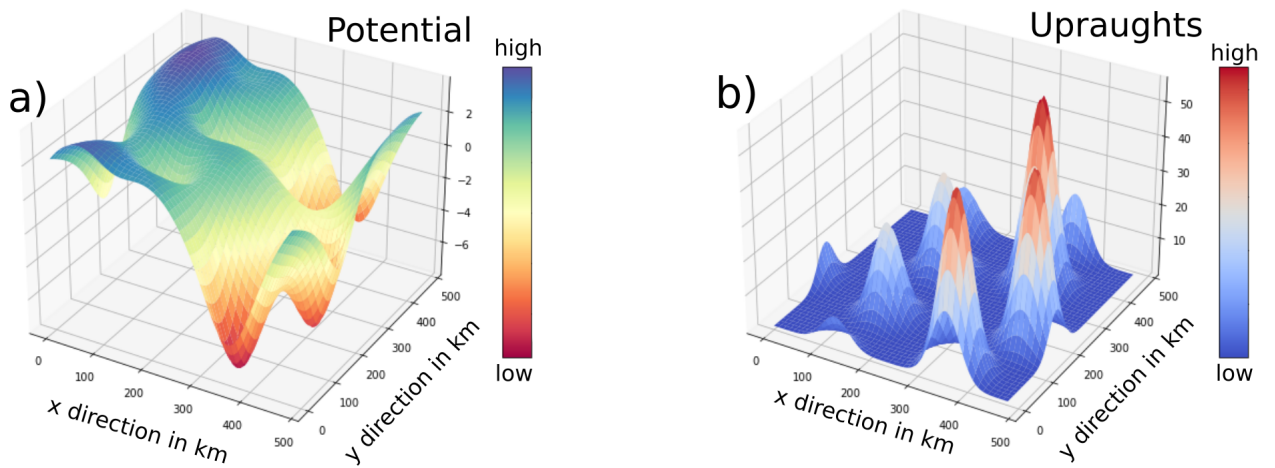


Figure 7.7.: Potential landscape and corresponding updraughts Every updraught peak (b) is influenced by the presence other updraughts. Linked to their finite spatial extend, this affects the corresponding potential (a). By design of the function, Potential minima are wider than the updraught peaks. The potential barrier between updraughts is reduced and the fixed-point in a potential minimum can become unstable when two updraughts are close.

We show, that our model does not only capture the direction of the simulated flow, but also displays the spatial pattern of surface wind speed (Fig. 7.6). Again, this speed is close to zero at the basin borders, areas of horizontal divergence where the direction of the horizontal flow changes. From these border towards the locations of the sinks, (x_i, y_i) , the flow accelerates, forming a near-perfect 'ring' of maximum speed around each updraught. Imagining velocity as height, this would correspond to a crater around the updraught. Stronger updraughts are encircled by higher craters of speed, as expected for an incompressible flow.

Upon closer inspection, the crater is often disrupted by a small side segment of lower wind speed, both in the data and the model. Imagining an aerial perspective, an open crater corresponds to a horseshoe shape with an opening that seems to point towards the nearest neighbour updraught.

To understand this segment of lower wind speed, we consider that local speed is determined by the local potential slope. By design, the width of a potential minima is wider than the width of the corresponding updraught peak, which is determined by the finite spatial extend of the cosine functions (Fig. 7.7, a). Thus, the presence of an updraught significantly influences the potential landscape in the vicinity of other updraughts and the potential depth between updraughts is reduced when the distance is small enough. Consequently, the potential's slope and thereby the speed reduces between close updraughts, causing the opening of the craters.

7. Model

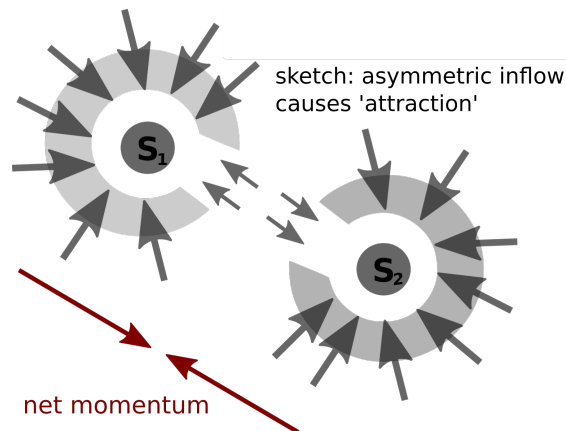


Figure 7.8.: Schematic sketch of two updraught-sinks (S_1 and S_2) with finite expand, encircled by an asymmetric inflow, which results in a net momentum towards each other. Larger arrows on the sides of S_1 and S_2 'turned away' from each other correspond to strong inflow areas with high horizontal wind speed (light grey shading, reminding of a horseshoe shape). In contrast the inflow between S_1 and S_2 is weaker (smaller arrows), corresponding to lower wind speed ('opening' of the horseshoe). Red arrows indicate the suggested net forcing, moving S_1 and S_2 towards each other. This sketch draws upon a hypothesis presented by Patrizio and Randall, Sec. 7.3.3 [PR19].

As an additional effect, the reduced potential barrier between close updraughts can cause the fixed point of a potential minima to become unstable. Usually, this is the case for the weaker updraught with the shallower potential minima. Especially on respect of working towards a dynamical model, we think of this rather as a feature than as an artefact.

If two updraught-sinks are close and the fixed point of the smaller minimum becomes unstable, the flow goes towards the deeper minimum resulting from the larger neighbouring updraught. Consequently, the two basins become one, which could be interpreted as a merge. Previous to a merge, the potential barrier between the two minima would decrease, leading to pronounced horseshoe shapes with openings towards each other. Concluding, the opening could forecast an approaching merge.

The following links the 'horseshoe shapes' we observe in surface wind speed to a study by Patrizio and Randall. Despite their much larger domains up to 6.144 km with 3 km resolution, some of the mechanism they discuss should be comparable to our NoEvap simulations, which miss typical small scale dynamics through the absence of CPs. Patrizio and Randall observed that the circulation seems to organise in a way, that produces weaker surface wind between neighbouring convective

cells⁴, while the wind stays strong on the opposites of the clusters. This asymmetry in surface wind is referred to as horseshoe shape. Such spatial patterns of surface winds were discussed in the context of cloud merging before [Tur82] [SK94]. Both studies suggest the parts of updraughts closest to neighbouring clouds are weakened, for example by interfering downdraughts causing evaporation and mixing. Weaker updraughts lead to weaker surface inflow along nearest edges of neighbouring convective cells, which we observe as reduced wind speed in our NoEvap data.

Patrizio and Randall suppose the relatively strong inflow at the outside of neighbouring convective cells (big arrows, Fig. 7.8) and the reduced inflow (small arrows, 'opening' Fig. 7.8) between them, could cause neighbouring cells to be 'advected' toward each other. Therefore, they link the horseshoe shaped wind speed pattern to the pairwise merging of neighbouring convective cells they observe.

Supposing that merging dominates the later stage in the approach of the final aggregated state, we aim to capture this dynamic in our reconstruction. In the following, a first approach towards a dynamical model is introduced which proposes a movement of the updraughts towards each other based on our phenomenological observations.

7.4. Towards a dynamical model

In order to turn the compressed reconstruction⁵ of the flow field into a dynamical model, we aim to introduce an update mechanism for the updraught-sink locations (x_i, y_i) . The idea is to propose a forcing on the updraught sinks that has the tendency to move them towards each other, so their corresponding basins fuse once the updraughts are close enough (Sec. 7.3.3). We assume, where and when a merging between two sinks occurs does not only depend on their locations $((x_1, y_1), (x_2, y_2))$ and updraught strengths factor (S_1, S_2) , but on the entire potential landscape.

To take the joint potential into account, we suggest a gravity-like update mechanism, that again draws on methods commonly used in electrodynamics (see 7.1.2). One by one, we remove the updraughts and calculate the potential that would originate from all remaining updraughts. We then take the potential of the location where the updraught under consideration would have been and derive the gradient for this one point only. This gradient is a hypothetical velocity vector resulting from the hypothetical field of all other updraughts. We multiply this vector with a hypothetical time factor that we call 'step-length' to obtain a displacement prognosis vector. In further work, this 'step-length' parameter can be optimised to correspond to typical displacement distances (Sec. 6.4.4) of updraughts in a given simulation period (for example one day, Sec. 7.4.1).

⁴which they refer to as 'clusters'

⁵Which does have a value on its own

7. Model

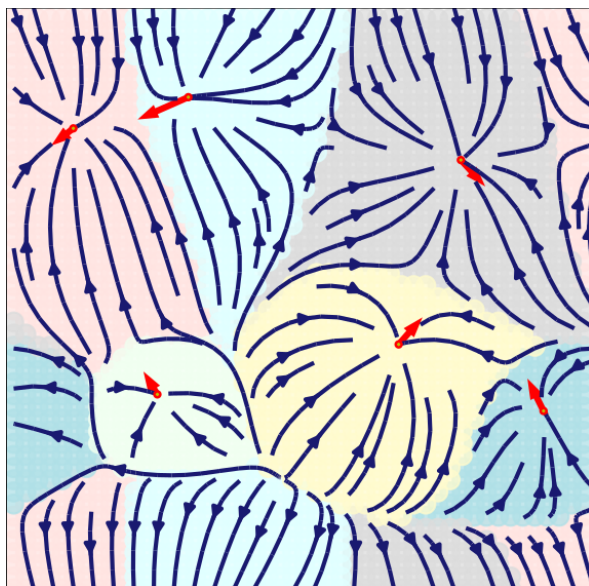


Figure 7.9.: Prognosis for updraughts' movements based on a gravity approach The streamlines of the reconstructed flow ($t = 240$ h) are plotted above the corresponding basins of attraction. The red arrows indicated the displacement prognosis based on the gravity approach with a 'step-length' factor of 50.

We work towards a dynamical model by taking the idealised reconstruction of a simulation snapshot and displace the updraughts based on the gravity approach, which offers an update dynamic for the location. To update the strength factors S_i we calculate the basin areas, that result from the model configuration of the previous 'step' $(x_i^{t-1}, y_i^{t-1}, s_i^{t-1})$. This implementation would create an autonomous dynamical model. Additionally, the development of this model could be compared to the actual development from the point of origin in the original simulation data (Sec. 7.4.1). Parameters could be optimised by taking multiple data sets into account, to prevent over-fitting to one simulation example.

7.4.1. Performance of the update dynamic

We developed a framework (Fig. 7.10) that extracts discrete information of the data, derives an idealised reconstruction of the velocity field and calculates a prognosis for the movement of each updraught, based on the gravity approach (Sec. 7.4). To evaluate the performance of our prognoses they are then compared to the actual updraught displacements in the data.

To obtain a quantitative comparison, we calculate the angle between our prognosis vector and the actual displacement vector derived from the simulation data, always referring to the smaller angle between zero and π . We observe that calculating the actual displacement after a longer simulation period (as one day) offers a clearer

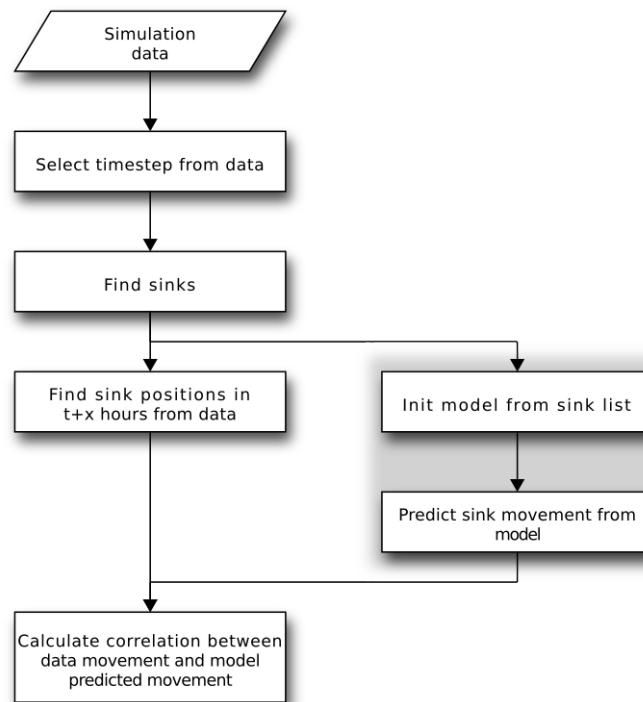


Figure 7.10.: Extended version of model implementation flow-chart (Fig. 5.6), illustrating how we aim to predict a displacement vector (Sec. 7.4) for each updraught-sink and compare this prognosis to the actual displacement in the data. To incorporate an update dynamic to our model, we could calculate such a prognosis vector, then place the updraught-sink to the resulting new location and calculate the resulting basin area, using the former basin areas as 'strength' factor input.

7. Model

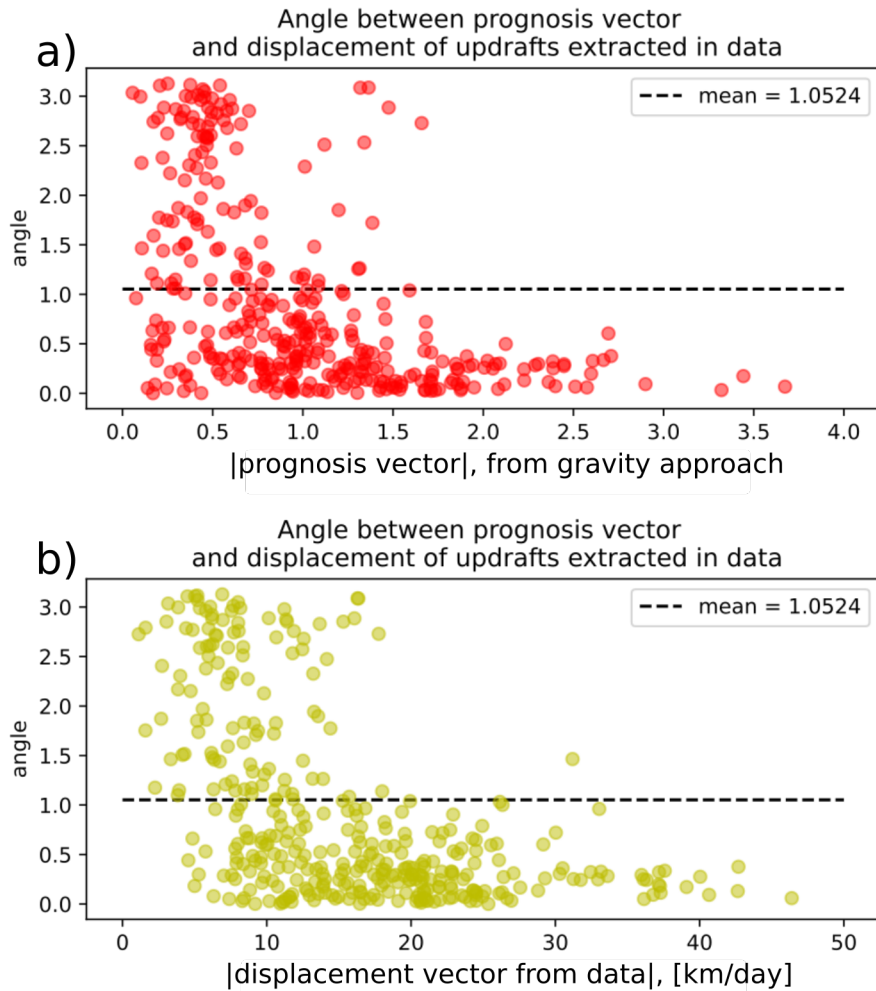


Figure 7.11.: Performance of updraught displacement prognosis The instantaneous flow fields of the simulation period 200 h-300 h is reconstructed and an arrow indicates the predicted direction of movement based on the gravity approach. We calculate the angle between our prognosis vector and the actual displacement in the data 24 h later. Always referring to the smaller angle, we obtain values between zero and π . The performance is significantly better than random with a mean angle below $\pi/2$. We further notice, that the quality of the prognosis increases with the absolute value of both the prognosis vector (a) and the actual displacement vector (b).

7.5. Contribution of our model to the state of the art

picture than calculating displacements only one hour ahead. Assuming that noise, fluctuations and maybe oscillations in the data could influence the short term displacement of updraughts in the simulation, we calculate the displacement after a time interval of 24 h.

Then we statistically evaluate the angle between displacement prognosis vector and actual displacement vector for 100 simulation snapshots (100 h-200 h) and find positive correlations between both directions. Plotting the angle against the absolute value of either the length of the actual (data) displacement vector or the length of the prognosis-vector shows, that the performance of the dynamic increase with both vector lengths (Fig. 7.11). In case of the data displacement vector, this could indicate small displacements to be dominated by noise, whereas larger displacements in the result from an actual drift. In case of the prognosis vector length (based on the gravity approach), a short prognosis vector could result from ambiguous forcing and thus lead to poorer performance than a longer prognosis vector.

Summarised, we developed and implemented the basis for a dynamical model that will capture the merging dynamic observed in the NoEvap simulations.

7.5. Contribution of our model to the state of the art

Until today, various models with different strengths and weaknesses can only capture aspects of CSA (Sec. 4.3.2, 4.3.3). Furthermore, these model approaches are usually field-like and divide the domain into coarse columns. In contrast, we aim to model CSA on the level of individual convective cells, describing each cell as discrete convective object.

Similar to previous coarsening models [CM13], [WH19], our approach captures a progressive increase in spatial length scale, corresponding to increasing basin area and a reduced number of updraught sinks (Sec. 6.5).

To extract the temporal scaling factor of the coarsening process from data, Craig and Mack suggest a manual and qualitative method (Sec. 4.3.3, Sec. 6.5). Our method automatically detects the number of convective cells using our two tracking algorithms. This becomes relevant when aiming towards a statistical evaluation of larger data sets (Sec. 8.5), which could then allow to derive simple rules for the 'competing' mechanisms between basins. Craig and Mack define the final aggregated state as the stage when only one circular moist area remains. In our approach, this corresponds to the stage with a singular convective region that owns the entire domain as its inflow basin. Concluding, their coarsening model and our model both suggest a 'winner takes it all' dynamic. But to be exact, our 'winner' has a history of merges and is a fusion of several former separated convective events. Therefore, we introduced an approach towards a dynamical model (Sec. 7.4) that captures this merging dynamic on the level of individual convective cells.

8. Discussion

In this chapter we discuss possible physical feedbacks ¹ and the possible origin of the oscillations observed (Sec. 8.1). If not denoted otherwise, this chapter always refers to the 480 km × 480 km NoEvap simulation, with 1 km resolution. We introduce the idea of convective oscillators (Sec. 8.2) as a possible framework to explore these oscillations further. The conclusion (Sec. 8.4) summarises the thesis and motivates an outlook on future work (Sec. 8.5), which could be based on this thesis.

8.1. Discussion of possible physical mechanism

In general, spreading CPs introduce a spatial scale $\mathcal{O}(10 \text{ km})$, at which the larger-scale flow is disrupted: CPs suppress ongoing convection by their negative buoyancy effect (Sec. 4.1.7). Numerical models of CPs should therefore resolve scales much finer than 10 km. One can make the reasonable assumption that if CPs are not resolved due to coarser numerical resolution, as done in "classical" simulations of convective self-aggregation [BPB04], the dynamics could have similarity to our NoEvap simulations, which have a relatively fine resolution but suppress CPs by design. Under the removal of CPs (NoEvap) the larger-scale flow is not disrupted and the requirement on resolution may therefore be less strict. In conclusion, we expect that physical mechanisms found to drive CSA at coarser resolutions ([WE14], [EWW14], [Win+17]) carry over to the current simulation setup, especially radiative feedbacks.

8.1.1. Transient stage

The temporal approach to the final steady state could be seen as a transient stage towards the aggregated equilibrium. In this transient stage (0 h-350 h), domain-mean precipitation $\overline{I(t)}$, lowest-layer wind speed $\overline{|v(t)|_s} = \overline{|v(t)|_{z \approx 50m}}$, surface heat fluxes $\overline{\text{lhf}(t)}$ and near-surface relative humidity $\overline{q(t)|_s} = \overline{q(t)|_{z \approx 50m}}$ oscillate with varying periods, which are all less than 10 h (Fig. 8.1). Concurrently, we observe oscillations with more regular and longer periods between 17 h-30 h in the outgoing longwave radiation $\overline{\text{OLR}(t)}_{TOA}$. Despite these longer periods, we assume the oscillations in $\overline{\text{OLR}(t)}_{TOA}$ can be associated with the oscillations in $\overline{I(t)}$.

Apart from their respective oscillations, each variable undergoes slow longer-term changes.

¹Ideas presented will be followed up with further research in the future

8.1. Discussion of possible physical mechanism

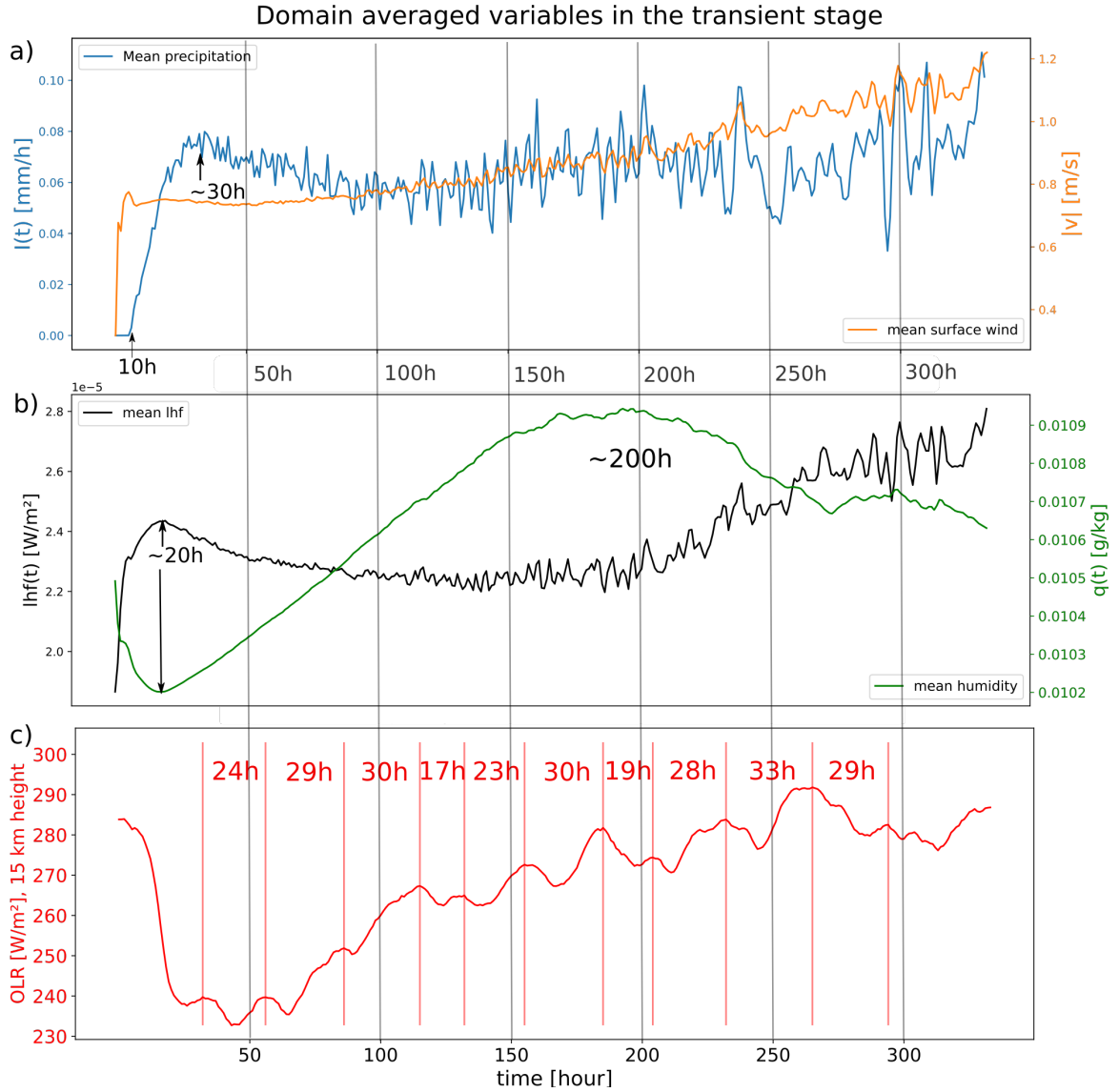


Figure 8.1.: Development of domain-mean variables

(a) Surface precipitation $\bar{I}(t)$ and lowest-layer horizontal wind speed $|\bar{v}(t)|_s = |\bar{v}(t)|_{z \approx 50m}$. (b) Surface latent heat flux $\bar{l}hf(t)$ and lowest-layer relative humidity $\bar{q}(t)_s = \bar{q}(t)|_{z \approx 50m}$. (c) Top-of-atmosphere $\bar{OLR}(t)_{TOA} = \bar{OLR}(t)|_{z \approx 17500m}$ drops with the onset of simulation, but increases again after ~ 70 h exceeding its initial value. Vertical dashed red lines indicate temporal $\bar{OLR}(t)_{TOA}$ maxima and times indicate periods between them.

8. Discussion

With the onset of the simulation, $|\overline{v(t)}|_s$ steeply increases, which we associate with the low-level convergence induced by the onset of convection. This steep increase in $|\overline{v(t)}|_s$ is consistent with a steep increase in $\overline{\text{lhf}(t)}$, which seems coupled to the steep decrease in $\overline{q(t)}_s$.

We hypothesise that this early convection could be Rayleigh-Bénard like convection in the boundary layer, which builds non-precipitating clouds roughly between 2100 m and 5500 m (Fig. A.3). These clouds moisten the atmosphere and allow very deep clouds to form (up to 8800 m height) about 10 h later. Contemporaneous, $\overline{I(t)}$ starts to rise with time lag of ~ 10 h towards $|\overline{v(t)}|_s$. Clearly, $\overline{I(t)}$ is not determined by $|\overline{v(t)}|_s$ only, but shows strong correspondence to $\overline{\text{lhf}(t)}$, which provides additional buoyant potential energy for convective activity (bulk formula, Eq. 4.9). $\overline{I(t)}$ peaks and then drops around 30 h, which is about 10 h after the peak in $\overline{\text{lhf}(t)}$. Notice that the peak in $\overline{\text{lhf}(t)}$ approximately coincides with the minimum in relative humidity $\overline{q(t)}_s$: As the saturation of the atmosphere increases with $\overline{q(t)}_s$ less moisture is absorbed from the sea surface, which decreases $\overline{\text{lhf}(t)}$ (with $\text{lhf}(t) \sim |v|(t) \cdot [q_s^{\text{sat}} - q(t)_s]$, Eq. 4.9, Fig. 8.1). After reaching its maximum (~ 200 h), $\overline{q(t)}_s$ decreases, which enables $\overline{\text{lhf}(t)}$ to enhance again.

Decreasing humidity is typical during the development of CSA: With aggregation, the majority of the domain is cloud free and the $\overline{\text{OLR}(t)}_{\text{TOA}}$ ² increases (Fig. 8.1, Sec.4.3.2, [WE14]). Initial high values in $\overline{\text{OLR}(t)}_{\text{TOA}}$ are associated with little cloud-coverage in the spin-up stage, thus $\overline{\text{OLR}(t)}_{\text{TOA}}$ first decreases as clouds develop until it increases again after about 70 h. Understanding the development of OLR (glossary) might require to consider additional variables such as the liquid water path, ice water path or water vapour path.

Furthermore, OLR is strongly influenced by the vertical cloud and water vapour distribution: According to the Stefan-Boltzmann law the thermal radiative flux is proportional to $\sim T^4$ (e.g [Lam07]), thus in the warmer lower atmosphere, clouds and water vapour cause stronger OLR than in the colder upper atmosphere (Sec. 4.2.3, [MH12], [EWV14]). Muller and Held suggested the height of clouds to be crucial for the development of CSA³.

8.1.2. Cloud effects

Boye Nissen and Haerter (2019) suggest that in the absence of CPs the main drivers of the NoEvap aggregation are deep convective clouds, which cause a net warming in the troposphere and thereby promote further convective activity in the surroundings, but we suppose that the detailed feedback mechanisms are more complex. Especially, we think understanding the horizontal and vertical distribution of clouds is

²to be exact, $\overline{\text{OLR}(t)}_z$ increases for all heights $z > 11\,000$ m

³They study simulations with CPs but at coarser resolutions of 2 km, which might not resolve all CP properties [MH12]

8.1. Discussion of possible physical mechanism

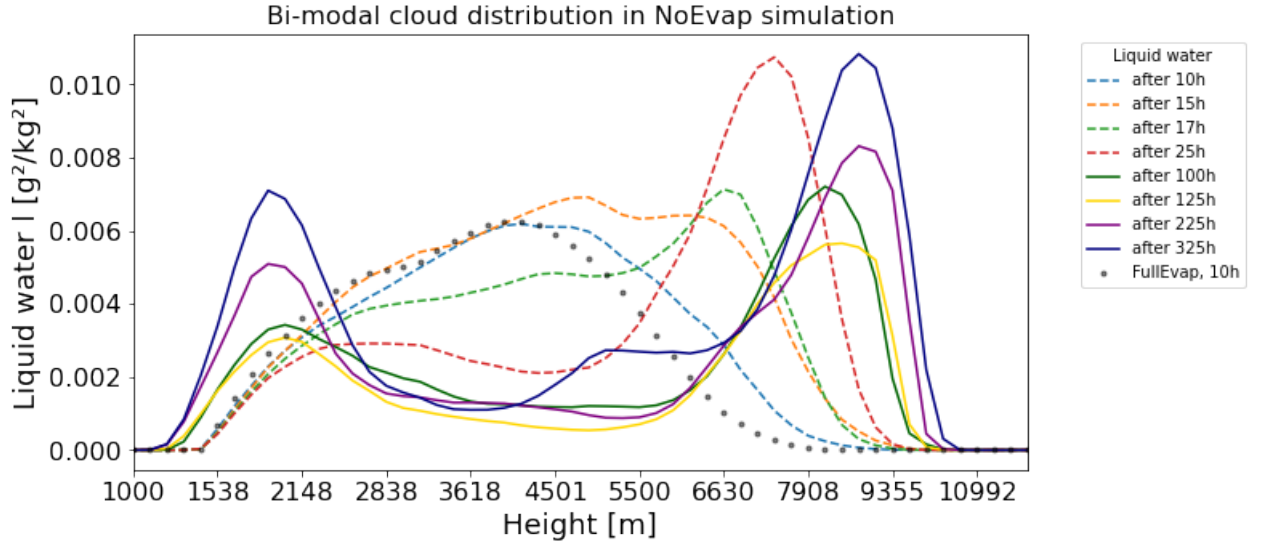


Figure 8.2.: Profiles of integrated liquid water of the NoEvap simulation

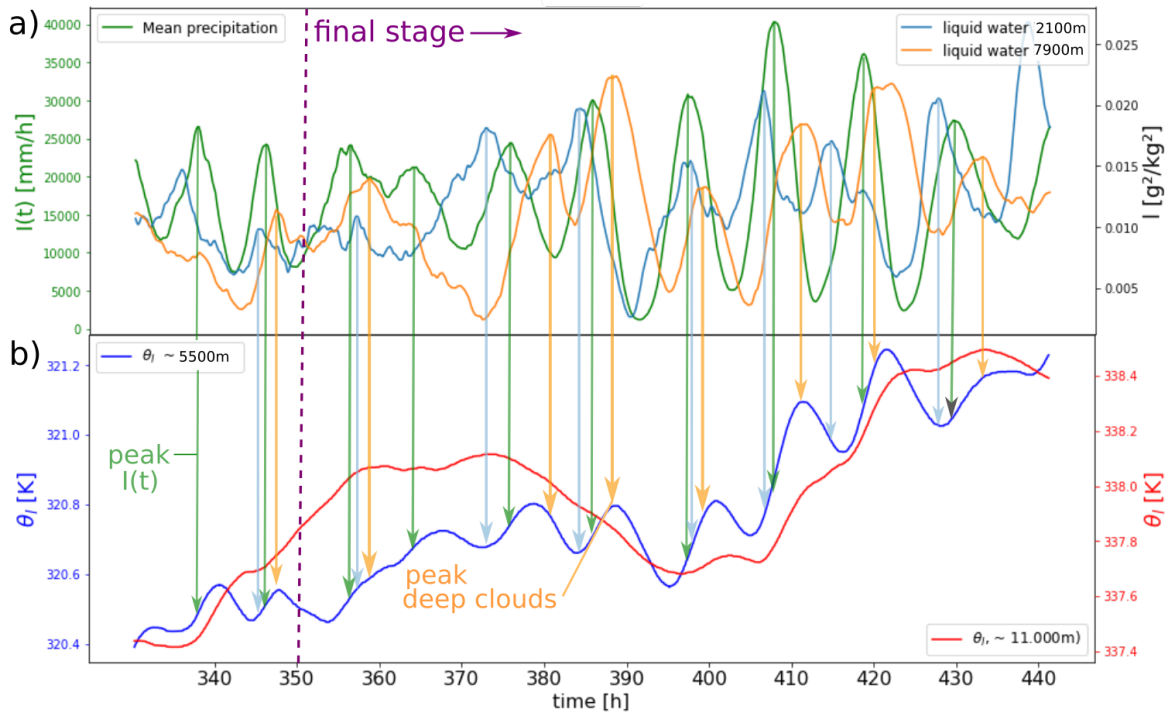


Figure 8.3.: Oscillations of deep cloud fraction $\overline{I(t)}|_{z \approx 8000m}$ (a, orange), shallow cloud fraction $\overline{I(t)}|_{z \approx 2100m}$ (a, blue), precipitation $I(t)$ (a, blue), and liquid water potential temperature $\overline{\Theta_l(t)}$ at heights $z \approx 5500m$ (b, blue) and $z \approx 11000m$

8. Discussion

crucial to comprehend the NoEvap aggregation process and the observed oscillations.

With the onset of convection ($t \approx 10$ h) clouds build at intermediate heights, in a symmetric distribution around $z = 4000$ m, which is similar to the cloud distribution in FullEvap (Fig. 8.2). In FullEvap, this distribution does not qualitatively change, but in NoEvap very high clouds up to 7000 m height build between 10 h and 25 h. The formation of such deep convection could be associated to the absence of CPs, which do not disrupt the convective circulation in NoEvap. For $25 \text{ h} < t < 125 \text{ h}$ this deep cloud peak decreases and splits in a deep cloud fraction above 8000 m and a low cloud fraction around 2100 m, leading to a noticeable bi-modal cloud distribution. It will be crucial to investigate what causes the intermediate clouds to disappear (Fig. A.3). One explanation could be increasing liquid water potential temperature $\overline{\Theta}_l(t)$ (Sec. 4.1.1) in the middle atmosphere: This warming could reduce the cloud water, for example $\overline{\Theta}_l(t)_{z \approx 5500 \text{ m}}$ increases from 313 K to 320 K over the first 330 h of the simulation.

Both cloud fractions appear in similar horizontal locations, supposing mechanisms of deep and shallow convection to be present in the same convective regions, beyond which little cloud formation occurs.

We further observe interesting anti-phase oscillations between high and low cloud fractions, as the liquid water content $\overline{l}(t)$ of individual layers changes. These anti-phase oscillations become very regular in the final stage, which seem highly correlated to the oscillations in precipitation $\overline{I}(t)$ (Fig. 8.3).

Within intervals of enhanced deep cloud fraction, $\overline{\Theta}_l(t)$ increases in the middle atmosphere (eg. $\overline{\Theta}_l(t)$ at 5500 m height in Fig. 8.3). Further, we observe $\overline{\text{OLR}}(t)_z$ oscillates quite in phase with the temperature of the respective height z , for example $\overline{\text{OLR}}(t)_{z \approx 5500 \text{ m}}$ is quite synchronised to $\overline{\Theta}_l(t)_{z \approx 5500 \text{ m}}$ (compare Fig. A.4). Temperature above the deep cloud height seems to vary with the longer periods of $\overline{\text{OLR}}(t)_{\text{TOA}}$ (compare Fig. A.4 to Fig. 8.3).

Concluding, we hypothesise cloud and radiative feedbacks cause intervals of deep, strongly precipitating convection, which warms the middle atmosphere, followed by intervals of shallow, less precipitating convection, where the atmosphere cools again. Further analysis is needed to understand how exactly these alternating intervals of high and low clouds provide a net positive or negative feedback on convection, respectively, but some known cloud properties and their relevance for our simulations are outlined in the following.

In general, clouds result from convective activity, which increases the temperature of the atmosphere by vertical heat transport and latent heat release aloft. Clouds can also have a cooling effect by reflecting incoming short wave radiation (SW), but this cooling will play a minor role because very little SW is absorbed by the atmosphere and we prescribe the SST in our simulations. Equally, the effect of surface warm-

8.1. Discussion of possible physical mechanism

ing caused by low clouds, which absorb heat and re-radiate it back down, can be neglected. Taken together, we suppose the crucial effect is the cloud-OLR interaction, which causes a net warming or cooling depending on cloud height⁴, condensate mixing ratio and temperature (sec. 4.3.2, [EWV14]). Thus, intervals of alternating cloud height could provide alternating forcing on convection and consequently cause oscillations in $\overline{I(t)}$.

8.1.3. Synchronised oscillations in the final aggregated stage

In the final stage, clear oscillations emerge in $\overline{I(t)} \approx I(\text{cluster}_{final,t})$ (Fig. A.4, Sec. 6.2) with periods of 8 h-11.8 h. Surface wind speed $|\overline{v(t)}|_s$ oscillates with very similar periods, but $\overline{I(t)}$ lags behind $|\overline{v(t)}|_s$ with a quite consistent delay of around 1 h (A.4, a, black arrows). We claim it is reasonable, that convergence $|\overline{v(t)}|_s$ precedes precipitation $\overline{I(t)}$ and not vice versa. Also $\overline{\text{lhf}(t)}$ shows strong similarity to $|\overline{v(t)}|_s$ and seems to follow the oscillations in $\overline{q(t)}_s$ (Fig. A.4). Remarkably, $\overline{q(t)}_s$ decreases just after the last merge at 350 h until it oscillates below 0.0105 g kg^{-1} , indicating an enhanced drying in the final aggregated stage. Overall, we assume the final stage to be an equilibrium state with a (quasi)-periodic solution and very regular circulation patterns.

8.1.4. Concluding hypotheses on physical feedbacks possibly contributing to CSA and observed oscillations

As discussed, precipitation increases in the final aggregated stage, which previous studies attribute to the increasing OLR and temperature (Sec. 8.1.1, 8.1.3). For example, Betts and Ridgway relate the observation that domain-mean precipitation increases with warming to radiative constraints (Sec. 4.1.5, other studies eg. [Tak09], [JR18]). We observe a warming of the atmosphere between 1000 m and 12 500 m height (Sec. 8.1.2), which is further indicated by an increasing $\overline{\text{OLR}(t)}$ above 11 000 m height. Warming of the middle atmosphere decreases the temperature gradient towards the boundary layer below. This decrease is in line with the observed decrease of CAPE over the simulation (Fig. 6.7). The decreasing CAPE is consistent with Xu and Emanuel (1989), who found that the tropical atmosphere organises along a moist adiabatic lapse rate (Sec. 4.1.5). In such a state, small deviations $\delta\text{CAPE} > 0$ could determine at which horizontal locations convection is possible.

We suspect dynamical forces to dominate convective processes at the late stages. This hypothesis is consistent with the strong correspondence between precipitation

⁴Usually higher clouds have a strong greenhouse effect which amounts to a net warming of the atmosphere.

8. Discussion

and the low-level wind speed (Fig. 8.1, Fig. A.4)⁵.

We ask whether there is some correspondence between feedback mechanisms that reduce CAPE over a long period, and feedback that causes shorter intervals of reduced convective activity, associated with the oscillations in $\overline{I(t)}$. Underlying these oscillations, we suspect a complex interaction chain of multiple physical variables. Therefore we focus on the final state with very regular oscillations and hypothesise the following:

The final stage has a bi-modal distribution of clouds (Sec. 8.1.2, Fig. 8.2), with alternating intervals of increasing deep clouds or low clouds, where we expect the deep clouds to cause enhanced precipitation.

Again, due to low CAPE, the energy needed for intervals of deep convection has to be provided dynamically, by an interval of strong circulation. A strong circulation is associated with strong low-level convergence. Therefore intervals of high $\overline{|v(t)|}_s$ and $\overline{\text{lhf}(t)}$ could enable deep clouds to form, which then precipitate about 1 h later (and subsequently dissipate). The deep clouds could warm the middle atmosphere due to a greenhouse effect. We then suppose increasing temperature and OLR⁶ will reduce the temperature gradient, which inhibits convective activity⁷. When deep clouds are removed by precipitation, the greenhouse warming of the atmosphere decreases. Once the atmosphere is colder, (deep) convection can build again and the circle repeats. Further work will examine these hypotheses carefully, and especially consider the two layer model suggested by Emanuel, Wing, and Vincent [EWW14] (Sec. 4.3.2).

Overall, the hypothesised feedback loop reminds us of typical predator-prey like dynamics, described by Lotka-Volterra equations [Lot25], [Vol36]), which show very regular oscillations in the steady state.

8.2. Interpretations as convective oscillators

Oscillations can be seen as, typically temporal, repetitive variations of a quantity [Str04], which can for example occur when a system is deflected from its equilibrium state. Then the system is pulled back towards the equilibrium, but exceeds it and starts to oscillate with a characteristic frequency, as for example the Brunt-Väisälä frequency for an air parcel that is vertically displaced from its level of neutral buoyancy (LNB) in stable atmospheric conditions (Sec. 4.1.1).

Such real physical oscillations are always 'damped', since frictional processes dissipate buoyant energy. Therefore oscillations decay over time unless there is a per-

⁵The importance of the dynamical convergence and the circulation of individual convective cells was discussed in Sec. 6.4.3

⁶Again, in the final stage, $\overline{\Theta_l(t)}$ and $\overline{\text{OLR}(t)}$ oscillate nearly synchronously to precipitation. (Fig. 8.3).

⁷It would be meaningful to investigate which temperature and OLR levels are most relevant

8.2. Interpretations as convective oscillators

sistent energy supply [LL67], which is, in our simulations, ensured by the constant heating by the sea surface.

In general, studies showed that systems tend to oscillate when interactions between components require finite time, hence providing a time-delayed feedback [CY99]. In our system, such time-delayed feedbacks could be at the core of the spatio-temporal oscillations of convective cells, because feedback processes in the atmosphere do take finite time.

To expand on an example mentioned earlier: when a sufficient temperature gradient is established in the atmosphere, convection starts to transport heat and humidity upwards with typical accelerations around 1 m s^{-1} (at 1200 m height). After a characteristic time ($\sim 1 \text{ h}$), clouds develop, which then influence the incoming short-wave radiation and the outgoing longwave radiation. Depending on their specific properties⁸, clouds then either provide a net negative or net positive feedback on convection (Sec. 8.1). Gravity waves laterally dissipate heat on much shorter time scales (Sec. 4.1.1), and thereby could provide a coupling between different convective cells. Droplet coagulation takes time too, leading to an additional time delay until precipitation starts. Since we suppress the re-evaporation of precipitation⁹, $I(x,y,t)$ will not have a major impact on the temperature, but condensed droplets will also feed back on radiation effects before they fall towards the ground.

As shown before (Sec. 6.2), we do not only observe oscillations in $\overline{I(t)}$, but also spatio-temporal oscillations in $I(\text{cluster}_m, t)$ for each rain cluster m (Fig. 8.4). Combined with the suggested interaction, these individual oscillations associate the picture of coupled oscillators.

To elucidate this picture, we hypothesise an interaction as follows: Each convective updraught u_i has a convective potential p_i , which is fed by near-surface moisture convergence potential. Cloud formation increases the stability of the atmosphere (Sec. 8.1, Sec. 6.4.3). Heat is transferred between convective cells via gravity waves¹⁰, which would provide an inhibitory coupling and reduces all potentials p_i , in the simplest case equally in line with the weak temperature gradient approximation (Sec. 4.1.6). If an updraught u_i is no longer strong enough to compensate this convective inhibition, u_i decays, which could be compared to an over-damped oscillator. To complete this picture, we will investigate further what rules the moisture convergence in each u_i 's inflow basin b_i and how this is related to the basin area.

Multiple coupled oscillators are generally complex to describe analytically (details in A.4.1). Therefore, one approach could be to start with the final steady state, for

⁸Which themselves are determined by the cooling and heating processes

⁹And thereby the evaporate cooling effect

¹⁰enabling interaction without mass exchange

8. Discussion

which we observe synchrony in oscillations of different physical variables¹¹ as $I(t)$, $|v(t)|$, $lhf(t)$, $q(t)$ and Θ_l (Sec. 8.1.3, Fig. A.4, Fig. 8.3).

We suggest to use a type of predator-prey system (eg. [Wan78]) to capture the observed oscillations of the last remaining convective cell in a phenomenological model of a 'convective oscillator'.

Such a convective oscillator could have a 'predator-like' variable $x(t)$ mimicking the precipitation, which increases when sufficient 'convective activity' builds up. The second 'prey-like' variable $y(t)$ could for example symbolise an abstract 'convective activity'. Alternatively, $y(t)$ could be associated with physical properties, as surface fluxes and low layer convergence or a sufficient cloud fraction. Some parameters could be extracted directly from the simulation data: for example $x(t)$ could lag behind $y(t)$ with the observed 1 h delay of $\overline{I(t)}$ towards $|\overline{v(t)}|_s$. In the resulting feedback circle, increasing convective activity $y(t)$ enhances precipitation $x(t)$, which then negatively feedbacks on the convective activity $y(t)$. Then $y(t)$ decreases, which reduces precipitation $x(t)$ until $x(t)$ is small enough so the convective activity $y(t)$ increases again. Details of the physical processes causing this feedback loops would have to be explored with further data analysis.

In a second step such convective oscillators could be coupled. To provide a starting point for a such a toy model, we now discuss the simulated oscillations of the two final remaining convective cells.

8.3. Two coupled oscillators

Coupled oscillators are a common association for two related, but different phenomena [Hel+21]. In the simplest case both oscillators affect each other mutually, which often leads to synchronisation.

If we suppose differences for the interaction between convective cells, for example depending on their respective rain intensity, this enables diverse dynamic¹². The resulting dynamic depends on the type of interaction, which can be a positive, a negative coupling, or an asymmetric coupling, and the coupling strengths [Hel+21]. Our discussion above, which invokes a global energy constraint, implies an two-way inhibitory coupling. In line with an inhibitory coupling, rain intensities of the last two remaining clusters $I(\text{cluster}_1, t)$ and $I(\text{cluster}_2, t)$ ¹³ oscillate with lower amplitude and reach lower maximum values than the final cluster $I(\text{cluster}_{final}, t)$ (Fig. 8.4).

¹¹Unfortunately, it is very difficult to discuss 'what causes what' in a complex synchronised system.

For example causal relations become quite invalid as soon as systems synchronise, since 'cause' and 'effect' can not really be separated anymore.

¹²from synchrony to chaos [Pik+03]

¹³Since it is a matter of definition which cluster continues after a merge, the 'last cluster' (cluster_1 , navy) is the one reaching a higher precipitation maximum ($\max(I(\text{cluster}_1, t))=22\,060\text{ mm/h} > \max(I(\text{cluster}_2, t))=15\,975\text{ mm/h}$) in this example.

8.3. Two coupled oscillators

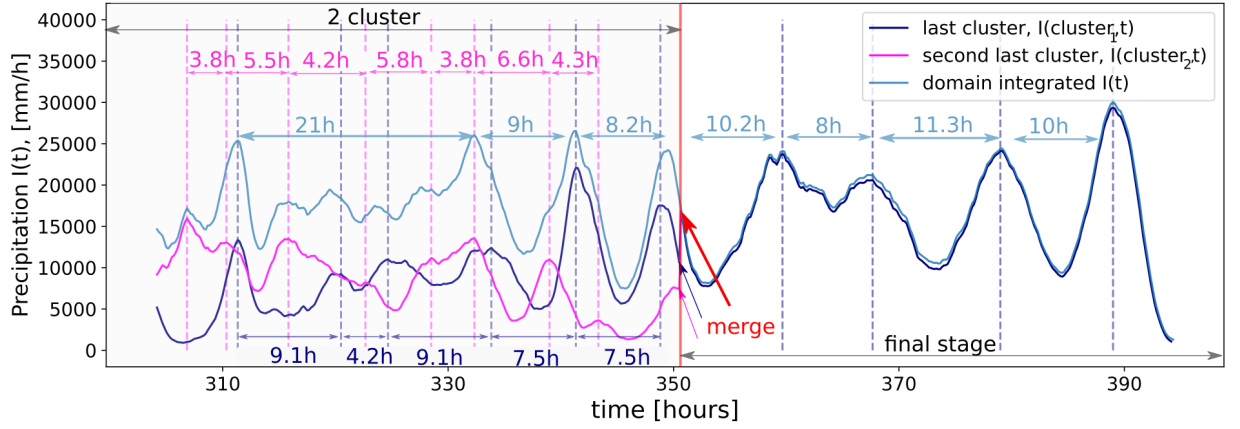


Figure 8.4.: Precipitation falling inside the last two remaining clusters ($I(\text{cluster}_1, t)$ navy, $I(\text{cluster}_2, t)$ magenta) and over the domain $I(t)$ (light blue). After 330 h two rain clusters are left (light grey shading), which then merge after 350 h (final stage, white background). Vertical dashed lines of respective colours indicate temporal maxima of the variables and times corresponds to the periods (arrows) between the respective maxima.

In general, we suppose convective cells with stronger precipitation to oscillate with longer periods and maybe larger absolute amplitudes. Consistent with this assumption, the stronger precipitating cluster₁ oscillates with a longer periods than cluster₂ and in the final stage $I(\text{cluster}_{final}, t)$ reaches the highest peaks and the longest periods (around 10 h, Fig. 6.5). As mentioned, $\bar{I}(t)$ increasing in the final state seems also linked to radiative constraints [Win+17].

The oscillations of $I(\text{cluster}_1, t)$ and $I(\text{cluster}_2, t)$ are not in phase and could rather be associated with an anti-phase oscillation between 305 h and 335 h. These out of phase oscillations cause varying and long intervals between two maxima of domain integrated precipitation $I(t)$ (8.1 h – 21 h, Fig. 8.4, light-blue). Interestingly, $I(\text{cluster}_1, t)$ and $I(\text{cluster}_2, t)$ show a tendency to synchronise already (around 345 h) before they merge (at 350 h). This synchronisation could be a hint that the corresponding convective cells fuse already, before the raining patches overlap.

Again, to investigate the observed interaction further, we suggest to model individual rain cells as convective oscillators. Each convective oscillator could have a predator-prey like dynamic (Sec. 8.2) and an inhibitory coupling could be provided between their precipitation variables $x_1(t)$ and $x_2(t)$.

Such coupling could correspond to the global energy constraint, limiting the global rainfall. Negatively coupled oscillators can develop anti-phase synchronisation [Ble88], which seems to match the interaction of the last two remaining convective cells (Fig. 8.4).

8. Discussion

To expand such a toy model on earlier simulation stages, we would have to consider $N_{cluster}$ convective cells and therefore model N coupled convective oscillators. Some aspects of the suggested two-oscillator model could transfer to such a framework, if we focus on nearest-neighbour interactions. Another possible simplification would be, to couple all convective oscillators equally (again corresponding to the weak temperature gradient approximation). The issue of ' N coupled oscillators' is discussed further with A.4.1.

8.4. Conclusion

This thesis contributes to the state of the art by providing a detailed analysis of CSA in the absence of CPs (NoEvap case), which has not been studied in depth before. When considering CSA as resulting from a bifurcation in a bi-stable moisture distribution [WH19], we claim that this bifurcation always occurs in the NoEvap case¹⁴, but is prevented if CPs are present (FullEvap). For NoEvap simulations, initial noise seems sufficient to 'break symmetry', thus creating a dry spot that triggers CSA [BH19].

To analyse our simulations, two tracking algorithms are presented, one based on the precipitation output and one based on the horizontal wind field in the lowest layer. We observe CSA as a decreasing number of convective cells that either decay or merge pairwise until only one cell is finally left. Whereas no final physical explanation can be given for the decaying process, we discuss possible feedback mechanisms and how these could be associated with the observed stability increase of the atmosphere (Sec. 6.3, Sec. 8.1).

Within the transient approach towards CSA we suppose that merging dominates the dynamic. A hypothetical physical explanation for this merging is an asymmetric inflow of convective updraughts (Sec. 7.3.3). Prospectively, we will capture this merging in a dynamical model, as suggested in Sec. 7.4. We already implement the foundation for such an update dynamic: based on a gravity approach a prognosis vector for the displacement of each convective updraught is calculated. Comparing this prognosis vector to the actual displacement vector, extracted from the data, yields significant positive correlations.

Our static model however offers a discrete and highly compressed reconstruction of the essence of simulated convection, on the level of individual convective cells: each updraught is characterised by its location and the area of its inflow basin only. Determining inflow areas in the reconstructed flow field provides very close accordance to the inflow areas of the original data in 464/520 samples. In this context we implement a framework allowing to divide the domain into 'basins' converging

¹⁴which is in line with Jeevanjee and Romps, [JR13]

towards each convective updraught for each time of the simulation. These basins allow to derive a dynamical length scaling, which we discuss in respect to established coarsening models that describe spatial development of CSA (Sec. 6.5, Fig. 6.12). The basin framework will be valuable for a systematic analysis of physical variables involved in CSA.

Furthermore, we describe interesting oscillations in precipitation (Sec. 6.2) and other physical variables, which we analyse (Sec. 8.1.3) and interpret as coupled oscillators (Sec. 8.2). Overall, we suspect these oscillations to relate to cloud-radiative feedbacks driving CSA and aim to investigate these results further within future work.

8.5. Outlook

We aim to perform a systematical data analysis and evaluate data statistically, based on the basin framework. Therefore we will simulate larger domains with a larger number of convective cells, for example a $960 \text{ km} \times 960 \text{ km}$ simulation with 1 km resolution has already been set up. To analyse such larger data sets, we suggest to improve the basin algorithm as discussed in Sec. 5.5.3. The resulting toolbox will not only allow to track the number and size of basins over time, but to take all physical variables of the simulation output into account. For example, we could detect the basin averaged rain intensity $\overline{I(\text{basin},t)}$ and investigate how it correlates with basin area and how these correlations change over the simulations.

Further, we aim to quantify how the low-level horizontal wind speed surrounding an updraught is linked to the updraught strength and connect this to a relation between basin area and the likelihood of an updraught to decay (Sec. 8.1). If we are able to derive a mathematical expression for this relation, we could include this as an additional update mechanism into our model. Foremost we aim to include an update dynamic that captures the merging of convective cells. One suggestion is presented in Sec. 7.4, which only needs a careful investigation of parameters to be implemented. An phenomenological alternative to mimic merging could be to simply move updraughts towards their nearest neighbour so the corresponding basins fuse once the updraughts are close enough.

Additionally, results of the data analysis might allow to derive an analytical model, capturing the essential physical mechanisms in the transient stage (Sec. 8.1.1). The final stage could be captured in an oscillator model, with a (quasi)-periodic solution. As discussed above, a model of two or more coupled oscillators could be sought (Sec. 8.3, Sec. A.4.1), in order to improve the understanding of oscillations in the transient stage.

We are aware that NoEvap is generally quite artificial and that our analysis can

8. Discussion

not be strictly attributed to either shallow or deep convective dynamics [Ste05]. This ambiguity is stressed by the bi-modal cloud distribution (Fig. 8.2). Therefore, it could be interesting to test multiple, very different perspectives on the NoEvap data. For example, detected basins could be analysed similar to the cellular network approach suggested by Glassmeier and Feingold, which analyses patterns in shallow convective clouds. Despite many of the geometric properties not being transferable, our basins offer domain filling objects with well defined areas and numbers of neighbours, which is required to apply their algorithm [GF17].

From a methodical perspective it would be interesting to investigate, how results between the two tracking methods coincide, especially since both algorithms have their own advantages: The basin approach allows wider analysis of simulation data, but in the theoretical our rain tracking algorithm could allow to analyse observational (precipitation) data¹⁵. Unfortunately the NoEvap simulations do not correspond to any observational data. Nevertheless, comparable merging dynamics could be investigated in larger scale observational data sets, since Patrizio and Randall found merging in coarse, large domain simulations, which do not resolve CPs, too (Sec. 8.1).

Further, it would be meaningful to investigate slightly more realistic simulations with low to intermediate re-evaporation (Fig. 6.1). We suppose multiple competing mechanisms determine when and how aggregation occurs. Evap01, Evap02 dynamics differ from the NoEvap case, which could limit the extent to which the basin approach would be appropriate. Such reduced re-evaporation analysis could hypothetically be associated to observational data, for example precipitation data from regions with little re-evaporation as in the very moist boundary layer above the tropical sea [Ste+05].

Within a much wider perspective, understanding aggregation and merging dynamics contribute to understanding how phenomena as tropical cyclones, possibly associated to CSA, form in the real atmosphere. To better understand such extreme weather and precipitation events enables to improve early warning systems. Such real time prognoses can prevent harm resulting from extreme events. Due to climate change, such extreme events are expected to even increase in frequency and severity [ONe+17] and, since they often impact vulnerable regions (as the tropics) most, should be a research focus of the coming years.

¹⁵For such an application the rain tracking would need to be improved as discussed in Sec. 5.3.3

Bibliography

- [AR15] Nathan P Arnold and David A Randall. “Global-scale convective aggregation: Implications for the Madden-Julian Oscillation”. In: *Journal of Advances in Modeling Earth Systems* 7.4 (2015), pp. 1499–1518.
- [BB08] Linge Bai and David Breen. “Calculating center of mass in an unbounded 2D environment”. In: *Journal of Graphics Tools* 13.4 (2008), pp. 53–60.
- [BBK05a] Christopher S Bretherton, Peter N Blossey, and Marat Khairoutdinov. “An energy-balance analysis of deep convective self-aggregation above uniform SST”. In: *Journal of the atmospheric sciences* 62.12 (2005), pp. 4273–4292.
- [BBK05b] Christopher S. Bretherton, Peter N. Blossey, and Marat Khairoutdinov. “An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST”. In: *Journal of the Atmospheric Sciences* 62.12 (2005), pp. 4273–4292. DOI: [10.1175/JAS3614.1](https://doi.org/10.1175/JAS3614.1). URL: <https://journals.ametsoc.org/view/journals/atsc/62/12/jas3614.1.xml>.
- [Bet82] Alan K Betts. “Saturation point analysis of moist convective overturning”. In: *Journal of the Atmospheric Sciences* 39.7 (1982), pp. 1484–1505.
- [BH19] Silas Boye Nissen and Jan O. Haerter. “How weakened cold pools open for convective self-aggregation”. In: *arXiv e-prints*, arXiv:1911.12849 (Nov. 2019), arXiv:1911.12849. arXiv: [1911.12849](https://arxiv.org/abs/1911.12849) [[physics.a0-ph](https://arxiv.org/abs/1911.12849)].
- [Ble88] Ilia Izrailevich Blekhman. *Synchronization in science and technology*. ASME press, 1988.
- [BOS11] Yann Brenier, Felix Otto, and Christian Seis. “Upper bounds on coarsening rates in demixing binary viscous liquids”. In: *SIAM journal on mathematical analysis* 43.1 (2011), pp. 114–134.
- [BPB04] Christopher S Bretherton, Matthew E Peters, and Larissa E Back. “Relationships between water vapor path and precipitation over the tropical oceans”. In: *Journal of climate* 17.7 (2004), pp. 1517–1528.
- [BR89] Alan K Betts and W Ridgway. “Climatic equilibrium of the atmospheric convective boundary layer over a tropical ocean”. In: *Journal of Atmospheric Sciences* 46.17 (1989), pp. 2621–2641.

Bibliography

- [Bra03] AJ Bray. “Coarsening dynamics of phase-separating systems”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1805 (2003), pp. 781–792.
- [Bro+12] Ilja N Bronstein et al. *Taschenbuch der mathematik*. Vol. 1. Springer-Verlag, 2012.
- [CHR10] F Couvreur, F Hourdin, and C Rio. “Resolved versus parametrized boundary-layer plumes. Part I: A parametrization-oriented conditional sampling in large-eddy simulations”. In: *Boundary-layer meteorology* 134.3 (2010), pp. 441–458.
- [CM13] GC Craig and JM Mack. “A coarsening model for self-organization of tropical convection”. In: *Journal of Geophysical Research: Atmospheres* 118.16 (2013), pp. 8761–8769.
- [Cra79] Robert K Crane. “Automatic cell detection and tracking”. In: *IEEE Transactions on geoscience electronics* 17.4 (1979), pp. 250–262.
- [CY99] Guanrong Chen and Xinghuo Yu. “On time-delayed feedback control of chaotic systems”. In: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 46.6 (1999), pp. 767–772.
- [DBB06] Christopher Davis, Barbara Brown, and Randy Bullock. “Object-based verification of precipitation forecasts. Part II: Application to convective rain systems”. In: *Monthly Weather Review* 134.7 (2006), pp. 1785–1795.
- [Ema+94] Kerry A Emanuel et al. *Atmospheric convection*. Oxford University Press on Demand, 1994.
- [Ema86] Kerry A Emanuel. “An air-sea interaction theory for tropical cyclones. Part I: Steady-state maintenance”. In: *Journal of Atmospheric Sciences* 43.6 (1986), pp. 585–605.
- [EWW14] Kerry Emanuel, Allison A Wing, and Emmanuel M Vincent. “Radiative-convective instability”. In: *Journal of Advances in Modeling Earth Systems* 6.1 (2014), pp. 75–90.
- [FH20] Herman F Fuglestedt and Jan O Haerter. “Cold pools as conveyor belts of moisture”. In: *Geophysical Research Letters* 47.12 (2020), e2020GL087319.
- [FR02] J Fröhlich and W Rodi. *Introduction to large eddy simulation of turbulent flows*. 2002.
- [FS76] Carl A Friehe and Kurt F Schmitt. “Parameterization of air-sea interface fluxes of sensible heat and moisture by the bulk aerodynamic formulas”. In: *Journal of Physical Oceanography* 6.6 (1976), pp. 801–809.
- [GC17] Françoise Guichard and Fleur Couvreur. “A short review of numerical cloud-resolving models”. In: *Tellus A: Dynamic Meteorology and Oceanography* 69.1 (2017), p. 1373578.

- [GF17] Franziska Glassmeier and Graham Feingold. “Network approach to patterns in stratocumulus clouds”. In: *Proceedings of the National Academy of Sciences* 114.40 (2017), pp. 10578–10583.
- [Gra+20] Leah D Grant et al. “Shear-parallel tropical convective systems: Importance of cold pools and wind shear”. In: *Geophysical Research Letters* 47.12 (2020), e2020GL087720.
- [Hae+19] Jan O Haerter et al. “Circling in on convective organization”. In: *Geophysical Research Letters* 46.12 (2019), pp. 7024–7034.
- [Hae19] Jan O Haerter. “Convective self-aggregation as a cold pool-driven critical phenomenon”. In: *Geophysical Research Letters* 46.7 (2019), pp. 4017–4028.
- [Hei+16] Reuven H Heiblum et al. “Characterization of cumulus cloud fields using trajectories in the center of gravity versus water mass phase space: 1. Cloud tracking and phase space description”. In: *Journal of Geophysical Research: Atmospheres* 121.11 (2016), pp. 6336–6355.
- [Hei+19] M. Heikenfeld et al. “tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets”. In: *Geoscientific Model Development* 12.11 (2019), pp. 4551–4570. DOI: [10.5194/gmd-12-4551-2019](https://doi.org/10.5194/gmd-12-4551-2019). URL: <https://gmd.copernicus.org/articles/12/4551/2019/>.
- [Hel+21] Mathias L Heltberg et al. “A tale of two rhythms: Locked clocks and chaos in biology”. In: *Cell systems* 12.4 (2021), pp. 291–303.
- [HHR93] Isaac M Held, Richard S Hemler, and V Ramaswamy. “Radiative-convective equilibrium with explicit two-dimensional moist convection”. In: *Journal of Atmospheric Sciences* 50.23 (1993), pp. 3909–3927.
- [Hir+20] Mirjam Hirt et al. “Cold pool driven convective initiation: using causal graph analysis to determine what convection permitting models are missing”. In: *Quarterly Journal of the Royal Meteorological Society* (2020).
- [HK76] Joseph Hoshen and Raoul Kopelman. “Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm”. In: *Physical Review B* 14.8 (1976), p. 3438.
- [Hol73] James R Holton. “An introduction to dynamic meteorology”. In: *American Journal of Physics* 41.5 (1973), pp. 752–754.
- [HS16] Cathy Hohenegger and Bjorn Stevens. “Coupled radiative convective equilibrium simulations with explicit and parameterized convection”. In: *Journal of Advances in Modeling Earth Systems* 8.3 (2016), pp. 1468–1482.
- [JR13] Nadir Jeevanjee and David M Romps. “Convective self-aggregation, cold pools, and domain size”. In: *Geophysical Research Letters* 40.5 (2013), pp. 994–998.

Bibliography

- [JR18] Nadir Jeevanjee and David M Romps. “Mean precipitation change from a deepening troposphere”. In: *Proceedings of the National Academy of Sciences* 115.45 (2018), pp. 11465–11470.
- [KR03] Marat F Khairoutdinov and David A Randall. “Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities”. In: *Journal of Atmospheric Sciences* 60.4 (2003), pp. 607–625.
- [KS79] Tiruvalam Natarajan Krishnamurti and Lydia Stefanova. *Tropical meteorology*. Springer, 1979.
- [Kur+18] Marcin J Kurowski et al. “Shallow-to-deep transition of continental moist convection: Cold pools, surface fluxes, and mesoscale organization”. In: *Journal of the Atmospheric Sciences* 75.12 (2018), pp. 4071–4090.
- [Lam07] Harvey SH Lam. “On the Effective Stefan-Boltzmann Temperature Exponent of the Earth”. In: (2007).
- [Lif62] IM Lifshitz. “Kinetics of ordering during second-order phase transitions”. In: *Sov. Phys. JETP* 15 (1962), p. 939.
- [LL67] LD Landau and EM Lifschitz. *Lehrbuch der theoretischen Physik, Band VIII: Elektrodynamik der Kontinua*. Akademie-Verlag, 1967.
- [Lot25] Alfred James Lotka. *Elements of physical biology*. Williams & Wilkins, 1925.
- [MB15] Caroline Muller and Sandrine Bony. “What favors convective aggregation and why?” In: *Geophysical Research Letters* 42.13 (2015), pp. 5626–5634.
- [Men02] Annette Menzel. “Phenology: its importance to the global change community”. In: *Climatic change* 54.4 (2002), p. 379.
- [MH12] Caroline J Muller and Isaac M Held. “Detailed investigation of the self-aggregation of convection in cloud-resolving simulations”. In: *Journal of the Atmospheric Sciences* 69.8 (2012), pp. 2551–2565.
- [MH20] Bettina Meyer and Jan O Haerter. “Mechanical Forcing of Convection by Cold Pools: Collisions and Energy Scaling”. In: *Journal of Advances in Modeling Earth Systems* 12.11 (2020), e2020MS002281.
- [MHH19] Christopher Moseley, Olga Henneberg, and Jan O Haerter. “A statistical model for isolated convective precipitation events”. In: *Journal of Advances in Modeling Earth Systems* 11.1 (2019), pp. 360–375.
- [MO54] AS Monin and AM Obukhov. “Basic laws of turbulent mixing in the atmosphere near the ground”. In: *Tr. Geofiz. Inst., Akad. Nauk SSSR* 24.151 (1954), pp. 163–187.

- [Mos+19] Christopher Moseley et al. “Impact of resolution and air temperature on Large Eddy Simulation of mid-latitude summer time convection”. In: *Atmospheric Chemistry and Physics Discuss* 2019 (2019), pp. 1–25.
- [MS90] Renato E Mirollo and Steven H Strogatz. “Synchronization of pulse-coupled biological oscillators”. In: *SIAM Journal on Applied Mathematics* 50.6 (1990), pp. 1645–1662.
- [Mur82] JD Murray. “Parameter space for Turing instability in reaction diffusion mechanisms: a comparison of models”. In: *Journal of Theoretical Biology* 98.1 (1982), pp. 143–163.
- [Nol16] Wolfgang Nolting. “Electrodynamics”. In: *Theoretical Physics 3*. Springer, 2016, pp. 207–419.
- [ONe+17] Brian C O’Neill et al. “IPCC reasons for concern regarding climate change risks”. In: *Nature Climate Change* 7.1 (2017), pp. 28–37.
- [Pal48] Erik Palmén. “On the formation and structure of tropical hurricanes”. In: *Geophysica* 3.1 (1948), pp. 26–38.
- [Pik+03] Arkady Pikovsky et al. *Synchronization: a universal concept in nonlinear sciences*. 12. Cambridge university press, 2003.
- [Pos+12] Derek J Posselt et al. “Changes in the interaction between tropical convection, radiation, and the large-scale circulation in a warming environment”. In: *Journal of climate* 25.2 (2012), pp. 557–571.
- [PPP01] Markus Pahlow, Marc B Parlange, and FERNANDO Porté-Agel. “On Monin–Obukhov similarity in the stable atmospheric boundary layer”. In: *Boundary-layer meteorology* 99.2 (2001), pp. 225–248.
- [PR19] Casey R Patrizio and David A Randall. “Sensitivity of convective self-aggregation to domain size”. In: *Journal of Advances in Modeling Earth Systems* 11.7 (2019), pp. 1995–2019.
- [PS09] Robert Pincus and Bjorn Stevens. “Monte Carlo spectral integration: A consistent approximation for radiative transfer in large eddy simulations”. In: *Journal of Advances in Modeling Earth Systems* 1.2 (2009).
- [PVS08] DJ Posselt, SC Van Den Heever, and GL Stephens. “Trimodal cloudiness and tropical stable layers in simulations of radiative convective equilibrium”. In: *Geophysical Research Letters* 35.8 (2008).
- [Ran09] David A Randall. *The Moist Adiabatic Lapse Rate*. 2009.
- [Ray+09] David J Raymond et al. “The mechanics of gross moist stability”. In: *Journal of Advances in Modeling Earth Systems* 1.3 (2009).
- [RE87] Richard Rotunno and Kerry A Emanuel. “An air–sea interaction theory for tropical cyclones. Part II: Evolutionary study using a nonhydrostatic axisymmetric numerical model”. In: *Journal of Atmospheric Sciences* 44.3 (1987), pp. 542–561.

Bibliography

- [ROH05] Juan G Restrepo, Edward Ott, and Brian R Hunt. “Onset of synchronization in large networks of coupled oscillators”. In: *Physical Review E* 71.3 (2005), p. 036151.
- [Rom14] David M Romps. “An analytical model for tropical relative humidity”. In: *Journal of Climate* 27.19 (2014), pp. 7432–7449.
- [RZ00] David J Raymond and Xiping Zeng. “Instability and large-scale circulations in a two-column model of the tropical troposphere”. In: *Quarterly Journal of the Royal Meteorological Society* 126.570 (2000), pp. 3117–3135.
- [SA95] Roland B Stull and C Donald Ahrens. *Meteorology today for scientists and engineers*. West Pub., 1995.
- [Sat+19] Masaki Satoh et al. “Global cloud-resolving models”. In: *Current Climate Change Reports* 5.3 (2019), pp. 172–184.
- [SB06] Axel Seifert and Klaus Dieter Beheng. “A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description”. In: *Meteorology and atmospheric physics* 92.1 (2006), pp. 45–66.
- [SBT13] Gregor Skok, Julio Bacmeister, and Joseph Tribbia. “Analysis of tropical cyclone precipitation using an object-based algorithm”. In: *Journal of climate* 26.8 (2013), pp. 2563–2579.
- [Sch+09] Eckehard Schöll et al. “Time-delayed feedback in neurosystems”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367.1891 (2009), pp. 1079–1096.
- [Sel00] A. P. S. Selvadurai. “Poisson’s equation”. In: *Partial Differential Equations in Mechanics 2: The Biharmonic Equation, Poisson’s Equation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 503–647. ISBN: 978-3-662-09205-7. DOI: [10.1007/978-3-662-09205-7_2](https://doi.org/10.1007/978-3-662-09205-7_2). URL: https://doi.org/10.1007/978-3-662-09205-7_2.
- [SK94] A Shapiro and YL Kogan. “On vortex formation in multicell convective clouds in a shear-free environment”. In: *Atmospheric research* 33.1-4 (1994), pp. 125–136.
- [Sma63] Joseph Smagorinsky. “General circulation experiments with the primitive equations: I. The basic experiment”. In: *Monthly weather review* 91.3 (1963), pp. 99–164.
- [SNP01] Adam H Sobel, Johan Nilsson, and Lorenzo M Polvani. “The weak temperature gradient approximation and balanced tropical moisture waves”. In: *Journal of the atmospheric sciences* 58.23 (2001), pp. 3650–3665.

- [Ste+05] Bjorn Stevens et al. “Evaluation of Large-Eddy Simulations via Observations of Nocturnal Marine Stratocumulus”. In: *Monthly Weather Review* 133.6 (2005), pp. 1443–1462. DOI: [10.1175/MWR2930.1](https://doi.org/10.1175/MWR2930.1). URL: <https://journals.ametsoc.org/view/journals/mwre/133/6/mwr2930.1.xml>.
- [Ste05] Bjorn Stevens. “Atmospheric moist convection”. In: *Annu. Rev. Earth Planet. Sci.* 33 (2005), pp. 605–643.
- [Str04] Steven Strogatz. *Sync: The emerging science of spontaneous order*. Penguin UK, 2004.
- [Tak09] Ken Takahashi. “Radiative constraints on the hydrological cycle in an idealized radiative–convective equilibrium model”. In: *Journal of the atmospheric sciences* 66.1 (2009), pp. 77–91.
- [TC98] Adrian M Tompkins and George C Craig. “Radiative–convective equilibrium in a three-dimensional cloud-ensemble model”. In: *Quarterly Journal of the Royal Meteorological Society* 124.550 (1998), pp. 2073–2097.
- [Tom01] Adrian M. Tompkins. “Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Water Vapor”. In: *Journal of the Atmospheric Sciences* 58.6 (1Mar. 2001), pp. 529–545. DOI: [10.1175/1520-0469\(2001\)058<0529:OOTCIL>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<0529:OOTCIL>2.0.CO;2). URL: https://journals.ametsoc.org/view/journals/atsc/58/6/1520-0469_2001_058_0529_ootcil_2.0.co_2.xml.
- [Tsu14] Toshitaka Tsuda. “Characteristics of atmospheric gravity waves observed using the MU (Middle and Upper atmosphere) radar and GPS (Global Positioning System) radio occultation”. In: *Proceedings of the Japan Academy, Series B* 90.1 (2014), pp. 12–27.
- [Tur82] Olli Turpeinen. “Cloud interactions and merging on day 261 of GATE”. In: *Monthly Weather Review* 110.9 (1982), pp. 1238–1254.
- [Vol36] Vito Volterra. “Leçons sur la théorie mathématique de la lutte pour la vie”. In: *Bull. Amer. Math. Soc* 42 (1936), pp. 304–305.
- [Wan78] Peter J Wangersky. “Lotka-Volterra population models”. In: *Annual Review of Ecology and Systematics* 9.1 (1978), pp. 189–218.
- [WBS17] RH White, DS Battisti, and G Skok. “Tracking precipitation events in time and space in gridded observational data”. In: *Geophysical Research Letters* 44.16 (2017), pp. 8637–8646.
- [WE14] Allison A Wing and Kerry A Emanuel. “Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations”. In: *Journal of Advances in Modeling Earth Systems* 6.1 (2014), pp. 59–74.

Bibliography

- [WH19] JM Windmiller and Cathy Hohenegger. “Convection on the edge”. In: *Journal of Advances in Modeling Earth Systems* 11.12 (2019), pp. 3959–3972.
- [Win+17] Allison A Wing et al. “Convective self-aggregation in numerical simulations: A review”. In: *Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity*. Springer, 2017, pp. 1–25.
- [Win+18] Allison A Wing et al. “Radiative–convective equilibrium model inter-comparison project”. In: *Geoscientific Model Development* 11.2 (2018), pp. 793–813.
- [Win17] Julia Miriam Windmiller. “Organization of tropical convection”. PhD thesis. Imu, 2017.
- [XE89] Kuan-man Xu and Kerry A Emanuel. “Is the tropical atmosphere conditionally unstable?” In: *Monthly Weather Review* 117.7 (1989), pp. 1471–1479.
- [YB09] Jun-Ichi Yano and Marine Bonazzola. “Scale analysis for large-scale tropical atmospheric dynamics”. In: *Journal of the atmospheric sciences* 66.1 (2009), pp. 159–172.
- [Zui+17] Paquita Zuidema et al. “A survey of precipitation-induced atmospheric cold pools over oceans and their interactions with the larger-scale environment”. In: *Surveys in Geophysics* 38.6 (2017), pp. 1283–1305.

A. Appendix

A.1. Methods supplements

A.1.1. TOBAC tracking algorithm, applicable to a wide variety of situations

The Tracking and Object-Based Analysis of Clouds (TOBAC, [Hei+19]) project aims to address the issue of limited compatibility between different existing approaches. The project claims, that most algorithms for resolving individual updraughts exist for high resolution data only, where in contrast grid spacing usually is up to several km when simulating larger domains and longer time intervals. TOBAC offers a toolbox for an object-based analysis of a variety of cloud types, including deep convection in LES and CRM data. The flexible framework is set up in a modular way and thus suitable for wide range of data sets. It provides modules concerned with handling input and output data, feature detection, segmentation of cloud areas, a trajectory linking and finally object-based analysis tools. It is implemented in python, uses established python libraries and is published under Creative Commons license.¹

Potentially, existing approaches could be integrated to the TOBAC framework. In theory this could include features of the algorithm presented in 5.3, which is implemented in python too, but stands on its own.

¹[Hei+19] Creative Commons Attribution 4.0 License, code available at <https://github.com/climate-processes/tobac>

A. Appendix

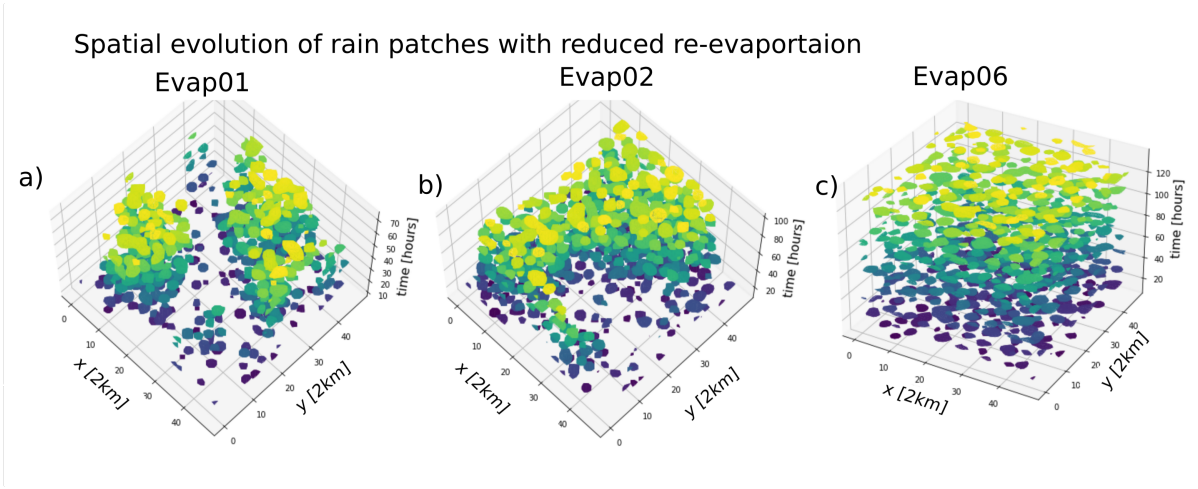


Figure A.1.: Rain patches over time, for simulations with **Evap01** (a), **Evap02** (b) and **Evap06** (c). With little re-evaporation (a and b), rain patches are restricted to localised areas quickly (after about three days, a and b) and a persistent dry area emerges, whereas no aggregation occurs for Evap06 (c). Interestingly, aggregation pattern occur faster with weak CPs (Evap01, Evap02) than in their absence (NoEvap) (Sec. 6.0.1).

A.2. Results supplements

A.3. Model supplements

A.3.1. Scale Analysis, estimating the Reynolds number

We insert a typical horizontal velocity of synoptic scales of $U \approx 10 \text{ m s}^{-1}$, which is comparable to the maximal velocities measured at our lowest layer (compare Fig. 8.1). For the typical length scale of $L \approx 10^4 \text{ m}$ we approximate twice the domain size ($L \approx 2 \cdot 480 \text{ km} \approx 10^7 \text{ m}$). The viscosity is a material constant, which is approximated with the magnitude $\nu \approx 10^{-5} \text{ m}^2 \text{ s}^{-1}$ for air. Applying these values in $\text{Re} = \frac{U \tilde{L}}{\nu}$ leads to a high Reynolds-number [Hol73]:

$$\text{Re} = \frac{U \cdot \tilde{L}}{\nu} = \frac{10 \text{ m} \cdot 10^7 \text{ m/s}}{10^{-5} \text{ m}^2/\text{s}} = 10^{13} \quad (\text{A.1})$$

A.3.2. Fourier expansion of the source function

$$\begin{aligned}
W &= \left(1 + \cos\left(2\pi \frac{x - x_0}{L}\right)\right)^N \text{ with : } \cos(x - x_0)^2 = \frac{1 + \cos(2(x - x_0))}{2} \\
&= 2\cos\left(\pi \frac{x - x_0}{L}\right)^{2N} \\
&= 2^N \left(\frac{1}{2} \cdot e^{i\pi \frac{x-x_0}{L}} + e^{-i\pi \frac{x-x_0}{L}}\right)^{2N} \text{ with : } (a + b)^n = \sum_{k=0}^{k=n} \binom{n}{k} a^{n-k} b^k \\
&= \frac{2^N}{2^{2N}} \sum_{k=0}^{k=2N} \binom{2N}{k'} e^{2i\pi \frac{(x-x_0)}{L} (2N-k')} e^{-i\pi \frac{(x-x_0)}{L} k'}
\end{aligned}$$

to centre around zero shift index by substituting $k' = N - k$,

$$\begin{aligned}
\longrightarrow k = N - k', \longrightarrow \sum_{k'=N-0}^{k'=N-2N} &= \sum_{k'=N}^{k'=-N} \\
&= \frac{2^N}{2^{2N}} \sum_{k=-N}^{k=N} \binom{2N}{N-k} e^{i\pi \frac{(x-x_0)}{L} 2k}
\end{aligned}$$

split the sum in $N > 0$ and $N < 0$ and re-write combine the exponential pairwise to cosine

$$= \frac{2^N}{2^{2N-1}} \sum_{k=0}^{k=N} \binom{2N}{N-k} \cos\left(2\pi k \frac{(x - x_0)}{L}\right) \chi_k$$

ensure to treat $k=0$ separately using χ_k :

$$\text{where } \chi_k \begin{cases} 1/2 & \text{for } k = 0 \\ 1, & \text{for } k > 0 \end{cases} \tag{A.2}$$

$$\tag{A.3}$$

Now we want to normalise the result so $\int_{-L/2}^{L/2} W = 1$. This will satisfy the continuity constrain. We first look at a single mode ($k \neq 0$) of the Fourier-expansion, which we integrate over the entire domain:

$$\begin{aligned}
&\frac{2^N}{2^{2N-1}} \int_{-L/2}^{L/2} \binom{2N}{N-k} \cos\left(\pi k \frac{(x - x_0)}{L}\right) dx \\
&= \frac{2^N}{2^{2N-1}} \cdot \binom{2N}{N-k} \int_0^L \cos\left(\pi k \frac{(x - x_0)}{L}\right) dx \\
&= 0
\end{aligned}$$

The cosine function integrated over the entire domain is always zero. Thus, regarding normalisation only the constant mode $k = 0$, $\rightarrow \cos(0) = 1$ has to be taken into

A. Appendix

account:

$$1 \stackrel{!}{=} c \cdot \int_{-L/2}^{L/2} \cos(0) dx \quad (\text{A.4})$$

$$1 = c \cdot \int_0^L 1 dx \quad (\text{A.5})$$

$$1 = c \cdot (L - 0) \quad (\text{A.6})$$

$$\longrightarrow c = 1/L \quad (\text{A.7})$$

Extension to 2D

To extend the problem to two dimensions, the function W has to be periodic and real valued $W(x, y) = W(x + L_x, y)$ and $W(x, y) = W(x, y + L_y)$, for $x, y, L \in \mathbb{R}$. Further, a square domain has sides with equal lengths: $L_x \stackrel{!}{=} L_y = L$. For the two-dimensional Ansatz we choose $W(x, y) = (1 + \cos(2\pi \frac{x-x_0}{L} k))^N \cdot (1 + \cos(2\pi \frac{y-y_0}{L} l))^N$.

Analogously to the one-dimensional derivation above, $W(x, y) = W(x) \cdot W(y)$ can be expressed:

$$W(x, y) = W(x) \cdot W(y) \quad (\text{A.8})$$

$$= \left[\cos\left(\pi k \frac{(x-x_0)}{L}\right) \cos\left(\pi l \frac{(y-y_0)}{L}\right) \right]^{2N} \quad (\text{A.9})$$

$$= \cos\left(\pi k \frac{(x-x_0)}{L}\right)^{2N} \cdot \cos\left(\pi l \frac{(y-y_0)}{L}\right)^{2N} \quad (\text{A.10})$$

$$= \frac{1}{(2^{2N-1})^2} \sum_{k,l=0}^{k,l=N} \binom{2N}{N-k} \cos\left(2\pi k \frac{(x-x_0)}{L}\right) \binom{2N}{N-l} \cos\left(2\pi l \frac{(y-y_0)}{L}\right) \frac{1}{(\delta_{k0})(\delta_{l0})} \quad (\text{A.11})$$

$$\text{where } \delta_{k,0}(\text{or } \delta_{l,0}) = \begin{cases} 1, & \text{for } k = 0 \text{ (or } l = 0) \\ 0, & \text{for } k \neq 0 \text{ (or } l \neq 0) \end{cases} \quad (\text{A.12})$$

$$(\text{A.13})$$

In order to write these in the form of a standard Fourier-Series the following relation is used:

$$\cos(a + b) = \cos(a)\cos(b) + \sin(a)\sin(b) \quad (\text{A.14})$$

This results to:

$$\begin{aligned}
& \cos\left(2\pi k \frac{(x-x_0)}{L}\right) \cos\left(2\pi l \frac{(y-y_0)}{L}\right) \\
&= \cos\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) \cos\left(\frac{2\pi}{L} kx\right) \cos\left(\frac{2\pi}{L} ly\right) \\
&+ \cos\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right) \cos\left(\frac{2\pi}{L} kx\right) \sin\left(\frac{2\pi}{L} ly\right) \\
&+ \sin\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) \sin\left(\frac{2\pi}{L} kx\right) \cos\left(\frac{2\pi}{L} ly\right) \\
&+ \sin\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right) \sin\left(\frac{2\pi}{L} kx\right) \sin\left(\frac{2\pi}{L} ly\right)
\end{aligned}$$

We are aiming to get a normalised form of a general Fourier expansion:

$$\begin{aligned}
f &= \sum_{k,l} a_{kl} \cos\left(\frac{2\pi}{L} x\right) \cos\left(\frac{2\pi}{L} y\right) + b_{kl} \cos\left(\frac{2\pi}{L} x\right) \sin\left(\frac{2\pi}{L} y\right) \\
&+ c_{kl} \sin\left(\frac{2\pi}{L} x\right) \cos\left(\frac{2\pi}{L} y\right) + d_{kl} \sin\left(\frac{2\pi}{L} x\right) \sin\left(\frac{2\pi}{L} y\right)
\end{aligned}$$

In order get the results of the normalised solution, we calculate the pre-factor for two dimensions, by using again that $W(x,y) = W(x)W(y)$:

$$\begin{aligned}
A &= \frac{1}{(2^{2N-1})^2} \binom{2n}{n-k} \binom{2n}{n-l} \chi_k \chi_l \frac{(2^{2N-1})^2}{\left(\binom{2N}{N}\right)^2 \frac{1}{4} L^2} \\
&= 4 \chi_k \chi_l \frac{\binom{2n}{n-k} \binom{2n}{n-l}}{\binom{2N}{N}^2}
\end{aligned}$$

The final solution for the Fourier component are:

$$a_{kl} = A \cos\left(\frac{2\pi}{L} x_0\right) \cos\left(\frac{2\pi}{L} y_0\right) \quad (\text{A.15})$$

$$b_{kl} = A \cos\left(\frac{2\pi}{L} x_0\right) \sin\left(\frac{2\pi}{L} y_0\right) \quad (\text{A.16})$$

$$c_{kl} = A \sin\left(\frac{2\pi}{L} x_0\right) \cos\left(\frac{2\pi}{L} y_0\right) \quad (\text{A.17})$$

$$d_{kl} = A \sin\left(\frac{2\pi}{L} x_0\right) \sin\left(\frac{2\pi}{L} y_0\right) \quad (\text{A.18})$$

$$W = \sum_{k,l} A \cos\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) + A \cos\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right) \quad (\text{A.19})$$

$$+ A \sin\left(\frac{2\pi}{L} kx_0\right) \cos\left(\frac{2\pi}{L} ly_0\right) + A \sin\left(\frac{2\pi}{L} kx_0\right) \sin\left(\frac{2\pi}{L} ly_0\right) \quad (\text{A.20})$$

$$\text{with } A = \frac{1}{L} \frac{\binom{2n}{n-k} \binom{2n}{n-l}}{\binom{2n}{n}^2} \frac{4}{(1 + \delta_{k,0})(1 + \delta_{l,0})} \quad (\text{A.21})$$

A. Appendix

A.3.3. Implementation: Deriving the gradient velocity field from a given updraught-sink distribution

1. Create Fourier coefficients from sink distribution:

Input : N : the inverse peak width, L : the domain size, (x_i, y_i) : the locations of all flow sinks, extracted from the data, $s = S'/c$: a strength factor for each (x_i, y_i) tuple,

Processing : Calculating the Fourier expansion of the joint W of the given sink list. The Fourier components of each sink are weighted with s , a factor scaling the modelled sink strengths with basin area extracted from the data

Return: $[a, b, c, d]$: All four Fourier coefficients for the 2D Fourier expansion of W , $[x, y]$: the $L \times L$ grid where the coefficients are calculated

2. Further derive Fourier coefficients for the model properties of interest, based on $[a, b, c, d]$. Different functions calculate the coefficients for the potential (A), and the components of the velocity gradient field u and v (B):

Input : Each function (A,B) takes all four Fourier coefficient arrays of W as input

Processing : (A) Derive the coefficients for the potential Φ , using $\phi_{kl} = \frac{-q_{kl}}{(k^2+l^2)}$ (and $q_{00} \stackrel{!}{=} 0$)
(B) Derive the Fourier coefficients for u and v , to later obtain the gradient field, corresponding to the velocity vector $\vec{v} = \begin{pmatrix} u \\ v \end{pmatrix}$. We use the following trick: calculating the gradient of a sum of products of trigonometric functions means rearranging the coefficients and multiplying by $\frac{\Omega k}{(k^2+l^2)}$

Return: Returns: Fourier coefficients for Φ , u or v

3. Back transformation from Fourier space to real space, applicable for all properties W , Φ , u and v

Input : Fourier coefficients of Φ , u , v or Q

Processing : Multiply derived Fourier coefficients with corresponding term of the Fourier series, and add the different Fourier components to one solution

Return: Real space solution of the field, the potential, components of the velocity vector field as $L \times L$ array.

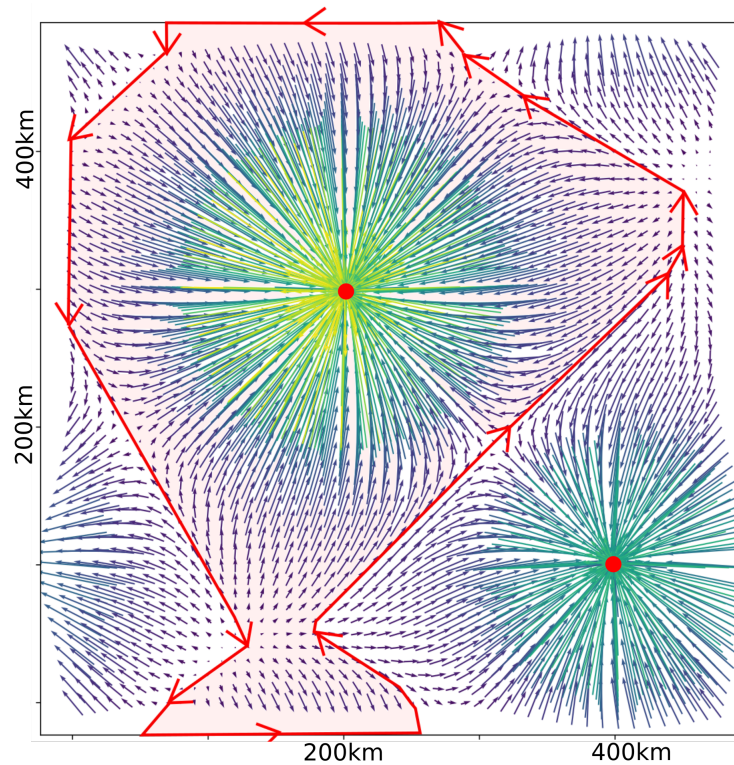


Figure A.2.: Illustration of the Gaussian theorem applied on a basins. Notice that the arrows are always parallel to the red line, so the orthogonal component vanishes and no flow crosses the basins' borders. The borders are areas of divergence. From there the flow increasingly converges towards the updraught centre.

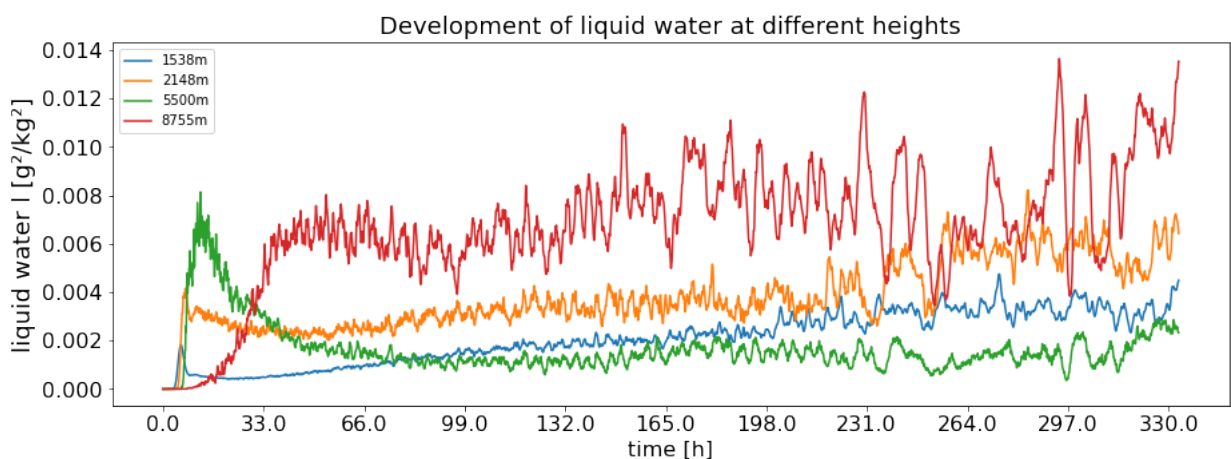


Figure A.3.: Temporal development of $\overline{l(t)}$ visualises how the vertical cloud distribution changes over the NoEvap simulation.

A. Appendix

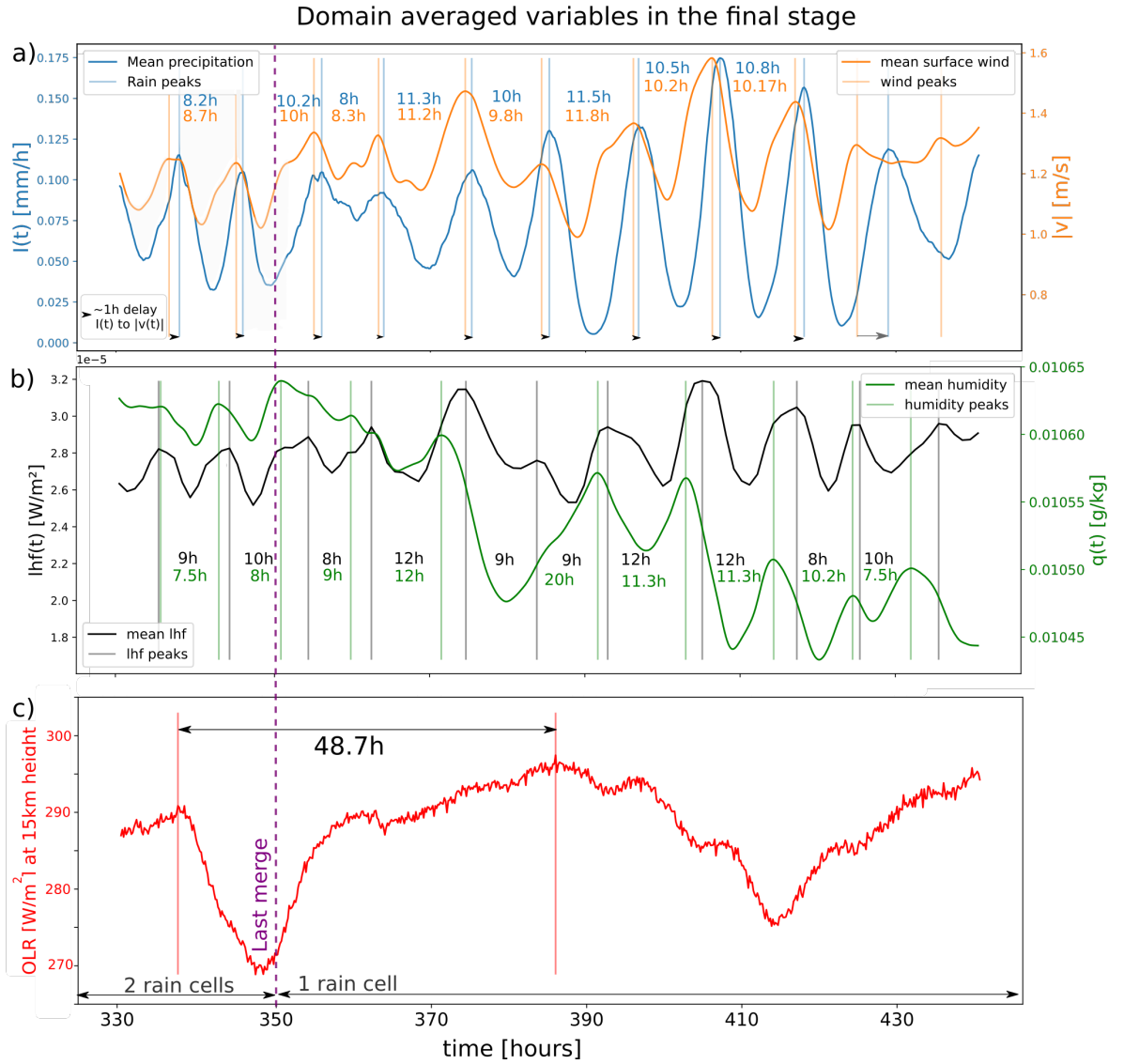


Figure A.4.: Domain averages of precipitation $\bar{I}(t)$ (a, blue), horizontal lowest-layer wind speed $|\bar{v}(t)|_{z \approx 50m}$ (a, orange), surface latent heat flux $\bar{l}hf(t)$ (b, black) and relative humidity $\bar{q}(t)|_{z \approx 50m}$ (b, green) and OLR at the top of the domain $\overline{OLR}(t)|_{z \approx 17500m}$ (c, red). After 330 h two convective cells are left (light grey shading), which then merge (purple line) to the final remaining one after 350 h (white background). All physical variables (a-c) show oscillations, with periods provided in the respective colour of their plot. Black arrows (a) visualise the delay time between maximum convergence $\bar{v}(t)|_{z \approx 50m}$ and maximum $\bar{I}(t)$.

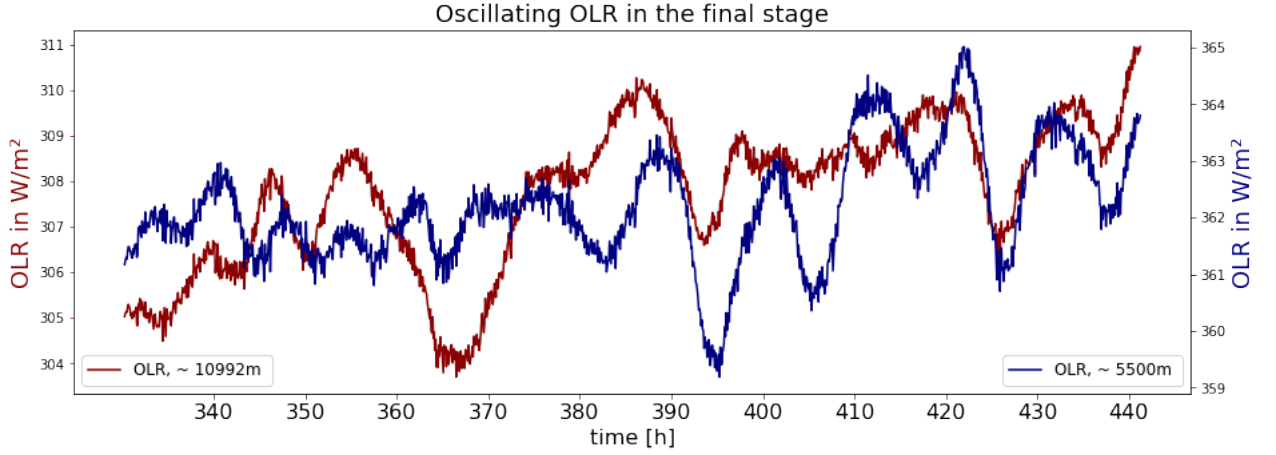


Figure A.5.: OLR oscillations in NoEvap simulations differ with heights.

A direct comparison showed, that oscillations between temperature and OLR at respective heights are quite in phase. For example $\overline{\text{OLR}(t)}_{z \approx 5500m}$ seems synchronised to $\overline{\Theta}_l(t)_{z \approx 5500m}$ (compare Fig. 8.3).

A.3.4. Proof: scaling of $A_{basin_i}^{model}$ with strength parameter S_i

A.4. Discussion supplements

A.4.1. Multiple coupled oscillators

In general N coupled oscillators can show very complex behaviour, which is often impossible to describe analytically. A simple case are (quasi)-periodic oscillations (Fig. A.4, Sec. 8.1.3) where different system components reach some kind of synchrony, which Strogatz describes as an order in time. But how and when coupled oscillators synchronise themselves is only studied for a limited number of simplified cases [Str04].

Assuming N coupled oscillators, dynamics become even more diverse and complex to describe or predict. An equal global coupling could provide a simplified framework, which could be associated to the interaction of N convective cells under the assumption that heat is equally dissipated over the entire domain (weak temperature gradient approximation, Sec. 4.1.6), so the spatial dependence could be neglected. Assuming that coupling would differ depending on the distance between convective cells demands to model a network of coupled oscillators (e.g [ROH05]) which raises the difficulty of finding an analytical representation².

In general, one well-studied type of oscillators are 'pulse-coupled' oscillators, whose interaction can be assumed to happen at discrete times, so their coupling can be

²Again, then a possible simplification is to focus on nearest neighbour interaction, Sec. 8.3

A. Appendix

described based on a delta function. Although we can not assume a perfect pulse interaction between convective cells, gravity waves travel on much faster than convective time-scales (Sec. 4.1.1), which might allow to estimate that the effect of an interaction dissipates before the next perturbation is received³. Mirollo and Strogatz studied the problem of N identical, pulse-coupled oscillators, where each oscillator is coupled to all others. Over time groups of oscillators synchronise, whereby a larger group 'absorbs' other oscillators until a single synchronised group remains, a picture which could be associated with 'competing' rain-cell dynamics observed in CSA (Fig. 8.4).

Suspecting an effect (as increasing temperature) would take finite time until convective dynamics are impacted (eg. convective activity is inhibited), would require to consider a coupling with time delayed feedback.

Schöll et al. studied a simple model of time delayed coupled neural oscillators [Sch+09] and found that synchronisation can be either enhanced or suppressed depending on the delay time. An additional application of time-delayed self-feedback caused complex scenarios, for example synchronised in-phase or anti-phase oscillations, bursting patterns or amplitude death, where amplitude variations are suppressed as a consequence of coupling. Developing our hypothesis further, we suggest that each convective cell (oscillator) would obtain such time-delayed self-feedback by warming the atmosphere. Therefore, it is possible that some of the oscillation and decaying phenomena observed in our NoEvap simulations could be interpreted as the complex interaction of inhibitory (self-)coupled oscillators and concluding a toy model should be sought to mimic such interaction of convective oscillators⁴.

Overall, despite the absence of CPs, NoEvap dynamics can be quite complex and simplifications will have to be made, to derive an analytical model.

³<http://www.scholarpedia.org/article/Pulse-coupled-oscillators>

⁴We are aware, that implementing such a model would be challenging, but encouragingly Heltberg et al. consider oscillations in the convective contexts to be much simpler to described than oscillations in biological processes, which are prevailing in the oscillator models discussed above [Hel+21]