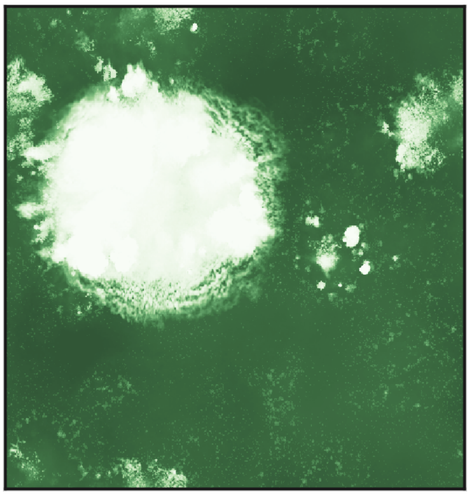
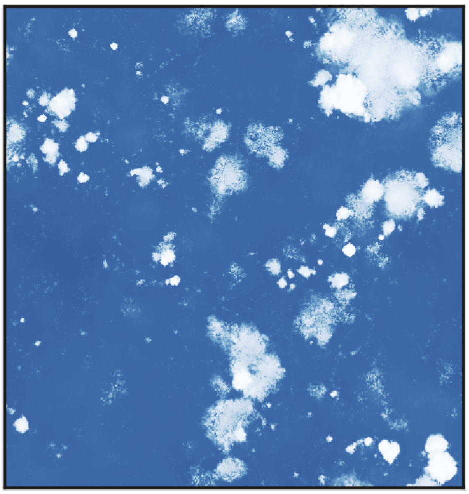
UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

PhD Thesis

Chasing the Storms

A Simulation and Observation-Based Exploration of Mesoscale Convective Systems and Cold Pools, from the Midlatitudes to the Tropics





Irene Livia Kruse

Supervised by Jan O. Haerter

March 2024

Irene Livia Kruse

Chasing the Storms

PhD Thesis, March 2024

Supervisor: Jan O. Haerter

This thesis has been submitted to the PhD School of:

The Faculty of Science, University of Copenhagen

To David, the original cumulonimbus-admirer

Acknowledgements

In the last 3 years and 3 months I've had the opportunity to belong to two institutes: the Niels Bohr Institute in Copenhagen and the Leibniz ZMT in Bremen. I'd like to thank the mentors that I met along the way in both places: my supervisor, Jan, for providing the resources to explore the realms of cloud research, from simulating storms on a supercomputer, to setting up a weather station in Senegal; the former Atmospheric Complexity team in Copenhagen (and dearest Bente!); the Complexity and Climate team in Bremen; and the strong scientific motivators scattered across the latitudes, in particular Adrian, Richard, Cathy, Conni, Kelly and Cristian. A big thank you to Bettina, for being my first mentor and lifelong friend since, Romain for supporting me throughout the PhD, and Leif for showing me the future.

Friends have been my sunny days in the stormy weather. Thanks in particular to Santi and Chrissy, for the lustige Zeiten in Deutschland. Gracias a Vicky por ser una compañera de piso increíble y enseñarme español. Thank you Ellie for being there from day 0 of this PhD which started in a tiny apartment in a locked-down Copenhagen; Miguelito for swimming with me in the randomness and the aperols; Marlies for the good laughs over the years; Dina for always being a text message away; Hauke and Nicola for our worldwide adventures; my beautiful cuz Sofie for all the pastries. Thank you Ilana and Evelyn, for being the ultimate cheerleaders and making me look forward to every new glass of wine á lá 'terrace under the sun in Crete'. A quick shout out to Lorelei and Rory Gilmore, for the timeless comfort.

Last but not least, I wish to thank my lovely mother Flavia and my lively father Peter for putting up with their sometimes stressed and hard working daughter (adapted from Kruse, 1997); grazie zia Manola e Ica per il tifo dall'Italia; tak til min farmor Inge for samtidig at være min største fan og største idol; and thank you Clara for joining me on this journey whilst showing me the silver lining to any cloud.

Abstract

Convective storms. We have most probably all experienced them. In Northern Europe, sporadically in the summer season. In the tropics, several times a week in the rainy season. The most organized storms highly contribute to extreme rainfall, both in Europe and in the tropics. The frequency of these very storms is increasing with a warmer climate, making the mechanistic understanding of the upscale growth of a storm, and the predictability of its onset essential today.

In this dissertation we explore the effect of a large diurnal cycle in surface temperatures, typical of tropical land, on organizing small convective clouds into large storms - mesoscale convective systems (MCSs) - and how these MCSs can in turn organize the surrounding moisture field into a breeding ground for even more organized storms, that can persist in less favorable environments, like over the ocean, hinting at a path to cyclone formation. We then focus on cold pools (CPs), a defining feature of precipitating convective clouds. CPs are both a key ingredient in the organization of these into larger systems, and an easy-to-interpret measure of storms from the ground. After learning key properties of real-world CPs by analyzing measurements over 10 years from a weather tower in Northern Europe, we turn to the practicality of portable automatic weather stations in regions of the world with less weather monitoring capacity. We set these up in two locations in Senegal, and develop an AI-based nowcasting tool with long short-term memory neural networks (LSTMs), trained to predict the advent of CPs in the city of Dakar.

With the combination of observation and simulations, we thus dive into the core of MCSs and their associated CPs, and turn our findings into an application useful to society in improving convective weather forecasting.

Resumé

Konvektive storme. Vi har sikkert alle sammen oplevet dem. I Nordeuropa sporadisk i sommerhalvåret. I troperne flere gange om ugen i regntiden. De mest organiserede storme bidrager i høj grad til ekstrem nedbør, både i Europa og i troperne. Hyppigheden af netop disse storme stiger med et varmere klima, hvilket gør den mekanistiske forståelse af en storms opskalering og forudsigeligheden af dens begyndelse vigtig i dag.

I denne afhandling undersøger vi effekten af en stor døgncyklus i overfladetemperaturer, som er typisk for tropisk land, på organiseringen af små konvektive skyer til store storme - mesoskala konvektive systemer (MCS'er) - og hvordan disse MCS'er igen kan organisere det omgivende fugtfelt til en grobund for endnu mere organiserede storme, der kan fortsætte i mindre gunstige miljøer, som over havet, hvilket antyder en vej til cyklon-dannelse. Derefter fokuserer vi på 'cold pools' (CP'er), som er et definerende træk ved udfældende konvektive skyer. CP'er er både en nøgleingrediens i organiseringen af disse i større systemer og et let fortolkeligt mål for storme fra jorden. Efter at have lært de vigtigste egenskaber ved virkelighedens CP'er ved at analysere målinger over 10 år fra et vejrtårn i Nordeuropa, vender vi os mod den praktiske anvendelighed af bærbare automatiske vejrstationer i regioner af verden med mindre vejrovervågningskapacitet. Vi sætter dem op to steder i Senegal og udvikler et nowcasting-værktøj med neurale netværk med lang korttidshukommelse (LSTM'er), der er trænet til at forudsige fremkomsten af CP'er i byen Dakar.

Med kombinationen af observationer og simuleringer dykker vi således ned i kernen af MCS'er og deres tilknyttede CP'er og omsætter vores resultater til en applikation, der er nyttig for samfundet og til at forbedre prognoser for konvektivt vejr.

Manuscripts

Manuscripts included in the chapters of this dissertation:

Kruse, Irene L., Romain Fiévet, and Jan O. Haerter (-). „Tipping to an Aggregated State by Mesoscale Convective Systems". Submitted to *Journal of Advances in Modeling Earth Systems.*

Kruse, Irene L., Jan O. Haerter, and Bettina Meyer (2022). „Cold pools over the Netherlands: A statistical study from tower and radar observations". In: *Quarterly Journal of the Royal Meteorological Society* 148.743, pp. 711–726.

Additional manuscripts I contributed to during my time as a phd student:

Vraciu, Cristian V., Irene L. Kruse, and Jan O. Haerter (2023). „The Role of Passive Cloud Volumes in the Transition From Shallow to Deep Atmospheric Convection". In: *Geophysical Research Letters* 50.23, e2023GL105996.

Nazari, Sara, Irene L. Kruse, and Nils Moosdorf (-). „Spatiotemporal Dynamics of Global Groundwater Recharge from 2001 to 2020". In review in *Journal of Hydrology*.

Contents

1 Rationale

This thesis encompasses three stories, which all have the same narrative thread: Convective storms.

It all starts with convection, a crucial process in meteorology that describes the vertical movement of heat and moisture within the Earth's atmosphere. In a physics laboratory, this phenomenon arises when you have a fluid trapped between two plates, with the lower plate being much warmer than the upper one. Initially, the fluid is still, but as it heats up, it starts to move in a pattern called Rayleigh-Bénard convection. In your kitchen, you'll see it in the form of rising of bubbles in a pot of water on a hot stove.

Figure 1.1: **The life cycle of a convective storm.** Solar radiation heats the surface; moist air rises in the form of a thermal; water vapor condenses forming a cloud and heating the atmosphere; rain forms and falls to the ground, evaporating, and extracting heat from the lower atmosphere; cold, dense air hits the ground and spreads out - forming a cold pool.

In the atmosphere, convection is driven by differential heating of the Earth's surface by the sun's radiation (Figure 1.1, starting from the left). When sunlight strikes the Earth's surface, it warms it up, heterogeneously depending on the type of surface, some which absorb and retain heat more effectively than other surfaces. As the surface heats up, it, in turn, heats up the air directly above it through a process called conduction. The warmer air near the ground becomes less dense than the cooler air surrounding it, causing it to rise in columns known as thermals. These thermals act as conduits for transporting heat and moisture from the surface into the atmosphere.

Now, we need to add an important physical fact here, which is necessary to understand the rest of this thesis. When water changes phase, energy is transferred. *When water vapor condenses, it releases energy (heat) to its surroundings. When liquid water evaporates, it takes energy (heat) from its surroundings.*

Condensation. Coming back to our thermals: when the warm air rises, it carries moisture with it, leading to the formation of clouds under the right atmospheric conditions. If the rising air contains sufficient moisture and continues to ascend, it cools as it reaches regions of lower atmospheric pressure. This cooling can cause the moisture in the air to *condense* into water droplets or ice crystals, releasing heat to its surroundings and further fueling the ascent. The condensed water droplets (which we know as clouds) are the visible manifestations of the convective process in the atmosphere and can take various forms, such as cumulus or cumulonimbus clouds. Cumulonimbus clouds, are associated with intense atmospheric phenomena such as thunderstorms, lightning, and heavy precipitation. This convective precipitation occurs when the convective clouds contain enough moisture to produce rain or other forms of precipitation.

Evaporation. When it rains, the falling rain can in turn *evaporate*, extracting heat from the air. This leads to the formation of a region of cold, dense air, that falls to the ground and spreads out, similarly to a density current. This cold air, called a convective **cold pool (CP)**, produces a strong wind gust spreading around the storm, potentially triggering new convection around it, and can linger for hours after the rain has stopped, as a footprint of the storm that prevents new convection from happening in its place.

This all seems straightforward, yet there is still a world of complexities to understand. Convective storms come in many sizes, starting from single isolated convective cells that are typically 1 to 10km in diameter (Figure 1.2, starting from the left). These isolated convective cells can organize into one large system, called a **mesoscale convective system (MCS)**, by definition larger than 100 km horizontally. If we add in the swirling Coriolis force, we end up with tropical cyclones (hurricanes, typhoons) which can be 1000 km in diameter. This upscale growth of convection is not yet fully understood, albeit years and years of research. Because these processes are not fully understood, their predictability is impacted. Weather forecasting models are often run at spatial resolutions that are too coarse to resolve convective processes, with the consequence that these processes need to be included as parameterizations learnt from higher resolution simulations. However, even with these parameterizations, the skill in predicting convective rain in specific locations is still poor, globally, due to the stochastic nature of convection. On top of that, the weather monitoring network in terms of radar coverage, is extremely unbalanced, with tropical countries experiencing the largest storms on Earth with the highest frequency often having limited weather monitoring capacity.

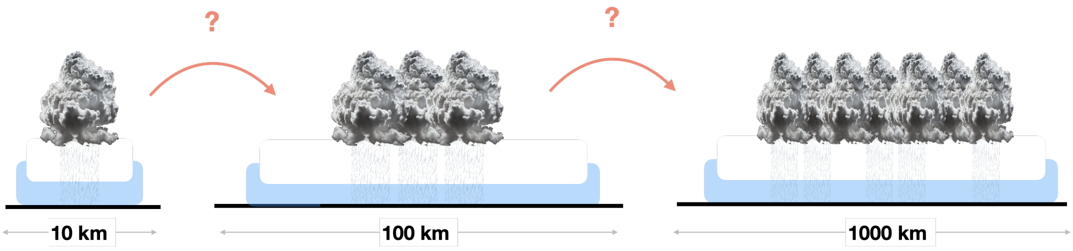


Figure 1.2: **Understanding the upscale growth of convection.** Spatial scales of: isolated deep convective cell; a mesoscale convective system; a tropical-cyclone like organized deep convective system. The upscale growth is yet to be fully understood.

In this dissertation we will explore the upscale growth of convection. Thanks to results from high resolution simulations based on observational data from the tropics, we will describe a theory on how convective storms organize into MCSs thanks to the diurnal cycle. We will define how these in turn affect the atmosphere they are embedded within, creating a state that favors clustering of convection, due to the interplay of processes mentioned above. We will present a conceptual model which encompasses the convective processes leading up to the formation of MCSs and the self-aggregated state. We will then dive into analysing CPs, a handy way to measure convective storms from the ground. We will find the typical properties of CPs thanks to 10 years of measurements from a unique 200-meter weather tower in Cabauw, the Netherlands. Thanks to this characterization of CPs from time series, we will turn our focus back to the tropics, to the West African country of Senegal. Here, where MCSs traverse the country regularly in the rainy season, we will develop a machine learning-based tool to predict the onset of CPs in the capital city of Dakar, that can be used as a warning mechanism before it starts raining (Figure 1.3).

The structure of the thesis is as follows. To set the stage, Chapter 2 will introduce the Groundwork that has laid the foundations to this thesis. We will get acquainted with MCSs and CPs, through the lense of historical observations; convective self-aggregation, a puzzling (but enlightening) phenomenon arising in simulations; and the concept of nowcasting including how to use AI for this purpose. After presenting the work that has preceded this dissertation, in the three novel studies contained here we will: (Chapter 3) Explore how MCSs developed over tropical land strongly contribute to convective self-aggregation; (Chapter 4) Characterize convective CPs in the midlatitudes from a weather tower; (Chapter 5) Describe the first pillars of a field campaign to monitor MCSs in West Africa with automatic weather stations, and explore how the new observational data can be used, with the help of machine learning, to nowcast CPs.

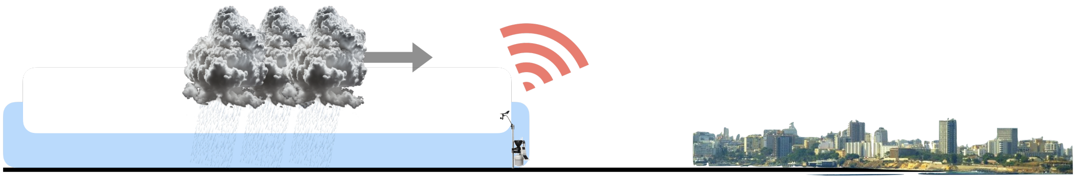


Figure 1.3: **Predicting convective storms ahead of time.** In the culmination of this thesis, we will explore a way to predict the arrival of a convective storm from its associated cold pool in the absence of radar, using automatic weather stations situated upstream of the location of interest (here Dakar, Senegal).

2 The Groundwork

2.1 Mesoscale Convective Systems (MCSs) and Cold Pools (CPs)

–Through the Lense of Historical Observations–

Excerpt from Nature, 1896:

> Symons's Monthly Meteorological Magazine, July.
> "The International Cloud Atlas." – Mr. Symons takes the opportunity provided by the publication of this work, of which only a very few copies have yet been distributed, to briefly mention the principal works on clouds that have recently preceded the present one. These include L. Weilbach's "Nordeuropas Skyformer" (Copenhagen, 1881), the "Wolken-Atlas" by H. H. Hildebrandsson, Koppen, and Neumayer (Hamburg, 1890), M. Singer's "Wolkentafeln" (Munich, 1892), the "Classificazione delle nubi" by the Specola Vaticana, which contains some excellent reproductions of M. Mannucci's photographs (Rome, 1893), and the Rev. W. Clement Ley's "Cloud Land" (London, 1894). The "International Cloud Atlas" (Paris, 1896) has been prepared under the supervision of the International Meteorological Committee and includes twenty-eight colored reproductions of clouds. Although none of them is from an English photograph, Mr. Symons believes that our countrymen may be satisfied to see how much the international system of 1896 is based on the work of Luke Howard and that the classification adopted is essentially that of the collaborative effort of Dr. Hildebrandsson and the Hon. Ralph Abercromby.

As I was digging into archives about definitions of Mesoscale Convective Systems, I found the *International Cloud Atlas* (1896), the first published photographic collection of clouds on an international level. The classification by which the clouds are named, follows that proposed by the Scottish meteorologist Ralph Abercromby and the Swedish meteorologist Hugo Hildebrand Hildebrandsson, and the publication of the first edition was arranged by Hildebrandsson himself and the members of the

Clouds Commission of the International Meteorological Organization (now: the World Meteorological Organization). The Atlas consists of color plates of clouds, with descriptive text in English, French, and German. The collaborative effort extended mere visual representation, serving as a resource for meteorological education and promoting uniform terminology in cloud description - marking a significant step forward in weather forecasting and fostering clearer communication and consistency among meteorologists. As you might have read above, several nation-bound works preceded this Cloud Atlas, however the truly revolutionary novelty behind this work, in my eyes, is the realization that clouds exist in the same general forms everywhere in the world.

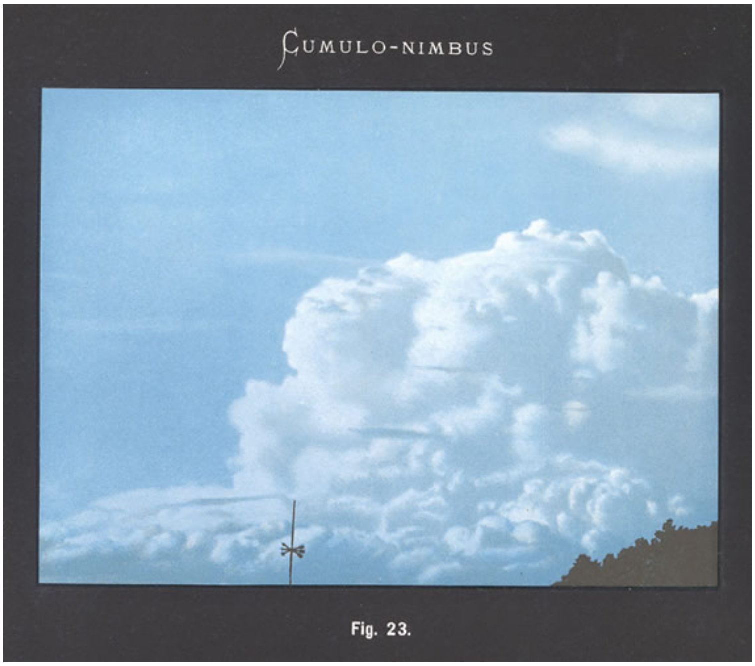


Figure 2.1: Cumulonimbus cloud, as depicted in the *International Cloud Atlas* (1896)

This entire dissertation will focus on the cloud first classified in the International Cloud Atlas as the cumulonimbus (Figure 2.1). In layman terms: A stormy cloud. These clouds have a dense, cauliflower-like appearance and can take the form of thunderstorms, with heavy rain, lightning, and hail. Cumulonimbus clouds form through the process of atmospheric convection, where warm, moist air rises rapidly, cools, and condenses into a towering cloud structure. The World Meteorological Organization today defines the cumulonimbus cloud in the following way: "A heavy and dense cloud, with a considerable vertical extent, in the form of a mountain or huge towers. At least part of its upper portion is usually smooth, or fibrous or striated, and nearly always flattened; this part often spreads out in the shape of an anvil or vast plume". Hildebrandsson in his times, described the upper part of the cumulonimbus as “towering up to colossal proportions as mountain ranges, or a gigantic mushroom, with a flat layer of 'false cirrus' around or on the top." However, at the end of the 1800s, without airplanes, radar or satellites, it was nearly impossible to fully recognize the horizontal dimensions of these storms.

Mesoscale Convective Systems. It took a World War and two Royal Air Force officers stationed in Nigeria in 1945, to figure out the dimensions that convective storms could actually take on, albeit within a questionable frame. Combining ground-based station data, soundings and aircraft flights in tropical Western Africa, it was established that storms could be more than 100 km wide and more than 10 km tall (Hamilton *et al.*, 1945). A few years later, the "Thunderstorm Project" was carried out in postwar US (Byers *et al.*, 1949). Led by Byers and Braham, from Chicago University, this was one of the largest meteorological field campaigns in history, using aircrafts, radiosondes and radars left over from the war, and solely focused on understanding thunderstorms. The pilots flew inside storms over Ohio and Florida over the summers of 1946 and 1947, gathering data. Notably, it became clear that thunderstorms in a mature phase comprised both of updrafts and downdrafts, as shown in Figure 2.2, and typical vertical dimensions of more than 10 km were quantified.

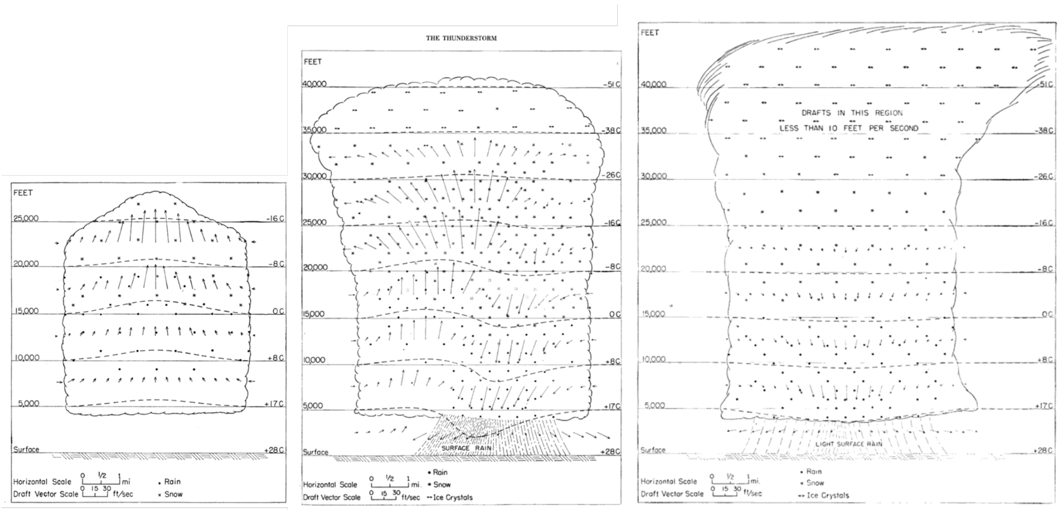


Figure 2.2: Conceptual drawing of a thunderstorm cell in its developing, mature and dissipating phases. From the Thunderstorm Project (1949) (Byers *et al.*, 1949). Note: The lower and upper boundaries of the mature cloud, 5000 and 40000 feet, correspond to approximately 1.5 and 12 km.

Using time series from US ground-based meteorological stations in 1955, Fujita was able to quantify the dimensions of storms without needing to fly into them with airplanes (Fujita, 1955). He was the first in the field to establish a very simple space-time conversion, to go from point-measured time series to the spatial dimensions of a storm, knowing the relative velocity of the storm with respect to the station (this could be extrapolated when there are two stations separated in space, measuring the storm at different times). With this approach he was able to estimate that storms could reach even 300 km in horizontal dimensions. This dimensional scale, larger than an isolated cloud, and smaller than synoptic weather systems, takes on the term meso-scale, and alludes to the fact that storm systems are often organized structures comprising of more than one cloud within. Referring to the mesoscale, Fujita coined the field of study of "Meso-meteorology".

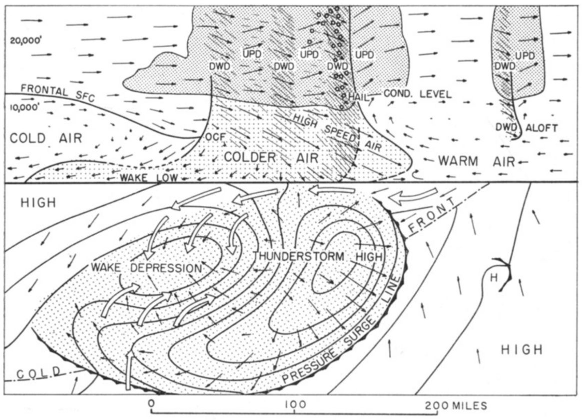


Figure 2.3: Conceptual drawing of a thunderstorm organized on the mesoscale. From Fujita (1955). Note: the horizontal extent of the storm, 200 miles, corresponds to approximately 320 km.

The next leap in storm research came with the advent of radar technology. The bridge between radar and weather research was built rather accidentally: Radar echoes were used for spotting ships and aircrafts during WWII, and echoes of rain, snow, and hail were seen as noise that needed to be filtered out. Only after the war, did military scientists in civilian costume start focusing on that "noise" for weather forecasting, finding that radar echoes can quantify the size and shape of the raindrops they encounter. With radar it became possible to spatially characterize the precipitation distribution in a passing storm (as in Figure 2.4, which distinguished between two types of precipitation in MCSs: convective and stratiform), and to quantify the precipitation intensity in space. With more and more mesoscale observations of storms, Mesoscale Convective Systems were coined in literature, and the general textbook definition by Houze (1993), that will be the definition used in this dissertation, is the following:

> *Mesoscale Convective System* (MCS): "A cumulonimbus cloud system that produces a contiguous precipitation area of ~ 100 km or more in at least one direction."

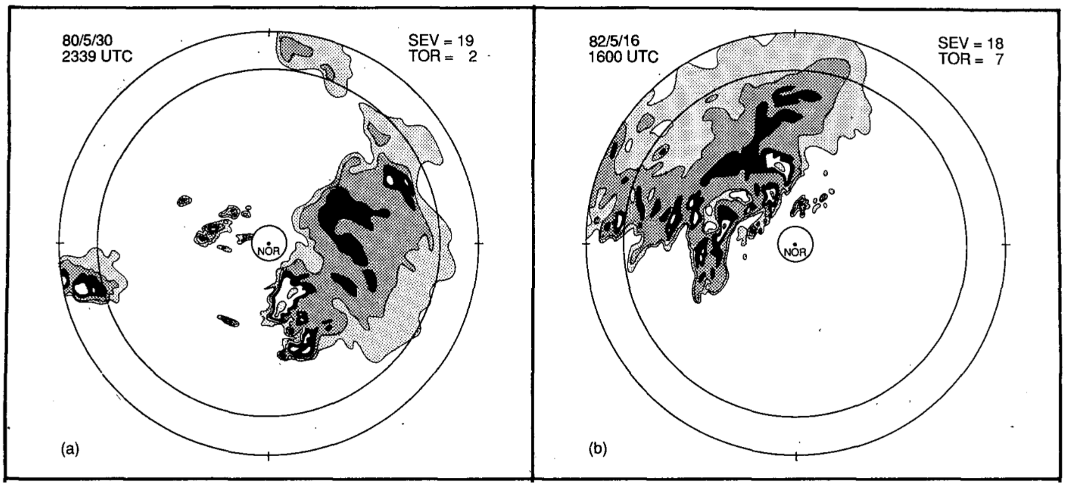


Figure 2.4: Example of radar echo of rainfall in two MCSs (top-down view) that lead to severe weather reports in 1980s Oklahoma. Darker shades represent stronger echoes (convective rain), and lighter shades represent weaker echoes (stratiform rain). Outer circle denotes 240 km range from radar. From Houze *et al.* (1990).

Radar-based studies were and are in nature regionally-based to this day, since radars are not present everywhere on the globe (and especially not over the oceans) and the coverage of a single radar tower is only within the 100s of kilometers. In fact, there is barely any radar coverage on the African continent. On the other hand, Europe is thoroughly covered, and we will use radar data from the Netherlands in Chapter 4 of this dissertation to characterize convective rainfall.

An enormous field campaign that combined all the above-mentioned ways to study an MCS, was the Global Atmospheric Research Program Atlantic Tropical Experiment (GATE), carried out in 1974 over the Eastern tropical Atlantic Ocean, off the coast of West Africa. Interestingly, the campaign was inspired by the still-to-be-proven hypothesis based on Lorenz's work on weather predictability (Lorenz, 1963), that a better understanding of tropical convection could lead to the prediction of global weather up to two weeks in advance. GATE included reasearch aircrafts and ships, equipped with radar, soundings and ground-based stations. This gave a means of comparison between MCSs measured in the continental US in the preceding radar studies, and MCSs measured over the tropical Atlantic. The common factor to both the midlatitude and tropical MCSs was the upscale growth of a region of deep convection so that the rain covered an area of mesoscale extent, and the existence of a mesoscale region of stratiform precipitation occurring in association with the deep convective rain (Houze, 2018). Thanks to measurements from GATE, the main life stages of an MCS were defined, as depicted in Houze's schematic (Figure 2.5) (Houze, 1982): the early stage consists of isolated precipitating convective towers, which merge into one large unified entity containing deep convective cells and stratiform cloud in the mature phase. The overall precipitation then weakens to being only stratiform before it totally disappears in the dissipating stage when the clouds break apart.

A significant finding in this conceptual study is that MCSs have the key property of substantially heating the upper troposphere through condensational heating and radiative effects, to the point that this can alter larger scale circulations. Quantifying these larger scale circulations due to MCSs remains a topic of interest.

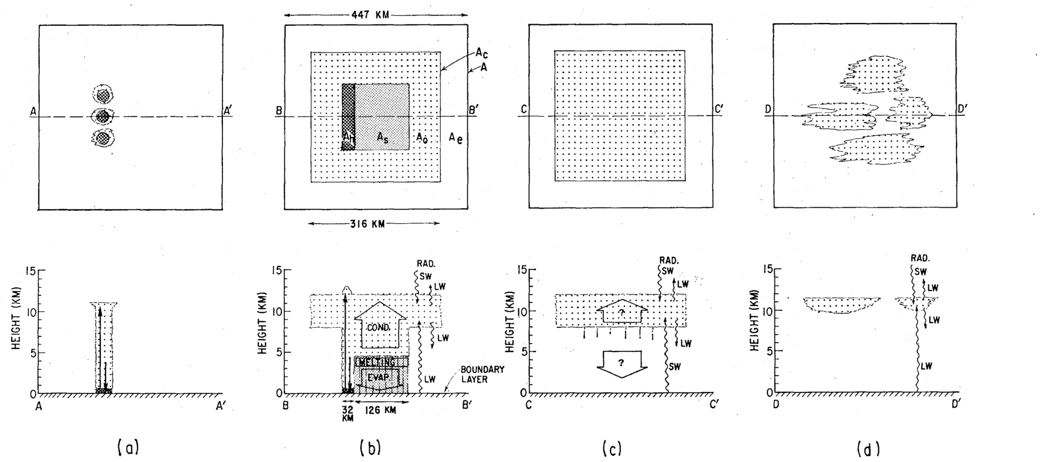


Figure 2.5: Schematic of an MCS in four stages of development, from Houze (1982). (a) Early stage; (b) Mature stage; (c) Weakening stage; (d) Dissipating stage. Top row shows the top-down view, and bottom row shows the side section.

To study MCSs on a global level, and to observe the full evolution throughout the lifetime of an MCSs, one more technological development was necessary: The launch of weather satellites from the 1960s and onwards. Satellites equipped with infrared imagers were particularly useful for storm research, as they could detect storms well beyond national scales, during the day and the night. MCSs have a typical footprint in the infrared retrievals: since they reach 10+ kilometer heights in the atmosphere, they result as very cold, large areas. Satellite-borne radar, in addition to infrared retrievals, has allowed the characterization of the precipitation within the MCSs, on a global scale since the 1990s. In recent times, thanks to 30 years of measurements, we are seeing the first MCS climatologies. In fact, with a threshold on infrared temperature and area, or on radar rainfall and area, MCSs can be detected - and with higher and higher temporal and spatial resolutions on geostationary satellites, they can be tracked and characterized globally, with methods that are still being improved to this day.

To track tropical MCSs in Chapter 3 of this dissertation we will use the "Tracking Of Organized Convection Algorithm through a 3-D segmentatioN" (TOOCAN) (Fiolleau and Roca, 2013) applied to infrared satellite retrievals. This innovative tracking algorithm from 2013 stands out, as it uses the definition of an MCS throughout its lifetime, as depicted in Figure 2.6, to construct a cloud volume in space and time around any detected deep convective core. It is quite fascinating to compare Figure 2.6 with the 1982 MCS schematic by Houze (Figure 2.5), showing how the conceptual understanding of MCSs from the 80s is still very much the groundwork of MCS research of the last decade.

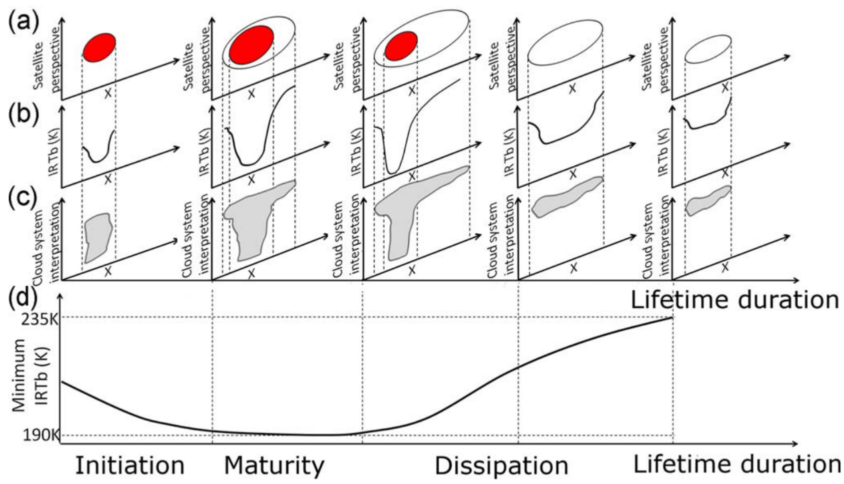


Figure 2.6: Schematic of an MCS for successive time steps and from satellite perspectives, as a base for TOOCAN algorithm. From Fiolleau and Roca (2013). (a) Red part corresponds to the convective core, black line represents the high cold cloud shield boundaries. (b) Minimum brightness temperature in an X cross section. (c) Associated convective system. (d) Minimum brightness temperature.

One of the most recent comprehensive MCS tracking studies is Feng et al. (2020), which quantifies for the first time the statistics of MCSs globally, along with the precipitation attributable to MCSs (Figure 2.7) and the average movement of MCSs across the globe (Figure 2.8). It is worth highlighting in these Figures, that tropical Africa stands out as a very large land-region that has a particularly high number of MCSs annually, which contribute to more than 70% of the total annual rainfall, and have a general westward motion. When focusing on extreme convective rainfall, there is a strong preference for extreme events to be located over land, as Zipser *et al.* (2006) found in an earlier satellite-based study focused on the tropics. They also found that there is a clear tendency for the most intense storms over oceans to be adjacent to land, in locations favoring storm motion from land to ocean, such as the tropical ocean west of west Africa (Zipser *et al.*, 2006).

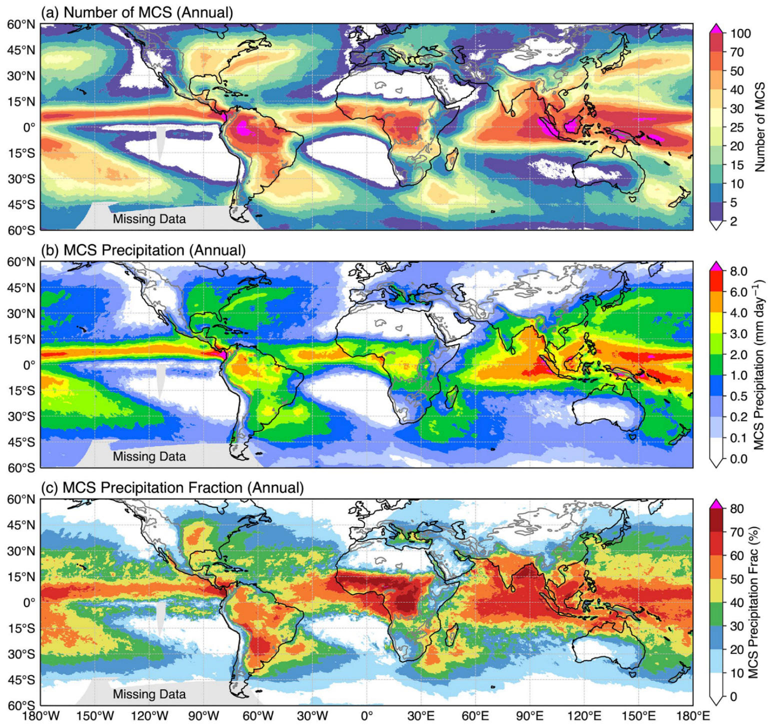


Figure 2.7: Annual mean global distribution of (a) number of MCSs, (b) MCS precipitation amount, and (c) percentage of MCS precipitation to total precipitation between 2001 and 2019. From Feng *et al.* (2021).

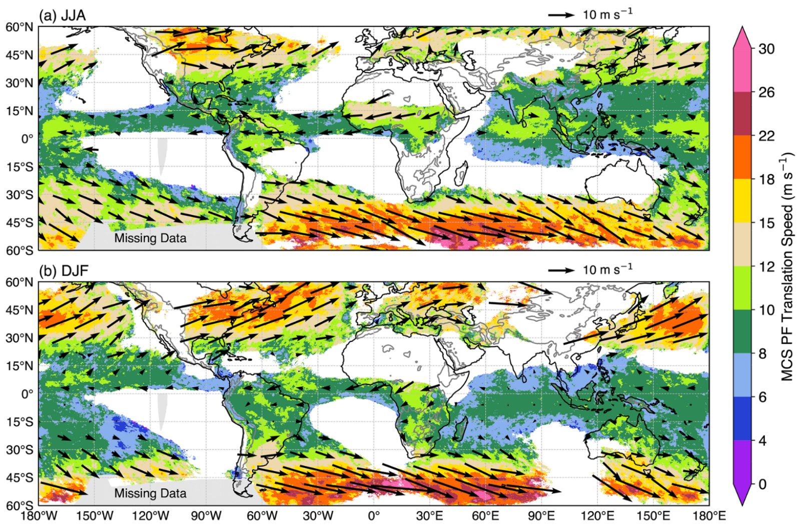


Figure 2.8: MCS translation speeds and directions for (a) June-July-August and (b) December-January-February, from satellite-tracked MCSs between 2001 and 2019. From Feng *et al.* (2021).

Cold pools. While the focus so far has been on the cumulonimbus cloud, and the organization of multiple convective clouds into an entity much larger than the single cloud, I will now lower the perspective to what happens below a precipitating convective cloud. When it rains into a subsaturated lower troposphere, the falling raindrops can evaporate, cooling the surrounding air. This cold, dense air, falls to the ground and spreads out around the precipitating cloud.

In the early conceptual drawings of thunderstorms, the region below the cloud was merely hinted at. In the conceptual drawing of a mature MCS from the 1946-1947 Thunderstorm Project (back to Figure 2.2), small diverging arrows are drawn below the thunderstorm cloud. Digging deeper into the Thunderstorm Project archives, I found this to be the first time that, thanks to the analysis of surface measurements during storms, thunderstorm downdrafts at surface level were quantitatively described in literature. The observations showed that a thunderstorm passing a ground-based station coincided with a sudden drop in temperature, and subsequent rain (see Figure 2.9). The cold air associated with the term "thunderstorm downdraft" had many interesting features already mentioned in the final document of the project, such as having a dry interior, and having the capacity to produce new convective updrafts on their leading edge.

In Fujita's drawing from 1955 (Figure 2.3), the downdraft receives its own outline and the words 'COLDER AIR' and 'HIGH SPEED AIR' appear below the thunderstorm cloud. The PhD work by Charba (1974) contained the first detailed analysis of an MCS downdraft from a network of ground-based weather stations spread over Oklahoma, combined with a 444 meter television tower equipped with meteorological instruments along its height. Many similarities were found between the dynamics of the gust front and those of a gravity current. This work was succeeded by Goff (1976) with an analysis of 20 gust fronts over Oklahoma from the TV-tower measurements. It became clear that the sequence of meteorological events measured during the passage of a thunderstorm downdraft were the following: a rise in pressure; a shift in wind direction; a sudden increase in wind speed; a drop in temperature; and finally rainfall.

Downdrafts are generally invisible to the eye, unless located in a desert-like environment, in which the outflow air forms a haboob, a literal moving wall of sand. In other environments, the border of a downdraft, or the "gust front", can be seen sometimes in radar echoes due to water particles, dust, and insects brought aloft in the convergence zone between the outflowing cold air, and the surrounding

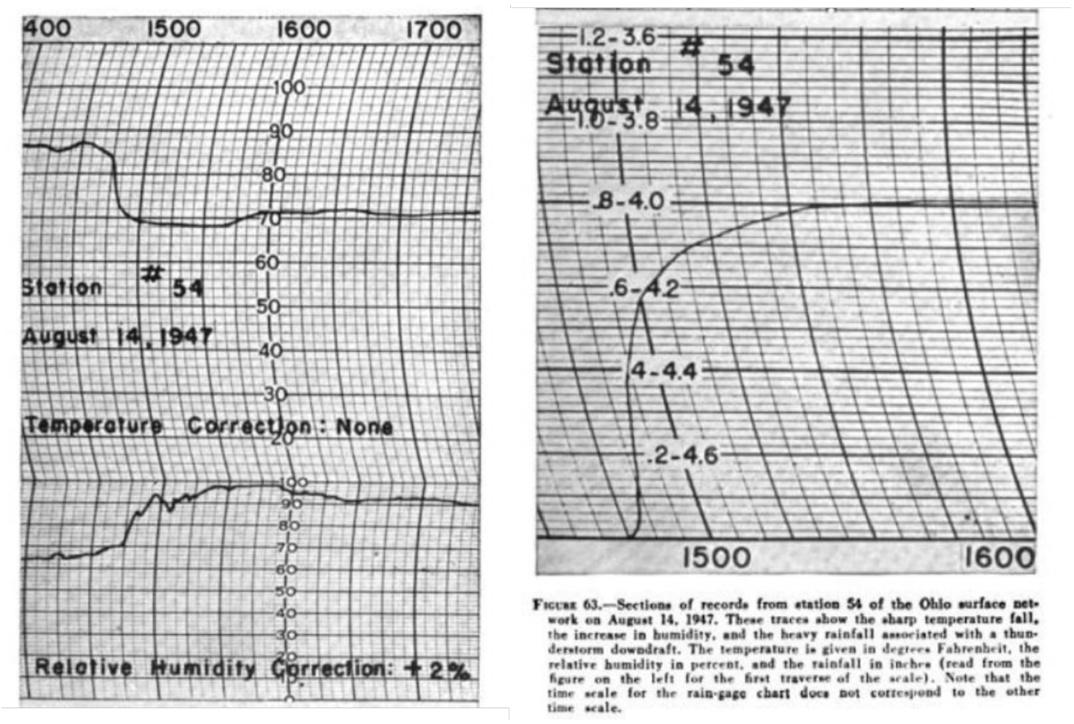


Figure 2.9: Time series of (upper left) measured temperature, (lower left) relative humidity and (right) precipitation at ground level, in correspondence with a passing thunderstorm. From the Thunderstorm Project (Byers *et al.*, 1949). Note: The temperature drops approximately from 86 to 68 F (30 to 20 °C), the relative humidity increases, but not to saturation, and precipitation is recorded in the 30 minutes succeeding the temperature drop.

warmer air. Thanks to this feature, Wakimoto, in the 1980s was able to study the internal velocities of gust fronts, resulting in a conceptual diagram of the stages of a gust front throughout time (Figure 2.10) (Wakimoto, 1982). In this work, it was also shown that the equation governing the propagation speed of a density current can serve to predict the movement of a gust front.

From the 2000's, the term cold pool (CP) started being used in literature, as a replacement for thunderstorm downdraft, convective outflow and gust front. In fact, in December 2022, about 30 atmospheric scientists from around the world sat in a castle in the mountains of southern Germany, namely Schloss Ringberg, discussing the weather-dictionary definition of the term "(convective) cold pool", to officially coin the downdrafts measured 75 years earlier in the Thunderstorm Project.

One of the most important effects of CPs is, as already suggested by the early studies, that they act as density currents and trigger new convection at their leading edges, an effect that is accentuated by the collision with one or more CPs. A beautiful

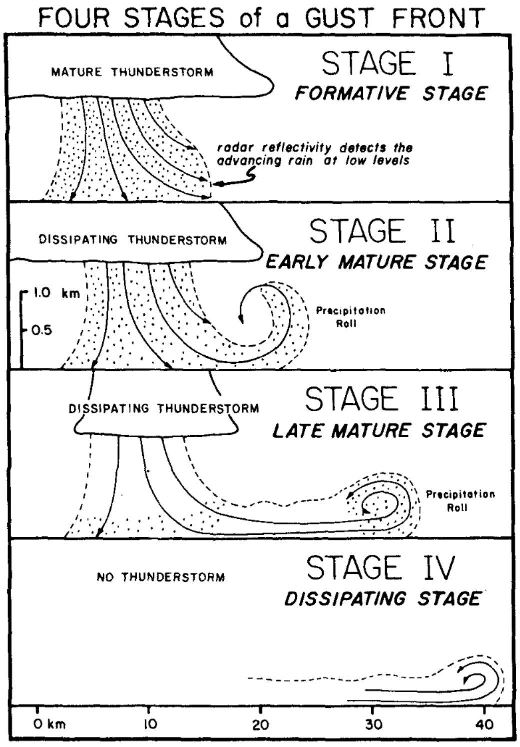


Figure 2.10: The four stages of a thunderstorm gust front. From Wakimoto (1982).

depiction and analysis of this effect is shown in Figure 2.11 from Meyer and Haerter (2020).

As Houze (2018) puts it, "cold pools can be thought of as a medium of communication between existing and future convection". CPs regained attention in the last decade due to their potential involvement in facilitating the transition from shallow to deep convection (suggested by e.g. Rowe and Houze Jr. (2015)), for organizing deep convective clouds into MCSs and for capturing the correct diurnal cycle of convection in models (Schlemmer and Hohenegger, 2014).

However, while the origin of CPs across the globe is quite similar in terms of general characteristics, i.e. in the form of precipitating convection, the surface and boundary layer with which a CP interacts varies greatly based on the location. The surface fluxes determine a CP's lifetime, and these can be very different over continental land, tropical ocean, or coastal midlatitudes. Moreover, the properties of the surrounding boundary layer, which can vary greatly based on the location, will determine the strength of the CP and its ability to create new convection. The observational CP studies mentioned until now were case studies of thunderstorms primarily over continental United States. A first statistical study of more than

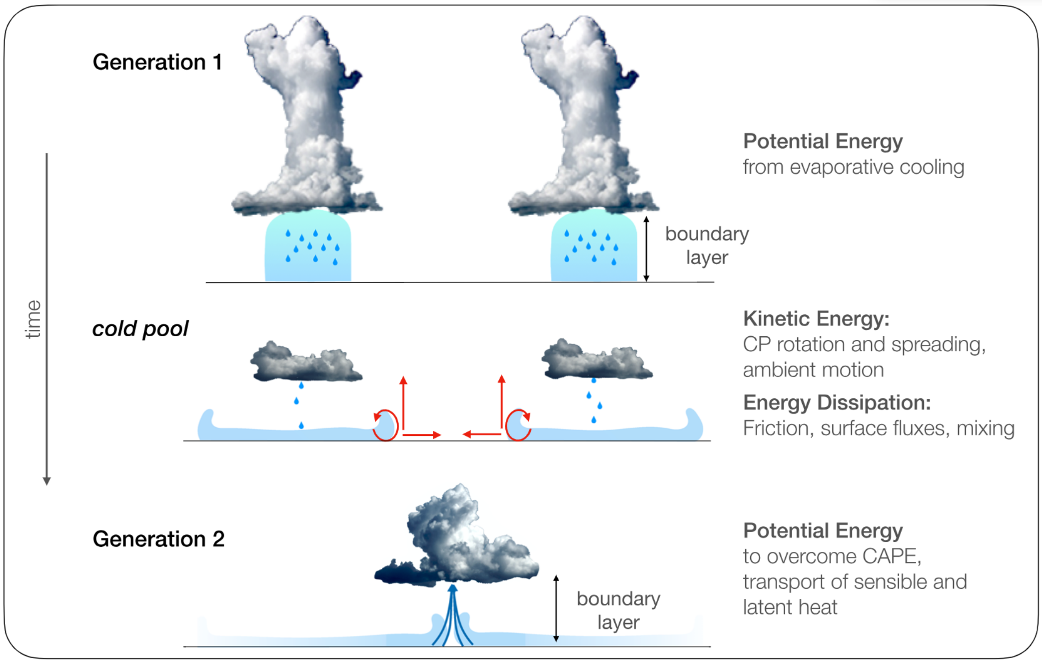


Figure 2.11: Simplified schematic of CPs linking convective events over space and time by converting the initial potential energy (Generation 1) from evaporatively cooled air into kinetic energy of a CP, which by triggering convection transfers it back to potential energy (Generation 2). From Meyer and Haerter, 2020.

200 CPs from observational data, over the central Indian Ocean, was presented in Szoeke *et al.* (2017), using the station data from the 2011 ship-based campaign DYNAMO. Composite time series of oceanic CPs resulting from this work can be seen in Figure 2.12. Inspired by this way of analyzing CPs, and by the lack of CP studies from coastal Europe, we conducted a statistical study of midlatitude (coastal) CPs from observational data, presented in Chapter 4, with the use of a 213 meter boundary layer measurement tower and collocated radar data.

The first field campaign solely focused on studying CPs over land was recently carried out in Germany with a spatially dense station network (FESSTVAL) (Hohenegger *et al.*, 2023). With the data from this campaign, we are seeing the first analyses of CPs observed throughout their lifetimes, with spatially and temporally high resolution measurements (Kirsch *et al.*, 2024). However, CPs have yet to be characterized in a global sense. Progress is underway in terms of using satellite retrievals innovatively to measure the shapes and sizes of CPs across the world. Scatterometer retrievals have been used to analyze patches of high surface roughness over the ocean, corresponding to CP gust fronts. The diurnal cycle of CPs and their typical area across the tropical oceans have been quantified this way (Garg

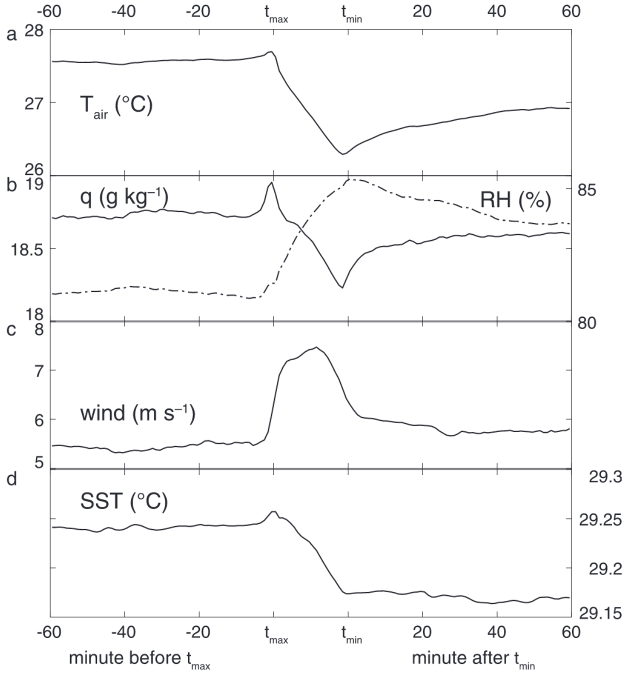


Figure 2.12: Time series analysis of 200+ cold pools measured over the Indian Ocean. Mean (a) 10-m air temperature, (b) specific and relative humidity, (c) wind speed, and (d) sea surface temperature composited on time elapsed from the start and end of the cold pool front. From Szoeke *et al.* (2017).

et al., 2021), however this technique can only be applied over the ocean. Thanks to AI algorithms, such as the one proposed by Hoeller, Fiévet, *et al.* (2024) which could be applied to geostationary infrared satellite imagery, we can expect to see global statistics of CPs, along the lines of the global studies of MCSs (like the one shown in Figure 2.8), coming out in the next few years, along with further enlightenment regarding the way CPs act in organizing convection across the globe.

Throughout this dissertation, CPs and MCSs will have a very strong association. The presence of an MCS, the organized form of deep convection, implies the presence of CPs covering areas proportional to the mesoscale size of the MCS. The presence of a CP does not necessarily imply the presence of an MCS, however in the right conditions, the CP is the essential ingredient for a deep convective cloud to become an MCS (as presented in Chapter 3). The presence of precipitating deep convection implies the presence of a CP, and the precipitation can be characterized by the measured CP itself (as presented in Chapter 4). Finally, the CP can be seen as a precursor to the precipitation of an MCS, a useful feature for short-term forecasting (as presented in Chapter 5).

2.2 Convective Self-Aggregation (CSA):

–A Puzzling Phenomenon Arising in Simulations–

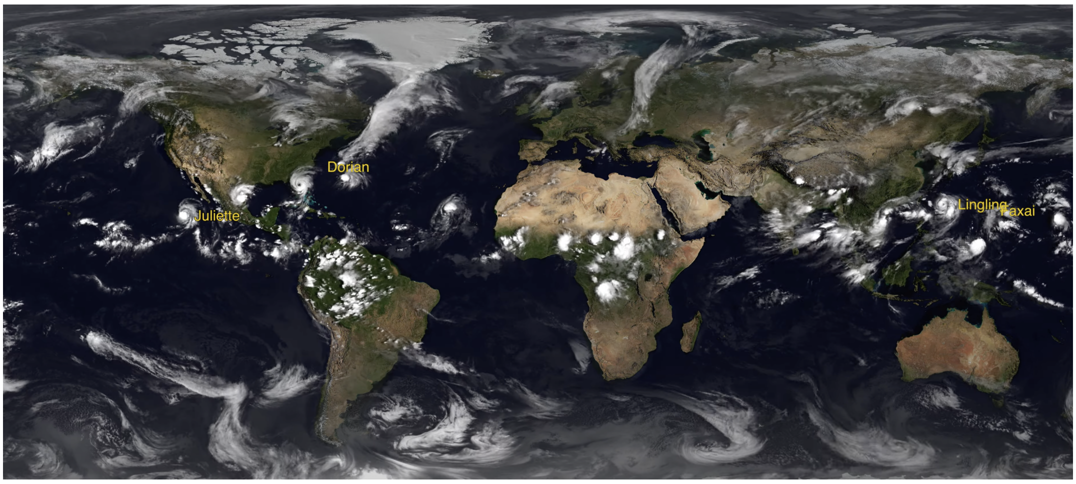


Figure 2.13: Snapshot of a September day, from "A Year of Weather 2019", a video by EUMETSAT. Composition of satellite infrared data layer, provided by Météo-France, superimposed over NASA's 'Blue Marble Next Generation' ground maps. Named hurricanes are labelled.

Convective clouds display a wide array of spatial arrangements, spanning from randomly distributed convective cells on the order of 10 km horizontally, to organized structures like MCSs on the order of 100 km horizontally, to cyclones on the order of 1000 km horizontally, and even planetary-scale cloud envelopes. In Figure 2.13, the global composition of infrared satellite retrievals from a day in September 2019, highlight precisely these different forms that convection can take on. If we focus on the African continent, it is daytime and there is a variety of small to enormous convective cloud clusters over the tropical African land belt, the larger ones being strongly organized MCSs. If we bring our focus westward across the Atlantic Ocean, we notice a cyclone formation hitting the East coast of the United States, namely hurricane Dorian. The fascinating question of what determines the upscale organization of deep convective cells to MCSs, and potentially on to cyclones, is, to this day, not fully resolved.

In meteorology, the search for an answer to this question, is driven by the evidence that the organization of convection significantly influences the prediction of severe weather events, with more organized events leading to more severe weather (as already hinted to with the analysis on the degree of MCS organization from radar imagery and the severity of weather reports in Houze *et al.* (1990)). In climatology, the grand challenge today is to understand how a change in convective organization might make a difference to the overall radiative budget, contributing to warming or cooling the climate.

But what is the intrinsic difference between a single deep convective cloud and a large organized system of deep convective clouds? Houze (1982), with simple calculations of condensation and evaporation rates within cloud clusters of different precipitating areas, revealed a fundamental distinction in the vertical distribution of diabatic heating between isolated convective towers and mature convective cloud clusters, the latter heating more aloft (see idealized profiles in Figure 2.14).

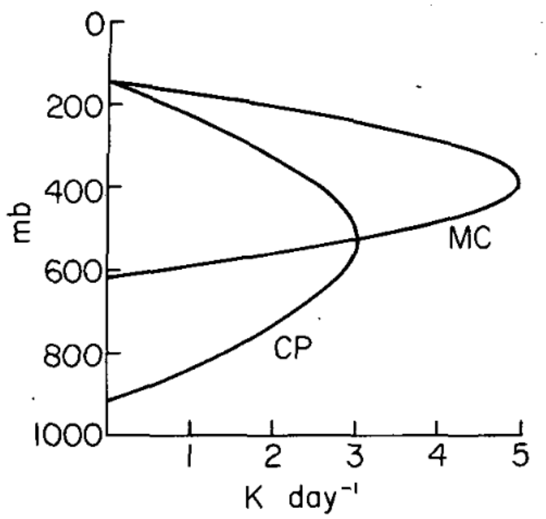


Figure 2.14: Vertical profiles of idealized heating distributions for one convective cloud (here CP for convective plume - not to be confused with cold pool!) and a mature cloud cluster (here MC). From Hartmann *et al.* (1984).

The characteristic top-heavy heating profile of larger convective clusters arises from the presence of the extensive cloud deck connecting the active cumulonimbus cells within a cloud cluster, which reinforces the heating of convective towers at higher altitudes and counteracts heating at lower levels. This difference in heating profiles in turn affects the large scale circulation around the cloud.

It is known that large organized convective clusters contribute to approximately half of the total tropical precipitation (Mapes and Houze, 1993), playing a crucial role in modulating the moisture distribution and hydrological cycle. Observational data suggests that the occurrence of organized convection has risen throughout the tropics over the last decades, and a majority of the regional increases in tropical precipitation are linked to this increase (Tan *et al.*, 2015). Moreover, apart from their impact on tropical cloud cover and rainfall, tropical cloud clusters serve as significant precursors to tropical cyclones, with approximately 6.4% of such clusters globally transitioning into tropical cyclones annually (Hennon *et al.*, 2013).

Efforts have been ongoing to reconcile the clustering of tropical convection with basic theoretical frameworks. Randall and Huffman (1980) suggested that clustering arises when clouds can generate a surrounding environment that is more conducive to future convection compared to more distant areas. But what exactly creates the prolific surrounding environment? What determines the transition of a single cumulonimbus, to an MCS, to a hurricane? There is not a clear consensus on this in the scientific community. If we knew the answer, we could predict the birth of a hurricane from the point in time that it was just a single, seemingly innocuous, cloud.

Until this point, we have primarily focused on MCS and CP characteristics derived from observational analyses from the last 80 years, i.e. measurements from weather stations, weather towers, soundings, radar and satellites. A milestone in weather and climate research was the possibility to simulate the evolution of the governing equations of the weather. Solving the Navier-Stokes equations and energy conservation laws on a discretized grid requires a potent calculator, and this became possible with the advent of the computer in the 1960s. The first models consisted of representing the atmosphere as one single column. With more computing power, this increased to more columns (higher resolutions), covering larger and larger areas (larger domains). Besides giving the ability to simulate the future, for weather forecasting and climate projections, the capacity to simulate weather phenomenon accurately is important diagnostically because some features of the weather are almost impossible to observe. In this dissertation, we will focus on simulations from cloud resolving models (CRMs). These models run at a high enough resolution horizontally and vertically to resolve many of the important kilometer-scale processes involved in the creation and development of convective clouds.

There is a simple fundamental principle which governs the global atmosphere: the time-averaged radiative cooling of the atmosphere must be balanced by latent heating from condensation and the supply of sensible heat from the surface. Manabe and Strickler (1964) were the first to use a single-column model to show that the radiative cooling balanced by convective overturning could explain the vertical temperature structure of the tropical atmosphere (Manabe and Strickler, 1964). This contributed to developing a framework for studying the atmosphere from a simplified point of view: Radiative-convective equilibrium (RCE), i.e. the statistical equilibrium state that the atmosphere and surface would be in, in the absence of lateral energy transport, consisting of a balance between net radiative cooling and convective heating. RCE is a valid approximation of the global atmosphere, and has in particular been used for investigating the behavior of moist convection in 2D and 3D CRMs (e.g. Held *et al.*, 1993; Robe and Emanuel, 1996; Tompkins and Craig, 1998). RCE simulations are often used to model the atmosphere over the tropical ocean with CRMs, because of the simple framework applicable here where the sea surface temperature (SST) variations can be relatively small, both in space and time, and there are fewer external perturbations to the atmosphere than in the midlatitudes (where frontal dynamics are predominant).

This is where an intriguing phenomenon arises: CRM simulations run in RCE with a constant SST and homogeneous boundary conditions, while initially characterized by randomly distributed convective cells, can spontaneously generate a highly localized and well-organized cluster of convection surrounded by a cloud-free domain after 10s of days of simulation time, as in Figure 2.15.

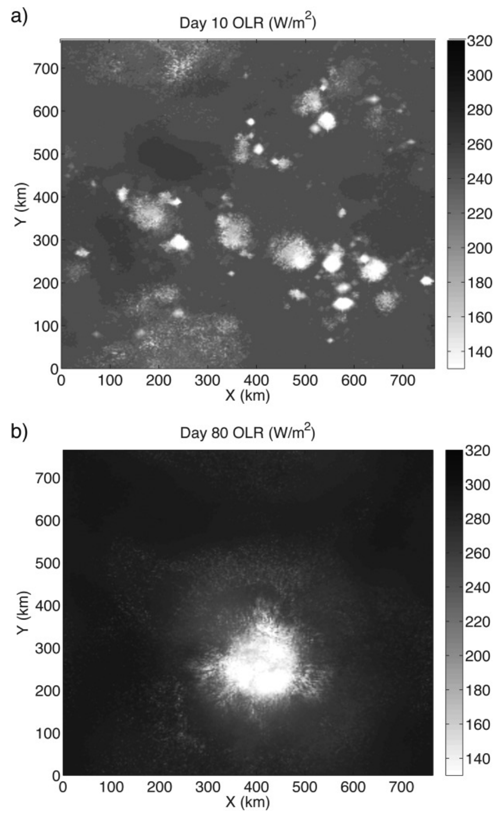


Figure 2.15: Snapshot of outgoing longwave radiation (OLR) at (a) day 10 and (b) day 80 of a radiative-convective equilibrium simulation at 305 K. From Wing and Emanuel (2014).

The individual convective rain cells at the beginning of the simulations measure only few kilometers horizontally, while the clusters they form after tens of days of simulation can span hundreds or even thousands of kilometers. This phenomenon is known as convective self-aggregation (CSA), and has been extensively studied in the last two decades - first, to understand the processes that lead to clumping of convection starting with otherwise homogenous boundary conditions (reviewed thoroughly in Wing *et al.* (2017)); second, to understand if the simulation-born phenomenon can be found in reality (a still emerging field of research presented in the review by Holloway *et al.* (2017)). While CSA was initially found in simple RCE set-ups with fixed sea surface temperatures and homogeneous boundary conditions, it has been shown to be robust to the presence of rotation (Bretherton *et al.*, 2005; Khairoutdinov and Emanuel, 2013; Wing *et al.*, 2016), vertical wind shear (Bretherton *et al.*, 2005), two-dimensional or three-dimensional settings (Jeevanjee and Romps, 2013), an interactive ocean mixed layer (Hohenegger and Stevens, 2016), and to occur in global climate simulations with parameterized convection in aquaplanet non-rotating settings (Coppin and Bony, 2015).

CSA spontaneously emerges from the very small scales of individual convective clouds, but it can show effects at the synoptic scale. The clumping of convection is in fact associated with changes in the large-scale state of the atmosphere, such as a drying of the atmosphere, a shrinking of upper-tropospheric clouds, and an enhanced ability of the atmosphere to loose heat to space (e.g., Wing and Emanuel, 2014; Wing and Cronin, 2016; Bony *et al.*, 2016). In the course of an aggregating simulation, dry regions that are initially small tend to expand and merge, often leading to only one persistent dry region that covers most of the domain. The remainder of the domain features an intensely convecting moist patch, with strong convective updrafts, reduced outgoing longwave radiation and heavy precipitation. The transient approach to the fully aggregated state can be described as “dry gets drier” and “moist gets moister” (Haerter and Muller, 2023). The exact processes behind CSA have been investigated in depth in simulations, and there is not one process that rises over the rest in literature. We will briefly approach them one by one, while deeper explanations can be found in the recent reviews by Muller *et al.* (2022) and Haerter and Muller (2023).

Key processes influencing CSA.

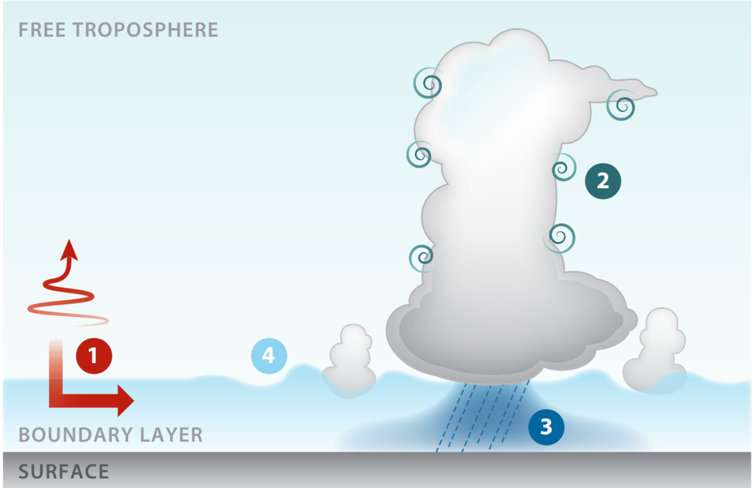


Figure 2.16: Schematic of key processes leading to CSA. (1) enhanced radiative cooling in dry regions and associated shallow divergent circulation (red arrow), (2) turbulent entrainment of environmental air at the edge of clouds, (3) evaporation-driven cold pools in the boundary layer, (4) boundary layer wave emission. From Muller *et al.* (2022).

1. Radiative feedbacks are critical for CSA (Bretherton *et al.*, 2005) to develop and persist in a CRM. Longwave radiative cooling is increased in dry regions: clear-sky areas don't obstacle outgoing longwave radiation emitted to space by the surface. On the other hand, longwave radiative cooling is reduced in moist regions, which are covered by high clouds which trap outgoing longwave radiation emitted to space by the surface. This results in a differential cooling between dry and moist regions, with cooling in the dry regions promoting subsidence (also known as a dry pool), and reduced cooling/heating in the moist regions promoting upward motion, thus generating a circulation with a near-surface flow directed from the dry areas to the moist areas. This flow acts by bringing more moisture to the moist areas, creating a positive feedback loop (process 1 in Figure 2.16).

2. Moisture feedbacks contribute to maintaining CSA in a CRM. Cloudy updrafts, through turbulent motions, entrain neighboring air at their edge: This decreases the buoyancy and upward motion of the updrafts (through drying and evaporatively driven latent cooling). In self-aggregated simulations, the air surrounding updrafts is relatively moist, since the updrafts are in moist areas of the domain, which reduces the negative effect of entrainment and

further favors the aggregation of updrafts in the moist region (Tompkins and Semie, 2017) (process 2 in Figure 2.16).

3. Cold pools have a dual role in CSA, notably through their ability to trigger new convection around a precipitating cloud and to redistribute moisture in the boundary layer (process 3 in Figure 2.16). The edge of a CP, with its internal vortical structure, can cause mechanical lifting of air around the original cloud, leading to the rapid development of new convective clouds in the vicinity of the precipitation, which can obtain even more energy where two or more CPs collide (Meyer and Haerter, 2020; Torri *et al.*, 2015). Furthermore, the presence of enhanced moisture on the edge of CPs ("moisture rings") as seen primarily in simulations following Tompkins (2001a) and in observations of CPs over tropical ocean (Szoeke *et al.*, 2017; Zuidema *et al.*, 2017), can also contribute to new convection forming on the edges of a CP, albeit on longer timescales than mechanical lifting. However, the CP also has a cold, dry interior, which tends to stabilize the atmosphere below the precipitating cloud, and suppress new convection in place of the previous precipitation. A CP can therefore act as a means to redistribute moisture back away from the moist convective area, thus acting against the moisture feedback (process 2) in the boundary layer. While the role of CPs in CSA is not entirely settled, it is important to note that large scale radiation and moisture feedbacks tend to dominate, once the domain-wide anomalies such as dry patches, exceed the typical size of individual CPs. We will explore this further in Chapter 3.

4. Gravity waves can trigger CSA in the absence of the other feedbacks (process 4 in Figure 2.16). As proposed by Yang (2021), CSA can spontaneously emerge through the formation of standing wave packets by convectively coupled gravity waves, which segregates the domain into regions of convective activity and inactivity.

The four key processes above have been extracted by Muller *et al.* (2022) from the many CRM simulations in literature focused on CSA. To complement the simulations, there is a growing pool of literature that focuses on building conceptual models that can explain the onset and development of CSA in the CRM simulations, with the aim of gaining deeper understanding through lower-complexity models. We will focus primarily on the radiative feedbacks and CPs.

Simplified models to explain CSA.

We will approach the simplified models by thinking of them in terms of the directions along which the processes act, i.e. in the vertical (radiation), and in the horizontal (CPs).

In the vertical.

Emanuel *et al.* (2014) coined the concept of "Radiative convective instability", exploring CSA as a linear instability of the tropical atmosphere with an elegant theoretical emissivity model, consisting of a two-layer atmosphere (Figure 2.17) which represents humidity in the upper and lower troposphere.

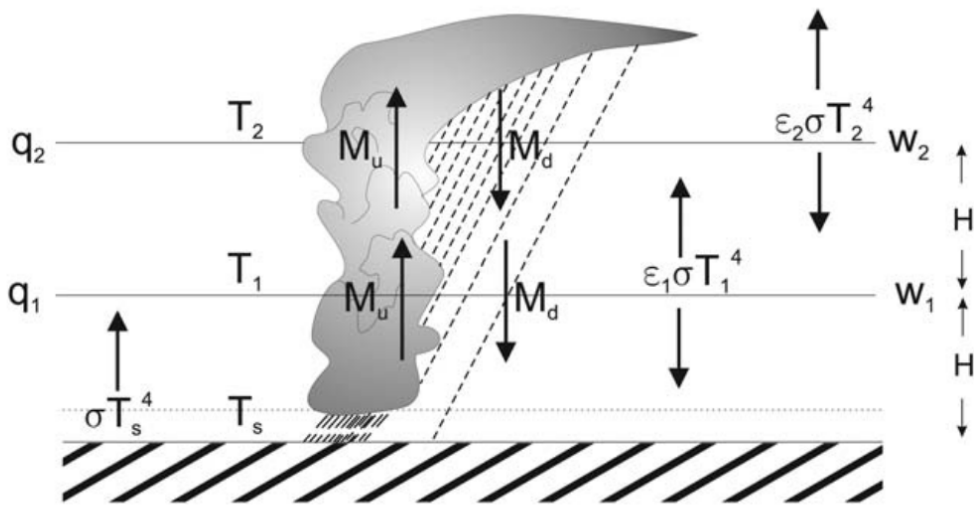


Figure 2.17: The two-layer model by Emanuel *et al.* (2014). Surface temperature and the temperatures of each layer are specified and constant. The emissivities, ϵ, updraft and downdraft mass fluxes, M_u and M_d, large-scale vertical velocities, w, and specific humidities, q, are variable. The vertical arrows depict the convective and radiative fluxes.

They found that above a critical SST threshold, the RCE state becomes linearly unstable, leading to the emergence of large-scale overturning circulations. This happens in particular when the lower troposphere becomes so optically thick due to high specific humidity, that its cooling to space depends only on the humidity in the upper troposphere.

The instability in Emanuel *et al.* (2014) represents a subcritical bifurcation of the RCE state (Figure 2.18). Below a critical SST, the RCE state is stable to small perturbations in humidity, but large enough perturbations can result in the emergence of either a dry state characterized by widespread subsidence or a moist state marked by overall ascent. Above the critical SST, these transitions are spontaneous. Since the Emanuel *et al.* (2014) model is a one-column model, CSA in a CRM can be seen as a combination of the two equilibria in distinct parts of the domain.

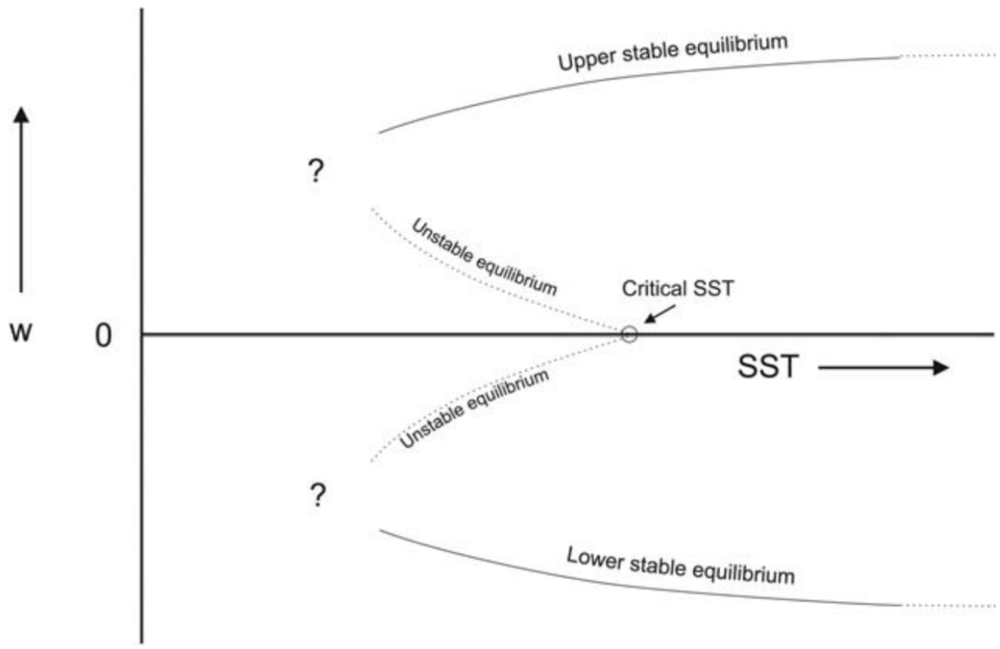


Figure 2.18: Schematic diagram of equilibrium states, showing the large-scale vertical velocity w as a function of SST. Above the critical SST, the RCE state is linearly unstable. Question marks denote unexplored regions. From Emanuel *et al.* (2014).

It is worth noting that there is no consensus regarding the effect of SST on CSA, as shown in the comprehensive RCE model intercomparison study by Wing *et al.* (2020). The model in Emanuel *et al.* (2014) focuses entirely on free-tropospheric moisture and radiation feedbacks, neglecting any boundary-layer processes that act to redistribute moisture (CPs). In Figure 2.18, there are question marks denoting the unexplored regions of the regime diagram: The transition areas between the unstable equilibria and the stable equilibria, which leads us to explore the horizontal direction.

In the horizontal.

A conceptual model to explain CSA emergence that focuses solely on the boundary-layer processes involved and disregards radiation feedbacks is found in Haerter *et al.* (2019). This model introduces a mechanism for phase separation similar to CSA, which relies solely on CP collisions and a global energy constraint. Here, CPs are conceptualised as growing circles, new rain cells originate at the intersections of the circles (mimicking CP collisions triggering new convection), and rain activity ceases when the global energy budget is "used up". With random initial conditions, where the CP center density is evenly distributed across the area, the model's dynamics progressively lead to the formation of increasingly larger clusters of convective activity, and inactive areas, like what is seen in classic CSA - all thanks to CPs. Nissen and Haerter (2020) built on to this by adding two properties of the circles: a radius R_{min} - within which rain is suppressed, and R_{max}, the maximum radius at which the CP can trigger new convection. Figure 2.19 shows 2D snapshots in time of various steps of the "circle model".

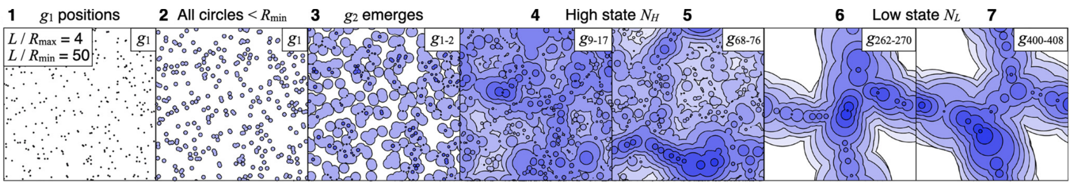


Figure 2.19: The "Circle model" capturing CP dynamics leading to CSA. Seven snapshots running forward in time from one model run: Snapshots 1–2 show the initial first generation rain-cell positions. Snapshot 3 shows the emergence of circles (CPs). Snapshots 4–5 and 6–7 show representative pictures of the high state and the low state. From Nissen and Haerter (2020).

By testing the two parameters, Nissen and Haerter (2020) find that smaller CPs (smaller R_{max}), and CPs with larger suppression areas (larger R_{min}), tend to facilitate the emergence of CSA, suggesting that large CPs actually hamper CSA. This work, in a way, focuses on the development and initial growth of the first dry patches in CSA, attributing it to CP effects, when the CPs have the "right" properties.

An incredibly simple conceptual model that manages to efficiently explain the effect of CPs on CSA in relation to the larger scale radiative feedbacks, is found in Yanase *et al.* (2020). They first compare 20 CRM simulations run in RCE with constant SST (set to 300K) for many different horizontal resolution and domain sizes, and they find a critical domain size around 500 km, above which all simulations with this SST (even at high resolutions) develop CSA, suggesting the existence in the real atmosphere of a characteristic length in the interaction of convective clouds with the larger scale environment. Their conceptual model is explained through the competition of opposite effects on the moisture variance in the boundary layer: CPs and the subsiding radiatively driven dry pool (see schematic in Figure 2.20).

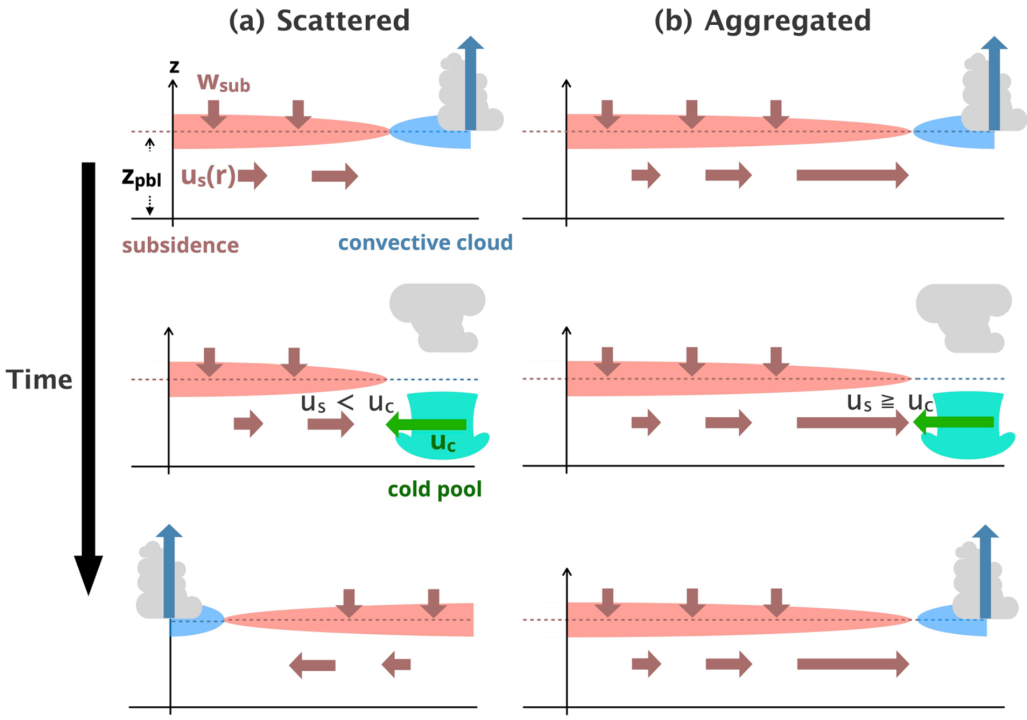


Figure 2.20: Schematic of the competition between the two opposite feedbacks by horizontal divergent flows in a subsidence area and a convective area accompanied by temporal evolution. (a) Scattered case and (b) aggregated case. From Yanase *et al.* (2020).

CPs, constantly redistributing moisture in the boundary layer and suppressing new convection where rain was present, tend to act against CSA in small domains (the "scattered" state in Figure 2.20). However, as the domain size increases, the positive feedback from the radiatively driven dry pool, and its associated low-level horizontal divergent flow strengthens. When the domain size exceeds the critical length, the positive feedback wins over the negative feedback due to the CPs in the convective region.

It seems to emerge from these conceptual models that CPs with specific properties and with a global constraint can determine the onset of CSA, and radiative feedbacks can maintain CSA once this wins over the CPs. We will combine these concepts and build our own toy model operating both in the horizontal and in the vertical, in Chapter 3, with the "Game of Cloud".

Time-varying surface temperatures. The majority of the presented studies mentioned until here, refer to simulations with constant SSTs, mimicking tropical sea surfaces. Haerter *et al.* (2020) instead applied diurnal variations to the surface temperature to mimic both land surfaces (5K diurnal cycle amplitude) and sea surfaces (2K diurnal cycle amplitude) - and found qualitatively different behaviors between the two. With the stronger surface temperature forcing, MCSs in fact emerge in the time scale of days, stretching over 100 kilometers horizontally, self-organizing within the domain and leaving other regions predominantly rain-free (Figure 2.21).

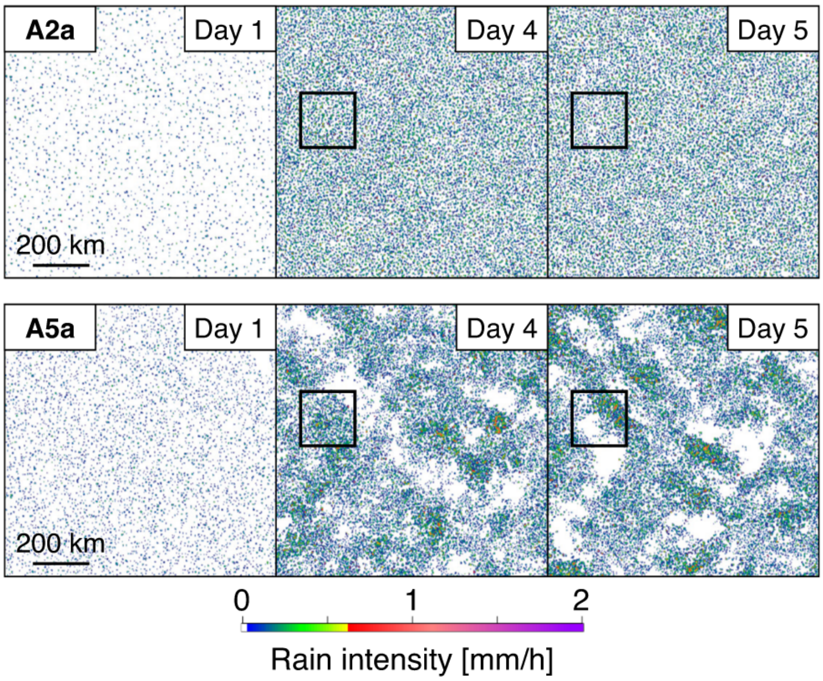


Figure 2.21: Day averages of surface rainfall during day 1, day 4 and day 5 for CRM simulations with (top row) 2K amplitude diurnal cycle and (bottom row) 5K amplitude diurnal cycle. From Haerter *et al.* (2020).

This phenomenon, referred to as "diurnal self-aggregation" in Haerter *et al.* (2020) is attributed to the higher spatial density of rain cells in a short window of time. Under the windows of high convective activity during the diurnal cycle, the CPs under nearby convective rain cells in fact tend to merge, forming a super CP of much larger extent, spatially and temporally. This larger CP then triggers more convection along its gust front, setting off a cascade of convective rain cells, adding to the combined CP. The thermodynamic anomalies generated by these combined CPs in the boundary layer dissipate relatively slowly, suppressing further convection in the same region on subsequent days. A beautiful analysis of the described CP-cascade happening already on the first day of a simulation with a 5K amplitude diurnal cycle was shown in the more recent work by Jensen *et al.* (2022), seen in Figure 2.22.

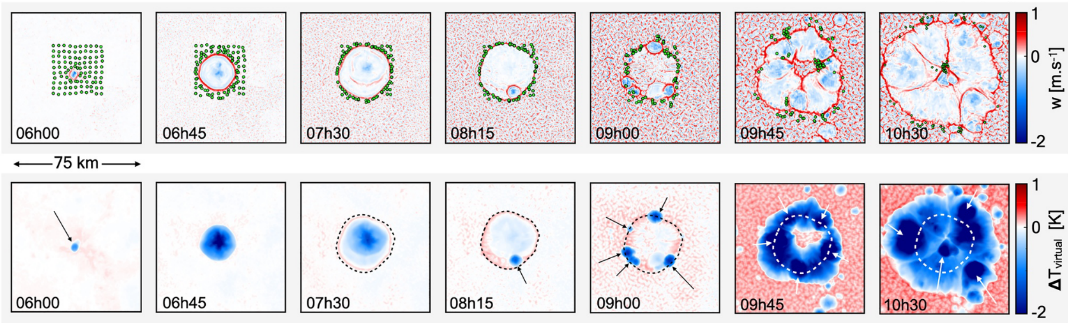


Figure 2.22: (top row) Instantaneous vertical velocity fields at z=50m during day 1 of a CRM simulation with 5K amplitude diurnal cycle, showing the evolution after the first CP. (bottom row) Same as top but showing virtual temperature anomaly field. Arrows highlight new CPs. From Jensen *et al.* (2022).

Jensen *et al.* (2022) took the Haerter *et al.* (2020) study further by running longer simulations with the 5K and 2K amplitude diurnal cycles. They find that with the stronger surface temperature forcing, not only do MCSs emerge, but also a domain-scale dry patch, reminiscent of CSA. The emergence of a dry patch is attributed to the action of MCSs and their associated CPs: the CP cascades are in fact shown capable of clearing large-scale (~ 100km) areas from moisture and suppressing convective activity, leading to a subsiding circulation over the already dry regions which then permanently drives moisture out toward the already moist (convective) regions - just like the radiatively driven CSA feedbacks. Jensen *et al.* (2022) checked whether eliminating the temporal oscillation in surface temperatures would result in the disappearance of the dry patches once they had formed due to the diurnal temperature fluctuation: however dry patches persisted even when the surface temperature was set to constant. This results in a form of hysteresis, which we will explore further in Chapter 3.

Haerter *et al.* (2020) concluded their study with a hypothesis on a shortcut to observed clustering over sea: "clustering may emerge over land surfaces, where a strong diurnal cycle prevails, and may then be advected over the sea, where it could grow further under more RCE-like conditions." Jensen *et al.* (2022) concluded their study with a hypothesis on a path to persistent clustering over sea: "when organized convective cloud clusters, produced under a high-amplitude surface temperature forcing, are eventually advected over regions with little surface temperature variation, the clustered pattern may persist and even intensify further. Such a situation could be found at the interface between tropical continents and oceans, for example, at the west coast of Africa."

We will follow these hypotheses, by focusing, in Chapter 3 on the transition from tropical land to tropical sea. We will impose realistic surface temperature forcings found over tropical African land and the adjacent Atlantic Ocean, and analyse the emergent hysteresis arising in the CSA developed by MCSs over land and advected over the ocean, hinting at a potential path to the formation of large scale hurricanes over the ocean.

2.3 Nowcasting and Recurrent Neural Networks (RNNs):

–Towards Real-World Applications–

We investigate the upscale growth of MCSs over tropical land and the transition to a self-aggregated state in cloud-resolving simulations in Chapter 3, and we measure CPs with a weather tower and radar in the Netherlands in Chapter 4. These MCSs and CPs, which have an obvious impact on society due to the extreme rainfall and high wind gusts associated with their occurrence, are incredibly difficult to predict in real time in the real world due to the apparently random nature of convection. Numerical models used for operational forecasting, also called Numerical Weather Prediction (NWP) models, are run at a resolution that is too coarse for capturing the initiation and development of convective clouds bearing rain. The "skill" for predicting convective rain in the short-term future is therefore low for NWPs. Even with high-resolution cloud resolving models, the predictability of convective rain is poor (Hohenegger and Schär, 2007).

Nowcasting. In Chapter 5, we will step out of the theoretical world and look at a more immediate application of atmospheric sciences in society. Nowcasting is a branch of forecasting that utilizes different techniques than the traditional NWP modeling, to predict weather conditions ranging from 0 to several hours ahead. A common technique used for locally nowcasting rainfall, which outperforms NWP modeling for the near future, is the extrapolation of rainfall trajectories from radar imagery, once a storm has formed. Radar-based nowcasting can be effectively implemented thanks to spatially dense radar networks present in many areas of the world, that provide data in real time with updates every 5 minutes. The problem is, in the locations of the world where there is sparse radar coverage, or no radar coverage at all, there is very little accessible short-term information about rainfall.

Tropical Africa is both a region of the world that hosts the largest MCSs in the world (as we saw in Figure 2.7) and with the absolute least radar coverage in the world (see Figure 2.23 showing the curent global radar coverage). Long-lived MCSs play a crucial role in extreme precipitation events in this part of the world (Roca and Fiolleau, 2020), particularly over the Sahel (Figure 2.24), where they contribute up to 90% of seasonal rainfall (Nesbitt *et al.*, 2006). Improving the prediction of these

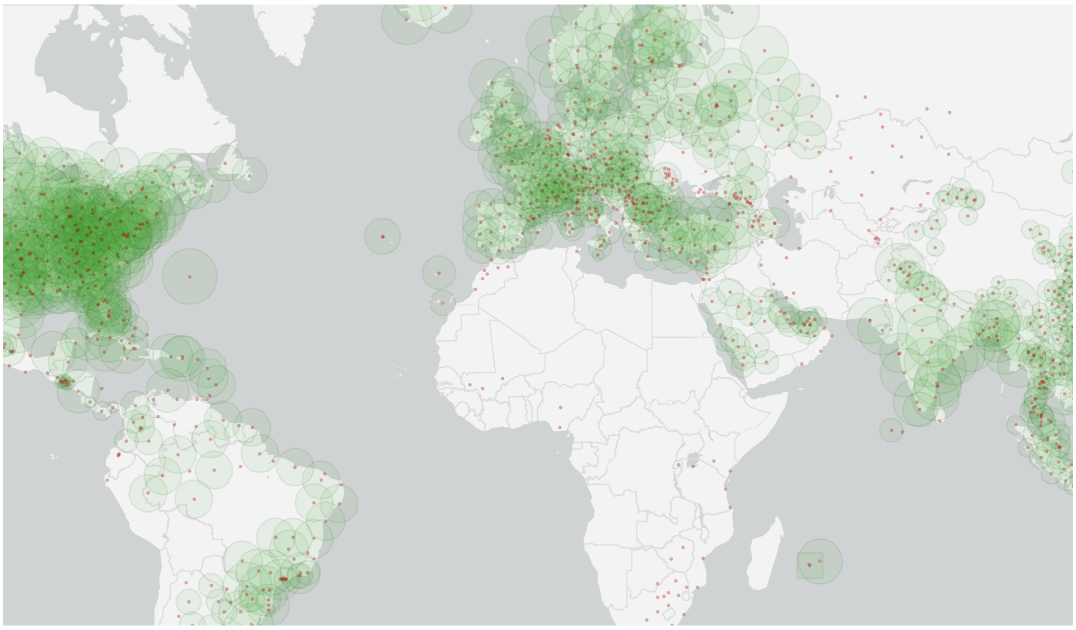


Figure 2.23: A snapshot of radar coverage in the world, as of February 2024 (*Tropical Globe Radar Database* Accessed: 2024-02-26). Locations of existing radars in red and range of publicly available radar data in green.

Figure 2.24: Sahel region of Africa. (*Encyclopædia Britannica* Accessed: 2024-02-28[a])

events is vital due to the rising frequency and intensity of extreme rainfall in the Sahel (Taylor *et al.*, 2017), which has led to escalating socio-economic impacts from floods in recent decades (Di Baldassarre *et al.*, 2010; Tramblay *et al.*, 2020). We put our focus on the Sahel in Chapter 3 due to the large MCSs that are formed in this region thanks to the strong diurnal cycle, and advected westward over land and towards the Atlantic Ocean. There, we use satellite data to monitor the MCSs and we run high-resolution simulations to gain insight about the interplay between MCSs and the larger-scale dynamics. What can we then do to nowcast these enormous storms, to provide an early-warning signal of incoming rain and wind gusts to the people living in flood-prone urban areas and fishermen that risk being caught at sea? These questions, alongside the research-driven questions about the initiation and development of MCSs, brought us to think of a field campaign to take place in Senegal, the Western-most country of the Sahel.

Rainfall in West Africa is closely associated with the West African Monsoon (WAM): A seasonal wind system that brings significant rainfall to the region annually (Figure 2.25). It typically occurs from June to September, peaking in July and August,

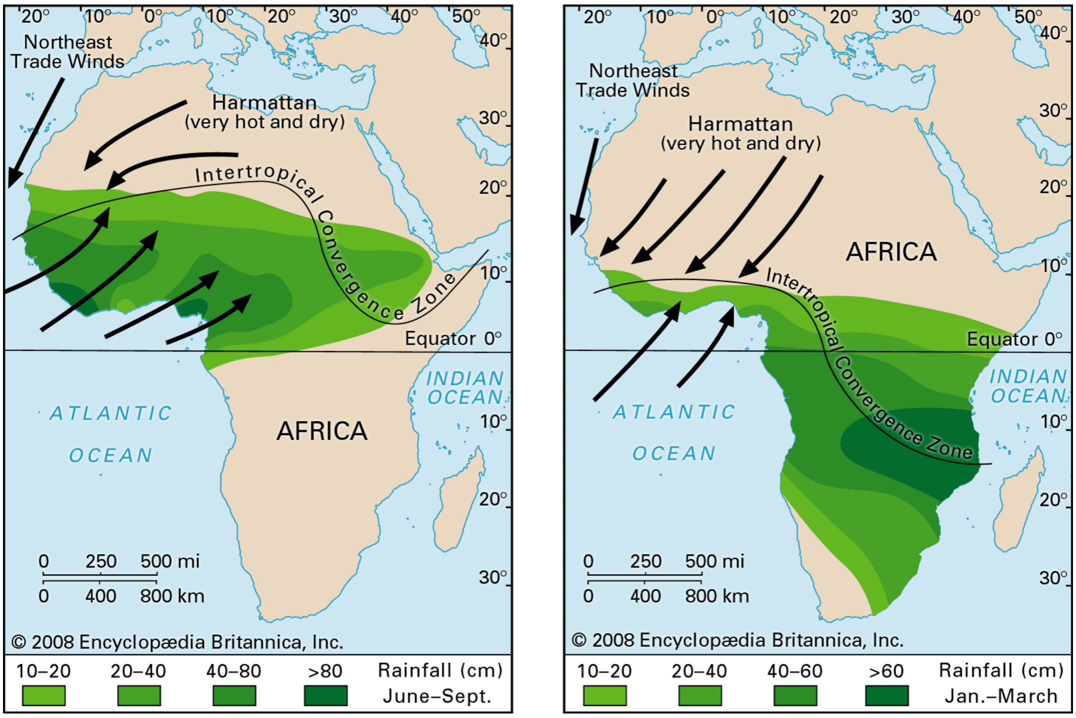


Figure 2.25: Wind and rainfall patterns of the WAM. (left) June-September; (right) January-March (*Encyclopædia Britannica* Accessed: 2024-02-28[b])

which we will call "rainy season". The WAM is driven by the differential heating between the African landmass and the surrounding oceans, particularly the Gulf of Guinea and the Atlantic Ocean. During the summer months, the land heats up more quickly than the ocean, creating a low-pressure area over land. This contrast in pressure causes moist air from the ocean to flow inland, where it rises and cools, leading to the formation of clouds and precipitation. The WAM interacts with the Intertropical Convergence Zone (ITCZ), a band of low pressure near the equator where the trade winds of the Northern and Southern Hemispheres converge. During the WAM season, the ITCZ also shifts northward, following the seasonal migration of the Sun. This northward movement brings the ITCZ and its associated band of deep convection into closer proximity to West Africa, enhancing the convergence of moist air masses from both the Atlantic Ocean and the Gulf of Guinea. As a result, the convergence of these air masses intensifies the monsoon circulation and contributes even more to the onset of the rainy season in West Africa. The rainy season in this region, a product of the WAM and the ITCZ, plays a crucial role in the climate and ecosystems of West Africa, providing much-needed rainfall for agriculture and sustaining river systems. However, it is also the season of severe weather events, bringing MCSs and associated floods. It is important to underline here that the deep convection in the ITCZ tends to be transported westward across the continent with the African Easterly Jet (AEJ), as highlighted by the average MCS velocity vectors over the Sahel in Figure 2.8.

A dense network of weather stations, especially in the absence of radar, is a way to monitor the WAM, the associated MCSs, study the heterogeneous rain rates within MCSs, and the abrupt changes in temperature and the wind gusts due to CPs - the variables that ultimately affect humans. However, as stated in Knippertz *et al.* (2020), "The meteorological station network in West Africa is sparse and existing data are not always available for research and operations, limiting evaluation of model and satellite products." Efforts have been made to install ground-based instruments in this region: Both the AMMA (African Monsoon Multidisciplinary Analysis) field campaign and the DACCIWA (Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa) project aimed to deepen the general understanding of atmospheric processes in West Africa, a region heavily influenced by the WAM. The AMMA initiative (Lebel *et al.*, 2011), spanning from 2002 to 2010, conducted extensive long-term monitoring efforts and deployed a network of instrumentation across multiple countries, with longer term ground-based measurements in Mali, Niger and Benin, enhancing West Africa's visibility in global meteorological systems. The DACCIWA project (Kohler *et al.*, 2022), active from 2013 to 2018, focused on enhancing understanding of meteorology and air quality in the region, particularly during the summer of 2016 when a comprehensive field campaign was conducted in southern West Africa. This campaign involved three research aircrafts and various surface-based instruments across sites in Ghana, Benin, and Nigeria, generating a comprehensive atmospheric dataset that supported the evaluation of dynamical models and satellite data. Besides the importance of installing more ground-based instruments with easy data-access, there is a need for improved forecasting tools and approaches (Fletcher *et al.*, 2023). The African Science for Weather Information and Forecasting Techniques (SWIFT) program (Fletcher *et al.*, 2023) is the most recent large-scale effort aimed at addressing this challenge, by enhancing forecasting capabilities through collaboration between researchers and operational forecast services in Kenya, Nigeria, Ghana, Senegal, and the UK. Notable steps forward in nowcasting for the Sahel are already being seen through the use of geostationary satellite data. Infrared cloud-top temperature retrievals cannot give direct measurements of precipitation, however, filtering algorithms can be applied to identify convective cores within MCSs, strongly linked to heavy precipitation (Klein *et al.*, 2018). The most recent studies on nowcasting in this region have shown novel approaches to produce probabilistic nowcasts of convective activity with satellite imagery (Anderson *et al.*, 2023). Another approach for nowcasting MCSs is to use land surface temperature retrievals from satellites as a proxy for soil moisture levels, since dry soils early in the day have shown a strong correlation with MCS activity in the Sahel (Taylor *et al.*, 2022). These exciting ideas have yet to be implemented in operational centers.

The path to a field campaign. Throughout my PhD, thanks to periodic meetings between the Laboratory of the Atmosphere and the Ocean at the University Cheikh Anta Diop (UCAD) of Dakar, the National Agency of Civil Aviation and Meteorology of Senegal (ANACIM), the UK Centre for Ecology and Hydrology (UKCEH) and our Bremen-based research group, a field campaign slowly came to the surface, to study MCSs and CPs over Senegal, with a dense weather station network sending real time measurements to a cloud-server at 1-minute update intervals. I had the opportunity to participate in the planning phase of what we will call the DakE (Dakar-East) campaign, the testing of weather stations in Bremen, the installation of the first two pillars of this campaign, namely two weather stations set up, one in Dakar, and one East of Dakar, in the last two weeks of September 2023. Due to the

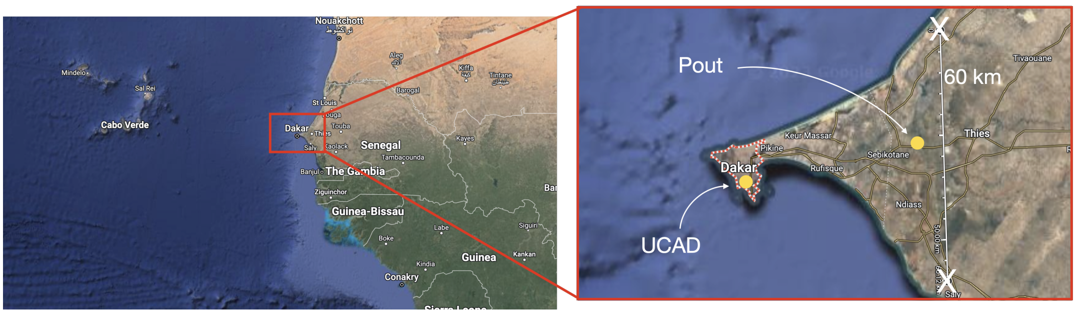


Figure 2.26: Locations of the first two stations of the DakE field campaign in Senegal: the town of Pout and the University of Cheikh Anta Diop (UCAD).

monsoonal dynamics, the rainy season in Senegal is approximately from June to September, and we successfully managed to measure a handful of rain events in the last two weeks of September 2023, before the dry season kicked in.

While the rest of the measurement network is being set up by my colleagues and we wait for the rainy season of 2024, in Chapter 5 we focus on creating a simple nowcasting tool to be used with the two first stations we set up in September 2023. Since we do not have a full rainy season of real-time data yet, we run a weather forecasting model at a 1 km horizontal resolution over Senegal, with realistic boundary conditions, to simulate a full rainy season from 2019. From this simulation we extract time series from the locations of the installed weather stations, thus a full rainy season of "synthetic station data". We know that an MCS produces CPs, and that CPs can be detected and measured from ground-based weather stations (Chapter 4). We can then detect CPs in the synthetic station data from Senegal and study the time series of measured data leading up to the CPs.

Here is where machine learning comes in: we can train a machine learning algorithm to identify certain patterns in the measured variables that are precursors to the CPs, so it can learn to predict, or nowcast, a CP in a specific location before its arrival. Using machine learning in weather forecasting is an exploding field of research, thanks to the abundant amount of weather data (measurements and simulation output) to train algorithms on. Think of the recent boom in machine learning applied to language models, for instance, ChatGPT. This language model is trained on books, articles and websites, to predict the next word in a sequence of text based on the words that came before it. By repeating this process millions of times with different sequences of text, the model also learns patterns and relationships in the data. We want to do something very similar, but with time series data. We will now break down the pieces to understand what algorithm we will use in Chapter 5 to learn to predict CPs in our time series data.

Neural networks. A neural network is a computational model inspired by the structure and function of the human brain's interconnected neurons. It consists of interconnected nodes, or "neurons," organized in layers in the case of feed-forward networks, which we will explore. Each neuron receives input, processes it, and produces an output that serves as input to other neurons. In a typical neural network, there are three main types of layers: the input layer (receives the initial data or input features typically in the form of multidimensional arrays, and produces a hidden state), the hidden layers (perform computations on the hidden state), the output layer (produces the network's output, which could be a prediction based on the input data). During training, a neural network learns to perform tasks by adjusting the strengths of connections between neurons, called "weights," based on examples of input-output pairs, mimicking the human brain's ability to make new connections. This process, known as "training" or "learning," involves forward propagation of input data through the network, comparing the predicted output to the actual output through a loss function, backpropagation of the loss function, and adjusting the weights to minimize the difference between them using optimization algorithms.

Recurrent Neural Networks. Recurrent Neural Networks (RNNs) are a type of neural network architecture commonly used for sequential data processing tasks, thanks to the presence of loops and a hidden memory state that allows for information to persist. While feedforward neural networks process input data sequentially without any memory of past inputs, recurrent neural networks can retain information about past inputs through their recurrent connections, enabling them to capture temporal dependencies in sequential data. Thus, RNNs are well-suited for tasks where the input and output data are sequences of variable length or where there is a temporal dependency between elements in the sequence. One can see an RNN as a sequence of many neural network cells. In each cell of the RNN, the input of the current time step x_t (present value) and the hidden state h_{t-1} of the previous time step (past value) are combined, and then limited by an activation function to determine the hidden state h_t of the current time step (see conceptual diagram in Figure 2.27).

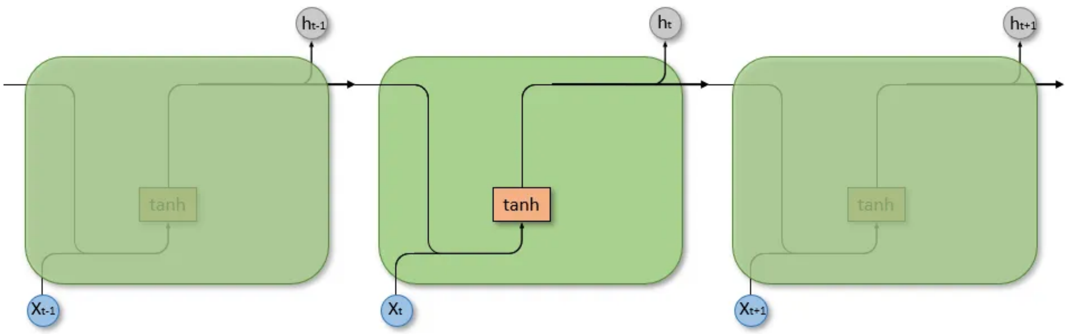


Figure 2.27: Schematic of an unfolded RNN. From Dancker (2022).

RNNs are used in various applications such as natural language processing for tasks like language modeling, machine translation, and text generation. There, RNNs process sequences of words or characters, capturing contextual information and dependencies between words. They are also used in speech recognition systems, where the input is an audio waveform represented as a sequence of samples over time. Most importantly, RNNs are used in time series analysis, where they can model and predict future values based on past observations. They are used in financial forecasting, weather prediction, and other domains where predicting future events based on historical data is important.

Long Short-Term Memory Networks. Long Short-Term Memory (LSTM) is a special kind of RNN architecture designed to overcome the limitations of traditional RNNs in capturing long-term dependencies in sequential data. First introduced in 1997 by Hochreiter and Schmidhuber (1997), LSTMs are particularly useful for tasks involving time series data, natural language processing, and other sequential data analysis tasks. In addition to a traditional RNN, an LSTM has a cell state, which represents the memory of the cell and is updated over time through a combination of forgetting old information, adding new information, and maintaining existing information (top horizontal line in the LSTM schematic, Figure 2.28).

At the core of an LSTM unit are three gates: the *forget gate* (decides what information to discard from the cell state), the *input gate* (decides what new information to store in the cell state), and the *output gate* (decides what information to output from the cell state).

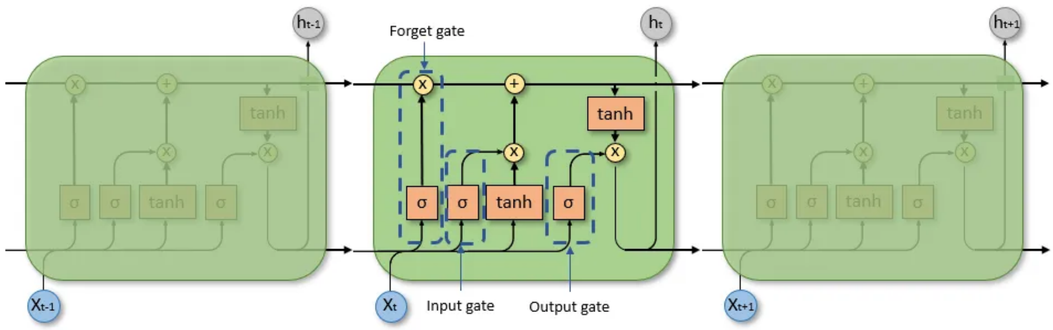


Figure 2.28: Schematic of an unfolded LSTM. From Dancker (2022).

These gates allow LSTMs to learn long-term dependencies by selectively remembering or forgetting information over time. Unlike traditional RNNs, LSTMs can maintain information over many time steps, making them well-suited for tasks requiring memory over time. When referring to observational time series, LSTMs take a sequence of data points as their input, where each data point represents an observation at a specific time step. The input sequence is then fed into one or more LSTM layers, which process the sequential input data and capture dependencies between time steps, generating an output array containing all desired time steps for prediction.

An auto-regressive LSTM, which we will use alongside a traditional LSTM in Chapter 5, furthermore combines the capabilities of LSTM cells with auto-regressive modeling. This means using the network's output from previous time steps as input to predict the next time step, instead of directly predicting the next time step's value based solely on the current input. In other words, the model learns to predict future values in the sequence based on its own past predictions. The final output of both an LSTM and an auto-regressive LSTM can be, for example, a sequence of predicted values for future time steps in the input sequence, and both are geared toward capturing complex temporal patterns and both long and short-term dependencies in sequential data.

With these methods in mind, we will try to answer the following question in Chapter 5: Can we train an LSTM with time series data from two weather stations situated upstream and downstream of the general path of MCSs, to predict incoming CPs, before they arrive to the capital city of Dakar?

3 Tipping to an Aggregated State by Mesoscale Convective Systems

Authors: Irene L. Kruse; Romain Fiévet; Jan O. Haerter
Journal: submitted to *Journal of Advances in Modeling Earth Systems*
My contributions: conceptualization; data curation; simulations; formal analysis; methodology; validation; visualization; writing – original draft.

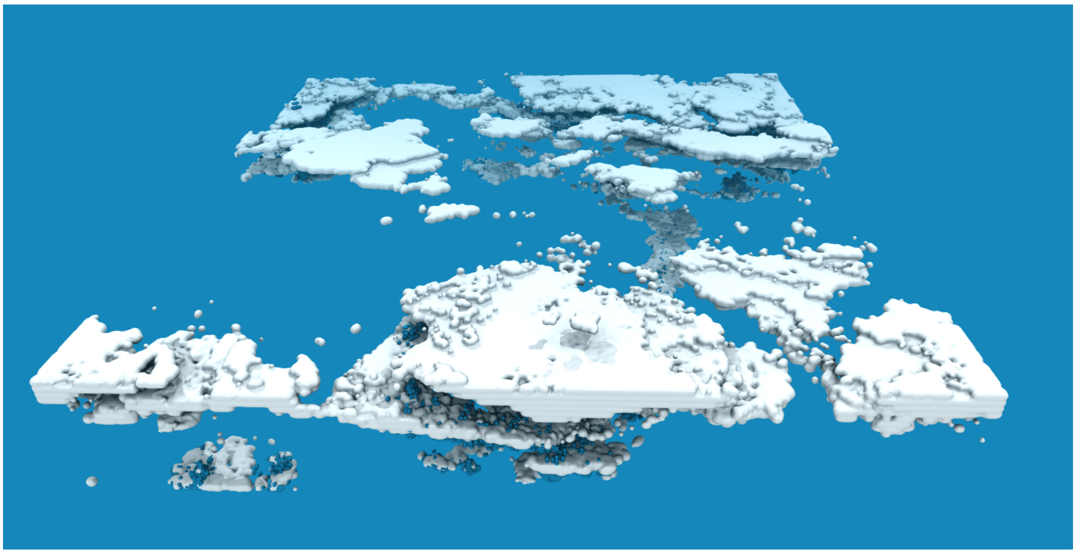


Figure 3.1: 3D rendering of deep convective clouds and rain in SAM simulation output (simulation DIU in Table 3.1). Thanks to Weria Pezeshkian for helping with this visualization.

Abstract

Radiative-convective equilibrium simulations were suggested to resist self-aggregation within a linearly-stable regime at low surface temperatures. Recent numerical work shows that this linearly-stable regime can rapidly transition to an aggregated state when exposed to realistic diurnal surface temperature variations. The resultant aggregated state is then stable, even when the surface temperature is set constant. Here we argue, by constructing a reaction-diffusion model, that this tipping process can be explained by the formation of mesoscale convective systems under the diurnal forcing. The model implies that strong cold pool interactions, invoked by the diurnal cycle, drive the self-organization of long-term buoyancy memory. Thus, whereas previous conceptual work disregarded the boundary layer, we here attribute key organizing mechanisms to it: namely the ability to cause rapid self-aggregation over continents and its advection over the ocean — with potential implications for hurricane formation.

3.1 Introduction

In Earth's atmosphere, convection can organize across many spatial scales, from local thunderstorms, to mesoscale convective systems (MCSs), tropical cyclones, and synoptic-scale waves. MCSs are organized clusters of thunderstorms spanning more than 100 km horizontally, persisting often for multiple hours (Houze, 2018). They are known to be the dominant source of rainfall in the tropics, and the longest-lived MCSs are shown to be largely responsible for tropical extreme precipitation (Tan *et al.*, 2015; Roca and Fiolleau, 2020). Globally, the most extreme storms tend to be located over land, and the most intense storms over oceans tend to be adjacent to land, where motion is favored from land to ocean, e.g. tropical West Africa and the adjacent Eastern Atlantic Ocean (Zipser *et al.*, 2006). However, the fundamental mechanisms driving the formation, intensification and dissipation of MCSs is not well established yet.

In an effort to better-understand the physics of convective storms, the atmospheric modeling community has been focusing for many years on the concept of Convective Self-Aggregation (CSA). In an idealized environment of radiative convective equilibrium, with homogeneous initial conditions and a constant-temperature tropical sea surface, convection can spontaneously clump, or aggregate, into domain-wide patterns of persistent dry areas and confined rainy areas over a temporal time-scale of weeks to months (Held *et al.*, 1993; Tompkins and Craig, 1998; Emanuel *et al.*, 2014; Muller *et al.*, 2022). CSA, albeit still a modeling paradigm, could reveal the mechanisms behind some of the convective organization observed in the tropics (Holloway *et al.*, 2017). The process of forming domain-wide structure can be achieved within few days by imposing oscillating surface temperatures with a large enough amplitude (Haerter *et al.*, 2020). The 'diurnally aggregated' cloud field is similar to CSA as it also constrains the surface rain field to certain parts of the domain. In fact, in the simulations by Haerter *et al.* (2020) mesoscale organization vanished when the diurnal cycle was removed.

The idealized tropical atmosphere, simulated under RCE conditions, is known to exhibit hysteresis effects (Khairoutdinov and Emanuel, 2010; Muller and Held, 2012). Recent work however describes a plausible route to such an aggregated state: under diurnally varying surface temperature forcing, multi-day spatiotemporally persistent dry patches can form. These first initiate in the uppermost atmospheric layers and subsequently penetrate through to the subcloud layer. When removing the diurnal forcing after a number of days, the persistently dry patches remain and are reminiscent of a classical aggregated state (Jensen *et al.*, 2022).

Conceptual models offer a tool to further understand convective self-aggregation (CSA) mechanisms (Emanuel *et al.*, 2014; Nissen and Haerter, 2021; Biagioli and Tompkins, 2023). Emanuel *et al.* (2014) explore CSA as a linear instability of the tropical atmosphere, focusing on radiative convective instability and free-tropospheric moisture, neglecting boundary-layer processes. In contrast, Haerter *et al.* (2020) and Nissen and Haerter (2021) emphasize the role of convective cold pools (CPs) in CSA emergence, with CP collisions driving phase separation akin to CSA formation. In Yanase *et al.* (2020), the interplay between CPs and radiatively driven dry pools in promoting or inhibiting CSA is investigated, highlighting the importance of the domain size. More recently, Biagioli and Tompkins (2023) build on Craig and Mack (2013) to simulate convective organization. They propose horizontal transport efficiency and subsidence rate as key parameters determining an "area of influence." Importantly, their model can estimate how a specific domain size and resolution is prone to aggregate, at least for constant surface temperatures. While the models above offer valuable insights, they individually overlook key aspects, lack the possibility to resolve both land and sea configurations, disregard boundary layer dynamics, or the integration of CPs contributing as both a negative feedback inside raincells and positive in their immediate surroundings. Together, these shortcomings point to the need of a conceptual model that integrates both vertical and horizontal dynamics influencing CSA, valid in different environments.

In this work we aim to strengthen the bridge between the modeling paradigm of CSA and the real atmosphere. Specifically, we build upon the reasoning of Jensen *et al.* (2022) that if an organized convective cloud field, produced under a high-amplitude surface temperature forcing, is eventually advected over regions with little surface temperature variation, the clustered pattern may persist and even intensify further. In the real atmosphere, the spatio-temporal scales of this phenomenon can be found over the tropical African continent and the adjacent Atlantic Ocean. There, land-born mesoscale convection is advected over the ocean, at times even maturing to tropical cyclones when reaching higher latitudes and thus acquiring rotation. Hereby, we explore how diurnal surface temperature amplitudes, typical of tropical land, affect the formation of persistent dry patches and the spatio-temporal extent of the emergent mesoscale convective systems.

To this end, we run a set of cloud resolving simulations initialized with typical vertical profiles of temperature and humidity. A large-amplitude diurnally oscillating surface temperature is imposed, which is then set to constant at different times, to see the effect on the diurnally aggregated cloud field. By tracking observed MCSs over West Africa and the adjacent Atlantic, we first show that our simulations give realistic deep convective diurnal cycles under the imposed surface temperature conditions. Analogous to Jensen *et al.* (2022), persistent mesoscale organization arises in the land simulations after running them for several days. When switching to constant surface temperatures, we find strong dependence on the degree of aggregation over 'land,' in determining its persistence over 'sea,' thus implicating a form of hysteresis. We capture this hysteresis by deriving a simplified conceptual model. Our model discretizes the boundary layer using a spatial grid scale corresponding to that of typical cold pools and captures basics of convective dynamics, radiative effects and cold pool interaction, both under diurnal and constant-temperature surface conditions.

The paper is organized as follows. Section 3.2 has two chapters which each encompass methods and relevant results: in section 3.2.1 we introduce the numerical model used to simulate MCSs over tropical Africa and the adjacent Atlantic Ocean and observation dataset used to calibrate and validate the former; subsequently we introduce the methods to study self-aggregation and show the emergence of bi-stability in the transition from land to sea; In section 3.2.2 we introduce the conceptual model that can capture the bi-stability in self-aggregation from land to sea and its parameter space. In Section 3.3 we discuss the implications of our results.

3.2 Methods and Results

3.2.1 Idealized Simulations of Organized Convection

Numerical methods

Our numerical simulations (Table 3.1) are conducted using the System for Atmospheric Modeling, version 6.11 (Khairoutdinov and Randall, 2003). The model uses a single-moment microphysics scheme, the RRTM radiation scheme and a 1.5-order sub-grid scale closure scheme for subgrid turbulent processes. Surface fluxes are evaluated based on the Monin-Obukhov similarity. The computational domain consists of 480 × 480 grid points in the horizontal, with doubly periodic lateral boundary conditions, and a horizontal resolution of 1 km, respectively 500 m for the high-resolution simulations. We use 64 levels in the vertical, with a logarithmic scaling stretching from 50 m at the lowest level to 1000 m at the top of the model, which is set at 27 km and has a sponge layer above. We do not account for Coriolis forces. The vertical profiles are initialized using a tropical sounding without horizontal wind derived from ERA5 dataset, averaged over the Atlantic and tropical West Africa (5:10N,-40:10E) in the months of June-August 2020. For all 'ocean' simulations (OCEAN), constant surface temperatures $T_s(t) = T_0 = const$ conditions are prescribed domain-wide. As is appropriate for a water surface, the specific humidity at the surface is assumed to be at saturation. For 'land' simulations (DIU), we impose sinusoidally-oscillating surface temperatures $T_s(t)$, that is:

$$T_s(t) = T_0 - \Delta T \cos(2\pi t/t_0)\,, \tag{3.1}$$

where T_0 is the average surface temperature, ΔT is the surface temperature amplitude and $t_0 = 24\,h$ is the one-day period. For DIU, surface latent heat fluxes are reduced to 0.7 of those obtained for the (saturated) oceanic surface, in an effort to mimic an unsaturated surface and to keep the Bowen ratio from becoming too unrealistically low. The temporally-averaged Bowen ratio in the DIU simulation is 0.2 with this LAND set-up, and 0.1 for the OCEAN simulation. These Bowen ratios, calculated from the time-averaged latent heat flux and sensible heat flux over the whole simulations, are approximately representative of tropical rain forest and tropical ocean, respectively.

Finally, as this study focuses on the effects surface temperature variation have on CSA, we remove the top-of-the-atmopshere radiative cycle. It is set to a constant average incoming radiation calculated from July 1st, at 10 degrees North.

Table 3.1: **Summary of simulations.** Table 3.1 indicates the horizontal domain sizes L_x and L_y, the horizontal grid resolution Δx, the duration of each simulation, as well as the imposed surface temperature T_0 and its diurnal amplitude ΔT. We also provide the start time of each simulation. For the branch cases, the initial condition is the state of DIU at the start time specified.

Case name	$\mathbf{L_x}$ $[km]$	$\mathbf{L_y}$ $[km]$	$\mathbf{\Delta x}$ $[km]$	**Duration** $[days]$	$\mathbf{T_0}$ $[K]$	$\mathbf{\Delta T}$ $[K]$	**Start time** $[day]$
DIU	480	480	1	56	305	10	0
OCEAN	480	480	1	42	300	0	0
DIU hires	480	480	0.5	28	305	10	0
DIU lores	480	480	4	28	305	10	0
OCEAN hires	480	480	0.5	28	300	0	0
OCEAN lores	480	480	4	28	300	0	0
OCEAN warm	480	480	1	28	305	0	0
branches:							
DIU2OCEAN A_1	480	480	1	28	300	0	DIU 13.25
DIU2OCEAN B_1	480	480	1	28	300	0	DIU 13.5
DIU2OCEAN C_1	480	480	1	28	300	0	DIU 13.75
DIU2OCEAN D_1	480	480	1	28	300	0	DIU 14
DIU2OCEAN A_2	480	480	1	42	300	0	DIU 27.25
DIU2OCEAN B_2	480	480	1	42	300	0	DIU 27.5
DIU2OCEAN C_2	480	480	1	42	300	0	DIU 27.75
DIU2OCEAN D_2	480	480	1	42	300	0	DIU 28

Observational analysis

In an effort to design realistic tropical conditions for our idealised simulations, we first scrutinize satellite imageries over tropical continental Africa, where MCSs are systematically formed over land and advected over the adjacent Atlantic Ocean. In the month of July, which coincides with the beginning of the local monsoon season as well as the Atlantic hurricane season, there is a peak of MCS activity between the latitudes of 0 and 20 degrees N, meaning the MCSs caught in the African Easterly Jet have the longest possible travel time over land. To obtain information on the environment that MCSs are embedded within, in their transition from tropical African land to the adjacent ocean, we use a 5 year-long database of MCSs tracked from geostationary infrared satellite observations (TOOCAN, Fiolleau and Roca, 2013). In Figure 3.2a, we show the area of interest along with TOOCAN-tracked MCS cloud shields on July 14 2016. The evolution of groups of MCSs is followed using a Lagrangian frame.

To do this, we place a 15 × 15 degree frame at the time of maximum convective activity of the Eulearian frame (18UTC), that is in the same initial location over tropical West Africa (10:25 N, 6:21 E) as shown in Figure 3.2a. Then, we follow the movement of the tracked MCSs for every day in July for the 5 years of available TOOCAN data (2012-2016). The frame moves along with the average movement of the areal centers of mass of all the MCSs contained within it. When there are no MCS cloud shields present in the frame, e.g., as is common at night over land, the frame continues moving with the average velocity of the previous 3 time steps. The velocity vector of the frame is approximately always westward and the latitudinal component near-zero. While the frame is fully over land, we classify it in the land regime, while if it is partially over land and partially over the ocean, we classify it as a transition regime. Once the frame is fully over the ocean, we classify it as in the ocean regime. We find that the average westward velocity is larger in magnitude when the frame is over land, compared to when it is over the ocean (Figure 3.2b). The mean longitudinal velocities of the frame while it is over land is -2.8 m/s, and -1.7 m/s while it is fully over the ocean. Within the Lagrangian frame tracks, we record domain-averaged values of skin temperature, using ERA5 reanalysis data, along with the TOOCAN MCS fraction within the box. We superimpose all the tracks, take diurnal composites of these, of which we plot the 24 hour timeseries showing the first diurnal cycle from midnight to midnight over land and then over ocean, in 3.2c.

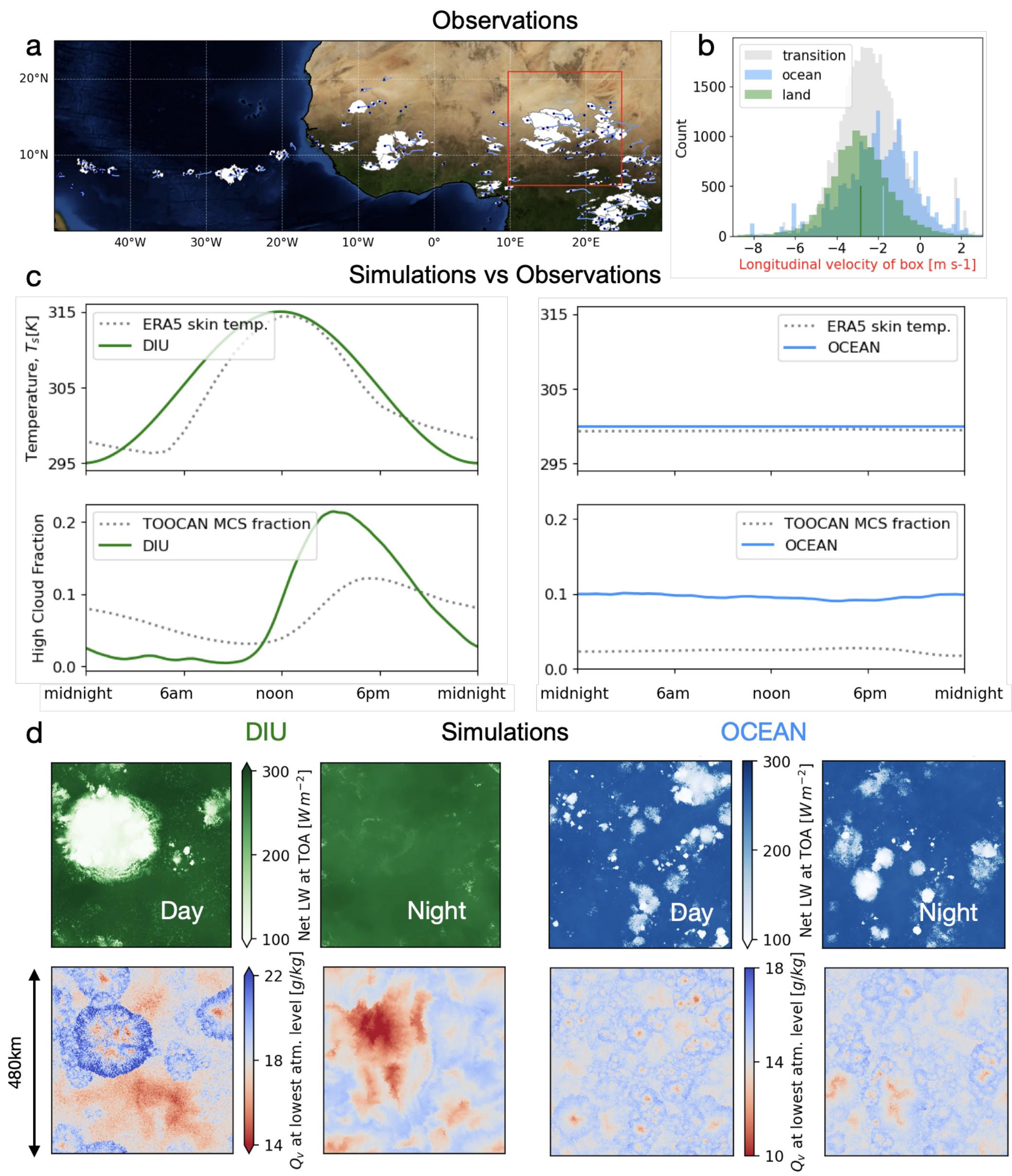


Figure 3.2: **Mimicking realistic diurnal cycle dynamics with idealised cloud resolving simulations. a**. Satellite-observed MCSs over tropical Africa (white areas) and 15 deg × 15 deg Lagrangian frame (thin red square); **b**. Longitudinal velocity of Lagrangian frame when over land, ocean, and in the transition zone. The mean values of the ocean and land velocity distributions are indicated by the small vertical lines. **c**. Imposed SSTs, and domain mean high cloud fraction. The latter is calculated within the simulation domain of DIU and OCEAN respectively. Dotted lines show domain-mean skin temperature (from ERA5 reanalysis) and MCS cover (from TOOCAN tracked MCSs), calculated within the Lagrangian frame that follows tracked MCSs across tropical African land and onto the adjacent Atlantic Ocean, in July 2012-2016. **d**. Representative instantaneous horizontal plots of outgoing longwave radiation, OLR, and near-surface specific humidity, q_v for DIU and OCEAN, as labeled.

From the Lagrangian tracks, we find a clear diurnal cycle in the frame-averaged skin temperature over land. It has an amplitude of approximately 10 K and oscillates from 295 K to 315 K. Conversely, the MCS cover over land has a clear diurnal cycle which lags behind the skin temperature diurnal cycle. Over the ocean, the frame-averaged temperature is constant at approximately 300 K. Compared to land, the MCS cover over the ocean is relatively low and does not exhibit a clear diurnal cycle. We note that the composites are calculated within a moving box, that moves towards the West, i.e. with the sun. This means that the diurnal cycles shown in Figure 3.2c have a slightly longer period than 24 hours.

Comparison between observational and numerical MCSs

We then seek to reproduce the observational findings in Figure 3.2c (dotted lines) using our idealised cloud resolving model output, calibrated with the aforementioned surface temperature time-varying conditions. To this end we calculate composite of consecutive one-day time series of the simulated variables. In Figure 3.2c, first row, we see the imposed SSTs for the DIU and OCEAN simulations, mimicking best the observations, and consequently the high cloud fraction, which should be compared with the TOOCAN MCS cover. In our idealised simulations, we do capture the diurnal cycle in deep convection seen over land, and a nearly constant deep convective time series over the ocean. However, the high cloud fraction in our simulations nearly reaches zero during the night hours, while MCSs on average persist through the night in the observations, likely aided by realistic phenomena such as wind shear which are not included in our idealised simulations. When investigating the other variables in the simulations, two striking differences are noticeable between DIU and OCEAN: 1) the patterns of the deep convective clouds and 2) their footprint in the cold pool field. In DIU, in the convective window of the afternoon, large organized deep convective structures appear that can easily take up a quarter of the domain, reaching sizes of MCSs (Figure 3.2 d). Below these MCS structures large cold pools are present — that is, cold patches of air with dry interiors and moist borders, as seen in the near surface specific humidity field in (3.2 d). These cloud structures disappear at night in the outgoing longwave radiation field, but persist in the moisture field as dry patches. In the OCEAN simulations, on the other hand, the deep convection remains relatively small and sparse around the domain, and there is no notion of "day" and "night." Finally, the collocated cold pools are also more homogeneous.

Further, we investigate the emergence of self-aggregation of convection in our suite of simulations. Similar to previous measures (Wing *et al.*, 2018) we use the normalized spatial variance of precipitable water,

$$Var(PW) \equiv \langle (PW - \langle PW \rangle)^2 \rangle / \langle PW \rangle^2 , \tag{3.2}$$

as a measure of aggregation of the moisture field. The pointed parentheses here indicate an average over all horizontal grid cells. The normalization ensures that overall changes in moisture do not appear as changes in variance. High values of $Var(PW)$ reflect a more aggregated field, since aggregated convection is characterized by the striking segregation of moist, strongly convecting regions embedded within an otherwise very dry and non-precipitating domain. We calculate $Var(PW)$ in each time step for all our simulations (Tab. 3.1).

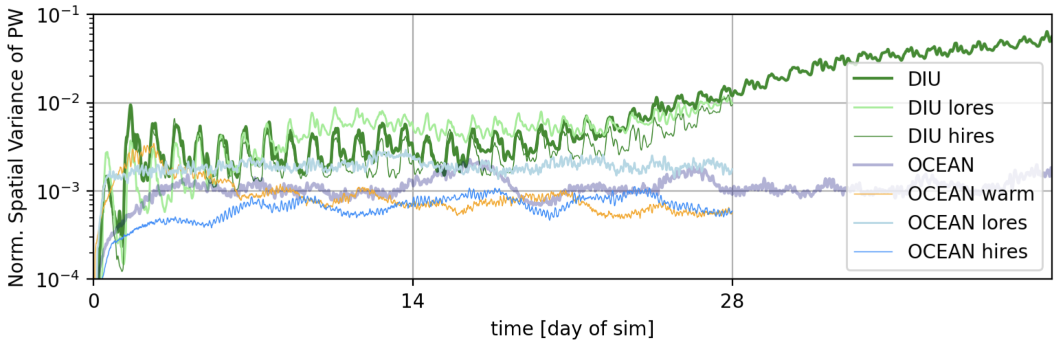


Figure 3.3: **Emerging self-aggregation of convection.** Normalized spatial variance (Eq. 3.2) of precipitable water (PW) for the cloud-resolving simulation results using the System for Atmospheric Modeling (SAM) as described in Tab. 3.1. Note that the DIU simulations (green curves) consistently show a systematic increase in PW over time whereas OCEAN (all other curves) consistently fluctuates around a state of low PW. As indicated in the legend, the simulations explore the sensitivity to numerical model grid resolution and SST in the case of OCEAN. Note the logarithmic vertical axis scaling.

In the first 14 days of simulation the main difference between DIU and OCEAN (Figure 3.3) is that DIU shows a large day-to-night variation, with peaks in $Var(PW)$ during the day, when the MCSs form, and minima during the night - where MCS-produced moisture anomalies slowly diffuse to even out (Figure 3.2d). However, during this time interval, the daily average $Var(PW)$ remains relatively constant for all simulations, even though it is already considerably larger for DIU. After day 14, DIU starts showing an increase in $Var(PW)$ whereas OCEAN continues to show near-zero $Var(PW)$. Repeating these simulations for two-folder higher and two-fold lower horizontal resolution shows qualitative similar results (Figure 3.3). We note that we ran an OCEAN warm simulation, to highlight that even with a higher SST that matches the average SST in DIU, the OCEAN simulation does not aggregate in this time frame (Figure 3.3).

Similar to Jensen *et al.* (2022), we mimic the land-sea transition of the deep convective cloud field, by branching off from the DIU simulation onto a constant SST (DIU2OCEAN). Here, however, we repeat the branching at different times of day (A: 6am, B: 12 noon, C: 6pm, D: 12am) and on different days of the simulation (days 14 and 28 (Figure 3.4). This allows us to study whether the diurnal phase of a branching — that is whenever a land-born MCS moves over the ocean — matters to its growth. Further, we also observe how MCS residence-time over land — akin to their "maturity" — matters to determine whether the system will aggregate over the ocean.

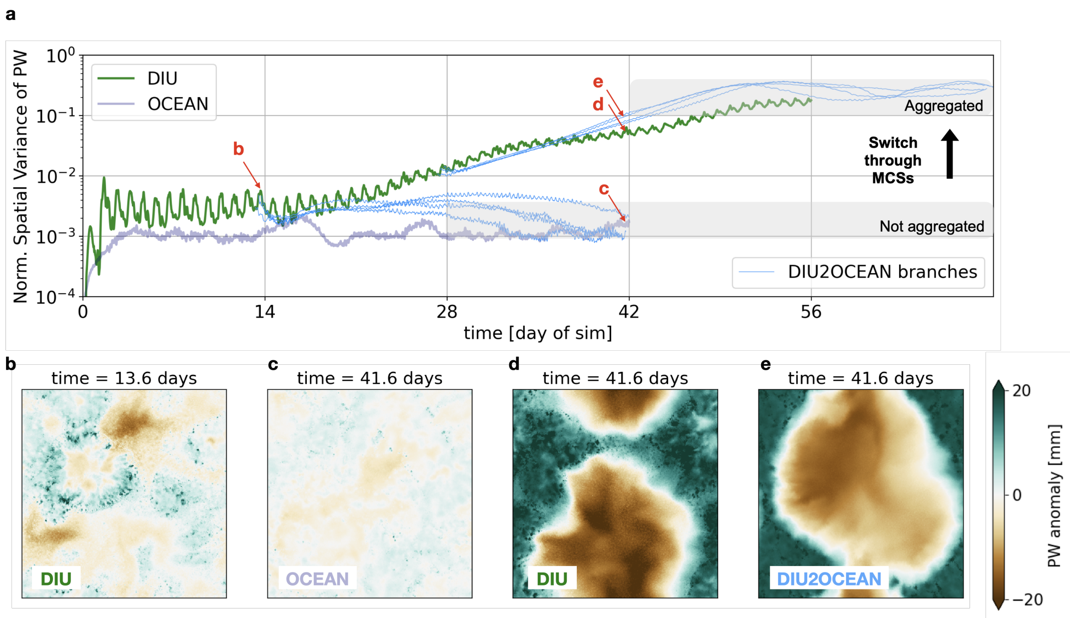


Figure 3.4: **Bistable switching through MCSs.** Analogous to Figure 3.3, however now allowing for a branching off between DIU and OCEAN simulations. DIU2OCEAN branches originate near day 14 and day 28, respectively and are shown in light blue colors. OCEAN simulations without branching are shown in light purple colors (see legend). Steady-state regions of high and low normalized PW variance are indicated by light gray shades. Insets b–e exemplify the spatial pattern of PW anomalies at specific times of the simulation (see red arrows and panel labels for the exact timing). Note again the logarithmic vertical axis scaling.

The effects of land-sea transition appears clearly when comparing the branches initialized from DIU day 14 (branches $A_1B_1C_1D_1$) to the branches initialized from DIU day 28 (branches $A_2B_2C_2D_2$). Branches $A_1B_1C_1D_1$ fail to maintain the degree of aggregation already acquired in the DIU simulations (Figure 3.4) and their $Var(PW)$ gradually returns to the values obtained by the original OCEAN simulation. Interestingly, the DIU2OCEAN D1-branching (occurring at midnight) does appear to resist the change better than the other phases, suggesting that MCSs crossing the coastline at this time of the day have better chances of strengthening over the ocean. On the other hand, branches $A_2B_2C_2D_2$, proceed to aggregate further and even more rapidly than the continuation of DIU (Figure 3.4), reaching a fully aggregated state with characteristic low frequency fluctuations, similar to those seen and described in Patrizio and Randall (2019). This time, the time-of-transition appears not to the matter, and all trajectories behave similarly, suggesting that the MCS was already strong enough over the land to quickly dissipate over the ocean. Thus, we can conclude that the system shows bi-stability, with one low and one high spatial variance state. A transition between the low and the high variance state can be achieved by diurnal surface temperature forcing and the resultant MCS, but not through constant temperature lower boundary conditions.

Notably, our current simulations differ from those in Jensen *et al.* (2022) in that we use a different CRM, namely SAM vs UCLA-LES in their study. Still, our results are in agreement with Jensen *et al.* (2022), as they also suggest that the diurnal cycle in surface temperature can induce persistent moisture patterns that profoundly impact on the multi-day precipitation distribution. Thus, our simulations add additional robustness to the mechanism proposed in their work. On a time scale of hours, simply imposing a realistic diurnal cycle in the surface temperature of a cloud resolving model can produce large MCS-like deep convective structures during the afternoon hours. These structures do not develop at these time scales in simulations with constant surface temperature conditions, which rather mimic the observed oceanic deep convection that occurs irrespective of day and night and has smaller spatial extent. On a time scale of days, the land-like simulations furthermore develop a typical self-aggregated state in the moisture field, where parts of the domain become too dry to support any new convection, whereas such persistent drying does not develop at these timescales in the ocean-like simulations. Ocean-like simulations are able to maintain the aggregated moisture field, once this is sufficiently developed in land-like simulations. The fact that an earlier branching after 14 days does not yield persistent aggregation suggests that a substantial perturbation away from a linearly stable, homogeneous RCE state is needed.

Bringing our results into a realistic context, if we consider self-aggregation to be present to some extent in the real atmosphere, and land-produced MCSs to be a path to self-aggregation, then the residence time of a volume of atmosphere over land is an important factor in determining the maintenance of an aggregated state once it is advected over the ocean. This could mean that MCSs — produced over tropical African land — could be the key player needed in pushing the moisture distribution of the volume of atmosphere they are embedded within to an aggregated state, which might then be able to persist over the ocean — given that spatial moisture variance is high enough. This could be a so-far-missing ingredient for tropical cyclogenesis from the CSA point of view.

3.2.2 Conceptualizing the bi-stability of self-aggregation

We implement a conceptual cellular automaton model, the "Game of Cloud", to study the basic physical mechanisms that can drive diurnal self-aggregation and the resulting bi-stability.

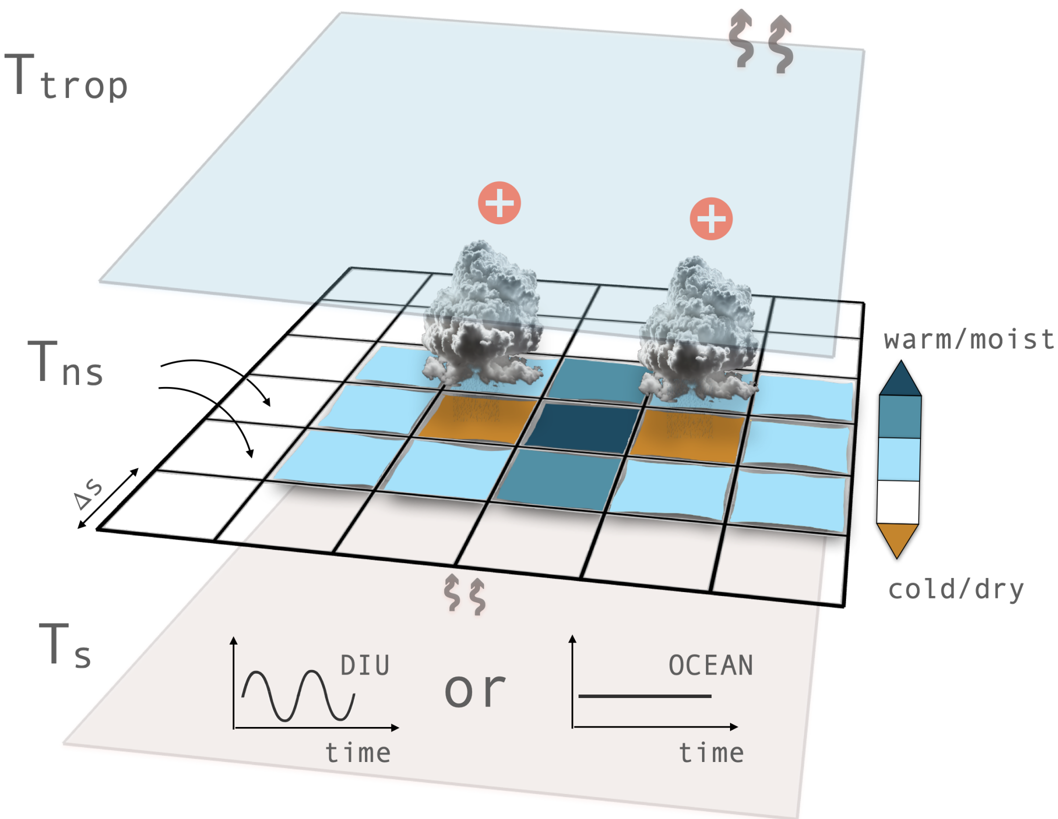


Figure 3.5: **The Game of Cloud.** Conceptual schematic of our model to study CSA developing from MCSs. From bottom layer to the top: surface temperature, T_s, is either prescribed as oscillating with a diurnal frequency (land) or constant (ocean); near-surface temperature, T_{ns}, is discretized on a 2D grid, where each grid cell assumes a value based on T_s, caused by surface heat fluxes, its neighbors, due to diffusion, and boundary-layer CP dynamics. Tropospheric temperature, T_{trop}, assumes one value for the entire troposphere. When $T_{ns} > T_{trop}$, deep convection can arise in the corresponding column, which triggers CP dynamics, decreasing T_{ns} beneath the rain cell and enhancing T_{ns} for its neighboring grid cells, and increases T_{trop}. T_{trop} gradually cools at a set rate.

We demand that the model captures: (i) the spatial segregation effect which is seen in DIU but is absent in OCEAN; (ii) long-term persistence for DIU; (iii) the hysteresis effect described for the DIU2OCEAN branches; (iv) realistic domain-mean diurnal cycles of convective rain. The model assumes three layers and describes the dynamics of buoyancy in each of the three. We use the term "temperature" in the following to refer to an appropriate measure of buoyancy for the moist atmosphere, such as entropy. We distinguish such temperature within (see also Figure 3.5): (i) the surface, which has a spatially-homogeneous and prescribed temperature $T_s(t)$ and is assumed to be at saturation, thus representing a water surface or a swamp. (ii) the near-surface atmosphere temperature $T_{ns}(x, y, t)$, where we retain a horizontal spatial variation, as described further below; (iii) the free-tropospheric temperature, which responds to $T_{ns}(x, y, t)$ but acts as a spatially equilibrated field $T_{trop}(t)$, such that it remains spatially-homogeneous.

Spatial and temporal discretization. The model uses for horizontal coordinate system (x, y), a uniformly-discretized square domain with periodic boundary conditions. A spatial resolution of $\Delta s \equiv 10\,km$ is chosen, which represents a typical spatial scale of isolated deep convective raincells and their individual cold pools. Therefore, each discrete coordinate $(x, y) = (x_i, y_j)$ simply corresponds to the position $(x, y) = (i, j)\Delta s$, where i and j take integer values between 0 and $N-1$, where N is the number of pixels in each horizontal dimension. The near-surface temperature $T_{ns}(x, y, t)$ is defined at the discrete spatial coordinates described above. The typical temporal scale is assumed to be $\Delta t \equiv 30 min$, representing the period between triggering of a given raincell and the initiation of the resultant rain event. For both model variables, T_{ns} and T_{trop}, the dynamics therefore evolve in discrete time steps of Δt, that is, each time point is defined as $k\Delta t$, where $k = 0, 1, \dots$ is an integer.

Boundary conditions. As explained before, the surface temperature, $T_s(t)$, is spatially homogeneous and prescribed as the harmonic function defined in Eq. 3.1. It is represented in the bottom layer of the schematic in Figure 3.5. Importantly, $T_{trop}(t)$ is globally increased by any deep convective event — wherever it occurs — which is assumed to release a fixed quantity c_0 of latent heat of condensation (the plus signs in Figure 3.5). Further, the troposphere continuously cools under radiative emission at a constant rate $R \approx 3K\,day^{-1}$, as a simple approximation of the daily averaged tropospheric cooling that one could expect in the tropics (Jeevanjee and Fueglistaler, 2020; Fildier *et al.*, 2023). The instantaneous rate of change of T_{trop} reads as:

$$\frac{dT_{trop}(t)}{dt} = C(t) - R, \tag{3.3}$$

where $C(t)$ would be the convective heat flux into the free troposphere at time t. In our discretized model dynamics, Eq. 3.3 becomes

$$\Delta T_{trop} = nc_0 - R\Delta t. \tag{3.4}$$

Equation 3.4 is applied at every timestep Δt, and the integer n counts the number of convective events within the given timestep. It is assumed that the horizontal domain size is small enough such that the weak temperature gradient approximation applies (Sobel *et al.*, 2001), thus justifying the assumption of a spatially homogeneous field $T_{trop}(t)$.

Near surface conditions and diffusivity. For the near-surface atmospheric (virtual) temperature, $T_{ns}(x,y,t)$, the continuum version of the basic model dynamics can be summarized as a prognostic equation, namely

$$\begin{aligned}
\frac{d}{dt}T_{ns}(x,y,t) &= \underbrace{\frac{1}{\tau}(T_s(t) - T_{ns}(x,y,t))}_{\text{surface-atmosphere heat flux}} + \underbrace{D_h\nabla^2 T_{ns}(x,y,t)}_{\text{horizontal diffusion}} && (3.5)\\
&+ \underbrace{\mathscr{R}(x,y,t;\{x_i\},\{y_j\},\{t_l\})}_{\text{reactive processes}} && (3.6)\\
&+ \underbrace{b_1\, T'_{ns}(x,y,t)(1 - b_2 T'^2_{ns}(x,y,t))}_{\text{moisture-radiation-circulation feedback}}, && (3.7)
\end{aligned}$$

In Eq. 3.7, $T'_{ns}(x,y,t) \equiv T_{ns}(x,y,t) - \overline{T_{ns}}(t)$, with $\overline{T_{ns}}(t)$ the spatial mean of $T_{ns}(x,y,t)$ at time t. The surface-atmosphere heat flux is a diffusive flux generated by the temperature difference between the surface and the atmosphere, and modulated by the time scale τ. Wind speed effects on surface-atmosphere heat fluxes are neglected for simplicity. The horizontal diffusion is implemented to capture all horizontal mixing processes taking place through eddy-diffusive motion within the boundary layer.

Positive feedback mechanism. The term in Eq. 3.7 describes a moisture-radiation-circulation feedback, mimicking the "rich-gets-richer" dynamics invoked in previous works on convective self-aggregation (Bretherton *et al.*, 2005; Wing *et al.*, 2018; Muller *et al.*, 2022). In effect, the resultant action of sustained subsidence is to dry the near-surface atmosphere, thus leading to further drying in regions that are already dry. Conversely, we allow moist regions to moisten further. The positive coefficients b_1 and b_2 in Eq. 3.7 prohibit unbounded decrease or increase of buoyancy, thus ensuring a naturally occurring maximum or minimum in buoyancy. In our cellular automaton model Eqs 3.5—3.7 are discretized in spatial and temporal steps of Δs and Δt to match the typical convective scales assumed.

Triggering of convection. We now turn to the reactive term $\mathscr{R}$ (Eq. 3.6), which introduces a measure of stochasticity into the model. $\mathscr{R}(x, y, t; \{x_i\}, \{y_j\}, t)$ models the threshold-like triggering of a convective event at any discrete location (x, y) at time t. The reactive term consists of two parts of which the first allows individual locations (x, y) to interact with other locations through the "mean field" $T_{trop}(t)$: during a given timestep Δt, all locations (x, y) with $T_{ns}(x, y, t) > T_{trop}(t)$ are considered potentially unstable. To each of these locations (x, y) we therefore assign a triggering probability

$$p(x, y, t) \propto T_{ns}(x, y, t) - T_{trop}(t) > 0\,, \tag{3.8}$$

which assumes that locations with buoyancy exceeding that of the free troposphere may be unstable to convection. In a random procedure, we now sequentially draw locations from the list of unstable locations according to their probabilities $p(x, y, t)$ and for each of them transfer condensation heat c_0 to the free troposphere, thus, for each of them, increasing $T_{trop}(t)$. We therefore then update all $p(x, y, t)$ as the increase in $T_{trop}(t)$ will often have stabilized a number of the previously unstable locations. During the given timestep this stochastic procedure is repeated until all locations are stable.

The second part of $\mathscr{R}$ models the suppression and activation of new convective raincells through cold pools. Cold pool effects are incorporated in a two-fold manner, reflecting known physical processes: at each timestep t, the temperature $T_{ns}(x, y, t)$ at any given location (x, y) which was occupied by raincells at time $t-\Delta t$ is lowered by ΔT_{cp}, mimicking the well-known near-surface evaporative cooling experienced beneath deep convective clouds. Further, to mimic thermodynamic and mechanical triggering at the edges of existing cold pools, such as by forced lifting or collision effects, the $T_{ns}(x, y, t)$ at any such location (x, y) is incremented by an enhancement that is proportional to the number of surrounding locations that had active rainfall at $t - \Delta t$, where the 8-neighborhood of the location (x, y) is used to define the neighborhood. This weighting ensures that locations surrounded by one or several cold pools are more likely to experience convective triggering. We note here that a rainy pixel is considered "active" for two time steps, that is, one hour, to increase the chance of nearby raincells to interact with each other.

In a cellular automaton step, denoted by an arrow below, a temperature enhancement is applied specifically to any pixel (x, y) that is inactive (not raining) at time t as:

$$T_{ns}(x, y, t) \to T_{ns}(x, y, t) + I\,|\Delta T_{cp}|\,\overline{T}_{ns}(t) \cdot n/8 \tag{3.9}$$

where n is the number of active neighbors of the pixel (x, y) at time t. The proportionality with $\overline{T}_{ns}(t)$ ensures that the triggering effect does not dominate when $\overline{T}_{ns}(t)$ is low, such as at night for DIU, or generally for OCEAN, and the proportionality with ΔT_{cp} ensures that the triggering effect is stronger (weaker) for strong (weak) cold pools when changing ΔT_{cp} (as in Figure 3.8).

Conceptual model properties

With the parameters chosen (Tab. 3.2), our model captures the mean diurnal cycle of convection of land and sea (Figure 3.6) by imposing $\Delta T = 10K$ for land and $T_0 = 305\,K$ for land and ocean.

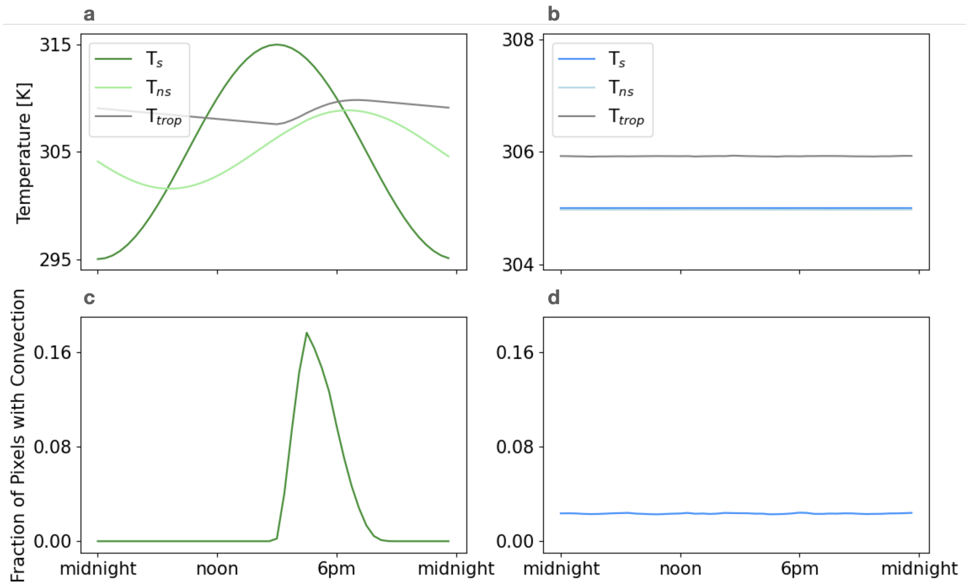


Figure 3.6: **Diurnal time series in the Game of Cloud. a**, Diurnal cycle of temperatures T_s, T_{ns} and T_{trop} for the DIU configuration. **b**, As (a) but for OCEAN. Note that $T_{ns} = T_s = 305\,K$ here. **c**, **d**, Convective precipitation diurnal cycle for DIU and OCEAN, respectively.

Table 3.2: **Parameters used in the Game of Cloud.** Table indicates the parameters used in the conceptual model, including the mathematical symbol, the numerical value and notes on the respective physical meaning.

Symbol	Numerical value	Notes
Δs	10 km	spatial scale of individual deep convective events
N	50	linear dimension in units of Δs (grid size)
Δt	30 min	time step approximating a convective life cycle
ΔT	10 K	diurnal surface temperature amplitude
ΔT_{cp}	-2 K	cooling induced by a CP
D_h	100 $m^2\,s^{-1}$	horizontal eddy diffusion coefficient
τ	10 hours	surface-atmosphere heat diffusion time scale
I	0.7 K^{-1}	scaling coefficient for CP triggering of convection
b_1	0.06 s^{-1}	moisture-radiation feedback coefficient 1
b_2	1 K^{-2}	moisture-radiation feedback coefficient 2
c_0	$4\,K/\Delta s^2$	tropospheric heating due to a single convective event
R	$3\,K\,day^{-1}$	free tropospheric diurnal cooling rate

We compute the normalized spatial variance of T_{ns} (Figure 3.7), which shows constantly small values throughout the timeseries for the $\Delta T = 0$ case, whereas $\Delta T = 10\,K$ yields a systematic increase over the course of 28 days and subsequent saturation. When switching to $\Delta T = 0$ after 28 days of $\Delta T = 10\,K$, our model also

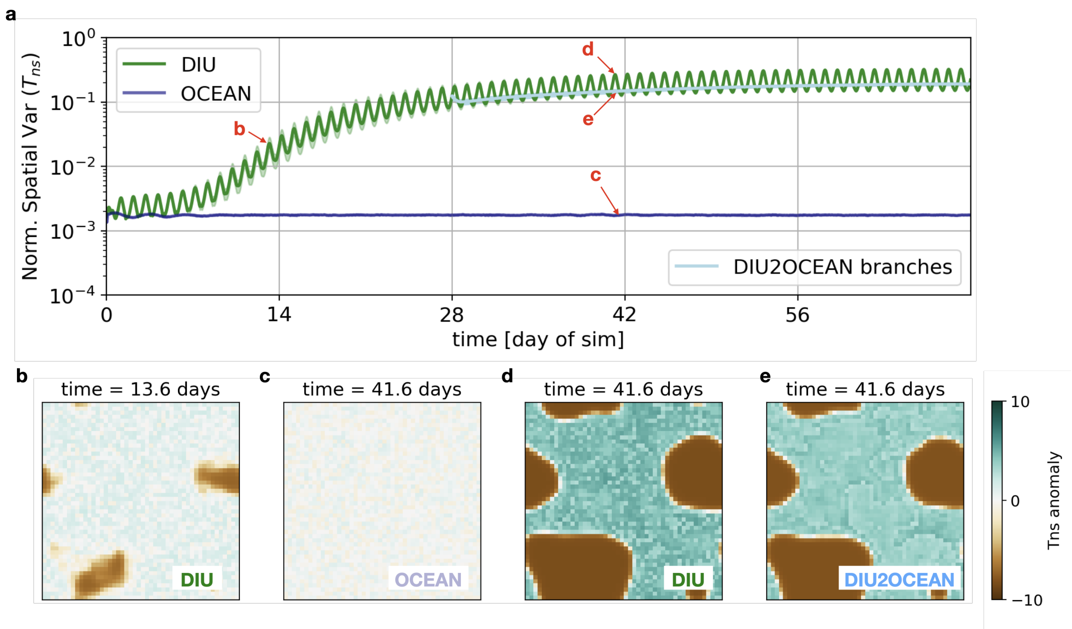


Figure 3.7: **Bistable switching through MCSs in the Game of Cloud.** Conceptual model analog of Figure 3.4 showing the normalized spatial variance of near surface temperature, T_{ns}, for DIU, OCEAN and DIU2OCEAN branches originating near day 28. Light shading indicates spread over ten simulations. Insets b–e show the spatial pattern for specific times and simulations. Note again the logarithmic vertical axis scaling.

shows that variance is preserved, thus capturing the hysteresis effect. Examining spatial patterns at day 42 reveals that T_{ns} is structured into mesoscale clusters for the oscillating case, whereas T_{ns} remains scattered for the constant-T_s counterpart.

Qualitatively, the conceptual model dynamics compares well to the variance of precipitable water for the SAM simulations, as the spatial segregation effect seen in DIU but not in OCEAN is reproduced, long-term persistence for DIU is captured and the hysteresis effect described for the DIU2OCEAN branches is mimicked. Furthermore, the diurnal cycles of convective rain are realistic (Figure 3.6).

Parametric sensitivity study

In an effort to better illustrate the model's driving mechanisms, a parametric study is carried out within the $(\Delta T, \Delta T_{CP}, D_h)$-space. As a reminder, these parameters respectively quantify 1) the diurnal cycle amplitude - which determines the surface temperature forcing regime (DIU/OCEAN) - and 2) the cold pool cooling strength and 3) the rate at which horizontal temperature is homogenized. These parameters are varied sequentially two-by-two while the third one retains its default value, as presented in Sec. 3.2. The parametric space was chosen as wide as possible while retaining physically-relevant bounds. Further, each parameter was varied 50 times, resulting in a total of 7,500 configurations, which all ran for 100 model days. Finally, over the course of these 100 days, two metrics were evaluated to quantify the model's response to a particular $(\Delta T, \Delta T_{CP}, D_h)$-solution: 1) the spatial variance of T_{ns} on day 100 and 2) the day when this quantity reached a threshold of 0.05.

The resulting contour maps are presented in Figure 3.8, with the default values represented with dashed red lines. The OCEAN configuration is marked by a dashed blue line, while the DIU configuration studied above is marked by a green dashed line.

Diurnal amplitude ΔT. An increase in ΔT consistently enhances both CSA metrics (Figure 3.8a—d): large ΔT overcomes the anti-aggregating dynamics caused by a high horizontal diffusion or strong cold pools. A ΔT_{CP} of 1.5 K can consistently prevent aggregation until the diurnal cycle amplitude reaches about 6 K, at which point it seems to be the dominating organizational driver (Figure 3.8d).

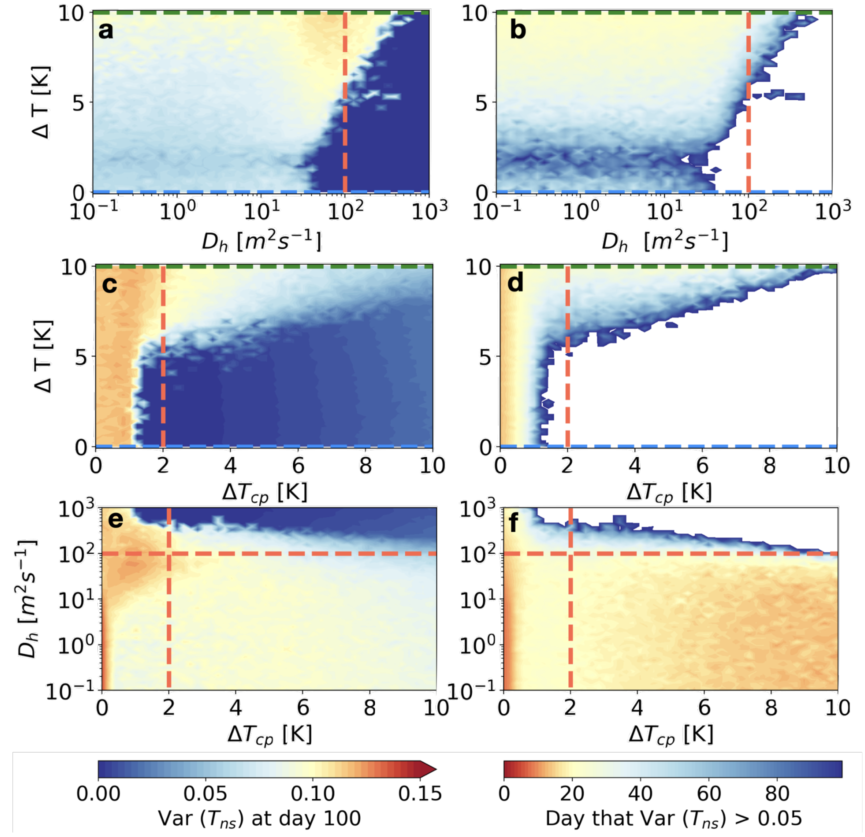


Figure 3.8: **Exploring the parameter space of the Game of Cloud.** Left column indicates the level of aggregation (measured by spatial variance of T_{ns}) as a function of parameter space after 100 days of simulation. Right column indicates day of the simulation that the level of aggregation has surpassed the value of 0.05. The parameter space explored comprises the diurnal cycle ΔT, cold pool strength ΔT_{CP} and horizontal diffusion D_h. Green dotted lines indicate the chosen ΔT for DIU, blue dotted lines indicate the chosen ΔT for OCEAN, red dotted lines indicate the chosen parameters for ΔT_{CP} and D_h in our default runs.

This is consistent with findings from Jensen *et al.* (2022), where strong surface temperature amplitude would trigger convective aggregation. Importantly, the ($\Delta T = 0$)-case - corresponding to OCEAN - is still capable of aggregating under specific circumstances, such as *weak* cold pools (*i.e.* low ΔT_{CP}) or little horizontal diffusion (*i.e.* low D_h). This captures findings from a range of RCE simulations (Wing *et al.*, 2018).

Cold pool cooling strength $\Delta \mathbf{T_{CP}}$. Strikingly, $\Delta T_{CP} = 0$ K - corresponding to no cold pool dynamics - invariably leads to aggregation, regardless of the other parameter values. The de-aggregating force of cold pools is consistent with the observation from idealized RCE studies (Jeevanjee and Romps, 2013; Muller and Bony, 2015; Nissen and Haerter, 2021) where disabling rain evaporation in the lower levels - thereby removing cold pools - enhanced CSA. Conversely, increasing ΔT_{CP} progressively overwhelms the other parameters' importance, suggesting that the de-aggregating force increases with cold pool strength. This is also consistent with the work of Meyer and Haerter (2020), who observed how both the size and propagating velocity of cold pools increases with the initial temperature reduction. Interestingly, stronger cold pools may speed up aggregation at low horizontal eddy diffusivity and in the presence of a diurnal cycle (Figure 3.8f). This, again, echoes the recent work of Jensen *et al.* (2022) who observed numerically how strong *macro*-cold pools could contribute to a form of self-aggregation by forming persistently dry regions.

Horizontal eddy diffusivity $\mathbf{D_h}$. Strong horizontal diffusion is capable of completely shutting down aggregation through brute-force damping of any spatial heterogeneities (Figure 3.8a,b,e,f). Overall, our model's response to enhanced D_h is similar to that of Biagioli and Tompkins (2023), for the evolution of column-integrated relative humidity. They also observe that horizontal diffusivity acts to counter-balance CSA by redistributing moisture away from rainy clusters into dry regions.

In summary, the model is shown to behave consistently with more complex and idealized studies of CSA. However, we note that all our simulations are deliberately idealized, to reveal the key mechanisms involved in the hysteresis-like effect of induced self-aggregation by the diurnal forcing. Important factors, such as wind shear, the Coriolis force, two-way coupling to the land or water surface, as well as topographic forcing and land-sea breezes, are left out and should be explored in follow-up works.

3.3 Conclusions

We have simulated tropical land-like MCSs by incorporating a realistic diurnal cycle in surface temperatures within a cloud resolving model, and we mimic the transition to the ocean by removing the diurnal cycle for ocean-like deep convection. Our results suggest that the diurnal cycle in surface temperature can induce persistent moisture patterns that profoundly impact the multi-day precipitation distribution, through the presence of MCSs, consistent with the recent literature (Jensen *et al.*, 2022). On a time scale of hours, simply imposing a realistic diurnal cycle in the surface temperature of a cloud resolving model can produce large MCS-like deep convective structures during the afternoon hours. These structures do not develop under constant surface temperature conditions, mimicking oceanic deep convection that occurs irrespective of day and night and has smaller spatial extent. On a time scale of days, the land-like simulations furthermore develop a typical self-aggregated state where parts of the domain become too dry to support any new convection, whereas such persistent drying does not develop at these timescales in the ocean-like simulations.

Expanding on the finding of hysteresis in cloud resolving simulations transitioning from a diurnal cycle to none (Jensen *et al.*, 2022) we here explore branching from the land-like to ocean-like simulations at different days and find that ocean-like simulations are able to maintain the aggregated moisture field only once this has reached a certain level of aggregation. To deepen our understanding of the key processes acting in these simulations, we develop a discrete reaction-diffusion-type model, characterized by three key parameters: the diurnal cycle, cold pool strength, and horizontal diffusion. This model successfully captures the bi-stability observed in the transition from dispersed to aggregated states.

We conclude that the MCSs produced over land can induce domain-wide CSA in the time scale of weeks, which persists even in ocean-like conditions that would not typically aggregate on these timescales, i.e., when the system has been "tipped" towards an aggregated state by the influence of diurnal cycle-induced MCSs. We thus propose a regime diagram (Figure 3.9). There are two stable equilibria that the volume of atmosphere (the grey ball in Figure 3.9 a and b) can be in: the "Not aggregated state" and the "Aggregated" state, identifiable by proxies like the spatial variance of precipitable water that we used in our study, as a function of the amplitude of the diurnal cycle in surface temperatures (ΔT). The "Not aggregated" state is the equilibrium state for low ΔT, such as over the tropical ocean. The "Aggregated" state is the equilibrium state for high ΔT, such as over tropical land. The blue ball can be transported from the "Not Aggregated" equilibrium to the "Aggregated" state via the action of MCSs that can help overcome the barrier.

It is also useful to understand the relation to the seminal model by Emanuel *et al.* (2014). There, classical convective self-aggregation was proposed to be the departure from a linearly stable regime. In their model, the spatially homogeneous steady state is linearly stable at low sea surface temperatures but is unstable at higher sea surface temperatures. Thus, as temperature is increased above a certain critical value, the linearly stable regime disappears at the benefit of a structured state, where a system-scale subdomain shows subsiding conditions, that is, negative vertical velocity w, whereas the remainder of the domain shows strongly convecting conditions with $w > 0$ (Figure 3.9 c). Our current results are obtained at relatively low sea surface temperatures where the cloud resolving model is linearly stable, that is, there is no classical CSA. The diurnally-induced mesoscale organization however allows the transition to a persistently organized state. We thus propose that the boundary layer moisture dynamics, disregarded in the elegant model by Emanuel *et al.* (2014), are key to capturing the departure from the linearly-stable regime proposed in their work. Indeed, it is the strong feedback between cold pool dynamics and the subsequent, long-lasting suppression of convection in mesoscale areas that allows for the tipping to take place.

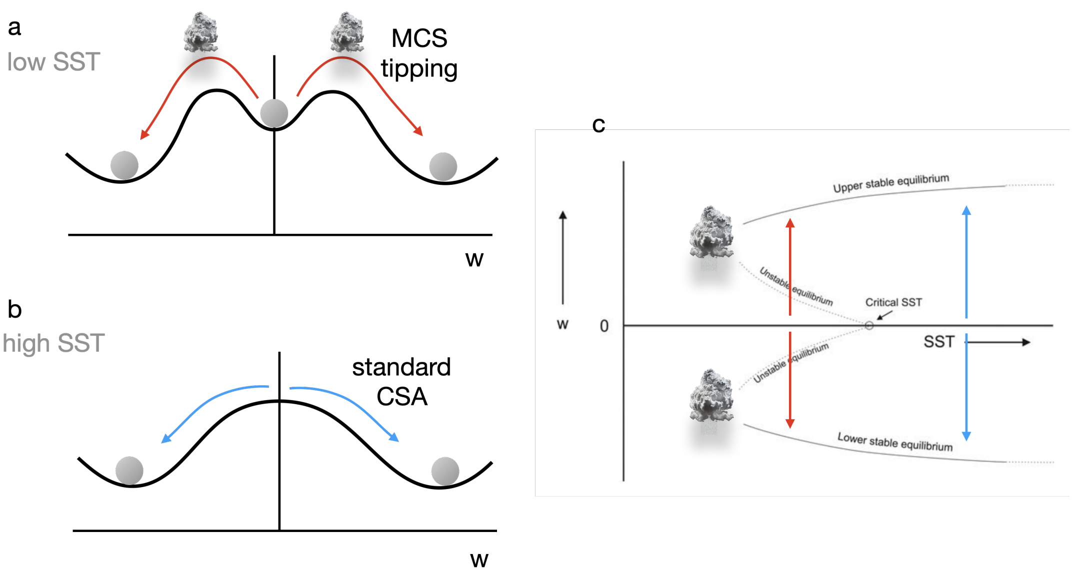


Figure 3.9: Schematic regime diagram for equilibrium states in (a) low SST regime, (b) high SST regime and their relation to (c) Figure 7 from Emanuel *et al.* (2014).

While we suggest that land-produced MCSs are the process by which a tropical atmospheric volume can "tip" to an aggregated state, the present work still leaves open which exact process within MCSs determines the emergence of persistent drying — and thus CSA. Haerter *et al.* (2020) suggested that so called 'diurnal self-aggregation' occurs with a large enough diurnal cycle in surface temperature, due to the higher spatial density of rain cells within a short time window during the diurnal cycle, leading to the merging of convective cold pools (CPs) and the formation of a super CP. This larger CP triggers a cascade of convective rain cells along its gust front, contributing to the combined CP, whose thermodynamic anomalies dissipate slowly, subsequently suppressing further convection in the same region on subsequent days, creating large dry areas that shut off convection and kick starting the radiative feedback. An alternate explanation could be that there is large subsidence around an MCS, which could, in principle, also create a large dry area that could first form in the higher levels of the troposphere, as seen in Jensen *et al.* (2022), then extend to the full atmospheric column and kickstart the radiative feedback.

In either case, the origin of CSA would be intimately tied to the presence of MCSs. Yet, as we cannot pinpoint the cause of the "first dry region" in our simulations, we suggest this aspect to be explored further. Assuming that CPs of MCSs over land initially contribute to *accelerating* the onset of CSA, we point out that once the system is aggregated, these MCS-scale CPs act to redistribute moisture and *oppose* the radiative feedback over land by counteracting the radiative dry pool in the boundary layer (Yanase *et al.*, 2020) and slow down self-aggregation. When the aggregated state then transitions to constant SSTs mimicking an ocean surface, the disappearance of MCSs replaced by isolated deep convection and their associated small CPs, cannot counteract the radiative dry pool - so at this stage the small sparse CPs are not an obstacle to aggregation at all. This is an explanation to the slowdown of the DIU curve in Figure 3.4 as opposed to the DIU2OCEAN branches, after about six weeks of simulation. The dual effect of CPs based on their size is not included in our conceptual model and could be further explored.

Follow-up work could include incorporating a westerly flow and wind-shear, to explore the effect of realistic wind profiles on the self-aggregated state transitioning from land to sea. A complementary observational study to ours could also be explored, to detect the footprint of CSA (as in high variance in precipitable water, or large dry areas in the upper troposphere) around MCSs created over land, to see if these exist and if they persist when advected westward.

Our CRM simulations as well as the conceptual model suggest that land-produced MCSs can be a path to self-aggregation and that the residence time of a volume of atmosphere over land is crucial in determining the maintenance of an aggregated state once it is advected over the ocean. MCSs — produced over tropical African land — may thus carry the moisture feedback mechanisms needed in eventually yielding an aggregated state which might then be able to persist over the ocean — given that spatial moisture variance is high enough. In contrast to the classical CSA mechanisms alone, where the process of full aggregation is rather slow, MCS-based aggregation may be the missing ingredient for more speedy tropical cyclogenesis.

3.4 Supplement

This section was not included in the manuscript, but serves as a suplementary analysis of the incorporation of realistic wind shear in the previously described cloud resolving simulations. Two extra simulations are described in Tab. 3.3, analogous to the DIU (OCEAN) simulations in Tab. 3.1 with (without) an imposed diurnal cycle in surface temperature, but with a wider domain and an imposed wind shear from realistic ERA5 reanalysis data averaged over the Atlantic and tropical West Africa (5:10N,-40:10E) (Figure 3.10). We call these simulations DIU Wind and OCEAN Wind.

Case name	$\mathbf{L_x}$ $[km]$	$\mathbf{L_y}$ $[km]$	$\mathbf{\Delta x}$ $[km]$	**Duration** $[days]$	$\mathbf{T_0}$ $[K]$	$\mathbf{\Delta T}$ $[K]$
DIU Wind	1440	480	1	42	305	10
OCEAN Wind	1440	480	1	42	300	0

Table 3.3: **Summary of simulations.** Horizontal domain sizes L_x and L_y, the horizontal grid resolution Δx, the duration of each simulation, imposed surface temperature T_0 and its diurnal amplitude ΔT.

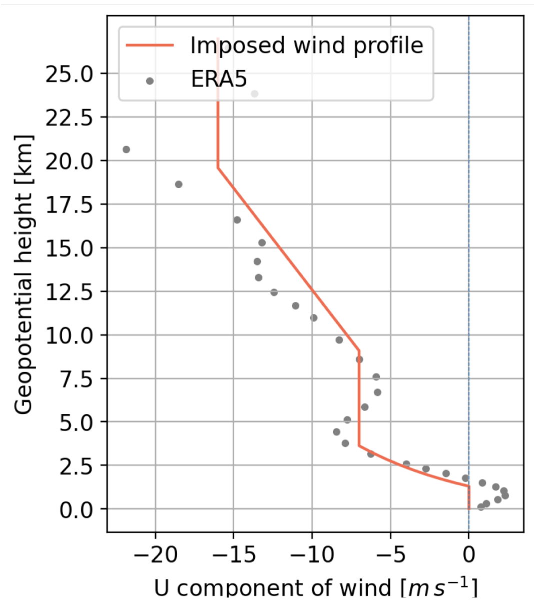


Figure 3.10: **Idealised westward wind shear profile**. ERA5 data averaged over the Atlantic and tropical West Africa (5:10N,-40:10E) over the month of July 2016, idealised wind profile (red) imposed in the simulations DIU Wind and OCEAN Wind from Tab. 3.3.

We first compare the simulated diurnal cycle of deep convection in the simulations with imposed wind and wind shear in Table 3.3, to the simulations without wind described earlier and observations, in Figure 3.11. We include a measure of precipitation from the simulations and from the observational Lagrangian frame (the latter is calculated from IMERG satellite data (Huffman *et al.*, 2019)). We notice that the inclusion of wind shear flattens the peak of MCSs in DIU Wind compared to DIU, moving the peak to later afternoon hours, and increases the night-time minimum, to a value that indicates a persistence of deep convection through the night thanks to wind shear. The same behavior can be seen for precipitation, lowering the maximum and shifting it forward in time, closer to the observed values. In the OCEAN Wind simulation, the high cloud fraction is lowered compared to the OCEAN simulation towards the observed value of MCS fraction, and the precipitation stays close to the observed value. We are thus towards a more realistic representation of convection even within a very idealised set up, thanks to the addition of imposed (realistic) wind shear.

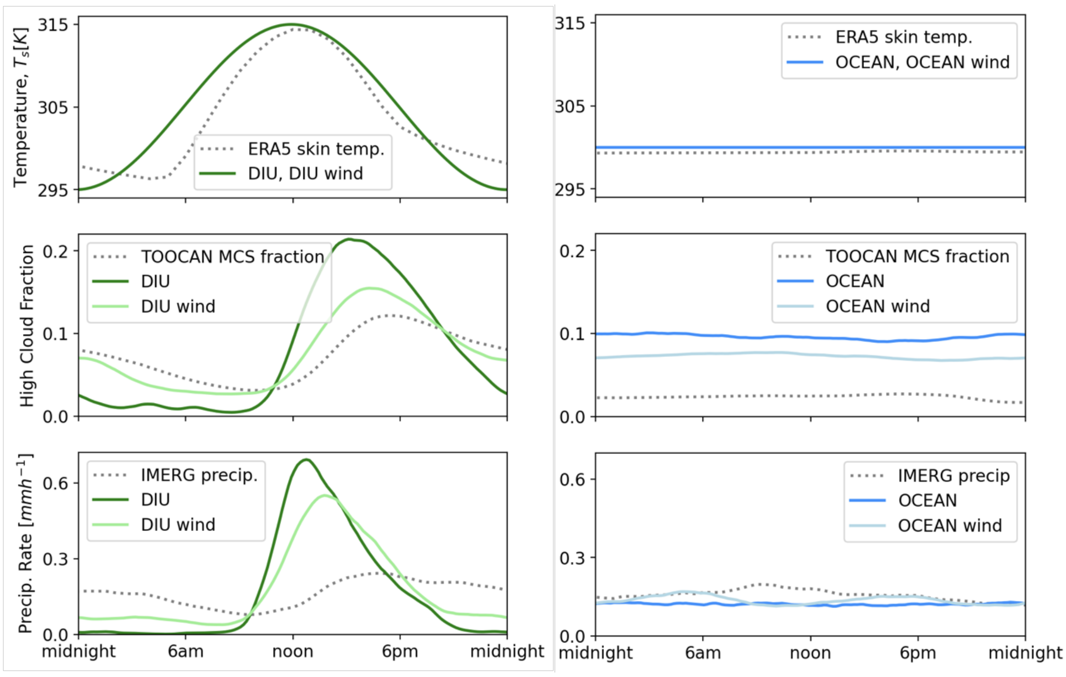


Figure 3.11: **Mimicking realistic diurnal cycle dynamics with idealised cloud resolving simulations.** Imposed SSTs, and domain mean high cloud fraction and precipitation. The latter two are calculated within the simulation domain of DIU and OCEAN respectively. Dotted lines show domain-mean skin temperature (from ERA5 reanalysis), MCS cover (from TOOCAN tracked MCSs), and satellite inferred precipitation rate (from IMERG) calculated within the Lagrangian frame that follows tracked MCSs across tropical African land and onto the adjacent Atlantic Ocean, in July 2012-2016.

When investigating the emergence of self-aggregation in the two simulations with wind shear, a similar property arises: namely, the spatial variance of precipitable

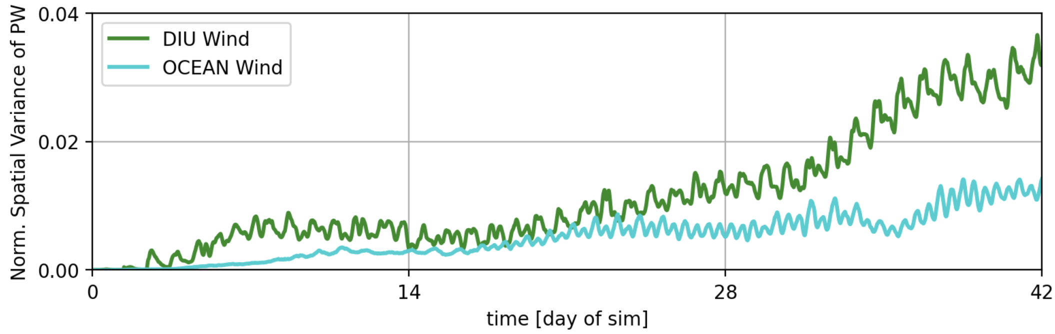


Figure 3.12: **Emerging self-aggregation of convection in simulations with wind.** Aggregation of convection for DIU Wind vs OCEAN Wind simulations described in Tab. 3.3, measured with the normalized spatial variance of precipitable water.

water starts increasing in the DIU Wind simulation and diverges from the OCEAN Wind simulation after about 3 weeks of simulation (see Figure 3.12), increasing at a similar rate in 3.3. So the addition of wind does not seem to affect the aggregation of a simulation with a large amplitude of diurnal cycle in surface temperatures. The self-aggregation of the OCEAN Wind simulation seems to be aided by the wind shear and slowly reaches values of normalized spatial variance of precipitable water of 0.01, as opposed to the low levels of 0.001 for all OCEAN simulations in Figure 3.3. This alludes to the fact that the wind shear does not necessarily impact the strong aggregation due to diurnal-cycle induced MCSs, but could act to aid the formation of convective aggregation in simulations with no diurnal cycle, through the sustenance of deep convection in the form of long lived squall lines (as described in Rotunno *et al.* (1988)), that then act like the diurnal MCSs on the moisture field.

Deeper insight into the different mechanisms of aggregation of the DIU and OCEAN simulations with wind, can be gained by calculating the spatial correlation between two moisture fields at different levels. We calculate this as the pixel-by-pixel Pearson correlation between the horizontal moisture distribution at the surface level of the model and at mid-cloud level (3500m) as in equation 3.10. For the moisture distribution at mid-cloud level we use the total specific humidity q_t (kg/kg) which includes water vapor, cloud condensate and ice; while for the moisture distribution at the surface we use the water vapor specific humidity q_v (kg/kg).

$$corr(t) = \sum_{i,j=1}^{N} qv_{SURF}(t) \cdot qt_{3500}(t) \tag{3.10}$$

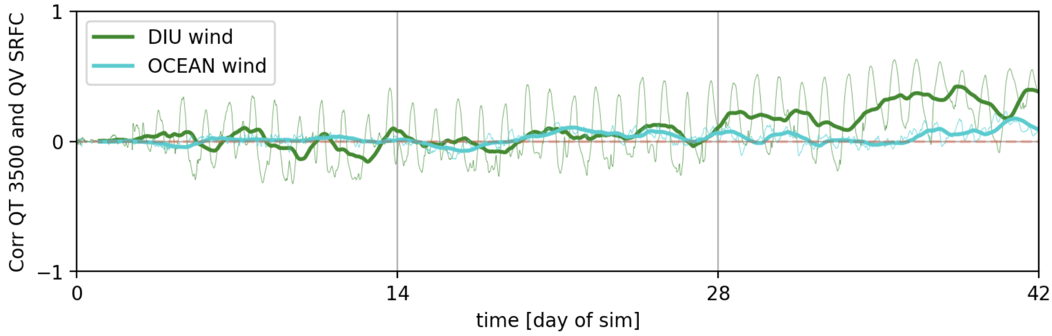


Figure 3.13: **Spatial correlation between two moisture levels in simulations with wind.** Spatial correlation between two moisture levels: total specific humidity at 3500 m (QT 3500) and specific humidity at surface level (QV SRFC) for DIU Wind vs OCEAN Wind simulations described in Tab. 3.3. Thick lines represent 24-h running mean, and thin lines represent all time steps.

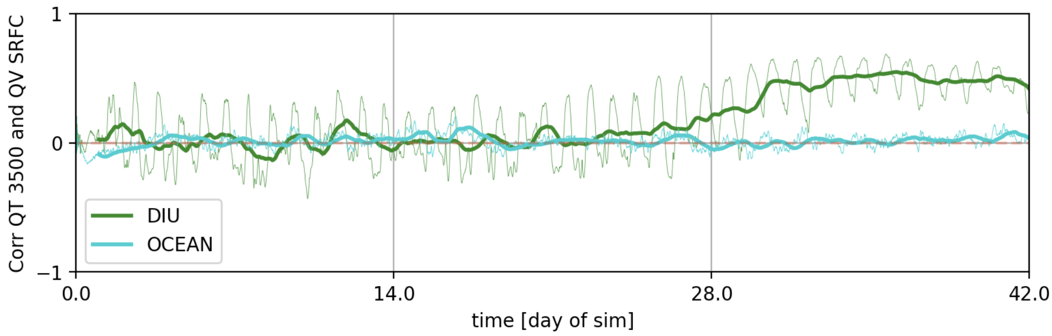


Figure 3.14: **Spatial correlation between two moisture levels in simulations without wind.** Spatial correlation between two moisture levels: total specific humidity at 3500 m (QT 3500) and specific humidity at surface level (QV SRFC) for DIU vs OCEAN simulations described in Tab. 3.1. Thick lines represent 24-h running mean, and thin lines represent all time steps.

The correlation essentially tells us when the mid cloud moisture matches up with the low-level moisture, and when it does not. In Figure 3.13, we see the correlation between moisture levels, for DIU Wind and OCEAN Wind simulations. The 24 hour running mean stays around the value 0 for both simulations for the first four weeks, indicating that on average, there is no spatial correlation between the two levels. A large difference appears when focusing on the individual time steps: DIU Wind, after the first few days of spin-up, exhibits large oscillations between positive and negative correlations between the two levels of moisture. The positive correlation during the day can be attributable to the formation of MCSs thanks to the diurnal cycle, which are collocated in the moist columns within which they are triggered during the day, exhibiting thus a positive correlation between the lowest level and the mid-cloud level. The negative correlation during the night can be attributable to the aftermath of large CPs created by MCSs, that dry the lowest level of moisture of the column they were produced in, while the remnants of the MCSs have moistened the top level of that same column, exhibiting thus a negative correlation between the lowest level and mid-cloud level. In the OCEAN Wind simulation, there are small oscillations towards positive and negative correlation values, but these stay small, intuitively due to the small size of the (unorganized) deep convective cells. This behavior was seen for DIU and OCEAN (see Figure 3.14), but seems to be robust with the inclusion of westward wind and shear. Things get interesting around week 4 of the simulation, when the moisture field starts to self-aggregate for DIU Wind, as we had seen for DIU. This is indicated by the increase in the 24 hour running mean correlation, which stays strongly positive. From analyzing the 2D fields of moisture, and as was found also in Jensen *et al.* (2022), this is due to dry moisture patch, that grows and extends to the whole column, kick-starting the radiative feedbacks akin to self-aggregation. As this dry region of the domain expands, the 24 hour average correlation between the mid-cloud moisture field and low-level moisture field stays positive (compare Figures 3.13 and 3.14) for DIU wind and DIU) - and the fact that this feature is also robust with the inclusion of wind (and thus, advection), suggests that an aggregated moisture field can both be formed and advected.

When observing the output of the simulations in Hovmöller diagrams (Figure 3.15 c and d), we observe an interesting phenomenon: deep convective clouds advected with an imposed wind profile are on average faster than the advected moisture field they are embedded within. However, when there is also a diurnal cycle, the large convective clusters tend to slow down, converging to the advection speed of the surrounding moisture field. To produce the Hovmöller diagrams in Figure 3.15 c and d, we average over the latitudinal dimension of our simulations and visualize each time step as a line. We compare the fields of Longwave Radiation at Top of Atmosphere (LW at TOA) as a proxy for deep convection and total specific humidity field at 3500 m (QT at 3500 m), to understand where the deep convective clouds form with respect to the moisture field at cloud height. Due to the imposed westward wind, and the associated advection of high clouds, the deep convection and collocated moisture travels from the East to the West of the domain, as indicated by the inclined lines in Figure 3.15 c and d.

We calculate the advection velocity of the various fields, in Figure 3.15 a and b. It emerges that the high cloud field in DIU Wind is much faster than the advected moisture field at mid-cloud level, but it slows down gradually throughout the

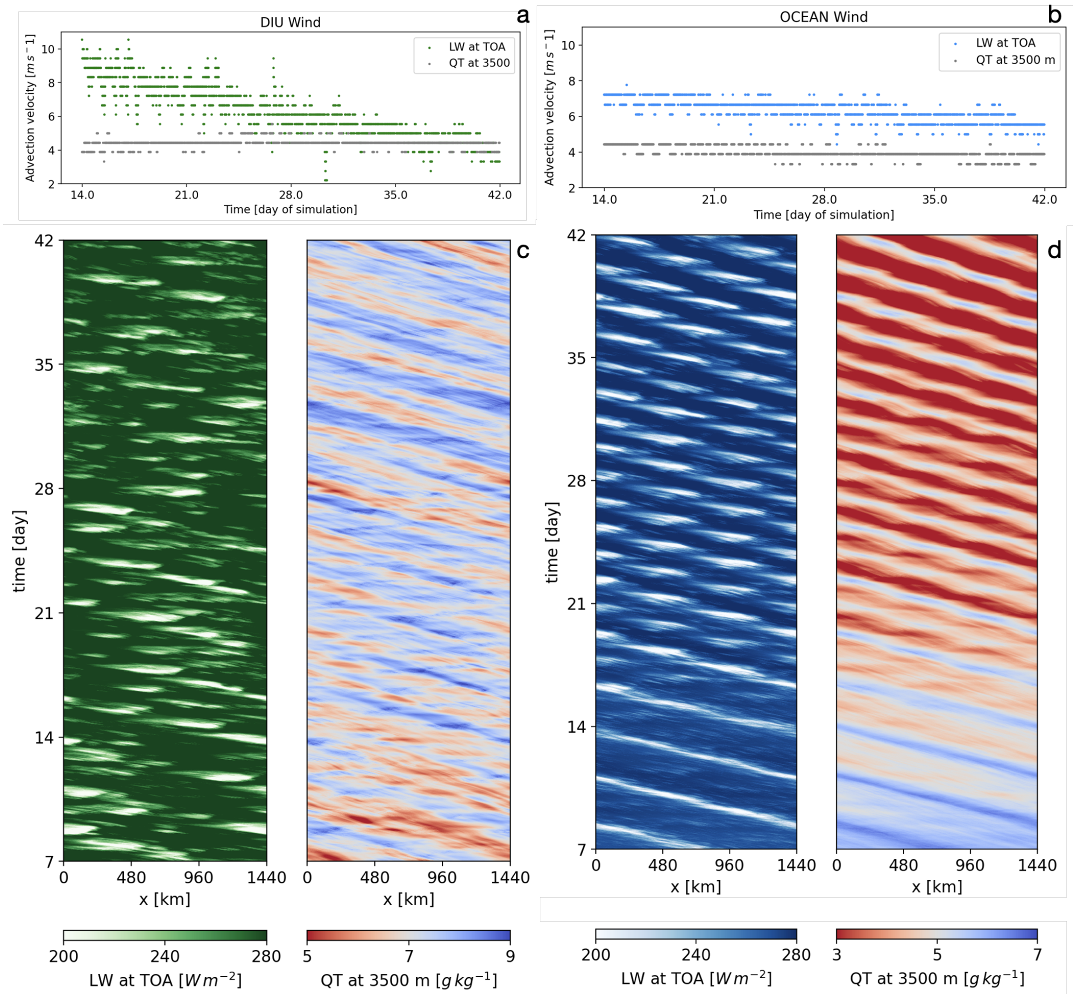


Figure 3.15: **Deep convection embedded in a moisture field, with imposed background wind shear.** a) Time series of advection speed of Longwave Radiation at Top of Atmosphere (LW at TOA) compared to advection speed of total specific humidity field at 3500 m (QT at 3500 m) for DIU Wind (a) and OCEAN Wind (b). Hovmöller diagrams of LW at TOA and QT at 3500 for DIU Wind (c) and OCEAN Wind (d).

simulation - converging fully with the moisture field after 6 weeks of simulation. MCSs moisten the upper troposphere and thus leave behind a moist trail when advected. For the OCEAN Wind simulation in Figure 3.15 d, the high clouds are initiated and advected in a more ordered manner, traveling as a wave packet and leaving behind a more slowly advected moist trail. There might eventually be convergence of the two advection speeds in OCEAN Wind, but not within this time frame. Jung *et al.* (2021) investigated how the presence of a mean flow affects the propagation of organized deep convection in an RCE framework with constant SSTs and found that the convective cluster initially moves slower than the pure advection imposed, eventually becoming stationary regardless of wind speed.

They find that the near surface wind responds to the mean flow by altering the surface fluxes, thereby decreasing the near-surface wind on the upwind side of the cluster and enhancing it on the downwind side, acting thus as a drag on the mean background wind. This could be an explanation to the DIU Wind simulation showing a more rapid slow-down, due to the large clusters created, with larger CPs and stronger near-surface winds (Figure 3.16 shows the surface winds in the DIU and OCEAN simulations without imposed winds - it is clear that the surface winds are inherently larger in DIU).

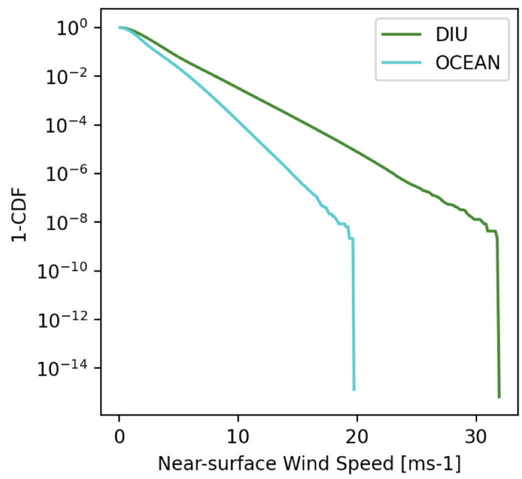


Figure 3.16: **Near Surface Wind Field in DIU and OCEAN.** Cumulative Distribution Function (1 - CDF) of near-surface wind values for DIU and OCEAN simulations.

We can assume this effect to explain why the DIU simulations show a slow-down of the deep convective cloud field - when this converges with the advective speed of the moisture they are embedded within, the atmospheric volume is free to aggregate with all the processes akin to CSA, as if there were no prescribed wind profile advecting the clouds. MCSs produced with the diurnal cycle, again, accelerate the way to this state, paving the way for a segregation of the (advected) simulation domain into persistent dry regions of the simulation domain, and moist regions where the MCSs continue to be produced.

That our results remain valid in simulations with imposed wind shear, thus idealised simulations that are brought even closer to reality, increases our confidence in CSA emerging in the real tropical atmosphere, where deep convection is strongly influenced by the key forcing conditions imposed, i.e. diurnally varying land surface temperatures, near-constant sea surface temperatures, and westward wind shear.

4

Cold Pools Over the Netherlands: A Statistical Study From Tower and Radar Observations

Authors: Irene L. Kruse; Jan O. Haerter; Bettina Meyer
Journal: Quarterly Journal of the Royal Meteorological Society
Vol./page: 148/711–726
DOI: 10.1002/qj.4223

My contributions: conceptualization; data curation; formal analysis; methodology; validation; visualization; writing – original draft, review and editing.
Notes in relation to Master's Thesis: My Master's thesis was a predecessor to this manuscript, presenting a first analysis of a subset of the data. This manuscript has been written during the PhD and includes new and more comprehensive data, results, and analyses.

Figure 4.1: Boundary Layer Measurement Tower in Cabauw, the Netherlands

Abstract

We provide a detailed analysis of convectively-generated cold pools (CPs) over flat mid-latitude land, combining a ten-year high-frequency time series of vertical measurements from the 213 m tower observatory at Cabauw, the Netherlands, with a collocated 2D radar rainfall dataset. This combination of data allows to relate observations of the CP's temporal and vertical structure with the properties of each CP's parent rain cell, which we identify by rain cell tracking. Using a new detection method, based on the anomalies of both the vertically-averaged wind and the temperature, we monitor the arrival and passing of 189 CPs during ten summers (2010-2019). The time series show a clear signature of vortex-like motion along the leading CP edge in the vertical and horizontal wind measurements. The arrival of the CP gust fronts is characterized by a steep decrease in both temperature and moisture with a recovery time of approximately two hours. We see no evidence of moisture rings on the gust front edge, and therefore no indications for thermodynamic convective triggering. From the tower data we obtain a median CP temperature drop of $T_{drop} \approx -2.9\,\mathrm{K}$ and a height-averaged horizontal wind anomaly of $\Delta u_{max} \approx 4.4\,\mathrm{m\,s^{-1}}$. Relating the individual CP horizontal wind anomalies and temperature drops, we confirm the validity of the theoretical, density current relationship $\Delta u_{max} \propto T_{drop}{}^{1/2}$. We propose a simple statistical model to relate CP strength defined by T_{drop}, to the environmental properties mostly influencing the CP: rain intensity and lower boundary layer saturation. A multi-variate linear regression suggests a 1 K colder CP for a $4\,\mathrm{mm\,h^{-1}}$ more intense rain cell (instantaneous area-averaged rain intensity) or for a 2.5 K larger pre-CP dew point depression.

4.1 Introduction

A convectively generated cold pool (CP) is a sub-cloud volume of air, which is cooled by the partial evaporation of precipitation. This dense, cold air descends as it is negatively buoyant relative to its surroundings and is often further accelerated by the drag exerted by the falling hydrometeors (Wakimoto, 2001). When hitting the ground, the CP spreads along the surface away from its source as a density current (Charba, 1974), introducing cold dense air underneath warmer environmental air. Thereby, vertical motion builds up, which is measurable as the combination of horizontal wind, distinctive horizontal convergence lines, and vertical mass fluxes — often referred to as the "CP outflow boundary", or "CP gust front." The amount of evaporation, and hence cooling within a CP, depends on rain intensity and area of the generating rain cell, as well as the environmental atmospheric profiles of temperature and relative humidity. The coldest outflows thus stem from high-based thunderstorms that precipitate into very dry boundary layers (Markowski and Richardson, 2010). Beyond these macrophysical conditions, microphysical parameters, crucially the drop size distribution, influence the rain evaporation (Seifert, 2008).

CP characteristics are often studied using numerical simulations that now approach the fine scales needed to resolve some CP properties, that is, horizontal grid resolutions of substantially less than one kilometer (Drager and Heever, 2017; Fournier and Haerter, 2019; Cafaro and Rooney, 2018; Meyer and Haerter, 2020; Drager *et al.*, 2020). Observational work is however indispensable as a mean of comparison and validation for numerical studies. Oceanic measurement campaigns have provided insight into the dynamics of ensembles of CPs over the tropical and subtropical oceans (Szoeke *et al.*, 2017; Zuidema *et al.*, 2012; Vogel, 2014; Young *et al.*, 1995; Chandra *et al.*, 2018). As in simulations, CPs observed over the tropical oceans show relatively weak temperature anomalies of typically $-1\,\mathrm{K}$ to $-1.5\,\mathrm{K}$ and wind gusts of $2-2.5\,\mathrm{m\,s^{-1}}$. An important morphological feature arising in simulated oceanic CPs are so-called moisture rings, that is, bands of enhanced water vapor, which build up near the gust front as the CP spreads. Oceanic studies find minor increases in moisture, $0.25\,\mathrm{g\,kg^{-1}}$ (Szoeke *et al.*, 2017), measured ahead of the gust front, whereas subsequent decreases in moisture, measured behind the gust front, are nearly an order of magnitude larger. CPs over land vary much more strongly than those over ocean and can reach temperature anomalies as deep as $-17\,\mathrm{K}$ and wind gusts larger than $15\,\mathrm{m\,s^{-1}}$ (Engerer *et al.*, 2008). Observational studies of CPs over land are often focused on case studies and are mostly based in the continental US (Mueller and Carbone, 1987; Wakimoto, 1982; Engerer *et al.*, 2008; Hitchcock *et al.*, 2019; Heever *et al.*, 2021). In contrast to oceanic CPs, the moisture signal is not always prevalent in observations over land: a recent statistical analysis of mid-latitude CPs over Hamburg, Germany, finds no strong signature of moisture rings near the gust front, but the authors do find a pronounced, longer-lasting moisture signal that increases up to one hour after the first detection of the CP (Kirsch *et al.*, 2021). Simulating continental CPs and detecting their edges through a buoyancy anomaly method, Drager *et al.* (2020) conclude that moisture rings over land may only occur in conditions of moist surfaces, but not for drier conditions.

A CP can interact with its environment and other CPs, producing updrafts and potentially triggering new convective clouds (Purdom, 1976; Weaver and Nelson, 1982; Droegemeier and Wilhelmson, 1985a; Droegemeier and Wilhelmson, 1985b; Wilson and Schreiber, 1986; Moncrieff and Liu, 1999; Tompkins, 2001a; Torri *et al.*, 2015). The triggering of new convection can happen through mechanical forcing, due to the lifting of air along the gust front of the CP, or thermodynamic forcing, due the above-mentioned moisture rings that provide additional buoyancy, favouring new convection at the edge of the CP (Tompkins, 2001b; Torri *et al.*, 2015; Drager *et al.*, 2020). An incessant sequence can unfold, where the new precipitating clouds themselves may create new CPs that in turn can result in a new precipitation event, which then produces a new CP, etc. CPs may therefore be a key ingredient to the understanding of how clouds organize spatially and temporally into larger-scale precipitating systems (Böing, 2016; Haerter *et al.*, 2019; Haerter, 2019). The "up-scale communication" between small-scale individual CPs (~ 10 km horizontally and ~ 1 h temporally) and the larger-scale spatial organization of the cloud field ($\gtrsim 100$ km, respectively $\gtrsim 3$ h, such as mesoscale convective systems) has been explored in both theoretical (Rotunno *et al.*, 1988; Jeevanjee and Romps, 2013; Haerter *et al.*, 2019; Haerter *et al.*, 2020) and observational studies (Zipser, 1977; Feng *et al.*, 2015; Zuidema *et al.*, 2017). All the more, there is a need for better understanding of the processes that determine the temporal evolution of CP properties, such as their spatial extent, lifetime, and strength (Drager and Heever, 2017). Defining the relationship between the properties of a given CP and the precipitation cell that produces this CP (termed “parent” rain cell) can serve as a useful benchmark for numerical simulations.

In this study, we provide an observation-based analysis of CPs over mid-latitude coastal land, relating CP properties to their parent rain cell and the environment.

We use weather measurements from the Netherlands, a region whose climate is a hybrid between a moist oceanic and a drier continental regime and thus interesting to study CPs, which are strongly affected by boundary layer moisture. We combine meteorological tower data and radar data and further develop a method to detect CPs from weather tower measurements based on a previous study by Szoeke *et al.* (2017). The measured horizontal wind and temperature at the time of CP detection allow us to empirically test the theoretical scaling between CP propagation speed and temperature depression (Karman, 1940; Benjamin, 1968). Radar imagery is then utilized to visualize and track the rain cells in the tower's surroundings. Using the distance of the rain cells to the measurement tower and the wind direction during CP passage, we attribute a specific parent rain cell to each CP. Combining the point measurements from the tower with the tracked rain cells, allows us to analyse the relation between CP temperature anomaly, the pre-event saturation of the boundary layer, and the rain intensity of the CP's parent rain cell.

4.2 Methods

4.2.1 Data

We use data from a 213 m boundary layer measurement tower at the Cabauw Experimental Site for Atmospheric Research (CESAR observatory), located in Cabauw (51.971 N, 4.927 E), the Netherlands. These data enable us to study the temporal evolution of lower boundary layer properties before, during and after the passage of a CP gust front over land, in a temperate maritime climate. We focus on the summer period (May to September), to capture the maximum convective activity in the region. The data sets used consist of one-min averaged measurements of temperature, dew point temperature, horizontal wind speed, and wind direction at six different heights (10 m, 20 m, 40 m, 80 m, 140 m, 200 m) for the period 2010-2019. Temperature is measured with Pt500 elements, placed in unventilated screens to minimize influence of radiation and precipitation; dew point temperature is measured with EplusE 33 polymer-based relative humidity sensors, which are heated to decrease measurement problems during humid conditions; horizontal wind speed and wind direction is measured with a cup-anemometer and vane combination that rotate with the direction of the wind and record the velocity of the propellers (Bosveld *et al.*, 2020). We additionally use 0.1-sec records of water vapor concentration and vertical wind speed measured at heights of 60 m and

180 m that are available only for the summer of 2019, measured with a Gill-R50 sonic anemometer (Bosveld *et al.*, 2020). Tower data was retrieved from the CESAR Data Portal (*Cesar (Cabauw experimental site for atmospheric research) Database* 2020).

Additionally, we use a rainfall radar dataset from the Royal Netherlands Meteorological Institute (KNMI) to track and study the rain events generating the CPs for the period 2010-2019. These 2D horizontal data are composites of radar reflectivities from both of the KNMI weather radars, Den Helder and Herwijnen. The resultant dataset encompasses five-minute precipitation heights on a 1 km × 1 km grid, which have been adjusted employing validated and complete rain gauge data from the KNMI rain gauge networks and provides data for the entire land surface of the Netherlands.

To look for visible gust fronts, we further use 5-min time resolution radar imagery from the Herwijnen C-band polarimetric Doppler radar tower (51.837 N, 5.138 E). The Herwijnen radar scans the surrounding atmosphere at 15 different elevation angles, starting from the horizontal and upwards. The finest radial resolution of the Herwijnen radar is approximately 225 m and the azimuthal resolution is approximately 1 degree.

All the radar data was retrieved from the open-access KNMI Data Portal (KMMI, 2020).

4.2.2 Algorithm for the Detection of Cold Pools from Tower Measurements

We build an algorithm to detect CPs from tower measurements, based on a temperature detection algorithm used previously for oceanic CPs (Szoeke *et al.*, 2017). We modify this algorithm by tailoring the temperature threshold to continental conditions, imposing a time constraint on the temperature anomaly and adding a criterion on the horizontal wind anomaly. This additional criterion is imposed to ensure the existence of a wind gust along with the detection of cold air, thus incorporating the two main characteristics of a CP. For any given day, we use the one-min time series of temperature at the 10 m tower level, and one-min time series of wind speed at the six tower levels (10, 20, 40, 80, 140, and 200 m).

Temperature Criterion. The algorithm by Szoeke *et al.* (2017) identifies CP gust fronts from one-min surface temperature over a tropical oceanic surface. It is “designed to be sensitive to asymmetric cooling events visible in the time series, yet insensitive to high-frequency noise, to exclude false positives". We use this algorithm as a first step in CP detection. We first smoothen the temperature time series with a running 11-min centered window. A series of threshold operations is then applied to the smoothed time series, to identify and “record" CPs along with their properties (Fig. 4.2a):

1. A CP candidate is identified when the smoothed temperature reaches a minimum within the preceding 20-min window. This minimum temperature is the first T_{min} of a possible CP event (left cross in Fig. 4.2a).

2. Multiple temperature minima are combined to one CP event, if they are consecutive (i.e., separated by one min), or if lying within 20 min of each other given that the temperature during that time window does not exceed either of them by 0.5 K. This way we allow for small temperature fluctuations within the CP interior without detecting it as two separate CP events. We choose these values following the algorithm proposed in Szoeke *et al.* (2017).

3. For each detected CP event, the temperature drop δT is defined as the absolute difference between the last T_{min} and the maximum smoothed temperature in the 20-min time window preceding the first T_{min} (Fig. 4.2a). A time interval Δt is defined as the time elapsed between the first and the last T_{min}.

4. An event is recorded if δT exceeds a 1.5 K threshold, and the time interval Δt does not exceed 60 min. The temperature threshold is raised compared to the one used in Szoeke *et al.* (2017) (aimed at detecting CPs over a tropical oceanic surface) to reduce the signal-to-noise ratio, caused by the higher temperature fluctuations over land compared to the (tropical) ocean.

5. A refined temperature drop T_{drop} is defined as the difference between the maximum unfiltered temperature within 10 min preceding the first T_{min}, and the minimum unfiltered temperature within the temperature drop. We use this stronger temperature drop to have a more accurate measure of the effective cooling due to the CP. We note that while the δT-values are systematically too low due to smoothing, may at times be very high if it includes local fluctuations.

Wind criterion. A further criterion is then added to the detection algorithm, as a novelty with respect to the algorithm used by Szoeke *et al.* (2017), to ensure that there is a wind gust associated with each detected temperature decrease. For this purpose we use the time series of horizontal wind speed at the six tower levels. For each day we smooth the one-min time series with a running two-hour centered window and subtract the smoothed time series from the original time series to obtain the "horizontal wind anomaly". We choose the two-hour smoothing window to extract short-term fluctuations from long-term wind variability. To extract the vertically-coherent signals in the data set, we compute the average wind speed of all six tower heights. This way, we reduce random (or turbulent) fluctuations, and a wind gust visible at all six heights will rise above the noise. We call this variable the "height-averaged horizontal wind anomaly" (Δu). For simplicity, we will name the maximum of this variable in a given time window "wind peak" Δu_{max}.

The detection algorithm then scans the events recorded by the temperature criterion, and ultimately saves an event as a "cold pool" if there is a wind peak within ten minutes preceding the first T_{min} and the time of the last T_{min}, that exceeds four standard deviations (4σ) of the daily one-min Δu time series. An example, illustrating the algorithm (Fig. 4.2), shows the ten-meter temperature time series and the height-averaged horizontal wind anomaly time series of a specific day where a CP was detected (August 27, 2019).

Parameter Sensitivity. The parameters for the algorithm were initially tuned based on two case studies, where the front of the CPs was clearly visible in the radar images due to dust and/or insects drafted up in the convergece zone (Herwijnen radar images, May 29 2018 14:30-15:30 UTC and Aug 27 2019 16:30-17:00 UTC (Fig. 4.4), presented in Kruse (2020)). This allowed to visually determine the time instance when the CP front should be detected at the measurement tower, revealing typical CP signals in the time series we should look for and against which the algorithm was calibrated. The thresholds on temperature and wind peak were chosen to capture these CP cases from the daily temperature and wind time series, and were kept as high as possible in order to find similar cases of strong, clear CP signals throughout the year, and to exclude sea breezes. Reducing the threshold in temperature anomaly from 1.5 K to 1 K adds only very few extra cases (additional 15% for year 2019). By contrast, reducing the threshold on the wind speed from 4σ to 3σ nearly doubles the number of cases detected. However, the additional cases are of very similar nature, i.e., the gust fronts are detected shortly prior to the presence of a rain cell over tower, with the only difference that the rain cells are weaker. This means that the additional CPs detected are generated by rain cells with intensities that are often lower than the threshold of 1 mm h^{-1}, set for the detection of convective rain cells (Section 4.2.3).

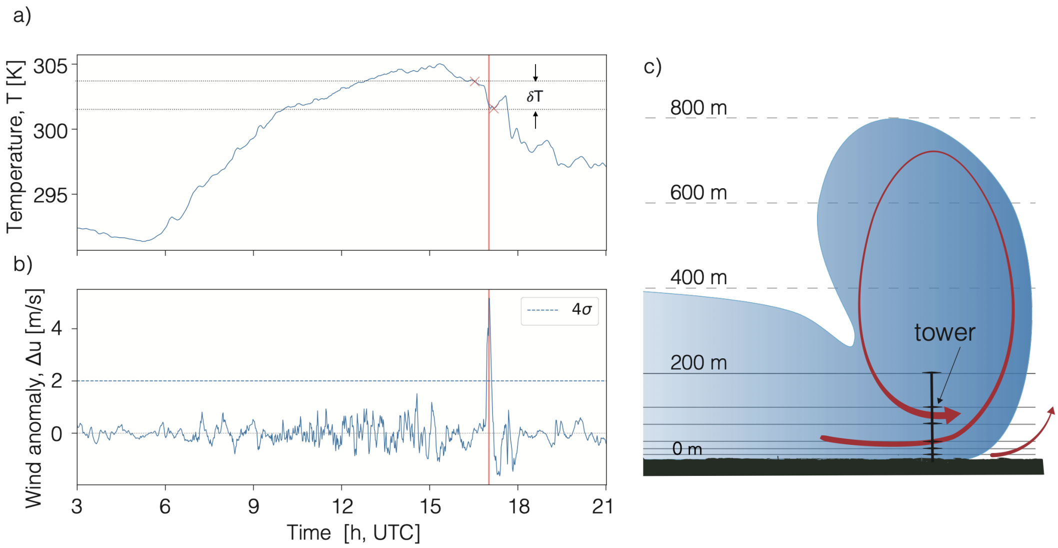


Figure 4.2: **Exemplifying cold pool detection by two-step criterion.** (a) Daily time series of temperature, measured at the 10 m-level at one-min temporal resolution and smoothed with an 11-min centered window. The red symbols and two horizontal lines mark the respective first and last values of T_{min} of a detected cold pool event, denoting the initial temperature drop δt; (b) Horizontal wind anomaly averaged over all tower heights (Sec. 4.2.2) measured at one minute temporal resolution. The anomaly is computed with respect to a two hour centered running temporal average. The horizontal dashed blue line indicates four standard deviations from the daily mean horizontal wind anomaly, exceedance of which is used as a criterion for the detection of strong wind anomalies. (c) Sketch of a cold pool (blue shaded area) crossing the Cabauw tower. The red arrows indicate the propagation velocity and internal circulation of the cold pool, together composing the measured horizontal wind anomaly. The levels of temperature and horizontal wind measurements (10 m, 20 m, 40 m, 80 m, 140 m, 200 m) are indicated by solid horizontal black lines.

4.2.3 Attribution of a Rain Cell from Radar Data

Rain Cell Tracking. We use an Iterative Raincell Tracking (IRT) method (Moseley *et al.*, 2014; Moseley *et al.*, 2019) to track rain cells in time and space from the gridded radar rainfall product. The IRT locates spatially contiguous areas of rainfall, termed *objects*, and tracks them in time if they overlap with objects in subsequent time steps. This allows the definition of rain cell *tracks*, extending over a time window $t \in [t_{\min}, t_{\max}]$. For the object identification, a threshold of 0.08 mm per 5 min, corresponding to $I \approx 1\,\mathrm{mm\,h^{-1}}$ is imposed and a minimum of 4 contiguous rainy pixels (one pixel corresponds to an area of approximately 1 km × 1 km). To each detected CP a rain cell track is then attributed based on a multi-step algorithm (Fig. 4.3): First, we select all rain tracks that exist for at least $\Delta = 10\,\mathrm{min}$ during the time interval $\delta = 30\,\mathrm{min}$ preceding the CP detection time t_0 at any point in the domain. The tracks are allowed to end before the CP is detected, defining a "rain cell timestep" $t_{\mathrm{RC}} = \min(t_0, t_{\max}) \in [t_0 - \delta + \Delta, t_0]$. In a second step, all rain tracks are discarded, whose closest edge is further than $r_{\max} = v_{\max} \cdot \delta = 18\,\mathrm{km}$ away from the location of the Cabauw tower ($\vec{x}_{\mathrm{Cabauw}}$) at time step t_{RC}. The choice of this radius is based on the assumption of an upper bound on CP propagation speed $v_{\max} = 10\,\mathrm{m\,s^{-1}}$, implying that gust fronts generated by rain that falls further away cannot reach the tower in time. In a final step, the domain is cut in half-planes based on the wind direction measured at the tower during CP passage (ϕ_{CP}): all rain cells with centers of mass (COM) located in a direction relative to Cabauw tower that deviates by more than 90° from ϕ_{CP} are discarded (Fig. 4.3). In the majority of cases this three-step process allows for the identification of a unique rain cell, which is assumed to be the parent rain cell of the gust front measured in the tower time series. We point out that the number of CPs with a unique attributable rain cell is 116 out of 189 - meaning the statistics involving CPs in connection to the rain, have fewer data points. We note that in general, the detected CPs are located close to the edge of their parent rain cell. This is quantified by the relative distance of the RC's COM to Cabauw tower, divided by the approximate radius of the RC when assuming circular shape, both taken at t_{RC}

$$d = \frac{\vec{x}_{\mathrm{RC,\,COM}} - \vec{x}_{\mathrm{Cabauw}}}{\sqrt{A_{\mathrm{RC}}/\pi}}, \tag{4.1}$$

where $\vec{x}_{\mathrm{RC,COM}}$ is the position of the rain cell's center-of-mass, and A_{RC} is the area of the rain cell. We find that d is distributed around a mean value of $\langle d \rangle = 1.7$ with a heavy tail to large values up to $d = 5$ (Appendix Fig. 4.9), where $d = 1$ corresponds to the CP gust front being detected directly at the edge of the rain cell. The relatively large number of cases with $d < 1$ is an artefact of the definition of d being based on the assumption of circular RCs. In reality, many CPs are generated by RCs of elongated shape that are oriented perpendicular to the vector $\vec{x}_{\mathrm{RC,\,COM}} - \vec{x}_{\mathrm{Cabauw}}$ (Appendix Fig. 4.10).

Qualitative Analysis of Weather Situations. The combination of radar data with the CP detection from tower measurements allows the observation of the large-scale weather situation in which the CPs live. As expected, the algorithm detects gust fronts in very diverse weather situations, which we qualitatively identify as isolated convection (Appendix Fig. 4.10a), elongated precipitation cells resembling squall-lines (Appendix Fig. 4.10b), large-scale fronts (Appendix Fig. 4.10c) and mesoscale convective systems, characterised by low large-scale wind (Appendix Fig. 4.10d). The different weather situations that the CPs are nested in are each defined by a particular boundary layer wind shear, background wind direction with respect to the direction of propagation towards the tower, soil moisture, etc. These factors are contributing to the spread in the measurements of CP properties (Sec. 4.3).

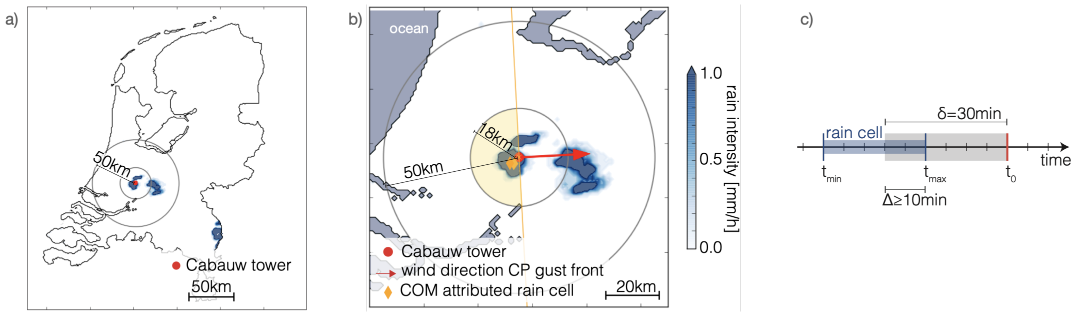


Figure 4.3: **Schematic of rain cell attribution algorithm.** (a) Overview of radar data domain with the location of the Cabauw tower marked. Black contour line marks the political boundary of the Netherlands.; (b) Schematic of geometrical requirements on the rain cell position during $t \in [t_0 - \delta \min(t_0, t_{\mathrm{min}})]$ based on the wind direction of the detected CP gust front at time t_0. Only rain cells within the yellow half-circle are accepted; (c) Schematic of requirement on temporal overlap of rain cells, that exist in the time window $t \in [t_{\mathrm{min}}, t_{\mathrm{max}}]$.

4.3 Results

4.3.1 Cold Pool Structure

Case Study. We first discuss an individual CP event as seen in C-band radar imagery. In the radar reflectivity, measured at the Herwijnen radar tower (Fig. 4.4, left panels), one can clearly see the development of a quasi-circular gust front (light green shades), the CP edge, spreading around an area of high-reflectivity (dark blue) that we associate with the rain event. The gust front is visible in the reflectivity due to insects and/or dust carried aloft by the convergence of air at the CP boundary (Markowski and Richardson, 2010).

In the radial Doppler velocity measurements (Fig. 4.4, right panels), one can see that the area delimited by the gust front is characterized by outward movement (positive radial velocities ranging from $5-10\,\mathrm{m\,s^{-1}}$.). If we approximate the spreading gust front with a temporally growing circle of radius $r(t)$, we obtain a near-constant horizontal propagation speed $v(r) \equiv dr(t)/dt \approx 7\,\mathrm{m\,s^{-1}}$. We note that this is a unique case, as we have a strong CP signal that spreads almost perfectly around the radar tower - this is generally not the case in the rest of our data set. Interestingly, the wind peak measured at Cabauw tower at the time when the radar gust front seems to pass by the tower, actually comes from the direction of the smaller rain cell, located South-West of the tower. We use this example to stress the added benefit of verifying the wind direction of the wind gust when attributing a generating rain cell to a CP signal.

CP Composite Time Series. The composites for horizontal wind anomaly and temperature anomaly (Fig. 4.5) are drawn from 189 CPs detected in the 10 summers from 2010 to 2019, while the composites for vertical wind and moisture anomaly (Fig. 4.6) are drawn from only 18 CPs detected in the summer of 2019. The discrepancy in CP count is due to the availability of data from different measurement instruments (Sec. 4.2.1). Most CPs are detected in the afternoon, coinciding with the time of day when convection is most active over land (Appendix Fig. 4.11).

We compute the composites by first centering the individual CP time series on their respective times of maximum horizontal wind anomaly (t_0) and retain the data for 60 min before and 120 min after t_0, resulting in a set of time series which

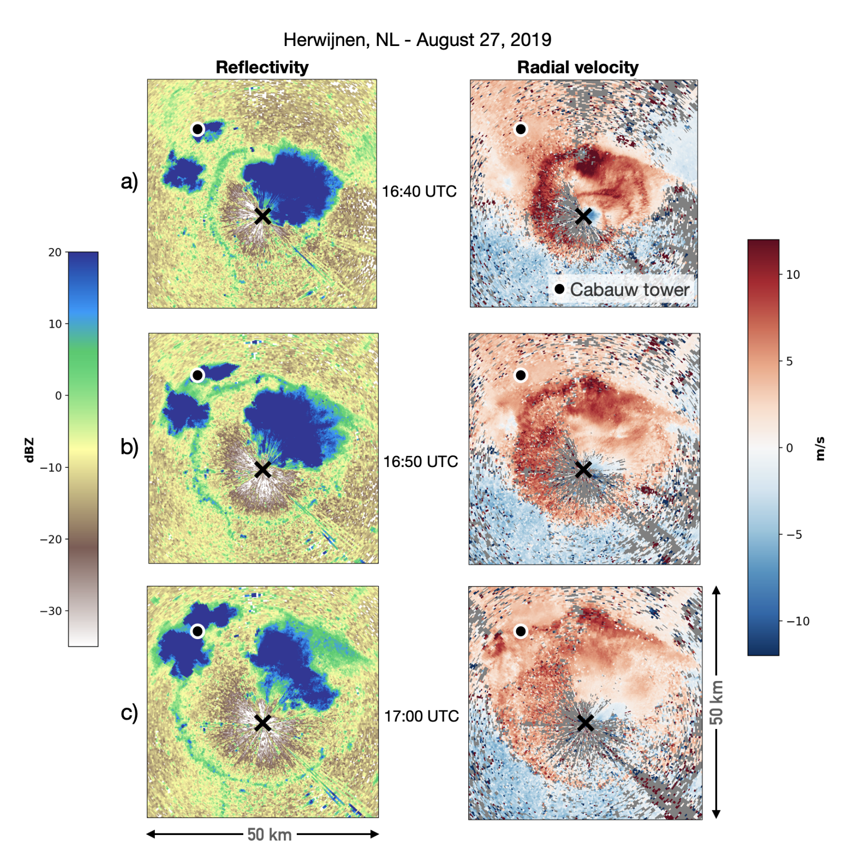


Figure 4.4: **Cold pool developing around Herwijnen radar tower.** The Herwijnen tower is marked as an "x" in each figure. Radar reflectivity (left column) and Doppler radial velocity (right column) were recorded on August 27, 2019 in ten-minute steps from 16:40 to 17:00 as marked in panels a-c, respectively. The gust front is clearly seen as a ring of low reflectivity values spreading around the the largest rain event. This ring corresponds to positive radial velocities. We note that the measurements shown are taken at an elevation angle of 1.20°. In panel a), the gust front is observed at approximately 200 m asl, whereas in panel c) the observed height is approximately 400 m asl.

each have the same number of time steps. For each time step we then average over all time series, yielding the mean time series. To compute the composites of the anomalies $x'(t)$ of a quantity $x(t)$, such as horizontal wind, temperature and water vapor concentration, we first remove the respective pre-CP temporal mean from the time series, that is,

$$x'(t) \equiv x(t) - \overline{x}, \tag{4.2}$$

where $\overline{x}$ is the time average over the 51 one-min time steps from $t_0 - 60\,min$ to $t_0 - 10\,min$, hence the time window preceding the arrival of the CP. The ten-min margin was chosen to ensure that the CP signal does not influence the mean. Furthermore, for the temperature anomaly, we remove the effects of the diurnal cycle by subtracting the two-hour running mean. We verified that the one year composites of temperature and wind, although they are somewhat more noisy, are comparable to the ten year composites of the same variables. This makes us confident that one year of data is representative of a larger data set.

The edge of the composite CP is characterized by a strong positive horizontal wind anomaly (gust front) seen at all tower heights (Fig. 4.5a). Before and after passage of the gust front, the wind anomaly increases monotonically with the tower height, as one would expect in the surface layer (see Fig. 4.5d). By contrast, within a window of approximately six minutes enclosing $t = t_0$, the largest value of horizontal wind anomaly is measured at an intermediate tower level, near $z = 80\,\mathrm{m}$ (Fig. 4.5c). Since the measured horizontal wind anomaly corresponds to the superposition

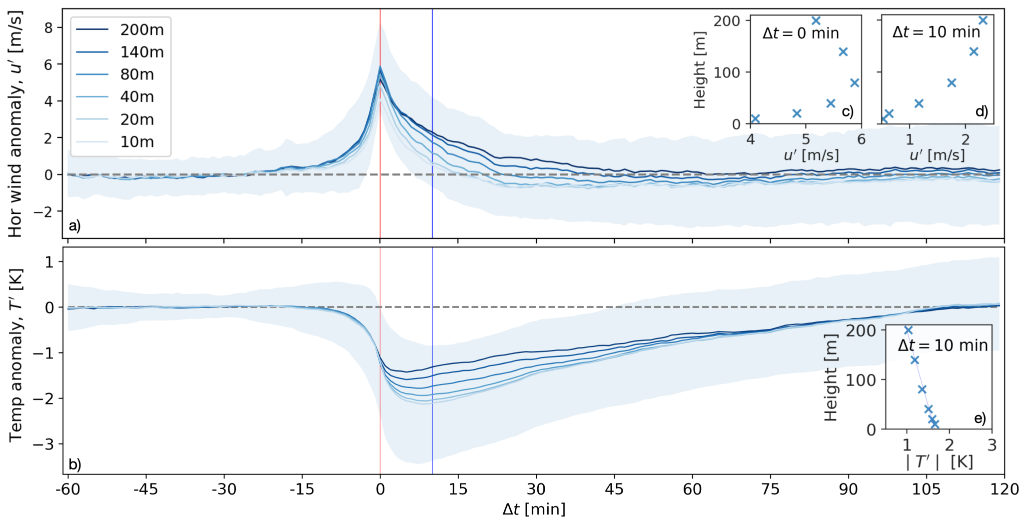


Figure 4.5: **Composite time series of horizontal wind and temperature anomalies.** Measured time series of composited cold pools, showing (a) the anomaly of horizontal wind u', and (b) the anomaly of temperature T'. These composites include 189 CPs detected in the summers 2010-2019. Blue lines show ensemble mean; blue shaded areas show the standard deviation, computed from the CP ensemble, indicating the ensemble spread between the different CPs at each given time, and computed the height of strongest signal for each variable (80 m for the horizontal wind, 10 m for the temperature). Vertical lines highlight times $\Delta t = 0\,\mathrm{min}$ (red) and $\Delta t = 10\,\mathrm{min}$ (blue). Insets show c) u' at the heights measured at the tower at $\Delta t = 0\,\mathrm{min}$; d) analogous to c) but at $\Delta t = 10\,\mathrm{min}$; e) T' at the heights measured at the tower at $\Delta t = 10\,\mathrm{min}$, with linear fit used to estimate CP height (intercept: 503.8, slope:-304.7).

of the propagation speed of the CP front and internal CP circulation (Rooney, 2018), we interpret this window as the time interval where the vortical circulation within the CP head affects the measured horizontal wind speeds. If we assume the average CP gust front to be propagating at $u \approx 0.67\, u'(t_0)_{10m} = 2.7\,\mathrm{m\,s^{-1}}$, following the relation found in Goff (1976), this would imply that the width of the CP head is approximately 1 km, considering the transit time of six minutes.

The horizontal wind anomaly is preceded by a negative temperature anomaly that occurs simultaneously and at the same rate at all tower heights. After t_0 the lower heights z show systematically deeper anomalies $T'(z)$. We linearly interpolate $T'(z)$ at 10 min towards $T'(z) = 0$, to obtain a rough estimate of the height z_0 where the temperature anomaly disappears (Fig. 4.5b, inset). We estimate $z_0 \approx 500$ m, as an indication for a height scale for the body of the composite CP. We note that the temperature anomaly is fully recovered at all measurement heights after approximately two hours from the beginning of the temperature drop.

The circulation within the CP head is further characterised by the vertical wind peak (updraft) that precedes the horizontal wind anomaly, as reflected in a positive vertical wind anomaly 1-2 minutes before t_0, that exceeds four standard deviations of the fluctuations in the time series (Fig. 4.6 b). Here, the standard deviation represents the fluctuation of the composite vertical velocity time series in the time window $[t_0 - 60 min, t_0 + 120 min]$, so although the vertical wind is noisy, the exceedance of this line indicates a clear signal of a strong updraft. The updraft signal is strongest at the highest measurement level (180 m).

The water vapor concentration starts decreasing at the same time as the vertical wind peak occurs, at both measurement heights, confirming the dry CP interior. There is not a clear signal of enhanced moisture before t_0, indicating that moisture rings may not be a evident characteristic of the CPs in this study. We note however that the water vapor concentration starts decreasing four to five minutes after the temperature has started decreasing, meaning that the CP head is more moist than the body. The moisture anomaly is recovered within one to two hours. While all previously discussed composite characteristics are mostly in line with the findings from Kirsch *et al.* (2021) for CPs over Hamburg, Germany, the moisture signal differs significantly, which show moistening rather than drying in the interior of the CPs. The absence of a wind criterion with a high threshold may lead to the inclusion of CPs measured at different points of their lifetime, or from different types of rain events, which would affect the moisture signal. Furthermore, Drager *et al.* (2020) show in their simulations that the moisture content in the interior and ahead of CP fronts crucially depends on the soil moisture: over dry soils, their CPs show an increase in moisture, similar to the observations in Kirsch *et al.* (2021), while for wet soils, the moisture signal shows the same characteristics as our measurements

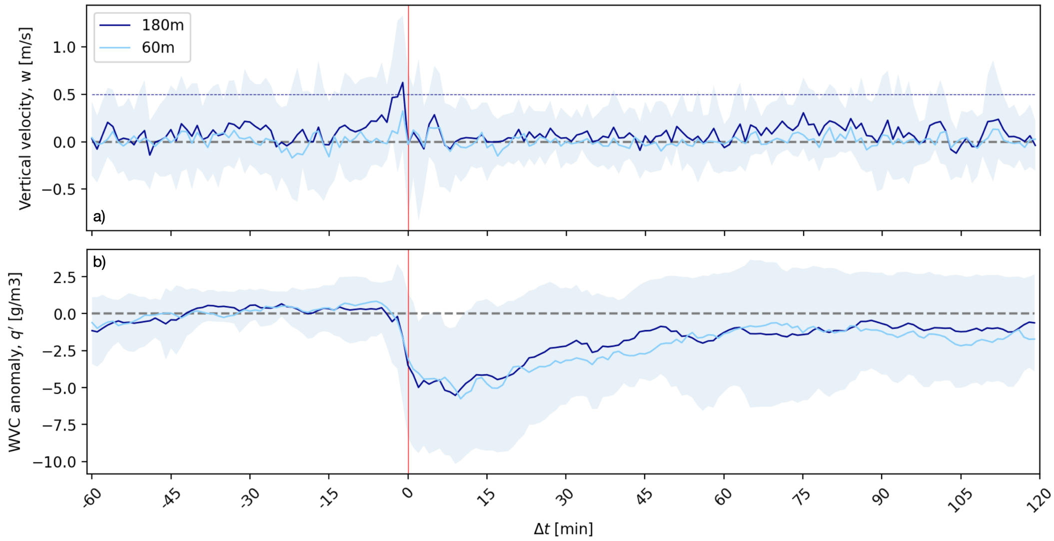


Figure 4.6: **Composite time series of vertical velocity and water vapor concentration anomaly.** Measured time series of composited cold pool, showing (a) the vertical velocity and (b) the anomaly of water vapor concentration (WVC). These composites include 18 CPs detected in the summer of 2019. Blue shaded areas in each panel show the standard deviation, computed from the CP ensembles, indicating the ensemble spread between the different CPs at a given time, and computed the height of strongest signal for each variable (180 m for the vertical velocity, 180 m for the WVC). Vertical red line highlights time $\Delta t = 0\,\mathrm{min}$.

with dry air in the interior, but with the addition of moisture rings ahead of the CP front.

In Fig. 4.7 we provide a sketched summary of the observed CP characteristics. The circulation, temperature and moisture signal indicate that the CP edge (measured at $t_0 - 5\,\mathrm{min}$) is characterized by a moist, cold, updraft; the CP head (measured at t_0) is characterized by cold, dry air and increased vorticity; and the body of the CP (measured at $t_0 + 10\,\mathrm{min}$) is characterized by dry air, with largest temperature anomalies at the surface. The thermodynamic structure of the CP interior is characterised by increased atmospheric stability of approximately $5\,\mathrm{K\,km^{-1}}$, as can be seen in the stratification of the temperature anomaly (Fig. 4.5b), where the lowest level shows the largest cooling. This stratification does not exist in the moisture signal, which appears to be homogeneous drying through the CP's height since the anomalies at 60 m and 180 m are similar in value (Fig. 4.6b). The recovery time for temperature and moisture after the passage of the CPs seems to vary strongly among CPs, as indicated by the ensemble variance in temperature and moisture anomaly (blue shading in Figures 4.5 and 4.6).

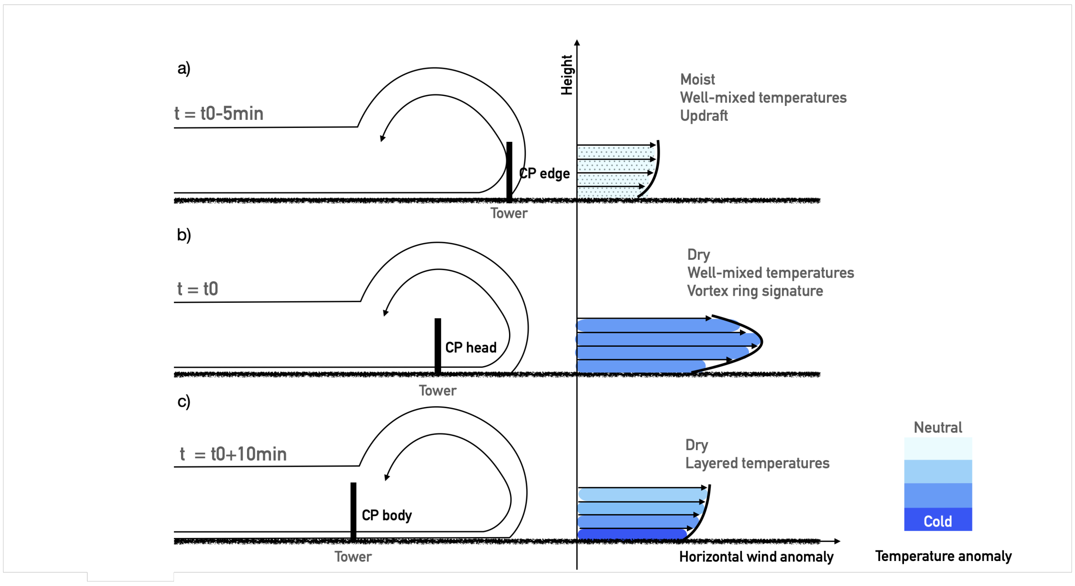


Figure 4.7: **Summarizing sketch of measured CP properties.** On the left, an approximate sketch of which part of the CP is being measured at the tower at a given point in time: (a) at $t = t_0 - 5$ min the edge of the CP is being measured at the tower; (b) at $t = t_0$ the head of the CP; (c) at $t = t_0 + 10$ min the body of the CP. On the right, a depiction of the measured CP properties corresponding to each part of the CP measured at its lowest 200 meters: the edge of the CP shows moist, well-mixed air, along with an updraft; the head of the CP shows dry, colder, well-mixed air, and the signature of a vortex ring is reflected in the horizontal wind anomaly; the body of the CP shows dry air and a layered temperature anomaly, with the largest cold anomaly within the bottom layers.

4.3.2 CP Strength

The "strength" of a CP can be characterizied dynamically, by its propagation speed, and thermodynamically, by its temperature anomaly. Here we wish to quantify the strength of an ensemble of CPs. Early studies show that for incompressible, inviscid, and irrotational (i.e., no internal motion) density currents in unstratified flows, the propagation speed u can be related to the relative density difference between the interior of the density current and its surrounding environment (Karman, 1940; Benjamin, 1968). Considering that the relative density difference can be approximated with the relative temperature difference, a general equation describing the relationship between the propagation speed u and the temperature anomaly ΔT is:

$$u = k\sqrt{gH\frac{\Delta T}{T_0}}. \tag{4.3}$$

where k is the 'internal Froude number' k (Benjamin, 1968; Wakimoto, 2001), g is the gravitational acceleration, H the CP height, ΔT the temperature difference between the CP and its environment and T_0 is the air temperature of the environment. Since CPs are density currents in a non-idealized environment, they are exposed to dissipation effects, such as surface friction and turbulent mixing, which are usually included within k. The inviscid case hereby represents a special case with $k = \sqrt{2}$, while meteorological studies have found values $k \approx 0.7$ to be more realistic (Markowski and Richardson, 2010; Wakimoto, 1982; Wakimoto, 2001). We here test the above relation, assuming a Froude number of $k = 0.7$.

For each CP, we estimate the environmental temperature T_0 from the temperature at 10 m, averaged over the time window $[t_0 - 60\,\text{min}, t_0 - 10\,\text{min}]$, as done previously. A scatter plot of the wind peak Δu_{max} against the relative temperature drop T_{drop}/T_0 for the 189 CPs detected in the ten summers 2010-2019 (Fig. 4.8) indeed suggests increasing gust front speed for larger temperature drops. Viewing Δu_{max} as a proxy for the CP's total kinetic energy density and T_{drop} as a representation of the CP's potential energy density, this indicates a monotonic relation between the kinetic and potential energy of the CP (Meyer and Haerter, 2020). We note that the median value of T_{drop} is -2.9 K and of Δu_{max} is $+4.4$ m/s. Comparing a linear least squares fit (green line), constrained to passing through zero ($\Delta u_{max} = 0$ should correspond to $T_{drop} = 0$), to a square-root least squares fit (red curve, Eq. 4.3) by inspecting the residuals (Figures 4.8b, c), indicates that the latter is a more appropriate fit, given that there is no trend in the residuals of the square root fit.

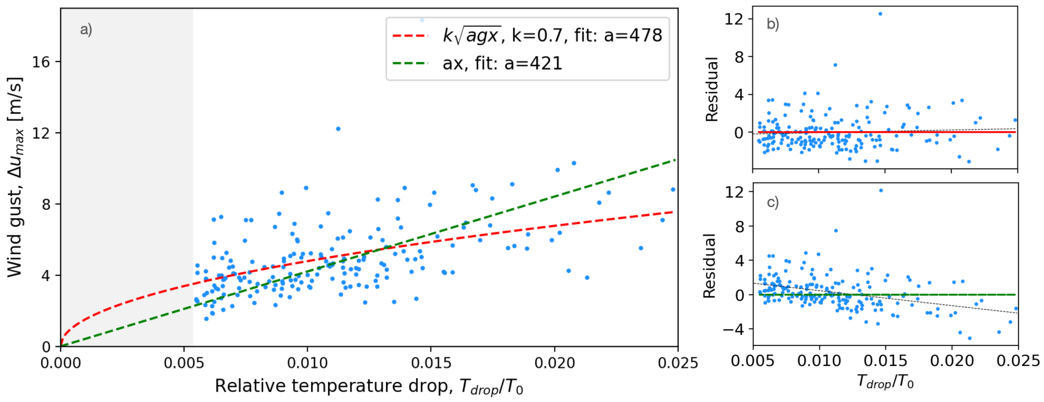


Figure 4.8: **Cold pool property relationship compared to theoretical model.** (a) Gust front strength $\Delta u_{\max}$ vs. relative temperature drop T_{drop}/T_0 for all CPs detected in 2010-2019. Considering equation 4.3, the fitting constant a can be understood as an estimate of the CP height. (b) Residuals from square-root fit (Eq. 4.3), trend shown as a thin dotted black line. Approximately symmetric spread around zero (red line). (c) Same as b, but for linear fit. Systematic tendency with an underestimation (overestimation) of $\Delta u_{\max}$ for low (high) T_{drop}/T_0.

By assuming a fixed internal Froude number $k = 0.7$, the fitting constant $a = 478\,\mathrm{m}$ can be understood as an estimate of the CP height H (Eq. 4.3). This estimate is highly sensitive to the chosen value for the Froude number k. Underestimating k will lead to an overestimation of H (and vice versa). Nevertheless, the estimated value is comparable to the CP height estimate based on the temperature anomaly extrapolation $z_0 \sim 500\,\mathrm{m}$ discussed earlier. Our value is higher than the average 300 m CP heights over tropical oceans inferred from aircraft pressure measurements (Terai and Wood, 2013) and lower than the height of 746 m found for CPs over Hamburg, Germany, with a pressure anomaly based extrapolation from tower measurements (Kirsch *et al.*, 2021). We point out that our CP height value is much lower than the 1.5 – 2 km heights found for early simulated thunderstorm outflows (e.g. Droegemeier and Wilhelmson, 1987; Liu and Moncrieff, 1996).

4.3.3 How Does Rain Intensity Influence CP Strength?

CPs develop primarily through the evaporation of rain in the sub-cloud layer (Kurowski *et al.*, 2018). The evaporation is enhanced in rain showers with high drop number density, though decreases with high ambient relative humidity (Seifert, 2008). We here test this bi-variate relationship between the CP, the parent rain cell, and their mutual environment.

Relative humidity can directly be estimated from the measured dew point depression $\tilde{T} \equiv \overline{T} - \overline{T}_d$, where $\overline{T}$ and $\overline{T}_d$ are averages of temperature and dew point temperature over the 50-min time interval preceeding CP detection ($[t_0 - 60\,\mathrm{min}, t_0 - 10\,\mathrm{min}]$) at the 140 m level of Cabauw tower. Whereas the data used in this study do not contain explicit microphysical information, rain intensity, I, is a rough proxy for rain drop number density. Its spatial average over the entire rain cell area, denoted $\langle I \rangle$, is computed at a single time step t_{RC} (Sec. 4.2.3).

While crude, this spatial average lowers the sensitivity to biases of the radar measurement, such as due to the presence of ice, artificially increasing the reflectivity locally.

Using linear regression on the data from the ten summers of CP data (2010-2019), we expand the temperature drop T_{drop} in terms of the low-order terms:

$$T_{drop} \sim \alpha_0 + \alpha_1 \langle I \rangle + \alpha_2 \tilde{T} + \alpha_3 \langle I \rangle^2 + \alpha_4 \tilde{T}^2 + \alpha_5 \langle I \rangle \tilde{T} + O(3). \tag{4.4}$$

The regression analysis indicates that significant non-linearity enters through the quadratic terms of rain intensity $\langle I \rangle^2$ and $\tilde{T}^2$, whereas the mixed term $I_{mean}\tilde{T}$ shows non-significant correlation, that is, α_5 shows very high standard error, and is thus neglected. To avoid that the simple model predicts CPs at vanishing rain intensities or in a totally saturated atmosphere, we impose the physically meaningful restriction of zero intercept, that is, $\alpha_0 = 0$. Together, we retain the fit function:

$$T_{drop} \sim \alpha_1 \langle I \rangle + \alpha_2 \tilde{T} + \alpha_3 \langle I \rangle^2 + \alpha_4 \tilde{T}^2 + O(3). \tag{4.5}$$

We here compare a linear regression, where $\alpha_3 = \alpha_4 = 0$ and a non-linear regression, where α_3 and α_4 may vary (Tab. 4.1). In both models, the CP strength is positively correlated with the dew point depression $\tilde{T}$ and mean rain intensity $\langle I \rangle$, confirming that CPs are measurably strengthened by drier environmental conditions and larger rain intensities. In the non-linear regression, all four remaining regression coefficients are found to be statistically significant. We speculate that the negative dependence on $\langle I \rangle^2$ ($\alpha_3 < 0$) stems from very strong rain quickly saturating the sub-cloud atmosphere, thus diminishing further rain evaporation.

To give a more tangible interpretation of the linear multivariate regression, we note that with the coefficients in Table 4.1, a CP becomes 1 K colder if the rain intensity of its parent cell is incremented by 4 mm h^{-1} (instantaneous area-averaged rain intensity) or if the ambient relative humidity is increased according to a 2.5 K larger dew point depression. We compare with a recent study by Kirsch *et al.* (2021) for CPs over Hamburg, Germany, hence a similar geographic region. The study shows two separate regressions to determine the relationship between CP temperature perturbation and point-measured accumulated rainfall, and CP temperature perturbation and pre-event saturation deficit. CP strength is found to increase with increasing point-measured rainfall, and with increasing pre-event saturation deficit, in line with our model.

	linear regression		non-linear regression	
	coeff.	std. error	coeff.	std. error
α_1 $[K\,h\,mm^{-1}]$	0.25	0.03	0.44	0.09
α_2 [-]	0.40	0.03	0.51	0.08
α_3 $[K\,h^2\,mm^{-2}]$			−0.020	0.007
α_4 $[K^{-1}]$			−0.021	0.007
R-squared (uncentered)	0.89		0.92	

Table 4.1: **Fitted coefficients for the linear and non-linear regression models (Eq.4.5)**. α_1 and α_2 represent the coefficients of the terms linear in rain intensity and temperature respectively, whereas α_3 and α_4 represent the coefficients of the terms quadratic in rain intensity and temperature respectively.

4.4 Conclusions

In this study we design and validate a methodology to detect convectively-generated cold pools (CPs) and their gust fronts over the Netherlands and relate them to their parent rain cell and environment. CP characteristics have been studied over many years from an observational point of view, however there are very few statistical studies of CPs over mid-latitude coastal land. Our study stands out by combining tower and radar measurements to analyze 100+ CPs in relation to their generating rain cells. We wish to highlight the following findings:

- the patterns in the horizontal and vertical wind measurements confirm the existence of a vortex ring in the CP head;
- the detected CPs show weak or absent moisture rings, while the CP interior shows a negative moisture anomaly;
- a simple model consisting in a multi-variate linear combination of pre-event dew point depression and area-averaged rainfall intensity allows a prediction of the generated CP's strength.

To study the evolution of CPs in time and their properties in relation to their parent rain cell and environment, we use the combination of local measurements from a 213-m meteorological tower located in Cabauw, the Netherlands, and precipitation radar output. To identify CPs from time series of point measurements, an existing algorithm for detecting CPs over the ocean (Szoeke *et al.*, 2017) is extended to capture CPs over land by (i) an increased threshold on the temperature anomaly (1.5 K) and (ii) an additional criterion on the horizontal wind speed anomaly. Studying a few exemplary CPs revealed a vertically coherent signal in horizontal wind velocity ("wind peak") at the tower during the passage of a CP gust front. Therefore, in our algorithm we record events as CPs if they are characterised both by a temperature drop and a coherent wind peak. This allows us to isolate the CPs from the temperature fluctuations over land which come without, or with a very weak, wind signal. The algorithm shows low sensitivity to the threshold on temperature anomaly but high sensitivity to the wind peak threshold. We thus recommend to keep the threshold on temperature at 1.5 K, whereas the threshold on the wind peak should be varied in accordance with the threshold on precipitation intensity of the parent rain cells under consideration. Our choice of parameters may bias the algorithm to detecting only strong CPs whose fronts are often found to be very close to the parent rain cell at the time of detection, indicating a young age of the CPs at the time of detection or a squall-line like system, where the CP is advected along with the cloud. However, we find that, by this constraint, it is ensured that all identified cases can be attributed to the passage of a CP rather than other forms of fluctuations. Our method's reliability is confirmed by the associated updrafts and succeeding dry moisture anomalies measured for all detected CPs.

The composites of 189 CPs from measurements taken during ten summers (May-September 2010-2019) allow to study statistical CP properties and their radial structure. The CP gust front is characterised by a strong updraft, which temporally coincides with the beginning of the negative temperature anomaly, and precedes the positive horizontal wind anomaly and negative moisture anomaly. The presence of a vortex ring on the edge of the CPs is confirmed by the changing signal in horizontal wind measurements at different heights, in the window of time around CP passage. In contrast to studies of oceanic CPs (Szoeke *et al.*, 2017; Zuidema *et al.*, 2017), we do not detect a clear signature of moisture rings, that is, areas of elevated moisture, on the edges of the CPs. This might be due to the smaller magnitude of latent heat fluxes over land with respect to the tropical ocean (Drager *et al.*, 2020). The absence of CP moisture rings in our observations suggests that for our ensemble of CPs, the thermodynamic triggering of new rain events is less important than the mechanical triggering of new rain events, particularly driven by the collision of strong gust fronts (Tompkins, 2001b; Torri *et al.*, 2015; Drager *et al.*, 2020). The interior of the CPs in our analysis is characterized by enhanced horizontal wind anomalies, and longer-lasting cold, dry air. The dry anomaly stands in contrast to a recent study of CPs over Germany (Kirsch *et al.*, 2021), which finds enhanced moisture in CP interiors. The thermodynamic anomaly recovers within two hours of the passage of the gust front, which is clearly shorter than the convectively active time in a day. It is, therefore, possible for multiple CPs to occur in the same location within one day, despite the inhibiting effect of a cold, dry lower boundary layer on the formation of new convection. We do indeed, in some cases, detect more than one CP on the same day. Regarding the scale of CPs, our current results indicate that typical dynamical gust fronts have a width of 1 km. For the CP height, we extrapolate two values from (i) the vertical gradient in temperature anomalies measured at different tower levels, and (ii) the relationship between CP temperature anomaly and CP speed. Both estimates (~500 m) are lower than the typical CP height over continental land, and higher than the typical CP height over the ocean. This might be due to the coastal location which provides an interesting environment to study CPs that are neither fully oceanic nor fully continental. Furthermore, the examination of radar images and attribution of generating rain events reveals that some CPs that we detect emerge from large-scale weather patterns with lower rain intensity than convective thunderstorms. We suspect that this contributes weaker and shallower CPs to the ensemble.

By combining the tower measurements with the radar data and using a wind direction criterion, we were able to attribute a parent rain cell to the majority of detected CPs. This allows to confirm the positive relation between CP strength and precipitation intensity with a bi-variate, linear regression of CP temperature anomalies against the pre-CP dew point depression and rain intensity averaged over the full rain cell area. Our simple model allows the prediction of CP strength (temperature anomaly), conditioned on both microphysical and environmental parameters. Knowledge on the relation between a given population of rain cells and the CPs initiated by any of them is important as a benchmark for numerical simulations, both idealised studies and comprehensive high-resolution climate models.

Outlook. Numerous recent studies have highlighted the importance of CP effects in structuring the convective cloud and precipitation field over space and time (Rio *et al.*, 2009; Böing *et al.*, 2012; Schlemmer and Hohenegger, 2014; Szoeke *et al.*, 2017; Böing, 2016; Haerter and Schlemmer, 2018; Haerter *et al.*, 2019; Haerter *et al.*, 2020). Further observational studies along the lines presented here may help clarify pressing modelling questions, such as how parameterized CPs are affected by large-scale weather conditions, or the numerical grid resolution required to appropriately resolve the gust fronts of spreading CPs. Future research is required to study the last step in the causal chain of convection, namely the attribution of the triggering of new convective cells to the detected CPs and their parent rain cells. The current study, and work following up on it, can help improve the realism of models and decipher how mesoscale convective systems build up dynamically and which role CPs play in correlating the dynamics of the individual rain cells involved.

4.4.1 Appendix Figures

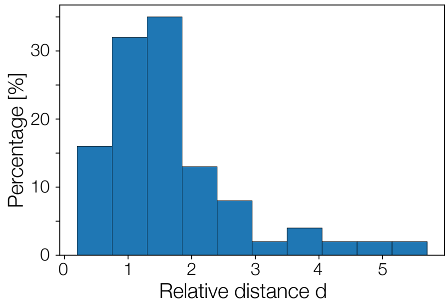


Figure 4.9: **Relative distance d between detected CP gust front and the rain cell edge for all CPs detected in the 10 summers 2010-2019.** The CP gust front is detected directly at the edge of the rain cell for $d = 1$, ahead of the rain cell for $d > 1$ or below the rain cell, i.e. it rains at the tower's location, for $d < 1$.

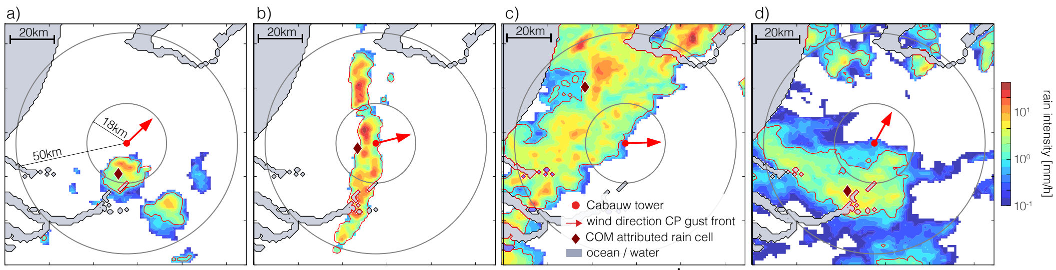


Figure 4.10: **Rain intensity from radar data, exemplifying different weather situations at moments of CP-detection.** Note the proximity of the rain cells to the Cabauw tower (red dot). Red contours indicate the intensity threshold used to track the rain cells. a) Isolated rain cells (2019/08/19, 15:25), b) Squall-line like convective cell (2019/05/08, 15:25, c) Extensive front approaching from the Atlantic (2018/08/24, 21:30), d) Mesoscale convective system (2019/06/07, 15:10). In cases a and b, the isolated rain cells with well-defined COM allow a unique attribution, whereas in cases c and d the COM can be ill-defined and the attribution is more ambiguous.

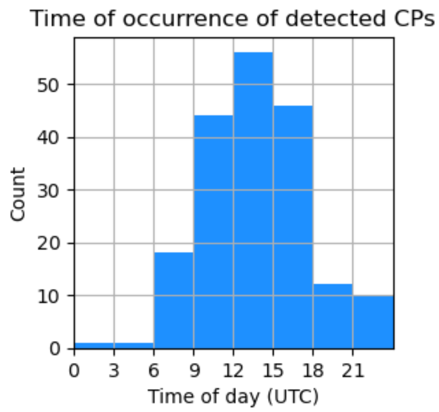


Figure 4.11: **Time of occurrence of all CPs detected in the 10 summers 2010-2019.** Note that the local summer time in the Netherlands is CEST, corresponding to UTC+2.

5

Nowcasting Cold Pools in Dakar with Long Short-term Memory Networks: The First Two Pillars

Before proceeding to the last study of this dissertation, I'd like to here acknowledge Leif Denby at the Danish Meteorological Institute for introducing me to the exciting field of machine learning and guiding me through my first AI-driven steps, Edward Engelbrecht for running the simulations in this chapter, Jan Haerter for making the DakE field campaign possible, Yahaya Bashiru for sparring with me in Bremen and Dakar, and the rest of the Complexity and Climate group at Leibniz ZMT for station testing and logistics. A large token of appreciation goes to the Physics Laboratory of the Atmosphere and the Ocean at UCAD, in particular Abdou Lahat Dieng, Salif Diedhiou, Dame Gueye and Noreyni Fall; and Abdoulahat Diop at ANACIM, for sharing their invaluable knowledge and expertise with me while in Dakar.

Figure 5.1: Mesoscale Convective System off the coast of Dakar, September 21 2023, 19:31 GMT. Photo taken while in front of the University of Dakar Cheikh Anta Diop after installing one of two weather stations discussed in this chapter.

Abstract

Frequent floods during the rainy season in Senegal underscore the need for advanced weather monitoring capabilities in tropical West Africa. Currently, the region lacks high-resolution temporal and spatial weather monitoring instruments with accessible data. In response, we have planned a field campaign in Senegal, of which the first two pillars are two automatic weather stations in Senegal, located in Pout and Dakar, operating at a one-minute temporal resolution and transmitting near real-time data to a cloud infrastructure.

Our research focuses on the development of a nowcasting tool to predict the arrival of cold pools (CPs) in Dakar. To achieve this, we employ Long-Short Term Memory networks (LSTMs) trained on simulated time series data generated from Weather Research and Forecasting (WRF) simulations of the 2019 rainy season over Dakar, mimicking automatic weather station data in Pout and Dakar. The LSTMs are thus trained to capture the intricate patterns associated with the onset of CPs and have notable skill in nowcasting the arrival of CPs in Dakar with a 30 min warning time, providing crucial lead time for mitigating the impact of strong wind gusts and heavy rain leading to potential floods. Our approach represents a novel application of advanced machine learning techniques to enhance short-term weather prediction capabilities in a region prone to weather-related hazards.

The integration of high-resolution weather stations and LSTMs holds promise for improving the accuracy and lead time of real weather predictions, particularly in regions with limited monitoring infrastructure.

5.1 Introduction

Meteorological monitoring and accurate short-term weather predictions are critical for various sectors globally, ranging from agriculture to disaster management. Sub-Saharan Africa in particular faces a pressing need to enhance predictions of high-impact weather events, primarily due to the frequent occurrence of intense convective storms in the tropical band leading to severe flooding, strong winds, and lightning. These hazards pose significant threats to humans, infrastructure, and the economy. Currently, numerical weather prediction in Africa lacks accuracy, especially for lead times less than 24 hours, emphasizing the critical necessity for immediate event prediction, known as nowcasting (Roberts *et al.*, 2022).

Nowcasting services are generally lacking across Africa, with a notable absence of automated nowcasting systems or tools. This deficiency hampers the ability of national meteorological services to issue timely warnings, potentially resulting in loss of life and substantial economic losses. Very recent research on nowcasting in this area has introduced innovative methods for generating probabilistic nowcasts of convective activity using satellite imagery (Anderson *et al.*, 2023). Another strategy for nowcasting mesoscale convective systems (MCSs) involves utilizing satellite-derived land surface temperature data as an indicator of soil moisture levels, given the strong correlation observed between dry soils in the early hours of the day and MCS activity in the Sahel (Taylor *et al.*, 2022). Our study takes on a more localized approach by harnessing novel measurements collected from two automatic weather stations installed in Pout and Dakar, Senegal, to enhance our understanding of local weather patterns, with a specific emphasis on nowcasting gust fronts from storms in Dakar.

Senegal, situated in West Africa and in the Sahel, experiences a diverse climate with a pronounced rainy season spanning from June to October, during which organized convective storms contribute to 90% of the seasonal rainfall (Mathon *et al.*, 2002). These storms lead to the formation of cold pools (CPs), cold downdrafts generated from the evaporation of precipitation, and their associated gust fronts, sudden increases in wind speed at the leading edge of a CP. Monitoring and predicting these gust fronts, especially in large urban areas like the capital Dakar, with a population of 3.9 million, are crucial for effective risk management and early response (Merz *et al.*, 2020).

Figure 5.2: Aftermath of rain events in Dakar on (left) the morning of September 20th, 2023 (photo taken at 8:45 UTC), and (right) the morning of September 25th, 2023 (photo taken at 09:30 UTC). Photos taken in the Point E district of Dakar.

It is noteworthy that Dakar and its surroundings are particularly vulnerable to flooding during the rainy season (Mbow *et al.*, 2008). The combination of convective storms and the geographical characteristics of the area contributes to a heightened risk of flooding, impacting local communities and infrastructure (Figure 5.2). Monitoring and predicting weather events, including gust fronts, hold particular significance in flood-prone regions for timely and effective disaster preparedness.

Our research leverages data obtained from two automatic weather stations installed in Senegal – the first two pillars of the ongoing field campaign DakE (Dakar East). The weather stations were installed in the town of Pout, East of Dakar, and the capital city of Dakar in September 2023. They are approximately 50 km from each other: the former upstream and the latter downstream of convective systems traveling westward with the African Easterly jet. The two stations installed provide high-resolution meteorological data, including atmospheric pressure, temperature, humidity, wind speed, and direction. The utilization of this detailed observational data opens avenues for a deeper understanding of local weather phenomena. Recent techniques have been developed and employed to detect and analyse CPs from observational time series from land-based weather stations (Kirsch *et al.*, 2021; Kruse *et al.*, 2022), the former used in particular for ground based automatic weather station networks set up to study CPs (Hohenegger *et al.*, 2023).

In this study, we use Weather Research and Forecasting (WRF) simulations at 1km resolution to generate synthetic station data for one rainy season, and we extract the time series for the locations of Pout and Dakar. The simulated data is used to train the Long Short-Term Memory (LSTM) networks. LSTMs, a type of recurrent neural network (RNN) known for capturing temporal dependencies, are employed here for nowcasting gust fronts associated with convective storms in Dakar. The LSTM network's ability to learn patterns and relationships within time series data makes it a promising tool for nowcasting. By training the network on simulated data, and successively testing it on observed data from the automatic weather stations, we aim to develop a robust model capable of accurately forecasting the onset of CPs, contributing to improved weather-related decision-making in Dakar.

5.2 Methods and Results

5.2.1 Measuring CPs with Automatic Weather Stations

The automatic weather stations we used to measure CPs are described in Table 5.1. Both stations were set up to send data to a cloud via the cellular network, at a 1-minute temporal resolution and are powered by a combination of batteries and solar panels, to have the option to be off the grid.

Testing the Automatic Weather Stations in Germany

The Davis VantagePro2 and Meter Atmos41 stations were first tested in a field in Bremen, Germany, for the summer season of 2023, before being transported to Senegal (Figure 5.3). During the testing phase in Germany, we investigated the

Figure 5.3: **Automatic Weather Stations.** Testing the automatic weather stations in Bremen, Germany. On the left: VantagePro2; On the right: Atmos41

ease of deployment and the differences in measurements capabilities. The stations are comparable in price range with the full set-up. With our set-up, both stations can measure at a 1-min temporal resolution, the following variables: temperature, wind speed, rain, pressure, humidity, solar radiation and soil moisture at several levels, with the Atmos41 additionally measuring lightning strikes. We note here that the Atmos41 has a drop counter for precipitation measurements, which captures low rain rates more accurately, but measures systematically lower rain rates, than the VantagePro2 which has a tipping spoon mechanism. Furthermore, the sonic wind anemometer on the Atmos41, during very high rain rates and strong wind, can experience measurement issues most probably due to water entering the sonic sensor. We found both stations to be adequate for further testing in Senegal, since the issues mentioned above were only discovered later, in tropical conditions. In

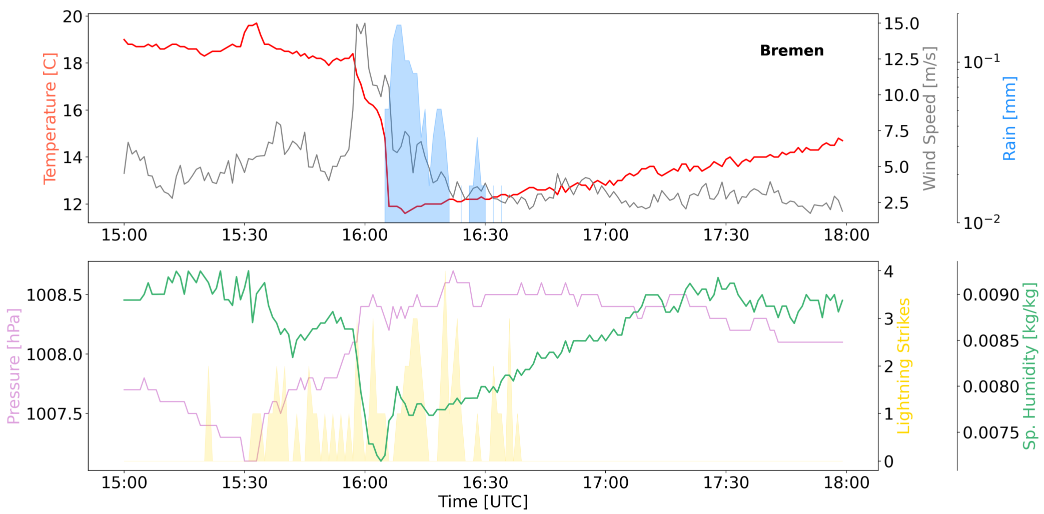


Figure 5.4: CP measured in Bremen. CP event measured in Bremen, August 27, 2023. Weather station: Atmos41. Temperature, wind speed and rain (top panel), pressure, lightning strikes and specific humidity (bottom panel).

Figure 5.4 we show a CP event measured in Bremen by the Atmos41 during the testing phase. The event was measured by both automatic weather stations, set up on 2-meter poles side-by-side in a field (coordinates: lat = 53.15, lon = 8.77). We see the temperature drop by about 6 degrees Celsius, a wind gust of about 10 m/s, followed by precipitation for about 30 minutes. The interior of the CP is dry, with the dry anomaly recovering after about 1.5 hours. There is a clear increase in pressure collocated with the decrease in temperature. Interestingly, the lightning strike measurements serve as a 30-minute warning to the rain event.

Measuring a CP event upstream and downstream of Dakar, Senegal

The stations were installed in Senegal in the second half of September 2023, capturing the last few weeks of the rainy season including several rain events, and continuing to measure to this day. The locations of the stations, Pout and Dakar, are shown in Figure 5.5 with details in Table 5.1. The town of Pout is upstream of Dakar with respect to the MCSs that tend to travel westward in the African Easterly Jet.

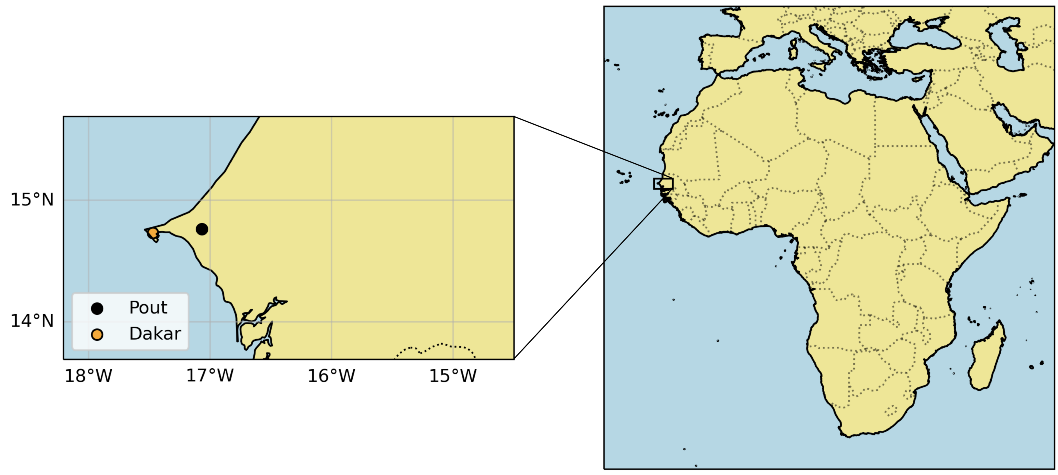


Figure 5.5: **Domain of study.** On the left, zoom over Senegal. Locations of automatic weather stations in Pout (black dot) and Dakar (orange dot).

Location	**Coordinates (lat, lon)**	**Station Type**	**Resolution**
Pout	(14.76, -17.07)	Davis VantagePro2	1 min
Dakar	(14.68, -17.47)	Decentlab+Meter Atmos41	1 min

Table 5.1: Coordinates, Station Type and Temporal Resolution for first two stations of DakE.

We successfully measured one CP event that occurred both in Pout and in Dakar, with the two respective weather stations, on September 25, 2023. The MCS that produced the CP, as seen from satellite, is shown in Figure 5.6. The cloud-top temperature shown is from thermal infrared images from Meteosat Second Generation (MSG). The convective cores, which identify the coldest areas within large storm clouds and responsible for intense rain rates are visualized with blue outlines thanks to the algorithm developed by Klein *et al.* (2018). The MCS, with a radius of roughly 200 km, traveled from East to West of Senegal in the early morning hours of September 25, 2023, reaching Dakar at around 7:30 UTC.

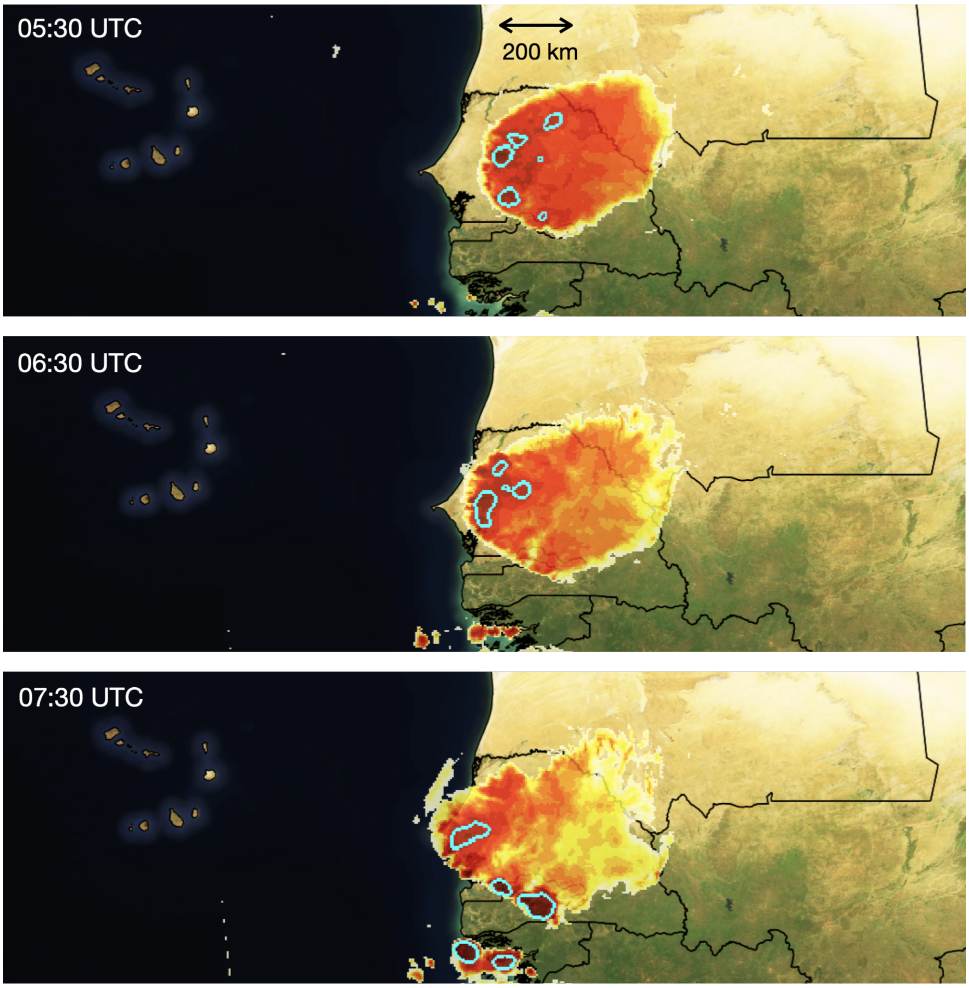


Figure 5.6: **MCS over Senegal from satellite imagery.** MCS over Senegal on September 25, 2023, as seen from infrared satellite data (UKCEH Nowcasting Portal, Accessed: [March 2024]). Deep convective cores from Klein *et al.* (2018), indicators of precipitating regions, are outlined in light blue. The associated CP was measured in Pout and Dakar in Figure 5.7. The associated flooding in Dakar can be seen in Figure 5.2.

From the station measurements in Figure 5.7, it is clear that the CP associated with the MCS in Figure 5.6, is measured first in Pout at around 7:00 UTC, and then in Dakar at around 7:30 UTC.

There is a temperature drop of about 6 K, an increase in wind speed of 10 m/s, and precipitation measured in both locations, lasting for 20-30 minutes. The interior of the CP is dry (the difference between the specific humidity before and after the onset of the CP is about 3 times larger than for the CP measured in Bremen),

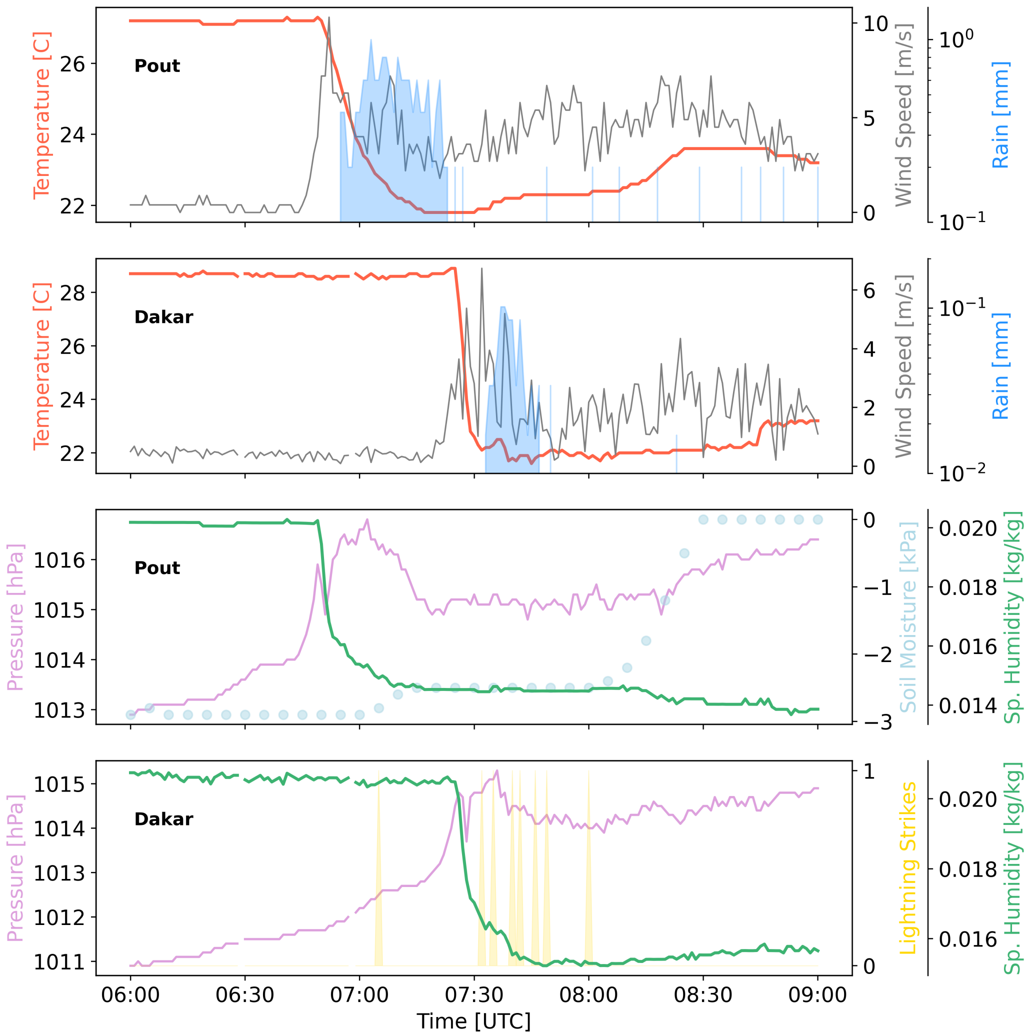


Figure 5.7: CP measured in Senegal. CP measured in Pout and Dakar with VantagePro2 and Atmos41 respectively, on September 25, 2023. Temperature, wind speed and rain in top two panels. Pressure and specific humidity in bottom two panels. Note: Soil moisture (measured at -0.1 m) is additionally shown for the Pout station and lightning strikes for the Dakar station, based on availability of measurements.

and this anomaly does not recover in the 2-hour time window visualized. The pressure increases, the soil moisture increases after the rain event, and there is a signal of lightning strikes in Dakar, that precedes the CP by 30 minutes. This is our first measurement of a CP in Senegal at 1-min resolution, and we are waiting to measure more. We want to build a nowcasting algorithm by training a machine learning model to nowcast these events in Dakar before they happen. However, to train a machine learning model like an LSTM, we need many more examples of CPs in Senegal. For this we turn to high-resolution cloud resolving simulations with realistic boundary conditions, to simulate a full rainy season over Dakar, and extract simulated "station data" to train our LSTM. The new measurements that will be obtained in the rainy season of 2024, will be further used for training, validation and testing.

5.2.2 Simulating Station Data with WRF for training LSTM networks

Simulation set-up

The simulation data used to train the LSTM networks were obtained using the non-hydrostatic Advanced Research WRF model version 4.3 (Skamarock *et al.*, 2021). The simulation period spans a 7 month period encompassing a full rainy season, from April 1 to November 30, 2019, over Senegal in West Africa, with the simulation domain shown in Figure 5.8. The rainy season of 2019 was chosen primarily for the high frequency of mesoscale convective systems. We applied a nesting approach in order to obtain numerical simulation data at a relatively high grid resolution, in this case 1 km. To achieve this, we created three domains at grid spacings 9, 3, and 1 km with a one-way nesting approach. We use only the data from the 1km resolution simulation in this study. We choose the following variables to output at a 1-min temporal resolution, to best-represent the weather stations: 2-meter temperature (T_{2m}), 2-meter specific humidity (q_{2m}), zonal and meridional components of the 10-meter wind speed (U_{10m} and V_{10m}), which can be combined to represent wind speed, surface pressure (p_{SURF}), and rain.

National Oceanic and Atmospheric Administration / National Centers for Environmental Prediction Global Forecasting Model (GFS) analysis and t+3 hourly historical forecast data were used to drive the initial and boundary conditions in the 9km outer domain. It is therefore one continuous forecast, restarted at weekly model simulation intervals, with the aim to preserve rainfall continuity. This set-up allows model nests to develop and maintain their internal dynamics, while boundary conditions at or close to the global initialization times are provided to the outer domain (D1).

To represent land surface fluxes, we used the Noah Land Surface Model scheme with soil temperature and moisture at four layers (Niu *et al.*, 2011). 55 vertical

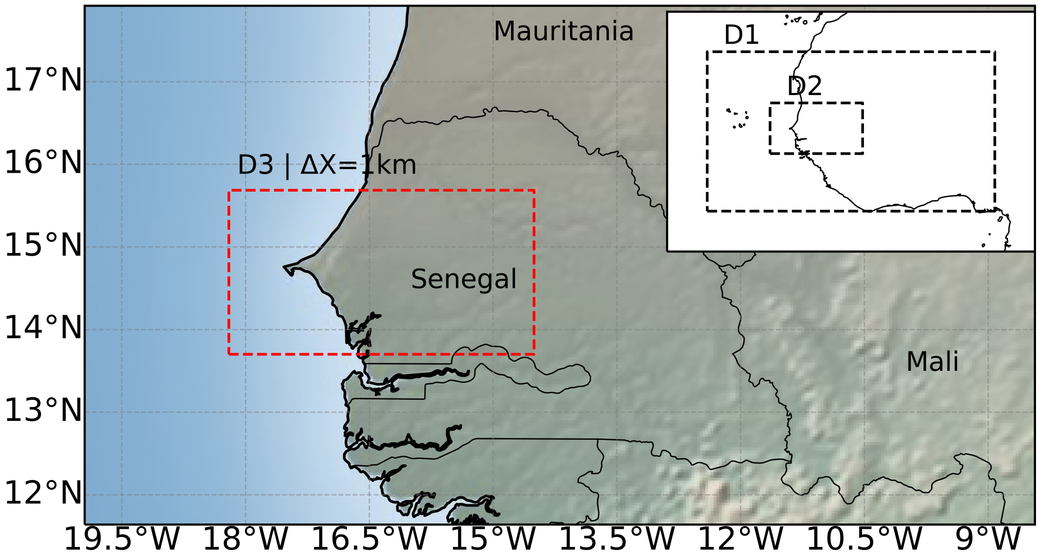


Figure 5.8: **WRF set-up.** Weather Research and Forecasting (WRF)-nested domains for the Senegal rainy season of 2019. Map of West Africa (top right corner) showing the outer domain (dashed), denoted as D1 at 9 km horizontal resolution and the second domain of 3km denoted as D2. The principal domain used in this study is D3 which denotes the nested domain boundaries (red dashed) at horizontal resolution of 1km.

atmospheric levels were chosen, such that their vertical spacing decreases closer to the surface. The outer 9km domain used the Kain-Fritsch convection scheme with a mass-flux approach, while domains 2 and 3 were run with no convective parameterization. To parameterize the Planetary Boundary Layer (PBL) we used the Yonsei University PBL scheme for all domains.

Deep moist convection, in particular MCSs, can be strongly influenced by heterogeneous soil moisture, particularly mesoscale dry soil moisture anomalies over the Sahel (Taylor *et al.*, 2012; Klein *et al.*, 2018). To maintain a high quality soil moisture state, we start the model simulation three months ahead of the rainy season to allow for spin-up time of modeled soil moisture. NOAA National Center for Environmental Prediction (NCEP) optimum interpolation sea surface temperature (SST) analysis data, provided at weekly intervals, was used for boundary conditions of the sea surface (NCEP, 1986).

Analyzing the simulated station data

We then extract two locations from the simulation, corresponding to the coordinates of the Pout and Dakar stations. The rain accumulated over the full simulation in the two locations, is shown in Figure 5.9. The simulated rainy season spans from mid-July to late October, with single rain events contributing to large increases in rain at a time. The step-like cumulative function is a clear signal of large MCS rain events contributing to most of the rain in a rainy season, as expected. Interestingly, the Dakar location received about half the amount of rain compared to Pout, in the simulation. This could be due to the fact that some MCSs passing over Pout do not make it to Dakar, either because they are deviated towards the North/South, due to interactions with sea-breezes or because they dissipate before reaching the coast. Most importantly, this means that the CPs of some of the MCSs measured in Pout, will not make it to Dakar.

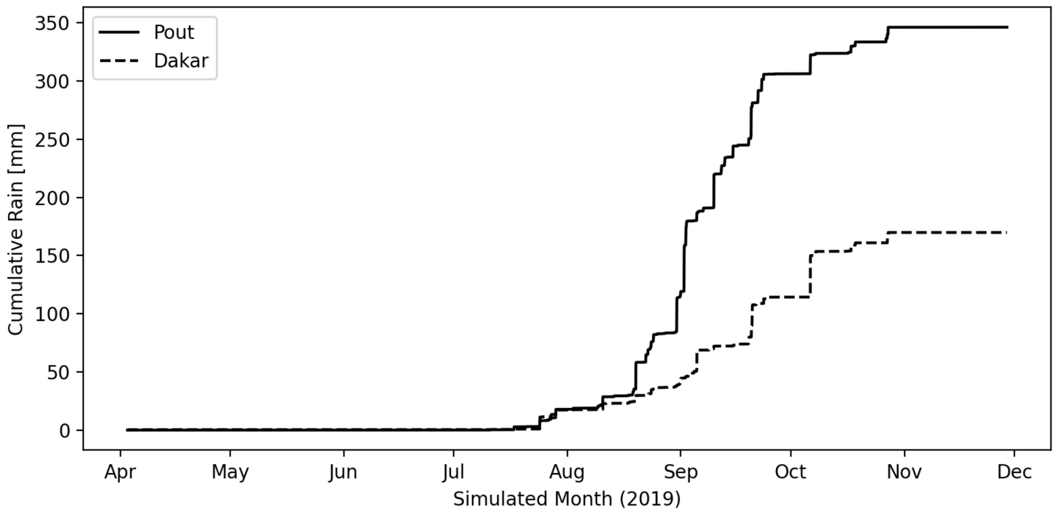


Figure 5.9: **Cumulative rain in Pout and Dakar from simulation.** Cumulative rain in WRF simulation of rainy season over Senegal in 2019, at the two locations of interest.

CPs in the simulation output. We now analyze the CPs in the WRF simulation output, as "measured" in Pout and Dakar. We run a CP detection algorithm tailored for ground station data (see Kirsch *et al.* (2021) for details on the algorithm) on the time series of 1-min simulated station output from both Pout and Dakar, and analyse the key variables related to CPs, as in Kruse *et al.* (2022). The algorithm detects CPs in time series when there is a temperature drop larger than 2K, with the criterion that this drop is followed by precipitation. In Figure 5.10 we see the composite time series of all the detected CPs in the simulation output for Pout and Dakar. We find a total of 56 CP events partitioned as follows: 42 in Pout and 14 in Dakar. The median temperature drop is -3.3 K. The significant disparity between the number of cold pools (CPs) detected in Pout and the notably fewer CPs detected in Dakar reflects the cumulative rainfall patterns illustrated in Figure 5.9, suggesting that Pout experiences a higher frequency of rain events than Dakar. The anomalies are computed with respect to the 60-min mean of the variable, preceding the temperature drop at time 0. The CPs exhibit the expected behavior for all variables: a sudden drop in temperature, of approximately 3 K, an increase in wind speed and an increase in surface pressure. The moisture also drops, indicating a dry interior, however the standard deviation is large, indicating CPs ranging from very dry interiors, to moist.

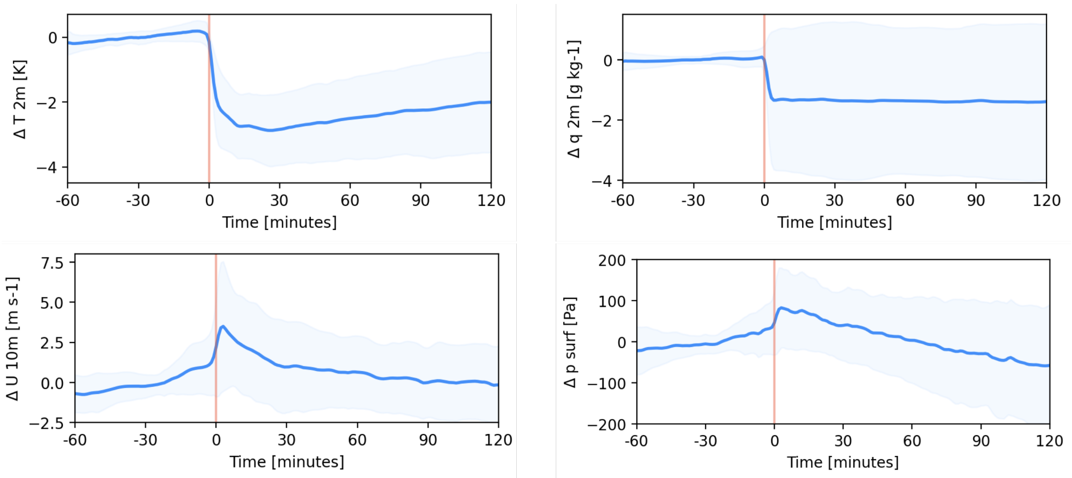


Figure 5.10: **Analysis of CPs from simulation.** Composite time series of 2-meter temperature, 2-meter moisture, 10-meter horizontal wind speed and surface pressure anomalies, related to the passage of 56 CP events (from WRF simulation output of rainy season of 2019). Shading indicates one standard deviation.

We focus on one case, to compare to the measured data in Figure 5.7. We take a CP from a large MCS over Senegal, occurring on September 23, 2019 in our simulation. The 2D snapshots of 2-meter temperature at three time steps are show in Figure 5.12. The CP is progressing from East to West, reaching Pout at approximately 13:30 UTC, and Dakar at approximately 14:00 UTC, so with a lag of 30 minutes, similar to what we had seen in the real measurements. When looking at the time series extracted from the simulation in the locations of Pout and Dakar, in Figure 5.12, we see a CP structure that is very similar to our one measured CP, albeit smoother curves due to the coarse nature of the simulation. In both locations, with a lag of 30 minutes, there is a sudden temperature drop in each location consecutively of approximately 5 K, a collocated wind gust of 10 m/s, an increase in pressure, a decrease in moisture, and subsequent rain for 30-60 minutes.

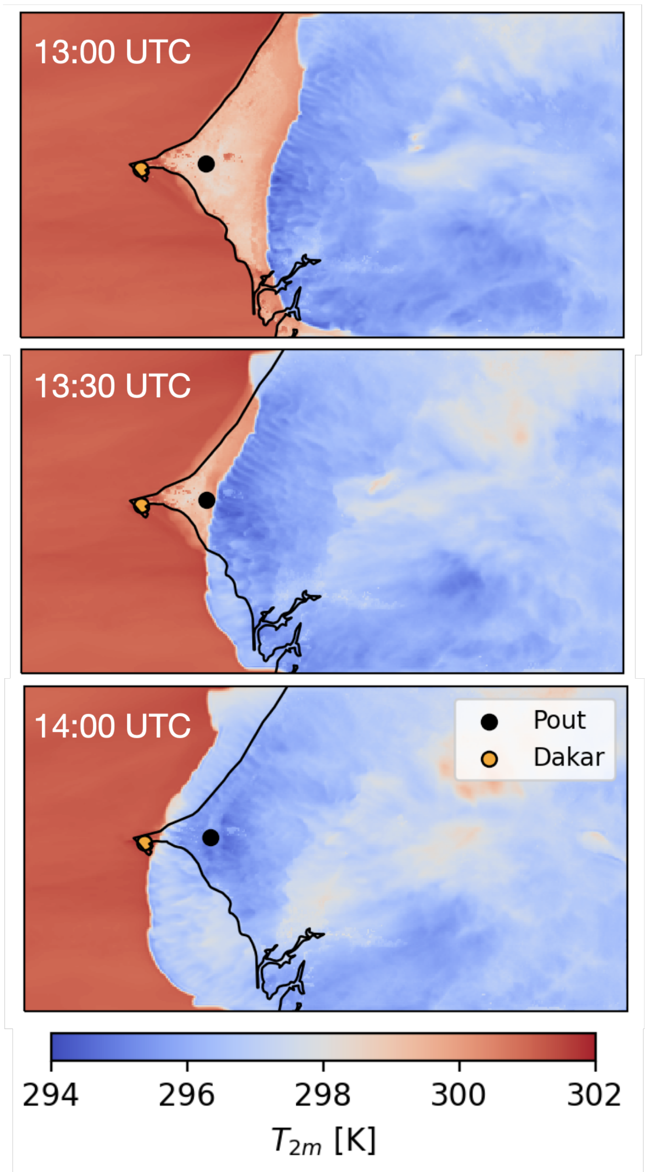


Figure 5.11: **CP from simulation over Senegal.** 2 meter temperature field of CP event over Senegal from WRF simulation output. Simulated date: September 23, 2019. Time series of same CP in Figure 5.12.

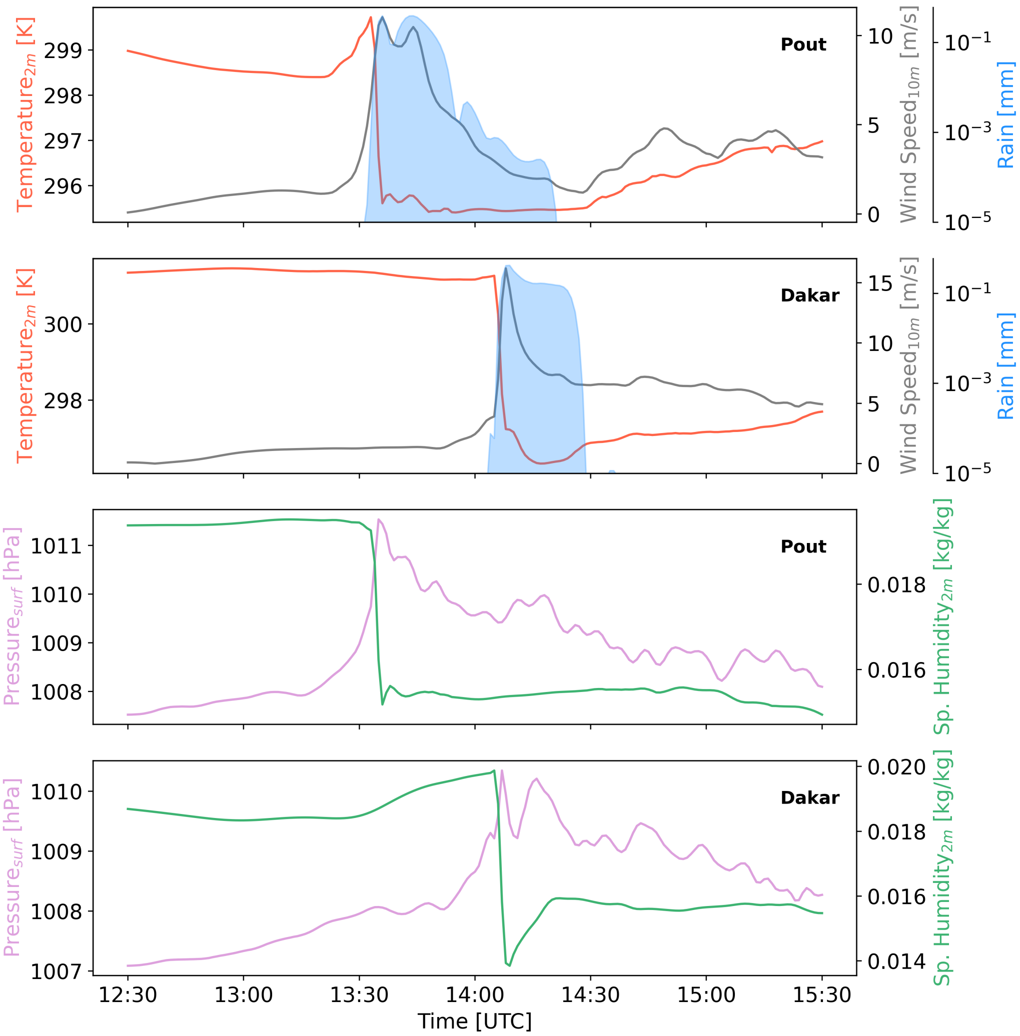


Figure 5.12: **CP from simulation over Senegal: time series.** Time series of CP event over Senegal from WRF simulation output. Simulated date: September 23, 2019. 2D snapshots of same CP in Figure 5.11.

5.2.3 Training an LSTM to Nowcast CPs

LSTM set-up

We train an LSTM to predict the advent of CPs in Dakar, with time series from Dakar and Pout. The dataset comprises the time series from the simulation output of the two locations corresponding to the stations of Pout and Dakar. Since we are interested solely in the rainy season, we extract the time series ranging from July 1, 2019 to October 31, 2019. We aggregate the 1-minute time series data by computing the average value over non-overlapping 10-minute intervals. This process involves grouping consecutive 10-minute time intervals and calculating the mean value of the data points within each interval. The resulting aggregated dataset represents a temporal resolution of $\Delta t = 10$ minutes, providing a broader overview of the underlying trends and patterns while reducing the granularity of the native 1-minute data.

Data Processing. We pre-process the dataset to make it compatible with the model architecture. This involves two primary steps: feature selection and feature normalization. Feature selection entails identifying and choosing the variables or features from the dataset that hold significance to detecting CPs. We choose the following five variables: 2-meter temperature (T2), 2-meter specific humidity (Q2), surface pressure (PSFC), 10-meter longitudinal wind velocity (U10) and 10-meter latitudinal wind velocity (V10). The latter two features are also combined into one positive wind speed (WS), with $WS = \sqrt{(U10)^2 + (V10)^2}$ to be used later. Feature normalization then ensures that these numerical features are scaled to a mean of 0 and a standard deviation of 1. This step is crucial for stabilizing the training process and enhancing the model's convergence speed. The distributions of the various normalized features can be seen in Figure 5.13. We here note the original means and standard deviations of the 10-min dataset:

Table 5.2: Means and Standard Deviations (Std) of Features Before Normalization

	T2 Pout [K]	T2 Dakar [K]	Q2 Pout [kg/kg]	Q2 Dakar [kg/kg]
Mean	300.6	300.5	0.017	0.018
Std	2.9	1.7	0.0021	0.0016
	PSFC Pout [Pa]	PSFC Dakar [Pa]	U10 Pout [m/s]	U10 Dakar [m/s]
Mean	100724	100734	2.0	2.1
Std	165	162	2.2	2.0
	V10 Pout [m/s]	V10 Dakar [m/s]	WS Pout [m/s]	WS Dakar [m/s]
Mean	-0.5	-0.7	3.5	3.6
Std	2.5	2.4	1.6	1.2

Data Windowing. We create input-output pairs, or windows, from the time series data, configured with the parameters *input width*, *label width*, and *shift* as visualized in Figure 5.14. We define these parameters so that we have an input window $\mathbf{x}$ of two hours (input width = 12 time steps) and label window $\mathbf{y}$ of one hour (label width = 6 time steps), which serves as the 'truth' for the prediction window $\hat{\mathbf{y}}$ of the same length. This complies with training the model to predict the advent

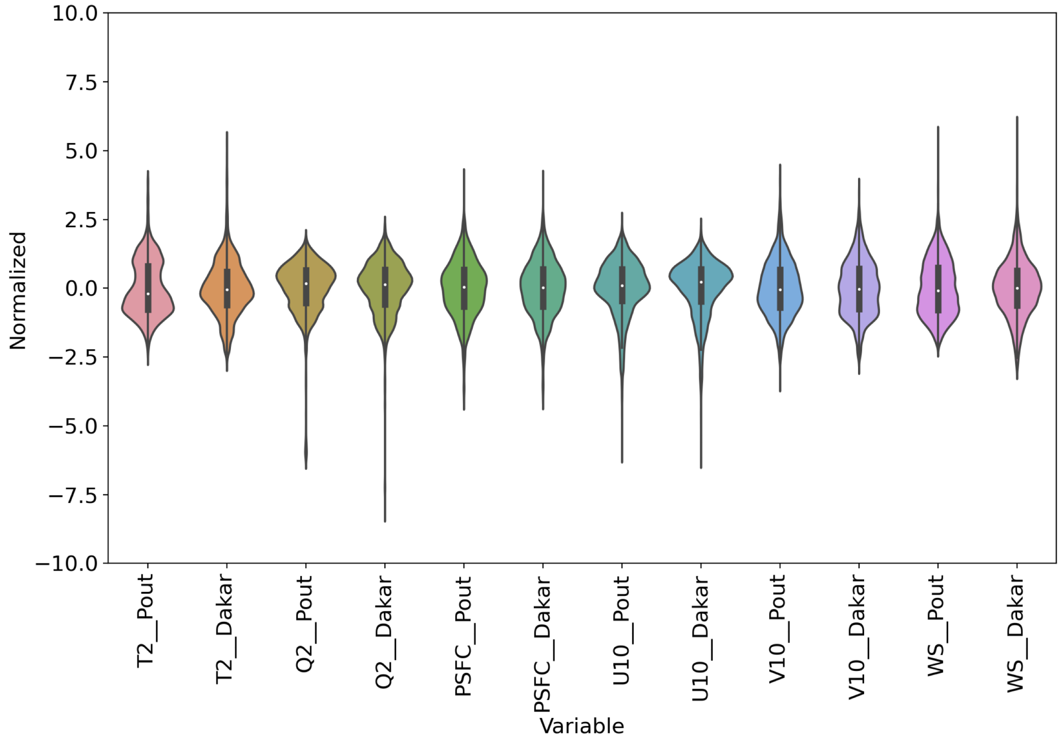


Figure 5.13: **Variable distributions.** Normalized variables from simulated station time series. These variables are then used as input features for the LSTM, to predict T2 Dakar.

and evolution of CPs in Dakar, given a two hour input of recent measurements. We insert a lag between input and the first prediction step (shift = 2 time steps), to demand a 30-minute predictive capability.

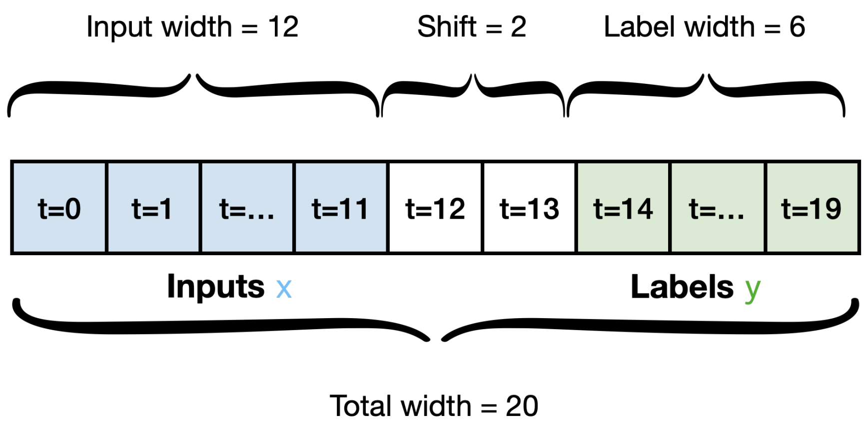


Figure 5.14: **Data Windowing.** Parameters for windowing as defined in this study: input width = 12 time steps, shift = 2 time steps and label width = 6 time steps. Note that one time step is $\Delta t = 10$ minutes.

The *inputs* used to train and evaluate the LSTM are the five features for each station (T2, Q2, PSFC, U10, V10), for a total of 10 features. The *label* is the 2-meter temperature in Dakar (T2 Dakar), as visualized in the schematic 5.15.

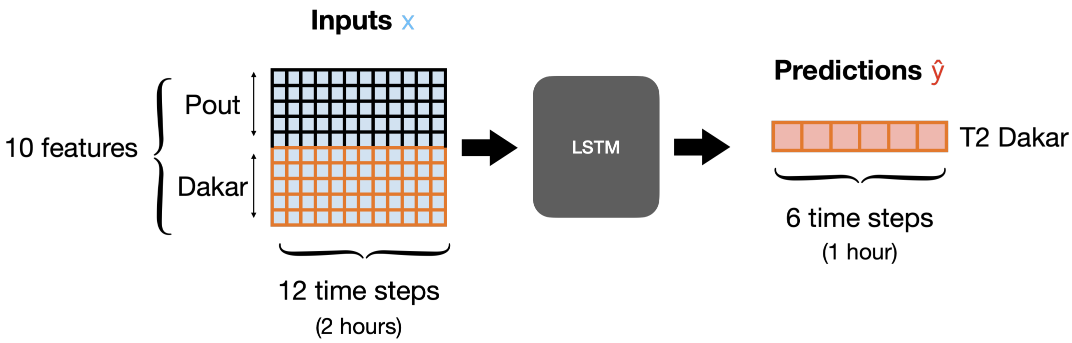


Figure 5.15: **Data Structure.** Schematic of input and output data for LSTM model.

To make sure that the dataset contains enough cases of CPs, we run the CP detection algorithm as before, on the native 1-min simulated time series of Pout and Dakar. We divide the dataset windows into three subsets, as schematized in Figure 5.16. The first subset comprises cold pools (CPs) that are detected first in Pout and then in Dakar (we call these "Propagating CPs"). The second subset consists of CPs detected in Pout but not in Dakar ("Not-Propagating CPs"). Lastly, the third subset comprises randomly sampled data points excluding the previous subsets ("Not CPs"). Since we only have 12 CP cases of propagating CPs, we randomly sample the windows related to these CPs, selecting different starting points between 10 steps before and after the temperature drop, to increase the size of the subset. We do the same for the not-propagating CPs. We randomly select an equal number of windows for each category, to ensure that each subset is of equal size.

Data Partitioning. We partition the collection of windows containing equal representation of the three CP categories, into distinct subsets, for training (70%), validation (20%), and testing (10%) purposes. The partitioning of windows consists of 980 windows, 280 windows, and 141 windows respectively. This partitioning is essential to assess the model's performance accurately and avoid overfitting, and ensures that the model learns from a diverse range of data while providing robust performance evaluations.

Model Compilation. With the data prepared and split, we can build and compile the model. Leveraging TensorFlow's Functional API, the architecture of the model

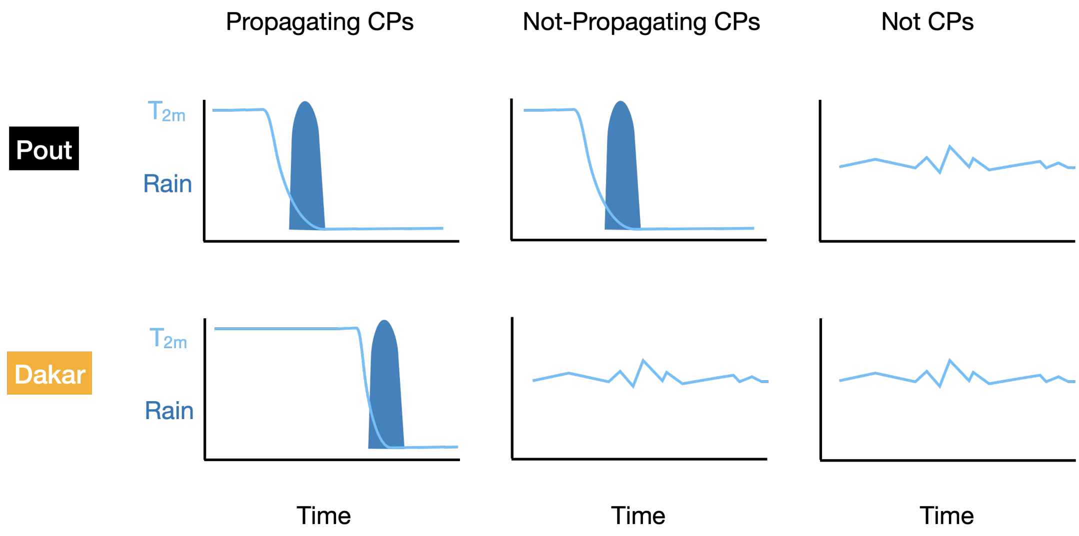


Figure 5.16: **CP classification.** Simplified schematic of classification between "Propagating CPs", "Not-Propagating CPs" and "Not CPs".

is defined. We here use two different LSTM architectures: "one shot" and "auto-regressive".

The one shot LSTM consists of an input layer, an LSTM cell layer with 100 hidden units for capturing temporal dependencies, and a dense layer which transforms the features learned by the LSTM layer into predictions for the target variable. It processes the input sequence and generates predictions for multiple future time steps by outputting a single vector representing the final prediction for the entire sequence. A schematic of the one-shot LSTM is seen in Fig. 5.17 a.

The auto-regressive LSTM similarly consists of an input layer, an LSTM cell layer with 100 hidden units and a dense layer for generating predictions, with the addition of a warmup method which initializes the model by processing the input data through the LSTM layer and generating an initial prediction, and a call method which extends the prediction over multiple time steps by iteratively feeding the previous prediction back into the model. Each iteration involves passing the prediction through the LSTM cell and generating an output prediction using the dense layer. Finally, the predictions are stacked and transposed to match the desired output shape, and the model returns the predicted values for each time step into the future. A schematic of the auto-regressive LSTM is seen in Fig. 5.17 b.

Irrespective of the selected architecture, we compile the LSTM models with a mean squared error loss function, Adam optimizer, and mean absolute error metric.

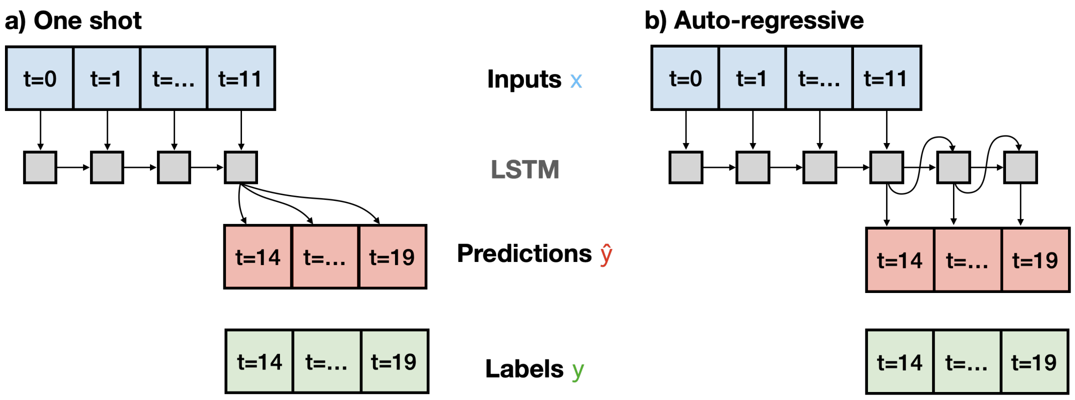


Figure 5.17: **LSTM structure.** Simplified schematic of (a) one shot, and (b) auto-regressive LSTM models, with input inputs and labels as defined in Figure 5.14.

The optimizer handles parameter updates during training, while the loss function computes the disparity between model predictions and true labels. We train the model in batches, with a batch size of 64. We fit the model to the training data after the appropriate number of epochs with validation data used for monitoring. We prevent overfitting by monitoring the validation loss and stopping training if the loss does not improve after a certain number of epochs, and halting the training when this has stopped decreasing for 15 consecutive epochs, indicating a potential convergence of the model. The model is thus trained on the training data using the fit method, iteratively adjusting its parameters to minimize the defined loss function. The model's performance is evaluated on the validation set, providing insights into its generalization ability and effectiveness in making predictions. We finally test its ability in making predictions with the test set, and check the capabilities of the model of accurately predicting windows from all three sets, "Propagating CPs", "Not-Propagating CPs" and "Not CPs" in Dakar. For prediction accuracy, we calculate the Mean Average Error (MAE), which measures the average absolute difference between the predicted and actual values.

Calculating importance of variables. To understand the learning process of the LSTM we calculate the "Relative Score" of each input feature after training, by applying the L1 normalization technique to the weights of the LSTM model. This process involves summing the absolute values of the weights associated with each input variable across all time steps. Subsequently, the weights for each variable are divided by the sum of these absolute values, resulting in normalized weights that reflect the relative importance of each input feature within the model. The resulting Relative Score represents the magnitude of influence that each feature exerts on the model's predictions, relative to other variables, without being constrained by specific units. A negative importance score in this context indicates that the corresponding feature has a negative impact on the prediction, i.e. an increase in the value of that feature is associated with a decrease in the predicted outcome.

LSTM evaluation

One shot LSTM vs auto-regressive LSTM. We first compare the one shot and the auto-regressive LSTM in terms of training. Both are able to improve with enough training epochs, reaching a training loss of 0.01 (Figure 5.18), and a validation loss of 0.03 after respectively 80 and 100 epochs. The Relative Importance Scores of the input variables are similar for both architectures: the 2-meter temperature (T2) in Dakar is highly influential on the predictions, along with the surface pressure (PSFC) in Pout, which gains even higher importance for the auto-regressive LSTM. The 10-meter longitudinal wind (U10) in Pout also wins in importance in both models, with anti-correlated influence that we might expect in the advent of CPs (increase in U10 in Pout indicates decrease in T2 in Dakar). The auto-regressive LSTM tends to use more information from all the features than the One-shot LSTM, which bases its predictions fully on the temperature in Dakar, the surface pressure in Pout, and the longitudinal winds of both locations. Strikingly, the 2-meter temperature in Pout (T2 Pout) has very little influence on the predictions, in both models. One could rather expect that the temperature drop in Pout, the characteristic property of cold pools, would have a large impact on the predictability of CPs in Dakar, but here other features seem to surpass in importance.

We evaluate our two LSTMs on the task they were designed to do: predicting CPs in Dakar. We want to check that the LSTMs have skill in predicting "Propagating CPs" measured first in Pout and then Dakar, but we also want to make sure that "Not-Propagating CPs" measured in Pout are not erroneously predicted as CPs in Dakar, and that the normal time windows without CPs are also predicted well. To do this, we calculate the Mean Average Error (MAE) between the predictions and the labels related to the three categories, in the test set (Figure 5.19). We compare the MAE of the one shot LSTM against the auto-regressive when asked to predict "Propagating CPs", "Not-Propagating CPs" and "Not CPs", and find that the one shot LSTM is best at predicting the temperature of "Not-Propagating CPs" in Dakar with an MAE of 0.1. The worst predictions being the ones by the

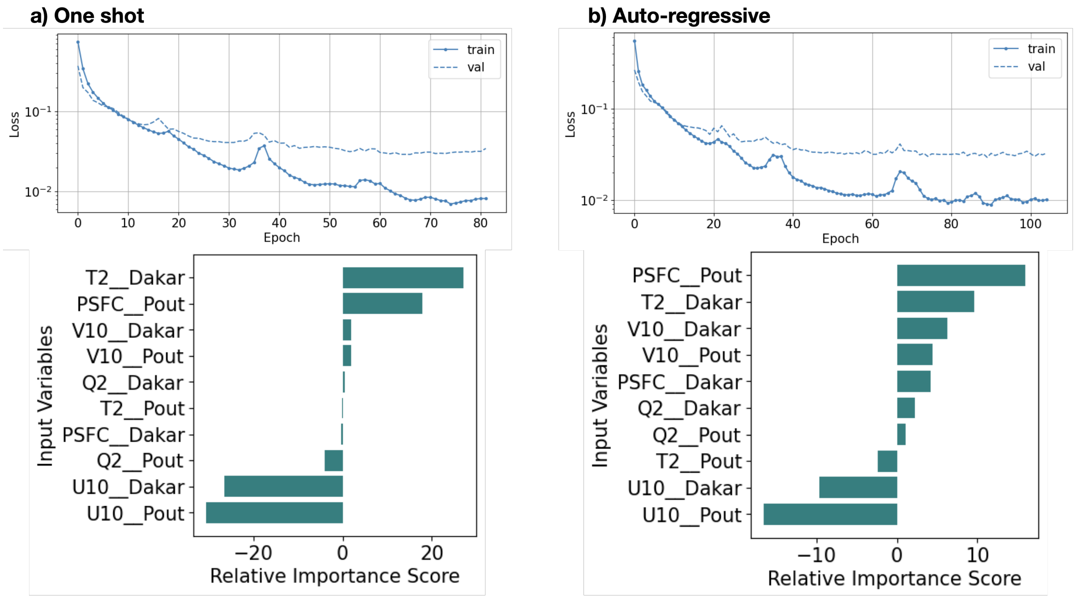


Figure 5.18: **Training the LSTMs.** Training and validation loss during training of a) one shot and b) auto-regressive LSTMs. Bottom row show Relative Importance Score for each Input Variable.

one shot LSTM on the temperature in Dakar of the "Propagating CPs", with an MAE of 0.25. To interpret MAE values in a physical sense, we need to consider the normalized scale of the output labels and then convert it back to the original scale. In the normalized scale (mean 0 and std 1), an MAE of x indicates that, on average, the predictions deviate from the actual values by x standard deviations. Now, to interpret this in the original scale: Given that the original label data had a mean of approximately 300K and a standard deviation of approximately 2K (see mean and standard deviation of "T2 Dakar" in Table 5.2), we can compute the equivalent deviation as $x \cdot 2K$. Thus, on average, the predictions deviate from the actual values by approximately $0.1 \cdot 2K = 0.2K$ in the best cases and $0.25 \cdot 2K = 0.5K$ in the worst, in the original temperature scale. The two algorithms are overall quite skilled in all three subcategories.

Predicting CPs in the simulated data.

We confirm the predictive skill of the LSTMs by visualizing how the auto-regressive algorithm predicts the time series relative to the three subcategories from the test set, in Figure 5.20 (the windows for the one shot LSTM are very similar). For this visualization, we un-normalize the inputs, labels and predictions to have relatable units. We show the inputs that the model sees, in blue. To make interpretation

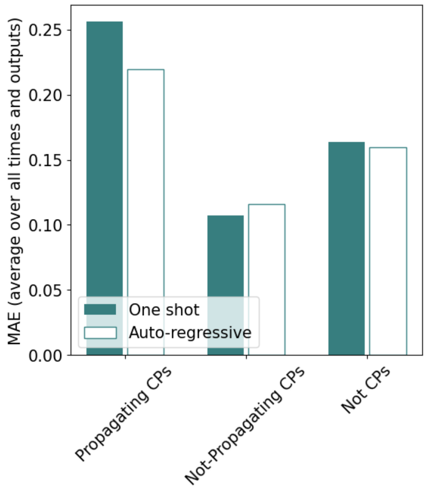


Figure 5.19: **Predictive skill of LSTMs on CPs.** Skill of LSTMs in predicting "Propagating CPs", "Not-Propagating CPs" and "Not CPs", evaluated as the Mean Average Error (MAE) of the one shot and auto-regressive LSTMs on all samples of each category contained in the test set.

easier, we include only two input features in the visualization: T2 Pout and T2 Dakar. (However, note that there are 8 additional input features!). We also show the original T_{2m} temperatures from the dataset, which the model does not see, in thin grey lines. The labels that the model is aiming to predict are shown with green dots and the actual predictions with red crosses. From Figure 5.20, it is clear that the LSTM is able to capture the time series from all three categories. Notably, it can predict the advent of a CP in Dakar (Figure 5.20 a), with a 30-minute warning. Interestingly, it predicts the CP temperature drop, even without knowing how much the temperature has dropped in Pout (the input is only up to time step 11 but we can see that full temperature drop occurs in time step 12, which the model has not seen yet). Furthermore, it is able to tell apart a "Propagating CP" from a "Not-Propagating CP" (Figure 5.20 b), where there is a clear temperature drop in Pout, and we deliberately show the instance where that first drop is included in the inputs. The fact that the prediction in Figure 5.20 b does not foresee a CP in Dakar, shows that the LSTM is not merely repeating the temperature pattern it sees in Pout, but finding patterns in all the variables that determine whether or not the CP measured in Pout will make it to Dakar. Finally, we note that the LSTM is also able to simply predict random snippets of temperature, that do not include CPs (Figure 5.20 c).

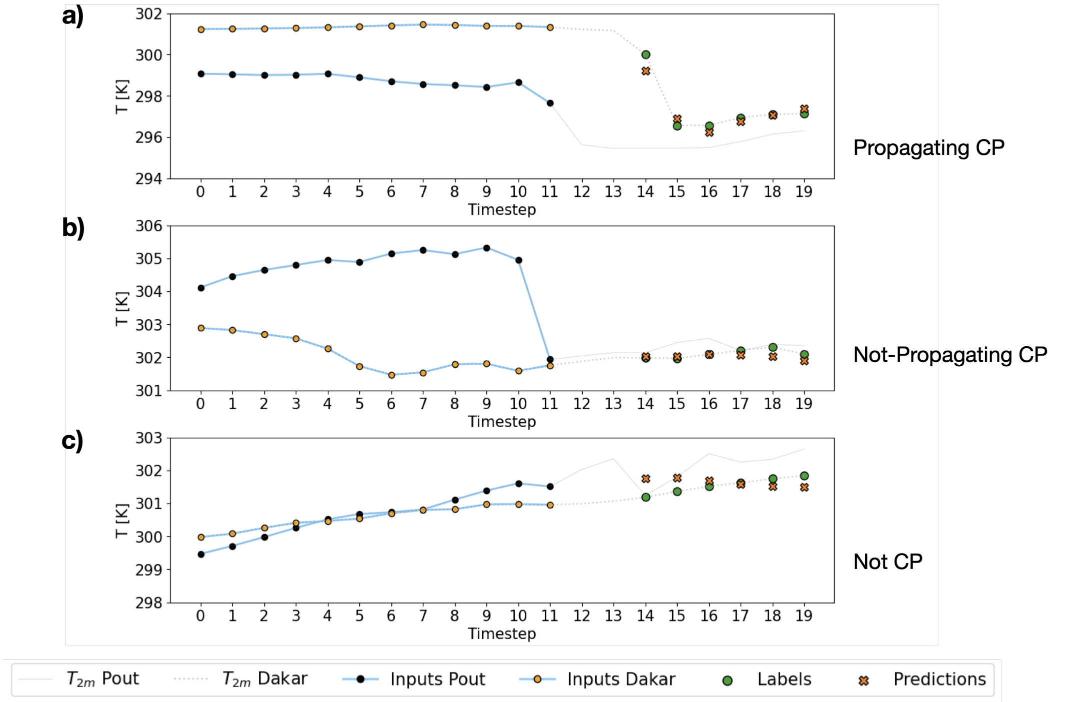


Figure 5.20: **Predictive skill of auto-regressive LSTM across different categories of simulated CPs.** a) "Propagating CP", b) "Not-Propagating CP" and c) "Not CP" from the *test* set, with inputs, labels and predictions of the auto-regressive LSTM. Note: The time steps here are $\Delta t = 10$ minute time intervals.

Towards predicting CPs in real weather station data.

That the LSTMs are good at predicting CPs in the simulation data, is a good step towards applying this method to real station data in Senegal. As we wait for our real stations to measure the upcoming rainy season to further test our LSTM, we do a simple test on our data from the end of the rainy season of 2023.

Since the two automatic weather stations measure wind differently, we simplify the feature space, by using wind speed (WS) instead of the latitudinal and longitudinal winds. We retrain our model on the simulated data, with only 4 input features per station (T2, Q2, PSFC and WS), for a total of 8 input features. The skill of both LSTMs on simulation data remains largely unaffected. However, interestingly, the relative importance scores of the variables change when U10 and V10 are removed from the inputs (Figure 5.21), with WS Pout joining the important features, and T2 Pout suddenly becoming an important feature, with a negative correlation, taking the previous spot of U10 Pout. While before, the LSTMs had directional information of the wind (in the form of U10 and V10), now there is no information about the direction, but only wind speed. The wind direction was clearly important for both LSTMs in determining the temperature (and temperature drops) in Dakar - possibly a signal of sea breeze and land breeze, which would be influential on the propagation of a CP. When removing this information, the temperature in Pout seems to be redeemed, potentially carrying similar information relative to sea breezes, that the wind direction previously contributed.

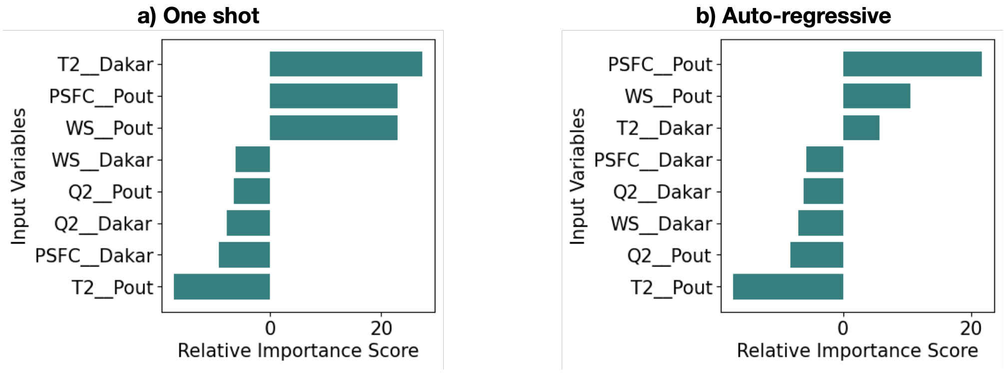


Figure 5.21: Relative Importance Score for each Input Variable, for a) one shot LSTM and b) auto-regressive LSTM, with 8 input features instead of 10.

We process the station data to match the structure of the data that the LSTMs need. We then take the two LSTMs, trained on simulation data with 8 input features, and test them on the real station data, using consecutive windows before and after the temperature drop of a "Real Propagating CP". The prediction of the auto-regressive LSTM shown in Figure 5.22, with three sequential instances visualized. About 30 minutes before the CP has been seen in either station (i.e., before the temperature drop is included in the input features), the LSTM is correctly predicting the 2-m temperature in Dakar (Figure 5.22 a). When the CP has been seen in Pout, the LSTM predicts a temperature drop in Dakar (2K), albeit with a third of the magnitude of the real temperature drop (6K). After the CP has been measured in both Pout and Dakar, the predictions are closer to the labels but still too warm by about 2K. This offset could be due to a large number of factors and a more thorough evaluation would need more real CP cases. It is important to remember that the median temperature drop of the simulated CPs that the LSTM has been trained on is -3.3 K, while this measured CP has a temperature drop of 6K both in Pout and in Dakar, so it is probably an extreme on the distribution of CPs that the LSTM has "seen before".

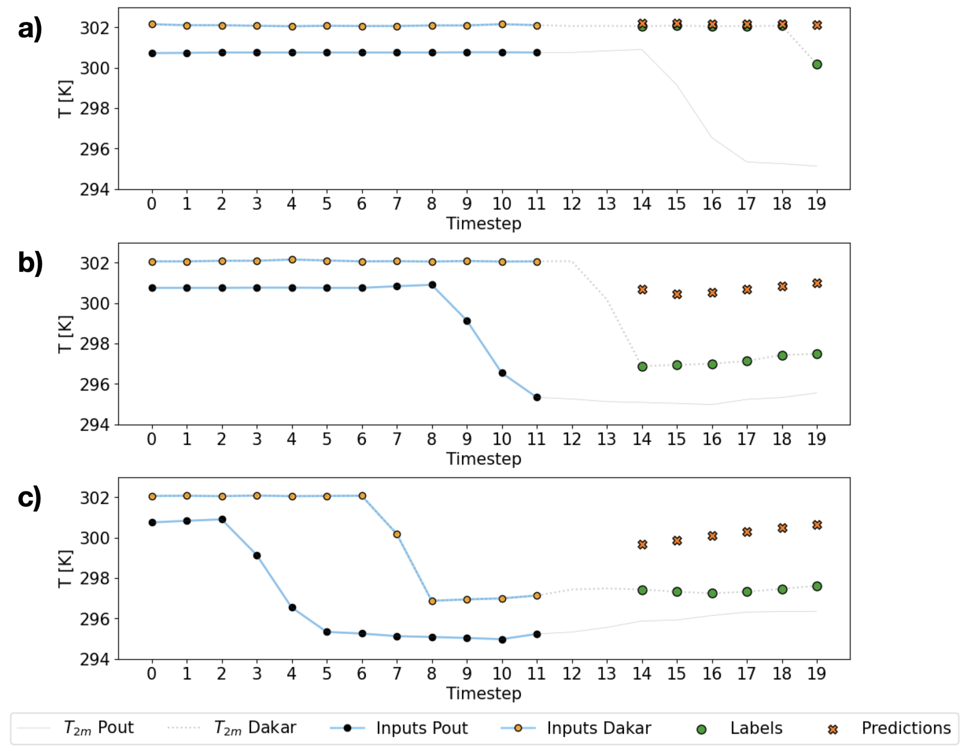


Figure 5.22: **Predictive skill of auto-regressive LSTM on real CP.** Windowing samples shown are a) before the CP has been measured in either station, b) when the CP has been measured in Pout, and b) after the CP has been measured in Pout and Dakar. The time steps here are $\Delta t = 10$ minutes.

5.3 Conclusions

This study represents a significant step toward the development of a nowcasting tool for predicting the onset of cold pools (CPs) in Dakar, Senegal, leveraging data from two automatic weather stations located approximately 50 kilometers apart, and delivering real-time measurements at 1-minute intervals.

Measuring CPs with automatic weather stations. Through careful analysis of CPs measured by our stations, both in Germany and Senegal, we identify a consistent sequence of events preceding the onset of rain, indicative of CPs. This sequence typically includes a sudden increase in windspeed, followed by a temperature drop, a decrease in moisture, and an increase in pressure. Notably, our observations in Senegal suggest that these events tend to occur first in the town of Pout and then (albeit not always) propagate to Dakar, with a lag of approximately 30 minutes.

Training LSTMs to predict CPs from simulated time series. In response to the challenge of predicting CPs in Dakar, we turn to machine learning methodologies, specifically leveraging the power of long-short term memory networks (LSTMs). LSTMs, being a type of recurrent neural network capable of learning temporal dependencies and patterns, offer a promising approach for forecasting CPs. To compensate our lack of historical measurements to train the LSTMs on, we simulate an entire rainy season over Senegal using the Weather Research and Forecasting (WRF) model at a resolution of 1km. From these simulations, we extract simulated "weather station" time series data for both Pout and Dakar. Subsequently, we train LSTM models on this simulated station data to predict the 2-meter temperature in Dakar. Encouragingly, our findings indicate that these LSTM models are capable of accurately capturing the onset of CPs in Dakar, providing a valuable 30-minute lead time for forecasters and disaster management authorities. The visualization of predictive outputs from our two LSTM models confirms their efficacy in capturing complex temporal patterns within simulated convective systems. The LSTM successfully anticipates the onset of convective phenomena in Dakar with a 30-minute lead time, showcasing its predictive skill on simulated data. Additionally, the model can discern between propagating and non-propagating convective events, illustrating its capacity to extract meaningful patterns from multivariate inputs. Moreover, our analysis reveals that the LSTM can accurately predict random temperature fluctuations, demonstrating its versatility across various scenarios.

Using LSTMs trained on simulated time series to predict CPs from real time series. In addition to our current findings, the LSTM models trained on simulated time series exhibit promising capabilities in predicting CPs before their occurrence in Dakar, based solely on real measured inputs from automatic weather stations. This provides compelling evidence for the efficacy of LSTMs trained on simulation data for nowcasting CPs, highlighting their potential for continued development in nowcasting applications.

In summary, our study underscores the potential of LSTM models in forecasting convective phenomena, bridging the gap between simulation and real-world applications. By leveraging insights from simulated data, we pave the way for robust applications of LSTM models in operational forecasting, providing valuable insights for weather monitoring and prediction.

Perspectives. To further enhance the algorithm's performance, several avenues for development are worth exploring. Firstly, expanding the dataset by simulating additional rainy seasons would augment the pool of training data, enabling the LSTM models to learn from a broader range of convective events. Moreover, the temporal resolution of the input and output data could be enhanced for a higher temporal accuracy in predicting CP onset, given that we have 1-minute resolution in both the simulations and the measurements. This would also require a larger training dataset, considering that, to accommodate 1-hour long predictions, we would need 60 output steps at $\Delta t = 1$ minute. Precipitation intensity could also be included as a feature to train on and predict. Finally, the integration of data from additional weather stations, such as those being established as part of the DakE network, presents an opportunity to enrich the input features and enhance the predictive skill of the models.

Looking ahead, the imminent availability of a full rainy season's worth of measured data in late 2024 offers a significant opportunity for further refinement of the LSTM models. By training, validating, and testing the models solely on measured data from Pout, Dakar, or the entire DakE network, we can tailor the algorithms to accurately predict real convective events, thereby advancing the capabilities of nowcasting techniques in weather monitoring and prediction.

6 Extrapolation

6.1 The New

In this dissertation, we have explored the dynamics of mesoscale convective systems (MCSs), and convective cold pools (CPs). Three distinct studies contribute to our understanding of these convective phenomena and the development of predictive tools:

Understanding the Impact of MCSs on Convective Self-Aggregation.
Our simulations of tropical land-like MCSs shed light on the profound influence of diurnal cycles in surface temperature on precipitation distribution. The presence of a diurnal cycle induces persistent moisture patterns, leading to multi-day precipitation distributions characterized by the formation of large MCS-like deep convective structures during afternoon hours. Notably, land-like simulations exhibit a self-aggregated state, which can then persist in less favourable conditions, highlighting the significance of the diurnal cycle in shaping convective behavior. We have developed a novel cellular automaton-based conceptual model, the "Game of Cloud", which encompasses the key processes related to MCSs, with a focus on the role of CPs. The conceptual model is able to replicate the self-aggregation of the atmosphere over land, and the persistence of this aggregated state over the ocean. Linking the abstract to reality, our findings could have implications for understanding hurricane formation, which could be regarded in this context as self-aggregation starting over tropical African land and intensifying over the Atlantic Ocean.

Detecting and Analysing Convective CPs from Weather Station Time Series.
We have developed and validated a methodology for detecting convective CPs over flat mid-latitude coastal land from weather station time series. Through statistical analysis of 189 composited CPs from 10 years of data from the Netherlands, we

have characterized CP properties over land, including strong updrafts, horizontal wind anomalies, a dry interior without moisture rings, and 1-2h recovery times of thermodynamic anomalies post-CP passage. Our findings furthermore lay the groundwork for creating a simple model for CP strength based on precipitation rate and atmospheric saturation.

Developing an AI-Based Tool for Nowcasting CPs in Senegal.
Leveraging machine learning methodologies, specifically LSTM networks trained on simulated data, we have developed a predictive tool capable of accurately anticipating the onset of CPs in Dakar, Senegal. The LSTM models demonstrate the ability to capture complex temporal patterns and provide a valuable 30-minute lead time for nowcasting. The nowcasting tool has been developed to be used with measurements from automatic weather stations recently set up in Senegal as a part of an emerging field campaign, DakE. Our findings underscore the potential of LSTM models in operational forecasting, bridging the gap between simulation and real-world applications.

In summary, our dissertation contributes to a deeper understanding of convective systems and their dynamics, while also paving the way for improved forecasting and nowcasting techniques. The insights gained from these studies have implications for weather monitoring and prediction in various regions around the world, offering valuable insights for advancing the field of meteorology. Future research endeavors may build upon these findings to further enhance our ability to forecast convective phenomena and mitigate the impacts of extreme weather events.

6.2 The Future

A Benchmark for CP Studies.

Our study of CPs over the Netherlands (Chapter 4, i.e., Kruse *et al.* (2022)) marked one of the first comprehensive statistical analyses of CPs over land, alongside the work of Kirsch *et al.* (2021), thereby supplementing the already existing studies focused on oceanic CPs. These two studies have thus established a benchmark for comparing other observed composites of land-based CPs. For example, Mai *et al.* (2023) recently used data from a 356-meter weather tower in Southern China to analyze a CP case study, with comparisons to our CPs in the Netherlands, notably finding a similarly dry interior at all tower levels, as opposed to the moister CPs in Kirsch *et al.* (2021). Hoeller, Haerter, *et al.* (2024), leveraging data from ground-based automatic weather stations located across several countries in Equatorial Africa (Cameroon, Democratic Republic of Congo, Nigeria and Uganda) adapted our CP detection algorithm for temporally lower-resolution data. Their analysis has revealed more robust and drier CPs than those observed in the Netherlands, reaffirming our finding of the absence of CP moisture rings over land. Very interesting data has emerged from Kirsch *et al.* (2024) thanks to the dense network of weather stations of the FESSTVAL field campaign in Germany. Their findings have challenged the traditional density current theory that we employed in Kruse *et al.* (2022). The spatially dense measurements, as opposed to point-data, have provided a clearer picture of the observed temperature difference between the inside and outside of a CP, and with the propagating velocity of the gust front, they show that the precipitation-driven volume increase of a CP describes the growth process better than its mean density excess.

Towards a Unified Parameterization Scheme for Deep Convection.

Models operating at lower resolutions, such as those utilized in climate projections spanning extended periods or in global models serving as boundary conditions for weather forecasts, encounter a dilemma, facing the choice of either awaiting advancements in computational technology to support high-resolution global models operating over centuries or developing suitable parameterizations for MCSs and CPs. Current convective parameterizations, which rely on distinguishing between convective and synoptic scales, prove inadequate due to the spatial dimensions of MCSs and CPs. The effective integration of MCSs and CPs into weather and climate models, whether through cloud-resolving modeling or parameterization, is therefore imperative. A conceptual model such as the "Game of Cloud" presented in Chapter 3, could be used as a parameterization for the formation of MCSs over land, advected over the ocean, for simulations coarser than the single isolated convective cell or CP. Since we only considered the organizational effect of CPs in this model, thus after the onset of rain (green shading in Figure 6.1), a unified parameterization could also include the upscale growth of thermals, that create the first deep convective cells to begin with, which has been interestingly theorized in our recent collaborative work, Vraciu *et al.* (2023).

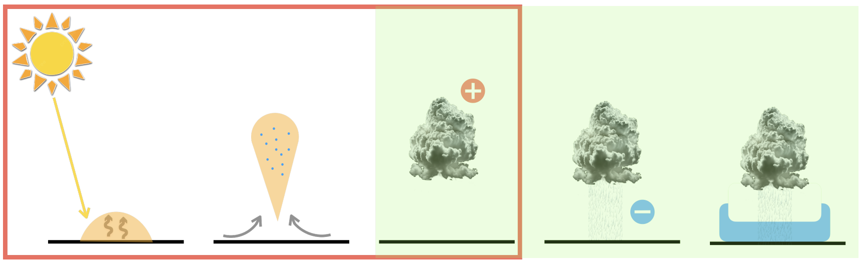


Figure 6.1: **Parameterizing the life cycle of a convective storm.** Revisiting Figure 1.1: In this dissertation, we have presented a conceptual model which describes the stages of the life cycle of convection involving precipitation and CP formation (green shaded box). Further work could include adding the first stages involving how thermals lead to deep convection (red box).

Encouraging the Link to Tropical Cyclogenesis.

While we have touched upon the relationship between the upscale growth of convective cells and tropical cyclones, a comprehensive exploration of tropical cyclogenesis was not within the scope of this thesis. The realm of tropical cyclones constitutes a distinct field of research, yet overlaps with our investigation arise when applying the Coriolis force to Radiative Convective Equilibrium (RCE) simulations, leading to the emergence of cyclones (Carstens and Wing, 2020; Carstens and Wing, 2022). Additionally, notable observational studies have delved into tropical cyclones originating off the coast of West Africa, particularly in connection to African Easterly Waves (Ocasio *et al.*, 2020). Future studies could merge these areas of research, potentially leveraging machine learning algorithms to discern complex patterns within observational and simulated data. This could lead to a deeper examination of the signatures of Mesoscale Convective Systems (MCSs) developed over African land potentially akin to convective self-aggregation, untangle the factors propelling them towards tropical cyclogenesis over the Atlantic Ocean, and ultimately enhance the ability to forecast their origins and landfall with a lead time of weeks.

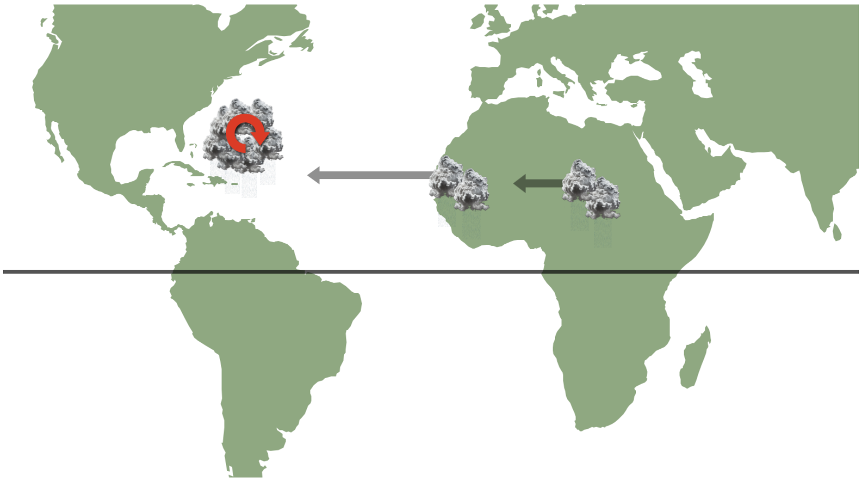


Figure 6.2: **The link to tropical cyclones?** We have found how MCSs produced over tropical Africa, impact the atmospheric state advected westward, which can lead to convective self-aggregation persisting over the ocean in simulations - indicating a possible path to developing tropical cyclones.

A Promising Tool for Nowcasting Deep Convection.
Our nowcasting tool for CPs in Dakar, presented, in Chapter 5, serves as an encouraging example of the significant advancements possible in convective weather forecasting. It demonstrates how with affordable, easy-to-install automatic weather stations, coupled with a weather simulation and a few AI techniques, it is possible to produce accurate predictions of convective events on the short-term time scale. This is already promising with just with two weather stations, and the ongoing addition of more stations to the DakE network in Senegal and the expected measurements of the 2024 rainy season will be an exciting test-bed for our nowcasting tool. While I am writing this (March 2024), 10 more automatic weather stations have been established, and inserted into an online, open-source portal maintained by the UK Centre for Ecology and Hydrology (Figure 6.3). It is worth noting, that while our CP nowcasting tool was developed in the context of a field campaign in Senegal, the concept could be tailored to any location on the globe experiencing deep convective events.

As a final note - MCSs and CPs persist as significant societal challenges worldwide, often implicated in severe weather events including flooding and strong wind gusts. With the Earth experiencing warming, the distribution patterns and frequency of MCS occurrences are expected to shift, increasing the frequency of the extreme events. It thus becomes increasingly important to fundamentally understand convective storms, and to be able to predict them ahead of time - as the fundamental review paper "100 Years of Research on Mesoscale Convective Systems" concluded

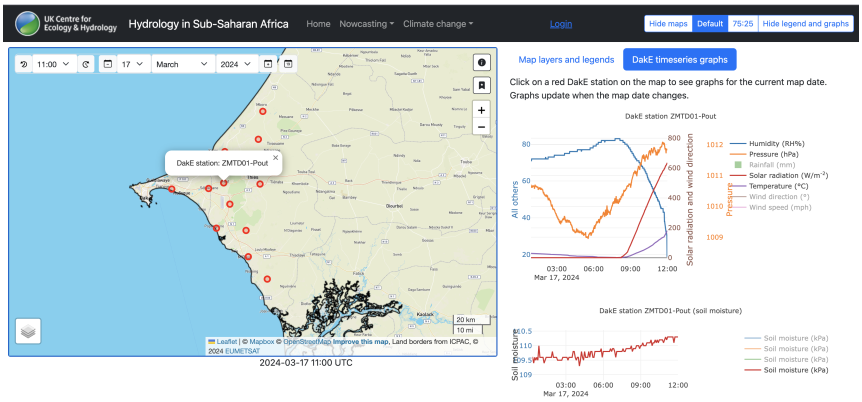


Figure 6.3: **Open-source forecasting portal by UKCEH.** This portal has been designed by the UK Centre for Ecology and Hydrology to help forecasters in Sub-Saharan Africa to predict severe convective storms (UKCEH Nowcasting Portal, Accessed: [March 2024]). It now includes the full DakE network with 1-min real-time updates (the station in Pout we use in chapter 5 is visualized). The portal could be the optimal location for the incorporation of our new CP nowcasting algorithm.

in 2018: "Forecasting MCSs both in real time and projecting their future occurrence in a changing climate remains a grand challenge for meteorology and climate. (Houze, 2018)" We have contributed to this grand challenge with this dissertation, and as we extrapolate our findings, the research continues.

I've looked at clouds from both sides now
From up and down and still somehow
It's cloud illusions I recall
I really don't know clouds at all.
– Joni Mitchell (Both Sides Now, 1967)

7 Bibliography

Anderson, Seonaid R., Steven J. Cole, Cornelia Klein, Christopher M. Taylor, Cheikh Abdoulahat Diop, and Mouhamadou Kamara (2023). „Nowcasting convective activity for the Sahel: A simple probabilistic approach using real-time and historical satellite data on cloud-top temperature“. en. In: *Quarterly Journal of the Royal Meteorological Society* n/a.n/a.

Benjamin, T. Brooke (1968). „Gravity currents and related phenomena“. In: *Journal of Fluid Mechanics* 31.2, pp. 109–248.

Biagioli, Giovanni and Adrian Mark Tompkins (2023). „A Dimensionless Parameter for Predicting Convective Self-Aggregation Onset in a Stochastic Reaction-Diffusion Model of Tropical Radiative-Convective Equilibrium“. en. In: *Journal of Advances in Modeling Earth Systems* 15.5, e2022MS003231.

Böing, Steven J (2016). „An object-based model for convective cold pool dynamics“. In: *Mathematics of Climate and Weather Forecasting* 2.1.

Böing, Steven J, Harm JJ Jonker, A Pier Siebesma, and Wojciech W Grabowski (2012). „Influence of the subcloud layer on the development of a deep convective ensemble“. In: *Journal of the Atmospheric Sciences* 69.9, pp. 2682–2698.

Bony, Sandrine, Bjorn Stevens, David Coppin, Tobias Becker, Kevin A. Reed, Aiko Voigt, and Brian Medeiros (2016). „Thermodynamic control of anvil cloud amount“. In: *Proceedings of the National Academy of Sciences* 113.32, pp. 8927–8932.

Bosveld, Fred C., Peter Baas, Anton C. M. Beljaars, Albert A. M. Holtslag, Jordi Vilà-Guerau de Arellano, and Bas J. H. van de Wiel (2020). „Fifty Years of Atmospheric Boundary-Layer Research at Cabauw Serving Weather, Air Quality and Climate“. In: *Boundary-Layer Meteorology* 177, pp. 583–612.

Bretherton, Christopher S., Peter N. Blossey, and Marat Khairoutdinov (2005). „An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST“. EN. In: *Journal of the Atmospheric Sciences* 62.12, pp. 4273–4292.

Byers, Horace Robert, Roscoe R. Braham, and United States Weather Bureau (1949). *The Thunderstorm: Report of the Thunderstorm Project*. en. Google-Books-ID: D8sJAQAAIAAJ. U.S. Government Printing Office.

Cafaro, C. and G. Rooney (2018). „Characteristics of colliding density currents: A numerical and theoretical study". In: *Quarterly Journal of the Royal Meteorological Society* 144.

Carstens, Jacob D. and Allison A. Wing (2020). „Tropical Cyclogenesis From Self-Aggregated Convection in Numerical Simulations of Rotating Radiative-Convective Equilibrium". en. In: *Journal of Advances in Modeling Earth Systems* 12.5, e2019MS002020.

Carstens, Jacob D. and Allison A. Wing (2022). „A Spectrum of Convective Self-Aggregation Based on Background Rotation". en. In: *Journal of Advances in Modeling Earth Systems* 14.5, e2021MS002860.

Cesar (Cabauw experimental site for atmospheric research) Database (2020). `https://ruisdael-observatory.nl/cesar-database/pages/datasetsKDC.html`. Accessed: 2020-08-01.

Chandra, Arunchandra S, Paquita Zuidema, Steven Krueger, Adam Kochanski, Simon P de Szoeke, and Jianhao Zhang (2018). „Moisture distributions in tropical cold pools from equatorial Indian Ocean observations and cloud-resolving simulations". In: *Journal of Geophysical Research: Atmospheres* 123.20, pp. 11–445.

Charba, Jess (1974). „Application of Gravity Current Model to Analysis of Squall-Line Gust Front". In: *Monthly Weather Review* 102.2, pp. 140–156.

Coppin, David and Sandrine Bony (2015). „Physical mechanisms controlling the initiation of convective self-aggregation in a General Circulation Model". en. In: *Journal of Advances in Modeling Earth Systems* 7.4, pp. 2060–2078.

Craig, G. C. and J. M. Mack (2013). „A coarsening model for self-organization of tropical convection". en. In: *Journal of Geophysical Research: Atmospheres* 118.16, pp. 8761–8769.

Dancker, Jonte (2022). *A Brief Introduction to Recurrent Neural Networks*. en.

Di Baldassarre, Giuliano, Alberto Montanari, Harry Lins, Demetris Koutsoyiannis, Luigia Brandimarte, and Günter Blöschl (2010). „Flood fatalities in Africa: From diagnosis to mitigation". en. In: *Geophysical Research Letters* 37.22.

Drager, A. J. and S. C. van den Heever (2017). „Characterizing convective cold pools." In: *J. Adv. Model. Earth Syst.* 9, pp. 1091–1115.

Drager, Aryeh J, Leah D Grant, and Susan C van den Heever (2020). „Cold pool responses to changes in soil moisture". In: *Journal of Advances in Modeling Earth Systems* 12.8.

Droegemeier, K. and R. Wilhelmson (1987). „Numerical Simulation of Thunderstorm Outflow Dynamics. Part I: Outflow Sensitivity Experiments and Turbulence Dynamics“. In: *Journal of the Atmospheric Sciences* 44.8, pp. 1180–1210.

Droegemeier, Kelvin K. and Robert B. Wilhelmson (1985a). „Three-Dimensional Numerical Modeling of Convection Produced by Interacting Thunderstorm Outflows. Part I: Control Simulation and Low-Level Moisture Variations“. In: *Journal of the Atmospheric Sciences* 42.22, pp. 2381–2403.

Droegemeier, Kelvin K. and Robert B. Wilhelmson (1985b). „Three-Dimensional Numerical Modeling of Convection Produced by Interacting Thunderstorm Outflows. Part II: Variations in Vertical Wind Shear“. In: *Journal of the Atmospheric Sciences* 42.22, pp. 2404–2414.

Emanuel, Kerry, Allison A. Wing, and Emmanuel M. Vincent (2014). „Radiative-convective instability“. en. In: *Journal of Advances in Modeling Earth Systems* 6.1, pp. 75–90.

Encyclopædia Britannica (Accessed: 2024-02-28[a]). `https://www.britannica.com/place/Sahel`. Accessed: 2024-02-28.

Encyclopædia Britannica (Accessed: 2024-02-28[b]). `https://www.britannica.com/science/West-African-monsoon`. Accessed: 2024-02-28.

Engerer, Nicholas A., David J. Stensrud, and Michael C. Coniglio (2008). „Surface Characteristics of Observed Cold Pools“. In: *Monthly Weather Review* 136.12, pp. 4839–4849.

Feng, Z., S. Hagos, A. K. Rowe, C. D. Burleyson, M. N. Martini, and S. P. de Szoeke (2015). „Mechanisms of convective cloud organization by cold pools over tropical warm ocean during the AMIE/DYNAMO field campaign“. In: *Journal of Advances in Modeling Earth Systems* 7.2, pp. 357–381.

Feng, Zhe, L. Ruby Leung, Nana Liu, Jingyu Wang, Robert A. Houze, Jianfeng Li, Joseph C. Hardin, Dandan Chen, and Jianping Guo (2021). „A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking“. en. In: *Journal of Geophysical Research: Atmospheres* 126.8, e2020JD034202.

Fildier, B., C. Muller, R. Pincus, and S. Fueglistaler (2023). „How Moisture Shapes Low-Level Radiative Cooling in Subsidence Regimes“. en. In: *AGU Advances* 4.3, e2023AV000880.

Fiolleau, Thomas and Rémy Roca (2013). „An Algorithm for the Detection and Tracking of Tropical Mesoscale Convective Systems Using Infrared Images From Geostationary Satellite“. In: *IEEE Transactions on Geoscience and Remote Sensing* 51.7, pp. 4302–4315.

Fletcher, J. K., C. A. Diop, E. Adefisan, *et al.* (2023). „Tropical Africa's First Testbed for High-Impact Weather Forecasting and Nowcasting“. EN. In: *Bulletin of the American Meteorological Society* 104.8, E1409–E1425.

Fournier, M. and J. O. Haerter (2019). „Tracking the Gust Fronts of Convective Cold Pools“. In: *Journal of Geophysical Research: Atmospheres* 124.

Fujita, Tetsuya (1955). „Results of Detailed Synoptic Studies of Squall Lines“. en-US. In: 7.44, p. 405.

Garg, Piyush, Stephen W. Nesbitt, Timothy J. Lang, and George Priftis (2021). „Diurnal Cycle of Tropical Oceanic Mesoscale Cold Pools“. EN. In: *Journal of Climate* 34.23, pp. 9305–9326.

Goff, R. C. (1976). „Vertical Structure of Thunderstorm Outflows“. In: *Monthly Weather Review* 104.11, pp. 1429–1440.

Haerter, Jan O (2019). „Convective Self-Aggregation As a Cold Pool-Driven Critical Phenomenon“. In: *Geophysical Research Letters* 46.7, pp. 4017–4028.

Haerter, Jan O, Steven J Böing, Olga Henneberg, and Silas Boye Nissen (2019). „Circling in on convective organization“. In: *Geophysical Research Letters* 46.12, pp. 7024–7034.

Haerter, Jan O and Linda Schlemmer (2018). „Intensified cold pool dynamics under stronger surface heating“. In: *Geophysical Research Letters* 45.12, pp. 6299–6310.

Haerter, Jan O., Bettina Meyer, and Silas Boye Nissen (2020). „Diurnal self-aggregation“. en. In: *npj Climate and Atmospheric Science* 3.1, pp. 1–11.

Haerter, Jan O. and Caroline Muller (2023). „Mechanisms for the Self-Organization of Tropical Deep Convection“. In: *Clouds and Their Climatic Impacts*. Geophysical Monograph Series, pp. 179–193.

Hamilton, R. A., J. W. Archbold, and C. K. M. Douglas (1945). „Meteorology of Nigeria and adjacent territory“. en. In: *Quarterly Journal of the Royal Meteorological Society* 71.309–310, pp. 231–264.

Hartmann, Dennis L., Harry H. Hendon, and Robert A. Houze (1984). „Some Implications of the Mesoscale Circulations in Tropical Cloud Clusters for Large-Scale Dynamics and Climate“. EN. In: *Journal of the Atmospheric Sciences* 41.1, pp. 113–121.

Heever, Susan C. van den, Leah D. Grant, Sean W. Freeman, *et al.* (2021). „The Colorado State University Convective CLoud Outflows and UpDrafts Experiment (C3LOUD-Ex)“. In: *Bulletin of the American Meteorological Society* 102.7, E1283–E1305.

Held, Isaac M., Richard S. Hemler, and V. Ramaswamy (1993). „Radiative-Convective Equilibrium with Explicit Two-Dimensional Moist Convection“. EN. In: *Journal of the Atmospheric Sciences* 50.23, pp. 3909–3927.

Hennon, Christopher C., Philippe P. Papin, Christopher M. Zarzar, *et al.* (2013). „Tropical Cloud Cluster Climatology, Variability, and Genesis Productivity“. EN. In: *Journal of Climate* 26.10, pp. 3046–3066.
Hitchcock, Stacey M., Russ S. Schumacher, Gregory R. Herman, Michael C. Coniglio, Matthew D. Parker, and Conrad L. Ziegler (2019). „Evolution of Pre- and Post-convective Environmental Profiles from Mesoscale Convective Systems during PECAN“. In: *Monthly Weather Review* 147.7, pp. 2329–2354.
Hochreiter, Sepp and Jürgen Schmidhuber (1997). „Long Short-Term Memory“. In: *Neural Computation* 9.8, pp. 1735–1780.
Hoeller, Jannik, Romain Fiévet, Edward Engelbrecht, and Jan O. Haerter (2024). „U-Net Segmentation for the Detection of Convective Cold Pools From Cloud and Rainfall Fields“. en. In: *Journal of Geophysical Research: Atmospheres* 129.1, e2023JD040126.
Hoeller, Jannik, Jan O. Haerter, and Nicolas A. Da Silva (2024). „Characteristics of Station-Derived Convective Cold Pools Over Equatorial Africa“. en. In: *Geophysical Research Letters* 51.6, e2023GL107308.
Hohenegger, Cathy, Felix Ament, Frank Beyrich, *et al.* (2023). „FESSTVaL: The Field Experiment on Submesoscale Spatio-Temporal Variability in Lindenberg“. EN. In: *Bulletin of the American Meteorological Society* 104.10, E1875–E1892.
Hohenegger, Cathy and Christoph Schär (2007). „Predictability and Error Growth Dynamics in Cloud-Resolving Models“. EN. In: *Journal of the Atmospheric Sciences* 64.12, pp. 4467–4478.
Hohenegger, Cathy and Bjorn Stevens (2016). „Coupled radiative convective equilibrium simulations with explicit and parameterized convection“. en. In: *Journal of Advances in Modeling Earth Systems* 8.3, pp. 1468–1482.
Holloway, Christopher E, Allison A Wing, Sandrine Bony, Caroline Muller, Hirohiko Masunaga, Tristan S L'Ecuyer, David D Turner, and Paquita Zuidema (2017). „Observing convective aggregation“. In: *Surveys in Geophysics* 38, pp. 1199–1236.
Houze, Robert A (1982). „Cloud Clusters and Large-Scale Vertical Motions in the Tropics“. In: *Journal of the Meteorological Society of Japan. Ser. II* 60.1, pp. 396–410.
Houze, Robert A. (1993). *Cloud Dynamics.* en. Academic Press.
Houze, Robert A. (2018). „100 Years of Research on Mesoscale Convective Systems“. EN. In: *Meteorological Monographs* 59.1, pp. 17.1–17.54.
Houze, Robert A., Bradley F. Smull, and Peter Dodge (1990). „Mesoscale Organization of Springtime Rainstorms in Oklahoma“. EN. In: *Monthly Weather Review* 118.3, pp. 613–654.

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, and Jackson Tan (2019). *GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06*. Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC).

Jeevanjee, Nadir and Stephan Fueglistaler (2020). „Simple Spectral Models for Atmospheric Radiative Cooling". EN. In: *Journal of the Atmospheric Sciences* 77.2, pp. 479–497.

Jeevanjee, Nadir and David M. Romps (2013). „Convective self-aggregation, cold pools, and domain size". en. In: *Geophysical Research Letters* 40.5, pp. 994–998.

Jensen, Gorm G., Romain Fiévet, and Jan O. Haerter (2022). „The Diurnal Path to Persistent Convective Self-Aggregation". en. In: *Journal of Advances in Modeling Earth Systems* 14.5, e2021MS002923.

Jung, Hyunju, Ann Kristin Naumann, and Bjorn Stevens (2021). „Convective self–aggregation in a mean flow". English. In: *Atmospheric Chemistry and Physics* 21.13, pp. 10337–10345.

Karman, Theodore von (1940). „The Engineer Grapples with Nonlinear Problems". In: *Bulletin American Meteorological Society* 46, pp. 615–683.

Khairoutdinov, Marat and Kerry Emanuel (2013). „Rotating radiative-convective equilibrium simulated by a cloud-resolving model". en. In: *Journal of Advances in Modeling Earth Systems* 5.4, pp. 816–825.

Khairoutdinov, Marat F and K Emanuel (2010). „Aggregated convection and the regulation of tropical climate". In: *29th conf. on hurricanes and tropical meteorology*. Amer. Meteor. Soc., P2–69.

Khairoutdinov, Marat F. and David A. Randall (2003). „Cloud Resolving Modeling of the ARM Summer 1997 IOP: Model Formulation, Results, Uncertainties, and Sensitivities". EN. In: *Journal of the Atmospheric Sciences* 60.4, pp. 607–625.

Kirsch, Bastian, Felix Ament, and Cathy Hohenegger (2021). „Convective Cold Pools in Long-Term Boundary Layer Mast Observations". In: *Monthly Weather Review* 149.3, pp. 811–820.

Kirsch, Bastian, Cathy Hohenegger, and Felix Ament (2024). „Morphology and growth of convective cold pools observed by a dense station network in Germany". en. In: *Quarterly Journal of the Royal Meteorological Society* 150.759, pp. 857–876.

Klein, Cornelia, Danijel Belušić, and Christopher M. Taylor (2018). „Wavelet Scale Analysis of Mesoscale Convective Systems for Detecting Deep Convection From Infrared Imagery". en. In: *Journal of Geophysical Research: Atmospheres* 123.6, pp. 3035–3050.

KMMI (2020). *KNMI (Koninklijk Nederlands Meteorologisch Instituut) Data Platform*. `https://dataplatform.knmi.nl/catalog/index.html`. Accessed: 2020-08-01.

Knippertz, Peter, John H Marsham, Angela Benedetti, *et al.* (2020). *Key lessons from the DACCIWA project for operational meteorological services*. eng.

Kohler, Martin, Geoffrey Bessardon, Barbara Brooks, *et al.* (2022). „A meteorological dataset of the West African monsoon during the 2016 DACCIWA campaign". en. In: *Scientific Data* 9.11, p. 174.

Kruse, Irene L. (2020). „Characterizing cold pool interactions over land with observational data from the Netherlands". MSc thesis. Utrecht University, Utrecht.

Kruse, Irene L., Romain Fiévet, and Jan O. Haerter (-). „Tipping to an Aggregated State by Mesoscale Convective Systems". Submitted to *Journal of Advances in Modeling Earth Systems*.

Kruse, Irene L., Jan O. Haerter, and Bettina Meyer (2022). „Cold pools over the Netherlands: A statistical study from tower and radar observations". In: *Quarterly Journal of the Royal Meteorological Society* 148.743, pp. 711–726.

Kruse, Peter (1997). „Experimental acute pancreatitis: Models for testing a free radical mechanism". PhD thesis. University of Copenhagen.

Kurowski, M. J., K. Suselj, W. W. Grabowski, and J. Teixeira (2018). „Shallow-to-Deep Transition of Continental Moist Convection: Cold Pools, Surface Fluxes, and Mesoscale Organization". In: *Journal of the Atmospheric Sciences* 75, pp. 4071–4090.

Lebel, Thierry, Douglas J. Parker, Cyrille Flamant, *et al.* (2011). „The AMMA field campaigns: accomplishments and lessons learned". en. In: 12, pp. 123–128.

Liu, C. and M. Moncrieff (1996). „A Numerical Study of the Effects of Ambient Flow and Shear On Density Currents". In: *Monthly Weather Review* 124, pp. 2282–2303.

Lorenz, Edward N. (1963). „Deterministic Nonperiodic Flow". EN. In: *Journal of the Atmospheric Sciences* 20.2, pp. 130–141.

Mai, Chuying, Yu Du, and Minghua Li (2023). „Processes of Colliding Cold Pools Derived from a 356-m-High Shenzhen Met-Tower during an Extremely Heavy Rainfall Event". EN. In: *Monthly Weather Review* 151.6, pp. 1571–1585.

Manabe, Syukuro and Robert F. Strickler (1964). „Thermal Equilibrium of the Atmosphere with a Convective Adjustment". EN. In: *Journal of the Atmospheric Sciences* 21.4, pp. 361–385.

Mapes, Brian E. and Robert A. Houze (1993). „Cloud Clusters and Superclusters over the Oceanic Warm Pool". EN. In: *Monthly Weather Review* 121.5, pp. 1398–1416.

Markowski, P. and Y. Richardson (2010). *Mesoscale Meteorology in Midlatitudes.* Wiley.

Mathon, Vincent, Henri Laurent, and Thierry Lebel (2002). „Mesoscale Convective System Rainfall in the Sahel". EN. In: *Journal of Applied Meteorology and Climatology* 41.11, pp. 1081–1092.

Mbow, C., A. Diop, A. T. Diaw, and C. I. Niang (2008). „Urban sprawl development and flooding at Yeumbeul suburb (Dakar-Senegal)". en. In: *African Journal of Environmental Science and Technology* 2.44, pp. 075–088.

Merz, Bruno, Christian Kuhlicke, Michael Kunz, *et al.* (2020). „Impact Forecasting to Support Emergency Management of Natural Hazards". en. In: *Reviews of Geophysics* 58.4, e2020RG000704.

Meyer, Bettina and Jan O. Haerter (2020). „Mechanical Forcing of Convection by Cold Pools: Collisions and Energy Scaling". In: *Journal of Advances in Modeling Earth Systems* 12.11.

Moncrieff, Mitchell W. and Changhai Liu (1999). „Convection Initiation by Density Currents: Role of Convergence, Shear, and Dynamical Organization". In: *Monthly Weather Review* 127.10, pp. 2455–2464.

Moseley, Christopher, Peter Berg, and Jan Olaf Haerter (2014). „Probing the convection life-cycle by Iterative Rain Cell Tracking." In: *Journal of Geophysical Research* 118.24, pp. 13, 361–13, 370.

Moseley, Christopher, Olga Henneberg, and Jan O Haerter (2019). „A statistical model for isolated convective precipitation events". In: *Journal of Advances in Modeling Earth Systems* 11.1, pp. 360–375.

Mueller, C. K. and R. E. Carbone (1987). „Dynamics of a Thunderstorm Outflow." In: *Atmos. Sci.* 44.15, pp. 1879–1898.

Muller, Caroline and Sandrine Bony (2015). „What favors convective aggregation and why?" In: *Geophysical Research Letters* 42.13, pp. 5626–5634.

Muller, Caroline, Da Yang, George Craig, *et al.* (2022). „Spontaneous Aggregation of Convective Storms". In: *Annual Review of Fluid Mechanics* 54.1, null.

Muller, Caroline J and Isaac M Held (2012). „Detailed investigation of the self-aggregation of convection in cloud-resolving simulations". In: *Journal of the Atmospheric Sciences* 69.8, pp. 2551–2565.

Nazari, Sara, Irene L. Kruse, and Nils Moosdorf (-). „Spatiotemporal Dynamics of Global Groundwater Recharge from 2001 to 2020". In review in *Journal of Hydrology.*

NCEP (1986). *NOAA NCEP Optimum Interpolation Sea Surface Temperature Analysis.* Boulder CO.

Nesbitt, Stephen W., Robert Cifelli, and Steven A. Rutledge (2006). „Storm Morphology and Rainfall Characteristics of TRMM Precipitation Features“. EN. In: *Monthly Weather Review* 134.10, pp. 2702–2721.

Nissen, S. B. and J. O. Haerter (2020). „How weakened cold pools open for convective self-aggregation“. In: *EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-1520*.

Nissen, Silas Boye and Jan O Haerter (2021). „Circling in on Convective Self-Aggregation“. In: *Journal of Geophysical Research: Atmospheres* 126.20, e2021JD035331.

Niu, Guo-Yue, Zong-Liang Yang, Kenneth E Mitchell, Fei Chen, Michael B Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, *et al.* (2011). „The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements“. In: *Journal of Geophysical Research: Atmospheres* 116.D12.

Ocasio, Kelly M. Núñez, Jenni L. Evans, and George S. Young (2020). „Tracking Mesoscale Convective Systems that are Potential Candidates for Tropical Cyclogenesis“. EN. In: *Monthly Weather Review* 148.2, pp. 655–669.

Patrizio, Casey R. and David A. Randall (2019). „Sensitivity of Convective Self-Aggregation to Domain Size“. en. In: *Journal of Advances in Modeling Earth Systems* 11.7, pp. 1995–2019.

Purdom, James F. W. (1976). „Some Uses of High-Resolution GOES Imagery in the Mesoscale Forecasting of Convection and Its Behavior“. In: *Monthly Weather Review* 104.12, pp. 1474–1483.

Randall, David A. and George J. Huffman (1980). „A Stochastic Model of Cumulus Clumping“. EN. In: *Journal of the Atmospheric Sciences* 37.9, pp. 2068–2078.

Rio, C, F Hourdin, J-Y Grandpeix, and J-P Lafore (2009). „Shifting the diurnal cycle of parameterized deep convection over land“. In: *Geophysical Research Letters* 36.7.

Robe, Françoise R. and Kerry A. Emanuel (1996). „Moist Convective Scaling: Some Inferences from Three-Dimensional Cloud Ensemble Simulations“. EN. In: *Journal of the Atmospheric Sciences* 53.22, pp. 3265–3275.

Roberts, Alexander J., Jennifer K. Fletcher, James Groves, *et al.* (2022). „Nowcasting for Africa: advances, potential and value“. en. In: *Weather* 77.7, pp. 250–256.

Roca, Rémy and Thomas Fiolleau (2020). „Extreme precipitation in the tropics is closely associated with long-lived convective systems“. en. In: *Communications Earth & Environment* 1.1, pp. 1–6.

Rooney, Gabriel Gerard (2018). „Similarity-based approximations for the evolution of a gravity current“. In: *Quarterly Journal of the Royal Meteorological Society* 144.716, pp. 2302–2310.

Rotunno, R., J. Klemp, and M. Weisman (1988). „A Theory for Strong, Long-Lived Squall Lines“. In: *Journal of The Atmospheric Sciences* 45, pp. 463–485.

Rowe, Angela K. and Robert A. Houze Jr. (2015). „Cloud organization and growth during the transition from suppressed to active MJO conditions“. en. In: *Journal of Geophysical Research: Atmospheres* 120.19, pp. 10, 324–10, 350.

Schlemmer, Linda and Cathy Hohenegger (2014). „The Formation of Wider and Deeper Clouds as a Result of Cold-Pool Dynamics“. EN. In: *Journal of the Atmospheric Sciences* 71.8, pp. 2842–2858.

Seifert, Axel (2008). „On the Parameterization of Evaporation of Raindrops as Simulated by a One-Dimensional Rainshaft Model“. In: *Journal of the Atmospheric Sciences* 65, pp. 3608–3619.

Skamarock, C., B. Klemp, Jimy Dudhia, *et al.* (2021). „A Description of the Advanced Research WRF Model Version 4.3“. en. In.

Sobel, Adam H, Johan Nilsson, and Lorenzo M Polvani (2001). „The weak temperature gradient approximation and balanced tropical moisture waves“. In: *Journal of the atmospheric sciences* 58.23, pp. 3650–3665.

Szoeke, S. P. de, E. D. Skyllingstad, P. Zuidema, and A. S. Chandra (2017). „Cold pools and their influence on the tropical marine boundary layer“. In: *Journal of the Atmospheric Sciences* 74.4, pp. 1149–1168.

Tan, Jackson, Christian Jakob, William B Rossow, and George Tselioudis (2015). „Increases in tropical rainfall driven by changes in frequency of organized deep convection“. In: *Nature* 519.7544, pp. 451–454.

Taylor, Christopher M, Richard AM de Jeu, Françoise Guichard, Phil P Harris, and Wouter A Dorigo (2012). „Afternoon rain more likely over drier soils“. In: *Nature* 489.7416, pp. 423–426.

Taylor, Christopher M., Danijel Belušić, Françoise Guichard, Douglas J. Parker, Théo Vischel, Olivier Bock, Phil P. Harris, Serge Janicot, Cornelia Klein, and Gérémy Panthou (2017). „Frequency of extreme Sahelian storms tripled since 1982 in satellite observations“. en. In: *Nature* 544.76517651, pp. 475–478.

Taylor, Christopher M., Cornelia Klein, Cheikh Dione, *et al.* (2022). „Nowcasting tracks of severe convective storms in West Africa from observations of land surface state“. en. In: *Environmental Research Letters* 17.3, p. 034016.

Terai, C. R. and R. Wood (2013). „Aircraft observations of cold pools under marine stratocumulus“. In: *Atmospheric Chemistry and Physics* 13.19, pp. 9899–9914.

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

Tompkins, A. M. (2001b). „Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Water Vapor." In: *Atmos. Sci.* 58.6, pp. 529–545.

Tompkins, Adrian M. and George C. Craig (1998). „Radiative–convective equilibrium in a three-dimensional cloud-ensemble model". en. In: *Quarterly Journal of the Royal Meteorological Society* 124.550, pp. 2073–2097.

Tompkins, Adrian M. and Addisu G. Semie (2017). „Organization of tropical convection in low vertical wind shears: Role of updraft entrainment". en. In: *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1046–1068.

Torri, G., Z. Kuang, and Y. Tian (2015). „Mechanisms for convection triggering by cold pools". In: *Geophysical Research Letters* 42.6, pp. 1943–1950.

Tramblay, Yves, Gabriele Villarini, and Wei Zhang (2020). „Observed changes in flood hazard in Africa". en. In: *Environmental Research Letters* 15.10, 1040b5.

Tropical Globe Radar Database (Accessed: 2024-02-26). `http://tropicalglobe.com/radar_database/`. Accessed: 2024-02-26.

UKCEH Nowcasting Portal (Accessed: [March 2024]). *Nowcasting Portal.* `https://eip.ceh.ac.uk/hydrology/sub-saharan-africa/nowcasting/`.

Vogel, R. (2014). „The influence of precipitation and cconvective organization on the structure of the trades". PhD thesis. Max-Planck-Institut für Meteorologie, Hamburg.

Vraciu, Cristian V., Irene L. Kruse, and Jan O. Haerter (2023). „The Role of Passive Cloud Volumes in the Transition From Shallow to Deep Atmospheric Convection". In: *Geophysical Research Letters* 50.23, e2023GL105996.

Wakimoto, R. M. (1982). „The Life Cycle of Thunderstorm Gust Fronts as Viewed with Doppler Radar and Rawinsonde Data." In: *Monthly Weather Review* 110, pp. 1060–1082.

Wakimoto, R. M. (2001). „Severe Convective Storms." In: Meteorological Monographs. American Meteorological Society, Boston, MA. Chap. Convectively Driven High Wind Events, pp. 255–298.

Weaver, John and Stephan Nelson (1982). „Multiscale Aspects of Thunderstorm Gust Fronts and Their Effects on Subsequent Storm Development". In: *Monthly Weather Review* 110, pp. 707–718.

Wilson, James W. and Wendy E. Schreiber (1986). „Initiation of Convective Storms at Radar-Observed Boundary-Layer Convergence Lines". In: *Monthly Weather Review* 114.12, pp. 2516–2536.

Wing, A. A., K. Emanuel, C. Holloway, and C. J. Muller (2017). „Convective self-aggregation in numerical simulations: A review". In: *Surveys in Geophysics* 142.694, pp. 1–15.

Wing, A. A., C. L. Stauffer, T. Becker, K. A. Reed, M.-S. Ahn, and N. P. Arnold (2020). „Clouds and convective self-aggregation in a multimodel ensemble of radiative-convective equilibrium simulations“. In: *Journal of Advances in Modeling Earth Systems* 12.9, e2020MS002138.

Wing, Allison A, Kerry Emanuel, Christopher E Holloway, and Caroline Muller (2018). „Convective self-aggregation in numerical simulations: A review“. In: *Shallow clouds, water vapor, circulation, and climate sensitivity*, pp. 1–25.

Wing, Allison A., Suzana J. Camargo, and Adam H. Sobel (2016). „Role of Radiative–Convective Feedbacks in Spontaneous Tropical Cyclogenesis in Idealized Numerical Simulations“. EN. In: *Journal of the Atmospheric Sciences* 73.7, pp. 2633–2642.

Wing, Allison A. and Timothy W. Cronin (2016). „Self-aggregation of convection in long channel geometry“. en. In: *Quarterly Journal of the Royal Meteorological Society* 142.694, pp. 1–15.

Wing, Allison A. and Kerry A. Emanuel (2014). „Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations“. en. In: *Journal of Advances in Modeling Earth Systems* 6.1, pp. 59–74.

Yanase, Tomoro, Seiya Nishizawa, Hiroaki Miura, Tetsuya Takemi, and Hirofumi Tomita (2020). „New Critical Length for the Onset of Self-Aggregation of Moist Convection“. en. In: *Geophysical Research Letters* 47.16, e2020GL088763.

Yang, Da (2021). „A Shallow-Water Model for Convective Self-Aggregation“. EN. In: *Journal of the Atmospheric Sciences* 78.2, pp. 571–582.

Young, G.S., S. M. Perugini, and Fairall C. W. (1995). „Convective Wakes in the Equatorial Western Pacific during TOGA“. In: *Monthly Weather Review* 123, pp. 110–123.

Zipser, E. J. (1977). „Mesoscale and Convective–Scale Downdrafts as Distinct Components of Squall-Line Structure“. In: *Monthly Weather Review* 105.12, pp. 1568–1589.

Zipser, E. J., Daniel J. Cecil, Chuntao Liu, Stephen W. Nesbitt, and David P. Yorty (2006). „WHERE ARE THE MOST INTENSE THUNDERSTORMS ON EARTH?“ en. In: *Bulletin of the American Meteorological Society* 87.8, pp. 1057–1072.

Zuidema, P., Z. Li, R. Hill, L. Bariteau, B. Rilling, C. Fairall, W. Brewer, B. Albrecht, and J. Hare (2012). „On Trade Wind Cumulus Cold Pools“. In: *Journal of Atmospheric Sciences* 69, pp. 258–280.

Zuidema, P., G. Torri, and C. Muller (2017). „A Survey of Precipitation-Induced Atmospheric Cold Pools over Oceans and Their Interactions with the Larger-Scale Environment.“ In: *Surv Geophys* 38, pp. 1283–1305.