Max Planck Graduate Center
mit der Johannes Gutenberg-Universität

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

Obtaining 3D convective characteristics from a machine learning-based integration of multi-sensor satellite observations

Dissertation

for the award of the academic degree of
"Doctor rerum naturalium" (Dr. rer. nat.)
in Atmospheric Science

at the faculties:
08 - Physics, Mathematics and Computer Science,
09 - Chemistry, Pharmacy, Geography and Geosciences,
10 - Biology,
and University Medicine

of the Max Planck Graduate Center
Johannes Gutenberg University Mainz

submitted by

Sarah Brüning
born 30.11.1993 in Greven

Mainz, 10 July 2025

1. Examiner: Prof. Dr. Holger Tost

2. Examiner: Prof. Dr. Michael Wand

Day of the oral examination: 08 September 2025

Declaration

I hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg University Mainz to insure standard scientific practice.

Sarah Brüning
Mainz, 10 July 2025

Acknowledgments

Removed for data protection reasons.

Abstract

Convective clouds are key players in the climate system of the Earth. They influence both incoming solar and outgoing terrestrial radiation, and they regulate the hydrological cycle through complex feedback mechanisms. Despite their importance, clouds remain one of the largest sources of uncertainty in climate models — posing persistent challenges to scientists around the globe. Among all cloud types, convective systems stand out due to their ability to evolve rapidly from harmless cumulus clouds into intense thunderstorms. Accurately predicting this evolution is vital for effective risk assessment and mitigation — especially as extreme weather events are expected to occur more frequently in a warming world.

Satellite observations offer profound insights into the behavior of convective clouds. Although the volume of satellite data has grown tremendously in recent decades, extracting meaningful patterns from these large and complex datasets remains a daunting task. However, recent advances in machine learning have introduced new tools that can help address this challenge. While satellite data often lack fine vertical resolution or have limited temporal coverage, machine learning techniques allow to bridge these gaps, revealing previously hidden patterns associated to the dynamic evolution of cloud systems.

This thesis addresses current challenges by developing a machine learning framework that combines multiple satellite datasets to improve our understanding of convective clouds. Specifically, it uses 2D imagery from the geostationary MSG SEVIRI satellite to predict 3D cloud structures as observed by the CloudSat cloud profiling radar, which provides 2D vertical cross-sections of the radar reflectivity. This approach helps overcome current limitations in vertical cloud profiling. The model evaluation demonstrates that it can accurately reconstruct both the vertical structure and the distribution of hydrometeors. By leveraging the high spatial and temporal resolution of MSG SEVIRI alongside the vertical detail from CloudSat, this method considerably enhances the availability of 3D cloud structures across broad regions on Earth.

Building on this foundation, the thesis applies the predicted 3D cloud fields to investigate tropical convective cloud behavior, with a focus on the role of convective cores and large-scale spatial clustering in shaping cloud structure. An adapted, object-based detection algorithm first identifies individual cloud objects and their cores, then tracks their evolution over time. This method may provide detailed insights into how cores influence the 3D structure of clouds and how such patterns relate to broader phenomena — such as convective organisation, a process closely tied to extreme weather events. The findings also reveal connections between cloud morphology, convective clustering, and larger-scale atmospheric dynamics, including the seasonal migration of the ITCZ.

Although this work centers on convective clouds, the machine learning framework developed here has broader applicability for questions related to atmospheric and climate sciences. By improving access to high-resolution, 3D atmospheric structures, it offers valuable tools for a wide range of studies on cloud processes — and lays the groundwork for supporting a more accurate assessment of climate-related risks in the future.

Zusammenfassung

Konvektive Wolken spielen eine zentrale Rolle im Klimasystem der Erde. Über ihre Feedback-Mechanismen beeinflussen sie sowohl die solare und terrestrische Strahlung als auch den globalen Wasserkreislauf. Trotz zahlreicher Studien bleiben Wolken eine der größten Unsicherheiten in Klimamodellen und stellen die Wissenschaft weiterhin vor große Herausforderungen. Besonders konvektive Wolken, die im Mittelpunkt dieser Arbeit stehen, bergen Risiken – vor allem durch ihre Fähigkeit, sich schnell von harmlosen Kumuluswolken zu schweren Gewittern zu entwickeln. Eine zuverlässige Vorhersage ist daher unerlässlich, gerade angesichts der Zunahme extremer Wetterereignisse im Zuge des Klimawandels.

Satellitendaten bieten herausragende Möglichkeiten, das Verhalten konvektiver Wolken besser zu verstehen. Zwar sind die Archive satellitengestützter Beobachtungen in den letzten Jahrzehnten massiv gewachsen, doch es bleibt herausfordernd, aus diesen komplexen Datensätzen verlässliche Erkenntnisse zu gewinnen. Fortschritte im Bereich des maschinellen Lernens – besonders in den letzten zehn Jahren – eröffnen hier neue Perspektiven. Solche Verfahren können helfen, typische Schwächen satellitengestützter Messungen, etwa eine eingeschränkte vertikale Auflösung oder begrenzte zeitliche Abdeckung, auszugleichen. Durch die Erkennung bisher verborgener Muster ermöglichen sie es, die Dynamik von Wolkenbildung und -entwicklung detaillierter zu erfassen.

Ziel dieser Arbeit ist es, unser Verständnis konvektiver Wolken mithilfe eines maschinellen Lernansatzes zu vertiefen, der verschiedene Satellitendaten kombiniert. Konkret werden zweidimensionale Bilder des geostationären MSG SEVIRI Satelliten verwendet, um daraus dreidimensionale Wolkenstrukturen vorherzusagen, wie sie vom CloudSat-Radar in Form vertikaler Querschnitte gemessen werden. Dieser Ansatz trägt dazu bei, bestehende Lücken im Bereich vertikaler Wolkenprofile zu schließen. Die Ergebnisse zeigen, dass das Modell sowohl die vertikale Struktur als auch die horizontale Verteilung von Hydrometeoren zuverlässig rekonstruieren und in die Fläche extrapolieren kann. Durch die Verbindung der hohen zeitlich-räumlichen Auflösung von MSG SEVIRI mit der vertikalen Detailtiefe von CloudSat kann die räumliche Abdeckung dreidimensionaler Wolkenstrukturen verbessert werden.

Mit Hilfe dieser Datensätze untersucht die Arbeit weiterhin das Verhalten konvektiver Wolken in den Tropen – mit besonderem Fokus auf die Rolle konvektiver Kernregionen sowie die Auswirkungen großräumiger Clusterbildung auf die 3D-Struktur von Wolken. Ein objektbasiertes Erkennungsverfahren identifiziert einzelne Wolken und ihre zugehörigen Kernregionen und verfolgt deren Entwicklung über die Zeit. So lassen sich präzise Rückschlüsse auf den Einfluss der konvektiven Kerne auf die dreidimensionale Struktur ziehen. Darüber hinaus wird untersucht, wie diese Muster mit großskaligen Phänomenen wie dem räumlichen Organisationsverhalten korrelieren – einem Prozess, der häufig mit extremem Wetter in Verbindung gebracht wird. Die Ergebnisse verdeutlichen Zusammenhänge zwischen Wolkenstruktur, verstärkter Clusterbildung und großräumigen Dynamiken wie der saisonalen Verlagerung der ITCZ.

Obwohl sich die Arbeit auf konvektive Wolken konzentriert, besitzt das entwickelte maschinelle Lernverfahren vielfältige Anwendungsmöglichkeiten. Die verbesserte Verfügbarkeit detaillierter 3D-Atmosphärenstrukturen kann künftige Studien zu Wolkenprozessen maßgeblich unterstützen – und damit einen Beitrag zu einer robusten Bewertung klimabedingter Risiken leisten.

Contents

Abstract **vii**

Zusammenfassung **viii**

1 Introduction **1**

2 Scientific background **3**

2.1 Clouds in the climate system 3

2.1.1 Atmospheric convection 3

2.1.2 Deep convective clouds 4

2.1.3 Convective organisation and self-aggregation 7

2.1.4 Convective clouds in a changing climate 8

2.2 Satellite-based remote sensing 9

2.2.1 Passive sensors 9

2.2.2 Active sensors 12

2.2.3 Limitations of remote sensing 14

2.3 Machine Learning 15

2.3.1 Theory and methods 15

2.3.2 Impact of machine learning applications 18

2.3.3 Machine learning in atmospheric sciences 19

2.3.4 Error sources and limitations 20

3 Results **21**

3.1 Overview 21

3.2 Machine learning-based reconstruction of 3D cloud fields 23

3.3 Relationship between convective clouds, their cores, and the cloud life-cycle 43

3.4 Investigating effects and spatial patterns of convective organisation 81

4 Conclusions and outlook **123**

List of Acronyms **127**

References **129**

1 Introduction

Motivation

Clouds are fundamental to the climate system of the Earth - they regulate the planet's radiative balance, shape the hydrological cycle, and influence climate dynamics across a range of spatial and temporal scales (e.g., Emanuel et al., 1994; Grabowski and Petch, 2009; Stevens, 2005). Despite major strides in understanding anthropogenic climate change, clouds remain one of the most persistent sources of uncertainty in climate models (e.g., Bony et al., 2015; Sherwood et al., 2020; Rolnick et al., 2022).

A key challenge lies in an accurate representation of processes at the convective scale (Gentine et al., 2018). Convective cloud systems, in particular, pose substantial risks to both ecosystems and human communities, often triggering extreme weather events such as intense winds, heavy rainfall, hail, and flash floods (Tippett et al., 2015). These phenomena result in significant loss of life and property every year. Between 1980 and 2020, meteorological and hydrological events were responsible for up to 10,000 deaths and an estimated €346–400 billion in damages across Europe alone (European Environment Agency, 2024). As such, accurately simulating cloud behavior is not only a scientific imperative - according to the Intergovernmental Panel on Climate Change (IPCC), it is also essential for mitigating the societal and environmental impacts of a warming world (Seneviratne et al., 2012).

To support climate adaptation strategies, researchers rely on a variety of data sources (Houze, 2018). In situ measurements offer high-resolution insights into atmospheric processes but are limited in both spatial and temporal coverage. Ground-based observations provide continuous time series yet remain geographically sparse, especially over oceanic regions. Satellite-based remote sensing, which is the focus of this thesis, addresses these limitations by offering near-global coverage with high spatial and temporal resolution (Young et al., 1998). The data allow monitoring the life-cycle of convective systems - from small cumulus clouds to intense thunderstorms - thereby helping us to understand cloud dynamics within the climate system.

Since the advent of operational satellite missions in the 1960s, the volume of atmospheric data has grown exponentially. However, extracting meaningful insights from these vast and complex data remains a challenge. The atmosphere is a highly dynamic and chaotic system, and even with abundant data, issues of coverage and quality can complicate analysis (Reichstein et al., 2019). To address these hurdles, atmospheric scientists are increasingly adopting tools from computer science. Among these, machine learning (ML) - and particularly deep learning (DL) - has emerged as a powerful approach for modeling complex, non-linear relationships in heterogeneous datasets (e.g., Beucler et al., 2024; Boukabara et al., 2019; Ebert-Uphoff and Hilburn, 2020). Owing to its flexibility and efficiency, ML is now a valuable tool in atmospheric and climate research (Rasp et al., 2018), enhancing predictive capabilities while often reducing computational costs compared to traditional modeling methods (Camps-Valls et al., 2014). By integrating diverse observational platforms and leveraging advanced data-driven techniques, researchers may continually improve our understanding of cloud behavior and its interactions with the climate. These efforts are crucial - not only for refining climate projections and improving weather forecasts, but also for strengthening societal resilience in the face of increasingly frequent and severe weather events.

Central objectives

This thesis aims to contribute to an advanced understanding of convective cloud behavior, which is crucial to assess climate risk around the globe. Specifically, the project explores how ML techniques may be leveraged to predict the three-dimensional (3D) structure of cloud systems from two-dimensional (2D) observational data. For instance, a focus lies on effectively using satellite data from active and passive remote sensing instruments to reconstruct the 3D cloud field, capturing both the horizontal spread and vertical evolution of individual clouds.

Following, the research is structured around three main objectives:

- *Reconstruction of a 3D cloud field*: The first objective involves designing and training a ML framework to infer contiguous 3D cloud structures from multi-sensor 2D satellite data, which may provide a more comprehensive view of cloud morphology than traditional approaches.
- *Identifying convective systems and their cloud life-cycle*: The second objective focuses on developing an object-based algorithm to detect and track convective cloud systems within the reconstructed 3D cloud field. By distinguishing between cloud and core regions, the study compares the development and life-cycle of convective cells from different convective regimes.
- *Analysing spatio-temporal patterns of convective organisation*: The third objective examines the spatial clustering among cloud objects, investigating a process known as "convective organisation". It aims on revealing how cloud properties relate to observed patterns of organisation. Special attention is given to understanding regional variations in convective organisation across tropical regions.

Together, these objectives may help to provide a novel perspective on studying cloud systems, based on a synergistic integration of open available observational data from multiple sensors. The results potentially offer new insights into the spatio-temporal variability of cloud structures, cloud development, and large-scale cloud organisation.

Structure of the thesis

This thesis provides an introduction to the topic, the key scientific objectives, and the structure of the thesis in Sect. 1, whereas Sect. 2 comprises the scientific background of the thesis. Section 2.1 describes the development of convective clouds and their role in the climate system. In Sect 2.2 and Sect. 2.3, the thesis provides an overview about principles of satellite-based remote sensing and the application of ML techniques, in particular in a meteorological context. Section 3 contains the results of the thesis. Section 3.2 has been published by the journal *Atmospheric Measurement Techniques*. It focuses on the first objective, dealing with a ML-based extrapolation of a contiguous 3D cloud field from 2D data. Section 3.3 is accepted for publication by the journal *Atmospheric Chemistry and Physics*. It presents the results for the second objective including the approach to detect and track convective clouds and their associated core regions in the predicted 3D field. The analysis focuses on the life-cycle and 3D properties of identified cloud objects. Likewise, Sect. 3.4 is accepted for publication by the journal *Atmospheric Chemistry and Physics*. The study is based on data from Sects. 3.2 and 3.3 and sheds a light on the third objective, which revolves around characterising regional patterns of convective organisation. The main conclusions, further challenges, and implications of this thesis are highlighted in Sect. 4.

2 Scientific background

About 67 % of the Earth's surface is covered by clouds at all times (King et al., 2013), making them one of the most prominent visible features of our atmosphere. They form when warm, moist air rises, cools, and condenses into water droplets (Stevens, 2005). Clouds come in a great variety of shapes and sizes along both the spatial and temporal scale. For example, on a hot summer day, the formation of shallow cumulus clouds is often followed by deep and precipitating convective clouds. Due to their ubiquitous occurrence, clouds play a critical role in shaping global weather and climate. They are accompanied by complex atmospheric and terrestrial feedback mechanisms which may alter precipitation patterns and large-scale weather dynamics (Grabowski and Petch, 2009). Given their profound impact on both natural ecosystems and human activities, understanding and accurately simulating cloud behavior is of vital importance.

2.1 Clouds in the climate system

Clouds may reflect, absorb, or scatter incoming solar and terrestrial radiation. In the solar wavelength regions, they cool the Earth by reflecting sunlight to space. In contrast, thermal wavelength regions are connected to the enhanced warming of the earth as clouds absorb and re-emit the radiation emitted by the surface and lower atmosphere (Wielicki et al., 1996). The cloud radiative properties are inherently connected to the cloud type. Optically thick, low-level clouds usually have a negative net radiative forcing as their thermal effect is small and reflection of solar radiance dominates. In contrast, the net radiative effect of high-level clouds is often positive due to a substantial thermal contrast between them and the surface (Boucher, 1999; Schumann et al., 2012). Monitoring clouds and quantifying their associated dynamics is essential to understand their role in the weather and climate system. Despite the progress research has achieved in the last decades, an insufficient representation of clouds in climate models may still considerably disturb the energy balance of global climate models (Wilcox et al., 2023; Beucler et al., 2021).

2.1.1 Atmospheric convection

Atmospheric convection operates across a wide range of spatial and temporal scales and plays a central role in the general circulation of the Earth. It facilitates the vertical transport of mass, momentum, and energy needed to offset imbalances caused by net radiative fluxes (Hartmann et al., 2019). For instance, convection can be defined as the vertical transport of sensible and latent heat driven by an unstable air mass distribution (Emanuel et al., 1994). In principle, it occurs in two forms: dry convection, which shapes the atmospheric boundary layer without cloud formation, and moist convection, which relates to the visible convection process. The latter occurs when moist air rises and cools adiabatically, gaining extra buoyancy from the release of latent heat (Figure 1). This additional supply of buoyancy is stored in the water vapour content, creating conditional instability and fueling cloud development (Houze Jr., 2004). In the following sections, this thesis will focus on moist convection as a fundamental driver of severe weather.

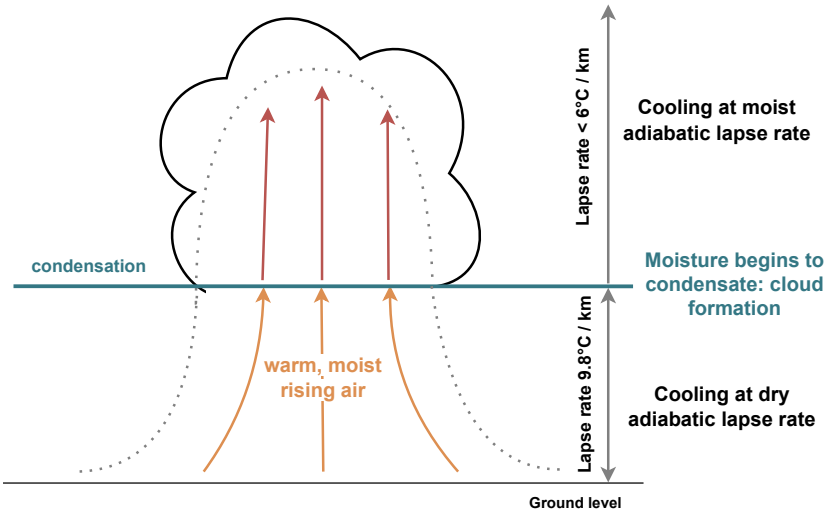

Figure 1: A simplified schematic for the formation of moist convection under diabatic and adiabatic conditions. Figure by the author based on Wielicki et al. (1996) and Emanuel et al. (1994).

2.1.2 Deep convective clouds

Clouds are not only integral to the large-scale atmospheric circulation but also influence the spatial distribution, frequency, and intensity of precipitation (Nesbitt et al., 2006). Among them, deep convective clouds are particularly important. Their formation and evolution are driven by moist convection and the release of latent heat. These cloud systems feature towering vertical structures with one or more active convective cores, strong updrafts and low temperatures. At higher altitudes, these updrafts often merge together inducing expansive anvil-shaped outflows near the cloud top (Fan et al., 2013).

Formation of mesoscale convective systems

Deep convective clouds may frequently merge together into large-scale systems. A prominent example of these are mesoscale convective systems (MCSs) - organised clusters of deep convection that contribute substantially to total precipitation in the tropics and subtropical midlatitudes (Aumann et al., 2018). These systems typically exhibit a horizontal extent exceeding 100 km in at least one dimension and persist for several hours (Houze, 2018). MCSs consist of one or more deep convective cores (hereafter: cores) embedded within extensive stratiform cloud regions. These cores are characterised by intense updrafts, high rainfall rates, and radar reflectivities indicative of strong convective activity, while the surrounding stratiform regions feature widespread moderate precipitation and extensive anvil cloud coverage (Leary and Houze, 1979). The vertical heating profiles associated with MCSs strongly influence large-scale circulations, including the Hadley and Walker cells. In the tropics, MCSs are the dominant contributors to latent heat release and significantly modulate the cloud radiative forcing, particularly through longwave cloud-top emissions and shortwave cloud albedo effects (Wilcox and Ramanathan, 2001). The temporal scale of extreme precipitation is closely tied to the structural and dynamical characteristics of convective systems. Short-duration extremes (e.g., flash floods) are often associated with isolated

or small-scale convective cells exhibiting high rainfall intensity and rapid development. In contrast, longer-duration events are typically linked to larger MCSs with prolonged lifespans and sustained moisture convergence. Hence, it is essential to understand the spatio-temporal distribution and variability of convective systems.

Convective clusters, such as MCSs, are frequently associated with hazardous weather, including hail, damaging winds, and tornadoes (Houze Jr., 2004). Here, thunderstorms display a particular threat. A thunderstorm is a specific type of MCS (Zipser and Liu, 2021; Lafore and Moncrieff, 1989), which requires sufficient moisture in the air, unstable air masses by a steep mid-tropospheric lapse rate, and a lifting force (Doswell III, 2001). They can be classified by their physical characteristics, such as the presence or absence of rotation, their anvil size, the number of locations of updrafts and downdrafts, or the occurrence of overshooting tops that penetrate the upper cloud layer (Figure 2). While single-cell thunderstorms are short-lived, typically lasting less than one hour, multicell systems involve organised clusters that can persist for several hours (Chan et al., 2023). Supercell thunderstorms, distinguished by their rotating updrafts (mesocyclones), are the most intense and dangerous convective storms due to their rapid development and complex dynamics (Sieglaff et al., 2011).

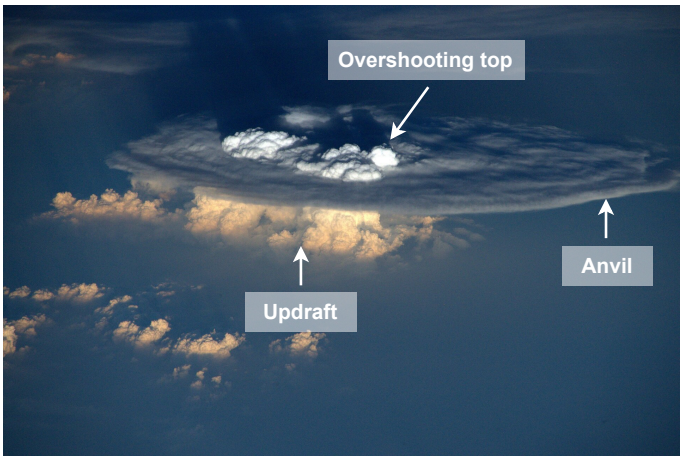

Figure 2: A cumulonimbus cloud recorded from the International Space Station (ISS) during the Principia mission between December 2015 and June 2016. The image shows the morphology of a mature thunderstorm including its updraft and anvil regions. Overshooting tops appear at the central part of the anvil where strong updrafts cause to penetrate the cloud top layer. Figure by the author, image credit: ESA/NASA/T. Peake.

The idealised life-cycle of a convective cloud follows three stages (Futyan and Genio, 2007). The first stage, the development stage, is characterised by enhanced updrafts that lead to the vertical growth of the cloud. Convective core regions start to form. They remain in the following maturity stage while the cloud reaches its maximum vertical expansion. An intense outflow at the cloud edges induces the formation of a stratiform cloud anvil. Ongoing updrafts may lead to the formation of overshooting tops that penetrate through the upper cloud layer (Machado et al., 1998). Downdrafts initiate the onset of precipitation, which is maintained in the following stage. During this dissipation stage, enhanced downdrafts cause the cloud to fade out slowly (Figure 3).

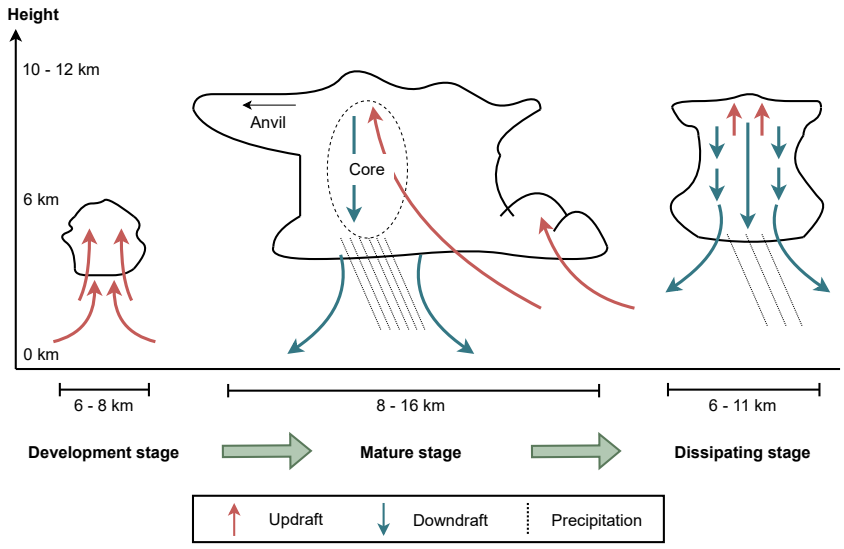

Figure 3: Simplified concept of the cloud life-cycle for an unicellular thunderstorm. It follows three stages, which comprise: **(a)** the development stage, **(b)** the mature stage, and **(c)** the dissipating stage. Figure by the author based on Doswell (1985) and Houze Jr. (1989).

Deep convection in the tropics

Overall, a high proportion of extreme events caused by convective activity is located in tropical and subtropical regions (Wu, Dong, et al., 2020), particularly over continental regions during the summer months (Houze Jr., 2004). However, the intensity and spatial organisation of deep convection are highly dependent on the surrounding dynamical and thermodynamic environment. Regional variability arises from differences in land–sea contrast, topography, and atmospheric circulation (Chen and Houze, 1997). Convective clouds over land are characterised by more rapid heating, variable moisture availability, and strong diurnal forcing. Peak activity typically occurs in the late afternoon and early evening when surface heating is at its maximum. For instance, orographic features can trigger intense but short-lived convective events. In contrast, oceanic convection is modulated by the ocean's high heat capacity, abundant and consistent moisture supply, uniform surface temperature, and large-scale atmospheric circulation patterns. These factors support more persistent and widespread convective systems - which appear often enhanced at night due to radiative cooling at the cloud top (e.g., Chen et al., 2021; Wu, Yuan, et al., 2020; Seidel et al., 2008).

Impact of severe weather

Extreme weather events, such as convective thunderstorms, are often linked to specific cloud microphysical processes (Ashley et al., 2019). For instance, towering convective clouds with very high cloud tops may transport near-surface hydrometeors and pollutants to the upper troposphere and even to the tropopause (Jensen and Pfister, 2004), affecting the stratosphere-troposphere exchange (Holton et al., 1995; Sherwood and Dessler, 2000). Through their impact on atmospheric conditions, convective storms may pose a peculiar threat to human health, infrastructure, and ecosystems (Bouwer, 2019). Accurate forecasting, early warning, and effective mitigation rely on a

detailed understanding of their microphysical characteristics and dynamic interactions (Wilson and Mueller, 1993). Despite advances in capturing the spatio-temporal variability of deep convection, key mechanisms remain poorly understood (Kunz, 2007). In response, a mischaracterisation of extreme events may continue to contribute to socio-economic and environmental damages around the world (e.g., Zipser and Liu, 2021; Chan et al., 2023).

2.1.3 Convective organisation and self-aggregation

Although the term "convective organisation" has become increasingly popular among researchers, it is often used vaguely. For example, Mapes and Neale (2011) broadly summarise convective organisation as "non-randomness in meteorological fields in convecting regions". The process describes the spatial clustering of deep convective cells, which is ubiquitous throughout the atmosphere, particularly in the tropics.

Based on this, clouds may appear in spatial patterns ranging from an unorganised or random to organised state. In an unorganised state, the clouds should be randomly distributed in space where their morphological and lifetime properties represent independent draws from the same probability density function. In contrast, convective clusters are classified as organised when they show a spatial or temporal correlation to other clouds (Muller et al., 2022). A spatial correlation arises when the clouds' positions are more clustered or regular than random. Moroever, a temporal correlation occurs when unique timesteps are affected by the attributes of previously existing convective cells (Bläckberg and Singh, 2022; Holloway, 2017). This organised state occurs on scales ranging between 100–1000 km, comprising multiple types of convective systems such as single-cell thunderstorms, linear squall lines, or mesoscale convective complexes (Tan et al., 2015). However, investigating convective organisation from observational data remains challenging - there exists no distinct quantification on how to identify and differentiate organisation. To this day, relatively little is known about the relationship of convective organisation with the large-scale state of the atmosphere. The morphological variability of convective organisation makes it hard to measure. In response, previous studies have developed various metrics aiming towards a deeper understanding about the underlying physical mechanisms. These indices analyse the spatial distribution of clouds within a defined area, contributing to approximating the strength of spatial clustering within this domain (Pscheidt et al., 2019).

In contrast, idealised model setups configured in radiative-convective equilibrium (RCE) may demonstrate a large-scale clustering of convective clouds, which occurs on a timescale between days and weeks (Wing, 2019). Despite homogeneous initial conditions, the domain eventually gets separated into moist and dry regions, with convection confined to the moist regions and a surrounding dry subsiding atmosphere (Hartmann et al., 1984). Studies suggest that an approximately random distribution of convective cells starts to differentiate into convecting and non-convecting regions that grow upscale over time (Figure 4). This process describing the transition from an initially random state to an organised state is known as "self-aggregation" of convection (e.g., Bretherton et al., 2005; Held et al., 1993; Wing et al., 2017). Numerous studies have been able to observe this process (e.g., Wing et al., 2018; Hohenegger et al., 2020; Tompkins, 2001). Research emphasises that self-aggregation may be driven by either internal dynamics, like cold pools and radiative feedback, or external forces, such as the land-sea-breeze and other large-scale flows (e.g.,

Muller and Held, 2012; Haerter et al., 2019; Tompkins, 2001). Moreover, the feedback between convection, surface fluxes, and radiation triggers and further drives aggregation (Windmiller and Hohenegger, 2019).

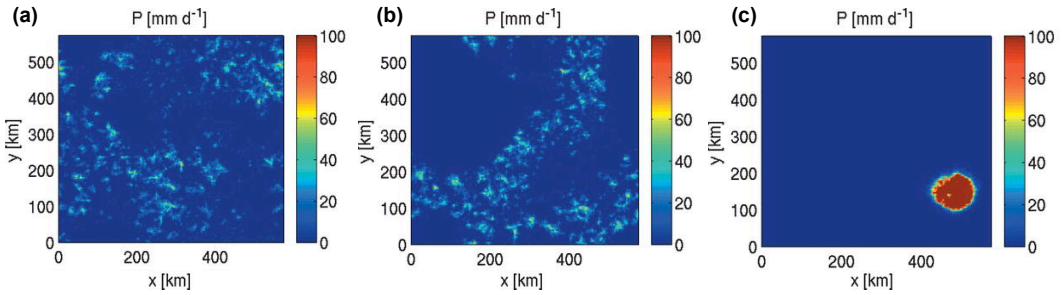

Figure 4: Time series of the daily accumulated precipitation [mm d^{-1}] from a model configured in RCE, used as an approximation for cloud activity. The panels show the temporal dependency of convective self-aggregation: While the spatial distribution of the clouds follows initially a more random pattern at day 10 **(a)** and 20 **(b)**, a single moist domain develops at day 50 **(c)**, suggesting a more organised state. Figure by the author, adapted from Bretherton et al. (2005).

These simulations induced new insights into how convection may interact with climate, e.g., helping to explain influences on the width and intensity of the Intertropical Convergence Zone (ITCZ) or the Madden-Julian oscillation (Wodzicki and Rapp, 2022). While Holloway (2017) found self-aggregation relevant to real-world convection, a simple transfer to a observational perspective is difficult as climate conditions (e.g., surface temperatures) are neither homogeneous nor constant (Tompkins and Semie, 2017). Previous findings highlight how further investigations are necessary to estimate the role and impact of convective organisation (Bao and Sherwood, 2019).

2.1.4 Convective clouds in a changing climate

Over the last century, increasing concentrations of greenhouse gases accelerated the warming of the Earth. Extreme events, particularly those connected to deep convection, shape the public's perception of climate change like no other (Wong and Teixeira, 2016; Singh et al., 2017). Despite substantial advances, existing uncertainties concern the response of convective clouds to a warming climate.

So far, simulations by Tan et al. (2015) showed that convective organisation and extreme precipitation will likely increase. Depending on the scenario, extreme events may become stronger, more frequent, or both in climate models (Rolnick et al., 2022). Notably, research widely agrees that future environments will contain favourable conditions for extreme events associated with deep convection (Feng et al., 2016; Westra et al., 2014). The predictions follow the Clausius-Clapeyron relation, stating that saturated atmospheric moisture increases almost exponentially with temperature (Clausius, 1850). For instance, the most extreme scenarios show an increase in the size of MCSs and the intensity of their associated maximum precipitation rates (Prein et al., 2017).

Open questions remain about the intensity and spatio-temporal distribution of future extremes (Agrawal et al., 2019). Research has shown that the occurrence of severe weather follows a non-uniform spatio-temporal distribution, with a focus on tropical regions around the ITCZ (Dougherty

and Rasmussen, 2020). In these areas, projected changes could be particularly devastating. They may induce a drying of the subtropics through increased subsidence and lead to an intensification of desertification (Seidel et al., 2008), which may affect agriculture, property, and the distribution of water resources (Spekkers et al., 2017; Bouwer, 2019). While extreme precipitation events are closely tied to organised convective systems, quantifying the overall connection between convective organisation and warming surface temperatures shows more ambiguous results (Nesbitt et al., 2006). Model limitations arise due to a coarse resolution and systematic bias in the distribution and intensity of convection (Sherwood et al., 2014). Nevertheless, a deeper understanding of convective organisation appears fundamental to accurately predict precipitation extremes in the present and future (Muller and Bony, 2015).

2.2 Satellite-based remote sensing

The launch of the first meteorological satellite in 1960, the Television InfraRed Observation Satellite (TIROS-1), marked the starting point for the operational usage of satellite-based atmospheric observations (Bluestein et al., 2022). Today, various sensors are in orbit to estimate the global cloud distribution, which is fundamental to diverse meteorological and climatological applications. Satellites provide detailed insights about the atmosphere, even above remote areas where it is challenging to deploy ground-based sensors (Lee et al., 2021; Fu et al., 2019). These data deepen our understanding of the Earth-atmosphere interactions, making satellites an invaluable asset to understanding and mitigating climate change and its consequences (Kidd et al., 2009).

Atmospheric observations from satellite systems provide diagnostic information for assessing and analysing meteorological features (Gómez-Chova et al., 2015). Data from satellites can be leveraged to measure radiation, surface temperature, wind, aerosols, clouds, or precipitation (Thies and Bendix, 2011). Typical cloud parameters derived from satellite comprise, e.g., the cloud-top height (CTH) (Schmetz et al., 2002), cloud optical thickness (COT) (Nakajima and King, 1990), cloud effective particle radius (CER) (Liou, 1992; Winker et al., 2017), cloud liquid water path (CWP) (Várnai and Marshak, 2002), or the cloud type (Bankert et al., 2009). Technical advances, like the increase in temporal and spatial resolution of the sensors, steadily advanced the quality and quantity of satellite data. Today, vast data archives display the basis for creating versatile applications for forecasting and monitoring cloud activities (Hilburn et al., 2020).

2.2.1 Passive sensors

Passive remote sensing sensors obtain global observations of the Earth and its atmosphere independent of the weather or daytime. They do not provide a source of energy, but detect natural energy (radiation) that is emitted, reflected, and scattered into space by Earth's atmosphere (Stephens and Kummerow, 2007). Optical sensors typically cover wavelengths from the visible light (VIS) between 0.4–0.6 μm to the near (NIR) and thermal infrared (IR) spectrum up to about 13 μm (Figure 5).

As clouds appear on various scales, data with a high spatial and temporal resolution is necessary to capture their occurrence around the globe. Here, passive satellite imagery may help to capture unprecedented detail of cloud fields occurring at different latitudes, e.g., over the Atlantic Ocean

(Figure 6). With a fast repeat cycle, the information from satellite channels can be leveraged to continuously monitor the spatio-temporal behaviour of clouds. They allow an in-depth analysis of cloud properties over land and sea, closing data gaps connected to ground-based or in situ measurements.

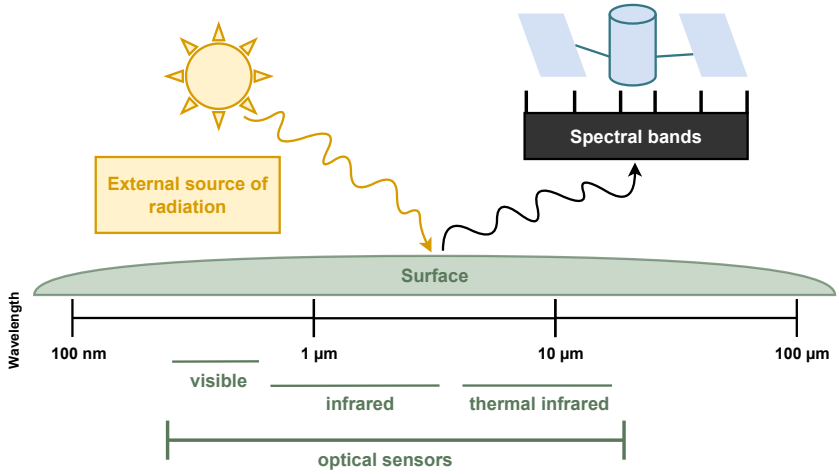

Figure 5: Schematic visualisation of multispectral optical sensors. Individual sensors can be distinguished by their number of spectral bands which lie in the visible, near-infrared, and thermal infrared spectrum across which the sensor measures radiance. Figure by the author based on Bluestein et al. (2022) and Pettorelli et al. (2018).

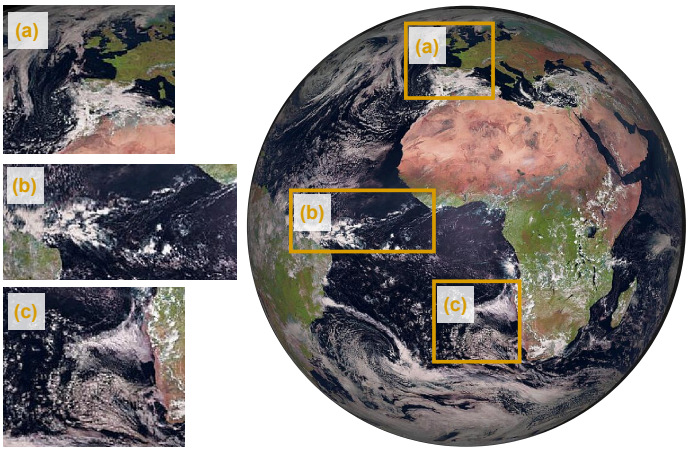

Figure 6: Satellite image received from the MSG SEVIRI sensor taken on April 22, 2022. The image covers the full disk of the sensor with a nadir centered over 0° longitude. Images **(a)** - **(c)** show oceanic cloud systems over Europe **(a)**, tropical convective clouds near the equator **(b)**, and marine cumulous fields near the tropic of Capricorn **(c)**. Figure by the author, image credit: ESA/METEOSAT.

There exist two types of sensors: low-Earth (including polar) orbiting system (LEO) and geostationary satellite system (GEO) (Figure 7). The LEO satellites - such as the National Aeronautics and Space Administration's (NASA) Moderate-resolution Imaging Spectroradiometer (MODIS) -

circle the Earth in a sun-synchronous orbit with consistent illumination of the Earth-scan view (Gorooh et al., 2023). Most LEO orbits are located between 500 and 2000 km altitude, providing global imagery through a cross-track scanner with rotating mirrors at least twice daily with a spatial resolution of less than 1 km (Salomonson et al., 2006; L'Ecuyer and Jiang, 2010). They are complemented by GEO systems - such as the Meteosat (MSG) Second Generation Spinning Enhanced Visible and Infrared Imager (SEVIRI) - which are positioned 35 786 km over the Equator. In contrast to polar-orbiting satellites, they appear quasi-stationary, covering large areas with a high temporal resolution and an orbital period of 24 h. Hence, they provide frequent observations covering the same area (Schmetz et al., 2002). Their spatial resolution typically ranges between 1 km to 3 km. Depending on the sensor, they receive information in intervals of 5 min to 15 min. Rapid scan techniques may enable a high temporal resolution of 1 minute, which can be used to monitor the dynamic development of even fast-evolving cloud clusters. Through a harmonisation of related sensors, GEO systems enable global investigations of clouds and their feedback mechanisms (Kidd et al., 2009).

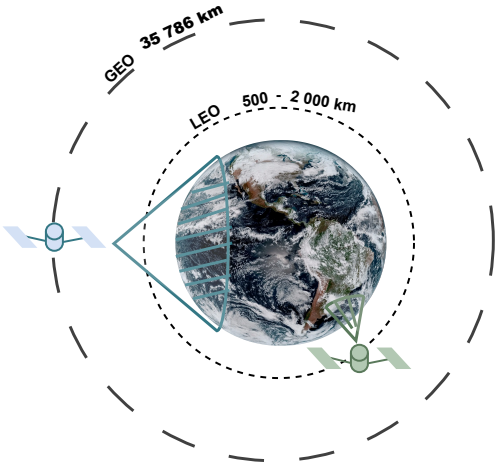

Figure 7: Comparison of the height and sensor coverage for satellites on the LEO and GEO orbit. The schematic illustration of the orbit is overlayed on a full disk image of the GOES geostationary satellite focused over the Americas taken on August 21, 2017. Figure by the author, image credit: NOAA/NASA.

Spectral and textural features captured by satellites play a key role in characterising clouds and their properties (Lee et al., 2021). To achieve a robust estimation of these properties, satellites analyse radiances across various spectral bands, also known as atmospheric window channels (Figure 8). In these channels, atmospheric gases absorb very little radiation, allowing clearer observations of the Earth's surface or cloud tops (Gao et al., 1998). The specific wavelengths used in these window channels vary between satellite missions (Zhang et al., 2019). However, each wavelength — alone or in combination — helps distinguish between surface types (e.g., vegetated vs. non-vegetated areas) and atmospheric features such as cloud type, cloud phase, or aerosol content (Amato et al., 2008). Using this information, satellites can detect cloud structures (Bankert et al., 2009), analyse their properties (Nakajima and King, 1990), and track their evolution over time (Menzel, 2001).

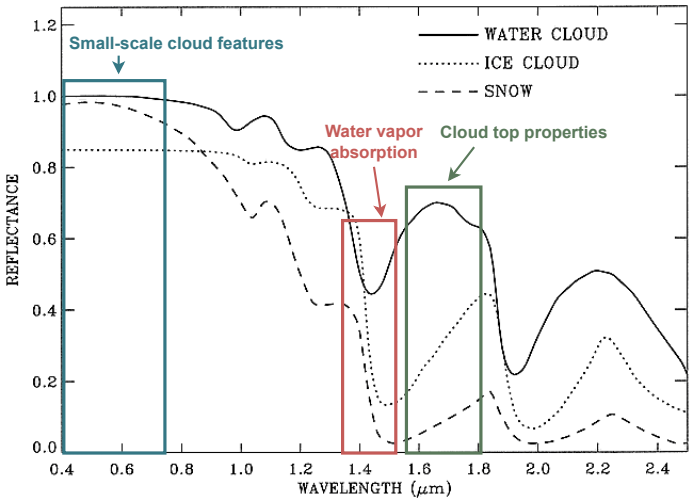

Figure 8: Overview of wavelengths used to extract information of different cloud types. The image shows the simulated directional–hemispherical reflectance spectra of a water cloud, an ice cloud, and a snow surface with the window channels along the atmospheric windows. Figure by the author based on Gao et al. (1998).

Various spectral bands are suitable for detecting cloud properties and approximating associated processes (King et al., 2013). In the visible range (VIS) between 300–700 nm, the channels receive part of the solar radiation reflected by the clouds or the Earth's surface. At daytime, they help to identify visual features, such as cloud edges, and structures connected to small-scale features, such as overshooting cloud tops or developing cumulus clouds (Schmetz et al., 2002). Near-infrared (NIR) channels span wavelengths from 0.7 to 3 μm (Hilburn et al., 2020). They provide insights into how clouds and surfaces reflect, transmit, and absorb NIR radiation, enhancing the contrast between water particles and their surroundings. For instance, the NIR channel at 1.37 μm shows a strong water vapour absorption, determining it suitable for detecting thin cirrus clouds during daytime. Meanwhile, the 1.61 μm channel helps determine the phase of cloud tops: liquid water clouds appear bright due to strong reflectivity, while ice clouds and snow seem darker because they absorb more radiation at this wavelength (Thies and Bendix, 2011; Menzel, 2001). In the IR regions, satellites measure brightness temperatures related to thermal radiation emitted by either the Earth's surface or clouds. As these IR channels are daytime independent, they can be used for contiguous cloud mapping, surface temperature determination, or to detect water vapour across different atmospheric layers (Stephens and Kummerow, 2007).

2.2.2 Active sensors

Most satellite-based cloud observations rely on passive sensors that detect solar (short-wave) radiation reflected by the Earth's atmosphere and thermal (long-wave) radiation emitted from it (Kidd et al., 2009). While valuable, these passive measurements primarily capture information about the cloud top and offer only limited insights into the vertical structure of clouds and hydrometeors. To address this limitation, space-borne active sensors offer a powerful complement. Unlike passive systems, active sensors emit their own energy and measure how the transmitted signal is reflected,

refracted, or scattered by the Earth's surface or atmosphere (Figure 9). This approach enables the identification of cloud droplets and ice crystals by passing through gases, clouds, and even solids to a limited degree (Stephens et al., 2008). As a result, active sensors augment the view on atmospheric processes and can retrieve valuable information on the vertical distribution of aerosols and precipitation particles under most atmospheric conditions (Okamoto and Sato, 2018; Dubovik et al., 2021).

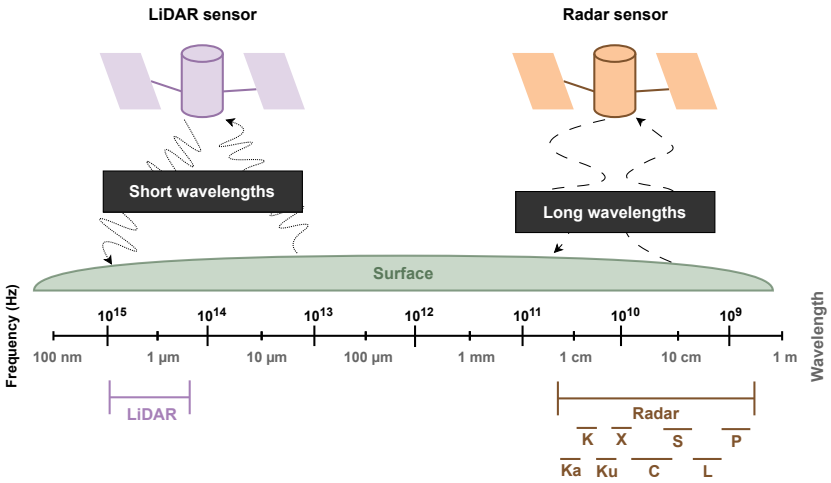

Figure 9: Overview of active sensors (LiDAR, radar) and the wavelengths at which they emit radiation. Figure by the author based on Bluestein et al. (2022) and Pettorelli et al. (2018).

The most common active remote sensing types are radar (radio detection and ranging) or LiDAR (Light Detection and Ranging) (Delanoë et al., 2013; Stubenrauch et al., 2010). Each instrument has a specific sensitivity that depends on the wavelength at which it emits radiation (Pettorelli et al., 2018). While LiDAR sensors primarily operate in the IR spectrum between 750 nm and 1.5 μm, radar sensors typically employ microwave and radio wavelength regions between 0.3 cm (111 GHz) and 120 cm (0.25 GHz).

For LiDAR, the active sensor sends out laser pulses to detect backscattered radiation. The sensor receives the reflection of the impulse it gave, whereas the distance from the backscatter location is determined using the time delay of the signal and the speed of light. The intensity of the backscattered energy depends on the number of particles, the particle size and the squared mass (Andersen et al., 2006). While LiDAR is sensitive to the concentration of hydrometeors and small particles like aerosols, it may be attenuated or extinguished in regions with high particle concentrations (Aubry et al., 2024).

A cloud profiling radar (CPR) - such as NASA's CloudSat - generates and emits radiation as regular energy pulses, with waves vibrating in a predetermined orientation. It typically operates between 35 GHz and 95 GHz, at which the radar may receive the echo from precipitation droplets to determine the rainfall rate of hydrometeors (Haynes et al., 2009). This data can be leveraged to discriminate vertical cloud profiles and the corresponding cloud type or phase (Battaglia et al., 2020). However, the radar signal can be considerably contaminated by orographic features at low

altitudes. Moreover, the sensor lacks sensitivity for ice clouds and shallow convection (Sassen and Wang, 2008). Nevertheless, the radar echo return from clouds provides information to reconstruct a three-dimensional profile of the cloud radar reflectivity, which measures the intensity of the cloud scattering over the Earth's surface. The data allow analysis of the cloud vertical column, e.g., to assess the updraft and axis length of tropical storms (Figure 10). Synergistic approaches have recognised potentials in combining data from passive and active sensors, aiming towards a deeper understanding of hydro-meteorological processes (Bellon et al., 1980; Jones et al., 2015).

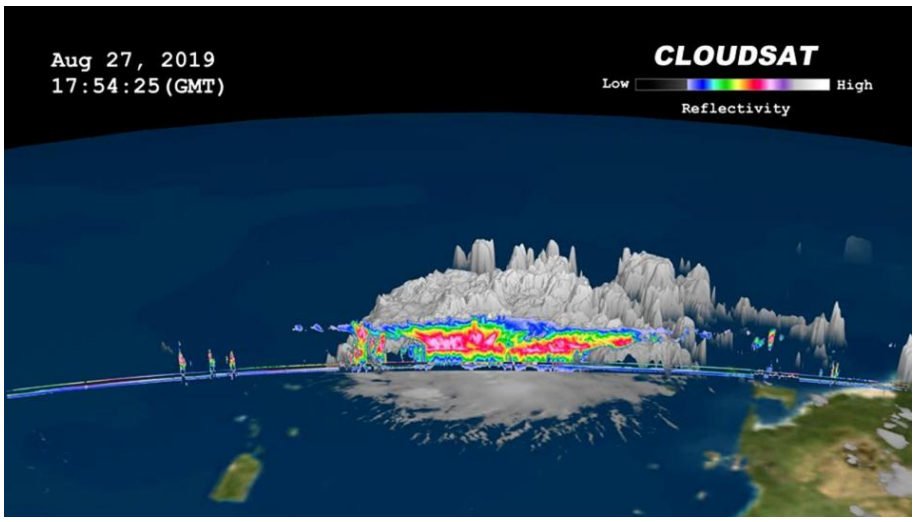

Figure 10: CloudSat image of tropical storm Dorian acquired on August 27, 2019, near Puerto Rico. The image shows the radar reflectivity in dBZ associated to each vertical level along the polar-orbiting CPR flight track. Image credit: NASA/JPL-Caltech.

2.2.3 Limitations of remote sensing

Despite technological progress, challenges remain. At first, observations do not fully constrain the retrieval problem. Satellite pixels typically cover an area with inhomogeneous clouds (Barker et al., 2011). Most information - such as from passive sensors - originates from the cloud top, whereas clouds are vertically extended and have complex structures. A single cloud layer can only approximate reality. Moreover, the thermal emission of the cloud comes from within the uppermost cloud layer. In the case of temperature inversions, the conversion from observed brightness temperature to CTH can be ambiguous and may lead to large displacements (Lee et al., 2021). In contrast, active sensors cover only a small horizontal extent. Cross-sections of the atmosphere may provide limited data on the vicinity and 3D properties of detected clouds. During the daytime, retrievals remain challenging for specific observation geometries such as a high solar zenith angle. Further uncertainties involve ice crystals, for which the shape determines their scattering and absorption properties, the state of the atmosphere, the albedo and emissivity of the surface, and the calibration and degradation of the satellite sensor (Zhao et al., 2023; Battaglia et al., 2020). Spectral similarities can distort the classification of different surface types (Henderson et al., 2013). All sensors underlie dimensional restrictions with retrievals only on a horizontal plane or along a limited 3D area (Dubovik et al., 2021). Although data from multiple sensors can be combined to close data gaps, no operating sensor with global 3D coverage exists to this day.

2.3 Machine Learning

In recent years, with the advent of big data and efficient supercomputers, the application of ML technologies across science registered an unprecedented growth in popularity (Sarker, 2021). Initially, the term "machine learning" described the "field of study that gives computers the ability to learn without being explicitly programmed" (Samuel, 1959). It comprises methods for a data-driven optimisation of complex functions that can be applied in artificial intelligence (AI) tasks to extract insights from diverse data sources. For instance, ML allows systems to learn and enhance from experience (LeCun et al., 2015). The following section provides a brief overview of different ML methods over the years, focusing on recent approaches, their application in atmospheric sciences, and their limitations.

2.3.1 Theory and methods

While some of the key concepts emerged more than 60 years ago, the popularity of ML varied over time. Following Goodfellow et al. (2016), it can be divided into three waves closely related to technological advances (Figure 11). The first wave (1970s - 1990s) focused on rule-based applications and feature engineering. These systems were good at reasoning but could not generalise well. In the following decades (1990s - 2000s), statistical models for specific domain problems exceeded those of the first wave in terms of perceiving and learning. However, they showed little ability to reason or generalise. The third wave of research started in 2006. Increasing computational power led to more explainable and applicable results. This wave popularised the term "deep learning" (DL), which refers to a subspace of ML and describes a technique that enables computer systems to improve with experience and data. At this time, DL outperformed competing AI systems. Since then, new concepts have emerged and steadily improved the performance of ML models on an ever-growing bandwidth of applications.

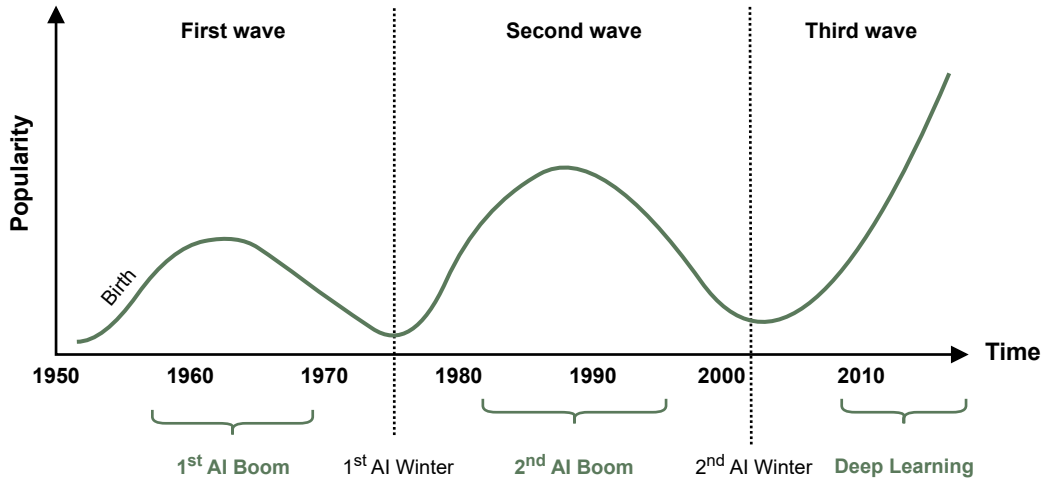

Figure 11: Simplified schematic for the three historical waves of AI research. Figure by the author based on Goodfellow et al. (2016).

Artificial neural networks

DL belongs to the family of artificial neural networks (ANN/NNs), a ML approach loosely inspired by the human brain (Ebert-Uphoff and Hilburn, 2020). Its primary goal is to learn from data by mapping inputs to outputs. The term "deep" refers to the multiple layers within the network, each adding a level of abstraction to the learned representation (Hanin, 2019). Compared to traditional ML methods, DL excels at uncovering complex patterns in high-dimensional data by structuring information into a hierarchical, layered representation (Lu et al., 2017).

A fundamental DL model, the multilayer perceptron (MLP), consists of an input layer, one or more hidden layers, and an output layer (Figure 12). Once trained, a neural network can process new inputs to generate predictions (Géron, 2017). It comprises interconnected neurons (or nodes) that pass signals via weighted connections called synapses. These neurons are arranged sequentially, transmitting signals in one direction — from the input layer, through hidden layers, to the output layer. Each neuron connects to one or more neurons in the preceding layer (LeCun et al., 2015).

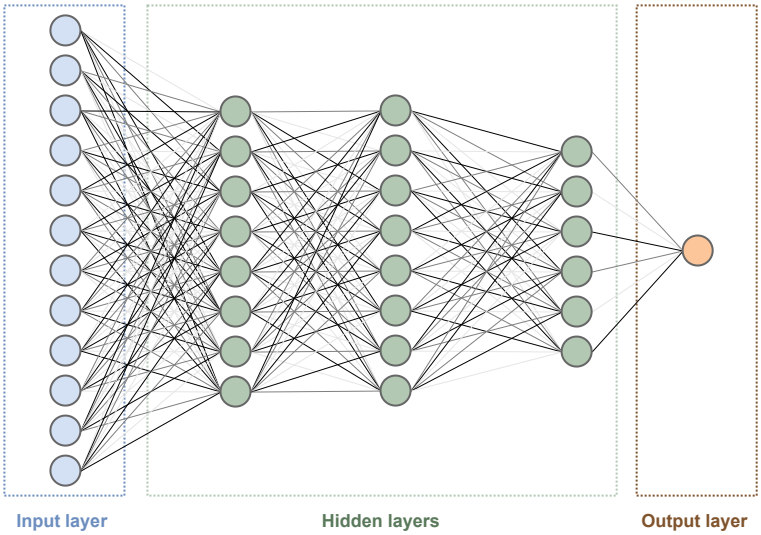

Figure 12: Schematic of a FCNN. The nodes of each layer are drawn as circles, the weights are drawn as straight lines between the nodes. The opacity of the lines encodes the weight values. The shown FCNN consists of an input layer with size 12, hidden layers of sizes 8, 8 and 6, and a single output neuron. Figure by the author based on Sarker (2021).

In an MLP, every node is linked to all nodes in the next layer with specific weights, making it a fully connected neural network (FCNN). Each neuron's state is determined by an activation value, a scalar computed through a nonlinear activation function. During training, the network optimises these weights to minimise prediction errors, guided by an objective cost (or loss) function (Sarker, 2021). Weight adjustments are performed iteratively using gradient descent, a process known as backpropagation, until the loss reaches a predefined threshold (LeCun et al., 1989). However, as FCNNs grow larger, their weight matrices scale with the number of neurons in adjacent layers. This increases training complexity and raises the risk of overfitting — where the model memorises training data rather than generalising to new inputs (Srivastava et al., 2014).

Convolutional neural networks

The success of DL in the 2010's - at the begin of the third wave of AI research - was largely driven by convolutional neural networks (CNNs) (Rawat and Wang, 2017). CNNs are specialised architectures designed to process grid-like data structures, such as images (LeCun et al., 1989). A typical CNN consists of three core layers: the convolutional layer, the pooling layer, and the fully connected layer (Figure 13).

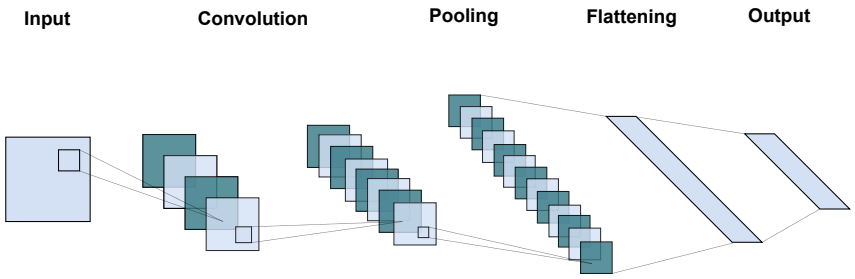

Figure 13: Schematic of a CNN. The size of the rectangles encodes the width and height of the image, the stacked rectangles refer to the color maps and feature maps, respectively. In the beginning of the CNN the small squares show the effect of filters acting at each position in the image individually through equivariant operations, e.g., convolution filters and local pooling operations. At one end of the CNN, a global pooling operation reduces the feature maps to a one-dimensional vector and a fully connected layer is applied to obtain the output. Figure by the author based on LeCun et al. (2015).

In the convolutional layer, the network applies a dot product between two matrices: one representing a small receptive field within the input and another containing a set of learnable parameters, known as kernel or filter. Instead of full matrix multiplications, these localised convolutions significantly reduce the number of parameters, as they share the same weights across the entire image (LeCun et al., 2015). As the kernel slides across the image, it produces an activation map which displays a two-dimensional representation that highlights relevant spatial features (Liu et al., 2016).

The pooling layer further refines this representation by summarising nearby outputs, reducing the spatial dimensions and, consequently, the computational complexity and number of weights (Ebert-Uphoff and Hilburn, 2020). Finally, in the fully connected layer, all neurons from the previous layer are connected via matrix multiplications, preserving learned representations for final predictions. Although CNNs require more computational power than MLPs, they significantly outperform them in image processing tasks (Rasp et al., 2020).

Neural networks for image segmentation

ML algorithms can be categorised into four main types: supervised, unsupervised, semi-supervised, and reinforcement learning (Sarker, 2021). This thesis focuses on supervised learning, where models are trained using labeled input-output pairs (Géron, 2017). While self-supervised learning approaches - a technique where the ML model trains itself to learn one part of the input from another part of the input - have gained popularity in recent years, supervised learning remains one of the most widely used ML techniques.

Supervised learning tasks fall into two primary categories: classification and regression. In classification tasks, the output is limited to a predefined set of discrete classes (binary or multi-class),

such as determining the cloud type as either convective or non-convective. In contrast, regression tasks predict continuous values, such as the rainfall intensity or CTH (Reichstein et al., 2019). While CNNs excel at classifying entire images, other tasks often require pixel-level predictions. To address this need, Ronneberger et al. (2015) developed the UNet, a convolutional network designed for fast and precise image segmentation. Its U-shaped architecture consists of two symmetrical paths: an encoder and a decoder (Figure 14). The encoder, composed of convolutional and pooling layers, captures the image's contextual features. The decoder, a symmetric expanding path, refines spatial details using transposed convolutions for precise localization (Lagerquist et al., 2021). As an image-to-image translation model, the UNet generates output images that typically match the dimensions of the input (Hilburn et al., 2020). Due to its flexibility and efficiency, the UNet has achieved widespread adoption across scientific disciplines (e.g., Chen et al., 2023; Grabowski and Petch, 2009; Agrawal et al., 2019). Since this thesis focuses on extracting image-based information, the following sections will explore a UNet-based approach which is described in detail in Sect. 3.2.

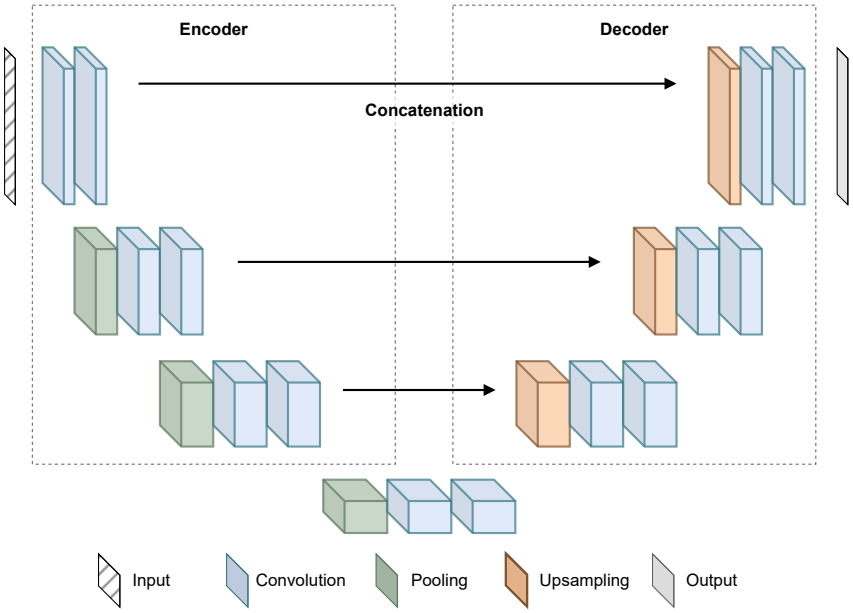

Figure 14: Schematic of a UNet. The size of the boxes encodes the height, width, and number of channels of the image. The input image is shown on the left side and fed into subsequent convolutional layers of the encoder side. Pooling operations reduce the height and width of the image while enhancing the number of channels. On the decoder side, the native size of the image is reconstructed using upsampling layers and convolutional layers. Skip connections are employed to concatenate the information from the encoder path. As a symmetric network, the UNet typically generates an output image with similar dimensions as the input image. Figure by the author based on Ronneberger et al. (2015).

2.3.2 Impact of machine learning applications

Thanks to their power and flexibility, ML applications have become invaluable across various scientific domains (Hanin and Rolnick, 2019; Lu et al., 2017). Today, NNs often drive groundbreaking advancements in tasks of image recognition, AI, and natural language processing (NLP).

One of the most impactful developments in recent years is OpenAI's chatbot, ChatGPT, which gained unprecedented attention in 2023 for its ability to generate human-like text and engage in natural conversations (Ray, 2023; Brown et al., 2020). ChatGPT is a generative AI which is built on the transformer architecture, leveraging the attention mechanism introduced by Vaswani et al. (2017). This mechanism rapidly became the foundation of modern NLP, powering large language models (LLMs) and generative AI systems like the Generative Pre-trained Transformer (GPT) series. Moreover, transformer networks have spread today across various domains and often dominate DL applications - in particular those dealing with large training datasets and extensive computational resources.

In image recognition, ML applications span tasks of classification (Krizhevsky, 2012), segmentation (Xie et al., 2021), and image generation (Goodfellow et al., 2020). Recent advancements have enabled high-resolution image creation and modification using text prompts (Rombach et al., 2022). Additionally, image recognition plays a critical role in diverse fields, from aiding medical diagnoses in public health (Litjens et al., 2017) to advancing autonomous driving technology (Feng et al., 2021).

2.3.3 Machine learning in atmospheric sciences

The growing computational power, increased data availability, and flexible frameworks have considerably boosted ML adoption in meteorology. The approaches offer valuable alternatives when direct measurements or observations fall short in quality or quantity (Boukabara et al., 2019).

In recent years, DL models have gained particular traction in atmospheric and climate science (Irrgang et al., 2021). Around Christmas 2022, Google DeepMind introduced GraphCast, a generative model for global medium-range weather forecasting (Lam et al., 2022). With a prediction time of under one minute, GraphCast quickly became a focal point in the field, outperforming the most accurate operational deterministic systems on 90 % of targets. The following year, the European Centre for Medium-Range Weather Forecasting (ECMWF) launched the Artificial Intelligence Forecasting System (AIFS), an ensemble of AI models excelling in short- to medium-range forecasting. AIFS is based on graph neural networks using a regression-based approach to predict future weather variables by leveraging ERA-5 reanalysis data (Keisler, 2022; Bi et al., 2023). With their performance and efficiency, those AI-driven models mark a potential revolution in traditional weather forecasting.

Beyond operational forecasting, ML has been applied across a wide range of meteorological challenges (McGovern et al., 2023). Use cases include improving bias correction in climate predictions (Beucler et al., 2024), integrating hybrid models with physical process simulations (Reichstein et al., 2019), and developing innovative climate process parameterisations (Beucler et al., 2021; Gentine et al., 2018). DL models have proven particularly effective in image-related meteorological tasks. They aid in detecting changes in satellite imagery time series (Pendergrass, 2020), improving severe thunderstorm predictions (Lagerquist et al., 2021), forecasting precipitation intensities (Sønderby et al., 2020), integrating LEO and GEO observations (Gorooh et al., 2023), and emulating radar imagery from passive sensors (Hilburn et al., 2020).

2.3.4 Error sources and limitations

ML approaches have proven highly effective in training specialized models for specific tasks. However, understanding the internal decision-making of a NN remains a challenge (Montavon et al., 2018). ML algorithms optimise an objective function based on the training dataset, minimising a loss function tied to the data distribution (Sarker, 2021). Yet, when faced with inputs outside this distribution, NNs rely on implicit assumptions, often with no inherent reason to generalise accurately (Srivastava et al., 2014). This can lead to unexpected outcomes, such as confidently incorrect classifications (out-of-distribution errors) (Beucler et al., 2024). Therefore, assessing prediction quality and associated uncertainty is crucial for evaluating the reliability of a NN (Boukabara et al., 2019; Eyring et al., 2024).

Several factors can degrade the model performance, including the selection of an appropriate loss function or training objective (Nguyen et al., 2014). Real-world datasets - e.g., from measurements - often contain biases, which are especially problematic in socially sensitive applications (Mehrabi et al., 2022). Additionally, data acquisition errors or sensor malfunctions can introduce noise or imbalances into training datasets. In particular in remote sensing applications, the quality and quantity of training data depend on the availability of annotated observations and pose a significant challenge to researchers (Sylolypavan et al., 2023; Rasp et al., 2020).

Many ML architectures rely on image datasets with relatively simple structures, such as MNIST (Deng, 2012) or ImageNet (Deng et al., 2009). In contrast, data from atmospheric sciences incorporates dynamic, thermodynamic, and radiative conditions, differing considerably from conventional training datasets. Observational data from passive remote sensing instruments typically spans multiple spectral wavelengths - which is fundamentally different from standard RGB (red, green, blue) images and limits the direct use of pre-trained ML models (Reichstein et al., 2019). As a result, researchers often have to develop custom frameworks tailored to their specific needs.

Lastly, the computational cost of ML should not be overlooked. Training large models requires significant energy, contributing to environmental concerns, including its potential role in accelerating fossil fuel exploration (Rolnick et al., 2022).

3 Results

3.1 Overview

This thesis builds upon three studies, one of which has been published in a peer-reviewed journal. The second and third study have been accepted for publication in a peer-reviewed journal. While the research questions were conceptualised in collaboration with my supervisors, I am the lead author of all studies presented in this thesis. I designed and implemented the ML model to predict the 3D cloud field, developed the algorithm for tracking convective clouds and their 3D properties, implemented the code for analysing convective organisation, produced the figures, interpreted the results, and wrote the original manuscripts and revised versions after peer-review. A more detailed overview of the author contributions are included in Sects. 3.2–3.4.

Figure 15 summarises the thesis following the three objectives presented in Sect. 1. Part **(a)** refers to the development of the ML model used to extrapolate a 3D cloud field from 2D satellite data (Section 3.2). The output of this study displays the input data for the study illustrated in Sect. 3.3, focusing on the connection between convective cores and the cloud life-cycle (Part **(b)**). Part **(c)** merges the output from Part **(a)** and **(b)** to investigate regional hotspots and spatio-temporal characteristics of convective organisation (Section 3.4).

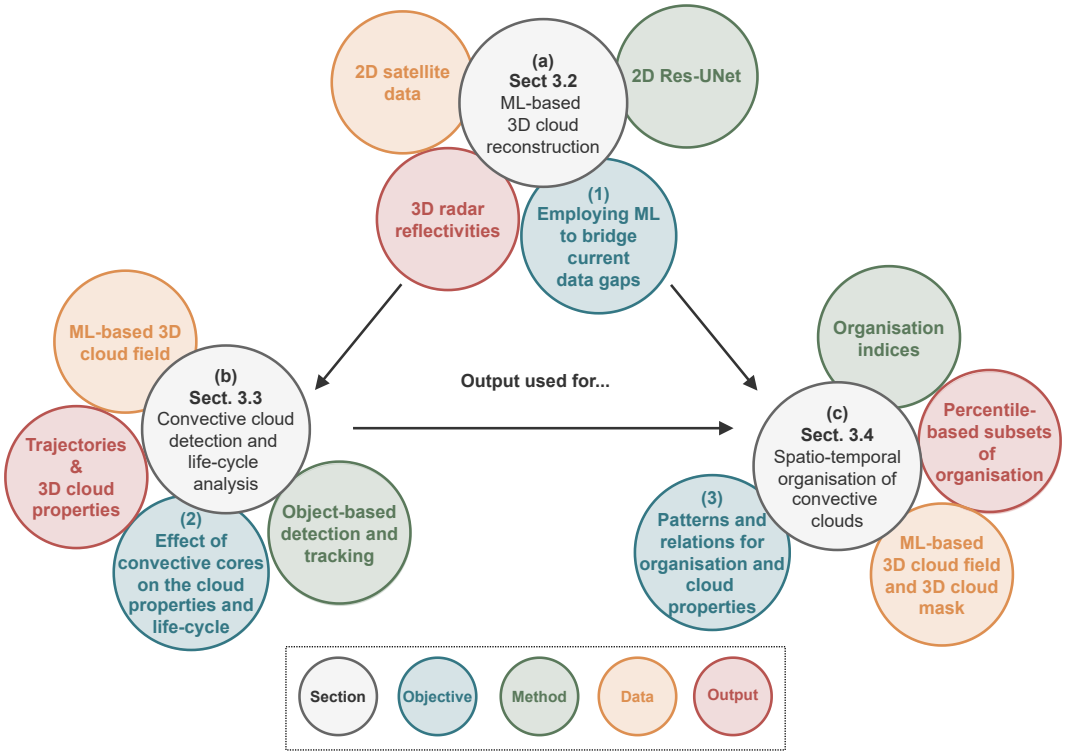

Figure 15: Summary of the content of this thesis, which includes **(a)** a ML-based extrapolation of a 3D cloud field from 2D data, **(b)** the detection and analysis of convective clouds and their cores, and **(c)** the identification of spatio-temporal patterns of convective organisation.

Section 3.2 presents the first study of this PhD project, which has been published in the peer-reviewed journal *Atmospheric Measurement Techniques* as a research article. In this study, a ResUNet was employed to reconstruct a contiguous 3D cloud field from 2D data, based on vertical cross sections of the CloudSat CPR and satellite imagery from MSG SEVIRI. The predictions were evaluated against the CTH derived from the operational CMSAF CLAAS-V002E1 dataset.

Section 3.3 includes a study which was accepted for publication in the peer-reviewed journal *Atmospheric Chemistry and Physics* as a research article. In this work, an adapted cloud tracking algorithm was implemented to detect convective clouds and their core regions in the ML-based 3D cloud field. Cloud objects are linked in time to follow their spatio-temporal evolution. The study focuses on tropical convection. It includes an analysis of cloud and core properties, and an evaluation of the connection between core regions and the cloud life-cycle.

Section 3.4 contains a study which was accepted for publication in the peer-reviewed journal *Atmospheric Chemistry and Physics* as a research article. This study examines the spatio-temporal distribution of convective organisation, approximated by three organisation indices. These indicators assess the spatial clustering of detected objected in the previously derived 3D cloud mask. A statistical analysis of the organisation indices unravels regional hotspots and their relationship to convective cloud and core properties in the tropics.

3.2 Machine learning-based reconstruction of 3D cloud fields

I am the lead author of this study, where I worked on the design of the ML model, including data processing and code implementations for a spatio-temporal matching scheme of different satellite data. Additionally, I trained the ML model, evaluated its performance, analysed the data, created figures, and wrote the manuscript. The co-authors contributed to implementing the code, interpreting the results, and proofreading the manuscript. The paper includes a section called *Author contributions* detailing individual contributions.

How to cite: Brüning, S., Niebler, S., and Tost, H.: Artificial intelligence (AI)-derived 3D cloud tomography from geostationary 2D satellite data, *Atmos. Meas. Tech.*, 17, 961–978, https://doi.org/10.5194/amt-17-961-2024, **2024**.

Submitted: 11 August 2023
Published: 9 February 2024

Atmos. Meas. Tech., 17, 961–978, 2024
https://doi.org/10.5194/amt-17-961-2024

Atmospheric Measurement Techniques
Open Access
EGU

Artificial intelligence (AI)-derived 3D cloud tomography from geostationary 2D satellite data

Sarah Brüning[1], **Stefan Niebler**[2], **and Holger Tost**[1]

[1]Institute for Physics of the Atmosphere, Johannes Gutenberg University Mainz, Johann-Joachim-Becher-Weg 21, Mainz 55128, Rhineland-Palatinate, Germany
[2]Institute of Computer Science, Johannes Gutenberg University Mainz, Staudingerweg 9, Mainz 55128, Rhineland-Palatinate, Germany

Correspondence: Sarah Brüning (sbruenin@uni-mainz.de) and Holger Tost (tosth@uni-mainz.de)

Received: 11 August 2023 – Discussion started: 16 August 2023
Revised: 13 December 2023 – Accepted: 22 December 2023 – Published: 9 February 2024

Abstract. Satellite instruments provide high-temporal-resolution data on a global scale, but extracting 3D information from current instruments remains a challenge. Most observational data are two-dimensional (2D), offering either cloud top information or vertical profiles. We trained a neural network (Res-UNet) to merge high-resolution satellite images from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) with 2D CloudSat radar reflectivities to generate 3D cloud structures. The Res-UNet extrapolates the 2D reflectivities across the full disk of MSG SEVIRI, enabling a reconstruction of the cloud intensity, height, and shape in three dimensions. The imbalance between cloudy and clear-sky CloudSat profiles results in an overestimation of cloud-free pixels. Our root mean square error (RMSE) accounts for 2.99 dBZ. This corresponds to 6.6 % error on a reflectivity scale between −25 and 20 dBZ. While the model aligns well with CloudSat data, it simplifies multi-level and mesoscale clouds in particular. Despite these limitations, the results can bridge data gaps and support research in climate science such as the analysis of deep convection over time and space.

1 Introduction

Clouds and their interdependent feedback mechanisms have been a source of uncertainty in Earth system models for decades. Their influence on atmospheric gases and general circulation patterns is evident (Rasp et al., 2018; Shepherd, 2014; Bony et al., 2015). In a world affected by climate change, we require an accurate representation of cloud dynamics today more than ever (Norris et al., 2016; Stevens and Bony, 2013; Vial et al., 2013).

In recent years, observational data from remote sensing instruments have been used to investigate cloud properties on multiple scales (Jeppesen et al., 2019). Nevertheless, techniques to detect three-dimensional (3D) cloud structures on a large scale are not yet established (Bocquet et al., 2015). Observations from passive sensors on geostationary satellites have a high spatiotemporal coverage, but they are limited to monitoring the uppermost atmospheric layer in 2D (Noh et al., 2022). By using the satellite's specificity at different wavelengths (Thies and Bendix, 2011) and subjective labeling or fixed thresholds (Platnick et al., 2017), we can estimate cloud physical properties like the cloud optical thickness (Henken et al., 2011) or the effective radius (Chen et al., 2020). In contrast, active radar penetrates the cloud top and delivers information on the subjacent reflectivity distribution (Barker et al., 2011). The radar receives detailed information on the cloud column along a 2D cross section with a high ground resolution and constant sun illumination. Due to its sun-synchronous orbit, it observes the same spot at the same local time. Compared to geostationary satellites, the active radar does not provide a continuous spatial and temporal coverage (Wang et al., 2023). Passive sensors can be used to deliver an approximation of the cloud vertical column, but their information density is reduced compared to active sensors (Noh et al., 2022). Combining data sources can fill current data gaps (Amato et al., 2020; Steiner et al., 1995). The combined use of different instruments has been investigated

Published by Copernicus Publications on behalf of the European Geosciences Union.

before. This research comprises the usage of statistical algorithms (Miller et al., 2014; Seiz and Davies, 2006; Noh et al., 2022), the integration of radiative transfer approaches (Forster et al., 2021; Zhang et al., 2012), or the derivation of the multi-angle geometry of neighboring clouds (Barker et al., 2011; Ham et al., 2015) to reconstruct the cloud vertical column.

Emerging facilitators of data availability, like open-data policies and improved technological standards, enable effective processing of memory-consuming data (Irrgang et al., 2021; Rasp et al., 2018). This development promotes a further integration of computer science methods in climate science (Jeppesen et al., 2019; Liu et al., 2016). Ever-growing quantities of data surpass the capability of the human mind to extract explainable information efficiently (Lee et al., 2021; Karpatne et al., 2019). Here, the usage of artificial intelligence (AI) has been assigned a primary role (Runge et al., 2019). Cloud properties have been analyzed before by machine-learning (ML) algorithms (Reichstein et al., 2019; Marais et al., 2020). The recent technological advances enable unprecedented operations (Amato et al., 2020). Deep-learning (DL)-based networks are suitable for identifying spatial, spectral, and temporal patterns in big data (Jeppesen et al., 2019; Hilburn et al., 2020). In contrast to traditional ML frameworks, they do not require manual feature engineering (Le Goff et al., 2017). Adapting DL frameworks to applications in climate science offers new perspectives for a gain in knowledge (Rolnick et al., 2022; Jones, 2017).

So far, cloud properties have been investigated by DL algorithms in various applications. These comprise the detection (Drönner et al., 2018) and segmentation of cloud fields (Jeppesen et al., 2019; Lee et al., 2021; Le Goff et al., 2017; Tarrio et al., 2020; Cintineo et al., 2020) or the classification of distinct cloud types from meteorological satellites and aerial imagery (Marais et al., 2020; Wang et al., 2023). Zantedeschi et al. (2022) used a neural network to bring together information from an active radar and high-resolution satellite images to reconstruct cloud labels. Regressive models were used to predict rain rates (Han et al., 2022) or convective onset (Pan et al., 2021) for an improved weather forecast. These studies are often limited to reflecting horizontal processes within the cloud field. Current studies by Hilburn et al. (2020) and Leinonen et al. (2019) use AI techniques such as convolutional neural networks (CNNs) and conditional generative adversarial networks (CGANs) to address this issue. They reconstruct the 1D cloud column (Hilburn et al., 2020) or the 2D cross section of the input data (Wang et al., 2023). To the best of our knowledge, no extrapolation of 2D radar data to a large-scale 3D perspective was conducted before (Wang et al., 2023; Dubovik et al., 2021). Clouds move within a 3D space. This limits the prediction of multi-layer and mesoscale events by a 1D or 2D pixel-wise reconstruction (Hilburn et al., 2020). Models that do not consider the spatial coherence between pixels fail to reconstruct comprehensive cloud structures (Hu et al., 2021). Image segmentation approaches like the UNet (Ronneberger et al., 2015; Jiao et al., 2020; Wieland et al., 2019) may reconstruct the ground truth data more adequately. They can be used to provide the indicators for predicting clouds in 3D with their adjacent boundaries, shadow locations, and geometries (Wang et al., 2023). This can lead to a more realistic representation of the predicted clouds (Jiao et al., 2020).

In this study, we employ a modified Res-UNet (Diakogiannis et al., 2020; Hu et al., 2021) to integrate 2D data from active (polar-orbiting satellite, radar) and passive (geostationary satellite, spectrometer) instruments to reconstruct a 3D cloud field. Previous studies focused on reconstructing the 1D cloud column or 2D cross section. In contrast, our approach utilizes a DL framework to predict the radar reflectivity, not only along the radar cross section, but also across the entire satellite full disk (FD). We use the radar height levels to extend 2D satellite channels to a 3D perspective. The goal is to establish a spatiotemporally consistent cloud tomography solely based on observational data. Predicted reflectivities can enhance the availability of 3D resolved cloud structures, particularly in regions with limited data.

2 Methods

2.1 Data overview

Our approach uses observational data from two different remote sensing sensors to predict a 3D cloud tomography. The input data for the neural network originate from a geostationary satellite. We use data from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) instrument on the Meteosat Second Generation (MSG) satellite (EUMETSAT Data Services, 2023). This sensor observes the Earth from a height of 36 000 km and provides 2D satellite images at a high spatial and temporal resolution. The ground truth of the study is derived from an active radar on board the CloudSat satellite which moves in a sun-synchronous orbit (CloudSat Data Processing Center, 2023). The 2D profiles along the track contain information on the cloud reflectivity. In our study, we feed the MSG SEVIRI data into a neural network to reconstruct the CloudSat radar reflectivity and extrapolate the 2D profiles to a 3D perspective.

2.1.1 Geostationary satellite images

Satellite images from the MSG SEVIRI instrument display the input for the network (later referred to as "imager data") (Schmetz et al., 2002). Observing the Earth's surface in intervals of 15 min and with a spatial resolution of 3 km at nadir, MSG SEVIRI provides information in 12 channels centered within wavelengths from 0.6 to 132 µm (Benas et al., 2017). Depending on the wavelength and daytime of retrieval, the channels are sensitive to reflected solar radiation or surface

Table 1. Overview of the MSG SEVIRI channels (Schmetz et al., 2002). "n/a" stands for "not available".

Channel	Center (µm)	Range (µm)	Type
VIS0.6	0.635	0.56–0.71	Solar reflective
VIS0.8	0.81	0.74–0.88	Solar reflective
NIR1.6	1.6	1.5–1.78	Solar reflective
IR3.9	3.92	3.48–4.36	Both
WV6.2	6.25	5.35–7.15	Thermal IR
WV7.3	7.35	6.85–7.85	Thermal IR
IR8.7	8.70	8.30–9.10	Thermal IR
IR9.7	9.66	9.38–9.94	Thermal IR
IR10.8	10.8	9.80–11.80	Thermal IR
IR12.0	12.0	11.00–13.00	Thermal IR
IR 13.4	13.4	12.40–14.40	Thermal IR
HRV	n/a	0.5–0.9	Solar reflective

emissions (Table 1). They can be applied to approximate cloud physical properties (Sieglaff et al., 2013). Our approach uses 11 channels. The high-resolution visible (HRV) channel is excluded due to its different resolution and uncertain added value. Three of the channels are sensitive to solar radiation, which restricts us to using only daytime data. We reformat all imager data onto a spatial grid with geographic coordinates, employing the global reference system WGS84 (Drönner et al., 2018). Each pixel has a resolution of 0.03° in both width (W) and height (H). To account for diminishing accuracy from the Equator to the poles, we exclude the areas near the sensor boundaries (Bedka et al., 2010). The designated area of interest (AOI) extends 60° in all directions, marking the boundaries of the new FD.

2.1.2 Radar data

Within the CloudSat (CS) GEOPROF-2B product, a nadir-looking 94 GHz cloud profiling radar (CPR) delivers information on the cloud reflectivity on the logarithmic dBZ (decibel relative to Z) scale (later referred to as "radar data") (Stephens et al., 2008). The radar receives a 2D cross section of the cloud column with a horizontal resolution of 1.1 km. The vertical dimension (Z) comprises 125 height levels with a bin size of 240 m (Guillaume et al., 2018). From the ground surface to the lower stratosphere, the vertical extent covers 30 km. We use the reflectivity transects as the ground truth to train and evaluate the model. In the subsequent steps, we adjust the height levels of the radar. The lower altitudes, specifically those between 0 and 3 km, are influenced by the topography and a radar signal weakening due to attenuation (Marchand et al., 2008). To enhance the model performance, we omit the 10 lowest height levels. Since we notice a significant imbalance between clear-sky and cloudy pixels, we exclude the predominantly cloud-free areas within the upper 25 height levels (Stephens et al., 2008). The final Z dimension encompasses 90 height levels ranging from 2.4 to 24 km. We note that, due to the sun-synchronous orbit of CloudSat, it has a reduced ability to account for diurnal variations within specific regions of the AOI (Stephens et al., 2008).

2.1.3 Matching scheme

We obtain training data for our study by aligning MSG SEVIRI scenes with CloudSat radar data as shown in Fig. 1a. To match the datasets, we compare their timestamps and locations. If the radar coordinates fall within the AOI, we determine the flight direction to identify whether CloudSat circles the Earth in ascending or descending orbit. We then extract images of 128×128 ($H \times W$) pixels from each MSG SEVIRI channel along the radar coordinates using a moving-window approach with a 50 % overlap between image–profile pairs (Denby, 2020; Jeppesen et al., 2019).

We prepare the matched image–profile pairs for further processing. To do this, we combine the 11 MSG SEVIRI channels into a single 3D array with dimensions $11 \times 128 \times 128$ [$C \times H \times W$] pixels. CloudSat flies across a horizontal transect within the satellite scene. It has a higher native resolution than MSG SEVIRI. To align the datasets, we downsample the radar pixels by aggregating them based on the local maximum reflectivity. This adjusts the CloudSat pixels to the MSG SEVIRI resolution of 0.03° but leads to some loss of sharp contrast in radar pixels (Jordahl et al., 2020). We standardize the data shape by transforming the 2D cross section into a sparse 3D array of $125 \times 128 \times 128$ [$Z \times H \times W$] pixels, representing reflectivities along the cross section. After downsampling, the transect becomes 1 pixel wide. We label pixels outside the transect as missing values to maintain the CloudSat data location during training. We use these pixel indices to compute the loss between the CloudSat data and the predicted cross section and to evaluate the model performance.

2.1.4 Data processing

Before training the model, we process the extracted image–profile pairs. We utilize a full year of data (2017) to incorporate seasonal variations into the modeling process. We split the 30 000 matched image–profile pairs, with 75 % (January to September) used for training and 25 % (October to December) for validation. Our test set is derived from data in May 2016, from which the matching algorithm extracts 1500 image–profile pairs. We impute missing data in the 3D MSG SEVIRI array by an interpolation of neighboring pixels (Troyanskaya et al., 2001). Afterwards, data from each satellite channel x were normalized between [0, 1] by

$$x' = \frac{x - \mu}{\sigma} \tag{1}$$

using the arithmetic mean μ and standard deviation σ of the training data (Leinonen et al., 2019). As described in Sect. 2.1.2, we reduce the height levels of the CloudSat profile from 125 to 90 (Fig. 1b). We use the CloudSat quality in-

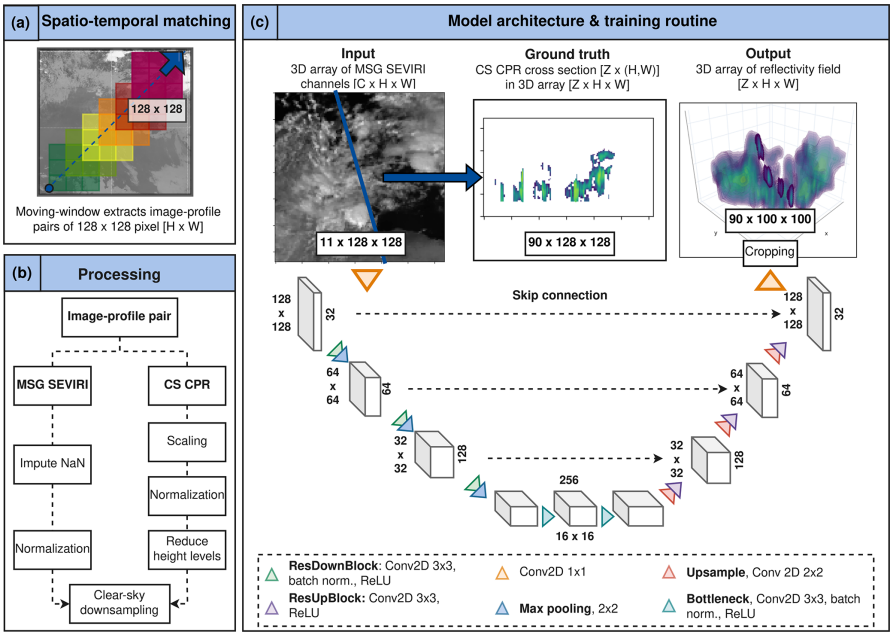

Figure 1. Workflow of the study. Panel **(a)** shows the moving-window approach used for matching the radar and the imager data. Steps needed for the processing of both datasets are depicted in panel **(b)**. In panel **(c)**, the architecture of the proposed Res-UNet is illustrated. The upper row shows an example of the input data, ground truth (with reduced 90 height levels and full transparency for values $\leqq -25$ dBZ), and output, respectively. The location of the radar transect within the 3D output image is pictured with full opacity. The numbers alongside the boxes in the architecture sketch refer to the feature channels (right) and image sizes (left) at the given model depth.

dex to identify noisy pixels. Pixels with a quality index lower than 6 were set to a background value of -25 dBZ to reduce noise (Marchand et al., 2008). All radar reflectivity values Z_{dB} were normalized between $[-1, 1]$ as follows:

$$Z'_{\mathrm{dB}} = 2\frac{Z_{\mathrm{dB}} + 35\,\mathrm{dB}}{55\,\mathrm{dB}} - 1, \tag{2}$$

where the maximum and minimum reflectivities are between $[-35, 20]$ (Stephens et al., 2008; Leinonen et al., 2019). The CloudSat data are highly skewed towards clear-sky samples. We limit the percentage of cloud-free profiles to 10 % to tackle this imbalance (Jeppesen et al., 2019).

2.2 Model architecture and training

Neural networks can capture highly complex relationships between input and output data (Lee et al., 2021). The Res-UNet used in this study displays a modified framework designed for remote sensing data (Dixit et al., 2021). Additional residual connections and continuous pooling operations aim to reduce the dependence of the network on the input's location (Diakogiannis et al., 2020). Former studies using the Res-UNet dealt with the classification of tree species (Cao and Zhang, 2020) or the prediction of precipitation (Zhang et al., 2023). The obtained results emphasize the ability of the Res-UNet to adequately address the importance of the spatial coherence in environmental research (Marais et al., 2020). In this study, we derive the cloud reflectivities (dBZ) from the satellite channels by a regression task (Hilburn et al., 2020; Zhang et al., 2023).

As introduced by Ronneberger et al. (2015), the UNet and its modifications provide an almost symmetrical architecture. The network architecture of the Res-UNet is shown in Fig. 1c. The parameters of the network are listed in Table A1 in Appendix A. Each box represents the layer sizes on the encoder and decoder sides. On the right-hand side of each box, the filter size is given. The respective height and width are given on the left-hand side.

The Res-UNet consists of six residual blocks, each including two 2D convolutions (3×3 kernel, stride 1) and rectified linear unit (ReLU) activation (Diakogiannis et al., 2020). On the encoder side, we add batch normalization. The output is merged with a skip connection that consists of one 2D convolution (3×3 kernel, stride 1) and a batch normalization. Adding the skip connection and the convolutional layer represents the output of a residual block.

We increase the channel dimension of the initial imager data from $11 \times 128 \times 128$ pixels with a 1×1 2D convolution to a feature map of size $32 \times 128 \times 128$. In the encoder branch, we then employ a sequence of three residual blocks with doubling filter sizes, each followed by a 2×2 maximum

pooling layer (Lee et al., 2021). We subsequently reduce the feature map size to $256 \times 16 \times 16$ pixels in the bottleneck layer. Here, we apply two 2D convolution layers, followed by batch normalization and ReLU activation.

The decoder side features three residual blocks, each with an upsampling layer (2D convolution, 2×2 kernel, stride 2) and a corresponding skip connection from the encoder. After upsampling, we apply a residual block with 2D convolution (3×3 kernel, stride 1) and ReLU activation, doubling the spatial extent to match the skip connection while halving the channel dimension (Li et al., 2018). The final 1×1 convolution maps the output to $90 \times 128 \times 128$ pixels, representing the 90 height levels of the radar cross section (Jeppesen et al., 2019). We remove the border pixels of the output, resulting in a final radar reflectivity output of $90 \times 100 \times 100$ pixels ($C \times W \times H$). Predicted reflectivities are scaled between -35 and 20 dBZ, with values below -25 dBZ considered cloud-free (Leinonen et al., 2019).

We conducted the training for 50 epochs with a batch size of 4 and a weight decay of 0.00001 (see Table A1 in Appendix A). We have 1 893 328 total trainable parameters. The estimated total size of the model is 194.27 MB (see Table B1 in Appendix B). We use the adaptive moment estimation (ADAM) method for model optimization due to its fast convergence rate (Kingma and Ba, 2014). The learning rate is initially set to 0.001 (see Table A1 in Appendix A). It is reduced by a learning rate scheduler during the training process when reaching a plateau. To enhance the number of training data, we give all input data a chance of 25 % of being rotated by 90° (Jeppesen et al., 2019). These flipped images are perceived as new samples. The goal is to increase the model invariance to the orientation of the radar cross section.

2.3 Evaluation

2.3.1 Analyzing and comparing the model performance

The model performance is quantified during training (loss function) and is evaluated afterwards by calculating the root mean square error (RMSE) (see Table A1 in Appendix A). The RMSE is equally able to penalize misses and false alarms (Lee et al., 2021). As described in Sect. 2.1.3, we preserve the pixel indices of the CloudSat cross section within each image–profile pair during training. We use the locations of these pixels to filter the observed and predicted transects. The RMSE is calculated along the filtered cross sections. Since it is only evaluated on a small subset of 10 % of all the pixels, we have a sparse regression task (Wang et al., 2020). We cannot quantify the model performance on the full 3D prediction of the cloud field.

The results of the Res-UNet are compared against two competitive methods (Drönner et al., 2018). First, we predict the radar reflectivity by an ordinary least squares model with multiple regression outputs (OLS). The 11 satellite channels were used as independent predictor variables. The output is a 1D cloud column. Second, a random-forest (RF) regression is applied (Breiman, 2001). The RF is a supervised ML algorithm suitable when working with environmental datasets in the natural sciences (Boulesteix et al., 2012). Its feasibility for complex meteorological data was investigated before. For example, McCandless and Jiménez (2020) used a RF regression to detect clouds. Our study uses a setup with 100 trees, each choosing a random subset of satellite channels to predict the reflectivity along a 1D cloud column. We use the same training, validation, and test split as for the Res-UNet. For each image–profile pair, we filter the 3D array to locate the radar cross section. This transect is separated into 1D cloud columns. For every pixel along the cross section, we receive a ground truth in the form of 90×1 $[Z \times (H, W)]$. The 3D array containing the satellite channels was filtered by the radar profile location and was divided into images of size 11×1 $[C, (H, W)]$. The OLS and RF map the imager data to an output size of 90×1 pixels $[Z \times (H, W)]$. We calculate the RMSE between the observed and predicted cloud columns and scale the output between -35 and 20 dBZ. We reconstruct the 2D transect by the preserved index of each pixel of the cross section. These profiles are compared to the output of the Res-UNet.

2.3.2 Merging 3D reflectivities on the FD

We predict the radar reflectivity for each MSG SEVIRI file in the test dataset (May 2016) using the trained Res-UNet. The result is a contiguous 3D cloud tomography for every 15 min time step. The MSG SEVIRI FD covers an extent of 2400×2400 pixels. For the FD prediction, we divide the FD into overlapping subsets of 128×128 pixels. These subsets are processed and fed into the network. The output is a 3D reflectivity image of $90 \times 100 \times 100$ pixels $[Z \times H \times W]$, which equals 2.5° on the MSG SEVIRI grid. We merge the tiles to cover the whole satellite AOI. Between the tiles, there is no overlap. The goal is to evaluate the network's ability to extrapolate a large-scale cloud field from single tiles.

2.3.3 Computing the cloud top properties

To the best of our knowledge, there exist no comparable datasets on the 3D cloud tomography in this study. Instead of a quantitative evaluation of the reflectivity, we evaluate the predictions based on their ability to derive the cloud top height (CTH) (Wang et al., 2023). We use the FD predictions for the test dataset (May 2016) for the computation. The CTH is defined as the distance between the ground surface and the uppermost cloud layer for every 1D vertical column (Huo et al., 2020). This calculation requires conversion of the height levels to a kilometer scale. We use a fixed threshold of -15 dBZ to differentiate a cloudy pixel from a clear-sky pixel (Marchand et al., 2008). The result is a binary classification for each pixel in the 3D cloud field. In this dataset, we extract the CTH as the cloud top signal in each 1D cloud col-

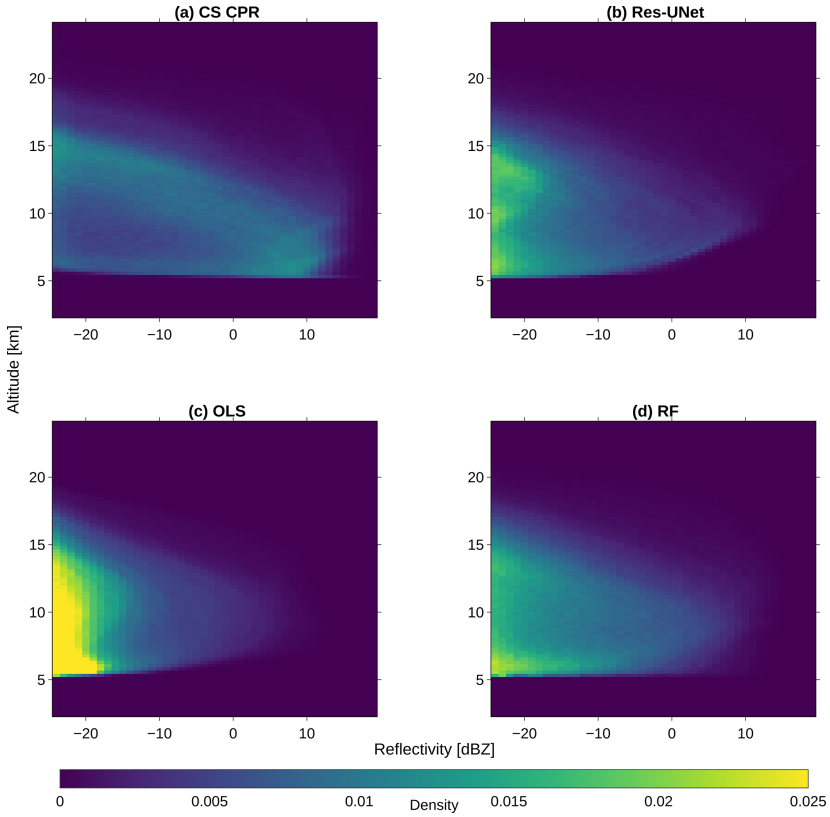

Figure 2. Height-dependent reflectivity distribution for every height bin between 2.4 and 24 km for the CloudSat data (CS CPR) **(a)**, the Res-UNet **(b)**, the ordinary least squares model (OLS) **(c)**, and the random-forest (RF) regression **(d)** ($n = 1500$).

umn of the FD. We aggregate the results on a monthly scale and compare the predicted CTH to the operational product CLAAS-V002E1 (CLoud property dAtAset using SEVIRI, Edition 2) (Finkensieper et al., 2020). It is based on the MSG SEVIRI channels and additional model data and provides information on the macrophysical and microphysical cloud properties. We use a monthly aggregate with a resolution of 0.05° on the MSG SEVIRI FD.

3 Results

3.1 Evaluating the reflectivity distribution

We analyze the ability of the three models (Res-UNet, OLS, and RF) to reconstruct the cloud vertical distribution for the test dataset in May 2016 (Sect. 2.3.1). The OLS and RF predict a 1D column, whereas the output of the Res-UNet comprises a 3D image of the cloud field. We filter all outputs by the preserved location of the radar cross sections to derive the original 2D transect. At first, we compute the height-dependent reflectivity distribution of the CloudSat data and the three models. Due to the applied quality flag (Sect. 2.1.4), we have few observations below 5 km height (Fig. 2a). This leads to a shift in low height levels. The models overestimate cloud-free values below −25 dBZ. The CloudSat reflectivities have a peak at 0–10 dBZ between 5 and 7 km. A second, weaker peak is observed between 12 and 15 km for reflectivities < 0 dBZ. All the predictions underestimate the first peak > 0 dBZ. Instead, they overestimate the occurrence of reflectivities < −20 dBZ (Fig. 2b–d). The OLS shows an especially high shift towards low reflectivities. The Res-UNet predicts low reflectivities < −20 dBZ along all the height levels between 5 and 15 km, whereas we observe a distinct peak at 5 km for the RF.

We analyze the difference between the observed and predicted reflectivities by a 2D joint distribution plot. For this purpose, we calculate the density distribution of the reflectivity between 2.4 and 24 km. Here, we use a bin size of 1 dBZ and 240 m height, respectively (Steiner et al., 1995). All the distributions are calculated on the test dataset and are normalized by the distribution size ($n = 1500$). Predictions differ from the original radar data, especially for values > 0 dBZ and at low altitudes (Fig. 3). In the joint plot,

the highest agreement appears in the form of a curved line between low reflectivities > 15 km and high reflectivities at 7 km. The results indicate an overestimation of high reflectivities and an underestimation of low reflectivities, especially for low-level clouds. Since we observe few clouds at high altitudes (Fig. 2a), the distribution differences become smaller above 15 km. The joint plot shows a similar distribution for the Res-UNet and the RF, whereas the error of the Res-UNet is slightly lower for reflectivities between −15 and 0 dBZ (Fig. 3a and c). We observe few predictions > 0 dBZ and a strong overestimation of reflectivities < −20 dBZ for the OLS (Fig. 3b).

3.2 Height-dependent model performance

We analyze the model error (RMSE) along the vertical cloud column. For all the models, we calculate the mean RMSE on the test dataset between 2.4 and 24 km. The results show an overall lower RMSE for the Res-UNet than for the OLS and RF (Fig. 4). The mean RMSE varies between 2.99 dBZ (Res-UNet), 4.1 dBZ (RF), and 4.58 dBZ (OLS). On a dBZ scale between −25 and 20 dBZ, this is equivalent to errors of 10.1 % (OLS), 9.1 % (RF), or 6.6 % (Res-UNet). Between 2.4 and 5 km, the RMSE is 0. This is due to the lack of CloudSat observations after filtering noisy pixels (Fig. 2a). Between 5 and 7 km, the RMSE increases to up to 8 dBZ for the Res-UNet, 10 dBZ for the RF, and 12 dBZ for the OLS (Fig. 4). At higher altitudes, the performance of all the models improves. The RMSE decreases to 4 dBZ (5.7 dBZ, 6 dBZ) for the Res-UNet (RF, OLS) at 15 km and reaches its minimum at 22 km (24 km for OLS and RF). Above 15 km, we have few CloudSat observations > 15 dBZ (Fig. 2a). We observe a lower model error (Fig. 4) and reduced the difference between the distributions (Fig. 3) at these height levels for all three models. The improved performance can be traced back to the superior number of background reflectivities or the presence of more uniform clouds, like extended tropical cirrus. Over all the height levels, the Res-UNet has the lowest RMSE of the three models. Compared to the OLS (RF), the mean RMSE of the Res-UNet is reduced by 34,8 % (27,1 %).

Figure 5 shows the predicted and observed reflectivities along the radar transect for four randomly chosen samples. For all the models, the reconstructed cloud signal is predicted at the right horizontal location along the cross section. Clear-sky situations of −25 dBZ are recognized without noise. The cross sections in Fig. 5a are created using processed CloudSat reflectivities with a resolution of 0.03°. Although the radar pixels lose some sharp contrasts after the downsampling, we observe a higher blurriness for the predictions. The edges of individual clouds smear out for all three models. Even though all the transects were labeled “cloudy”, we see a high percentage of background pixels.

For the Res-UNet, we observe a RMSE between 3.3 and 8.2 dBZ. The overall shape and increased intensification towards the cloud’s core follow the radar, even though edges are blurred, and peak reflectivities remain underestimated (Fig. 5b). This issue is reflected within the reflectivity distribution of the DL model (Fig. 2b). While the Res-UNet accurately identifies single-layer clouds, it misses sharp edges of multi-layer clouds, especially at mid altitudes. Clouds over multiple height levels are blurry and show a reduced small-scale accuracy in the vertical dimension (Fig. 5III). The lower height levels of multi-layer clouds are only partly represented (Fig. 5II and IV). Instead, we observe a simplification of these cloud layers.

For the OLS and the RF, the underestimation of the cloud core reflectivities resembles the Res-UNet (Fig. 5II and III). All four examples show a higher RMSE for the OLS and RF than for the Res-UNet (Fig. 5c and d). The difference varies between 0.1 (I) and 2.7 (IV) dBZ. While the error is predominantly similar for all three models, the shape of the predicted clouds differs (I, III). The OLS (RF) fails to accurately reconstruct the vertical extent in all the transects. Instead, the reflectivity is uniform along the cloud column. We see a continuous cloud signal between 5 and 15 km (Fig. 5c). In contrast, the Res-UNet predicts the vertical variability more precisely (Fig. 5b). While the 2D profiles of the Res-UNet are smooth, the RF and OLS lead to a fragmented structure with a high value variability between the single pixels of the transect (I, IV). The examples show an inaccurate reconstruction of shallow clouds and multi-layer clouds for the OLS and RF.

3.3 Geographic analysis of the 3D cloud tomography

With the trained Res-UNet, we predict clouds on the MSG SEVIRI AOI. Since the network was trained using visible-spectrum (VIS) channels, we cannot provide an accurate representation of the nocturnal cloud field. An exemplary 3D cloud tomography is predicted for 6 May 2016 at 13:00 UTC. For that purpose, the satellite scene was divided into small subsets of overlapping 128 × 128 pixel images as described in Sect. 2.3.2. After feeding each subset into the network, the output tiles of 90 × 100 × 100 pixels were merged into a FD scene of 90 × 2400 × 2400 pixels for the whole AOI (Fig. 6a).

The results contain a 3D cloud field along 90 height bins between 2.4 and 24 km. As shown in Fig. 6a, the top view of the maximum reflectivity per cloud column demonstrates the absence of hard borders between single prediction tiles. Even though CloudSat data are only available at the radar transects, we can extrapolate smooth cloud structures on the FD. The example tiles (b)–(d) show a fluid transition between the edges of single prediction tiles (Fig. 6). Each example spans a horizontal extent of > 2.5° (100 × 100 pixels) to demonstrate the absence of artifacts between the tiles. High-reaching convective complexes (b) and large-scale structures (c, d) are extrapolated at the FD scale regardless of their location. Even though the overall reflectivity is underestimated, low-level and multi-layer clouds are displayed as contiguous entities.

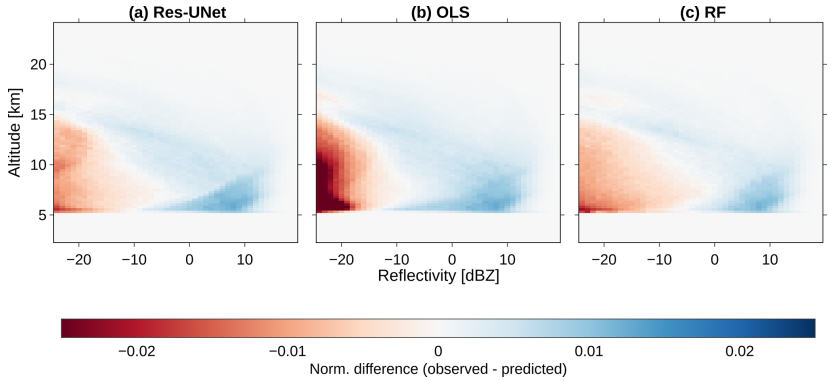

Figure 3. Joint plot of the normalized difference between the observed and predicted reflectivities on the test dataset ($n = 1500$). For each height bin between 2.4 and 24 km, we compare the CS CPR to the distribution of the Res-UNet **(a)**, OLS **(b)**, and RF regression **(c)**.

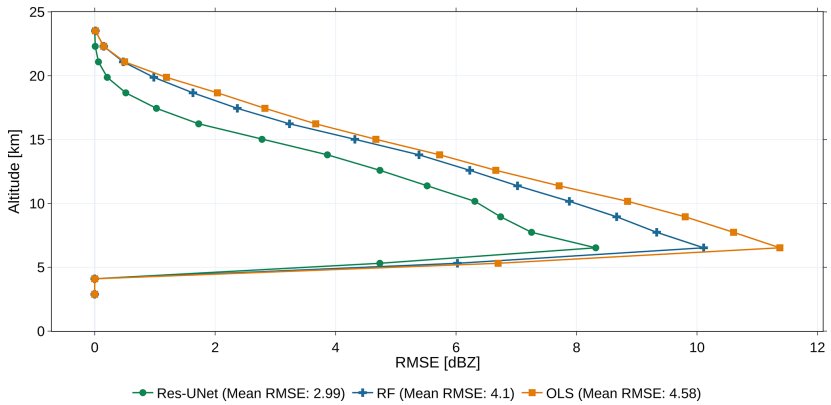

Figure 4. Height-dependent RMSE for every height bin and the mean RMSE for all the models calculated on the test dataset ($n = 1500$).

We visualize the mean RMSE between 60° N and 60° S to investigate zonal variations for the test dataset. The geographic analysis is used to evaluate the reliability of the 3D cloud tomography. The RMSE shows a high-latitudinal variability. At 30–50° N, we observe the highest RMSEs of 6–7 dBZ (Fig. 7a). The RMSE at mid latitudes in the Southern Hemisphere is lower than in the Northern Hemisphere. Nevertheless, the lowest RMSE is achieved in the tropics between 20° N and 20° S. We analyze the RMSE in relation to the number of image–profile pairs originating in the matching scheme in Sect. 2.1.3. Most image–profile pairs are located around the Equator and at the mid latitudes (Fig. 7b). Few pairs are matched around 10° N and 30° S. Regions at the mid latitudes have the highest RMSE and the highest number of observations. In the tropics, the RMSE is lower. Here, we obtain a high number of image–profile pairs from the matching scheme. The predicted cloud field shows high geographic variability of the RMSE. We observe a higher RMSE for the Northern Hemisphere than for the Southern Hemisphere. Clouds in the subtropics are more accurately represented than clouds at high latitudes. At the same time, we lack observations here. The analysis emphasizes the importance of the geographic location for the model predictions as well as the influence of the CloudSat orbit.

3.4 Comparison of the predicted CTH

To evaluate the quality of the Res-UNet predictions, we compare the reflectivity distribution between CloudSat and the Res-UNet predictions. In a second step, we calculate the CTH on the test dataset. The reflectivities in Fig. 8a are provided on a logarithmic scale due to the high proportion of cloud-free pixels around −25 dBZ (Fig. 8a and c). Values < -15 dBZ are visualized with a grey background. They lie below the threshold used to determine a cloud signal for the calculation of the CTH (Sect. 2.3.3). The distribution of CloudSat and predicted reflectivities is similar for up to −10 dBZ. For higher reflectivities, the distributions diverge. As demonstrated in Fig. 2, the network fails to accurately reconstruct high reflectivities. The difference increases for reflectivities > 0 dBZ. Both reflectivity distributions are dom-

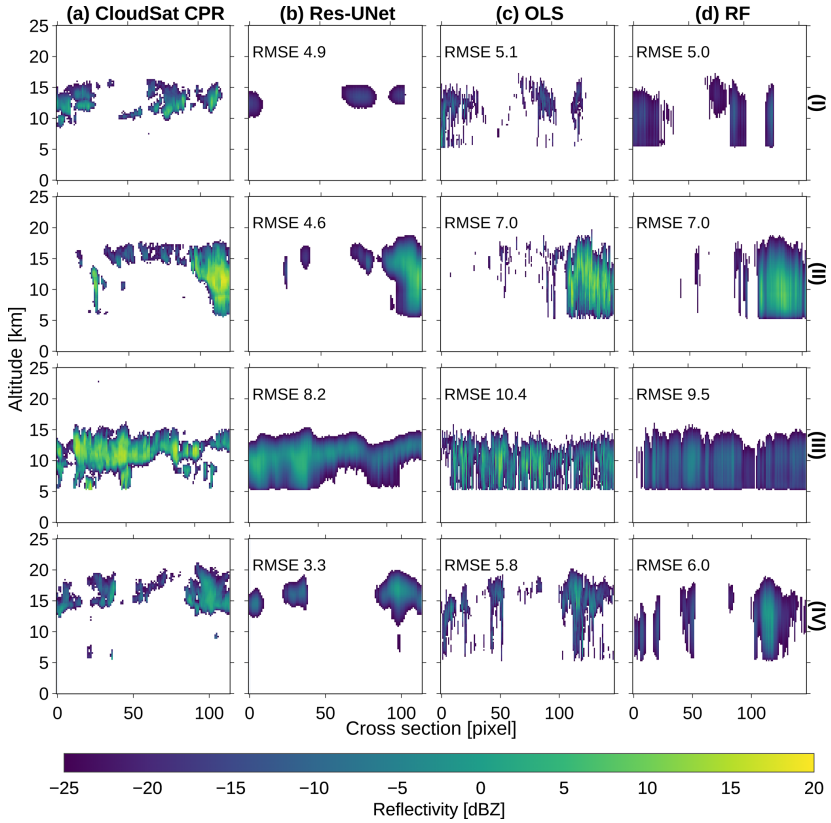

Figure 5. Reconstructing the radar cross section for four random examples of the test dataset ($n = 1500$). Values $\leqq -25$ dBZ are displayed transparently. We compare the reflectivity between the processed CloudSat cross sections **(a)** and the predictions of the Res-UNet **(b)**, OLS **(c)**, and RF **(d)** for each transect **(I)**–**(IV)**. The RMSE describes the error between the CloudSat data and the predicted profile.

inated by cloud-free pixels of −25 dBZ (Fig. 8c). The comparison shows the importance of the background value for the whole distribution. For values > -15 dBZ, the difference between the distributions decreases. The shift of the distribution is reflected in the CTH in Fig. 8b. Both datasets display a maximum CTH at 7 km height. This first peak is overestimated by the model. A second peak around 12–15 km height is underestimated by the Res-UNet. The difference between the predicted CTH is reflected within Fig. 8d. The mismatch between the two peaks is about the same size. The underestimated second peak can be traced back to the inaccuracy of the predicted reflectivities. The Res-UNet overestimates reflectivities < -15 dBZ at all height levels up to 15 km. It misses high reflectivities responsible for the peak of the CTH at 12–15 km (Fig. 3b). Instead, we see an overall surplus of background values in the FD prediction.

Calculating the CTH on the FD predictions substantially increases the number of available data points compared to the CloudSat data. Predicted images surpass the radar observations by a factor of 10 000. We use the 3D cloud tomography to derive the FD CTH on the test dataset. For each time step, we calculate the CTH on the FD and aggregate the results to a monthly mean. These values are compared to the CLAAS-V002E1 product with a resolution of 0.05° (Finkensieper et al., 2020). The predicted CTH has a resolution of 0.03°. Due to this mismatch, our predictions show more fragmented structures (Sect. 2.3.3). CloudSat faces sensor limitations at low and high altitudes of the troposphere (Sect. 2.1.2). While our analysis reveals an overall high agreement, the lack of, e.g., thin clouds within the radar data can lead to a reduced similarity between the CLAAS-V002E1 data and the predicted CTH. We observe a connection between the similarity of the datasets and the hemisphere. In the Northern Hemisphere, the highest number of image–profile pairs and the highest CTH difference occur between 0 and 20° N. Between the tropics of the Southern Hemisphere, the number of observations is similar, whereas the CTH difference is considerably lower. The variability between the hemispheres can be traced back to the distribution of land masses. A higher proportion of oceans in the Southern Hemisphere and a modified solar zenith angle affect the formation of clouds (Bruno et al., 2021). The result is an increased model performance which

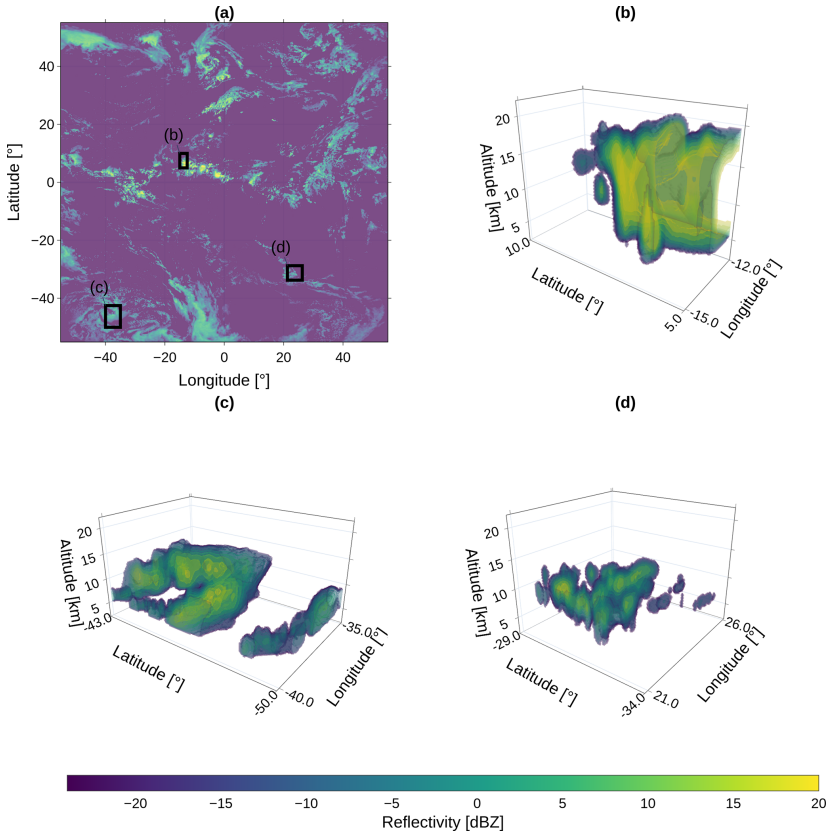

Figure 6. Prediction of 3D cloud structures from the Res-UNet along the FD MSG SEVIRI domain with a top view of the maximum cloud column reflectivity for each pixel on 6 May 2016 at 13:00 UTC **(a)**. The detailed views in panels **(b–d)** span several tiles of 100×100 pixels (2.5° on the geographic grid) to show the absence of artifacts between predictions.

might be caused by the existence of either more uniform or less complex clouds.

Although the small-scale accuracy of the predicted CTH is improvable, the results allow an investigation of regional differences on the large scale. These differences arise especially around the Equator and at mid to high latitudes (Fig. 9). At mid latitudes, the CTH over water bodies is overestimated in the Southern Hemisphere and underestimated in the Northern Hemisphere. These differences can be traced back to an increased RMSE in these regions (Fig. 7). In contrast, a low RMSE in the subtropics increases the accuracy of the predicted CTH. The model is biased toward predicting lower clouds than the observational data. Overall, the Res-UNet overestimates the occurrence of clouds in 6–8 km while underestimating high clouds (Fig. 8b).

This issue is reflected within Fig. 10. Here, we visualize the geographic distribution of the CTH difference (Fig. 10a). The mean difference over all the pixels accounts for 1.28 km. While the data show an overall agreement, the pixel-wise difference rises to a maximum of 10 km. This applies especially to regions in the subtropics. We observe an underestimation of the predicted CTH over land. Above the Atlantic Ocean, especially in the tropics, the predictions are too high. The highest difference occurs in the subtropics in the Northern Hemisphere (Fig. 10b). At 20° N, the mean difference accounts for 5 km. Around the tropics and mid latitudes, both datasets are in higher agreement. The distribution of the CTH difference is inversely proportional to the number of matched image–profile pairs (Fig. 7b). The CTH difference decreases with an increasing number of observational data from CloudSat. This applies to predictions over land and sea. Since we lack ground truth in the subtropics, the performance of the predictions decreases. The geographical differences are only partly in accordance with the distribution of the RMSE (Fig. 7a). While the RMSE is lower in the northern subtropics, the error of the predicted CTH reaches its peak (Fig. 10b). The RMSE alone does not appear to be a suitable measure for defining the reliability of the predicted reflectivity on the FD. This is due to the influ-

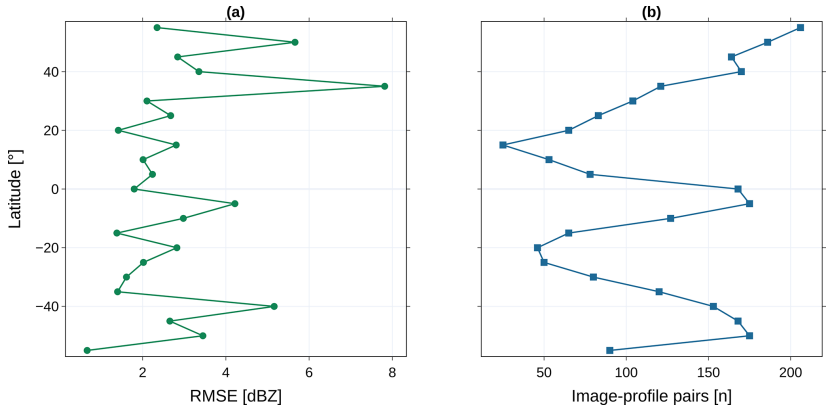

Figure 7. Zonal RMSE of the Res-UNet **(a)** and number of matched image–profile pairs used for model evaluation **(b)** for latitudes between 60° N and 60° S ($n = 1500$).

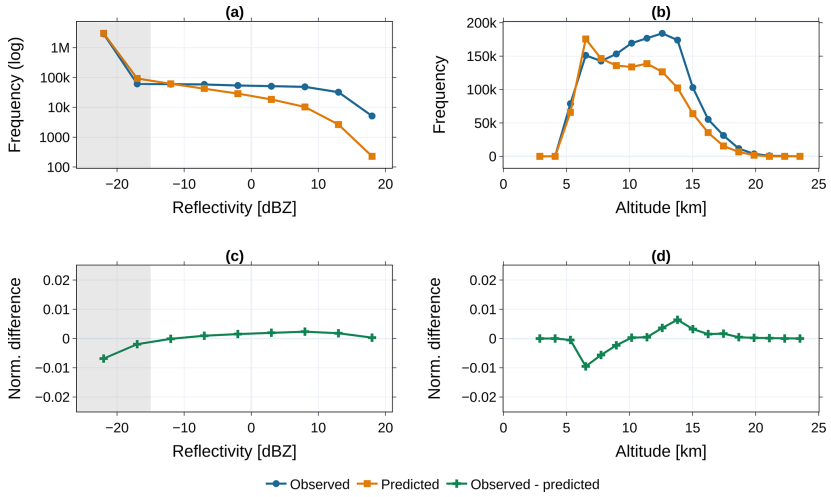

Figure 8. Comparing the reflectivity distribution and the derived CTH for CloudSat and the Res-UNet predictions. Data are calculated on the test dataset and aggregated to a monthly mean ($n = 1500$). The upper-row frequencies **(a)** and **(b)** display the dBZ and computed CTH for observed and predicted data. Lower-row images **(c)** and **(d)** show the difference between the observed and predicted data. Grey areas in plots **(a)** and **(c)** contain reflectivities below the threshold of −15 dBZ applied for the CTH analysis.

ence of the skewed reflectivity distribution on the RMSE and its geographic variability.

Even though the comparison of the CTH shows regional differences, the predictions can be used to represent the CTH pattern on the FD. The CLAAS-V002E1 data are computed using the MSG SEVIRI imager data as well as derived products and additional data. Each of them brings its own bias, potentially multiplying their effects on the final CTH. In contrast, our CTH is only based on the predicted reflectivity. In that way, we can minimize the influence of additional data sources.

4 Discussion

The Res-UNet makes predictions based on the MSG SEVIRI channels, preserving the spatial details and global context during training (Wang et al., 2022). The error of the model varies depending on cloud structure within the radar cross section. Compared to pixel-based approaches like OLS, the Res-UNet better reconstructs the pixel connectivity. The OLS and RF operate on 1D cloud columns. This limits their ability to extrapolate cloud information to a larger scale, resulting in fragmented reconstructions (Fig. 5c and d). While the RMSE and the reflectivity distribution are similar across all the models, only the Res-UNet predicts a contiguous radar

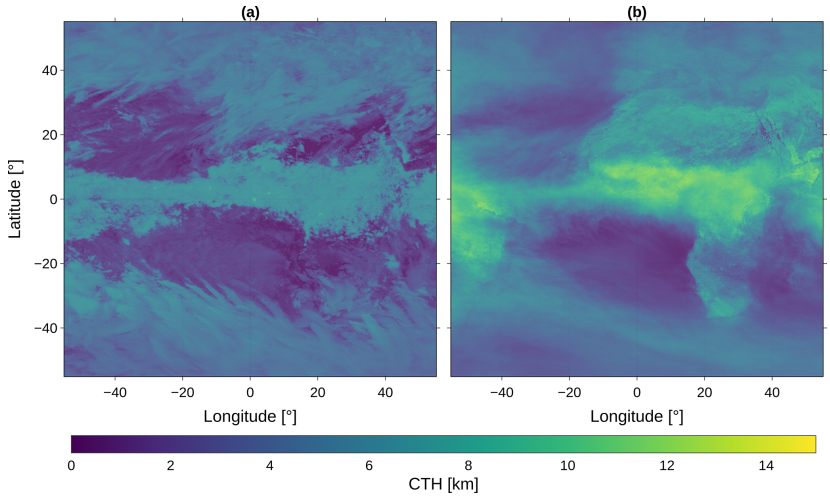

Figure 9. Monthly aggregation for the derived CTH for May 2016 **(a)** compared to CLAAS-V002E1 CTO **(b)**.

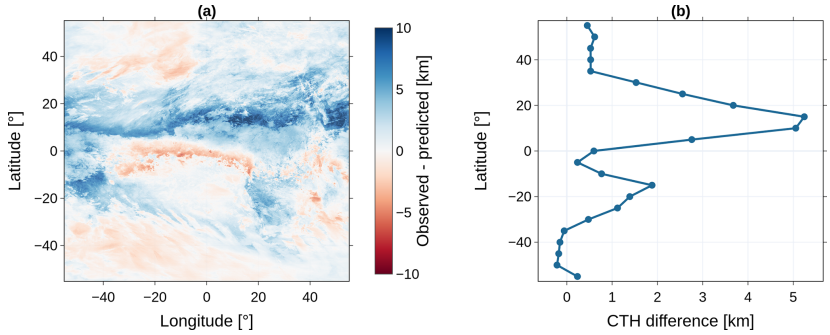

Figure 10. Difference between CLAAS-V002E1 CTO and the computed CTH for May 2016. Panel **(a)** shows the geographic distribution on the FD, panel **(b)** the zonal error.

cross section. Choosing a DL framework eliminates the need for prior predictor variable selection (Kühnlein et al., 2014; Leinonen et al., 2019). This can reduce the user bias in the input data (Jeppesen et al., 2019; Jiao et al., 2020).

The Res-UNet shows a 30 % improvement in the mean RMSE (Fig. 4). We could potentially further enhance the model performance by utilizing a more complex architecture. Our input data differ from typical grey-scale or RGB images, as they comprise multiple input channels and result in 3D output (Drönner et al., 2018). Given the demands of our data and resource constraints, we adapted a standard UNet architecture rather than using a pre-trained model (Amato et al., 2020). Selecting the RMSE as a loss function can increase the blurriness in the results, particularly as model bias grows (Mathieu et al., 2016). This issue becomes apparent as all the models struggle to predict high reflectivities (Fig. 2b–d). The predictions are influenced by an imbalance within the CloudSat data, with the distribution of all the models skewed toward low reflectivities. A resolution mismatch between CloudSat and MSG SEVIRI exacerbates this imbalance, causing peak reflectivities to blur out (Fig. 5b–d).

Our study covers a large-scale AOI spanning 60° in all directions. In contrast to the studies of Leinonen et al. (2019) or Hilburn et al. (2020), we incorporate a diverse landscape into our training. While Hilburn et al. (2020) focused on radar signal reconstruction over the USA using land-based radar, Leinonen et al. (2019) concentrated on radar cross-sectional prediction over the sea. In our study, we match image–profile pairs over land and the sea to achieve model invariance in the topography. The performance of our Res-UNet is similar to their results. Nevertheless, we observe regional difference, especially for the CTH. We use a geographical analysis to highlight the importance of the topography and land–sea distribution and their impact on cloud microphysics (Wang et al., 2023).

We emphasize the influence of the geographic location and the CloudSat orbit, particularly in regions farther from the Equator, where sensor accuracy diminishes (Fig. 7). Currently, we estimate the model error to be mainly influenced by the data imbalance and chosen loss function. In the future, addressing the cloud parallax shift in high-angle satellite observations could enhance the results by a more accurate image–profile matching (Bieliński, 2020). The most accurate predictions fall between 25° N and 25° S (Fig. 7), while mid and high latitudes exhibit a higher RMSE. This is likely due to the land–sea distribution and connected cloud patterns. Over ocean bodies, the model overestimates the reflectivity (Fig. 10a). Using the water vapor channels could lead to this distortion. Improved predictions are evident over the Southern Hemisphere. Since CloudSat operates in a sun-synchronous orbit, it misses diurnal variations in each region (Sect. 2.1.2). In this study, we only derive daytime predictions. This is due to the influence of solar radiation in VIS channels (Hilburn et al., 2020; Jeppesen et al., 2019). Additional distortions may arise from VIS channels, as imager data only represent the uppermost cloud layer. Depending on the location, they can be highly influenced by the surface albedo (Drönner et al., 2018). Training a model without the VIS channels can help to achieve predictions independent of the daytime. Reducing the extent of the AOI can mitigate the geographic performance differences but limits the applicability of the network. Training regional models and adjusting the loss function and model architectures offer potential solutions to improve the results of the 3D cloud tomography.

The reconstructed CloudSat cross sections are comparable to results achieved by Leinonen et al. (2019). For both studies, the RMSE varies between 0 and 1 dBZ for cloud-free samples, between 3 and 7 dBZ for more uniform clouds, and by more than 10 dBZ for multi-layer clouds. A common limitation is accurately representing multi-layer clouds. Using the satellite channels to derive this information may be limited (Schmetz et al., 2002; Thies and Bendix, 2011). High reflectivities tend to be underestimated due to noise near the ground (Stephens et al., 2008). To mitigate this, we exclude affected height levels, but this results in incomplete model predictions between 0 and 5 km (Fig. 2). Reducing noise is crucial for improving the performance of DL applications in remote sensing (Enitan and Ilesanmi, 2021). Our results are significantly influenced by the resolution difference between CloudSat and MSG SEVIRI as well as the choice of the loss function (Sect. 2.1.3). The aggregation of CloudSat pixels blurs the contrast within individual clouds (Fig. 5a), which is further reflected in the increased RMSE. In contrast, Leinonen et al. (2019) use data from the MODIS satellite. It has a higher spatial resolution than the MSG SEVIRI data, allowing for sharper predictions along the radar transect. However, polar-orbiting satellites like MODIS lack the spatiotemporal coverage of geostationary satellites (Dubovik et al., 2021). In their study, Wang et al. (2023) derive 24 000 training samples for matching CloudSat and MODIS over 6 years. By using MSG SEVIRI data, we amplify the volume of the training data. We have extracted approximately 30 000 training samples from 1 year of imager data, which results in a ratio of about 1 : 7.

Currently, a compromise on the resolution is necessary to obtain predictions for Europe and Africa. However, promising new instruments are emerging. While data from comparable sources like the GOES-R series and the Himawari 8/9 satellites already offer a 1 km resolution, the recently launched Meteosat Third Generation satellite by EUMETSAT allows us to close the gap and enables a more precise representation of individual clouds (Holmlund et al., 2021). Although our approach currently focuses on a region centered around 0° longitude, we can apply the same framework to other geostationary satellites, potentially achieving global 3D cloud coverage throughout the troposphere. The predicted cloud field can be valuable for time series analysis, enabling the tracking of clouds in four dimensions across space and time. Our results facilitate the identification of large-scale cloud patterns. They offer various applications, such as analyzing cloud organizational structures, pinpointing lightning locations, or conducting precipitation onset analyses. While we use CloudSat radar data as our ground truth, we need to evaluate whether this approach can be adapted to other 2D transect data sources, such as aerosol measurements. Former studies already derived aerosol properties from imager data (Carrer et al., 2010). The DL framework could help to achieve a full 3D retrieval of aerosols.

5 Conclusions

With the help of a neural network, we demonstrate for the first time the potential to infer comprehensive 3D radar reflectivities from 2D geostationary satellite images. While former studies were confined to a smaller region or the reconstruction of the 2D radar transect, we provide a framework to model the 3D cloud field at a high spatiotemporal resolution. The study is focused on Africa and Europe, but the approach can be used to predict the radar reflectivity on a global scale. Using only the predicted reflectivity, we derive the CTH without external data sources. Overall, the approach accurately reconstructs cloud structures under varying environmental conditions on the FD. Although the results are affected by sensor-specific and technical limitations, a vast potential for applications in atmospheric and climate sciences is apparent. With steadily growing data and the emergence of improved instruments, the results can close the existing global data gap. We emphasize the benefit of extrapolating a 3D cloud field, especially in remote oceanic regions. Future work will focus on extending the proposed network by data with an enhanced spatial and temporal resolution and investigating 3D cloud processes in active applications.

Appendix A: Overview of Res-UNet parameters

Table A1. Hyperparameters and training parameters of the Res-UNet.

Type	Parameter	Value
Hyperparameters	Depth	4
	Input channels	11
	Output channels	90
	Filter size	3×3
	Pooling size	2×2
	Dropout	0
	Activation function	ReLU
Training parameters	Number of epochs	50
	Batch size	4
	Input size	128×128
	Crop size	100×100
	Initial learning rate	0.001
	Learning rate (LR) scheduler (factor)	0.1
	Optimizer	ADAM
	Weight decay	0.00001
	Loss function	RMSE
	Augmentation (horizontal flip)	Randomness = 50 %

Appendix B: Summary of trainable model parameters

Table B1. Total number of Res-UNet model parameters.

Total number of trainable parameters	Estimated total size (MB)
1 893 328	194.27

Code and data availability. The source code for the imager data-matching scheme and the model framework is available at https://doi.org/10.5281/zenodo.8238110 (Brüning, 2023). The Meteosat SEVIRI level 1.5 data used in this study have been downloaded from the EUMETSAT Data Centre at https://navigator.eumetsat.int (last access: 27 July 2023; Schmetz et al., 2002). The level 2B-GEOPROF CloudSat data used in this study have been downloaded from the CloudSat Data Processing Center at https://www.cloudsat.cira.colostate.edu/ (last access: 27 July 2023, Marchand et al., 2008).

Author contributions. SB and HT designed the study. SB and SN developed the model code. SB performed the modeling and visualization. SB and HT contributed to the model validation and analysis of cloud properties. SB and HT wrote the draft of the paper. All the authors have read and agreed to the published version of the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We acknowledge the infrastructure provided by the Max Planck Graduate Center Mainz. We acknowledge EUMETSAT for providing access to the Meteosat SEVIRI imager data. We acknowledge the Cooperative Institute for Research in the Atmosphere, CSU, for providing access to the CloudSat GEOPROF-2B data. We acknowledge CM SAF for providing access to the CLAAS-2.1 data. We thank Peter Spichtinger for useful discussions and comments on the manuscript.

Financial support. This research has been supported by the Carl Zeiss Foundation (grant no. P2018-02-003).

This open-access publication was funded by Johannes Gutenberg University Mainz.

Review statement. This paper was edited by Cuiqi Zhang and reviewed by three anonymous referees.

References

Amato, F., Guignard, F., Robert, S., and Kanevski, M.: A novel framework for spatio-temporal prediction of environmental data using deep learning, Sci. Rep.-UK, 10, 22243, https://doi.org/10.1038/s41598-020-79148-7, 2020.

Barker, H. W., Jerg, M. P., Wehr, T., Kato, S., Donovan, D. P., and Hogan, R. J.: A 3D cloud-construction algorithm for the EarthCARE satellite mission, Q. J. Roy. Meteor. Soc., 137, 1042–1058, https://doi.org/10.1002/qj.824, 2011.

Bedka, K., Brunner, J., Dworak, R., Feltz, W., Otkin, J., and Greenwald, T.: Objective Satellite-Based Detection of Overshooting Tops Using Infrared Window Channel Brightness Temperature Gradients, J. Appl. Meteorol. Clim., 49, 181–202, https://doi.org/10.1175/2009JAMC2286.1, 2010.

Benas, N., Finkensieper, S., Stengel, M., van Zadelhoff, G.-J., Hanschmann, T., Hollmann, R., and Meirink, J. F.: The MSG-SEVIRI-based cloud property data record CLAAS-2, Earth Syst. Sci. Data, 9, 415–434, https://doi.org/10.5194/essd-9-415-2017, 2017.

Bieliński, T.: A Parallax Shift Effect Correction Based on Cloud Height for Geostationary Satellites and Radar Observations, Remote Sens.-UK, 12, 365, https://doi.org/10.3390/rs12030365, 2020.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmos. Chem. Phys., 15, 5325–5358, https://doi.org/10.5194/acp-15-5325-2015, 2015.

Bony, S., Stevens, B., Frierson, D. M. W., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P., Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, circulation and climate sensitivity, Nat. Geosci., 8, 261–268, https://doi.org/10.1038/ngeo2398, 2015.

Boulesteix, A.-L., Janitza, S., Kruppa, J., and König, I. R.: Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics, WIREs Data Min. Knowl., 2, 493–507, https://doi.org/10.1002/widm.1072, 2012.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Brüning, S.: AI-derived 3D cloud tomography, Zenodo [code], https://doi.org/10.5281/zenodo.8238110, 2023.

Bruno, O., Hoose, C., Storelvmo, T., Coopman, Q., and Stengel, M.: Exploring the Cloud Top Phase Partitioning in Different Cloud Types Using Active and Passive Satellite Sensors, Geophys. Res. Lett., 48, e2020GL089863, https://doi.org/10.1029/2020GL089863, 2021.

Cao, K. and Zhang, X.: An Improved Res-UNet Model for Tree Species Classification Using Airborne High-Resolution Images, Remote Sens.-UK, 12, 1228, https://doi.org/10.3390/rs12071128, 2020.

Carrer, D., Roujean, J.-L., Hautecoeur, O., and Elias, T.: Daily estimates of aerosol optical thickness over land surface based on a directional and temporal analysis of SEVIRI MSG visible observations, J. Geophys. Res.-Atmos., 115, D10, https://doi.org/10.1029/2009JD012272, 2010.

Chen, Y., Chen, G., Cui, C., Zhang, A., Wan, R., Zhou, S., Wang, D., and Fu, Y.: Retrieval of the vertical evolution of the cloud effective radius from the Chinese FY-4 (Feng Yun 4) next-generation geostationary satellites, Atmos. Chem. Phys., 20, 1131–1145, https://doi.org/10.5194/acp-20-1131-2020, 2020.

Cintineo, J. L., Pavolonis, M. J., Sieglaff, J. M., Wimmers, A., Brunner, J., and Bellon, W.: A Deep-Learning Model for Automated Detection of Intense Midlatitude Convection Using Geostationary Satellite Images, Weather Forecast., 35, 2567–2588, https://doi.org/10.1175/WAF-D-20-0028.1, 2020.

CloudSat Data Processing Center: Level 2B GEOPROF, Data Products, CloudSat DPC, https://www.cloudsat.cira.colostate.edu/data-products/2b-geoprof (last access: 27 July 2023), 2023.

Denby, L.: Discovering the Importance of Mesoscale Cloud Organization Through Unsupervised Classification, Geophys. Res. Lett., 47, e2019GL085190, https://doi.org/10.1029/2019GL085190, 2020.

Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C.: ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data, ISPRS J. Photogramm., 162, 94–114, https://doi.org/10.1016/j.isprsjprs.2020.01.013, 2020.

Dixit, M., Chaurasia, K., and Kumar Mishra, V.: Dilated-ResUnet: A novel deep learning architecture for building extraction from medium resolution multi-spectral satellite imagery, Expert Syst. Appl., 184, 115530, https://doi.org/10.1016/j.eswa.2021.115530, 2021.

Drönner, J., Korfhage, N., Egli, S., Mühling, M., Thies, B., Bendix, J., Freisleben, B., and Seeger, B.: Fast Cloud Segmentation Using Convolutional Neural Networks, Remote Sens.-UK, 10, 1782, https://doi.org/10.3390/rs10111782, 2018.

Dubovik, O., Schuster, G., Xu, F., Hu, Y., Bösch, H., Landgraf, J., and Li, Z.: Grand Challenges in Satellite Remote Sensing, Front. Remote Sens., 2, 619818, https://doi.org/10.3389/frsen.2021.619818, 2021.

Enitan, I. and Ilesanmi, T.: Methods for image denoising using convolutional neural network: a review, Complex & Intelligent Systems, 7, 2189–2198, https://doi.org/10.1007/s40747-021-00428-4, 2021.

EUMETSAT Data Services: High Rate SEVIRI Level 1.5 Image Data – MSG – 0 degree, available at https://navigator.eumetsat.int/product/EO:EUM:DAT:MSG:HRSEVIRI (last access: 27 July 2023), 2023.

Finkensieper, S., Meirink, J. F., van Zadelhoff, G.-J., Hanschmann, T., Benas, N., Stengel, M., Fuchs, P., Hollmann, R., Kaiser, J., and Werscheck, M.: CLAAS-2.1: CM SAF CLoud property dAtAset using SEVIRI – Edition 2.1, EUMESTST [data set], https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V002_01, 2020.

Forster, L., Davis, A. B., Diner, D. J., and Mayer, B.: Toward Cloud Tomography from Space Using MISR and MODIS: Locating the "Veiled Core" in Opaque Convective Clouds, J. Atmos. Sci., 78, 155–166, https://doi.org/10.1175/JAS-D-19-0262.1, 2021.

Guillaume, A., Kahn, B. H., Yue, Q., Fetzer, E. J., Wong, S., Manipon, G. J., Hua, H., and Wilson, B. D.: Horizontal and Vertical Scaling of Cloud Geometry Inferred from CloudSat Data, J. Atmos. Sci., 75, 2187–2197, https://doi.org/10.1175/JAS-D-17-0111.1, 2018.

Ham, S.-H., Kato, S., Barker, H. W., Rose, F. G., and Sun-Mack, S.: Improving the modelling of short-wave radiation through the use of a 3D scene construction algorithm, Q. J. Roy. Meteor. Soc., 141, 1870–1883, https://doi.org/10.1002/qj.2491, 2015.

Han, L., Liang, H., Chen, H., Zhang, W., and Ge, Y.: Convective Precipitation Nowcasting Using U-Net Model, IEEE T. Geosci. Remote, 60, 1–8, https://doi.org/10.1109/TGRS.2021.3100847, 2022.

Henken, C. C., Schmeits, M. J., Deneke, H., and Roebeling, R. A.: Using MSG-SEVIRI Cloud Physical Properties and Weather Radar Observations for the Detection of Cb/TCu Clouds, J. Appl. Meteorol. Clim., 50, 1587–1600, https://doi.org/10.1175/2011JAMC2601.1, 2011.

Hilburn, K. A., Ebert-Uphoff, I., and Miller, S. D.: Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations, J. Appl. Meteorol. Clim., 60, 3–21, https://doi.org/10.1175/JAMC-D-20-0084.1, 2020.

Holmlund, K., Grandell, J., Schmetz, J., Stuhlmann, R., Bojkov, B., Munro, R., Lekouara, M., Coppens, D., Viticchie, B., August, T., Theodore, B., Watts, P., Dobber, M., Fowler, G., Bojinski, S., Schmid, A., Salonen, K., Tjemkes, S., Aminou, D., and Blythe, P.: Meteosat Third Generation (MTG): Continuation and Innovation of Observations from Geostationary Orbit, B. Am. Meteorol. Soc., 102, 990–1015, https://doi.org/10.1175/BAMS-D-19-0304.1, 2021.

Hu, K., Zhang, D., and Xia, M.: CDUNet: Cloud Detection UNet for Remote Sensing Imagery, Remote Sens.-UK, 13, 4533, https://doi.org/10.3390/rs13224533, 2021.

Huo, J., Lu, D., Duan, S., Bi, Y., and Liu, B.: Comparison of the cloud top heights retrieved from MODIS and AHI satellite data with ground-based Ka-band radar, Atmos. Meas. Tech., 13, 1–11, https://doi.org/10.5194/amt-13-1-2020, 2020.

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., and Saynisch-Wagner, J.: Towards neural Earth system modelling by integrating artificial intelligence in Earth system science, Nat. Mach. Intell., 3, 667–674, https://doi.org/10.1038/s42256-021-00374-3, 2021.

Jeppesen, J. H., Jacobsen, R. H., Inceoglu, F., and Toftegaard, T. S.: A cloud detection algorithm for satellite imagery based on deep learning, Remote Sens. Environ., 229, 247–259, https://doi.org/10.1016/j.rse.2019.03.039, 2019.

Jiao, L., Huo, L., Hu, C., and Tang, P.: Refined UNet: UNet-Based Refinement Network for Cloud and Shadow Precise Segmentation, Remote Sens.-UK, 12, 2001, https://doi.org/10.3390/rs12122001, 2020.

Jones, N.: How machine learning could help to improve climate forecasts, Nature, 548, 379–379, https://doi.org/10.1038/548379a, 2017.

Jordahl, K., Bossche, J. V. D., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., Maxalbert, Bilogour, A., Rey, S., Ren, C., Arribas-Bel, D., Wasser, L., Wolf, L. J., Journois, M., Wilson, J., Greenhall, A., Holdgraf, C., Filipe, and Leblanc, F.: geopandas/geopandas: v0.8.1, Zenodo [code], https://doi.org/10.5281/ZENODO.3946761, 2020.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., and Kumar, V.: Machine Learning for the Geosciences: Challenges and Opportunities, IEEE T. Knowl. Data En., 31, 1544–1554, https://doi.org/10.1109/TKDE.2018.2861006, 2019.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv [preprint], https://doi.org/10.48550/arXiv.1412.6980, 2014.

Kühnlein, M., Appelhans, T., Thies, B., and Nauss, T.: Improving the accuracy of rainfall rates from optical satellite sensors with machine learning – A random forests-based approach applied to MSG SEVIRI, Remote Sens. Environ., 141, 129–143, https://doi.org/10.1016/j.rse.2013.10.026, 2014.

Le Goff, M., Tourneret, J.-Y., Wendt, H., Ortner, M., and Spigai, M.: Deep learning for cloud detection, in: 8th International Conference of Pattern Recognition Systems (ICPRS 2017), 1–6, https://doi.org/10.1049/cp.2017.0139, Madrid, Spain, 11–13 July 2017, 2017.

Lee, Y., Kummerow, C. D., and Ebert-Uphoff, I.: Applying machine learning methods to detect convection using Geostationary Operational Environmental Satellite-16 (GOES-16) advanced baseline imager (ABI) data, Atmos. Meas. Tech., 14, 2699–2716, https://doi.org/10.5194/amt-14-2699-2021, 2021.

Leinonen, J., Guillaume, A., and Yuan, T.: Reconstruction of Cloud Vertical Structure With a Generative Adversarial Network, Geophys. Res. Lett., 46, 7035–7044, https://doi.org/10.1029/2019GL082532, 2019.

Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F., and Li, W.: DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation, IEEE J. Sel. Top. Appl., 11, 3954–3962, https://doi.org/10.1109/JSTARS.2018.2833382, 2018.

Liu, Y., Racah, E., Prabhat, M., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M., and Collins, W.: Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets, arXiv [preprint], https://doi.org/10.48550/arXiv.1605.01156, 2016.

Marais, W. J., Holz, R. E., Reid, J. S., and Willett, R. M.: Leveraging spatial textures, through machine learning, to identify aerosols and distinct cloud types from multispectral observations, Atmos. Meas. Tech., 13, 5459–5480, https://doi.org/10.5194/amt-13-5459-2020, 2020.

Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G.: Hydrometeor Detection Using Cloudsat–An Earth-Orbiting 94 GHz Cloud Radar, J. Atmos. Ocean. Tech., 25, 519–533, https://doi.org/10.1175/2007JTECHA1006.1, 2008.

Mathieu, M., Couprie, C., and LeCun, Y.: Deep multi-scale video prediction beyond mean square error, arXiv [preprint], https://doi.org/10.48550/arXiv.1511.05440, 2016.

McCandless, T. and Jiménez, P. A.: Examining the Potential of a Random Forest Derived Cloud Mask from GOES-R Satellites to Improve Solar Irradiance Forecasting, Energies, 13, 1671, https://doi.org/10.3390/en13071671, 2020.

Miller, S. D., Forsythe, J. M., Partain, P. T., Haynes, J. M., Bankert, R. L., Sengupta, M., Mitrescu, C., Hawkins, J. D., and Haar, T. H. V.: Estimating Three-Dimensional Cloud Structure via Statistically Blended Satellite Observations, J. Appl. Meteorol. Clim., 53, 437–455, https://doi.org/10.1175/JAMC-D-13-070.1, 2014.

Noh, Y.-J., Haynes, J. M., Miller, S. D., Seaman, C. J., Heidinger, A. K., Weinrich, J., Kulie, M. S., Niznik, M., and Daub, B. J.: A Framework for Satellite-Based 3D Cloud Data: An Overview of the VIIRS Cloud Base Height Retrieval and User Engagement for Aviation Applications, Remote Sens.-UK, 14, 5524, https://doi.org/10.3390/rs14215524, 2022.

Norris, J. R., Allen, R. J., Evan, A. T., Zelinka, M. D., O'Dell, C. W., and Klein, S. A.: Evidence for climate change in the satellite cloud record, Nature, 536, 72–75, https://doi.org/10.1038/nature18273, 2016.

Pan, X., Lu, Y., Zhao, K., Huang, H., Wang, M., and Chen, H.: Improving Nowcasting of Convective Development by Incorporating Polarimetric Radar Variables Into a Deep-Learning Model, Geophys. Res. Lett., 48, e2021GL095302, https://doi.org/10.1029/2021GL095302, 2021.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE T. Geosci. Remote, 55, 502–525, https://doi.org/10.1109/TGRS.2016.2610522, 2017.

Rasp, S., Pritchard, M. S., and Gentine, P.: Deep learning to represent sub-grid processes in climate models, P. Natl. Acad. Sci. USA, 115, 9684–9689, https://doi.org/10.1073/pnas.1810286115, 2018.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., and Bengio, Y.: Tackling Climate Change with Machine Learning, ACM Comput. Surv., 55, 1–96, https://doi.org/10.1145/3485128, 2022.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation., in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, edited by: Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., Springer International Publishing, Cham, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., and Zscheischler, J.: Inferring causation from time series in Earth system sciences, Nat. Commun., 10, 2553, https://doi.org/10.1038/s41467-019-10105-3, 2019.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An Introduction to Meteosat Second Generation (MSG), B. Am. Meteorol. Soc., 83, 977–992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, 2002.

Seiz, G. and Davies, R.: Reconstruction of cloud geometry from multi-view satellite images, Remote Sens. Environ., 100, 143–149, https://doi.org/10.1016/j.rse.2005.09.016, 2006.

Shepherd, T. G.: Atmospheric circulation as a source of uncertainty in climate change projections, Nat. Geosci., 7, 703–708, https://doi.org/10.1038/ngeo2253, 2014.

Sieglaff, J., Hartung, D., Feltz, W., Cronce, L., and Lakshmanan, V.: A Satellite-Based Convective Cloud Object Tracking and Multipurpose Data Fusion Tool with Application to Developing Convection, J. Atmos. Ocean. Tech., 30, 510–525, https://doi.org/10.1175/JTECH-D-12-00114.1, 2013.

Steiner, M., Houze, R. A., and Yuter, S. E.: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data, J. Appl. Meteorol. Clim., 34, 1978–2007, https://doi.org/10.1175/1520-0450(1995)034<1978:CCOTDS>2.0.CO;2, 1995.

Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., Reinke, D., Partain, P., Mace, G. G., Austin, R., L'Ecuyer, T., Haynes, J., Lebsock, M., Suzuki, K., Waliser, D., Wu, D., Kay, J., Gettelman, A., Wang, Z., and Marchand, R.: CloudSat mission: Performance and early science after the first year of operation, J. Geophys. Res.-Atmos., 113, D8, https://doi.org/10.1029/2008JD009982, 2008.

Stevens, B. and Bony, S.: What Are Climate Models Missing?, Science, 340, 1053–1054, https://doi.org/10.1126/science.1237554, 2013.

Tarrio, K., Tang, X., Masek, J. G., Claverie, M., Ju, J., Qiu, S., Zhu, Z., and Woodcock, C. E.: Comparison of cloud detection algorithms for Sentinel-2 imagery, Science of Remote Sens., 2, 100010, https://doi.org/10.1016/j.srs.2020.100010, 2020.

Thies, B. and Bendix, J.: Satellite based remote sensing of weather and climate: recent achievements and future perspectives, Meteorol. Appl., 18, 262–295, https://doi.org/10.1002/met.288, 2011.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B.: Missing value estimation methods for DNA microarrays, Bioinformatics, 17, 520–525, https://doi.org/10.1093/bioinformatics/17.6.520, 2001.

Vial, J., Dufresne, J.-L., and Bony, S.: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates, Clim. Dynam., 41, 3339–3362, https://doi.org/10.1007/s00382-013-1725-9, 2013.

Wang, F., Liu, Y., Zhou, Y., Sun, R., Duan, J., Li, Y., Ding, Q., and Wang, H.: Retrieving Vertical Cloud Radar Reflectivity from MODIS Cloud Products with CGAN: An Evaluation for Different Cloud Types and Latitudes, Remote Sens.-UK, 15, 816, https://doi.org/10.3390/rs15030816, 2023.

Wang, S., Chen, W., Xie, S. M., Azzari, G., and Lobell, D. B.: Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery, Remote Sens.-UK, 12, 207, https://doi.org/10.3390/rs12020207, 2020.

Wang, Z., Zhao, J., Zhang, R., Li, Z., Lin, Q., and Wang, X.: UATNet: U-Shape Attention-Based Transformer Net for Meteorological Satellite Cloud Recognition, Remote Sens.-UK, 14, 104, https://doi.org/10.3390/rs14010104, 2022.

Wieland, M., Li, Y., and Martinis, S.: Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network, Remote Sens. Environ., 230, 111203, https://doi.org/10.1016/j.rse.2019.05.022, 2019.

Zantedeschi, V., Falasca, F., Douglas, A., Strange, R., Kusner, M. J., and Watson-Parris, D.: Cumulo: A Dataset for Learning Cloud Classes, arXiv [preprint], https://doi.org/10.48550/arXiv.1911.04227, 2022.

Zhang, D., He, Y., Li, X., Zhang, L., and Xu, N.: PrecipGradeNet: A New Paradigm and Model for Precipitation Retrieval with Grading of Precipitation Intensity, Remote Sens.-UK, 15, 227, https://doi.org/10.3390/rs15010227, 2023.

Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R., and Xue, H.: Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations, J. Geophys. Res.-Atmos., 117, D19, https://doi.org/10.1029/2012JD017655, 2012.

3.3 Relationship between convective clouds, their cores, and the cloud life-cycle

The preprint was accepted for publication in *Atmospheric Chemistry and Physics* as a research article.

I am the lead author of this study, where I worked on the implementation of a cloud and core tracking routine designed for the ML-based predictions described in Sect. 3.2. I analysed the data, created the figures, and wrote the manuscript. The co-author contributed to interpreting the results and proofreading the manuscript. The paper includes a section called *Author contributions* detailing individual contributions.

How to cite: Brüning, S. and Tost, H.: A machine learning-based perspective on deep convective clouds and their organisation in 3D. Part I: Influence of deep convective cores on the cloud life-cycle, *EGUsphere [preprint]*, https://doi.org/10.5194/egusphere-2025-374, **2025**.

Submitted: 05 February 2025
Revised: 23 May 2025
Accepted: 03 July 2025

A machine learning-based perspective on deep convective clouds and their organisation in 3D. Part I: Influence of deep convective cores on the cloud life-cycle

Sarah Brüning[1] and Holger Tost[1]

[1]Institute for Physics of the Atmosphere, Johannes Gutenberg University Mainz, Johann-Joachim-Becher-Weg 21, Mainz, 55128, Rhineland-Palatinate, Germany

Correspondence: Sarah Brüning (sbruenin@uni-mainz.de)

Abstract. In this two-part study, we examine spatio-temporal patterns of convective clouds, their properties, and organisation. We use a machine learning-based method to extrapolate a contiguous 3D cloud field of 2D satellite data. The predicted data are used to simultaneously track both the horizontal and vertical development of clouds. Our research focuses on West Africa, a region known for frequent convective events and severe weather. In Part 1, this study compares cloud and core properties and the cloud life-cycle over land and ocean during a six-month period from March to August 2019. Our analysis reveals that 65 % of tracked cloud systems contain only a single core and persist for less than three hours. Despite their shorter lifespan compared to multi-core clusters, single-core clouds exhibit stronger changes in the radar reflectivity and a higher vertical growth. In contrast, multi-core clouds show greater horizontal growth, encompassing larger cloud and core areas, higher cloud-top heights (CTH), and higher average reflectivity at 10 km altitude. We also find that, in systems with more cores, both the maximum number of cores and the peak core area occur later during the cloud life-cycle. Notably, the differences in cloud characteristics between land and ocean are smaller than those associated with the number of convective cores. However, the results may not fully capture climatological differences. Further research using longer time series is needed to quantify the observed variability of tropical convection.

1 Introduction

Convective clouds play a vital role in the hydrological cycle of the Earth through their radiative forcing and feedback mechanisms (Wielicki et al., 1995). Despite growing evidence for the connection between clouds and climate warming, they remain one of the greatest sources of uncertainty in climate sensitivity assessments (e.g., Bony et al. (2015); Sherwood et al. (2020)). Additionally, convective clouds are key drivers of severe weather, particularly large-scale systems like mesoscale convective systems (MCSs), which are linked to extreme events such as hailstorms, damaging winds, and intense rainfall (e.g., Houze

and Hobbs (1982); Leary and Houze (1980); Maddox (1980)). Because of their societal and environmental impacts, accurately representing convective clouds remains of particular interest.

MCSs are typically defined as convective storm complexes with an axis length of at least 100 km (Houze Jr., 2004). These systems often feature a broad cold cloud shield, one or more deep convective cores (hereafter, "cores"), strong vertical updrafts that connect these cores at higher altitudes, and widespread anvils extending from the convective region (Zipser and LeMone, 1980). While cores drive intense precipitation, the stratiform anvil and cirrus canopy generally produce lighter rain (e.g., Houze Jr. (1989); Hartmann et al. (1984)). Core sizes typically range from 10 to 100 km, with lifespans of 1–3 hours, whereas anvils can persist for up to 10–20 hours. The idealised MCS life-cycle includes three stages: development, maturity, and dissipation (Futyan and Genio, 2007). During development, deep convective cells form and transport condensate upward. In the maturity stage, the anvil and associated mesoscale circulation develop while convection continues. In the dissipation stage, deep convection ceases, and the system gradually fades (e.g., Houze and Hobbs (1982); Machado et al. (1998)). The MCS life-cycle is influenced by location, time of day, and surface type. For instance, small to medium MCSs commonly form over land in the afternoon due to local thermal instability and, potentially, sea breeze circulations (Chen and Houze, 1997). In contrast, oceanic MCSs experience weaker diurnal variability because of the surface's stable thermal properties (Nesbitt and Zipser, 2003).

Our current understanding of convective clouds largely stems from satellite observations (Haynes et al., 2009). We may identify convective clouds from satellite data by distinguishing core regions and the surrounding cloud field (Steiner et al., 1995). When derived from passive satellite observations, cores are typically characterised by cold peaks in brightness temperature, surrounded by warmer anvil regions. Morphological features such as aspect ratio, length, width, and area may further classify convective systems (Ganetis et al., 2018). For instance, passive and active sensors provide valuable insights into the temporal evolution of clouds. Passive sensors, especially those measuring infrared (IR) radiation, help identify cloud-top features (Mecikalski et al., 2010). However, they lack vertical resolution, making it difficult to distinguish between deep convection, stratiform clouds, and cirrus (Liu and Zipser, 2008). In contrast, active sensors like radar can resolve vertical cloud structures and hydrometeor distributions (e.g., Bacmeister and Stephens (2011); Oreopoulos et al. (2017)). Still, both sensor types offer only limited spatial or temporal coverage. As Masunaga and Luo (2016) point out, a global, continuous 3D view of convective clouds remains unavailable from current satellite missions.

Early studies relied on manual tracking, but automated detection algorithms now enable the processing of large datasets. Most of these algorithms are centroid-based, linking cloud objects across time steps (Prein et al., 2024). One of the earliest and most influential tools is *TITAN* Dixon and Wiener (1993), later adapted into *TINT* Raut et al. (2021), which is optimized for tracking fast-evolving storm cells. *TOOCAN*, developed by Fiolleau and Roca (2013), focuses specifically on convective cores and associated anvils in MCSs. More recently, general-purpose tools such as *PyFLEXTRKR* (Feng et al., 2023) and *tobac* (Heikenfeld et al., 2019) have emerged. *PyFLEXTRKR* offers flexible 2D tracking, while *tobac* supports 4D analysis, enabling a more comprehensive view of convective systems.

Despite decades of research, knowledge of the 3D structure of convective cores remains limited. In the absence of high-resolution, global 3D observational data, our understanding of the relationship between cores and overall cloud evolution relies

heavily on 2D observations and simulations. Active and passive sensors contain important vertical or horizontal information, but are limited in their spatial and temporal coverage (active) or offer only an approximation of the vertical column (passive) (Masunaga and Luo, 2016; Taylor et al., 2017). To address this gap, we apply a machine learning (ML) framework to reconstruct contiguous 3D radar reflectivity fields from 2D satellite data (Brüning et al., 2024). Our goal is to simultaneously capture the horizontal and vertical evolution of convective clouds and their cores. We use imagery from the Meteosat-11 SEVIRI sensor as input to the ML model, which is trained to reconstruct vertical cross sections based on CloudSat Cloud Profiling Radar (CPR) observations. This approach allows us to extrapolate a continuous 3D cloud field between 2.4 and 24 km altitude. The resulting dataset combines the spatial and temporal characteristics of the SEVIRI input with the vertical structure from CPR. We then use the *tobac* package to identify convective clouds and track them over time in 15-minute intervals. This enables the analysis of 3D convective cloud and core properties over both land and ocean. A focus is comparing the life-cycles of clouds with single versus multiple core regions, offering insights into the spatial clustering and organization of convection.

We organise the article as follows. In Sect. 2, we present the data used in this study. Section 3 provides details on details the ML-based 3D reconstruction and tracking methodology. Section 4 presents the results focusing on the temporal variability of convective cloud and core characteristics and the connection between the cores and cloud life-cycle. Section 5 compares our findings to known tropical convection characteristics and outlines limitations. Finally, Sect. 6 contains the concluding remarks.

2 Data

The area of interest (AOI) for this study spans a tropical region over central and western Africa, extending from 30° N to 30° S and 30° W to 30° E. This region is characterised by environmental conditions that contribute to the development of convective clouds (Takahashi et al., 2023). Our objective is to detect and analyse convective clouds and their life-cycles by a six month period between March and August 2019. This period was selected to highlight key characteristics of 3D cloud structures across different surface types within the AOI. Particular attention is given to the seasonal northward migration of the Inter-Tropical Convergence Zone (ITCZ) and the onset of the West African Monsoon (WAM). Since the WAM plays a critical role in shaping West Africa's climate and is responsible for a significant portion of the annual rainfall in the AOI, its arrival is expected to enhance the frequency of convective cloud formation (Andrews et al., 2024; Kniffka et al., 2019).

To investigate these phenomena, we employ a ML algorithm that generates time series of 3D radar reflectivity fields based on 2D satellite observations, as described in Brüning et al. (2024). The input data are derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the Meteosat-11 (MSG) satellite (Schmetz et al., 2002). The AOI is situated near the nadir of SEVIRI, which is positioned above the Equator at 0° longitude. SEVIRI captures multispectral imagery across 12 channels in the visible, near-infrared, and thermal-infrared ranges. Eleven of these channels offer a temporal resolution of 15 minutes and a spatial resolution of 3 km, while one high-resolution visible channel provides 1 km resolution at nadir (Table 1).

To validate our ML-based predictions, we use vertical cross-sections of radar reflectivity from the 94-GHz Cloud Profiling Radar (CPR) onboard the polar-orbiting CloudSat satellite. This active radar instrument transmits microwave pulses toward

Table 1. Overview of Meteosat SEVIRI channels (Schmetz et al., 2002).

Channel	Wavelength (μm)	Description	Spatial resolution at nadir	Retrieval at nighttime
VIS0.6	0.56-0.71	Visible channel	3 km	No
VIS0.8	0.74-0.88	Visible channel	3 km	No
NIR1.6	1.5-1.78	Near infrared window	3 km	No
IR3.9	3.48-4.36	Near infrared window	3 km	Yes
WV6.2	5.35-7.15	Upper-troposphere water vapour	3 km	Yes
WV7.3	6.85-7.85	Lower-troposphere water vapour	3 km	Yes
IR8.7	8.30-9.10	Mid infrared window	3 km	Yes
IR9.7	9.38-9.94	Ozone sensitivity	3 km	Yes
IR10.8	9.80-11.80	Clean longwave window	3 km	Yes
IR12.0	11.00-13.00	Dirty longwave window	3 km	Yes
IR 13.4	12.40-14.40	CO2 sensitivity	3 km	Yes
HRV	0.5-0.9	High-resolution visible	1 km	No

Earth to detect vertical profiles of cloud hydrometeors. The CPR achieves a vertical resolution of 240 m (distributed across 125 bins) and a horizontal resolution of 1.4 km across-track and 1.8 km along-track (Stephens et al., 2008). For this study, we use the level-2 2B-GEOPROF product. To address signal attenuation at lower altitudes, we limit the vertical analysis to 90 height levels ranging from 2.4 km to 24 km (Sassen and Wang, 2008). Additionally, due to reduced sensor sensitivity at high altitudes, the CPR may underrepresent certain cloud types, particularly thin ice clouds like cirrus. To mitigate noise, we filter the 2B-GEOPROF dataset using the CloudSat cloud mask quality flag (Marchand et al., 2008).

3 Method

3.1 Machine learning-based reconstruction of a 3D cloud field

In the following section, we briefly outline the method used to reconstruct a 3D cloud field, based on the framework developed by Brüning et al. (2024). Our approach employs a ML algorithm built on a 2D Res-UNet architecture — a modified version of a convolutional neural network specifically designed for image segmentation tasks (Ronneberger et al., 2015). While the model is primarily trained to reconstruct vertical cross-sections of the CloudSat CPR radar reflectivity using imagery from the MSG SEVIRI satellite, its output represents full 3D radar reflectivity volumes rather than just 2D slices.

The reconstructed 3D cloud field spans an area from 60° W to 60° E and from 60° S to 60° N, corresponding to 2400 × 2400 pixels in the horizontal dimensions. SEVIRI satellite imagery serves as input to the Res-UNet model, hence the horizontal resolution of the 3D data is 3 km x 3 km. Initially, 11 channels covering the visible, near-infrared, and thermal-infrared spectra

Table 2. Modifications to the Res-UNet applied in this study originally proposed in Brüning et al. (2024)

Parameter	Original configuration	Modification
Number of input channels	11	8
Loss function	L2	L1
Nighttime predictions	No	Yes
Average RMSE	3.05	2.99

were used (Table 1). For this study, we exclude the visible channels to ensure the model can make predictions independent of daylight conditions (Table 2).

Training data consist of 128 × 128 pixel patches of SEVIRI imagery that are spatially and temporally aligned with CloudSat overpasses. Each training sample includes a diagonal CPR cross-section. Due to the spatial resolution mismatch between MSG SEVIRI and CloudSat, we downsample the SEVIRI data to match the CPR's horizontal resolution. To address the strong class imbalance between cloudy and cloud-free conditions, we limit cloud-free samples to a maximum of 10 % of the training data. The model is trained on nine months of data and validated on a separate three-month period. The Res-UNet is trained to reconstruct CloudSat-like 3D reflectivity volumes with a horizontal size of 100 × 100 pixels and a vertical size of 90 levels. The predicted radar reflectivity values range from –25 to 20 dBZ and retain the 15-minute temporal resolution of the original SEVIRI input. We use an L1 loss function (mean absolute error) during training to evaluate the model's performance. Notably, direct validation is possible only for the diagonal cross-section, which accounts for about 10 % of each training sample. For the three-month test period, the modified daylight-independent model achieves a root mean square error (RMSE) of 2.99 dBZ — an improvement over the original model (Table 2). This level of accuracy is comparable to the 5 dBZ precision reported for other CloudSat products (Tomkins et al., 2024).

To generate complete coverage of the domain between 60° W to 60° E and 60° S to 60° N, the individual 3D output patches are stitched together, producing a unified output volume of with a size 2400 × 2400 × 90 pixels. This method may help to obtain a consistent spatial coverage, particularly over remote oceanic regions where active sensors are sparse (Prein et al., 2024). Visual inspection confirms the absence of artifacts at tile boundaries, indicating a seamless reconstruction of the 3D cloud field in different parts of the domain. To further assess model performance, we compute cloud top heights (CTH) from the predicted radar reflectivity and compare them to CTH values from the CMSAF CLAAS-V002E1 dataset (Finkensieper et al., 2020). The comparison reveals that the model captures realistic spatial patterns of CTH in both tropical and mid-latitude regions. However, model accuracy tends to decline with increasing distance from the MSG SEVIRI nadir. Finally, the time-series of 3D radar reflectivity volumes is merged along the temporal axis to generate a 4D cloud field, which is used to detect and track convective clouds. For the purposes of this study, we crop the domain to consist of 1200 × 1200 pixels, covering the region between 30° W–30° E and 30° N–30° S — effectively focusing on the area between the Tropic of Cancer and the Tropic of Capricorn.

3.2 Tracking convective clouds in 4D

In this study, we analyse the development and properties of convective clouds by employing the *tobac* package, a modular Python-based package for tracking atmospheric objects in 4D time series (Heikenfeld et al., 2019). In this study, we use the recently released version 1.5 of the software package (Sokolowsky et al., 2024). We merge the predicted 3D radar reflectivity fields along the temporal dimension and feed the 4D time series into the tracking algorithm to create continuous trajectories. The workflow to identify possibly convective trajectories consists of three steps: detecting cloud features by their centroid's position, segmenting the associated cloud field for each centroid, and linking segmented objects through time (Figure 1, a–c). Moreover, we aim to separate cloud clusters that are only connected by a few pixels in the horizontal and vertical dimensions (Oreopoulos et al., 2017). The workflow of this object-based approach is depicted in Fig. 1 and will be explained in the following paragraphs.

The framework, while enabling detailed analysis of convective cores, has limitations. The predicted 3D cloud fields represent model-based approximations rather than direct observations, reflecting patterns learned by the ML model. Additionally, using fixed thresholds in the object-based detection may oversimplify complex structures associated to clouds in the atmosphere. Nonetheless, we may employ the data to enable a large-scale, high-resolution tracking of convective systems over the tropical Atlantic and continental Africa.

3.2.1 Identifying cloud features

Radar reflectivity does not directly measure vertical air velocity, but it can serve as a valuable proxy for detecting hydrometeors associated with convective cloud development (Luo et al., 2008). To identify potential cloud structures, we apply a fixed threshold of –15 dBZ to distinguish signals of hydrometeors from background noise in the radar reflectivity data (Marchand et al., 2008). While this threshold is only moderately restrictive — allowing for the inclusion of short-lived or weak features — it is intentionally chosen to capture the full spatio-temporal evolution of convective clouds between development and dissipation stage (Esmaili et al., 2016).

The detection process begins by applying a Gaussian filter with a sigma value of 0.5 to smooth the input data and reduce noise (Kukulies et al., 2021). We then compute the centroid of each potential cloud using a weighted center-of-mass approach. Here, each point's weight is defined by its reflectivity value above the –15 dBZ threshold (Heikenfeld et al., 2019). These centroid positions are each assigned a unique identifier, which is maintained throughout the subsequent tracking and segmentation steps.

Next, we apply a 3D watershed segmentation algorithm to delineate the spatial extent of individual cloud structures associated with each centroid. In this approach, the 3D radar reflectivity field is interpreted as a topographic surface, where higher reflectivity values represent peaks and surrounding areas are segmented like catchment basins divided by ridges (Meyer, 1994). We initialise the algorithm by placing markers at the detected centroids in a binary 3D volume, where all other grid points are set to zero. From each marker, the algorithm expands through the volume, assigning reflectivity-based pixels to the corresponding cloud until the threshold of –15 dBZ is reached. The result is a labeled 3D cloud mask, where each pixel is either zero (indicating no cloud) or an integer label corresponding to a specific cloud object (Fiolleau and Roca, 2013). This mask allows

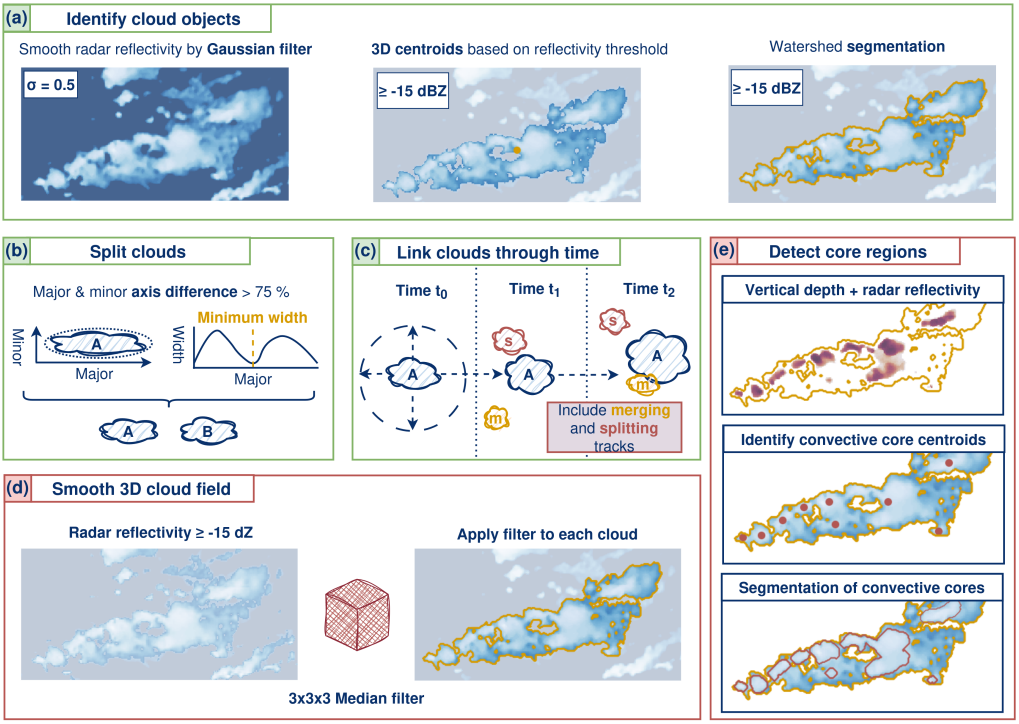

Figure 1. Workflow for tracking convective clouds **(a)**–**(c)** and their convective cores **(d)**–**(e)** using an object-based algorithm and ML-based 4D radar reflectivities. The routine consists of **(a)** the identification and segmentation of convective cloud centroids and volumes using a radar reflectivity threshold of -15 dBZ, **(b)** splitting shallow connected objects along their major axis, and **(c)** linking the labelled cloud objects through time. For detecting cores, we apply a Median filter with a kernel size of 3x3x3 pixels to smooth the 3D radar reflectivities associated to each cloud label **(d)**. We aim to identify convective core regions by adding **(e)** the number of pixels higher than 0 dBZ and the average cloud radar reflectivity in each vertical cloud column. This step provides a 2D layer of the combined radar reflectivity and potential core depth, which is used to search for local maxima displaying the centroid locations of convective cores. Subsequently, we apply a watershed segmentation to algorithm to derive the core area.

us to quantify the volumetric structure of clouds: the total number of labeled pixels per centroid corresponds to the cloud's volume rather than its area. For analyses requiring horizontal cloud coverage, we compute the 2D projected area by taking the column-wise maximum across vertical levels.

3.2.2 Split shallow connected clouds

After identifying the cloud centroids and their associated areas, we analyse the morphology of each cloud object to determine whether it may represent a merger of multiple cloud systems (Figure 1, b). To do this, we examine the labeled cloud mask to locate local minima in the cloud area, which may indicate potential split points between merged cloud structures. The shape of each cloud is characterised using the best-fitting ellipse (Ganetis et al., 2018). We then calculate the aspect ratio — i.e., the ratio of the major to minor axis lengths. If the major axis is more than 75 % longer than the minor axis, we classify the cloud as elongated (Cui et al., 2021). The orientation of the major axis provides the direction of elongation, which guides the search for potential split locations. Next, we examine the aggregated 2D cloud area along this direction and analyse the area distribution to detect change points. If the distribution is unimodal — featuring a single peak — we perform no split. However, if the distribution is multimodal, we apply a split at the local minimum, provided that this minimum deviates by more than 75 % from the mean size of the cloud shield. Then, we update the segmentation results by assigning a unique label to each newly separated cloud object.

3.2.3 Spatio-temporal linking

We track 3D cloud objects over time by linking them based on their estimated movement speed, following the method of Heikenfeld et al. (2019). The 3D perspective allows for detailed analysis of both horizontal and vertical cloud evolution—an essential aspect for understanding core development. At each 15-minute interval, we predict the expected location of a cloud object using its velocity from previous time steps (Figure 1, c). To streamline the linking process, we define a maximum search radius between time steps. Only cloud objects within this radius are considered potential matches, significantly reducing computational effort. When new clouds form, we assign them the average velocity of nearby clouds to estimate their likely movement (Sokolowsky et al., 2024). Due to computational limitations, we apply the linking algorithm to only two consecutive time steps at a time. We assume a successful link is established when a cloud object maintains a 15-minute temporal overlap and shares a consistent identifier across steps. For instance, we compare the cloud areas of linked objects, assuming that genuinely connected trajectories may exhibit more similar area changes than unrelated clouds (Prein et al., 2024). Finally, we evaluate the movement patterns before and after each time step to infer whether cloud objects may be merging or splitting and include the information in the final cloud trajectories.

3.3 Detect convective core regions

Convective clouds often contain one or more core regions, which are typically associated with stronger updrafts and intense precipitation that can penetrate above the freezing level. Because the formation and dissipation of these cores are closely linked to severe weather events, analysing their behavior is of particular interest (Takahashi et al., 2017).

To detect convective cores, we use the previously generated labelled 3D cloud mask (Section 3.2.1), derived from the ML–based radar reflectivity data. There are different approaches to identify convective cores from radar reflectivities. These methods may comprise the detection of convective precipitation, which may be associated to core regions in hydrometeors (Haynes et al., 2009; Pilewskie and L'Ecuyer, 2022) or the analysis of the radar reflectivity employing fixed thresholds along the vertical column (Luo et al., 2008; Bacmeister and Stephens, 2011; Igel et al., 2014). In this study, we focus on combining the latter with an object-based detection algorithm to identify centroids of convective cores in the predicted 3D radar reflectivity field. The approach is applied to each labelled cloud for each time step along the cloud trajectory (Figure 1, d,e). We begin by smoothing the radar reflectivity data associated with each cloud label using a 3×3×3 median filter. Core centroids are identified by locating local maxima in a combined metric that incorporates both smoothed radar reflectivity and the vertical extent of a contiguous potential core layer. Specifically, we calculate the mean radar reflectivity for each vertical cloud column, then determine the height of this core layer by counting the number of pixels with reflectivity values higher than 0 dBZ located at more than 5 km height. In our study, we do not include cores occurring lower than cloud base layer, where our model predictions may be less robust. We aim to fill isolated gaps for otherwise vertical continuous cores by expanding the threshold from 0 dBZ to –5 dBZ in columns that contain at least one pixel higher than 0 dBZ (Igel et al., 2014; Luo et al., 2008). We apply a minimum vertical extent of 5 km for a column to be considered part of a core; otherwise, its value is set to zero. The approach is visualised in Figure 1 (e). We add both indicators (average reflectivity and potential core vertical depth) for each pixel associated to a cloud label, resulting in a 2D layer in which we search for local maxima to serve as candidate core centroids. If no local maxima are found - e.g., in case no columns contain pixels higher than 0 dBZ at more than 5 km height - the cloud is recorded as having zero cores for that time step (Feng et al., 2023). When one or more core centroids are identified, we use a 3D watershed segmentation to delineate the core volumes. This process is repeated for every cloud object at each time step throughout its life-cycle, whereas a cloud may contain multiple cores at the same time.

3.4 Extract cloud and core properties

For each detected cloud trajectory, we extract both horizontal and vertical characteristics to describe the cloud and its internal structure (Table 3). Cloud properties include cloud area, CTH, CBH, duration, eccentricity, and the ratio of core to total cloud area. The cloud lifetime displays the total lifetime of the trajectory in hours, beginning from the first detection. Cloud area is calculated from the column-wise maximum of the 3D cloud mask, while vertical (CTH, CBH) metrics come from the number of pixels in the vertical column associated to each cloud label. Eccentricity is derived from the best-fitting ellipse, with values closer to 1 indicating a more circular shape (Cui et al., 2021). We also record the cloud's travel distance and assign a surface type using a binary land-sea mask and the modal value for the locations of the cloud trajectory within this land-sea mask.

225 The ratio of core to cloud area may provide a measure of convective compactness and intensity (Haberlie and Ashley, 2018). For clouds with one or more cores, we calculate the number, mean area, height, lifetime, eccentricity, and average distance between cores. The core area and height are derived from the column-wise maximum horizontal extent and vertical extent of the previously identified cores, similar to the cloud area and CTH. These metrics may help characterise the structural properties of detected cloud systems.

Table 3. Features used to describe the properties and life-cycle statistics of detected convective clouds and cores.

Feature type	Feature name	Definition
Cloud	Cloud area	Area of the cloud (km^2)
	Cloud top height (CTH)	Maximum height of the cloud (km)
	Cloud base height (CBH)	Minimum height of the cloud (km)
	Area ratio	Ratio between cloud area & core area
	Eccentricity	Roundness of the best fitting ellipse (cloud)
	Reflectivity	Average radar reflectivity of the cloud at 10 km height (dBZ)
	Location	Longitude and latitude of the cloud centroid (°)
	Travel distance	Euclidean distance for coordinates at initiation and dissipation (°)
	Cloud lifetime	Lifetime of the cloud trajectory (h)
	Surface type	Modal value from a binary land-sea mask for the cloud trajectory
Core	Number of cores	Number of identified convective core regions
	Core area	Average area of convective cores (km^2)
	Core vertical depth	Depth of the core in the vertical column (km)
	Mean distance	Average distance between cores in a cloud cluster (km)
	Core lifetime	Average lifetime of the cores (h)
	Core eccentricity	Roundness of the best fitting ellipse (core)
Life-cycle	Reflectivity gradient	Reflectivity change rate at 10 km height (dBZ)
	Area growth	Relative cloud area expansion (%)
	Vertical growth	Vertical growth of the cloud (km)

230 ### 3.5 Filter convective cloud trajectories

We filter the cloud trajectories to exclude possibly non-convective tracks from the analysis (Figure 2). For that purpose, we require the cloud tracks to have at least one core and a radar reflectivity of higher than 0 dBZ at 10 km height for at least 15 min along the trajectory. Additionally, we apply a minimum CTH of 10 km and a maximum CBH of lower than 5 km for the cloud during at least one time step (Igel et al., 2014; Luo et al., 2008). While we do not require the convective clouds to have a CTH

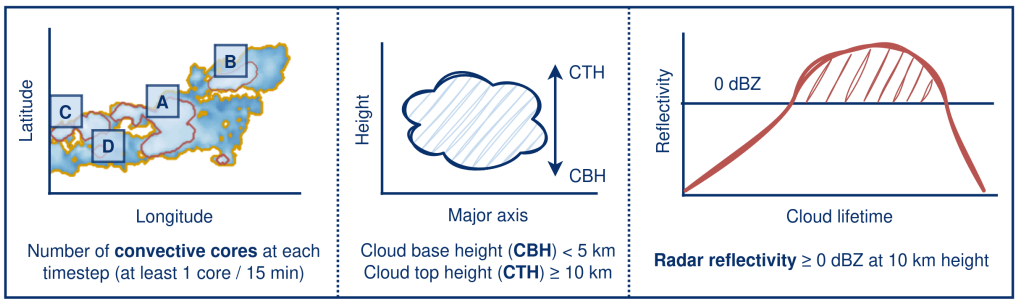

Figure 2. Criteria for filtering potentially convective cloud trajectories. The criteria consist of counting the number of convective cores (A–D) for each labelled cloud along the cloud lifetime. Moreover, we check the cloud base height (CBH) and cloud top height (CTH) of the cloud, and the radar reflectivity at 10 km height along the cloud trajectory. Following, we estimate the trajectory to belong to a convective cloud if we detect at least one core, a CBH lower than 5 km, a CTH higher 5 km, and a radar reflectivity higher than 0 dBZ at 10 km height for at least 15 minutes along the cloud lifetime.

higher than 10 km at every time step during their trajectory, we discard the trajectories that never reach the CTH threshold. The criteria may help to identify convective clouds with an evolved cloud base and vertical height that may be typically associated to tropical convection (Li et al., 2021; Takahashi et al., 2023).

3.6 Investigating the cloud life-cycle

We analyse the temporal evolution of detected clouds to explore how variations in the cloud life-cycle relate to the number of convective cores. For this purpose, we divide each cloud's life-cycle into three idealised stages, following the framework proposed by Futyan and Genio (2007). Each stage corresponds to distinct spatio-temporal changes in cloud structure, as simplified illustrated in Figure 3. The first time step of each trajectory marks the beginning of the development stage. Unlike methods that assess cloud stages using a cooling induced by temperature changes, the ML-based radar reflectivity does not provide information on temperatures. As an alternative, we approximate the life-cycle using temporal changes in radar reflectivity at 10 km height and the resulting vertical and horizontal cloud characteristics. For estimating the vertical growth of the cloud, we compute the difference between CTH and CBH (i.e., to display the height of the cloud layer) for every point in time. For the horizontal growth of the cloud, we calculate changes of the cloud area derived as proportional differences to the cloud area at the first timestep of detection.

- *Development stage*: Building on the approach by Luo et al. (2008), we use a radar reflectivity threshold of 0 dBZ at 10 km altitude as a proxy for potential cloud-top cooling, which may be indicative of convective growth. We calculate the temporal gradient of radar reflectivity at 10 km for each cloud trajectory, identifying the time of maximum increase to mark the cloud development stage. This stage may be associated with a high cloud vertical layer and strong updrafts

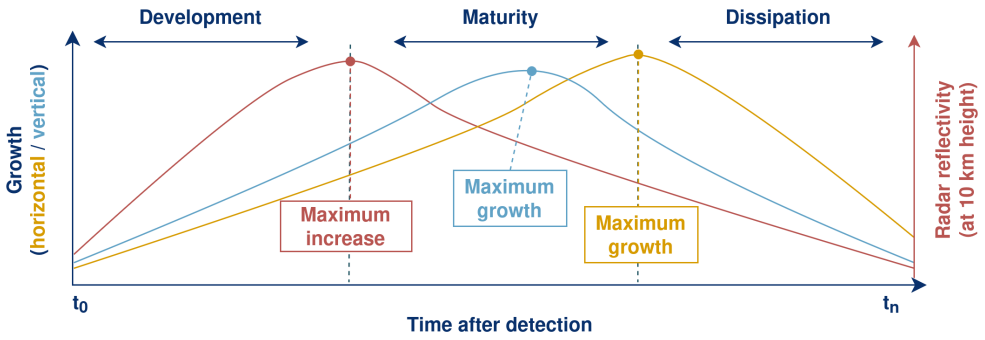

Figure 3. Schematic visualisation of the three stages (development, maturity, and dissipation) of the idealised convective life-cycle. Here, we show how changes of the radar reflectivity at 10 km height, the horizontal growth of the cloud area, and the vertical growth of the cloud vertical layer may be connected to the transition between life-cycle stages.

that support continued vertical growth (Kikuchi and Suzuki, 2019; Chen et al., 2021).The transition from development to maturity is defined by the time of maximum radar reflectivity increase (Takahashi et al., 2023; Hu et al., 2021).

- *Maturity stage*: Following the time of maximum radar reflectivity, the reflectivity gradient at 10 km height may gradually decrease. Instead, both the vertical thickness and horizontal extent of the cloud typically may increase in the maturity stage, indicating a sustained growth of the cloud (Gupta et al., 2024).

- *Dissipation stage*: Dissipation begins when vertical growth slows and the cloud reaches its maximum horizontal size. We continue tracking the cloud until the reflectivity falls below the -15 dBZ threshold and no centroids are identified during feature detection, indicating cloud decay (Crook et al., 2019).

For each trajectory, we determine key time points that may approximate changes in the cloud life-cycle: the moment of maximum reflectivity gradient at 10 km, peak area (horizontal) growth, maximum vertical extent, and the onset of dissipation (Table 3). We also record when the highest number of cores is detected and when cores reach their maximum area. These markers are used to compare life-cycle characteristics between clouds with a single core and convective systems containing multiple cores (i.e., more than one core). We note that the life-cycle statistics derived for this study are based on the ML-based radar reflectivities and inherit the uncertainties connected to these predictions. Hence, they may only provide an approximation of distinct changes occurring within the cloud trajectories over time.

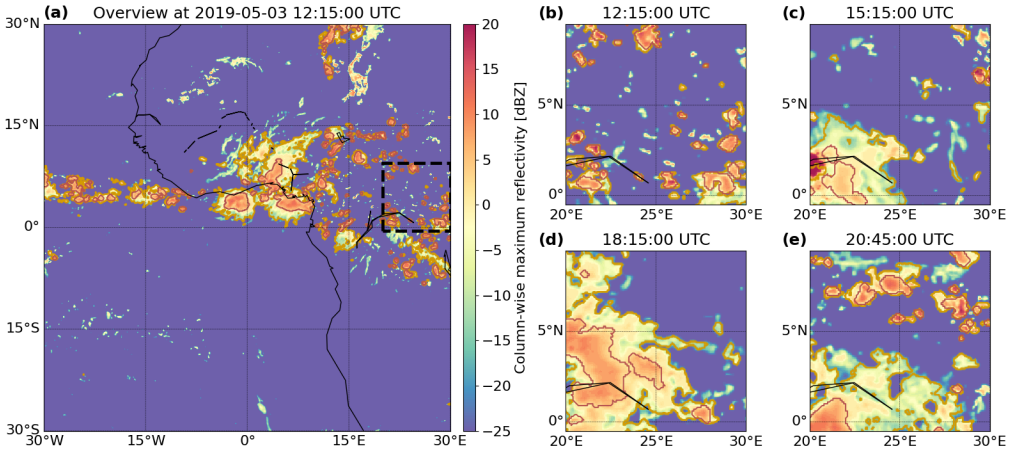

Figure 4. An example of convective clouds (orange outline) and their cores (red outline) detected in the ML-based 3D cloud field on the 03.05.2019, 12:15:00 UTC. The cloud mask is plotted over the 3D radar reflectivity, which shows the aggregated column-wise maximum. All times are given in UTC. In **(a)**, we see an overview of the AOI, **(b)** to **(e)** show a zoomed perspective in 3 h intervals (black square).

4 Results

4.1 Distribution of convective cloud and core properties

Between March and August 2019, we detected approximately 375,000 convective clouds using ML-based 3D radar reflectivity data. After excluding tracks with a lifetime of only one timestep, 338,142 cloud trajectories remained for analysis. Figure 4 shows an example from May 3, 2019, at 12:15 UTC, highlighting tracked convective clouds and cores. While regions over Morocco and Mauritania showed radar reflectivity higher than 0 dBZ, no vertically continuous convective systems were identified there. Instead, numerous convective clouds appeared over the Gulf of Guinea, the equatorial rainforest, and the Atlantic Ocean. Figures 4 (b)–(e) illustrate the development and dissipation of cores, often lasting only a short time. Some clusters contained multiple cores, potentially indicating mesoscale convective systems (MCSs) (Takahashi et al., 2017).

Our 3D framework allows us to simultaneously track horizontal and vertical cloud development. For core statistics, we separate the core region from the anvil cloud, as shown in Figure 5. Clouds are grouped by core count to distinguish potentially more isolated systems (one core) from clustered systems (multiple cores). For statistical purposes, clouds with 6–9 cores and those with 10 or more cores are combined into respective categories (Jones et al., 2024).

Single-core clouds make up roughly 65 % of all trajectories (Figure 6, a), with the frequency decreasing as core count increases. Only about 5 % of clouds have 10 or more cores. Most clouds (80 %) have lifespans between 0–6 hours (Figure 6, b). Surface type distribution reveals that 65 % of clouds form over the ocean and 35 % over land—about a 10 % shift toward ocean compared to land-sea coverage (Figure 6, c). Among single-core clouds, 70 % occur over the ocean, while for multi-core clouds, the figure is 75 %. This imbalance — 249,484 oceanic clouds vs. 88,658 continental — may reflect differences

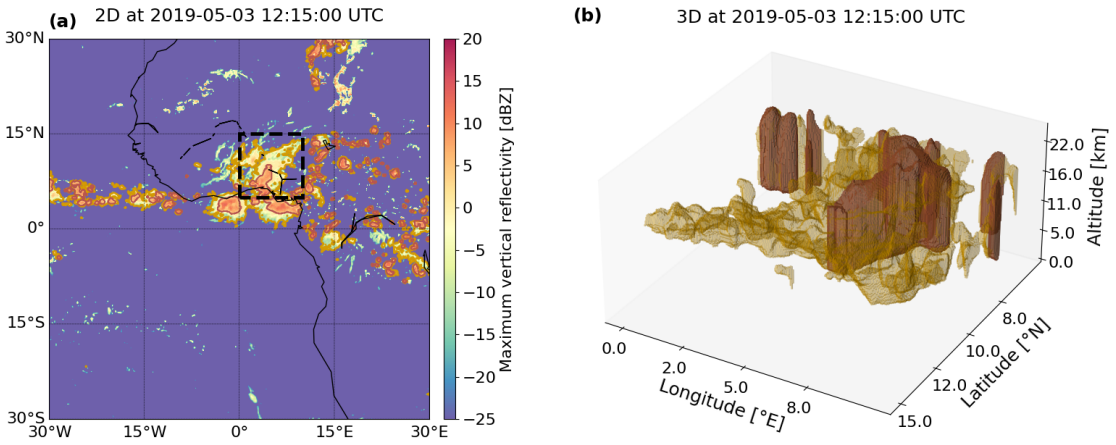

Figure 5. An example of the convective clouds (orange outline) and their cores (red outline) detected in the predicted 3D cloud field on the 03.05.2019, 12:15:00 UTC. The cloud mask is plotted over the 3D radar reflectivity, which shows the aggregated column-wise maximum. In **(a)**, we see an overview of the AOI, **(b)** shows the zoomed perspective (black square) in 3D for the cloud volume (orange) and core volume (red).

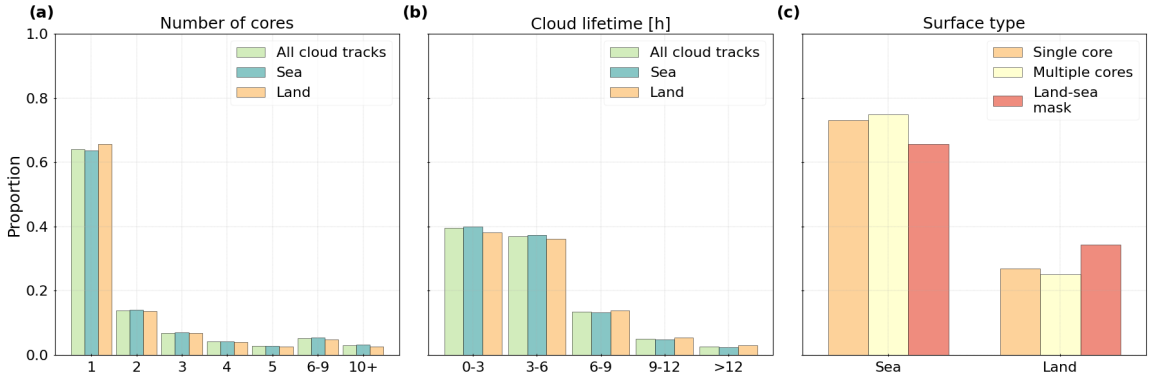

Figure 6. Distribution of **(a)** the number of associated cores, **(b)** the average cloud lifetime, and **(c)** the surface type derived from a land-sea mask compared to the modal locations of detected clouds with a single core or multiple cores. We show the distribution in **(a)** and **(b)** for all cloud tracks (n = 338,142), clouds over the ocean (n = 249,484), and clouds over land (n = 88,658).

in tropical landmass distribution and the eastward propagation of convective systems. Oceans may also offer more favorable conditions for multi-core development (Cui et al., 2021).

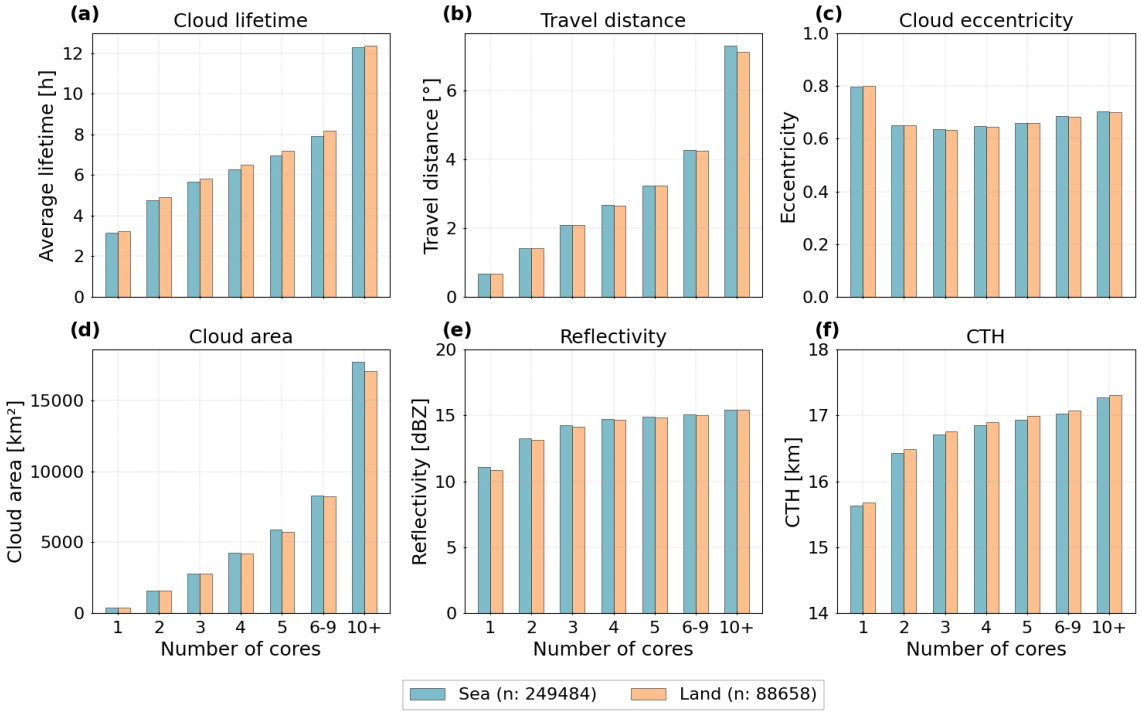

Figure 7. Distribution of cloud statistics grouped by the number of associated cores for **(a)** the cloud lifetime, **(b)** the travel distance between first and last detection, **(c)** the cloud eccentricity, **(d)** the cloud area, **(e)** the radar reflectivity at 10 km height, and **(f)** the CTH.

We assess how the 3D cloud properties described in Table 3 may vary with core count and surface type. Our findings show that single-core clouds have shorter lifetimes and travel distances than multi-core systems (Figure 7, a–b). Eccentricity exhibits a weak variation across all groups, mostly ranging between 0.6–0.7 (Figure 7, c). Cloud area increases significantly with core count, especially for clouds with 10 and more cores (Figure 7, d). CTH is 10–20 % greater over land, yet radar reflectivity at 10 km height and cloud area are slightly higher over the ocean (Figure 7, d–f). CTH increases from 15.5 km for single-core clouds to 17.25 km for multi-core ones (Figure 7, f). Land–sea differences are more pronounced for single-core clouds. Despite expectations based on previous tropical studies (Deng et al., 2016; Takahashi et al., 2017), oceanic clouds often show stronger reflectivity and larger areas — though overall surface-related differences remain small. The lower number of land-based clouds may exaggerate statistical noise.

The analysis of core properties (Figure 8) shows average core lifetimes of 0.3–0.4 hours for single-core clouds, increasing to about 0.8 hours for clouds with more than 10 cores (Figure 8, a). Core eccentricity shows little variability and ranges from 0.5–0.6 (Figure 8, b). Core area is slightly larger for single-core clouds than for those with 2–9 cores but increases considerably

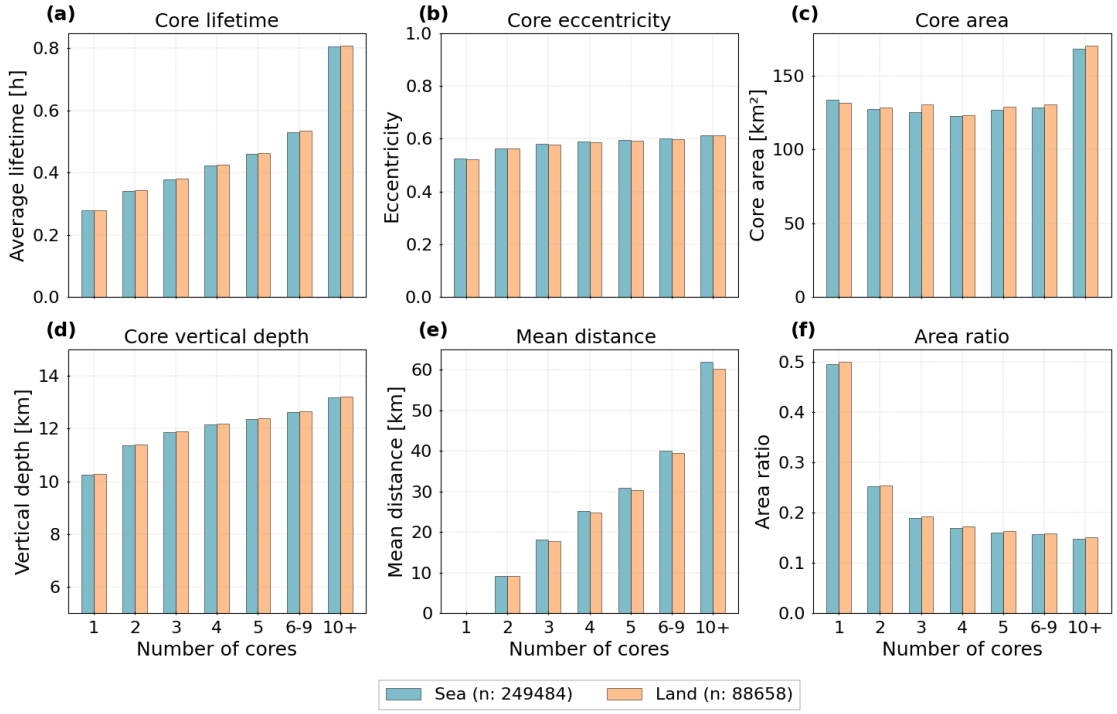

Figure 8. Distribution of core statistics grouped by the number of associated cores for **(a)** the core lifetime, **(b)** the core eccentricity, **(c)** the core area, **(d)** the vertical depth of the core, **(e)** the mean distance between individual cores, and **(f)** the area ratio between the cloud and the core.

for clouds with 10 and more cores. For single-core clouds, we detect a larger core area over the ocean, while cores for multi-core clouds are larger over land (Figure 8, c). Core height and distance between cores both increase with core count (Figure 8, d–e). The largest distances, which may indicate the least compact core morphology, occur for clouds with 10 and more cores (Figure 8, e). The area ratio between clouds and cores is highest for single-core clouds and declines with more cores. The sharp decrease may be connected to a faster increase of the cloud area compared to core area for multi-core clouds (Figure 7, d, Figure 8, f).

In summary, clouds with more cores exhibit longer lifetimes, larger areas, greater heights, and increased core size and distances between cores. Core properties are broadly consistent across surface types, except for core area and core distances (Figures 7a, 8c). However, these findings may be influenced by the land–sea imbalance in our dataset and inter-annual variability.

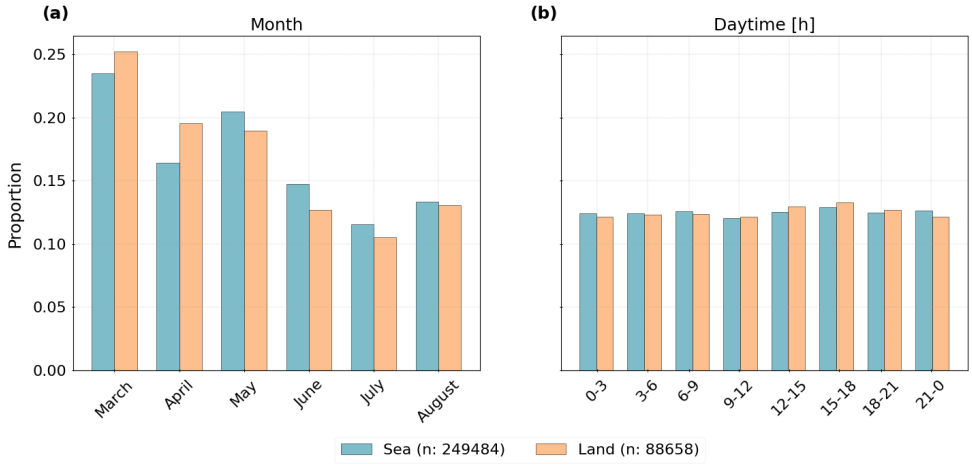

Figure 9. Distribution of cloud tracks over land and sea for **(a)** the months between March–August, and **(b)** the daytime of the first detection.

4.2 Temporal characteristics of convective clouds

We observe a higher proportion of cloud tracks between March to May, and notably less tracks between June to August. Over land, we find more clouds in March and April. Over the ocean, the proportion of cloud tracks is higher between May and August (Figure 9, a). For the diurnal distribution, we observe slightly more tracks between 12:00 to 18:00 UTC than at other time intervals. During nighttime (21:00–06:00 UTC), more cloud tracks occur over the ocean. In contrast, we detect a higher proportion of cloud occurrences over land during the day. However, the variability connected to the diurnal cycle and surface type remain weak (Figure 9, b).

4.2.1 Diurnal cycle over land and sea

We analyse the diurnal cycle of cloud properties for single-core and multi-core clouds over land and ocean by computing a 2D density distribution displaying hourly changes of the cloud properties. Figure 10 illustrates these variations in cloud lifetime (a–d), cloud area (e–h), and radar reflectivity at 10 km height (i–l). Over land, single-core clouds show an afternoon peak (12:00–16:00 UTC) in both radar reflectivity and cloud area, while cloud lifetime displays two peaks: one at night and one in the morning (Figure 10, b, d). Over the ocean, the diurnal cycle is weaker or less distinct. Cloud lifetime lacks a clear diurnal peak (Figure 10, a, c), whereas cloud area and reflectivity show nocturnal and daytime peaks (Figure 10, e, i, k). Despite similar diurnal patterns for the cloud lifetime and radar reflectivity, multi-core clouds consistently exhibit higher mean values than single-core clouds. These differences may reflect environmental contrasts between land and ocean. As suggested by Cui et al. (2021), local circulations over land in the tropics often trigger afternoon convection, producing the observed peaks in

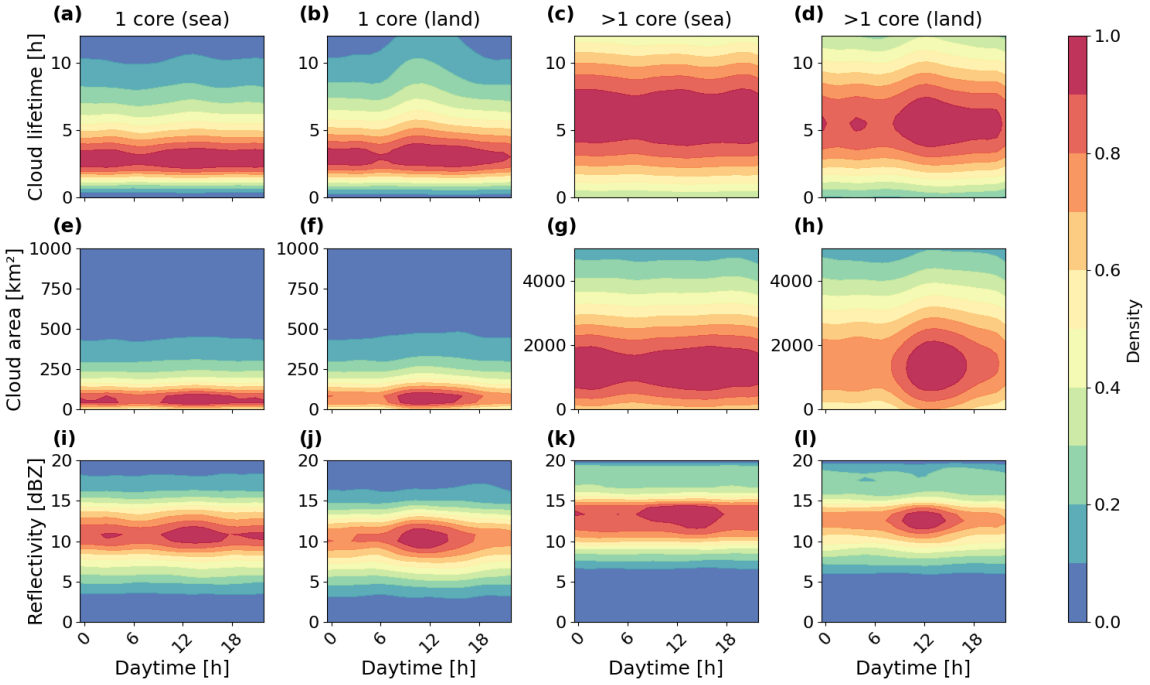

Figure 10. Diurnal cycle for cloud properties grouped by the number of associated cores and surface type. We display the hourly changes (in UTC) regarding **(a)**–**(d)** the cloud lifetime, **(e)**–**(h)** the cloud area, and **(i)**–**(l)** the radar reflectivity at 10 km height for single- (1 core) and multi-core (more than one core) clouds over sea and land. The values show the density distribution normalized between 0 and 1.

Figure 10 (f), (h), and (j). In contrast, more constant ocean temperatures may suppress strong diurnal variations (Figure 10, a, c, g).

The diurnal patterns of core properties (Figure 11) largely mirror those of the cloud properties. Over land, core area peaks between 12:00–18:00 UTC for both single- and multi-core clouds. Over the ocean, single-core clouds show two peaks between 00:00–06:00 and 14:00–20:00 UTC (Figure 11, a–d). The core lifetime follows a similar pattern for single-core clouds. For multi-core clouds, cores over land show two peaks, while oceanic cores point out no clear diurnal variation for multi-core clouds (Figure 11, e–h). For single-core clouds, peaks of the core lifetime resemble the core area (Figure 11, a, e). The distribution of the core height follows those of the core area (Figure 11, m–p). On average, clouds with multiple cores have higher and more variable values for core area, lifetime, and height. In contrast, the area ratio is lower and has a weaker variability for multi-core systems. For single-core clouds, we observe an afternoon peak over land and nocturnal and afternoon peaks over the ocean. Multi-core clouds show a weak diurnal variation, particularly over the ocean (Figure 11, i–l).

Overall, the diurnal cycle highlights a pronounced afternoon peak over land and a two peak, at nighttime and in the afternoon, over the ocean. These patterns align with observed differences in tropical convective behavior over land and sea (Vondou, 2012).

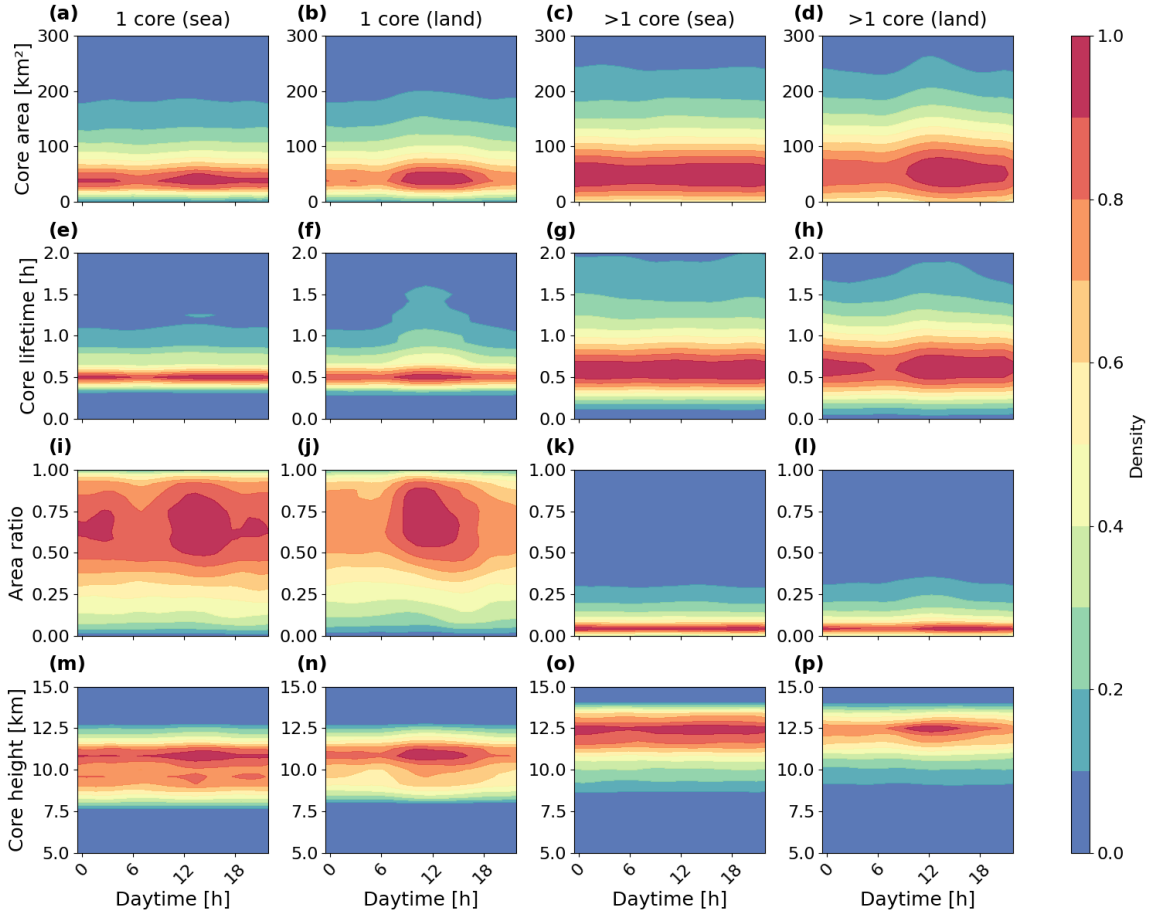

Figure 11. Diurnal cycle for core properties grouped by the number of associated cores and surface type. We display the hourly changes (in UTC) regarding **(a)**–**(d)** the core area, **(e)**–**(h)** the core lifetime, **(i)**–**(l)** the core eccentricity, and **(m)**-**(p)** the area ratio between the cloud core and anvil area for single- (1 core) and multi-core (more than one core) clouds over sea and land. The values show the density distribution normalized between 0 and 1.

While we find those patterns mostly for single-core clouds, results for multi-core clouds — especially over the ocean - are less distinct.

4.2.2 Monthly variability of convective properties

For different months, the value variability may considerably influence the development of convective clouds and their core structures within the tropics (Andrews et al., 2024). We explore these changes in Figure 12 by comparing monthly averages of

the cloud area, CTH, cloud lifetime, number of cores, core area, and area ratio over land and sea for single-core and multi-core clouds.

From March to August, the cloud area shows a gradual increase for single- and multi-core cloud systems over the ocean. For clouds over land, the cloud area slightly decreases (Figure 12, a). In contrast, CTH generally decreases, though month-to-month fluctuations appear to be higher than a consistent decrease (Figure 12, b). Cloud lifetime shows a higher variability between months with increases in April for single- and multi-core clouds and in April and June for single-core clouds. Overall, clouds with multiple cores exhibit a slight decline in lifetime over land. Over the ocean, lifetime rises from March to April, decreases in May, and increases again in June — returning to near-initial values by August (Figure 12, c). The number of cores per cloud increases over the ocean from March to July, followed by a sharp drop in August. Initially higher over land, core counts shift in favor of oceanic clouds after April. Convective systems over land show a decrease of core numbers from March to June, a peak in July, and another decline in August (Figure 12, d). The core area steadily increases over the ocean but fluctuates more over land (Figure 12, e). The area ratio shows a slight decrease for multi-core clouds throughout the period, while remaining higher for single-core clouds. We observe a high monthly variability over land and ocean, whereas the area ratio remains to be higher over continental Africa (Figure 12, f).

To quantify the effect of these changes, we compare average values across two periods: March–May (MAM) and June–August (JJA). Metrics include the cloud area, CTH, cloud lifetime, number of cores, core area, and area ratio (Table 4). We calculate Cohen's D to measure effect sizes, with thresholds defined as small (< 0.2), medium (0.2–0.5), and large (> 0.8) (Cohen, 2013). Over the ocean, cloud area, number of cores, and core area are higher in JJA, while CTH, cloud lifetime, and area ratio are greater in MAM. A similar pattern emerges over land, except cloud area and number of cores are higher in MAM. Overall, observed differences between the two seasons and over land and sea remain weak. Most effect sizes are small, indicating high internal variability rather than distinct temporal trends within the period. These results highlight the importance of analysing longer time periods to account for the inherent variability and imbalance between cloud tracks over land and sea (Figures 4 and 9), which may influence the representativeness of the findings.

4.3 Impact of convective cores on the cloud life-cycle

4.3.1 Relationship between life-cycle statistics and cloud properties

To analyse the cloud life-cycle (as outlined in Section 3.6), we check the point of time when three key events occur in each cloud trajectory: the maximum radar reflectivity gradient at 10 km altitude ("reflectivity gradient"), the maximum cloud area growth ("area growth"), and the maximum vertical growth ("vertical growth"). Figure 13 shows the distribution of these indicators grouped by the surface type and number of cores. The average maximum reflectivity gradient ranges from 10 to 16 dBZ. Clouds with 2–3 cores show the highest gradients (14.5–16 dBZ), while the gradient for single-core clouds averages around 14 dBZ. It decreases with further increasing core count, dropping to around 10 dBZ for clouds with 10 or more cores. Surface type has little impact overall, though values are slightly higher over the ocean for clouds with 1–3 cores (Figure 13, a). In contrast, cloud area growth is slightly higher over land. More important, clouds with multiple cores grow considerably more

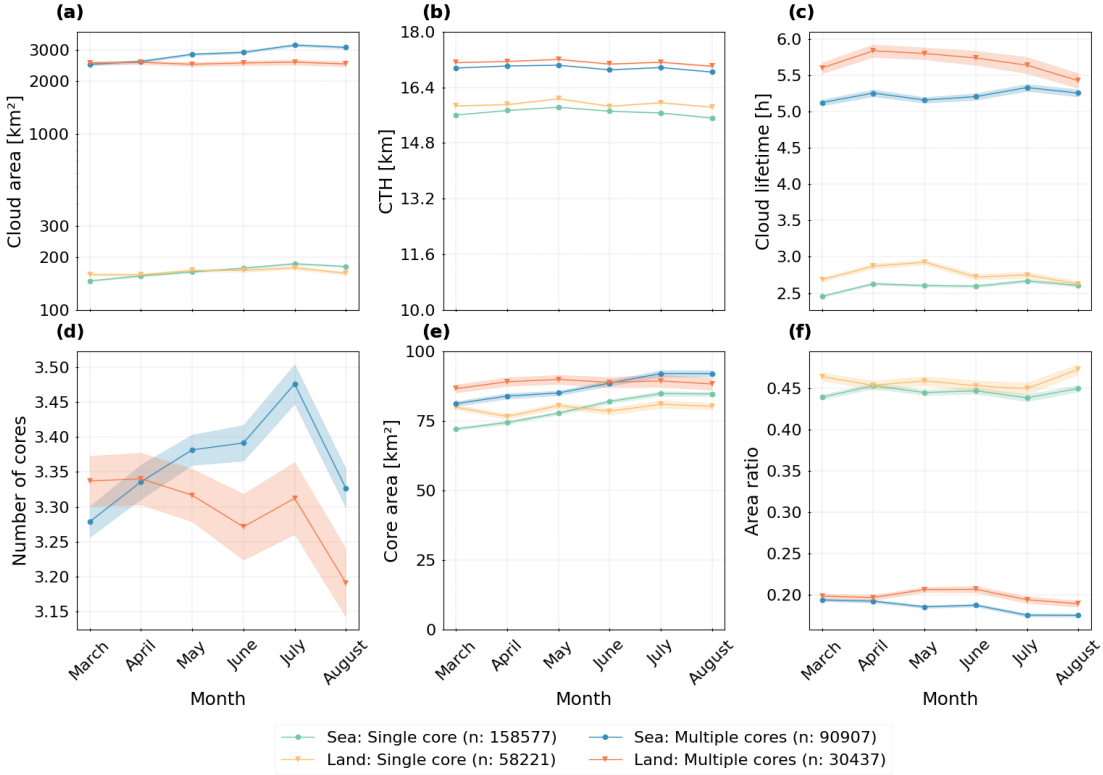

Figure 12. Monthly variability of cloud and core statistics between March and August for single-core and multi-core clouds grouped by the surface type for **(a)** the cloud area, **(b)** the CTH, **(c)** the cloud lifetime, **(d)** the number of cores (only multi-core clouds), **(e)** the core area, and **(f)** the area ratio between the cloud and the core. Line plots show the mean value with a confidence interval of 95 %.

in area than single-core clouds - ranging from 22 % (single-core) to 52 % (10 and more cores) (Figure 13, b). For the vertical growth, we observe average values between 5–8 km. Single-core clouds tend to grow higher than multi-core clouds, with only minor differences between land and sea (Figure 13, c).

We use Spearman's rank correlation coefficent R to examine relationships between life-cycle metrics and cloud/core properties for all cloud tracks, single-core, and multi-core clouds. Overall, cloud and core properties show predominantly positive correlations. The strongest correlation across all datasets appears between CTH and core height. Additional strong correlations include the number of cores with the cloud area and the core height. In contrast, cloud lifetime shows weak to moderate correlations (Figure 14, a). For single-core clouds, the correlation between cloud area and core height weakens, while the link between core area and cloud area strengthens (Figure 14, b). Multi-core clouds exhibit similar patterns to all cloud tracks, though with slightly weaker correlations (Figure 14, c). The number of cores correlates positively with area growth but negatively with both

Table 4. Comparison of mean values for cloud (cloud area, CTH, cloud lifetime) and core (number of cores, core area, area ratio) properties between March–May (MAM) and June–August (JJA) grouped by the surface type (Sea, Land). We calculate the effect size measured by Cohen's D to assess the difference between distributions over sea and land.

	March–May			June–August		
	Sea	Land	Cohen's D	Sea	Land	Cohen's D
Cloud area	1605.673	1742.597	0.017	2634.334	1681.849	0.107
CTH	16.061	16.229	0.145	15.982	16.147	0.138
Cloud lifetime	4.327	4.900	0.01	4.356	4.231	0.003
Number of cores	2.128	2.166	0.009	2.459	2.062	0.136
Core area	125.118	128.249	0.013	143.647	131.638	0.035
Area ratio	0.398	0.408	0.032	0.374	0.407	0.115

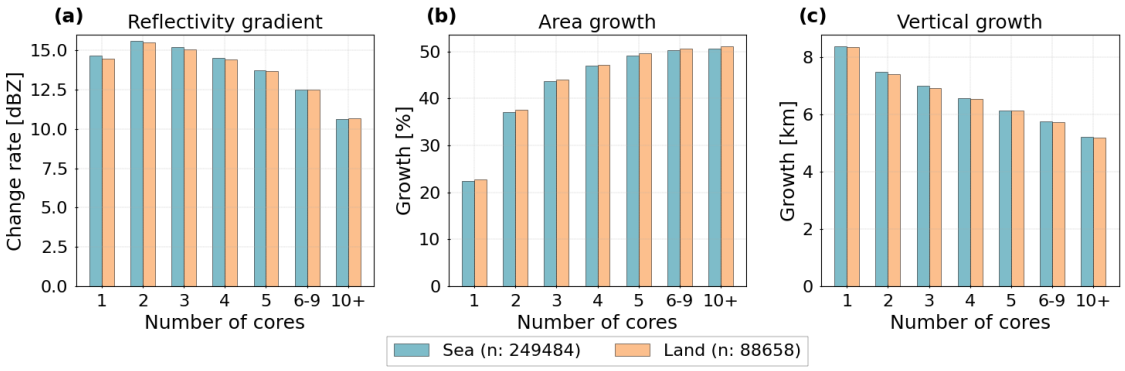

Figure 13. Life-cycle statistics grouped by the surface type and the number of associated cores for **(a)** the radar reflectivity gradient at 10 km height, **(b)** the relative area growth, and **(c)** the vertical growth of the cloud.

reflectivity gradient and vertical growth. Reflectivity gradient shows a weak positive correlation with area growth (< 0.25) and a stronger one with vertical growth (up to 0.5 for multi-core clouds). Area growth and vertical growth are negatively correlated, with medium to strong correlations. The relationship between the reflectivity gradient and other cloud properties is weak, and its direction differs by cloud type — positive for single-core, negative for multi-core clouds. Area growth correlates positively with cloud and core properties, especially the cloud lifetime. Vertical growth shows a negative correlation with cloud lifetime and mixed, generally weak correlations with cloud area, CTH, and core metrics.

Figure 15 presents the average timing (post-detection) of life-cycle statistics, grouped by surface type and core count. Here, we compute the average time after detection when the clouds reach their maximum for the reflectivity gradient, anvil growth,

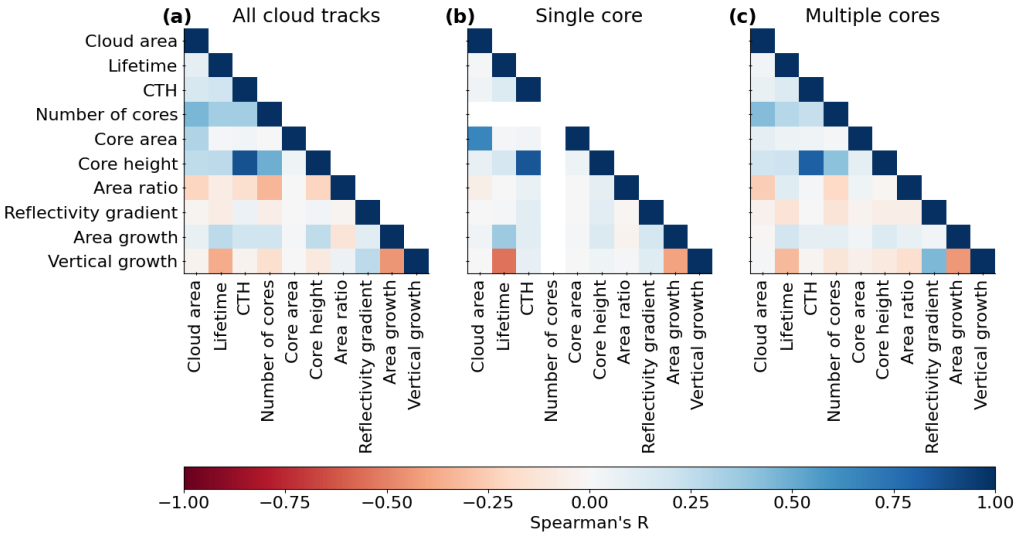

Figure 14. Correlation matrix for the life-cycle statistics, cloud and core properties. We calculate Spearman's R to quantify the correlation coefficient on a scale between -1 and 1 for **(a)** the whole dataset (n = 338,142), **(b)** clouds with a single core (n = 216,798), and **(c)** clouds with multiple cores (n = 121,344).

vertical growth, number of cores, and core area. Moreover, we derive the time of dissipation which marks the end of the cloud life-cycle.

- Reflectivity gradient: Peaks around 0.85 h on average, with little difference for the median between land and sea. However, the distribution broadens over both surface types with increasing core numbers. The arithmetic mean increases up to 2.5–3 h for clouds with 10 and more cores (Figure 15, a).
- Area growth: Occurs at 1.64 h on average. Single-core clouds peak earlier (< 1.5 h), while multi-core clouds range from 1.7 h (2 cores) to 4 h (10 and more cores). For clouds with 3–5 cores, we find a predominantly similar distribution with only slight variations of the arithmetic mean (Figure 15, b).
- Vertical growth: Peaks around 1.19 h. This occurs earlier for single-core and oceanic clouds. For clouds with 4 (5) or more cores over land (sea), the distributions are similar and show only a weak variability (Figure 15, c).
- Core area & core number: The maximum core area typically occurs 1.25 h after detection, while the core number peaks slightly later at 1.37 h. Surface type has little effect on the time of the maximum core area, though the timing increases with more cores (Figure 15, d). Clouds over land - especially those with 2–5 cores - reach their maximum number of cores later than oceanic clouds. Despite this observation, we find less of a linear timing increase compared to other life-cycle statistics (Figure 15, e).

– Cloud dissipation: Clouds over land generally dissipate later than oceanic ones. Lifetime extends further with more cores, and value variability is also higher over land. Single-core clouds typically dissipate within 1–3 hours, whereas multi-core clouds last longer, with broader distributions and higher mean lifetimes (Figure 15, f).

Notably, the analysis shows that vertical growth may peak after the reflectivity gradient but before area growth. The times stretch for multi-core clouds, while single-core systems exhibit more compact timelines. However, outliers may distort observed mean values. Hence, the consecutive order of the life-cycle statistics may be affected by a high variability in the distribution. Across all cloud tracks, core number and core area tend to peak between vertical and area growth maxima. However, the distributions show a high variability, especially for multi-core clouds and clouds over land. While we observe life-cycle statistics occurring on average later for clouds over land, the differences induced by the surface type remain overall low.

4.3.2 Temporal variability of life-cycle statistics

Figure 16 illustrates the diurnal patterns of the reflectivity gradient, the area growth, and the vertical growth for single-core and multi-core clouds, grouped by surface type. Similar to results from Sect. 4.2.1, the diurnal cycle is more pronounced over land than over the ocean. The reflectivity gradient exhibits short-term fluctuations with noticeable nocturnal peaks around 20:00–21:00 UTC and 00:00–01:00 UTC, followed by decreases. During the day, peaks occur between 09:00–12:00 UTC and around 16:00 UTC, with a negative dip around noon (land) and 18:00 UTC (both land and sea). Over the ocean, the reflectivity gradient is generally higher and shows a slightly weaker variability than over land. Over land, multi-core clouds exhibit stronger gradients than single-core clouds. Distinct land-based negative peaks occur around 03:00–06:00, 08:00, and 11:00–15:00 UTC. Over the ocean, we find a weaker nocturnal peak at 01:00 UTC and a gradual increase between 06:00–20:00 UTC. Overall diurnal variability ranges from 0.5 dBZ (ocean) to 1 dBZ (land), or roughly 8–16 % of the mean gradient range (10–16 dBZ) (Figure 16, a, d). Multi-core clouds show significantly greater area growth (42–48 %) than single-core clouds (20–27 %). Over the ocean, we see sporadic daytime peaks that occur around 15:00, 20:00, and 01:00 UTC. Over land, area growth increases steadily for single-core clouds in the morning and peaks between 12:00–14:00 UTC. For multi-core clouds, we find several sporadic peaks during the day, similar to clouds over the ocean. Evening peaks appear around 18:00 and 22:00 UTC for multi-core clouds, and 20:00 UTC for single-core clouds. Diurnal variability remains low, ranging up to 5 % over both land and sea (Figure 16, b, e). Diurnal variation in vertical growth is weak over the ocean with values below 0.5 km. Over land, morning peaks are evident, particularly for single-core clouds, while all cloud types show a distinct dip around noon. Afternoon and evening values rise again, with multiple peaks (e.g., 17:00, 21:00, 06:00 UTC) and troughs (18:00, 22:00, 03:00 UTC), typically within a 1–1.5 km variability range. Over both surface types, single-core clouds reach higher altitudes (7.5–8.75 km) than multi-core clouds (6.25–7.5 km) (Figure 16, c, f).

Figure 17 presents monthly changes for the reflectivity gradient, the area growth, and the vertical growth between March and August 2019 for single- and multi-core clouds over land and sea. Overall, reflectivity gradients increase from March to August. For single-core clouds, the increase is more pronounced over the ocean (from 14.3 to 15.8 dBZ) than over land (13.5 to 14.25 dBZ). For multi-core clouds, values over land rise more (14.3 to 14.6 dBZ) than over the ocean (14.5 to 14.6 dBZ).

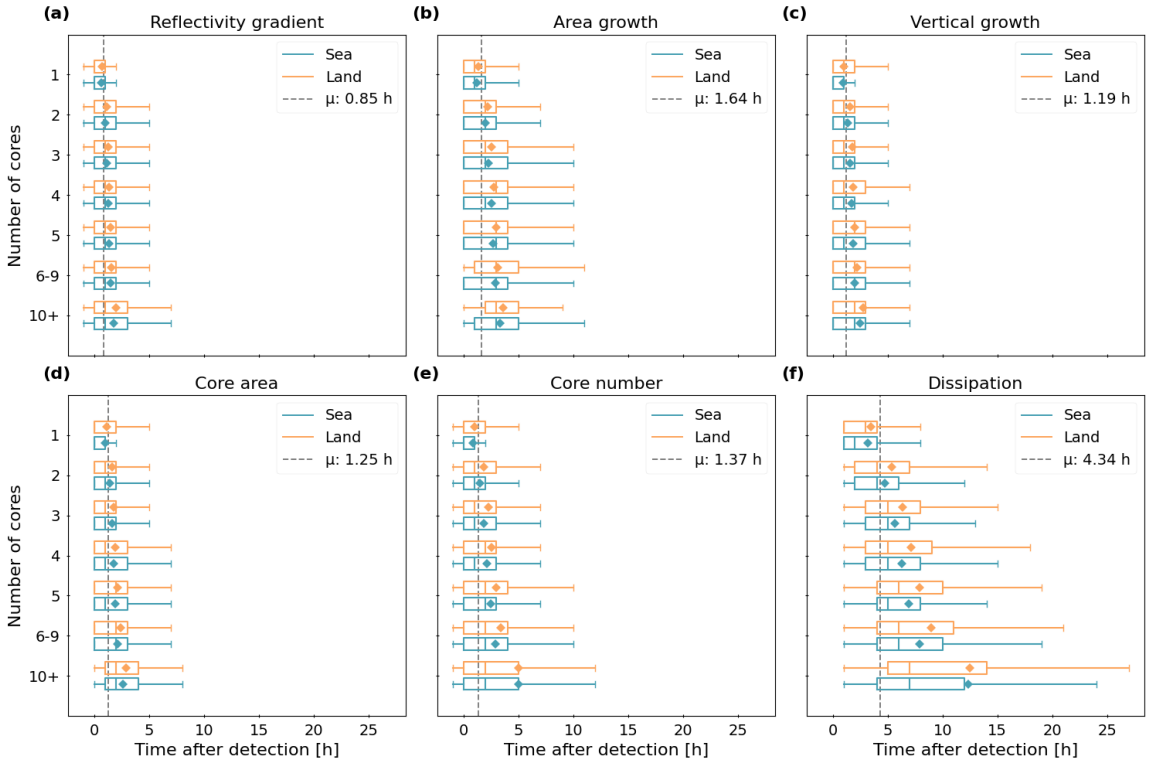

Figure 15. Distribution of the time dependency (x-axis) for the life-cycle statistics. The boxplot diagrams show the time after detection (in h) when the detected clouds reach on average the maximum value for **(a)** the radar reflectivity gradient at 10 km height, **(b)** the vertical growth, **(c)** the anvil growth, **(d)** the core area, and **(e)** the core number. Moreover, we show the average **(f)** dissipation time. Clouds are grouped by the surface type and number of cores (y-axis). The boxplot contains the median (bold blue or orange lines) and the arithmetic mean (blue or orange diamonds). The grey vertical lines show the mean time dependency averaged over all clouds tracks.

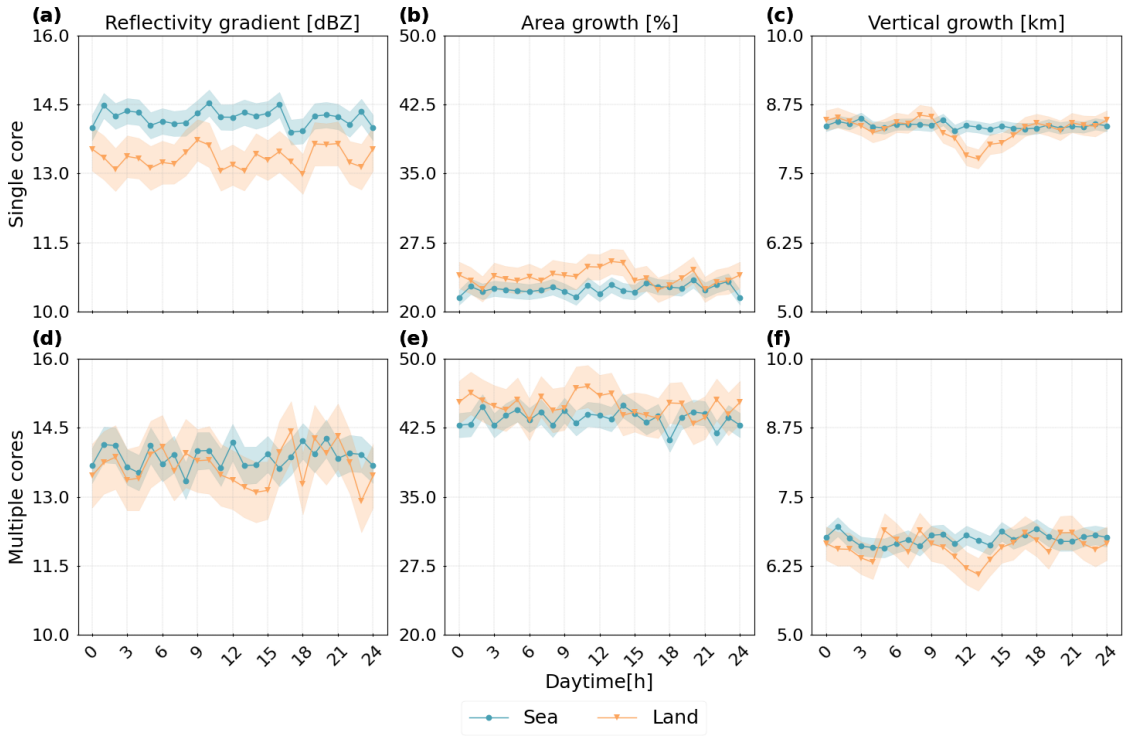

Figure 16. Diurnal cycle for the life-cycle statistics. We display the hourly changes (in UTC) regarding **(a)** & **(d)** the maximum cooling, **(b)** & **(e)** the cloud area growth, and **(c)** & **(f)** the cloud vertical growth for single- (1 core) and multi-core (>1 core) clouds over sea and land. Line plots show the mean value with a confidence interval of 95 %.

A notable dip occurs across all cloud types in April. Month-to-month variability is high (Figure 17, a, d). Between March and June, area growth is higher over land, peaking in May for single-core and in June for multi-core clouds. After June, area growth becomes higher over the ocean. Over the period, single-core clouds over land show a net decline (around 3 %), while values increase slightly over the ocean (around 5 %) and for multi-core clouds over land (around 3 %). Monthly changes appear to be nonlinear and fluctuate considerably (Figure 17, b, e). The vertical growth of single-core clouds peaks in March, drops sharply in April. From May onward, it rises again over land and decreases over the ocean. The total change ranges between 0.5–1 km. Early in the period, oceanic clouds have a higher vertical growth than clouds over land. From June onward, this pattern reverses (Figure 17, c, f). Throughout the study period, single-core clouds consistently show higher reflectivity gradients and vertical growth, while multi-core clouds exhibit greater area growth. Though the surface type may influence these statistics, the observed effect in our study remains relatively small. In contrast, the number of cores plays a more substantial role in shaping the cloud life-cycle.

Figure 17. Monthly changes for the life-cycle statistics. In **(a)** & **(d)**, we show the maximum cooling, in **(b)** & **(e)** the anvil area growth, and in **(c)** & **(f)** the vertical growth for clouds with a single (1 core) core or multiple (>1 core) cores. Line plots show the mean value with a confidence interval of 95 %.

5 Discussion

5.1 Summary of results

In this study, we use a ML–based 3D extrapolation of radar reflectivity, derived from 2D satellite imagery, to characterise single- and multi-core convective clouds in the tropics. Our analysis shows that the majority of detected clouds are short-lived (under three hours) and dominated by a single core (65 %). Compared to multi-core clouds, these single-core systems tend to have a smaller cloud area, lower reflectivity at 10 km, and reduced CTH. Their cores are also typically smaller, shorter-lived, and lower in altitude — though they exhibit a higher area ratio between cloud and core area. In contrast, multi-core systems a larger in area, persist longer, reach a greater CTH and core height, and form larger individual cores (Section 4.1). Correlation analysis reveals a high interdependence among cloud and core properties. The strongest positive relationships are found between CTH and core height, and between the number of cores and both cloud area and core height (Figure 14).

Over land, the diurnal cycle exhibits a midday-to-afternoon peak (12:00–16:00 UTC) for the cloud and core area, the reflectivity at 10 km, the core height, and area ratio, with a secondary overnight peak for the cloud and core lifetimes. Over the ocean, the diurnal cycle is less pronounced but still shows afternoon, evening, and nighttime maxima. These patterns are especially distinct for single-core clouds (Section 4.2) and are largely consistent with prior tropical observations (e.g., Chen et al. (2021); Takahashi et al. (2017); Gupta et al. (2024)). Derived life-cycle statistics indicate that peaks for the reflectivity gradient typically precede vertical growth and subsequent area expansion — mainly around local noon and early evening over continental Africa. In oceanic regions, a less consistent pattern emerges, with a high variability throughout the day (Section 4.3.2). Compared to Wilcox et al. (2023), we observe weaker diurnal amplitudes over both land and ocean; however, the differences between single- and multi-core clouds are more pronounced than those between land- and sea-based convection.

We also find that changes in cloud area, core number, and cloud height often evolve in line with an idealised convective life cycle described in Sect. 3.6. Longer-lived clouds tend to exhibit more cores and larger maximum core areas. Multi-core systems reach their peak core number and core size later in their life cycle than single-core clouds (Section 4.3.1). The reflectivity gradients correlate positively with vertical growth and negatively with area expansion, reflecting transitions from development to maturity, as noted by Hu et al. (2021). Single-core clouds display stronger vertical ascent and higher reflectivity gradients, though most correlations are weak — aside from a strong negative relationship between cloud lifetime and vertical growth, and a moderate positive link between lifetime and area growth. For multi-core systems, average area growth is about 20–30 % higher than for single-core clouds. Reflectivity gradient and vertical growth are negatively correlated with most other properties, while the area growth shows positive correlations with the core count, core area, and core height. However, correlations differ between single-core and multi-core systems (Figure 13, Figure 14). Finally, our results underscore how differences in core structure and number may influence convective cloud development and the associated life-cycle.

5.2 Influences on convective clouds in the tropics

Compared to previous studies by Takahashi et al. (2023) and Hu et al. (2021), our analysis identifies a significantly higher number of potentially convective cloud tracks. The derived cloud characteristics align well with aircraft observations (Zipser and LeMone, 1980) and precipitation-based studies (Zipser et al., 2006). Over tropical Africa, our core distribution results are consistent with those derived for geostationary satellite data (Jones et al., 2024) or the CloudSat CPR (Deng et al., 2016), both of which found a high prevalence of clouds with one to three cores. Similarly, Pilewskie and L'Ecuyer (2022) reported that one-third to half of convective systems observed globally by the CloudSat CPR contain a single core. For the tropics, however, our results are in closer agreement. In line with these findings, we observe that cloud area generally increases with the number of cores. However, this relationship exhibits substantial variability, especially in multi-core systems (Section 4.3.1).

Our results show that convective cloud properties over land typically peak during the day, while over the ocean, we observe two peaks during daytime and at night (Sect. 4.2.1). These findings align with the diurnal cycle of tropical convection (Vondou, 2012; Takahashi et al., 2023). A nocturnal enhancement over the ocean may be linked to the diurnal cycle of free-tropospheric humidity, which peaks overnight and supports convection (Wall et al., 2020). After sunrise, solar heating may stabilize the atmosphere. A weakening land breeze may lead to the dissipation of night-time clusters (Houze Jr., 2004). While these diurnal

patterns may be reflected by the cloud properties (Sections 4.2.1), differences in the daytime of the first detection for the cloud tracks appear weaker (Figure 9). Throughout the day, we observe several peaks for the reflectivity gradient, followed by phases of vertical and horizontal cloud growth (Section 4.3.2). Notably, as the number of cores increases, both reflectivity gradient and vertical growth decline, while area growth becomes more pronounced (Section 4.3.1). This finding may correspond to a higher reflectivity at 10 km and broader spatial extent seen for multi-core systems (Section 4.1). Our observations may point to a self-sustaining mechanism where cores are regenerated in response to diurnal heating (e.g., Deng et al. (2016); Hartmann et al. (2018); Takahashi et al. (2017)). However, we did not explicitly investigate this process.

Seasonal variability in tropical convection has been highlighted in past studies. For example, multi-core systems often persist overnight during the onset of the West African Monsoon (Futyan and Genio, 2007). During this period, convection may frequently initiate over high terrain and propagate downslope at night under katabatic flow (Nicholson, 2018). While our results may be influenced by interannual variability, as the dataset spans only one year and does not capture a full annual cycle, we may observe temporal changes in cloud and core properties between March and August. For instance, we find an increase of the cloud area, cloud lifetime, number of cores, and core area over the ocean. Over land, these properties slightly decrease (Section 4.2.2). Reflectivity gradients increase along the period, in particular over the ocean, while vertical and area growth vary more considerably (Section 4.3.2). Between June and August, we detect overall a lower proportion of cloud tracks than between March and May. However, we see a shift regarding the distribution of cloud tracks over land and sea along the period: While we detect a higher proportion of cloud tracks over continental Africa in March and April, the number of detected clouds over the ocean exceeds those over land from May to August. The differences between monthly averages over both surface types remain small. Here, our results may diverge from studies that report more pronounced spatial and seasonal variations for convective clouds over land and sea (e.g., Takahashi and Luo (2012); Wilcox et al. (2023)). More striking than these surface-type induced differences in mean cloud properties are the contrasts between single- and multi-core clouds. Longer-lived, multi-core systems often exhibit repeated phases of growth (Takahashi et al., 2017). Consistent with Taylor et al. (2017), we observe slightly larger cores, especially between March and May, and enhanced area growth for continental convection (Sections 4.2.2 and 4.3.2).

5.3 Limitations and future challenges

Our analysis is based on ML-derived 3D radar reflectivity fields. While the results may help to extend the data availability of global radar reflectivities, they possibly contribute to mitigating limitations in approximating the cloud vertical extent from geostationary satellite observations. However, we point out several important limitations. The input data for the ML model are based on observations from the Cloud Profiling Radar (CPR) aboard the CloudSat satellite, which has known limitations in detecting ice clouds due to its tendency to underestimate the height of upper-level outflow (Wang et al., 2014). Additionally, signal attenuation near the surface caused by topography reduces the CPR's sensitivity to shallow convection. As a result, our analysis underrepresents both shallow convective and cirrus cloud types. Emerging satellite systems, such as the Flexible Combined Imager on the Meteosat Third Generation (MTG) platform (Holmlund et al., 2021) and the enhanced CPR on the EarthCARE mission (Eisinger et al., 2024), offer improved spatial and temporal resolution. These instruments are expected to enhance the detection and characterisation of the convective cloud life-cycle. Moroever, our study does not account for

several potentially important influences on convection, such as aerosol interactions, vertical wind shear, and entrainment rates (Masunaga and Luo, 2016). Incorporating these factors in future analyses could lead to a more comprehensive understanding of convective processes. Our study focuses on a domain within the tropical band from 30° W to 30° E and 30° N to 30° S. As extratropical influences may blur the statistics at the domain's northern and southern edges, it may be beneficial to explore the intra-tropical variability and the role of large-scale dynamics beyond this region to distinguish tropical from midlatitude convective processes.

To identify isolated and clustered convective systems, we employ the *tobac* framework (Sokolowsky et al., 2024). However, this object-based approach relies on manually set thresholds for the detection, introducing a degree of subjectivity that may influence the resulting cloud statistics. This includes potential limitations in the analysis of the cloud life-cycle, which itself does not provide insights on actual changes of the cloud temperature. It is important to emphasise that no universally optimal detection algorithm exists (Lakshmanan and Kain, 2010); each method has context-specific strengths and limitations depending on the intended application and geographic domain (Prein et al., 2024). While our approach may enable an approximation of the vertical cloud column, radar reflectivity alone does not substitute for measurements of vertical wind shear (Luo et al., 2008). Although vertical shear may have a smaller impact on convective processes in the tropics than in mid-latitude regions, its role still warrants further investigation (Takahashi et al., 2017). Finally, to better assess current and future convective risks, particularly those posed by multi-core systems, future work should explore their associated precipitation patterns in more depth (Atiah et al., 2023).

6 Conclusions

This study analyses the properties and life-cycle of convective clouds and their deep convective cores over a tropical region covering the Atlantic Ocean and West Africa. Using an ML-based extrapolation of radar reflectivities, we may enhance the number of detected clouds compared to retrievals from CloudSat CPR alone. Hence, the approach may help to close current data gaps. In this study, we aim to showcase the 3D data and their ability to track convective clouds and cores along their life-cycle. Compared to using data from only either a passive or active sensor, our perspective may allow a simultaneous coverage of cloud development in the horizontal and vertical dimension.

The results suggest that differences based on the number of cores are higher than the surface-type induced variability. Single-core clouds develop and dissipate on shorter timescales. They have a smaller cloud and core area, and lower CTH and core height than multi-core systems. The longer cloud lifetime of multi-core clouds may be associated to a later occurrence of the maximum number of cores and core area. Between single-core and multi-core clouds, we find considerable differences in the cloud life-cycle statistics regarding the changes in the radar reflectivity at 10 km height, the vertical growth, and the area growth of the cloud. While the former two are higher for clouds with a single core, multi-core cloud clusters with a larger cloud area tend to grow more along the horizontal dimension. The more cores we find, the later the maximum number of cores and the maximum core area occur. While the differences between the convective clouds over land and ocean are lower

than expected, we emphasise our analysis uses six months of data and may not represent the annual cycle of convection. Nevertheless, expanding the approach to investigate a longer time series may account for current uncertainties.

In this work, we use the number of convective cores to compare the effects of spatial clustering. However, we think that it is worth comparing these results to a quantification of convective organisation using more advanced metrics, as done in the accompanying manuscript.

Code and data availability. The level 2B-GEOPROF CloudSat data used in this study are available at the CloudSat Data Processing Center at CIRA/Colorado State University and can be retrieved from http://www.CloudSat.cira.colostate.edu/order-data (CloudSat Data Processing Center, 2024). The Meteosat SEVIRI level 1.5 data used in this study is freely and openly available via the EUMETSAT Data Store at https://navigator.eumetsat.int/product/EO-:EUM:DAT:MSG:HRSEVIRI (EUMETSAT Data Services, 2024). The code used in this study will be released upon publication. The dataset of convective cloud tracks created in this study is available at the following repository: https://zenodo.org/records/14724869 (Brüning, 2025b). The material used to prepare this paper, including code used to perform analysis and that needed for the preparation of figures, is archived at https://zenodo.org/records/15607483 (Brüning, 2025a).

Author contributions. S.B and H.T. designed the study. S.B developed the code for the analysis and visualisation. S.B. and H.T. contributed to the analysis and evaluation of cloud tracks and properties. S.B. and H.T. wrote the draft of the paper. All authors have read and agreed to the published version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the project “Big Data in Atmospheric Physics (BINARY)”, funded by the Carl Zeiss Foundation (grant P2018-02-003), and the Max Planck Graduate Center with the Johannes Gutenberg University of Mainz (MPGC). We thank EUMETSAT for providing access to the Meteosat SEVIRI imager data and the Cooperative Institute for Research in the Atmosphere, CSU, for providing access to the CloudSat 2B-GEOPROF data.

References

Andrews, P. C., Cook, K. H., and Vizy, E. K.: Mesoscale convective systems in the Congo Basin: seasonality, regionality, and diurnal cycles, Clim. Dyn., 62, 609–630, https://doi.org/10.1007/s00382-023-06903-7, 2024.

Atiah, W. A., Amekudzi, L. K., and Danuor, S. K.: Mesoscale convective systems and contributions to flood cases in Southern West Africa (SWA): A systematic review, Weather and Climate Extremes, 39, 100 551, https://doi.org/10.1016/j.wace.2023.100551, 2023.

Bacmeister, J. T. and Stephens, G. L.: Spatial statistics of likely convective clouds in CloudSat data, J. Geophys. Res. Atmos., 116, D04 104, https://doi.org/10.1029/2010JD014444, 2011.

Bony, S., Stevens, B., Frierson, D., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T., Sherwood, S., Siebesma, A., Watanabe, M., and Webb, M.: Clouds, circulation and climate sensitivity, Nature Geoscience, 8, 261–268, https://doi.org/10.1038/ngeo2398, 2015.

Brüning, S.: Detecting ML-based convective clouds using 3D observational data, Zenodo [code], https://doi.org/10.5281/zenodo.15607393, 2025a.

Brüning, S.: Convective cloud trajectories from 3D radar reflectivities, Zenodo [data set], https://doi.org/10.5281/zenodo.14724401, 2025b.

Brüning, S., Niebler, S., and Tost, H.: Artificial intelligence (AI)-derived 3D cloud tomography from geostationary 2D satellite data, Atmos. Meas. Tech., 17, 961–978, https://doi.org/10.5194/amt-17-961-2024, 2024.

Chen, P.-J., Chen, W.-T., Wu, C.-M., and Yo, T.-S.: Convective Cloud Regimes From a Classification of Object-Based CloudSat Observations Over Asian-Australian Monsoon Areas, Geophys. Res. Lett., 48, e2021GL092 733, https://doi.org/10.1029/2021GL092733, 2021.

Chen, S. S. and Houze, R. A.: Diurnal variation and life-cycle of deep convective systems over the tropical pacific warm pool, Q. J. R. Meteorol. Soc., 123, 357–388, https://doi.org/10.1002/qj.49712353806, 1997.

CloudSat Data Processing Center: Data Products, CloudSat DPC [data set], https://www.cloudsat.cira.colostate.edu/data-products, accessed: 2024-12-12, 2024.

Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, Routledge, 2 edn., https://doi.org/10.4324/9780203771587, 2013.

Crook, J., Klein, C., Folwell, S., Taylor, C. M., Parker, D. J., Stratton, R., and Stein, T.: Assessment of the Representation of West African Storm Lifecycles in Convection-Permitting Simulations, ESS, 6, 818–835, https://doi.org/10.1029/2018EA000491, 2019.

Cui, W., Dong, X., Xi, B., and Feng, Z.: Climatology of Linear Mesoscale Convective System Morphology in the United States Based on the Random-Forests Method, J. Clim., 34, 7257–7276, https://doi.org/10.1175/JCLI-D-20-0862.1, 2021.

Deng, M., Mace, G. G., and Wang, Z.: Anvil Productivities of Tropical Deep Convective Clusters and Their Regional Differences, Journal of the Atmospheric Sciences, 73, 3467 – 3487, https://doi.org/10.1175/JAS-D-15-0239.1, 2016.

Dixon, M. and Wiener, G.: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A Radar-based Methodology, J. Atmos. Oceanic Tech., 10, 785–797, https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2, 1993.

Eisinger, M., Marnas, F., Wallace, K., Kubota, T., Tomiyama, N., Ohno, Y., Tanaka, T., Tomita, E., Wehr, T., and Bernaerts, D.: The Earth-CARE mission: science data processing chain overview, Atmos. Meas. Tech., 17, 839–862, https://doi.org/10.5194/amt-17-839-2024, 2024.

Esmaili, R. B., Tian, Y., Vila, D. A., and Kim, K.-M.: A Lagrangian analysis of cold cloud clusters and their life cycles with satellite observations, J. Geophys. Res. Atmos., 121, 11,723–11,738, https://doi.org/10.1002/2016JD025653, 2016.

EUMETSAT Data Services: High Rate SEVIRI Level 1.5 Image Data - MSG - 0 degree, https://navigator.eumetsat.int/product/EO:EUM:DAT:MSG:HRSEVIRI, accessed: 2024-12-12, 2024.

Feng, Z., Hardin, J., Barnes, H. C., Li, J., Leung, L. R., Varble, A., and Zhang, Z.: PyFLEXTRKR: a flexible feature tracking Python software for convective cloud analysis, Geosci. Model Dev., 16, 2753–2776, https://doi.org/10.5194/gmd-16-2753-2023, 2023.

Finkensieper, S., Meirink, J. F., van Zadelhoff, G.-J., Hanschmann, T., Benas, N., Stengel, M., Fuchs, P., Hollmann, R., Kaiser, J., and Werscheck, M.: CLAAS-2.1: CM SAF CLoud property dAtAset using SEVIRI - Edition 2.1, https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V002_01, 2020.

Fiolleau, T. and Roca, R.: An Algorithm for the Detection and Tracking of Tropical Mesoscale Convective Systems Using Infrared Images From Geostationary Satellite, IEEE Trans. Geosci. Remote. Sens., 51, 4302–4315, https://doi.org/10.1109/TGRS.2012.2227762, 2013.

Futyan, J. M. and Genio, A. D. D.: Deep Convective System Evolution over Africa and the Tropical Atlantic, J. Clim., 20, 5041–5060, https://doi.org/10.1175/JCLI4297.1, 2007.

Ganetis, S. A., Colle, B. A., Yuter, S. E., and Hoban, N. P.: Environmental Conditions Associated with Observed Snowband Structures within Northeast U.S. Winter Storms, Mon. Weather Rev., 146, 3675–3690, https://doi.org/10.1175/MWR-D-18-0054.1, 2018.

Gupta, S., Wang, D., Giangrande, S. E., Biscaro, T. S., and Jensen, M. P.: Lifecycle of updrafts and mass flux in isolated deep convection over the Amazon rainforest: insights from cell tracking, Atmos. Chem. Phys., 24, 4487–4510, https://doi.org/10.5194/acp-24-4487-2024, 2024.

Haberlie, A. M. and Ashley, W. S.: A Method for Identifying Midlatitude Mesoscale Convective Systems in Radar Mosaics. Part II: Tracking, JAMC, 57, 1599–1621, https://doi.org/10.1175/JAMC-D-17-0294.1, 2018.

Hartmann, D. L., Hendon, H. H., and Houze, R. A.: Some Implications of the Mesoscale Circulations in Tropical Cloud Clusters for Large-Scale Dynamics and Climate, Journal of Atmospheric Sciences, 41, 113 – 121, https://doi.org/10.1175/1520-0469(1984)041<0113:SIOTMC>2.0.CO;2, 1984.

Hartmann, D. L., Gasparini, B., Berry, S. E., and Blossey, P. N.: The Life Cycle and Net Radiative Effect of Tropical Anvil Clouds, Journal of Advances in Modeling Earth Systems, 10, 3012–3029, https://doi.org/https://doi.org/10.1029/2018MS001484, 2018.

Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., and Tanelli, S.: Rainfall retrieval over the ocean with spaceborne W-band radar, J. Geophys. Res. Atmos., 114, https://doi.org/10.1029/2008JD009973, 2009.

Heikenfeld, M., Marinescu, P. J., Christensen, M., Watson-Parris, D., Senf, F., van den Heever, S. C., and Stier, P.: tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets, Geosci. Model Dev., 12, 4551–4570, https://doi.org/10.5194/gmd-12-4551-2019, 2019.

Holmlund, K., Grandell, J., Schmetz, J., Stuhlmann, R., Bojkov, B., Munro, R., Lekouara, M., Coppens, D., Viticchie, B., August, T., Theodore, B., Watts, P., Dobber, M., Fowler, G., Bojinski, S., Schmid, A., Salonen, K., Tjemkes, S., Aminou, D., and Blythe, P.: Meteosat Third Generation (MTG): Continuation and Innovation of Observations from Geostationary Orbit, BAMS, 102, E990–E1015, https://doi.org/10.1175/BAMS-D-19-0304.1, 2021.

Houze, R. A. and Hobbs, P. V.: Organization and Structure of Precipitating Cloud Systems, Adv. Geophys., 24, 225–315, https://doi.org/10.1016/S0065-2687(08)60521-X, 1982.

Houze Jr., R. A.: Observed structure of mesoscale convective systems and implications for large-scale heating, Q.J.R. Meteorol. Soc., 115, 425–461, https://doi.org/https://doi.org/10.1002/qj.49711548702, 1989.

Houze Jr., R. A.: Mesoscale convective systems, Rev. Geophys., 42, RG4003, https://doi.org/10.1029/2004RG000150, 2004.

Hu, X., Ge, J., Li, W., Du, J., Li, Q., and Mu, Q.: Vertical Structure of Tropical Deep Convective Systems at Different Life Stages From CloudSat Observations, Journal of Geophysical Research: Atmospheres, 126, e2021JD035 115, https://doi.org/https://doi.org/10.1029/2021JD035115, 2021.

Igel, M. R., Drager, A. J., and van den Heever, S. C.: A CloudSat Cloud-Object Partitioning Technique and Assessment and Integration of Deep Convective Anvil Sensitivities to Sea Surface Temperature, J. Geophys. Res. Atmos., 119, 10 515–10 535, https://doi.org/10.1002/2014JD021717, 2014.

Jones, W., Stengel, M., and Stier, P.: A Lagrangian perspective on the lifecycle and cloud radiative effect of deep convective clouds over Africa, Atmos. Chem. Phys., 24, 5165–5180, https://doi.org/10.5194/acp-24-5165-2024, 2024.

Kikuchi, M. and Suzuki, K.: Characterizing Vertical Particle Structure of Precipitating Cloud System From Multiplatform Measurements of A-Train Constellation, Geophysical Research Letters, 46, https://doi.org/10.1029/2018GL081244, 2019.

Kniffka, A., Knippertz, P., and Fink, A. H.: The role of low-level clouds in the West African monsoon system, Atmos. Chem. Phys., 19, 1623–1647, https://doi.org/10.5194/acp-19-1623-2019, 2019.

Kukulies, J., Chen, D., and Curio, J.: The Role of Mesoscale Convective Systems in Precipitation in the Tibetan Plateau Region, J. Geophys. Res. Atmos., 126, e2021JD035 279, https://doi.org/10.1029/2021JD035279, 2021.

Lakshmanan, V. and Kain, J. S.: A Gaussian Mixture Model Approach to Forecast Verification, WAF, 25, 908–920, https://doi.org/10.1175/2010WAF2222355.1, 2010.

Leary, C. A. and Houze, R. A.: The Contribution of Mesoscale Motions to the Mass and Heat Fluxes of an Intense Tropical Convective System, Journal of Atmospheric Sciences, 37, 784 – 796, https://doi.org/10.1175/1520-0469(1980)037<0784:TCOMMT>2.0.CO;2, 1980.

Li, W., Zhang, F., Yu, Y., Iwabuchi, H., Shen, Z., Wang, G., and Zhang, Y.: The semi-diurnal cycle of deep convective systems over Eastern China and its surrounding seas in summer based on an automatic tracking algorithm, Clim. Dyn., 56, 357–379, https://doi.org/10.1007/s00382-020-05474-1, 2021.

Liu, C. and Zipser, E. J.: Diurnal cycles of precipitation, clouds, and lightning in the tropics from 9 years of TRMM observations, Geophysical Research Letters, 35, L04 819, https://doi.org/https://doi.org/10.1029/2007GL032437, 2008.

Luo, Z., Liu, G. Y., and Stephens, G. L.: CloudSat adding new insight into tropical penetrating convection, Geophys. Res. Lett., 35, L19 819, https://doi.org/10.1029/2008GL035330, 2008.

Machado, L. A. T., Rossow, W. B., Guedes, R. L., and Walker, A. W.: Life Cycle Variations of Mesoscale Convective Systems over the Americas, Monthly Weather Review, 126, 1630 – 1654, https://doi.org/10.1175/1520-0493(1998)126<1630:LCVOMC>2.0.CO;2, 1998.

Maddox, R. A.: Mesoscale Convective Complexes, BAMS, 61, 1374–1387, https://doi.org/https://doi.org/10.1175/1520-0477(1980)061<1374:MCC>2.0.CO;2, 1980.

Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G.: Hydrometeor Detection Using Cloudsat—An Earth-Orbiting 94-GHz Cloud Radar, J. Atmos. Oceanic Tech., 25, 519–533, https://doi.org/10.1175/2007JTECHA1006.1, 2008.

Masunaga, H. and Luo, Z. J.: Convective and large-scale mass flux profiles over tropical oceans determined from synergistic analysis of a suite of satellite observations, J. Geophys. Res. Atmos., 121, https://doi.org/10.1002/2016JD024753, 2016.

Mecikalski, J. R., MacKenzie, W. M., Koenig, M., and Muller, S.: Cloud-Top Properties of Growing Cumulus prior to Convective Initiation as Measured by Meteosat Second Generation. Part I: Infrared Fields, JAMC, 49, 521–534, https://doi.org/10.1175/2009JAMC2344.1, 2010.

Meyer, F.: Topographic distance and watershed lines, Signal Process., 38, 113–125, https://doi.org/10.1016/0165-1684(94)90060-4, 1994.

Nesbitt, S. W. and Zipser, E. J.: The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements, Journal of Climate, 16, 1456 – 1475, https://doi.org/10.1175/1520-0442(2003)016<1456:TDCORA>2.0.CO;2, 2003.

Nicholson, S. E.: The ITCZ and the Seasonal Cycle over Equatorial Africa, BAMS, 99, 337–348, https://doi.org/10.1175/BAMS-D-16-0287.1, 2018.

Oreopoulos, L., Cho, N., and Lee, D.: New insights about cloud vertical structure from CloudSat and CALIPSO observations, J. Geophys. Res. Atmos., 122, 9280–9300, https://doi.org/10.1002/2017JD026629, 2017.

Pilewskie, J. A. and L'Ecuyer, T. S.: The Global Nature of Early-Afternoon and Late-Night Convection Through the Eyes of the A-Train, Journal of Geophysical Research: Atmospheres, 127, e2022JD036 438, https://doi.org/https://doi.org/10.1029/2022JD036438, 2022.

Prein, A. F., Feng, Z., Fiolleau, T., Moon, Z. L., Núñez Ocasio, K. M., Kukulies, J., Roca, R., Varble, A. C., Rehbein, A., Liu, C., Ikeda, K., Mu, Y., and Rasmussen, R. M.: Km-Scale Simulations of Mesoscale Convective Systems Over South America—A Feature Tracker Intercomparison, J. Geophys. Res. Atmos., 129, e2023JD040 254, https://doi.org/10.1029/2023JD040254, 2024.

Raut, B. A., Jackson, R., Picel, M., Collis, S. M., Bergemann, M., and Jakob, C.: An Adaptive Tracking Algorithm for Convection in Simulated and Remote Sensing Data, JAMC, 60, 513–526, https://doi.org/10.1175/JAMC-D-20-0119.1, 2021.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, edited by Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., vol. 9351, pp. 234–241, Springer International Publishing, Cham, 2015.

Sassen, K. and Wang, Z.: Classifying clouds around the globe with the CloudSat radar: 1-year of results, Geophys. Res. Lett., 35, L04 805, https://doi.org/10.1029/2007GL032591, 2008.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An introduction to Metetsat second generation (MSG), BAMS, 83, 977–992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, 2002.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of Geophysics, 58, e2019RG000 678, https://doi.org/https://doi.org/10.1029/2019RG000678, 2020.

Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., Heikenfeld, M., Brunner, K. N., Bruning, E. C., Collis, S. M., Jackson, R. C., Leung, G. R., Pfeifer, N., Raut, B. A., Saleeby, S. M., Stier, P., and van den Heever, S. C.: *tobac* v1.5: introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena, Geosci. Model Dev., 17, 5309–5330, https://doi.org/10.5194/gmd-17-5309-2024, 2024.

Steiner, M., Houze, R. A., and Yuter, S. E.: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data, JAMC, 34, 1978–2007, https://doi.org/10.1175/1520-0450(1995)034<1978:CCOTDS>2.0.CO;2, 1995.

Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., Reinke, D., Partain, P., Mace, G. G., Austin, R., L'Ecuyer, T., Haynes, J., Lebsock, M., Suzuki, K., Waliser, D., Wu, D., Kay, J., Gettelman, A., Wang, Z., and Marchand, R.: CloudSat mission: Performance and early science after the first year of operation, J. Geophys. Res. Atmos., 113, https://doi.org/10.1029/2008JD009982, 2008.

Takahashi, H. and Luo, Z.: Where is the level of neutral buoyancy for deep convection?, Geophysical Research Letters, 39, L15 809, https://doi.org/https://doi.org/10.1029/2012GL052638, 2012.

Takahashi, H., Luo, Z. J., and Stephens, G. L.: Level of neutral buoyancy, deep convective outflow, and convective core: New perspectives based on 5 years of CloudSat data, J. Geophys. Res. Atmos., 122, 2958–2969, https://doi.org/10.1002/2016JD025969, 2017.

Takahashi, H., Luo, Z. J., Stephens, G., and Mulholland, J. P.: Revisiting the Land-Ocean Contrasts in Deep Convective Cloud Intensity Using Global Satellite Observations, Geophys. Res. Lett., 50, e2022GL102 089, https://doi.org/10.1029/2022GL102089, 2023.

Taylor, S., Stier, P., White, B., Finkensieper, S., and Stengel, M.: Evaluating the diurnal cycle in cloud top temperature from SEVIRI, Atmos. Chem. Phys., 17, 7035–7053, https://doi.org/10.5194/acp-17-7035-2017, 2017.

Tomkins, L. M., Yuter, S. E., and Miller, M. A.: Dual adaptive differential threshold method for automated detection of faint and strong echo features in radar observations of winter storms, Atmos. Meas. Tech., 17, 3377–3399, https://doi.org/10.5194/amt-17-3377-2024, 2024.

Vondou, D. A.: Spatio-Temporal Variability of Western Central African Convection from Infrared Observations, Atmos., 3, 377–399, https://doi.org/10.3390/atmos3030377, 2012.

Wall, C. J., Norris, J. R., Gasparini, B., Smith, W. L., Thieman, M. M., and Sourdeval, O.: Observational Evidence that Radiative Heating Modifies the Life Cycle of Tropical Anvil Clouds, J. Clim., 33, 8621–8640, https://doi.org/10.1175/JCLI-D-20-0204.1, 2020.

Wang, X., Cui, C., Cui, W., and Shi, Y.: Modes of mesoscale convective system organization during Meiyu season over the Yangtze River basin, Acta Meteorol. Sin., 28, 111–126, https://doi.org/10.1007/s13351-014-0108-4, 2014.

Wielicki, B. A., Cess, R. D., King, M. D., Randall, D. A., and Harrison, E. F.: Mission to Planet Earth: Role of Clouds and Radiation in Climate, BAMS, 76, 2125 – 2154, https://doi.org/10.1175/1520-0477(1995)076<2125:MTPERO>2.0.CO;2, 1995.

Wilcox, E. M., Yuan, T., and Song, H.: Deep convective cloud system size and structure across the global tropics and subtropics, Atmos. Meas. Tech., 16, 5387–5401, https://doi.org/10.5194/amt-16-5387-2023, 2023.

Zipser, E. J. and LeMone, M. A.: Cumulonimbus Vertical Velocity Events in GATE. Part II: Synthesis and Model Core Structure, Journal of Atmospheric Sciences, 37, 2458–2469, https://doi.org/10.1175/1520-0469(1980)037<2458:CVVEIG>2.0.CO;2, 1980.

Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., and Yorty, D. P.: Where are the most intense thunderstorms on earth?, BAMS, 87, 1057–1072, https://doi.org/10.1175/BAMS-87-8-1057, 2006.

3.4 Investigating effects and spatial patterns of convective organisation

The preprint was accepted for publication in *Atmospheric Chemistry and Physics* as a research article.

I am the lead author of this study, where I implemented the calculation of the convective organisation indices and the detection of regional hotspots regions based on the previously introduced ML-based predictions of a 3D cloud field (Section 3.2) and the cloud detection algorithm (Section 3.3). I analysed the data, created the figures, and wrote the manuscript. The co-author contributed to interpreting the results and proofreading the manuscript. The paper includes a section called *Author contributions* detailing individual contributions.

How to cite: Brüning, S. and Tost, H.: A machine learning-based perspective on deep convective clouds and their organisation in 3D. Part II: Spatial-temporal patterns of convective organisation, *EGUsphere [preprint]*, https://doi.org/10.5194/egusphere-2025-376, **2025**.

Submitted: 05 February 2025
Revised: 23 May 2025
Accepted: 03 July 2025

A machine learning-based perspective on deep convective clouds and their organisation in 3D. Part II: Spatial-temporal patterns of convective organisation

Sarah Brüning[1] and Holger Tost[1]

[1]Institute for Physics of the Atmosphere, Johannes Gutenberg University Mainz, Johann-Joachim-Becher-Weg 21, Mainz, 55128, Rhineland-Palatinate, Germany

Correspondence: Sarah Brüning (sbruenin@uni-mainz.de)

Abstract. This series of papers explores spatio-temporal patterns of convective cloud occurrence and organisation. We use a machine learning-based method to extrapolate a contiguous 3D cloud field of 2D satellite data. In Part 2, we focus on convective organisation in tropical West Africa between March and August 2019, examining how it relates to the 3D properties of convective clouds and their core structures. We quantify organisation using three indices (SCAI, COP, ROME) to capture different aspects of spatial cloud clustering. Our findings highlight how cloud properties may interact with organisation. Hence, strong organisation may occur with larger cloud areas, lower cloud tops and core heights, and shorter lifespans compared to the average convective system. In contrast, weak organisation may be associated with smaller clouds, fewer cores, but similarly shorter lifespans. We find an increasing frequency of convective organisation in the northern hemisphere during boreal summer months, likely linked to the northward migration of the Intertropical Convergence Zone (ITCZ). From March to May, patches of strong convective organisation emerge along the African coastlines and over the southern Atlantic Ocean. Between June and August, hotspots shift inland, particularly across the Sahel and wider West African plains. Notably, oceanic regions show slightly stronger organisation overall. However, overlapping regions of strong and weak organisation may complicate the interpretation of regional statistics. While the machine learning-based 3D perspective helps bridge observational gaps in the representation of cloud structures, the inherent complexity and variability of convective organisation highlight the need for continued investigation.

1 Introduction

Atmospheric convection plays an essential role in the climate system through its contribution to weather and climate variability (Brune et al., 2020). In the tropics, we observe convective clouds forming as spatially connected structures of extensive size (Houze, 1977). These mesoscale convective systems (MCSs) are one of the main drivers for the transport of heat and moisture through the atmosphere. Furthermore, they affect the hydrological and radiative variability on Earth (Hartmann et al., 1984).

The spatial clustering of convective systems - also known as convective organisation - may promote the occurrence of severe weather events such as hail and floods (Becker et al., 2021). However, a robust assessment of the connection between convective organisation and extreme weather, in particular in a future climate under global warming, expresses the need for further research.

Although the term convective organisation has become increasingly popular in climate research, it is often used vaguely. Mapes and Neale (2011) broadly summarise organisation as "non-randomness in meteorological fields in convecting regions". This definition induces a clustering of deep convective cells which is ubiquitous in the atmosphere, particularly in the tropics. However, the underlying mechanisms remain insufficiently understood (Muller and Bony, 2015). While convective organisation is difficult to quantify in observational data, idealised model configured in radiative-convective equilibrium (RCE) could demonstrate a large-scale clustering of convective clouds which is known as self-aggregation of convection (e.g, Nakajima and Matsuno (1988); Held et al. (1993); Wing et al. (2017)). Following Bretherton et al. (2005), it occurs on a timescale between days and weeks and describes the transition of an approximately random distribution of convective cells into convecting and non-convecting regions that grow upscale over time. Convective aggregation is driven by either internal dynamics, like cold pools and radiative feedback, or external forces, such as the land-sea-breeze (e.g., Haerter et al. (2019), Coppin and Bony (2015), Dauhut et al. (2016)). Self-aggregation increases with the size and proximity of convective clouds and affects the radiative feedback, large-scale circulation, and moisture distribution in the vicinity of a cloud cluster (Hartmann et al., 1984). For instance, an idealised model setup shows that an aggregated state consists of a single moist region surrounded by dry regions. Moreover, the feedback between convection, surface fluxes, and radiation further drives aggregation (Tobin et al., 2012). Research shows that self-aggregation may increase with a warming climate (Wing et al., 2020). However, there remain uncertainties connected to a large model spread (Bläckberg and Singh, 2022).

Despite these insights derived from models, identifying and quantifying convective organisation in observational data persists to be a challenge. This may be due a high variability in the quality and quantity of observations. In response, previous studies have developed various metrics aiming towards a deeper understanding about the underlying physical mechanisms. The indices analyse the spatial distribution of the clouds within a defined area to estimate the strength of convective organisation (Pscheidt et al., 2019). For instance, they help differentiate a regularly distributed, randomly distributed, or organised cloud field by using morphological attributes such as the number of clouds, their nearest-neighbour distances, size, shape, pattern, and timing (Pendergrass, 2020; Retsch et al., 2020).

Providing timely forecasts and a robust climate risk assessment requires even more a correct representation of convective organisation. While satellite observations has shown that organisation within the tropics may increase overall with extreme precipitation (Semie and Bony, 2020), we have limited knowledge about convective organisation on a regional level. In this study, we aim to provide a deeper understanding of the relationship between cloud properties and convective organisation on this regional scale, comparable to the work of Bao et al. (2024). The area of interest (AOI) covers West Africa and the tropical Atlantic Ocean between 30° N–30° S and 30° W–30° E and lies within the Inter-Tropical Convergence Zone (ITCZ). Here, the environmental conditions favour the development of deep convective clouds, which are often associated to heavy rain (Takahashi et al., 2023). A heterogeneous landmass distribution in the northern and southern hemispheres and land-ocean

contrasts may affect the development of convection (Zipser et al., 2006). Over the tropical Atlantic Ocean, a weaker large-scale forcing may induce lower cloud tops and less intense rain rates than over continental Africa (Futyan and Genio, 2007). The rainfall variability between the individual regions of the AOI substantially depends on the moisture availability and thermal gradients (Berthou et al., 2019). Overall, the West African monsoon (WAM) dominates the West African climate. A strong temperature gradient between the warm Sahara and the colder waters of the Gulf of Guinea drives the WAM (Fontaine and Philippon, 2000). Stronger convection generally leads to an increase in heavy rain, a larger detrainment, and a slightly smaller thick anvil emissivity. For instance, Stubenrauch et al. (2023) found a distinct annual cycle of convective organisation connected to seasonal shifts of the convective cloud properties.

In Part 1 of this sequence of papers, we derived contiguous trajectories of convective clouds and their deep convective core regions (hereafter: cores) in 15-minute intervals for a six-month period between March to August 2019 (Brüning and Tost, 2025). In this study, we examined cloud and core properties of tropical convection and the life-cycle of single-core and multi-core convective clouds. In this paper, we aim to complement the findings by an in-depth analysis of spatio-temporal patterns of convective organisation. Moreover, we aim to investigate the connection between convective organisation and cloud properties within the AOI. For this purpose, we quantify convective organisation at each point in time by employing three organisation indices. The goal is to derive spatial patterns of organisation and compare their spatio-temporal variability (Biagioli and Tompkins, 2023). Our study employs convective cloud trajectories derived from a 4D time series of contiguous 3D radar reflectivities, which we predict from a machine learning (ML)-based extrapolation of 2D satellite data (Brüning et al., 2024). We employ an object-based algorithm to detect and track convective clouds in the predicted radar reflectivity field. This perspective allows a simultaneous coverage of the horizontal (cloud and core area) and vertical (cloud and core height) properties in the AOI, including remote oceanic regions over the Atlantic Ocean. Our aim is to showcase how convective organisation is distributed in the AOI within the six-month period. Furthermore, we strive to quantify how differences in the cloud and core properties are connected to a weak or strong convective organisation.

We have divided this article into five further sections. In Sect. 2, we describe the dataset used in this study. Section 3 presents an overview of metrics employed to quantify convective organisation. Section 4 contains an overview of the results comprising the spatio-temporal variability of organisation indices and cloud properties. Section 5 relates our key findings to other studies. Moreover, we discuss some limitations we encountered and evaluate the role of the ITCZ and other environmental drivers for the development of tropical organisation. Finally, Sect. 6 contains a summary and the main conclusions.

2 Data

To quantify convective organisation over tropical West Africa, we use a ML-based 3D cloud mask build on the 3D cloud reconstruction method described in Brüning et al. (2024) and the convective cloud detection framework by Brüning and Tost (2025). The following section outlines the workflow for producing the 3D radar reflectivity dataset, detecting convective clouds and cores, and extracting cloud properties (Figure 1).

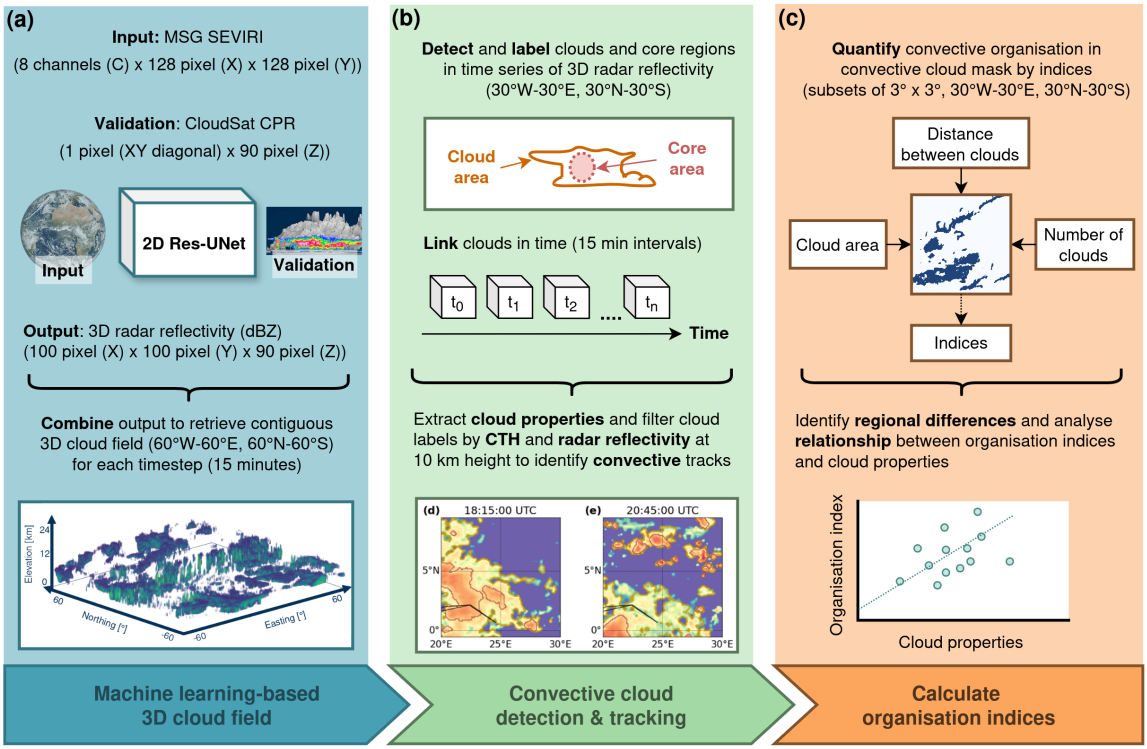

Figure 1. Overview of the workflow for this study. In **(a)**, we show how to derive a contiguous 3D cloud field from 2D data by a machine learning-based extrapolation (Brüning et al., 2024). For this purpose, 2D satellite imagery from the MSG SEVIRI sensor is fed into a 2D Res-UNet and trained to predict a 3D image of radar reflectivities validated against vertical cross sections of the CloudSat CPR. The predictions cover 100 x 100 pixels along 90 vertical bins of 240 m. These patches are combined to cover an area between 60° W–60° E and 60° N–60° S. In **(b)**, an object-based algorithm is employed to detect convective clouds and their cores within the predicted 3D radar reflectivity field. The temporal resolution of the data is 15 minutes. Through each point in time, we link identified cloud objects and filter the trajectories by the cloud top height (CTH), cloud base height (CBH), and number of cores to identify possible convective tracks (Brüning and Tost, 2025). In the current study **(c)**, we aim to quantify convective organisation by calculating organisation indices that are based on the area, distance, and number of objects in each cloud mask. The indices are calculated in 15-minute intervals for a period between March–August 2019. The results are used to analyse regional differences of convective organisation and to describe the relationship between convective organisation and cloud properties.

Table 1. Overview of MSG SEVIRI channels used to predict 3D radar reflectivities in this study.

Channel	Wavelength (μm)	Description	Spatial resolution at nadir	Retrieval at nighttime
IR3.9	3.48–4.36	Near infrared window	3 km	Yes
WV6.2	5.35–7.15	Upper-troposphere water vapour	3 km	Yes
WV7.3	6.85–7.85	Lower-troposphere water vapour	3 km	Yes
IR8.7	8.30–9.10	Mid infrared window	3 km	Yes
IR9.7	9.38–9.94	Ozone sensitivity	3 km	Yes
IR10.8	9.80–11.80	Clean longwave window	3 km	Yes
IR12.0	11.00–13.00	Dirty longwave window	3 km	Yes
IR 13.4	12.40–14.40	CO2 sensitivity	3 km	Yes

2.1 Satellite data

To identify, track, and analyse convective clouds, we employ a machine learning (ML) algorithm that generates time series of 3D radar reflectivity fields based on 2D satellite observations, as described in Brüning et al. (2024). The input data are derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the Meteosat-11 (MSG) satellite (Schmetz et al., 2002). The AOI is situated near the nadir of SEVIRI, which is positioned above the Equator at 0° longitude. SEVIRI captures multispectral imagery across 12 channels in the visible, near-infrared, and thermal-infrared ranges. Eleven of these channels offer a temporal resolution of 15 minutes and a spatial resolution of 3 km, while one high-resolution visible channel provides 1 km resolution at nadir. From these, we use eight channels to train our ML model (Table 1).

We employ vertical cross-sections of radar reflectivity from the 94-GHz Cloud Profiling Radar (CPR) onboard the polar-orbiting CloudSat satellite to validate our ML-based predictions. The CPR is an active radar instrument, which transmits microwave pulses toward Earth to detect vertical profiles of cloud hydrometeors. It has a vertical resolution of 240 m (distributed across 125 bins) and a horizontal resolution of 1.4 km across-track and 1.8 km along-track (Stephens et al., 2008). Our study employs data from the level-2 2B-GEOPROF product. While the CPR has a reduced sensor sensitivity at high altitudes, thin ice clouds like cirrus may be underrepresented. Moreover, the radar may be affected by signal attenuation at low altitudes caused by the topography (Sassen and Wang, 2008). To address these limitations, we limit the data to contain 90 height levels ranging from 2.4 km to 24 km. To improve the ML model performance, we filter the radar reflectivities by the CloudSat cloud mask quality flag to reduce the number of noisy pixels (Marchand et al., 2008).

2.2 3D cloud field reconstruction

In the following section, we describe the methodology used to reconstruct a 3D cloud field, based on the framework developed by Brüning et al. (2024). Our approach utilises a ML algorithm built on the 2D Res-UNet architecture — a modified convolutional neural network specifically designed for image segmentation tasks (Ronneberger et al., 2015). The model is primarily

Table 2. Modifications applied in this study to the Res-UNet originally proposed in Brüning et al. (2024)

Parameter	Original configuration	Modification
Number of input channels	11	8
Loss function	L2	L1
Nighttime predictions	No	Yes
Average RMSE	3.05	2.99

trained to reconstruct vertical cross-sections of the CloudSat CPR using data from the MSG SEVIRI satellite. Due to the U-Net architecture, the model is capable of producing full 3D radar reflectivity volumes rather than just 2D slices.

The AOI for the reconstructed 3D cloud field spans from 60° W to 60° E and from 60° S to 60° N, corresponding to 2400 × 2400 pixels in the horizontal dimensions. MSG SEVIRI satellite imagery serves as input to the Res-UNet model, setting the horizontal resolution of the 3D data to 3 km × 3 km. Initially, we used 11 spectral channels covering the visible, near-infrared, and thermal-infrared ranges. However, the visible channels were excluded in this study to enable daylight-independent predictions (Tables 1 and 2).

The training data consist of 128 × 128 pixel patches of MSG SEVIRI imagery, spatially and temporally aligned with CloudSat overpasses. Each training sample includes a diagonal CPR cross-section. To address the resolution mismatch between MSG SEVIRI and CloudSat, the CPR data are downsampled to match the horizontal resolution of MSG SEVIRI pixels. To mitigate the strong class imbalance between cloudy and cloud-free conditions, cloud-free samples are limited to a maximum of 10 % of the training data. The model is trained on nine months of data and validated on a separate three-month period. It is optimised to reconstruct CloudSat-like 3D reflectivity volumes with a horizontal resolution of 100 × 100 pixels and 90 vertical levels. Predicted radar reflectivity values range from –25 to 20 dBZ and maintain the 15-minute temporal resolution of the MSG SEVIRI input. An L1 loss function (mean absolute error, MAE) is used during training to evaluate performance. Direct validation is only possible for the diagonal cross-section, which constitutes about 10 % of the data points in each training sample. During the three-month test period, the modified daylight-independent model achieves a root mean square error (RMSE) of 2.99 dBZ, improving upon the original model's average RMSE of 3.05 dBZ (Table 2).

To achieve complete spatial coverage of the domain (60° W to 60° E and 60° S to 60° N), individual 3D output patches are stitched together to form a contiguous volume of 2400 × 2400 × 90 pixels (Figure 1, a). This approach may enable consistent spatial coverage, especially over remote oceanic regions where active sensors are scarce (Prein et al., 2024). Visual inspection confirms that no artifacts are present at tile boundaries, indicating seamless reconstruction across the domain. However, model accuracy tends to decrease with increasing distance from the MSG SEVIRI nadir. Finally, the 3D radar reflectivity volumes are concatenated along the temporal axis to create a 4D cloud field, which is then used to detect and track convective clouds. For the purposes of this study, we crop the domain to 1200 × 1200 pixels, covering the region from 30° W to 30° E and 30° N to 30° S — effectively focusing on the area between the Tropic of Cancer and the Tropic of Capricorn.

2.3 Detection and tracking of convective clouds and cores

Convective clouds are detected and tracked using the *tobac* package (Sokolowsky et al., 2024), which supports an object-based analysis of 3D meteorological data. The detection framework - as described in Brüning and Tost (2025) - proceeds in three stages: the cloud detection and tracking, the core detection, and the classification of potentially convective clouds. We use the ML-based predictions of the radar reflectivity as input data for the detection framework. While radar reflectivity does not directly measure vertical velocity, it may provide information for detecting hydrometeors associated with convective cloud development (Luo et al., 2008). By merging the 3D data fields along the temporal dimension, we receive a 4D time series that is fed into the tracking algorithm to create continuous trajectories with a temporal resolution of 15 minutes.

We identify potential candidates of convective clouds within the 3D cloud field by applying a fixed radar reflectivity threshold of -15 dBZ. This threshold is used to distinguish hydrometeors from background noise in the radar reflectivity data (Marchand et al., 2008). Although moderately restrictive, this threshold is intended to capture the full spatio-temporal evolution of convective clouds throughout their life cycle, thereby supporting the formation of contiguous trajectories (Esmaili et al., 2016). To reduce noise, we first apply a smoothing Gaussian image filter with an effective scale of half a standard deviation (sigma = 0.5) on the 3D radar reflectivity field. Next, we compute the centroids of potential cloud structures using a weighted center-of-mass approach, where the weight of each point is determined by its reflectivity value above the -15 dBZ threshold. Each identified centroid is assigned a unique identifier, which is retained throughout the subsequent tracking and segmentation processes. We then apply a 3D watershed segmentation algorithm to delineate the volume of individual cloud structures associated with each centroid. The algorithm places markers at the detected centroids within a binary 3D volume, where all other grid points are set to zero. From these markers, the algorithm expands outward through the volume, assigning reflectivity-based pixels to the corresponding cloud until the –15 dBZ threshold is reached. This process produces a labeled 3D cloud mask. Subsequently, we analyse the morphology of each cloud to determine whether any structures might represent a merger of multiple cloud systems. Each cloud's shape is characterised using the best-fitting ellipse, and we compute the aspect ratio — that is, the ratio of the major to the minor axis length. If the major axis is more than 75 % longer than the minor axis, we split the identified cloud into separate objects for further analysis. We track the labeled 3D cloud objects over time by linking them based on their estimated movement speed. At each 15-minute interval, we predict the expected position of a cloud object using its velocity from previous time steps. To streamline this linking process, we define a maximum search radius between time steps, within which only cloud objects are considered potential matches. When new clouds form, we assign them the average velocity of nearby clouds to estimate their likely movement (Heikenfeld et al., 2019). We require a minimum area overlap of 50 % to determine similarity between clouds across consecutive 15-minute intervals.

We aim to detect convective cores for each cloud object at every time step throughout its life cycle. For this purpose, we use the previously generated labeled 3D cloud mask. Core centroids are identified by locating local maxima in a combined metric that incorporates both smoothed radar reflectivity and the vertical extent of a contiguous potential core layer. Specifically, we calculate the mean radar reflectivity for each vertical cloud column, and determine the height of the core layer by counting the number of pixels with reflectivity values greater than 0 dBZ located above 5 km altitude. To fill isolated gaps in otherwise

Table 3. Cloud and core properties derived from the contiguous convective cloud trajectories.

Feature type	Feature name	Definition
Cloud	Cloud area	Area of the cloud (km^2)
	Cloud top height (CTH)	Height of the cloud (km)
	Lifetime	Lifetime of the cloud trajectory (h)
	Surface type	Value of land-sea mask
Core	Number of cores	Number of identified convective core regions
	Core area	Average area of convective cores (km^2)
	Core height	Depth of the core in the vertical column (km)

vertically continuous cores, we expand the threshold from 0 dBZ to -5 dBZ in columns that contain at least one pixel exceeding 0 dBZ (Luo et al., 2008; Igel et al., 2014). We then combine both indicators —average reflectivity and potential core vertical depth — for each pixel associated with a cloud label, resulting in a 2D layer where we search for local maxima. If at least one local maximum is detected, the corresponding locations are considered candidate core centroids. If no local maxima are found — for example, if no columns contain pixels above 0 dBZ at altitudes higher than 5 km — the cloud is recorded as having zero cores for that time step. Otherwise, we use a 3D watershed segmentation algorithm to delineate the core volumes surrounding each centroid, allowing for multiple cores to exist within a single cloud at the same time.

2.4 Extraction of cloud properties

We use the labelled cloud masks to extract cloud and core properties at each point in time. Moreover, we compute average properties across the cloud's lifetime to derive distinct key properties that may characterise the trajectory. These properties include the cloud lifetime, cloud area, cloud top height (CTH), number of cores, and mean core area and height (Table 3). The cloud area is computed from the column-wise maximum horizontal extent of the 3D cloud mask, while CTH is derived from the vertical extent. For the cloud lifetime, we extract the time (in hours) between the first and last detection of each trajectory of the labelled pixels. Each cloud track is classified as either marine (sea) or continental (land) using a binary land-sea mask. For this purpose, we determine the most frequent (modal) surface type across all grid points along the cloud trajectory. While this method does not capture changes in surface type throughout the cloud's life-cycle, it may provide insights on the effect of the most frequently occurring surface type. For clouds with one or more cores, we count the maximum number of cores associated to the trajectory. Moreover, the core area and height are derived from the column-wise maximum horizontal extent and vertical extent of the previously identified cores, similar to the cloud area and CTH.

2.5 Filter convective cloud trajectories

We filter the cloud trajectories to exclude possibly non-convective tracks from the analysis. For that purpose, we employ three criteria occurring for at least a single timestep of 15 minutes: (a) One or more core regions, (b) radar reflectivity of higher than 0 dBZ at 10 km height, and (c) minimum CTH of 10 km and maximum cloud base height (CBH) of less than 5 km. While we do not require the convective clouds to have a CTH higher than 10 km at every time step during their trajectory, we discard trajectories that never reach the CTH threshold. After filtering the dataset, we receive 375,000 uniquely labeled 3D cloud objects, each associated with a continuous time trajectory and structural information about cloud and core properties (Figure 1, b).

For further analysis, we exclude cloud tracks detected for a single time step of 15 minutes. This results in a refined dataset of 354,073 convective cloud trajectories between March and August 2019. In Fig. 2, we showcase the spatio-temporal distribution of the cloud trajectories. Most clouds are located between 5° S and 20° N, with peak activity from 5°–10° N (Figure 2, a). Approximately 75 % of cloud tracks occur over ocean, with land-based tracks comprising the remaining 25 % (Figure 2, b). Compared to the land-sea distribution of grid points across the AOI, we observe a 10 % shift toward ocean for detected clouds (Brüning and Tost, 2025). Most trajectories contain a single convective core (70 %), while the proportion of multi-core systems declines with increasing core count (Figure 2, c). Cloud frequency is higher in March–May (MAM) than in June–August (JJA) (Figure 2, d). The diurnal variability is less pronounced than these monthly differences along the period (Figure 2, e). We observe a high proportion of clouds have a lifetime between 0–3 h (42 %) or 3–6 h (37 %). Hence, about 80 % of the cloud tracks last for less than 6 h. The proportion of cloud tracks with a longer lifetime is considerably lower (Figure 2, f).

While this framework enables a seamless tracking of convective systems along the ML-based 4D time series, it remains subject to several limitations. The predicted data display a ML-based extrapolation of the received CloudSat CPR reflectivities. Hence, they include uncertainties connected to the ML model, such as the blurriness of predictions induced by the loss function which optimizes towards the mean. We receive few information on thin ice clouds due to a reduced sensitivity of the CloudSat CPR to ice clouds in high altitudes (Sassen and Wang, 2008). Moreover, the detection framework rests on an object-based perspective to investigate atmospheric processes. We note the identified trajectories may underlie simplifications caused by an inherent subjectivity of the thresholds applied in the cloud detection step. Nevertheless, the approach may help to bring further insights into the structure and organisation of convective clouds.

3 Method

3.1 Quantifying convective organisation

Convective organisation describes the contrast between convective cells randomly distributed in space and time from those clustering together inducing a stronger convective organisation (Pendergrass, 2020). While there exist various organisation indices to quantify the spatial clustering, each index alone may not sufficiently characterise convective organisation (Stubenrauch et al., 2023). Instead, all indices have specific limitations, such as a sensitivity to the mean cloud area or to the number

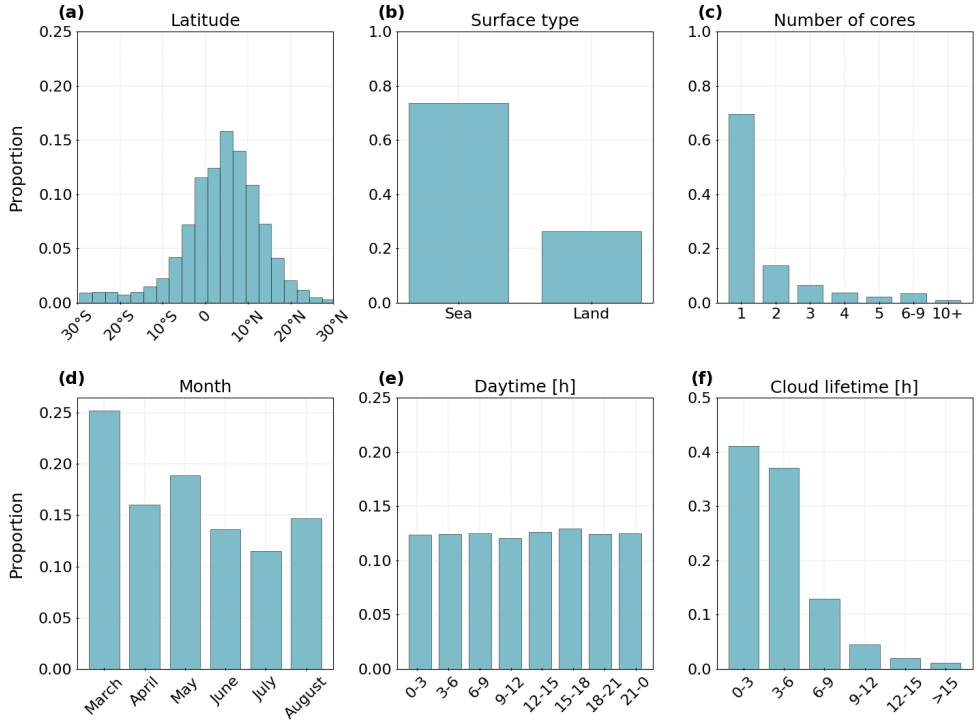

Figure 2. Summary of cloud tracks retrieved between March–August 2019 (n = 354,073). We show the spatial and temporal distribution of the data based on **(a)** the latitude grouped in 3° intervals between 30° S and 30° N, **(b)** the surface type derived from a land-sea mask, **(c)** the number of cores, **(d)** the month, **(e)** the daytime, and **(f)** the cloud lifetime.

of individual objects. In response, we chose a combination of three organisation indices (SCAI, COP, ROME). All indices are designed to work on 2D data. Their input is a binary field, in this case the cloud mask derived in Sect. 2.3, representing the location of labeled convective objects (Semie and Bony, 2020). We calculate the three organisation indices for the AOI between 30°W–30°E and 30°N–30°S at each timestep of 15 minutes (Figure 1, c).

The first index is the simple-to-compute and straightforward Simple Convective Aggregation Index (SCAI). SCAI describes the ratio of the degree of convective disaggregation to a potential maximal disaggregation within a domain (Tobin et al., 2012). The index is unitless and inversely proportional to the number of grid boxes. SCAI compares the number of objects in the domain (N) and the geometric mean distance (D_0) between the centroid positions of all possible object pairs to the possible maximum number of objects that can exist in the domain (N_{max}) and the characteristic domain size (L).

$$\mathrm{SCAI} = \frac{N D_0}{N_{max} L} 1000. \quad (1)$$

SCAI is a unitless index between 0 and infinity whereas lower values point towards a stronger convective organisation. By design, calculating SCAI requires the presence of multiple cloud clusters.

The Convective organisation Potential (COP) was developed by White et al. (2018) as an adaptation of the I_{org} index. It assumes objects that are larger and closer together are more likely to interact with each other. In contrast to SCAI, the index takes the cloud size into account. COP uses the number of objects (N), the area of the i-th object (A_i) and the j-th object (A_j), and the distance between the centroids of the i-th and the j-th object (d_{ij}). It adds the characteristic domain size (L) and the total image size (L_2). The index is defined by

$$\mathrm{COP} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{\sqrt{A_i/\pi} + \sqrt{A_j/\pi}}{d_{ij}} \tag{2}$$

which is the mean over all the possible pairs of the interaction potential. COP is a positive and unitless index between 0–1 whereas higher values indicate a stronger convective organisation. Larger and closer objects have a higher increase in COP than small and widespread objects (Pscheidt et al., 2019).

Additionally, we calculate the Radar Organisation MEtric (ROME). The index considers the average size, proximity, and size distribution of convective clouds. Initially, it was designed to analyse radar observations. However, it also worked well with other data (Bläckberg and Singh, 2022). The index assesses connections between pairs of continuous convective regions and assigns a weight to each pair that increases with their respective areas and decreases with their separation distance. The weight is equal to the area of the larger contiguous convective region plus a contribution from the smaller contiguous convective region that depends on the separation distance (Retsch et al., 2020). It employs the smallest distance between the edges of the i-th and the j-th object in the domain ($\tilde{d}_{ij}$) to define

$$\mathrm{ROME} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \cdot \left[A_{ij}^{(max)} + A_{ij}^{(min)} \cdot min\left(1, \frac{A_{ij}^{(min)}}{\tilde{d}_{ij}^2}\right)\right] \tag{3}$$

where $A_{ij}^{(max)}$ = max(A_i,A_j) and $A_{ij}^{(min)}$ = min(A_i,A_j). ROME is a positive index measured in units of area. Its value consists of a contribution from the mean area of contiguous convective regions and the distribution of sizes and interaction between different contiguous convective regions. The index is positive, with an increasing ROME value corresponding to a stronger aggregation.

While SCAI and COP are easy to compute, the calculation of ROME is less convenient. Since it has been designed to retrieve information from radar reflectivities, we include the index in our study. In contrast to SCAI and COP, ROME may also be computed when only a single object is present. As evaluated by, e.g., Mandorli and Stubenrauch (2024) and Biagioli and Tompkins (2023), each index has its own strengths and weaknesses. SCAI is insensitive to the size of the objects and mainly dominated by the variability in the number of clouds. However, it is less affected by shifts in time and space which induce high fluctuations of the index values, e.g, due to changes in the resolution of the input image or between two consecutive time steps. In contrast, the calculation of COP includes the object area. While COP correctly increases with the proximity and size, it is sensitive to noise caused in a domain with only a few objects. The index is correlated to the image resolution and shows a high variability for consecutive time steps. While ROME is more noise-safe and independent of the dataset resolution, it strongly

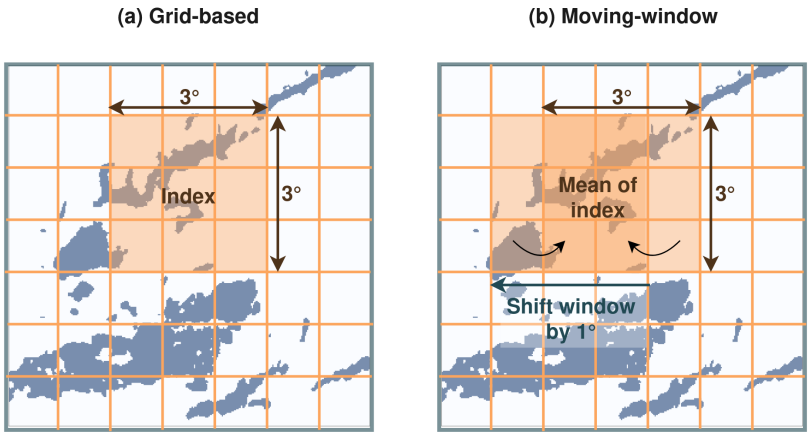

Figure 3. Visualisation of the moving-window approach used to calculate the organisation indices from the labeled 3D cloud mask. When using a fixed grid cell size **(a)**, clouds may be split at the borders leading to an enhanced small-scale value variability between the subsets. In this study, we employ a moving-window which iterates along the grid cells with a kernel size of 1° x 1°. At each iteration, we update the index value by calculating the mean between the index at the former and current subset **(b)**. In contrast to **(a)**, the indices are less influenced by a single grid cell and rather represent the average composed of all window locations.

connects to the object size. Compared to SCAI and COP, ROME shows a lower variability along consecutive time steps and it is less sensitive to the proximity of objects. Despite these limitations, we employ these indices that have been applied before in our studies to retrieve comparable results. However, building an adapted methodology for assessing convective organisation may benefit future research.

3.2 Calculating grid-based organisation indices

To assess regional variability in convective organisation, we refrain from computing organisation indices over the entire domain. Instead, the AOI is partitioned into overlapping 3° × 3° grid cells (e.g., Semie and Bony (2020); Tobin et al. (2012)). Given that the spatial extent and number of convective cloud elements affect the resulting index values, it may be beneficial to mitigate artifacts arising from cloud systems intersecting grid boundaries. In response, we implement a moving-window approach. The initial window is anchored at the northwestern corner of the AOI (27°–30° N, 27°–30° W) and is incrementally shifted by 1° in both the zonal and meridional directions (Figure 3). For each time step, the spatial organisation indices (SCAI, COP, and ROME) are computed within a 3° × 3° window. To enhance statistical robustness and reduce sensitivity to window placement, we calculate a local mean across adjacent overlapping windows, assigning the averaged value to the central grid cell. This approach may reduce boundary-related discontinuities and contribute towards a more stable representation of convective structure, particularly in regions where cloud systems span multiple windows (Jin et al., 2022).

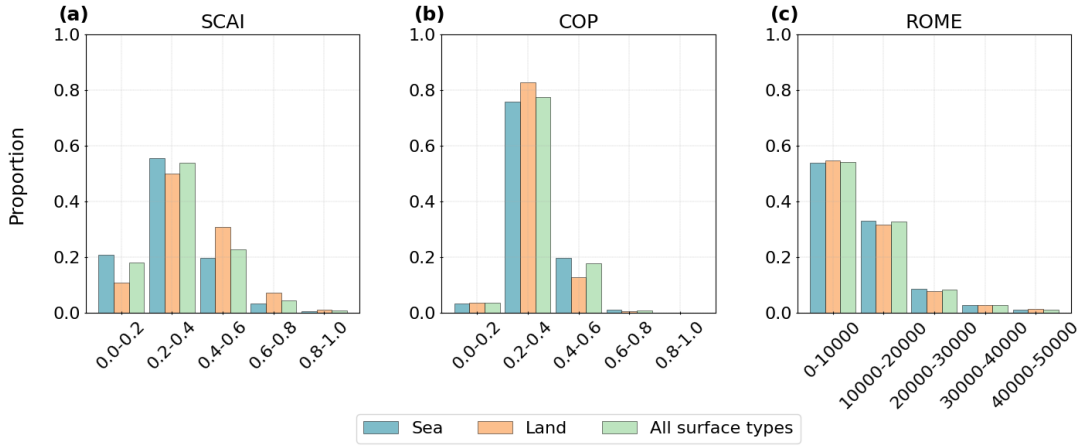

Figure 4. Distribution of convective organisation indices grouped by the surface type (for clouds over the sea, over land, and for all cloud tracks independent of the surface type). We see the frequencies of **(a)** SCAI, **(b)** COP, and **(c)** ROME.

4 Results

4.1 Distribution of organisation indices

This section analyses the spatial and temporal distributions of the three convective organisation indices: SCAI, COP, and ROME. Lower SCAI values (or higher COP and ROME values) are indicative of enhanced convective clustering, reflecting stronger spatial organisation. Conversely, high SCAI (low COP or ROME) values correspond to more scattered convective structures, implying weaker organisation (Biagioli and Tompkins, 2023).

Figure 4 (a) shows that SCAI values range between 0 and 1, with a peak concentration between 0.2–0.4. Oceanic regions have a slightly higher frequency of SCAI values lower than 0.4, whereas values higher than 0.4 are more common over land. This finding may suggest SCAI detects stronger convective organisation over water. COP values are mainly distributed between 0.2 and 0.6. Over the ocean, values above 0.4 are more frequent, whereas over land, lower values dominate — again pointing to stronger convective organisation over the ocean (Figure 4, b). ROME displays a right-skewed distribution, with most values falling below 20,000. Differences between land and ocean are minor compared to SCAI or COP (Figure 4, c). Overall, the results may indicate a marginally stronger convective organisation over oceanic regions, with ROME showing the weakest land–sea contrast.

Figure 5 compares the diurnal cycle, changes to core numbers, and latitudinal averages of the indices over land and ocean within the 30° S–30° N domain. For SCAI, we find predominantely lower values over land throughout the day. The diurnal cycle exhibits minima between 09:00–12:00 UTC and 21:00–00:00 UTC, particularly over land. SCAI increases between 00:00–06:00 UTC and 12:00–21:00 UTC (Figure 5, a). COP shows a weaker temporal variability than SCAI but with values

consistently suggesting higher organisation over the ocean (Figure 5, d). Diurnal variations in SCAI and COP reach up to 10 % of the indices' scales. ROME shows daytime (06:00–18:00 UTC) and nocturnal (00:00–03:00 UTC) peaks over land and mostly nocturnal peaks (21:00–06:00 UTC) over the ocean (Figure 5, g). Collectively, the indices indicate maximum convective organisation occurs over land in the afternoon and over the ocean in the night and early morning; minima occur at night over land and from noon to afternoon over the ocean. SCAI and ROME decrease with increasing numbers of convective cores (Figure 5, b, h). For ROME, organisation decreases up to five cores but increases beyond six, particularly over land (Figure 5, h). COP, by contrast, remains largely unaffected by core number as it points out only a slight decrease of convective organisation with increasing core numbers and an increase for clouds with more than six cores (Figure 5, e). This finding suggests for SCAI a stronger convective organisation for higher core numbers, which opposes the results for COP and ROME. Latitudinally, all indices show stronger organisation near the equator, although the spatial variability differs for the three indices. As SCAI is sensitive to object numbers, a higher frequency of detected clouds near the equator and less clouds near the borders of the AOI may contribute to the variability of the index (Figure 2, Figure 5, c). COP varies less with latitude, whereas we observe slightly higher values between 20° S–20° N (Figure 5, f). For ROME, we find the highest variability between latitudinal averages and surface types with peaks over land between 20° S and the equator, and over oceanic regions near the equator and between 20°–30° S (Figure 5, i). Compared to other regions in the domain, the results show a considerably stronger convective organisation over the southern Atlantic Ocean (30° S) for SCAI and ROME.

4.2 Spatial patterns and statistical relationships

Figure 6 presents the spatial distribution of the three organisation indices (SCAI, COP, ROME), along with associated cloud and core properties, interpolated onto a 3° × 3° grid and displayed as latitudinal cross-sections. Distinct regional patterns emerge across the AOI, highlighting potential links between convective organisation and cloud structure. Near the equator - particularly over continental Africa - higher SCAI values may coincide with a smaller cloud area, elevated cloud top height (CTH), and taller convective cores. In contrast, lower SCAI values are found primarily over the Atlantic Ocean and in subtropical zones of northern and southern Africa (15°–30° N/S). These regions are characterized by larger cloud areas, a lower CTH and lower core heights (Figure 6, a, d, e, i). For the cloud lifetime, the number of cores, and the core area, we observe a less distinct connection. They show a high spatial variability along the AOI, whereas a longer cloud lifetime, a higher number of cores, and a larger core area may be related to a smaller cloud area, higher CTH, and higher core height in near-equator regions (< 15° N/S) and to a larger cloud area, lower CTH, and lower core height near the tropics (> 15° N/S) (Figure 6, d–i). COP exhibits low spatial variability, with most values between 0.2–0.5 (Figure 6, b). ROME, in contrast, displays pronounced spatial differences: high values occur between 15°–30° N/S, particularly over the Atlantic Ocean and near African coastlines, and near the equator over the Gulf of Guinea and continental Africa (Figure 6, c). Over the Sahel, clouds tend to be large, with numerous, wide but relatively shallow cores. Over the South Atlantic (15°–30° S), cloud systems exhibit large areas, long lifetimes, and a high number of cores. This pattern may reflect cloud clustering in the AOI may be influenced by oceanic circulation and adjacent landmasses (Atiah et al., 2023). Overall, regions with stronger convective organisation - indicated by low SCAI and high COP or ROME - tend to exhibit smaller clouds with low CTH and core heights. For the number of cores, the core area, and cloud

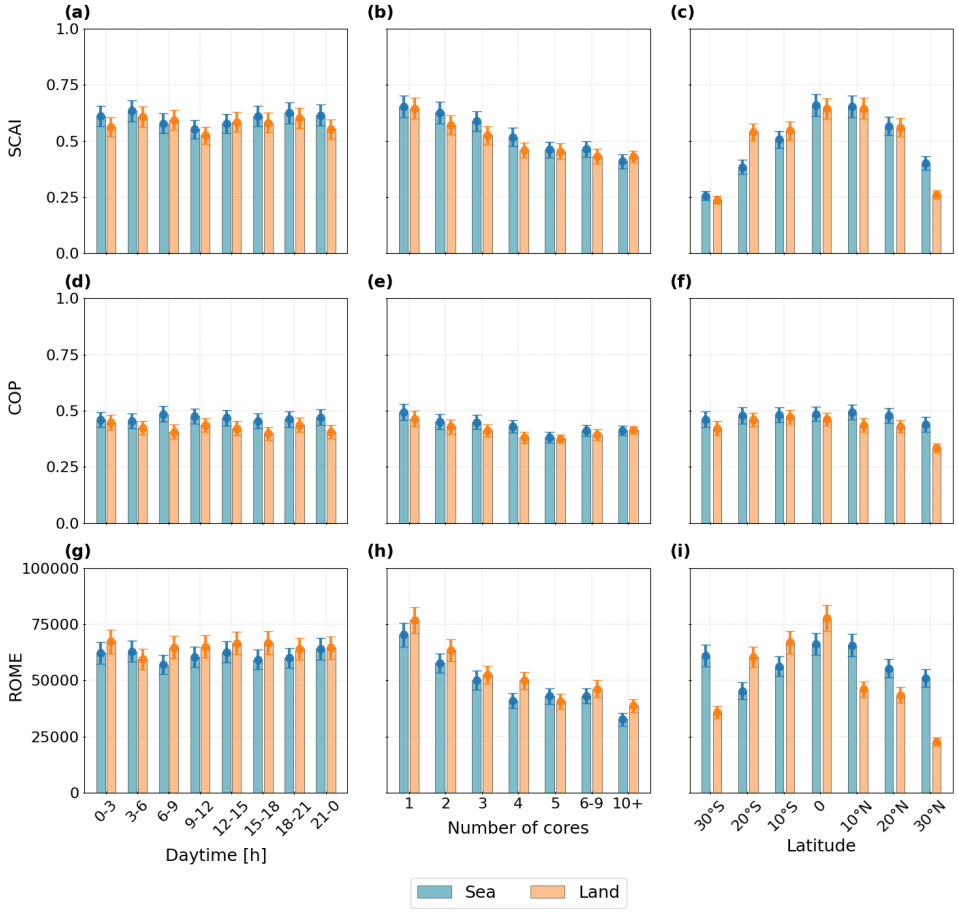

Figure 5. Comparison of the organisation indices **(a)–(c)** SCAI, **(d)–(f)** COP, and **(g)–(i)** ROME. The columns show the diurnal cycle (grouped in 3 h intervals), the number of cores, and the latitude (in 10° intervals) grouped by the surface type (land, sea). Vertical errorbars show the standard error of the mean.

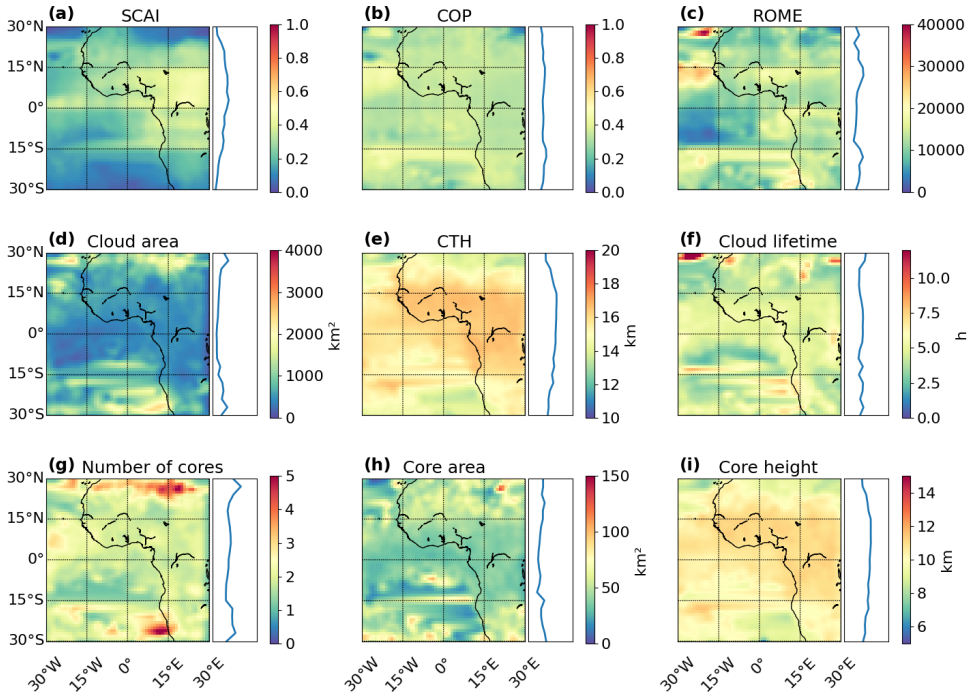

Figure 6. Mean values for **(a)** SCAI, **(b)** COP, **(c)** ROME, **(d)** the cloud area, **(e)** the CTH, **(f)** the cloud lifetime, **(g)** the number of cores, **(h)** the convective core size, and **(f)** the core height. The plot shows the spatial distribution in the AOI interpolated on a 3° x 3° grid (left) and the average for each latitude between 30° N and 30° S (n = 354,073).

lifetime, a higher regional variability may be apparent. These contrasts are most apparent between equatorial and subtropical regions.

To quantify the relationship between organisation indices and cloud properties, we compute Spearman's rank correlation coefficient R using data from all cloud tracks (Figure 7). The logarithmic distributions reveal a general skew toward low values for SCAI, ROME, cloud area, lifetime, number of cores, and core area. The correlation analysis shows that COP and ROME may be positively associated with cloud area, lifetime, CTH, number of cores, and core height (Figure 7, g–r). In contrast, SCAI is negatively correlated with all of these properties except for CTH and the core height (Figure 7, a–f). For the core area, we see a weak negative correlation to all indices. The findings suggest that stronger convective organisation may be statistically linked to larger, longer-lived cloud systems, a higher CTH and core height, and more cores. Interestingly, these statistical relationships contrast with some spatial patterns in Fig. 6. For instance, while higher ROME values spatially co-occur with smaller clouds and shorter lifetimes in some regions, correlation coefficients suggest that, overall, organisation increases with cloud area and

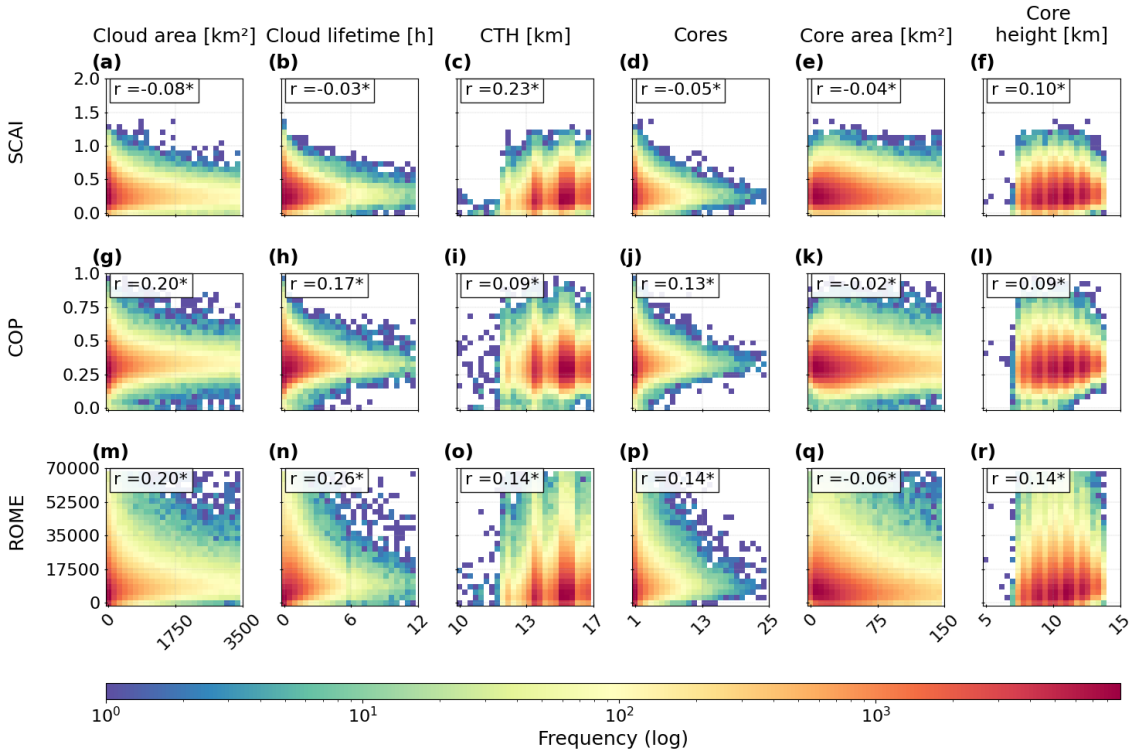

Figure 7. Histogram showing the logarithmic frequency distribution for the SCAI, COP, and ROME against the **(a, g, m)** the cloud area, **(b, h, n)** the cloud lifetime, **(c, i, o)** the CTH, **(d, j, p)** the number of cores, **(e, k, q)** the convective core size, and **(f, l, r)** the core height. In each histogram, we add the Spearman correlation coefficient R to quantify the strength of the relationship (n = 354,073) (significant with *: p < 0.1).

duration. However, most correlations are weak, with maximum coefficients around 0.26 between ROME and the cloud lifetime. They highlight the complex and regionally variable nature of these relationships.

4.3 Temporal variability of cloud properties and organisation indices

The previous analysis suggests the overall correlation between convective organisation indices and cloud/core properties is generally weak. In this section, we aim to capture changes in convective behaviour along the period that may help to explain observed patterns. For this purpose, we filter the dataset into two subsets between March to May (MAM, n = 212,984) and June to August (JJA, n = 141,089). Here, we analyse monthly means over land and ocean (Figure 8). Overall, differences between land and ocean typically span up to 10 % of each index's dynamic range (Figure 4). For the monthly changes, most variables do not exhibit a linear trend. SCAI, COP, and the number of cores remain relatively stable, while the CTH and core height

vary non-monotonically (Figure 8, a, b, e, g, i). SCAI generally decreases over the ocean and increases slightly over land until June, returning to near-March values by August (Figure 8, a). COP displays similar changes over land, while over the ocean, it increases marginally throughout the period (Figure 8, b). ROME exhibits the strongest variability, increasing over both surface types, especially over the ocean (Figure 8, c). Notably, average CTH, cloud lifetime, and core height are consistently higher over land, whereas cloud and core areas are larger over the ocean, particularly from May to August (Figure 8, d–f, h–i). The number of cores remains fairly constant across the time series (Figure 8, g). Over the ocean, we observe a steady increase in cloud and core area and a decrease in CTH. Core height peaks in May and July, followed by a decline in August. Over land, temporal changes are less pronounced, though the core area shows a slight dip until May and then rises again by August.

Figure 9 illustrates the mean differences between boreal spring (MAM) and boreal summer (JJA), calculated as MAM minus JJA. The data are interpolated onto a $3° \times 3°$ grid and averaged along latitudes. While SCAI shows only a weak monthly variability (Figure 8, a), we observe regional differences of up to $\pm$ 0.4 across the AOI. Notably, SCAI increases between 15°–30° N and decreases south of 15° N, especially over the Gulf of Guinea and central Africa (0°–15° S) (Figure 9, a). COP tends to increase between 0°–15° N and decrease north of 20° N during JJA, although these changes are generally small, remaining within $\pm$ 0.2. More pronounced decreases of up to -0.4 are seen south of 15° S, over the Sahel, and near the Canary Islands (Figure 9, b). ROME shows small localised decreases during JJA across northern Africa, the Canary Islands, and coastal southern Africa. In contrast, it increases between 15° N and 15° S, especially near the equator and around 15° S (Figure 9, c). The spatial patterns of the cloud properties partly align with (cloud area, cloud lifetime) or oppose (CTH, core height) those observed for the organisation indices. For instance, both cloud area and lifetime tend to increase in the Southern Hemisphere during JJA, though the magnitude and intensity of these changes vary considerably across the AOI. Over northern continental Africa and the Congo River basin, cloud area and lifetime decline from MAM to JJA (Figure 9, d, f). In contrast, the CTH increases north of 15° N and decreases south of 15° S during JJA (Figure 9, e). The number of cores reveals a less consistent pattern, with a high spatial variability. Increases are observed during JJA over the Atlantic Ocean, the West African coast, northern continental Africa, and the equatorial rainforests. Conversely, declines are noted over coastal areas north of 15° N and south of the equator (Figure 9, g). Similarly, the core area displays a rather fragmented spatial pattern across the AOI, with slightly larger values in the Northern Hemisphere and a particular increase south of 20° S during JJA (Figure 9, h). The core height broadly follows the same pattern as CTH, rising north of 15° N and declining south of 15° S. Additionally, core heights increase between 0°–10° S in boreal summer (Figure 9, i). Observed increases of the cloud area, core area, and cloud lifetime may coincide with a reduction in CTH, core height, and core number. However, there appear spatial and temporal variations which may reflect the influence of, e.g., local circulations and land–sea contrasts on convective development across the AOI.

We evaluate how the relationships between organisation indices and cloud/core properties evolve along the two seasonal subsets by comparing the correlation coefficients between MAM and JJA (Table 4). Overall, SCAI maintains negative correlations with cloud properties, while COP and ROME remain positively correlated. The direction of correlation does not change along the period, though some coefficients vary in strength. From boreal spring to summer, correlations between SCAI and cloud properties increase slightly - except for the CTH and core height. Correlations between COP and cloud properties predominantly increase, whereas the differences are lower than for SCAI. For ROME, we see an increase for the correlation to

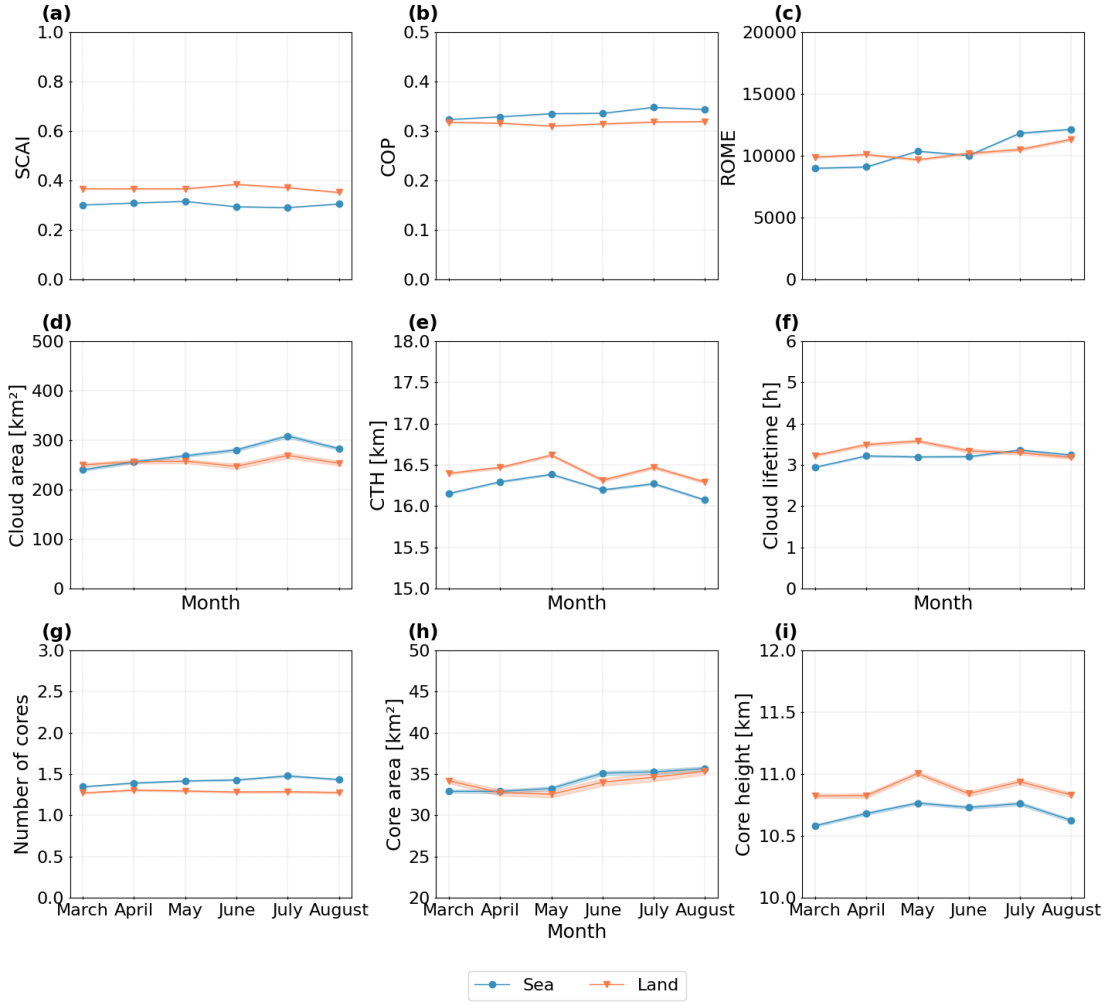

Figure 8. Monthly changes of the organisation indices and cloud and core properties between March and August 2019. We show **(a)** the SCAI, **(b)** the COP, **(c)** the ROME, **(d)** the cloud area, **(e)** the CTH, **(f)** the cloud lifetime, **(g)** the number of cores, **(h)** the convective core size, and **(f)** the core height grouped by the surface type. Line plots show the mean value (solid line) with a confidence interval of 95 %.

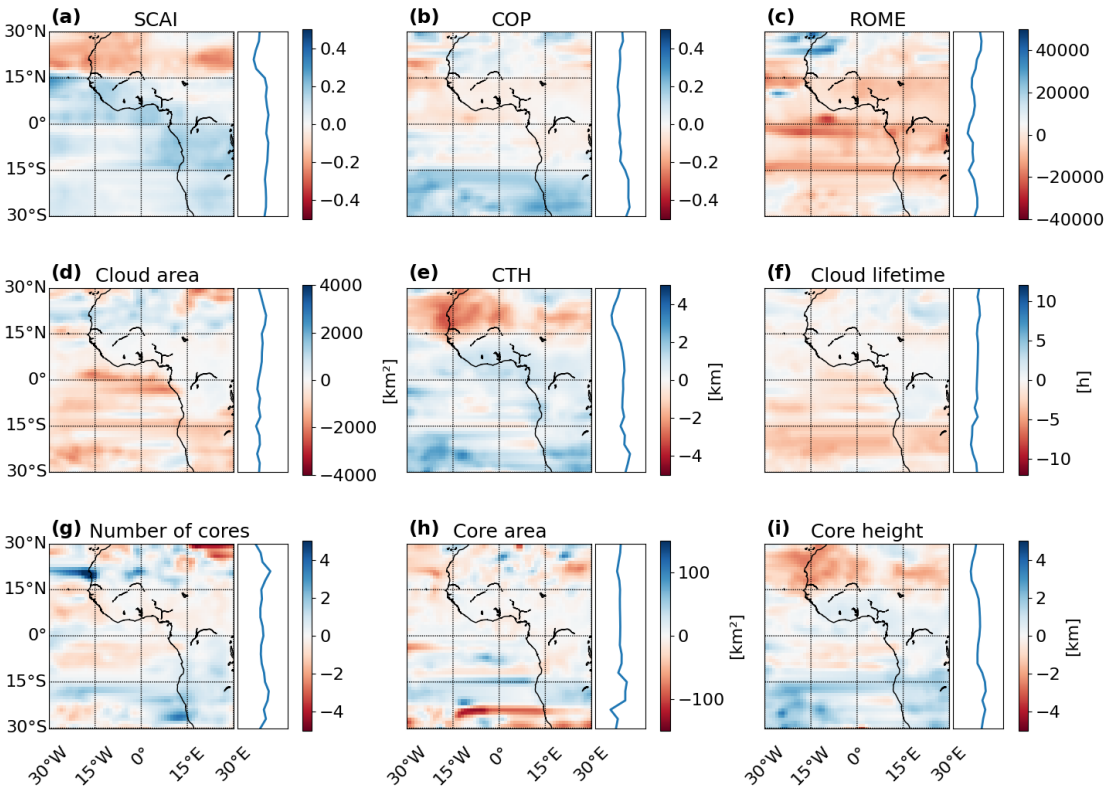

Figure 9. Changes between boreal spring (MAM, n = 212,984) and summer (JJA, n = 141,089) showing the average differences for MAM minus JJA. The plot shows the spatial distribution in the AOI interpolated on a 3° x 3° grid (left) and the average value for each latitude between 30° N and 30° S (right). Values are derived for **(a)** SCAI, **(b)** COP, **(c)** ROME, **(d)** the cloud area, **(e)** the CTH, **(f)** the cloud lifetime, **(g)** the number of cores, **(h)** the convective core size, and **(f)** the core height.

Table 4. Spearman's R for the SCAI, COP, and ROME against the cloud area, the cloud lifetime, the CTH, the number of cores, the convective core size, and the core height. The table shows the correlation coefficient for boreal spring (MAM) and summer (JJA) and the difference (MAM - JJA).

	SCAI			COP			ROME		
	MAM	JJA	Difference	MAM	JJA	Difference	MAM	JJA	Difference
Cloud area	-0.06	-0.13	0.07	0.12	0.12	0.00	0.09	0.07	0.02
Cloud lifetime	-0.01	-0.12	0.11	0.05	0.10	-0.05	0.08	0.11	-0.03
CTH	0.24	0.12	0.11	0.04	0.07	-0.03	0.04	0.07	-0.03
Number of cores	-0.04	-0.08	0.05	0.07	0.09	-0.02	0.05	0.05	0.0
Core area	-0.04	-0.04	0.00	0.02	0.01	0.01	0.01	-0.03	0.02
Core height	0.12	-0.01	0.11	0.05	0.08	-0.03	0.05	0.07	-0.02

the cloud lifetime, CTH, and core height, and a decrease to the cloud area and core area. However, these shifts are small, with changes up to 0.11 (SCAI vs. cloud lifetime, CTH, and core height). Despite apparent spatial patterns and temporal shifts in convective cloud organisation and structure as seen in Figs. 8 and 9, statistical relationships remain overall weak. These weak correlations suggest that relations may be affected by additional factors which were not integrated in our analysis, such as the large-scale circulation, interannual variations (caused by, e.g., El Niño-Southern Oscillation (ENSO)), or local topography.

4.4 Investigating effects of convective organisation

To identify regional patterns of convective organisation and their effects on cloud properties, we adopt a percentile-driven approach. There exist no universally defined thresholds to distinguish between weak and strong convective organisation. In response, we compute the 10th, 25th, 75th, and 90th percentiles based on the distribution of each organisation index (SCAI, COP, and ROME) using the cloud tracks between March to August 2019 (Table 5). These percentiles serve as thresholds to classify the data into subsets of weak and strong convective organisation, as induced by the interpretation of the indices: strong organisation may be related to low SCAI and high COP/ROME, weak organisation to high SCAI and low COP/ROME (Biagioli and Tompkins, 2023; Semie and Bony, 2020). Following, regions of strong convective organisation are defined as cloud tracks with an index value below the 10th percentile for SCAI or above the 90th percentile for COP and ROME. Conversely, regions of weak organisation correspond to values that lie above the 90th percentile for SCAI or below the 10th percentile for COP and ROME. To identify spatial and temporal patterns of convective organisation, we create two subsets from all data points in the dataset, whereas one represents the 10 % strongest convective organisation (Q10 for SCAI; Q90 for COP and ROME, hereafter: P90), and the other representing the 10 % weakest convective organisation (Q90 for SCAI; Q10 for COP and ROME, hereafter: P10). These may represent so-called "hotspots". We also define the interquartile range (IQR, values between the

Table 5. Percentiles for the organisation indices (SCAI, COP, ROME) derived from the time series between March and August 2019. The table contains the percentiles Q10, Q25, Q75, and Q90 which are used as thresholds to filter subsets of strong or weak convective organisation.

	Q10	Q25	Q75	Q90
SCAI	0.165	0.224	0.418	0.528
COP	0.237	0.278	0.381	0.443
ROME	3260.327	5652.496	14695.356	22659.608

25th–75th percentile) to represent a baseline, which is used to contrast the spatial distribution of average organisation against the identified hotspot regions.

4.4.1 Characteristics of percentile-based subsets

We filter the dataset by the percentiles from Table 5 to create the subsets of weak (P10) and strong (P90) convective organisation. Both subsets include 84,132 samples. Our analysis reveals that the frequency and location of convective clouds — and their strength of organisation — are not evenly distributed spatially or temporal. The majority of cloud tracks was detected between 10° S and 20° N (Figure 2). However, we observe distinct temporal and land–sea contrasts reflected in both P90 and P10. During March–May (MAM), strong convective organisation (P90) is more prevalent over land in the southern hemisphere and over ocean regions between 10°–30° S and 5° S–10° N. From June–August (JJA), P90 occurrences shift northward, peaking over land between 10° S–5° N and over ocean between 5°–15° N. A persistent local minimum appears around 0°–5° N in both seasons (Figure 10, a, c). In contrast, weak convective organisation (P10) is rare north of 15° N in boreal spring and south of 15° S in JJA. In MAM, it is more frequent over land from 10° S–10° N and over ocean between 0°–10° N. In JJA, we see an overall northward shift of the distribution (Figure 10, b, d).

Comparing the surface types of all cloud tracks and both percentile subsets, we observe a higher proportion of clouds over the ocean than over land for all datasets. However, there are differences within the surface-type distribution for the organisation-based subsets: when comparing all three datasets (all cloud tracks, P90, P10), strong convective organisation occurs about 5–15 % more frequently over the ocean, whereas the proportion of cloud systems with a weak convective organisation is about 10–15 % higher over land (Figure 11, a). P10 clouds are generally associated with fewer cores and shorter lifetimes than both P90 and the full dataset. They may be associated to a higher proportion to single-core clouds (15 % higher than P90) and clouds with a lifetime between 0–3 h (30 % higher than P90). We observe more clouds from P90 with a cloud lifetime of more than 3 hours. However, the longest lifetimes in the dataset may be found for clouds not connected to the percentile subsets (Figure 11, b, d). Clouds were detected slightly more frequently in MAM than JJA. In March, the proportion is especially high for P10 (15 % higher than P90). In contrast, occurrences of P90 are less common in MAM and increase in JJA (10 % higher than P10). These findings may indicate an increase of strong convective organisation during boreal summer (Figure 11, c).

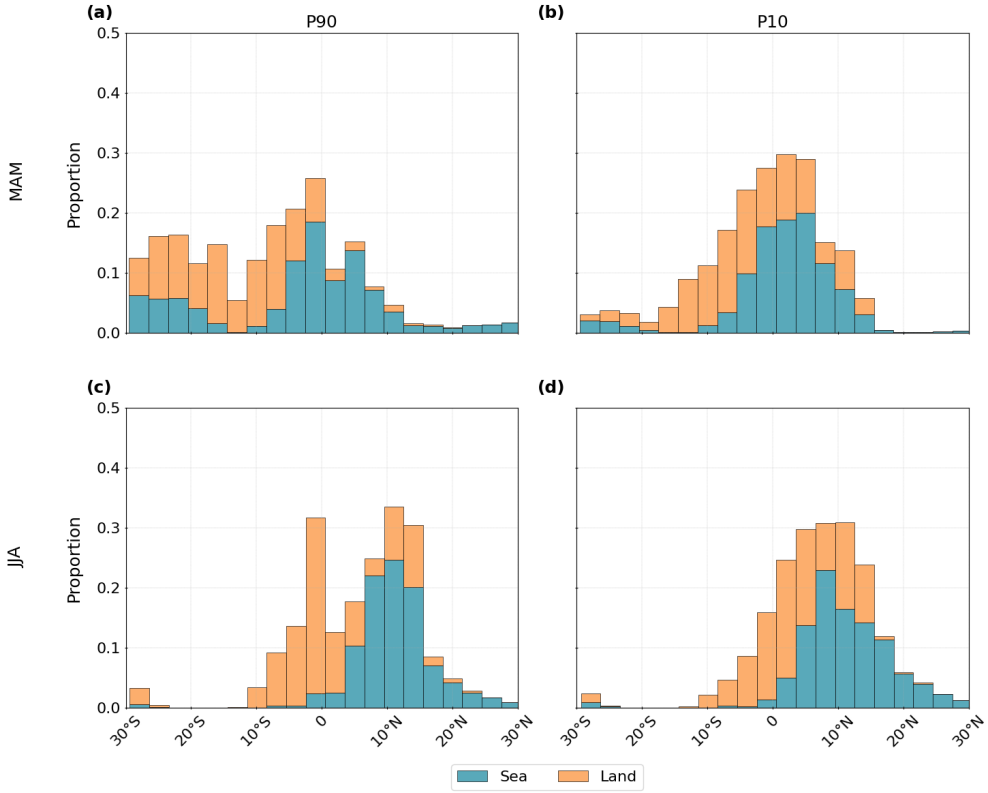

Figure 10. Distribution of detected clouds grouped in 3° intervals between 30° S and 30° N. The histograms show the proportions of cloud tracks grouped by the surface type (land, sea) for **(a)** the 10 % strongest convective organisation (P90, n = 84,132) and **(b)** the 10 % weakest convective organisation (P10, n = 84,132) between March–May (MAM) and June–August (JJA).

4.4.2 Relationship between organisation subsets and cloud properties

To explore how the relationship between cloud and core properties differs for weak (P10) and strong (P90) convective organisation, we compare the correlation coefficients between all cloud tracks and the two subsets. As noted in Sect. 4.2, SCAI tends to correlate negatively with cloud properties, while COP and ROME show positive associations. For all cloud tracks, correlations between the indices range from -0.08 to 0.26 (Figure 7). Figure 12 highlights that inter-index and intra-cloud property correlations are stronger than those between indices and cloud properties. Here, COP and ROME exhibit moderate to strong positive correlation, while COP and SCAI are moderately negatively correlated (Figure 12, a). SCAI and ROME show a weak to moderate inverse relationship. Among cloud and core properties, the strongest positive correlation is between cloud area and number of cores, followed by CTH and core height. The number of cores, core area, and core height are also moderately

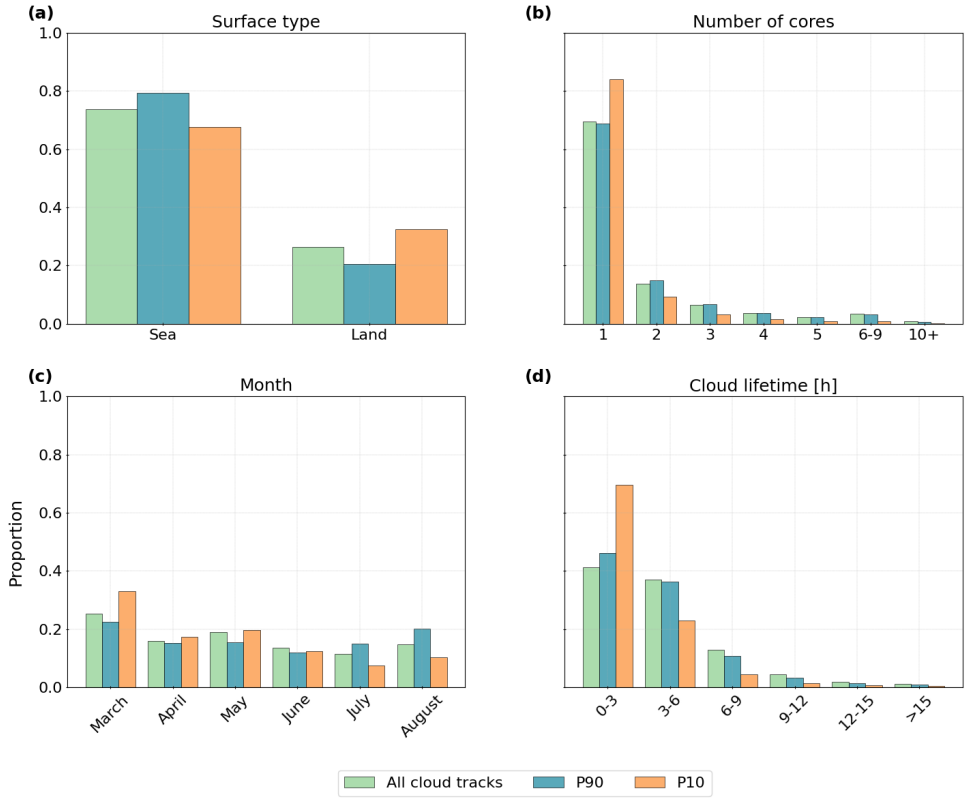

Figure 11. Comparison of the distributions between all cloud tracks, the 10 % strongest convective organisation (P90), and the 10 % weakest convective organisation (P10) for **(a)** the surface type derived from a land-sea mask, **(b)** the number of cores, **(c)** the month, and **(d)** the cloud lifetime after first detection.

correlated. Cloud lifetime, however, shows only weak to moderate associations with these properties. In the P90 subset, all three indices are positively correlated - a departure from the expected negative SCAI–COP/ROME relationship seen for all cloud tracks. Correlations between cloud and core properties in P90 remain largely similar to the full dataset, though some relationships (e.g., between COP/ROME and core height or area) strengthen slightly (Figure 12, b). In P10, we find similar property-to-property correlations, though the strength varies more. The strongest correlation remains between the number of cores and core area and the core and cloud height (Figure 12, c). For all data, we remain to find the strongest correlation between SCAI for the indices and CTH for the cloud/core properties. Uniquely, SCAI and ROME show a high positive correlation in both P10 and P90, despite being theoretically opposed in their interpretation of convective organisation (Section 3, Section 4.1). This apparent contradiction underscores the complexity of the indices, particularly when filtered by percentiles.

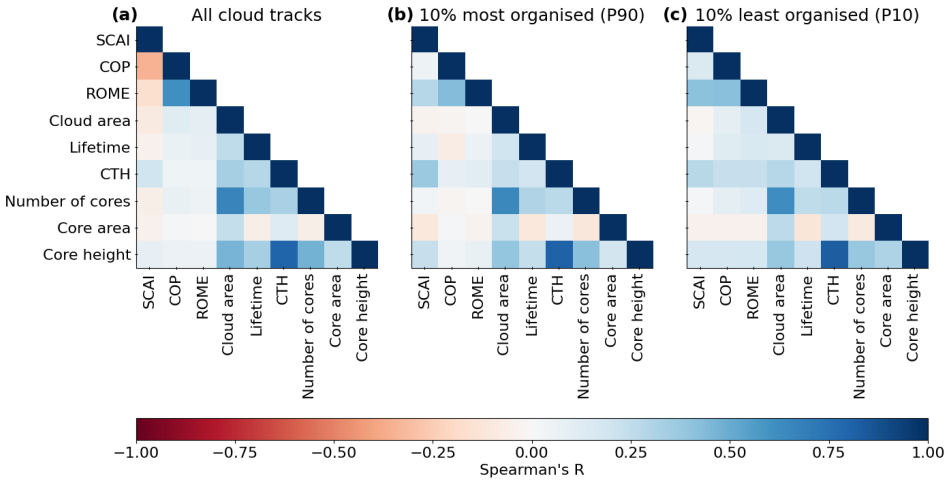

Figure 12. Correlation matrix for the organisation indices, cloud and core properties. We calculate Spearman's R to quantify the correlation coefficient on a scale between -1 and 1 for **(a)** the whole dataset of cloud tracks (n = 354,073), **(b)** the 10 % most organised clouds (P90, n = 84,132), and **(c)** the 10 % least organised clouds (P10, n = 84,132).

To assess whether differences between datasets are statistically significant, we compare parameter distributions for all cloud tracks, P90, and P10 subsets. We apply Welch's t-test, which may be more robust for unequal sample sizes (Derrick and White, 2016). For instance, we complement this with Cohen's D to estimate the effect size as small (< 0.2), medium (0.2–0.5), or large (higher than 0.8) (Cohen, 2013). The organisation indices show statistically significant differences across the three subsets, with large effect sizes for SCAI, COP, and ROME. Here, the effect size is largest between all data and P10 (Figure 13, a–c). Cloud and core properties exhibit more nuanced differences. The cloud area shows the largest effect size between all data and P10, while the CTH shows higher differences between all data and P90 (Figure 13, d, e). Compared to all cloud tracks, P90 clouds tend to be larger, with lower CTH, slightly shorter lifetimes, and slightly less, larger, and lower cores. P10 clouds are smaller, with a higher CTH, shorter lifetimes, fewer cores, and a larger core area and lower core height than clouds in the full dataset (Figure 13, d–i). For the number of cores, we find very low differences between all data and P90. As seen in Fig. 2, single-core clouds dominate the dataset. This skewness may affect statistics - in particular of data in P10 - which are heavily weighted toward fewer cores. Core area is larger in P90 and P10, whereas core height is lower in P90 and P10. However, for the core area, we observe only very small differences between the subsets (Figure 13, f–i).

While we detect statistically significant differences between percentile-based subsets and the dataset with all cloud tracks, the effect sizes for cloud and core properties remain mostly small to moderate. Our results indicate that strong convective organisation (low SCAI, high COP and ROME) tends to co-occur with larger cloud and core areas, slightly less and lower

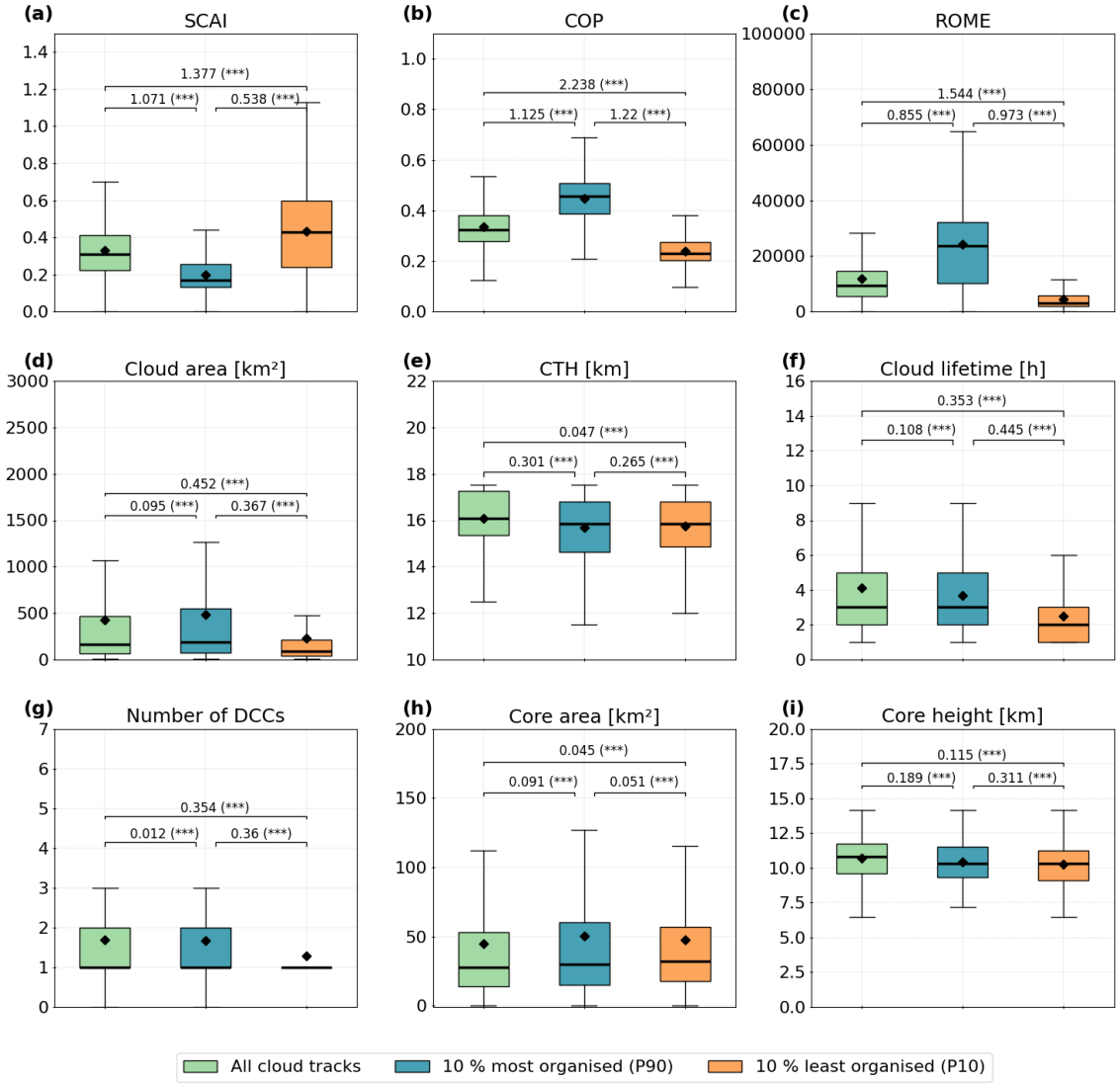

Figure 13. Boxplot showing the distribution of the convective organistaion indices, cloud and core properties for the whole dataset (n = 354,073), the 10 % most organised clouds (P90, n = 84,132), and the 10 % least organised clouds (P10, n = 84,132). We show the distribution of **(a)** the SCAI, **(b)** the COP, **(c)** the ROME, **(d)** the cloud area, **(e)** the CTH, **(f)** the cloud lifetime, **(g)** the number of cores, **(h)** the convective core area, and **(i)** the core height. The boxplot contains the median (bold black lines) and the arithmetic mean (black diamonds). Annotations depict the effect size measured by Cohen's D and the p-value derived from Welch's t-test (not significant (ns): p higher than 0.1, significant with *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$).

Table 6. Summary of differences between cloud and core properties for all cloud tracks against the percentile-based classification of weak (P10) and strong (P90) convective organisation. The table contains the arithmetic mean of all properties for the three datasets. We show in which direction the subset mean differs from all tracks (Direction) and the effect strength (Cohen's D) for all tracks compared to P90 or P10.

	All tracks		P90			P10	
	Arithmetic mean	Direction	Arithmetic mean	Cohen's D	Direction	Arithmetic mean	Cohen's D
SCAI	0.331	-	0.198	large	+	0.433	large
COP	0.334	+	0.449	large	-	0.240	large
ROME	11858.977	+	24274.897	large	-	4413.081	large
Cloud area	421.225	+	485.431	small	**-**	223.537	medium
CTH	16.077	**-**	15.696	medium	-	15.755	small
Cloud lifetime	4.119	-	3.691	small	-	2.501	medium
Number of cores	1.774	-	1.740	small	**-**	1.305	medium
Core area	44.976	+	50.074	small	+	47.636	small
Core height	10.704	-	10.428	small	-	10.266	small

cores, and slightly shorter lifetimes. The highest effect sizes may be found for the CTH, core height, and cloud lifetime. Weak organisation (high SCAI, low COP and ROME) is associated with smaller clouds, lower CTH, fewer cores, a smaller core area, lower core height, and shorter lifetimes. Here, we observe the highest effect sizes for the cloud area, number of cores, and cloud lifetime (Table 6). These findings - and the differences between the two percentile-based subsets - suggest that different aspects of cloud and core morphology may contribute to the strength of convective organisation.

4.4.3 Spatial distribution of percentiles

To identify how convective organisation may be spatially distributed for each of the three organisation indices (SCAI, COP, ROME), we filter the dataset of all cloud tracks by the percentiles (Q10, Q90) and we map the frequency of cloud occurrences across the area of interest (AOI) between 30° N–30° S and 30° W – 30° E. The data is interpolated using a 3° × 3° grid and smoothed with a Gaussian filter (sigma = 0.5). In addition to the percentiles Q10 and Q90, we visualise the interquartile range (IQR; 25th–75th percentile) for each organisation index. Frequency values (0–140 per grid cell) are colour-coded to represent absolute counts.

As shown in Sects. 4.1 and 4.2, high SCAI values - indicating weak convective organisation - are typically concentrated near the equator. In MAM, low SCAI values (Q10) occur over the equatorial Atlantic Ocean and land/sea areas south of 15° S. High values (Q90) appear over equatorial Africa (0°–15° N), especially in rainforest zones, and Cameroon. The IQR peaks near the equator, particularly over the Ivory Coast, Guinea, Benin, Angola's coast, and Lake Victoria (Figure 14, a–c). In boreal summer, values shift north to 0°–15° N, with SCAI Q10 regions over the Atlantic and coastal West Africa. High SCAI (Q90) values occur in MAM over the Congo and Central African Republic. The IQR also shifts north in JJA, with hotspots over the

West African plains, Jos Plateau, and Congo River basin (Figure 15, a–c). COP exhibits weaker spatial variability than SCAI or ROME. We detect clusters of low (Q10) COP near the equator in both seasons, over the Atlantic in MAM and across continental Africa in JJA. For high values (Q90) in MAM, strong peaks are found along West and Central African coasts and offshore in the Atlantic Ocean - many overlapping with regions of low COP, suggesting coexisting weak and strong organisation (Figure 14, d–f). In boreal summer, peaks of high COP (Q90) are concentrated over the Atlantic Ocean near Cape Verde and coastal zones between Senegal and Sierra Leone. Secondary peaks appear inland across West Africa (Figure 15, d–f). The IQR aligns closely around the equator but shifts northward in JJA, with dominant peaks over Central Africa's rainforest and minor peaks across the West African plains (Figures 14 b, e, 15, b, e). ROME shows greater latitudinal variability than SCAI and COP. In MAM, low values (Q10) values focus primarily along 15°–30° W near the equator, and secondarily between 15° S–15° N. High values (Q90) values are concentrated along the West African coast and between Cameroon and Gabon. IQR peaks are dispersed over the equatorial rainforest and coastlines (Figure 14, g–i). In JJA, low ROME (Q10) clusters around the Congo River and more diffusely across continental Africa. Peaks for high values of ROME (Q90) appear over the Jos Plateau, Congo River, and Atlantic. Like COP, ROME shows overlapping regions of weak and strong organisation over rainforests and oceans. IQR values peak between 0°–15° N and extend to coastal West Africa (Figure 15, g–i).

The spatial patterns of COP and ROME are closely aligned, with the 10th and 90th percentiles showing often spatial overlaps. SCAI has an inverse pattern due to its opposing index scale: regions with high COP/ROME may correspond to low SCAI (and vice versa). This inverse relationship is evident throughout the period, with all three indices exhibit consistent spatial patterns. The IQR maps, consistent across indices, reveal a northward shift of the indices which aligns with convective cloud occurrences during boreal summer as depicted for the percentile-based subsets in Fig. 10.

4.4.4 Identifying hotspots of convective organisation

In contrast to the former analysis, we examine the spatial distribution for clouds in the two subsets (P90, P10) (Section 4). These subsets of the 10 % strongest (P90) and the 10 % weakest (P10) convective organisation may help to identify cumulative hotspot regions averaged over the three indices. The data may allow us to analyse spatial patterns and temporal changes of convective organisation across two seasons from boreal spring (March to May, MAM) to summer (June to August, JJA). The occurrences are interpolated onto a 3° × 3° grid between 30° N–30° S and 30° W–30° E and smoothed using a Gaussian filter with a kernel size of 0.5.

In MAM, the highest proportion of strong convective organisation (P90) occurs over the Atlantic Ocean, with a notable concentration near the equator and between 15° and 30° S. Additional hotspots are found along the West African coastlines, the Gulf of Guinea. Moreover, we observe small peaks over the equatorial rainforest, Angola, and parts of the Sahel. Overall, most of the data points for the 10 % strongest convective organisation during boreal spring are located south of the equator (Figure 16, a). Weak convective organisation during MAM displays two primary clusters. The first is located over the equatorial Atlantic Ocean, particularly between 15° and 30° W. The second spans continental Africa, where more dispersed peaks emerge between Cameroon and the Congo River. Across the belt from 15° N to 15° S, the frequency of the 10 % weakest convective organisation is generally high (Figure 16, b). In boreal summer, the spatial distribution of strong convective organisation shifts

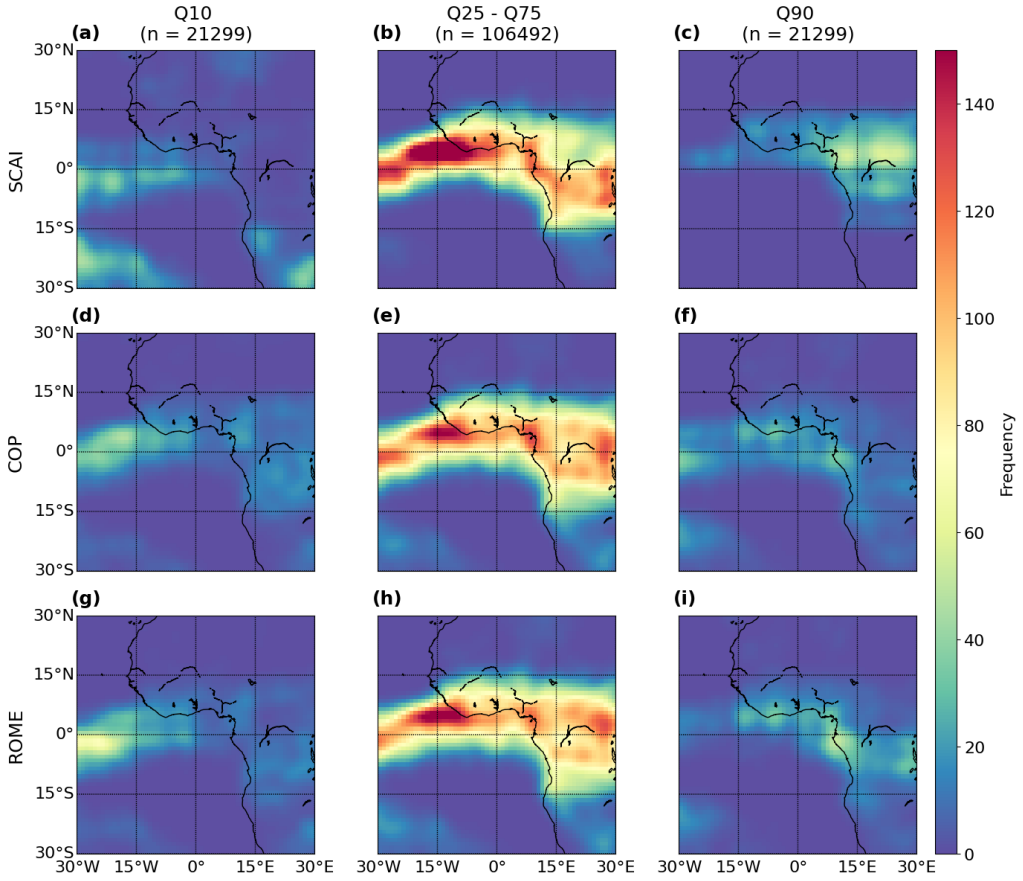

Figure 14. Spatial distribution of the percentiles Q10, Q25–Q75, and Q90 for the convective organisation indices **(a–c)** SCAI, **(d–f)** COP, and **(g–i)** ROME between March and May (MAM, n = 212,984). The values represent the frequency distribution interpolated on a 3° x 3° grid.

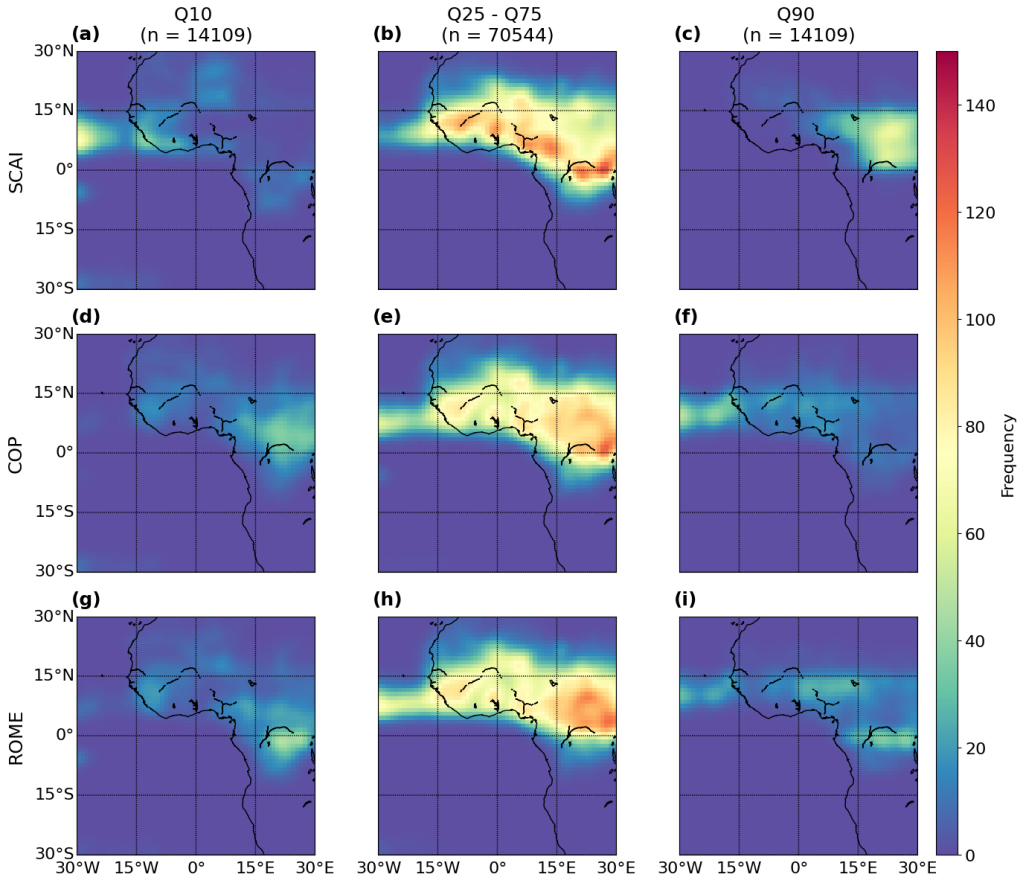

Figure 15. Spatial distribution of the percentiles Q10, Q25 - Q75, and Q90 for the convective organisation indices **(a–c)** SCAI, **(d–f)** COP, and **(g–i)** ROME between June and August (JJA, n = 141,089). The values represent the frequency distribution interpolated on a 3° x 3° grid.

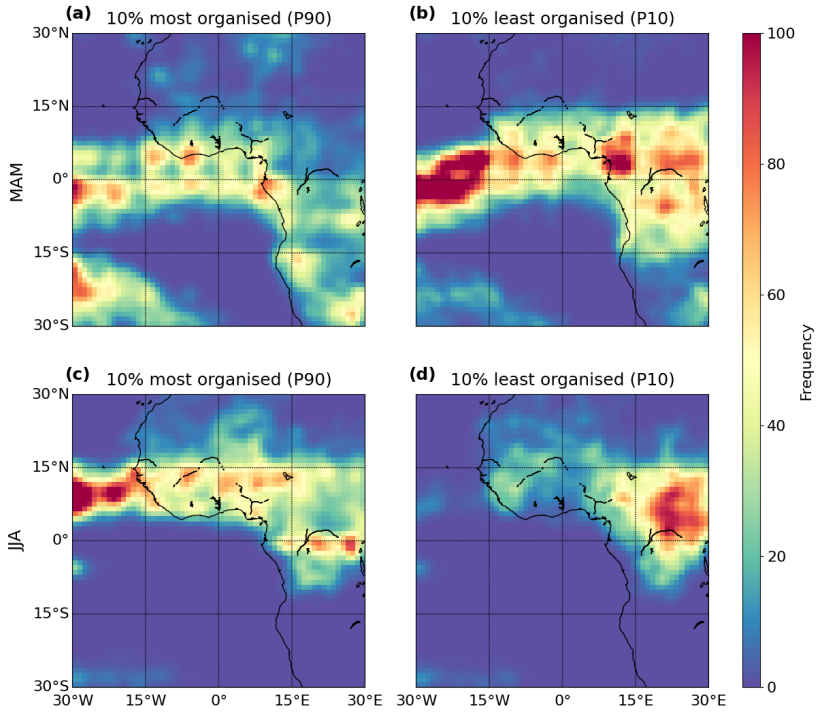

Figure 16. Spatial distribution of convective organisation based on an aggregation of the percentiles for SCAI, COP, and ROME in boreal spring (MAM, upper row) and summer (JJA, lower row). Clouds are grouped as the **(a) & (c)** strong organised or **(b) & (d)** weak organised using Q10 lowest (SCAI) or Q90 (COP, ROME). The values represent the frequency distribution interpolated on a 3° x 3° grid.

northward. Regions with a frequent occurrence of strong convective organisation emerge over the Atlantic Ocean and become more widespread across the West African plains, including areas around the Niger and Congo rivers (Figure 16, c). Weak organisation, on the other hand, is concentrated primarily over continental Africa, especially between 15° and 30° E, with a peak located just north of the Congo River (Figure 16, d).

As suggested in Sect. 4.4.3, we observe overlapping regions of weak and strong convective organisation throughout the period. In MAM, this overlap is evident over both ocean and land, whereas in JJA, it is mainly confined to continental Africa. Overall, cloud occurrences and spatial patterns suggest a shift between MAM and JJA which is consistent to the imbalance between cloud tracks over land and ocean observed in Fig. 11. In boreal spring, strong convective organisation is more frequently observed over the ocean, while weak organisation is distributed across both land and sea. By boreal summer, strong organisation becomes more prominent over land, and weak organisation is largely confined to the African continent. This migration of convective hotspots appears consistent with the northward movement of the ITCZ, as described in Atiah et al. (2023).

5 Discussion

5.1 Summary of key findings

Our analysis reveals that convective cloud occurrence and convective organisation vary considerably across both space and time. While the study spans only six months and does not provide a full climatology, the results highlight spatial and temporal changes of convective organisation during the period. Notably, the frequency of the 10 % strongest convective organisation increases during the boreal summer months (June to August, JJA), particularly north of the equator (Figure 11, Figure 16, c–d). In boreal spring (March to May, MAM), it may occur more frequently south of the equator (Figure 16, a–b). Between March and May, we observe a higher concentration of strongly organised convection over the Atlantic Ocean, primarily between 0° and 30° S. This peak shifts northward to between 0° and 15° N in JJA, with additional hotspots appearing over the equatorial rainforest and West Africa. Meanwhile, weakly organised convection tends to dominate over the Atlantic Ocean in boreal spring and shifts to continental Africa in JJA. These findings suggest a broader northward movement of convective cloud occurrences throughout the study period (Section 4.2, Section 4.4).

Correlations between convective organisation indices (SCAI, COP, ROME) and cloud/core properties suggest generally weak to medium relationships for all cloud tracks. SCAI is negatively correlated to the properties, except for the CTH and core height, while COP and ROME show positive correlations except for the core area. In all cases, the coefficients remain below 0.3 (Section 4.2). These correlations partly change for the 10 % strongest and 10 % weakest convective organisation. Within these subsets, we find the highest corelation coefficient between SCAI and the CTH for clouds with a strong convective organisation. However, the relationships between the cloud and core properties remain similar over all subsets (Section 4.4.2). In contrast, we observe pronounced changes in both the indices and the associated cloud characteristics along the period and across the AOI. These changes reflect a high variability for average values, though the correlation strength remains limited (Section 4.3). The distribution of cloud and core properties within identified hotspot regions differ from those observed in the full dataset of all cloud tracks. We analyse the effect size using Cohen's D to reveal how organisation strength may influence cloud characteristics. Compared to all cloud tracks, the cloud systems of the 10 % strongest organisation tend to have larger cloud and core areas, a lower CTH and core height, a shorter lifetime, and a lower number of convective cores. In contrast, weaker convective organisation may be typically associated with smaller clouds and larger cores, fewer cores, shorter lifetimes, and lower vertical extent. Strong convective organisation differs the most from all cloud tracks regarding the CTH and from weak organisation regarding the cloud lifetime. Between weak convective organisation and all cloud tracks, we identify the cloud area to have the highest effect size (Section 4.4.2). Hence, the cloud area appears to be more important to identify weak convective organisation, whereas strong convective organistaion may be stronger driven by the CTH. Despite these differences in the distribution of cloud and core properties, we detect partly the same direction for correlations in case of strong and weak convective organisation, highlighting the complexity of involved processes.

5.2 Spatio-temporal drivers of organisation

Our results show that convective organisation tends to be stronger for cloud properties typically associated with large convective systems containing multiple core regions, such as MCSs (Stubenrauch et al., 2023). In line with Brüning and Tost (2025), we observe that cloud area, lifetime, cloud top height (CTH), core area, and core height all grow with the number of convective cores (Figure 12). While multiple cores may enhance cloud longevity, promote cloud area growth, and strengthen vertical updrafts, the number of cores may also be a key factor in determining the strength of convective organisation. Interestingly, our findings contrast with Takahashi et al. (2017), as we observe stronger convective organisation - reflected in higher COP and ROME values and lower SCAI values - more frequently over the ocean. Over continental Africa, spatial patches of weak convective organisation appear in both seasons (Section 4.2, Section 4.4.4). However, the difference between land and ocean remain small and may partly stem from an uneven distribution of cloud tracks (Figure 2).

Spatial patterns of convective hotspots show differences over land and ocean. Around the equator, we observe a great share of cloud systems with a weak and strong organisation. This spatial overlap occurs between March–May, in particular over the ocean, and between June–August, especially over continental Africa (Section 4.4.4). Overall, the distribution of convective organisation varies notably between hemispheres and between equatorial and tropical zones (Section 4.3). These differences may be driven by a mix of local surface features (Vondou, 2012), monsoonal dynamics (Futyan and Genio, 2007), and topographic influences such as katabatic flows (Nicholson, 2018). Although our study reveals distinct geographical patterns, isolating the role of topography will require more targeted analysis.

We also detect a link between convective core occurrence and organisation that may follow the northward migration of the ITCZ in boreal summer. As the ITCZ shifts, it may alter regional circulation, surface energy balance, and moisture availability — particularly influencing cloud development over the northern Sahel and southern Sahara, as observed by, e.g., the spatial distribution of SCAI between June and August (Section 4.4.3). These changes may be associated with increased humidity, reduced subtropical subsidence, and deeper ascent within the tropical rainbelt (Fontaine and Philippon, 2000). Together with strengthened meridional pressure gradients (Lavaysse et al., 2009), they may contribute to the occurrence of large convective systems with multiple cores. This observation may be reflected in our results as a northward displacement of convective clouds and an increase in cloud area, core area, and core number over continental Africa in July and August (Section 4.2, Section 4.3). While our findings highlight the variability of convective organisation, the limited six-month time frame prevents a climatological interpretation. Extending this analysis across multiple years may provide deeper insights into the annual cycle of convective organisation and help refine operational forecasting and early-warning systems (Pendergrass, 2020).

5.3 Uncertainties and limitations

Our analysis may offer additional insights into the spatio-temporal distribution of convective organisation in the tropics. However, overall statistical relationships between convective organisation indices and cloud properties remain weak as we observe mostly small to medium effect sizes and low to moderate correlation coefficients (Section 4.2, Section 4.4.2). They highlight the complexity of quantifying convective organisation across space and time. Although our study may help to map patterns

of convective organisation across the AOI, gaining a deeper understanding of the underlying processes may require incorporating additional cloud parameters — such as cloud radiative properties - or associated precipitation rates (e.g., Stauffer and Wing (2024); Stubenrauch et al. (2023)). Moreover, addressing the imbalance between land and ocean cloud occurrences could strengthen the robustness of our findings. Currently, the cloud track distribution is skewed, with a heavy concentration near the equator. Notably, all indices indicate overlapping occurrences of both weak and strong organisation within the same regions — particularly over the Atlantic Ocean and continental Africa in boreal spring, and over the Congo River basin in summer. These spatial overlaps may obscure clearer statistical signals (Section 4.4).

Our dataset describes the three-dimensional structure of the tracked clouds, which may enable segmentation of cloud and core regions across horizontal and vertical dimensions at each point in time. Still, it is constrained by the performance of the ML model and the underlying tracking algorithm. Based on evaluations from Brüning et al. (2024) and Brüning and Tost (2025), the ML-predicted radar reflectivities exhibit a mean error of 2.99 dBZ. While suitable for building contiguous 3D cloud fields, the predictions struggle to accurately represent shallow cumulus and cirrus clouds — limitations inherited from the CloudSat CPR (Sassen and Wang, 2008). Incorporating higher-resolution satellite data or ground-based radar could enhance prediction accuracy. Other sources of uncertainty include the chosen thresholds for the detection algorithm and the skewed distributions underlying our percentile-based classifications of convective organisation. The indices themselves are sensitive to cloud object count (SCAI) or area (COP, ROME), which may affect spatial patterns, especially since equatorial convective clouds tend to be smaller and more frequent than those near the tropics (Section 2.5, Section 3.1). Additional uncertainties involve the influence of the terrain on cloud organisation (Biagioli and Tompkins, 2023). Future research could benefit from using combined indices or integrating temporal and spatial factors into a unified metric for 3D data. Our current method uses a moving-window, grid-based approach (Section 3.2), differing from past studies that partitioned the AOI into equal-area subsets (e.g., Tobin et al. (2012); Stubenrauch et al. (2023); Retsch et al. (2020)). While a moving window may reduce noise from small-scale fluctuations, its kernel size is manually chosen. To address this, we plan to explore unsupervised clustering techniques such as the Density Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996) or the extended Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) (Campello et al., 2013) as a more data-driven alternative. Zuo et al. (2022) successfully applied DBSCAN to identify cloud clusters in 3D radar data, while Kim et al. (2023) used the approach to derive precipitation probabilities from geostationary satellites. In future work, we aim to test whether such algorithms may reliably quantify convective organisation across space and time.

6 Conclusions

This study explores the spatial and temporal patterns of convective organisation in tropical West Africa using ML-based 3D radar reflectivities. We focus on the relationship between convective organisation, cloud structure, and core properties, using three organisation indices to statistically identify regional hotspots through a percentile-based classification.

Our analysis reveals that convective organisation tends to be slightly stronger over the ocean. However, differences between the indices over different surface types and along the period remain low and average around 10–15 %. We observe a consid-

erable spatial variability and a temporal shift in the distribution of strong convective organisation which appears linked to the northward migration of the ITCZ. From March to August, COP and ROME values increase while SCAI decreases, especially in the northern hemisphere, indicating an enhanced spatial clustering of convective clouds. Our regional analysis shows that the most cloud systems with a strong convective organisation during boreal spring (March–May) are concentrated over the Atlantic Ocean and coastal West Africa, predominantly in the southern hemisphere. In boreal summer (June–August), these hotspots shift inland toward the equatorial rainforest, West African Plains, and Sahel region. Notably, both weak and strong convective organisation frequently co-occur in the same regions, complicating statistical interpretation and underscoring the complexity of convective systems.

While correlations between organisation indices and cloud or core properties are generally weak to moderate, we observe that clouds with the 10 % strongest convective organisation tend to have larger cloud areas, lower cloud top and core heights, and more less but larger cores than the average cloud trajectory. In contrast, the 10 % weakest convective organisation are associated with smaller cloud areas, shorter lifetimes, fewer but larger cores, and a lower cloud and core height. Differences in CTH appear to be the most important for identifying cases of strong convective organisation. In contrast, the cloud area, cloud lifetime and number of cores appear to be a driver of weak convective organisation. Despite these findings, observed relationships and spatial patterns vary notably across the indices. The indices themselves often yield opposing results, reflecting their individual sensitivities and limitations. This variability is further influenced by the characteristics of the ML-based dataset. As the current study relies on 2D indices, developing a 3D organisation metric could provide a more accurate and holistic view. In summary, our findings highlight substantial variability in convective organisation across time and space. Given its influence on extreme weather, understanding these variations - and the mechanisms behind them - is crucial for improving climate risk assessments and forecasting capabilities in West Africa and beyond.

Code and data availability. The level 2B-GEOPROF CloudSat data used in this study are available at the CloudSat Data Processing Center at CIRA/Colorado State University and can be retrieved from http://www.CloudSat.cira.colostate.edu/order-data (CloudSat Data Processing Center, 2024). The Meteosat SEVIRI level 1.5 data used in this study is freely and openly available via the EUMETSAT Data Store at https://navigator.eumetsat.int/product/EO-:EUM:DAT:MSG:HRSEVIRI (EUMETSAT Data Services, 2024). The dataset of convective cloud tracks and organisation indices used in this study is available at the following repository: https://zenodo.org/records/14724869 (Brüning, 2025b). The material used to prepare this paper, including code used to perform analysis and that needed for the preparation of figures, is archived at https://zenodo.org/records/15607483 (Brüning, 2025a).

Author contributions. S.B and H.T. designed the study. S.B developed the code for performing the analysis and visualisation. S.B. and H.T. contributed to analysing and evaluating spatio-temporal patterns of convective cloud organisation. S.B. and H.T. wrote the draft of the paper. All authors have read and agreed to the published version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the project "Big Data in Atmospheric Physics (BINARY)", funded by the Carl Zeiss Foundation (grant P2018-02-003), and the Max Planck Graduate Center with the Johannes Gutenberg University of Mainz (MPGC). We thank EUMETSAT for providing access to the Meteosat SEVIRI imager data and the Cooperative Institute for Research in the Atmosphere, CSU, for providing access to the CloudSat 2B-GEOPROF data.

References

Atiah, W. A., Amekudzi, L. K., and Danuor, S. K.: Mesoscale convective systems and contributions to flood cases in Southern West Africa (SWA): A systematic review, Weather and Climate Extremes, 39, 100 551, https://doi.org/10.1016/j.wace.2023.100551, 2023.

Bao, J., Stevens, B., Kluft, L., and Muller, C.: Intensification of daily tropical precipitation extremes from more organized convection, Sci. Adv., 10, eadj6801, https://doi.org/10.1126/sciadv.adj6801, 2024.

Becker, T., Bechtold, P., and Sandu, I.: Characteristics of convective precipitation over tropical Africa in storm-resolving global simulations, Q. J. R. Meteorol. Soc., 147, 4388–4407, https://doi.org/10.1002/qj.4185, 2021.

Berthou, S., Rowell, D. P., Kendon, E. J., Roberts, M. J., Stratton, R. A., Crook, J. A., and Wilcox, C.: Improved climatological precipitation characteristics over West Africa at convection-permitting scales, Clim. Dyn., 53, 1991–2011, 2019.

Biagioli, G. and Tompkins, A. M.: Measuring Convective Organization, Journal of Atmospheric Sciences, 80, 2769–2789, https://doi.org/10.1175/JAS-D-23-0103.1, 2023.

Bläckberg, C. P. O. and Singh, M. S.: Increased Large-Scale Convective Aggregation in CMIP5 Projections: Implications for Tropical Precipitation Extremes, Geophys. Res. Lett., 49, e2021GL097 295, https://doi.org/10.1029/2021GL097295, 2022.

Bretherton, C. S., Blossey, P. N., and Khairoutdinov, M.: An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST, Journal of the Atmospheric Sciences, 62, 4273–4292, https://doi.org/10.1175/JAS3614.1, 2005.

Brune, S., Buschow, S., and Friederichs, P.: Observations and high-resolution simulations of convective precipitation organization over the tropical Atlantic, Q. J. R. Meteorol. Soc., 146, 1545–1563, https://doi.org/10.1002/qj.3751, 2020.

Brüning, S. and Tost, H.: A machine learning-based perspective on deep convective clouds and their organisation in 3D. Part I: Influence of deep convective cores on the cloud life-cycle, EGUsphere, 2025, 1–31, https://doi.org/10.5194/egusphere-2025-374, 2025.

Brüning, S.: Convective organisation indices based on 3D radar reflectivities, Zenodo [code], https://doi.org/10.5281/zenodo.14724869, 2025a.

Brüning, S.: Analysing spatial-temporal patterns of convective organisation from 3D data, Zenodo [data set], https://doi.org/10.5281/zenodo.15607483, 2025b.

Brüning, S., Niebler, S., and Tost, H.: Artificial intelligence (AI)-derived 3D cloud tomography from geostationary 2D satellite data, Atmos. Meas. Tech., 17, 961–978, https://doi.org/10.5194/amt-17-961-2024, 2024.

Campello, R. J. G. B., Moulavi, D., and Sander, J.: Density-Based Clustering Based on Hierarchical Density Estimates, in: Advances in Knowledge Discovery and Data Mining, pp. 160–172, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-642-37456-2, 2013.

CloudSat Data Processing Center: Data Products, CloudSat DPC [data set], https://www.cloudsat.cira.colostate.edu/data-products, accessed: 2024-12-12, 2024.

Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, Routledge, 2 edn., https://doi.org/10.4324/9780203771587, 2013.

Coppin, D. and Bony, S.: Physical mechanisms controlling the initiation of convective self-aggregation in a General Circulation Model, J. Adv. Model. Earth Syst., 7, 2060–2078, https://doi.org/10.1002/2015MS000571, 2015.

Dauhut, T., Chaboureau, J.-P., Escobar, J., and Mascart, P.: Giga-LES of Hector the Convector and Its Two Tallest Updrafts up to the Stratosphere, Journal of Atmospheric Sciences, 73, 5041–5060, https://doi.org/10.1175/JAS-D-16-0083.1, 2016.

Derrick, B. and White, P.: Why Welch's test is Type I error robust, TQMP, 12, 30–38, https://doi.org/10.20982/tqmp.12.1.p030, 2016.

Esmaili, R. B., Tian, Y., Vila, D. A., and Kim, K.-M.: A Lagrangian analysis of cold cloud clusters and their life cycles with satellite observations, J. Geophys. Res. Atmos., 121, 11,723–11,738, https://doi.org/10.1002/2016JD025653, 2016.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, August 2-4, Portland, Oregon, KDD'96, p. 226–231, AAAI Press, 1996.

EUMETSAT Data Services: High Rate SEVIRI Level 1.5 Image Data - MSG - 0 degree, https://navigator.eumetsat.int/product/EO:EUM:DAT:MSG:HRSEVIRI, accessed: 2024-12-12, 2024.

Fontaine, B. and Philippon, N.: Seasonal evolution of boundary layer heat content in the West African monsoon from the NCEP/NCAR reanalysis (1968–1998), Int. J. Climatol., 20, 1777–1790, https://doi.org/10.1002/1097-0088(20001130)20:14<1777::AID-JOC568>3.0.CO;2-S, 2000.

Futyan, J. M. and Genio, A. D. D.: Deep Convective System Evolution over Africa and the Tropical Atlantic, J. Clim., 20, 5041–5060, https://doi.org/10.1175/JCLI4297.1, 2007.

Haerter, J. O., Böing, S. J., Henneberg, O., and Nissen, S. B.: Circling in on Convective Organization, Geophys. Res. Lett., 46, 7024–7034, https://doi.org/10.1029/2019GL082092, 2019.

Hartmann, D. L., Hendon, H. H., and Houze, R. A.: Some Implications of the Mesoscale Circulations in Tropical Cloud Clusters for Large-Scale Dynamics and Climate, Journal of Atmospheric Sciences, 41, 113–121, https://doi.org/10.1175/1520-0469(1984)041<0113:SIOTMC>2.0.CO;2, 1984.

Heikenfeld, M., Marinescu, P. J., Christensen, M., Watson-Parris, D., Senf, F., van den Heever, S. C., and Stier, P.: tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets, Geosci. Model Dev., 12, 4551–4570, https://doi.org/10.5194/gmd-12-4551-2019, 2019.

Held, I. M., Hemler, R. S., and Ramaswamy, V.: Radiative-Convective Equilibrium with Explicit Two-Dimensional Moist Convection, Journal of Atmospheric Sciences, 50, 3909–3927, https://doi.org/10.1175/1520-0469(1993)050<3909:RCEWET>2.0.CO;2, 1993.

Houze, R. A.: Structure and Dynamics of a Tropical Squall–Line System, Mon. Wea. Rev., 105, 1540–1567, https://doi.org/10.1175/1520-0493(1977)105<1540:SADOAT>2.0.CO;2, 1977.

Igel, M. R., Drager, A. J., and van den Heever, S. C.: A CloudSat Cloud-Object Partitioning Technique and Assessment and Integration of Deep Convective Anvil Sensitivities to Sea Surface Temperature, J. Geophys. Res. Atmos., 119, 10 515–10 535, https://doi.org/10.1002/2014JD021717, 2014.

Jin, D., Oreopoulos, L., Lee, D., Tan, J., and Kim, K.-m.: A New Organization Metric for Synoptic Scale Tropical Convective Aggregation, J. Geophys. Res. Atmos., 127, e2022JD036 665, https://doi.org/10.1029/2022JD036665, 2022.

Kim, D., Kim, H.-J., and Choi, Y.-S.: Unsupervised Clustering of Geostationary Satellite Cloud Properties for Estimating Precipitation Probabilities of Tropical Convective Clouds, J. Appl. Meteorol., 62, 1083–1094, https://doi.org/10.1175/JAMC-D-22-0175.1, 2023.

Lavaysse, C., Flamant, C., Janicot, S., Parker, D. J., Lafore, J.-P., Sultan, B., and Pelon, J.: Seasonal evolution of the West African heat low: a climatological perspective, Clim. Dyn., 33, 313–330, https://doi.org/10.1007/s00382-009-0553-4, 2009.

Luo, Z., Liu, G. Y., and Stephens, G. L.: CloudSat adding new insight into tropical penetrating convection, Geophys. Res. Lett., 35, L19 819, https://doi.org/10.1029/2008GL035330, 2008.

Mandorli, G. and Stubenrauch, C. J.: Assessment of object-based indices to identify convective organization, Geosci. Model Dev., 17, 7795–7813, https://doi.org/10.5194/gmd-17-7795-2024, 2024.

Mapes, B. and Neale, R.: Parameterizing Convective Organization to Escape the Entrainment Dilemma, J. Adv. Model. Earth Syst., 3, M06 004, https://doi.org/10.1029/2011MS000042, 2011.

Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G.: Hydrometeor Detection Using Cloudsat—An Earth-Orbiting 94-GHz Cloud Radar, J. Atmos. Oceanic Tech., 25, 519–533, https://doi.org/10.1175/2007JTECHA1006.1, 2008.

Muller, C. and Bony, S.: What favors convective aggregation and why?, Geophys. Res. Lett., 42, 5626–5634, https://doi.org/10.1002/2015GL064260, 2015.

Nakajima, K. and Matsuno, T.: Numerical Experiments Concerning the Origin of Cloud Clusters in the Tropical Atmosphere, Journal of the Meteorological Society of Japan. Ser. II, 66, 309–329, https://doi.org/10.2151/jmsj1965.66.2_309, 1988.

Nicholson, S. E.: The ITCZ and the Seasonal Cycle over Equatorial Africa, BAMS, 99, 337–348, https://doi.org/10.1175/BAMS-D-16-0287.1, 2018.

Pendergrass, A. G.: Changing Degree of Convective Organization as a Mechanism for Dynamic Changes in Extreme Precipitation, Curr. Clim. Change Rep., 6, 47–54, https://doi.org/10.1007/s40641-020-00157-9, 2020.

Prein, A. F., Feng, Z., Fiolleau, T., Moon, Z. L., Núñez Ocasio, K. M., Kukulies, J., Roca, R., Varble, A. C., Rehbein, A., Liu, C., Ikeda, K., Mu, Y., and Rasmussen, R. M.: Km-Scale Simulations of Mesoscale Convective Systems Over South America—A Feature Tracker Intercomparison, J. Geophys. Res. Atmos., 129, e2023JD040 254, https://doi.org/10.1029/2023JD040254, 2024.

Pscheidt, I., Senf, F., Heinze, R., Deneke, H., Trömel, S., and Hohenegger, C.: How organized is deep convection over Germany?, Q. J. R. Meteorol. Soc., 145, 2366–2384, https://doi.org/10.1002/qj.3552, 2019.

Retsch, M. H., Jakob, C., and Singh, M. S.: Assessing Convective Organization in Tropical Radar Observations, J. Geophys. Res. Atmos., 125, e2019JD031 801, https://doi.org/10.1029/2019JD031801, 2020.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, edited by Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., vol. 9351, pp. 234–241, Springer International Publishing, Cham, 2015.

Sassen, K. and Wang, Z.: Classifying clouds around the globe with the CloudSat radar: 1-year of results, Geophys. Res. Lett., 35, L04 805, https://doi.org/10.1029/2007GL032591, 2008.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An introduction to Metetsat second generation (MSG), BAMS, 83, 977–992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, 2002.

Semie, A. G. and Bony, S.: Relationship Between Precipitation Extremes and Convective Organization Inferred From Satellite Observations, Geophys. Res. Lett., 47, e2019GL086 927, https://doi.org/10.1029/2019GL086927, 2020.

Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., Heikenfeld, M., Brunner, K. N., Bruning, E. C., Collis, S. M., Jackson, R. C., Leung, G. R., Pfeifer, N., Raut, B. A., Saleeby, S. M., Stier, P., and van den Heever, S. C.: *tobac* v1.5: introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena, Geosci. Model Dev., 17, 5309–5330, https://doi.org/10.5194/gmd-17-5309-2024, 2024.

Stauffer, C. L. and Wing, A. A.: How Does Organized Convection Impact Explicitly Resolved Cloud Feedbacks in the Radiative-Convective Equilibrium Model Intercomparison Project?, Journal of Advances in Modeling Earth Systems, 16, e2023MS003 924, https://doi.org/https://doi.org/10.1029/2023MS003924, 2024.

Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., Reinke, D., Partain, P., Mace, G. G., Austin, R., L'Ecuyer, T., Haynes, J., Lebsock, M., Suzuki, K., Waliser, D., Wu, D., Kay, J., Gettelman, A., Wang, Z., and Marchand, R.: CloudSat mission: Performance and early science after the first year of operation, J. Geophys. Res. Atmos., 113, https://doi.org/10.1029/2008JD009982, 2008.

Stubenrauch, C. J., Mandorli, G., and Lemaitre, E.: Convective organization and 3D structure of tropical cloud systems deduced from synergistic A-Train observations and machine learning, Atmos. Chem. Phys., 23, 5867–5884, https://doi.org/10.5194/acp-23-5867-2023, 2023.

Takahashi, H., Luo, Z. J., and Stephens, G. L.: Level of neutral buoyancy, deep convective outflow, and convective core: New perspectives based on 5 years of CloudSat data, J. Geophys. Res. Atmos., 122, 2958–2969, https://doi.org/10.1002/2016JD025969, 2017.

Takahashi, H., Luo, Z. J., Stephens, G., and Mulholland, J. P.: Revisiting the Land-Ocean Contrasts in Deep Convective Cloud Intensity Using Global Satellite Observations, Geophys. Res. Lett., 50, e2022GL102 089, https://doi.org/10.1029/2022GL102089, 2023.

Tobin, I., Bony, S., and Roca, R.: Observational Evidence for Relationships between the Degree of Aggregation of Deep Convection, Water Vapor, Surface Fluxes, and Radiation, J. Clim., 25, 6885–6904, https://doi.org/10.1175/JCLI-D-11-00258.1, 2012.

Vondou, D. A.: Spatio-Temporal Variability of Western Central African Convection from Infrared Observations, Atmos., 3, 377–399, https://doi.org/10.3390/atmos3030377, 2012.

White, B. A., Buchanan, A. M., Birch, C. E., Stier, P., and Pearson, K. J.: Quantifying the Effects of Horizontal Grid Length and Parameterized Convection on the Degree of Convective Organization Using a Metric of the Potential for Convective Interaction, Journal of Atmospheric Sciences, 75, 425–450, https://doi.org/10.1175/JAS-D-16-0307.1, 2018.

Wing, A. A., Emanuel, K., Holloway, C. E., and Muller, C.: Convective Self-Aggregation in Numerical Simulations: A Review, Surv. Geophys., 38, 1173–1197, https://doi.org/10.1007/s10712-017-9408-4, 2017.

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M.-S., Arnold, N. P., Bony, S., Branson, M., Bryan, G. H., Chaboureau, J.-P., De Roode, S. R., Gayatri, K., Hohenegger, C., Hu, I.-K., Jansson, F., Jones, T. R., Khairoutdinov, M., Kim, D., Martin, Z. K., Matsugishi, S., Medeiros, B., Miura, H., Moon, Y., Müller, S. K., Ohno, T., Popp, M., Prabhakaran, T., Randall, D., Rios-Berrios, R., Rochetin, N., Roehrig, R., Romps, D. M., Ruppert Jr., J. H., Satoh, M., Silvers, L. G., Singh, M. S., Stevens, B., Tomassini, L., van Heerwaarden, C. C., Wang, S., and Zhao, M.: Clouds and Convective Self-Aggregation in a Multimodel Ensemble of Radiative-Convective Equilibrium Simulations, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 138, https://doi.org/https://doi.org/10.1029/2020MS002138, 2020.

Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., and Yorty, D. P.: Where are the most intense thunderstorms on earth?, BAMS, 87, 1057–1072, https://doi.org/10.1175/BAMS-87-8-1057, 2006.

Zuo, Y., Hu, Z., Yuan, S., Zheng, J., Yin, X., and Li, B.: Identification of Convective and Stratiform Clouds Based on the Improved DBSCAN Clustering Algorithm, Adv. Atmos. Sci., 39, 2203–2212, https://doi.org/10.1007/s00376-021-1223-7, 2022.

4 Conclusions and outlook

This section contains a summary of the presented thesis and an outlook on future potentials and challenges. Overall, the project explores how ML techniques can address open questions in atmospheric sciences, focusing specifically on improving the representation of convective clouds. Modern ML algorithms, combined with large volumes of satellite data, are used to enhance our understanding of convective cloud development and their 3D characteristics. Currently, vertically resolved cloud properties are mainly obtained from active sensors on polar-orbiting satellites like the CloudSat CPR, which suffer from limited spatial and temporal coverage. In contrast, passive instruments, such as MSG SEVIRI, offer high-resolution observations across a broad domain but lack vertical detail. Since clouds are dynamic phenomena, an accurate depiction of their evolution — both horizontally and vertically — is essential for advancing a physical processes understanding (e.g., Seneviratne et al., 2012; Dubovik et al., 2021; Beucler et al., 2024).

Summary
To address its objectives, this thesis integrates 2D satellite datasets, creating a synergistic view that supports deeper insights into the representation and variability of convective clouds. ML-based predictions are used to identify convective cloud and core regions, which are linked over time into trajectories. These data reveal key features of cloud development, advancing our perspective on the relationship between 3D cloud structure and spatial organisation.

Section 3.2 deals with the first objective and introduces a simple ML architecture — a 2D Res-UNet — that maps geostationary MSG SEVIRI data to the vertically resolved observations of the CloudSat CPR. The model reconstructs vertical cross-sections and extrapolates them into 3D radar reflectivity images across nearly the entire MSG SEVIRI disk. Due to the limited swath and revisit times of CloudSat, only about 10 % of the predicted pixels can be validated. Despite this, the model effectively reconstructs vertical cloud structures and hydrometeor locations, achieving an average RMSE of 3.05 dBZ within a reflectivity range of -25 to 20 dBZ. This performance is on par with generative DL models (Leinonen et al., 2019) but offers spatial continuity in three dimensions. The Res-UNet outperforms the linear regression and random forest models in both accuracy and spatial coherence. By combining the high resolution of MSG SEVIRI with vertical information from CloudSat, the ML approach significantly expands the availability of vertically resolved cloud profiles. Although this study focuses on Africa, Europe, and the Atlantic Ocean, the method is transferable to other regions and satellite systems (e.g., NASA's GOES over the Americas) or to other applications like aerosol tracking. Harmonising datasets could lead to globally available - and thus highly valuable - resources for the broader scientific community. Nonetheless, questions remain regarding the physical consistency and interpretability of ML models (e.g., Beucler et al., 2021; Gentine et al., 2018). Overall, the results demonstrate the potential to extract insights from complex, large-scale atmospheric datasets even for ML models with a relatively simple architecture, such as the UNet (Reichstein et al., 2019).

Section 3.3 addresses the second objective which is to examine the relationship between 3D cloud structures, convective cores, and cloud life-cycle, using the radar reflectivity output from Sect. 3.2. An object-based detection algorithm identifies and tracks convective clouds and their cores over time. While the ML model covers a domain of 60° N–S by 60° E–W, analysis focuses on

tropical convection within a 30° N–S by 30° E–W region. Using six months of data (March–August 2019), the study reproduces known convective cloud characteristics — such as the differing diurnal cycles of land and ocean convection and the prevalence of short-lived, single-cell clouds (e.g., Chen and Houze, 1997; Takahashi et al., 2023; Jones et al., 2024). The findings suggest that a higher number of cores is linked to larger, deeper, and longer-lived cloud systems, possibly indicating feedback mechanisms like self-sustaining convection (e.g, Deng et al., 2016; Hartmann et al., 2018). Generated cloud trajectories offer simultaneous vertical and horizontal information over time, enabling new insights — particularly over data sparse remote oceanic regions. However, land-sea differences in this domain are less pronounced than expected. A more extended climatological analysis and validation of object-based methodologies inherently connected to a subjectivity of chosen thresholds would strengthen these results.

Section 3.4 sheds a light on the third objective and applies a percentile-based approach to quantify the spatial and temporal distribution of convective organisation across the tropics. Using the cloud tracks obtained from Sect. 3.3 (March–August 2019, 30° N–S × 30° E–W), the study aims to reveal "hotspot regions" of strong or weak convective organisation. Analysing the relationship between three established organisation indices and the previously extracted cloud and core properties emphasises how stronger organisation is typically associated with larger cloud areas, lower tops and core heights, and shorter lifetimes, while weaker organisation corresponds to smaller clouds with fewer cores. These results align with known traits of MCSs in the tropics (Stubenrauch et al., 2023). Moreover, the 3D dataset offers unique insights into both horizontal and vertical drivers of organisation. Seasonal variations show a shift in convective hotspots — from the Gulf of Guinea and the southern Atlantic in spring to West Africa and the equatorial rainforest in summer — possibly linked to the northward migration of the ITCZ (Stephens et al., 2024). However, overlaps between strong and weak organisation patterns may blur the statistics. The study concludes that long-term datasets and adaptive indices that incorporate 3D cloud structures are essential for a robust assessment of convective organisation.

To summarise, this thesis reveals that a further integration of ML models offers a novel perspective on cloud behavior and with this, may bring added value to the research community. The evaluation of predicted data and subsequent applications in this work highlight the possibility to derive deeper insights on convective cloud dynamics, which is of uttermost importance for an accurate assessment of their hazard potential. Despite remarkable potentials, the usage of ML comes along noteworthy limitations and uncertainties - in particular regarding the feasibility of state-of-the-art approaches. The robustness of the output is inherently connected to the quality of the input, which is reflected in this work, e.g., by an insufficient representation of ice clouds passed on from the CloudSat CPR. The model architecture and training design are crucial for deriving meaningful results (e.g, Ebert-Uphoff and Hilburn, 2020; Sarker, 2021). Although the integration of ML and DL in atmospheric and climate sciences is currently at an early stage, the approaches may help to address data gaps and improve cloud coverage compared to using observational data alone. By uncovering complex patterns and dynamics, ML-based analyses can potentially reduce uncertainty in weather forecasting and climate risk assessments (Rolnick et al., 2022). Hence, the findings demonstrate the value of even simple ML architectures in enhancing our understanding of convective processes and suggest promising avenues for further research.

Outlook

The ML framework developed in this thesis enables a contiguous prediction of 3D cloud fields from 2D satellite data, providing valuable insights into both the horizontal and vertical distribution of cloud properties. Despite some limitations — such as the under-representation of ice clouds, or the blurring introduced by the chosen loss function — the framework captures the evolution of cloud properties across a wide domain spanning the tropics, subtropics, and mid-latitudes.

Subsequent analyses include identifying distinct life-cycle stages and associated physical changes. While the thesis focuses on convective cloud development in the tropics, the data itself is neither limited to convective clouds or the chosen domain. Potential applications further include identifying different convective regimes or tracking of tropical cyclones (McGovern et al., 2023). Moreover, the data may provide suitable information to analyse the evolution of various cloud types and assessing rain rates in connection with 3D cloud properties. Such applications may enhance understanding of cloud formation, support climate modeling, and aid in mitigating severe weather impacts. The data can also serve as observational constraints for validating high-resolution atmospheric models.

Currently, the ML model uses geostationary satellite input images to reconstruct vertical cross-sections from CloudSat CPR and generate 3D cloud fields. It serves as a compelling example of how ML can integrate the strengths of multiple datasets while addressing their individual limitations. Although designed for MSG SEVIRI and CloudSat, the framework's flexible architecture can be adapted to other 2D multi-sensor datasets to produce synergistic outputs. For instance, Ceamanos et al. (2021) showcase how a combination of five geostationary satellites (the so-called GEO-Ring) may induce a simultaneous full globe coverage - together with the framework introduced in this thesis, global predictions of 3D cloud fields may become feasible.

Looking ahead, the robustness of the results could be enhanced using data from next-generation geostationary satellites with a higher native resolution, like EUMETSAT's Meteosat Third Generation (MTG) (Holmlund et al., 2021). The ML model may also be extended to predict vertical-profile datasets such as attenuated backscatter from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite (Winker et al., 2007). Further opportunities lie in integrating data from EarthCARE, a mission equipped with multiple active and passive instruments that succeeds CALIPSO and CloudSat. EarthCARE promises deeper insights into clouds, aerosols, precipitation, and radiative fluxes (Illingworth et al., 2015), including an improved retrieval of cloud-scale updrafts. A dataset combining EarthCARE and MTG, built using the approach presented here, may offer exciting avenues to enhance the accuracy and applicability of the framework. Such integration of ML methods can hold great promise for advancing scientific understanding and achieving an improved representation of Earth system processes (Beucler et al., 2023).

The findings illustrate that even relatively simple ML architectures can extract meaningful patterns from complex or sparse datasets, as shown in the analyses in Sects. 3.3 and 3.4. Still, advanced architectures could further improve cloud representation. For example, Girtsou et al. (2025) have built upon this work by applying self-supervised learning with Masked Autoencoders and Vision Transformer backbones to reconstruct 3D cloud structures, aiming towards a near real-time, global representation. Feeding the data from such ML models into operational systems could substantially enhance climate predictions, risk assessment, and decision-making capabilities in the future.

List of Acronyms

2D Two-dimensional

3D Three-dimensional

AI Artificial intelligence

AIFS Artificial Intelligence Forecasting System

ANN/NN Artificial neural network

CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations

CER Cloud effective radius

CMSAF Satellite Application Facility on Climate Monitoring

CNN Convolutional neural network

CPR Cloud profiling radar

CTH Cloud top height

CWP Cloud waterpath

DL Deep learning

ECMWF European Centre for Medium-Range Weather Forecasting

ESA European Space Agency

FCNN Fully connected neural network

GEO Geostationary earth orbit

GOES Geostationary Operational Environmental Satellite

GPT Generative pre-trained transformer

IPCC Intergovernmental Panel on Climate Change

IR Infrared

ITCZ Inter-Tropical Convergence Zone

LEO Low earth orbit

LiDAR Light Detection and Ranging

LLM Large language model

MCS Mesoscale convective system

ML Machine learning

MLP Multilayer perceptron

MODIS Moderate Resolution Imaging Spectroradiometer

MSG Meteosat Second Generation

MTG Meteosat Third Generation

NASA National Aeronautics and Space Administration

NIR Near infrared

NLP Natural language processing

NOAA National Oceanic and Atmospheric Administration

radar radio detection and ranging

RCE Radiative-convective equilibrium

SEVIRI Spinning Enhanced Visible and InfraRed Imager

VIS Visible light

References

Agrawal, S. et al. (2019). *Machine Learning for Precipitation Nowcasting from Radar Images*. arXiv: 10.48550/ARXIV.1912.12132.

Amato, U. et al. (2008). "Statistical cloud detection from SEVIRI multispectral images". In: *Remote Sens. Environ.* 112.3, pp. 750–766. DOI: 10.1016/j.rse.2007.06.004.

Andersen, H.-E., S. E. Reutebuch, and R. J. McGaughey (2006). "Active remote sensing". In: *Computer Applications in Sustainable Forest Management*. Springer Netherlands, pp. 43–66. ISBN: 978-1-4020-4305-5. DOI: 10.1007/978-1-4020-4387-1_3.

Ashley, W. S., A. M. Haberlie, and J. Strohm (2019). "A Climatology of Quasi-Linear Convective Systems and Their Hazards in the United States". In: *WAF* 34.6, pp. 1605–1631. DOI: 10.1175/WAF-D-19-0014.1.

Aubry, C. et al. (2024). "Lidar–radar synergistic method to retrieve ice, supercooled water and mixed-phase cloud properties". In: *Atmos. Meas. Tech.* 17.12, pp. 3863–3881. DOI: 10.5194/amt-17-3863-2024.

Aumann, H. H., A. Behrangi, and Y. Wang (2018). "Increased Frequency of Extreme Tropical Deep Convection: AIRS Observations and Climate Model Predictions". In: *Geophys. Res. Lett.* 45.24, pp. 13, 530–13, 537. DOI: 10.1029/2018GL079423.

Bankert, R. L. et al. (2009). "Comparison of GOES Cloud Classification Algorithms Employing Explicit and Implicit Physics". In: *J. Appl. Meteorol.* 48.7, pp. 1411–1421. DOI: 10.1175/2009JAMC2103.1.

Bao, J. and S. C. Sherwood (2019). "The Role of Convective Self-Aggregation in Extreme Instantaneous Versus Daily Precipitation". In: *J. Adv. Model. Earth Syst.* 11.1, pp. 19–33. DOI: 10.1029/2018MS001503.

Barker, H. W. et al. (2011). "A 3D cloud-construction algorithm for the EarthCARE satellite mission". In: *Q. J. R. Meteorol. Soc.* 137.657, pp. 1042–1058. DOI: 10.1002/qj.824.

Battaglia, A. et al. (2020). "Spaceborne Cloud and Precipitation Radars: Status, Challenges, and Ways Forward". In: *Rev. Geophys.* 58.3, e2019RG000686. DOI: 10.1029/2019RG000686.

Bellon, A., S. Lovejoy, and G. L. Austin (1980). "Combining Satellite and Radar Data for the Short-Range Forecasting of Precipitation". In: *Mon. Wea. Rev.* 108.10, pp. 1554–1566. DOI: 10.1175/1520-0493(1980)108⟨1554:CSARDF⟩2.0.CO;2.

Beucler, T. et al. (2021). "Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems". In: *Phys. Rev. Lett.* 126.9, p. 098302. DOI: 10.1103/PhysRevLett.126.098302.

Beucler, T. et al. (2023). "Machine Learning for Clouds and Climate". In: *Clouds and Their Climatic Impacts*. American Geophysical Union (AGU). Chap. 16, pp. 325–345. ISBN: 9781119700357. DOI: https://doi.org/10.1002/9781119700357.ch16.

Beucler, T. et al. (2024). "Climate-invariant machine learning". In: *Sci. Adv.* 10.6, eadj7250. DOI: 10.1126/sciadv.adj7250.

Bi, K. et al. (2023). "Accurate medium-range global weather forecasting with 3D neural networks". In: *Nature* 619.7970, pp. 533–538. DOI: 10.1038/s41586-023-06185-3.

Bläckberg, C. P. O. and M. S. Singh (2022). "Increased Large-Scale Convective Aggregation in CMIP5 Projections: Implications for Tropical Precipitation Extremes". In: *Geophys. Res. Lett.* 49.9, e2021GL097295. DOI: 10.1029/2021GL097295.

Bluestein, H. B., F. H. Carr, and S. J. Goodman (2022). "Atmospheric Observations of Weather and Climate". In: *Atmosphere-Ocean* 60.3-4, pp. 149–187. DOI: 10.1080/07055900.2022.2082369.

Bony, S. et al. (Mar. 2015). "Clouds, circulation and climate sensitivity". In: *Nature Geoscience* 8, pp. 261–268. DOI: 10.1038/ngeo2398.

Boucher, O. (1999). "Air traffic may increase cirrus cloudiness". In: *Nature* 397.6714, pp. 30–31. DOI: 10.1038/16169.

Boukabara, S.-A. et al. (2019). "Leveraging Modern Artificial Intelligence for Remote Sensing and NWP: Benefits and Challenges". In: *BAMS* 100.12, ES473–ES491. DOI: 10.1175/BAMS-D-18-0324.1.

Bouwer, L. M. (2019). "Observed and Projected Impacts from Extreme Weather Events: Implications for Loss and Damage". In: *Loss and Damage from Climate Change: Concepts, Methods and Policy Options*. Cham: Springer, Cham, pp. 63–82. DOI: 10.1007/978-3-319-72026-5_3.

Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov (2005). "An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST". In: *Journal of the Atmospheric Sciences* 62.12, pp. 4273–4292. DOI: 10.1175/JAS3614.1.

Brown, T. B. et al. (2020). *Language Models are Few-Shot Learners*. arXiv: 10.48550/ARXIV.2005.14165.

Camps-Valls, G. et al. (2014). "Advances in Hyperspectral Image Classification: Earth Monitoring with Statistical Learning Methods". In: *IEEE Signal Process. Mag.* 31.1, pp. 45–54. DOI: 10.1109/MSP.2013.2279179.

Ceamanos, X., B. Six, and J. Riedi (2021). "Quasi-Global Maps of Daily Aerosol Optical Depth From a Ring of Five Geostationary Meteorological Satellites Using AERUS-GEO". In: *Journal of Geophysical Research: Atmospheres* 126.20, e2021JD034906. DOI: https://doi.org/10.1029/2021JD034906.

Chan, S. C. et al. (2023). "Large-scale dynamics moderate impact-relevant changes to organised convective storms". In: *Commun. Earth Environ.* 4.1, pp. 1–10. DOI: 10.1038/s43247-022-00669-2.

Chen, M. et al. (2023). "ResU-Deep: Improving the Trigger Function of Deep Convection in Tropical Regions With Deep Learning". In: *J. Adv. Model. Earth Syst.* 15.11, e2022MS003521. DOI: 10.1029/2022MS003521.

Chen, P.-J. et al. (2021). "Convective Cloud Regimes From a Classification of Object-Based CloudSat Observations Over Asian-Australian Monsoon Areas". In: *Geophys. Res. Lett.* 48.10, e2021GL092733. DOI: 10.1029/2021GL092733.

Chen, S. S. and R. A. Houze (1997). "Diurnal variation and life-cycle of deep convective systems over the tropical pacific warm pool". In: *Q. J. R. Meteorol. Soc.* 123.538, pp. 357–388. DOI: 10.1002/qj.49712353806.

Clausius, R. (1850). "Ueber die bewegende Kraft der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen". In: *Annalen der Physik* 155.3, pp. 368–397. DOI: 10.1002/andp.18501550306.

Delanoë, J. et al. (2013). "Comparison of Airborne In Situ, Airborne Radar–Lidar, and Spaceborne Radar–Lidar Retrievals of Polar Ice Cloud Properties Sampled during the POLARCAT Campaign". In: *Journal of Atmospheric and Oceanic Technology* 30.1, pp. 57–73. DOI: 10.1175/JTECH-D-11-00200.1.

Deng, J. et al. (2009). "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

Deng, L. (2012). "The MNIST Database of Handwritten Digit Images for Machine Learning Research". In: *IEEE Signal Process. Mag.* 29.6, pp. 141–142. DOI: 10.1109/MSP.2012.2211477.

Deng, M., G. G. Mace, and Z. Wang (2016). "Anvil Productivities of Tropical Deep Convective Clusters and Their Regional Differences". In: *Journal of the Atmospheric Sciences* 73.9, pp. 3467–3487. DOI: 10.1175/JAS-D-15-0239.1.

Doswell, C. (1985). *The Operational Meteorology of Convective Weather: Storm scale analysis.* NOAA technical memorandum ERL ESG Bd. 2. U.S. Department of Commerce, NOAA, Environmental Research Laboratories.

Doswell III, C. (2001). "Severe Convective Storms—An Overview". In: *Meteorological Monographs (METEOR)*. American Meteorological Society, pp. 1–26. DOI: 10.1007/978-1-935704-06-5_1.

Dougherty, E. and K. L. Rasmussen (2020). "Changes in Future Flash Flood–Producing Storms in the United States". In: *J. Hydrometeorol.* 21.10, pp. 2221–2236. DOI: 10.1175/JHM-D-20-0014.1.

Dubovik, O. et al. (2021). "Grand Challenges in Satellite Remote Sensing". In: *Front. Remote Sens.* 2. DOI: 10.3389/frsen.2021.619818.

Ebert-Uphoff, I. and K. Hilburn (2020). "Evaluation, Tuning, and Interpretation of Neural Networks for Working with Images in Meteorological Applications". In: *BAMS* 101.12, E2149–E2170. DOI: 10.1175/BAMS-D-20-0097.1.

Emanuel, K. A., J. David Neelin, and C. S. Bretherton (1994). "On large-scale circulations in convecting atmospheres". In: *Q. J. R. Meteorol. Soc.* 120.519, pp. 1111–1143. DOI: 10.1002/qj.49712051902.

European Environment Agency, E. (2024). *The First European Climate Risk Assessment.* DOI: 10.2800/204249.

Eyring, V. et al. (2024). "Pushing the frontiers in climate modelling and analysis with machine learning". In: *Nat. Clim. Chang.* 14.9, pp. 916–928. DOI: 10.1038/s41558-024-02095-y.

Fan, J. et al. (2013). "Microphysical effects determine macrophysical response for aerosol impacts on deep convective clouds". In: *PNAS* 110.48, E4581–E4590. DOI: 10.1073/pnas.1316830110.

Feng, D. et al. (2021). "Deep Multi-modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges". In: *IEEE Transactions on Intelligent Transportation Systems* 22.3, pp. 1341–1360. DOI: 10.1109/TITS.2020.2972974.

Feng, Z. et al. (2016). "More frequent intense and long-lived storms dominate the springtime trend in central US rainfall". In: *Nat. Commun.* 7.1, p. 13429. DOI: 10.1038/ncomms13429.

Fu, L.-L. et al. (2019). "50 Years of Satellite Remote Sensing of the Ocean". In: *Meteorological Monographs* 59.1, pp. 5.1–5.46. DOI: 10.1175/AMSMONOGRAPHS-D-18-0010.1.

Futyan, J. M. and A. D. D. Genio (2007). "Deep Convective System Evolution over Africa and the Tropical Atlantic". In: *Journal of Climate* 20.20, pp. 5041–5060. DOI: 10.1175/JCLI4297.1.

Gao, B.-C. et al. (1998). "Cloud Detection over the Arctic Region Using Airborne Imaging Spectrometer Data during the Daytime". In: *Journal of Applied Meteorology* 37.11, pp. 1421–1429. DOI: 10.1175/1520-0450(1998)037⟨1421:CDOTAR⟩2.0.CO;2.

Gentine, P. et al. (2018). "Could Machine Learning Break the Convection Parameterization Deadlock?" In: *Geophys. Res. Lett.* 45.11, pp. 5742–5751. DOI: 10.1029/2018GL078202.

Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media, Sebastopol, CA. ISBN: 978-1-4919-6229-9.

Girtsou, S. et al. (2025). *3D Cloud reconstruction through geospatially-aware Masked Autoencoders*. arXiv: 10.32388/57QYAW.

Gómez-Chova, L. et al. (2015). "Multimodal Classification of Remote Sensing Images: A Review and Future Directions". In: *Proceedings of the IEEE* 103.9, pp. 1560–1584. DOI: 10.1109/JPROC.2015.2449668.

Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep Learning*. MIT Press, Cambridge, MA. ISBN: 0-262-03561-8.

Goodfellow, I. et al. (2020). "Generative adversarial networks". In: *Commun. ACM* 63.11, pp. 139–144. DOI: 10.1145/3422622.

Gorooh, V. A. et al. (2023). "Integrating LEO and GEO Observations: Toward Optimal Summertime Satellite Precipitation Retrieval". In: *J. Hydrometeorol.* 24.11, pp. 1939–1954. DOI: 10.1175/JHM-D-23-0006.1.

Grabowski, W. W. and J. C. Petch (2009). "Deep Convective Clouds". In: *Clouds in the Perturbed Climate System: Their Relationship to Energy Balance, Atmospheric Dynamics, and Precipitation*. Cambridge, London: The MIT Press. DOI: 10.7551/mitpress/8300.003.0011.

Haerter, J. O. et al. (2019). "Circling in on Convective Organization". In: *Geophysical Research Letters* 46.12, pp. 7024–7034. DOI: https://doi.org/10.1029/2019GL082092.

Hanin, B. (2019). "Universal Function Approximation by Deep Neural Nets with Bounded Width and ReLU Activations". In: *Mathematics* 7.10, p. 992. DOI: 10.3390/math7100992.

Hanin, B. and D. Rolnick (2019). *Complexity of Linear Regions in Deep Networks*. arXiv: 10.48550/arXiv.1901.09021.

Hartmann, D., H. Hendon, and J. Houze (1984). "Some Implications of the Mesoscale Circulations in Tropical Cloud Clusters for Large-Scale Dynamics and Climate". In: *American Meteorological Society* 41, pp. 113–121. DOI: 10.1175/1520-0469(1984)041⟨0113:SIOTMC⟩2.0.CO;2.

Hartmann, D. L., P. N. Blossey, and B. D. Dygert (2019). "Convection and Climate: What Have We Learned from Simple Models and Simplified Settings?" In: *Curr. Clim. Change Rep.* 5.3, pp. 196–206. DOI: 10.1007/s40641-019-00136-9.

Hartmann, D. L. et al. (2018). "The Life Cycle and Net Radiative Effect of Tropical Anvil Clouds". In: *Journal of Advances in Modeling Earth Systems* 10.12, pp. 3012–3029. DOI: https://doi.org/10.1029/2018MS001484.

Haynes, J. M. et al. (2009). "Rainfall retrieval over the ocean with spaceborne W-band radar". In: *Journal of Geophysical Research: Atmospheres* 114.D8. DOI: 10.1029/2008JD009973.

Held, I. M., R. S. Hemler, and V. Ramaswamy (1993). "Radiative-Convective Equilibrium with Explicit Two-Dimensional Moist Convection". In: *Journal of the Atmospheric Sciences* 50.23, pp. 3909–3927. DOI: 10.1175/1520-0469(1993)050⟨3909:RCEWET⟩2.0.CO;2.

Henderson, D. S. et al. (2013). "A Multisensor Perspective on the Radiative Impacts of Clouds and Aerosols". In: *Journal of Applied Meteorology and Climatology* 52.4, pp. 853–871. DOI: 10.1175/JAMC-D-12-025.1.

Hilburn, K. A., I. Ebert-Uphoff, and S. D. Miller (2020). "Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations". In: *J. Appl. Meteorol.* 60.1, pp. 3–21. DOI: 10.1175/JAMC-D-20-0084.1.

Hohenegger, C. et al. (2020). "Climate Statistics in Global Simulations of the Atmosphere, from 80 to 2.5 km Grid Spacing". In: *J. Meteorol. Soc. Jpn. Ser. II* 98.1, pp. 73–91. DOI: 10.2151/jmsj.2020-005.

Holloway, C. E. (2017). "Convective aggregation in realistic convective-scale simulations". In: *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1450–1472. DOI: https://doi.org/10.1002/2017MS000980.

Holmlund, K. et al. (2021). "Meteosat Third Generation (MTG): Continuation and Innovation of Observations from Geostationary Orbit". In: *BAMS* 102.5, E990–E1015. DOI: 10.1175/BAMS-D-19-0304.1.

Holton, J. R. et al. (1995). "Stratosphere-troposphere exchange". In: *Rev. Geophys.* 33.4, pp. 403–439. DOI: 10.1029/95RG02097.

Houze, R. A. (2018). "100 Years of Research on Mesoscale Convective Systems". In: *Meteorological Monographs* 59.1, pp. 17.1–17.54. DOI: 10.1175/AMSMONOGRAPHS-D-18-0001.1.

Houze Jr., R. A. (1989). "Observed structure of mesoscale convective systems and implications for large-scale heating". In: *Quarterly Journal of the Royal Meteorological Society* 115.487, pp. 425–461. DOI: https://doi.org/10.1002/qj.49711548702.

— (2004). "Mesoscale convective systems". In: *Rev. Geophys.* 42.4. DOI: 10.1029/2004RG000150.

Illingworth, A. J. et al. (2015). "The EarthCARE Satellite: The Next Step Forward in Global Measurements of Clouds, Aerosols, Precipitation, and Radiation". In: *BAMS* 96.8, pp. 1311–1332. DOI: 10.1175/BAMS-D-12-00227.1.

Irrgang, C. et al. (2021). "Towards neural Earth system modelling by integrating artificial intelligence in Earth system science". In: *Nat. Mach. Intell.* 3.8, pp. 667–674. DOI: 10.1038/s42256-021-00374-3.

Jensen, E. and L. Pfister (2004). "Transport and freeze-drying in the tropical tropopause layer". In: *J. Geophys. Res. Atmos.* 109.D2. DOI: 10.1029/2003JD004022.

Jones, T. A. et al. (2015). "Simultaneous Radar and Satellite Data Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011". In: *Mon. Wea. Rev.* 143.1, pp. 165–194. DOI: 10.1175/MWR-D-14-00180.1.

Jones, W., M. Stengel, and P. Stier (2024). "A Lagrangian perspective on the lifecycle and cloud radiative effect of deep convective clouds over Africa". In: *Atmos. Chem. Phys.* 24, pp. 5165–5180. DOI: 10.5194/acp-24-5165-2024.

Keisler, R. (2022). *Forecasting Global Weather with Graph Neural Networks.* arXiv: 10.48550/arXiv.2202.07575.

Kidd, C., V. Levizzani, and P. Bauer (2009). "A review of satellite meteorology and climatology at the start of the twenty-first century". In: *Progress in Physical Geography: Earth and Environment* 33.4, pp. 474–489. DOI: 10.1177/0309133309346647.

King, M. D. et al. (2013). "Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites". In: *IEEE Trans. Geosci. Remote Sensing* 51.7, pp. 3826–3852. DOI: 10.1109/TGRS.2012.2227333.

Krizhevsky, A. (2012). *Learning Multiple Layers of Features from Tiny Images*. URL: https://api.semanticscholar.org/CorpusID:18268744.

Kunz, M. (2007). "The skill of convective parameters and indices to predict isolated and severe thunderstorms". In: *Nat. Hazards Earth Syst. Sci.* 7.2, pp. 327–342. DOI: 10.5194/nhess-7-327-2007.

L'Ecuyer, T. S. and J. H. Jiang (2010). "Touring the atmosphere aboard the A-Train". In: *Physics Today* 63.7, pp. 36–41. DOI: 10.1063/1.3463626.

Lafore, J.-P. and M. W. Moncrieff (1989). "A Numerical Investigation of the Organization and Interaction of the Convective and Stratiform Regions of Tropical Squall Lines". In: *Journal of the Atmospheric Sciences* 46.4, pp. 521–544. DOI: 10.1175/1520-0469(1989)046⟨0521:ANIOTO⟩2.0.CO;2.

Lagerquist, R. et al. (2021). "Using Deep Learning to Nowcast the Spatial Coverage of Convection from Himawari-8 Satellite Data". In: *Mon. Weather Rev.* 149.12, pp. 3897–3921. DOI: 10.1175/MWR-D-21-0096.1.

Lam, R. et al. (2022). *GraphCast: Learning skillful medium-range global weather forecasting*. arXiv: 10.48550/ARXIV.2212.12794.

Leary, C. A. and R. A. Houze (1979). "Melting and Evaporation of Hydrometeors in Precipitation from the Anvil Clouds of Deep Tropical Convection". In: *Journal of the Atmospheric Sciences* 36.4, pp. 669–679. DOI: 10.1175/1520-0469(1979)036⟨0669:MAEOHI⟩2.0.CO;2.

LeCun, Y. et al. (1989). "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Comput.* 1.4, pp. 541–551. DOI: 10.1162/neco.1989.1.4.541.

LeCun, Y., Y. Bengio, and G. Hinton (2015). "Deep learning". In: *Nature* 521.7553, pp. 436–444. DOI: 10.1038/nature14539.

Lee, Y., C. D. Kummerow, and I. Ebert-Uphoff (2021). "Applying machine learning methods to detect convection using Geostationary Operational Environmental Satellite-16 (GOES-16) advanced baseline imager (ABI) data". In: *Atmos. Meas. Tech.* 14.4, pp. 2699–2716. DOI: 10.5194/amt-14-2699-2021.

Leinonen, J., A. Guillaume, and T. Yuan (2019). "Reconstruction of Cloud Vertical Structure With a Generative Adversarial Network". In: *Geophys. Res. Lett.* 46, pp. 7035–7044. DOI: 10.1029/2019GL082532.

Liou, K. N. (1992). *Radiation and Cloud Processes in the Atmosphere: Theory, Observation, and Modeling*. Oxford University Press, New York, NY. DOI: 10.1093/oso/9780195049107.001.0001.

Litjens, G. et al. (2017). "A survey on deep learning in medical image analysis". In: *Med. Image Anal.* 42, pp. 60–88. DOI: 10.1016/j.media.2017.07.005.

Liu, Y. et al. (2016). *Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets*. arXiv: 10.48550/ARXIV.1605.01156.

Lu, Z. et al. (2017). *The Expressive Power of Neural Networks: A View from the Width*. arXiv: 10.48550/arXiv.1709.02540.

Machado, L. A. T. et al. (1998). "Life Cycle Variations of Mesoscale Convective Systems over the Americas". In: *Monthly Weather Review* 126.6, pp. 1630–1654. DOI: 10.1175/1520-0493(1998)126⟨1630:LCVOMC⟩2.0.CO;2.

Mapes, B. and R. Neale (2011). "Parameterizing Convective Organization to Escape the Entrainment Dilemma". In: *J. Adv. Model. Earth Syst.* 3.2. DOI: 10.1029/2011MS000042.

McGovern, A. et al. (2023). "A Review of Machine Learning for Convective Weather". In: *Artificial Intelligence for the Earth Systems* 2.3, e220077. DOI: 10.1175/AIES-D-22-0077.1.

Mehrabi, N. et al. (2022). "A Survey on Bias and Fairness in Machine Learning". In: *ACM Comput. Surv.* 54.6, pp. 1–35. DOI: 10.1145/3457607.

Menzel, W. P. (2001). "Cloud Tracking with Satellite Imagery: From the Pioneering Work of Ted Fujita to the Present". In: *BAMS* 82.1, pp. 33–48. DOI: 10.1175/1520-0477(2001)082⟨0033:CTWSIF⟩2.3.CO;2.

Montavon, G., W. Samek, and K.-R. Müller (2018). "Methods for interpreting and understanding deep neural networks". In: *Digit. Signal Process.* 73, pp. 1–15. DOI: 10.1016/j.dsp.2017.10.011.

Muller, C. and S. Bony (2015). "What favors convective aggregation and why?" In: *Geophys. Res. Lett.* 42.13, pp. 5626–5634. DOI: 10.1002/2015GL064260.

Muller, C. and I. Held (2012). "Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations". In: *Journal of Atmospheric Sciences* 69, pp. 2551–2565. DOI: 10.1175/JAS-D-11-0257.1.

Muller, C. et al. (2022). "Spontaneous Aggregation of Convective Storms". In: *Annual Review of Fluid Mechanics* 54, pp. 133–157. DOI: 10.1146/annurev-fluid-022421-011319.

Nakajima, T. and M. D. King (1990). "Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory". In: *J. Atmos. Sci.* 47.15, pp. 1878–1893. DOI: 10.1175/1520-0469(1990)047⟨1878:DOTOTA⟩2.0.CO;2.

Nesbitt, S. W., R. Cifelli, and S. A. Rutledge (2006). "Storm Morphology and Rainfall Characteristics of TRMM Precipitation Features". In: *Mon. Wea. Rev.* 134.10, pp. 2702–2721. DOI: 10.1175/MWR3200.1.

Nguyen, A., J. Yosinski, and J. Clune (2014). *Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images*. arXiv: 10.48550/ARXIV.1412.1897.

Okamoto, H. and K. Sato (2018). "Cloud Remote Sensing by Active Sensors: New Perspectives from CloudSat, CALIPSO and EarthCARE". In: *Remote Sensing of Clouds and Precipitation*. Cham: Springer, Cham, pp. 195–214. DOI: 10.1007/978-3-319-72583-3_8.

Pendergrass, A. G. (2020). "Changing Degree of Convective Organization as a Mechanism for Dynamic Changes in Extreme Precipitation". In: *Curr. Clim. Change Rep.* 6.2, pp. 47–54. DOI: 10.1007/s40641-020-00157-9.

Pettorelli, N. et al. (2018). "SATELLITE REMOTE SENSING FOR CONSERVATION". In: *Conservation Technology Series* 4. DOI: 10.13140/RG.2.2.25962.41926.

Prein, A. F. et al. (2017). "Increased rainfall volume from future convective storms in the US". In: *Nature Clim. Change* 7.12, pp. 880–884. DOI: 10.1038/s41558-017-0007-7.

Pscheidt, I. et al. (2019). "How organized is deep convection over Germany?" In: *Q. J. R. Meteorol. Soc.* 145.723. DOI: 10.1002/qj.3552.

Rasp, S., M. S. Pritchard, and P. Gentine (2018). "Deep learning to represent subgrid processes in climate models". In: *Proceedings of the National Academy of Sciences* 115.39, pp. 9684–9689. DOI: 10.1073/pnas.1810286115.

Rasp, S. et al. (2020). "Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection". In: *BAMS* 101.11, E1980–E1995. DOI: 10.1175/BAMS-D-19-0324.1.

Rawat, W. and Z. Wang (2017). "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review". In: *Neural Comput.* 29.9, pp. 2352–2449. DOI: 10.1162/neco_a_00990.

Ray, P. P. (2023). "ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope". In: *Internet Things Cyber-Phys. Syst.* 3, pp. 121–154. DOI: 10.1016/j.iotcps.2023.04.003.

Reichstein, M. et al. (2019). "Deep learning and process understanding for data-driven Earth system science". In: *Nature* 566.7743, pp. 195–204. DOI: 10.1038/s41586-019-0912-1.

Rolnick, D. et al. (2022). "Tackling Climate Change with Machine Learning". In: *ACM Comput. Surv.* 55.2, 42:1–42:96. DOI: 10.1145/3485128.

Rombach, R. et al. (2022). *High-Resolution Image Synthesis with Latent Diffusion Models*. arXiv: 10.48550/arXiv.2112.10752.

Ronneberger, O., P. Fischer, and T. Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Vol. 9351. Cham: Springer, Cham, pp. 234–241. ISBN: 978-3-319-24574-4.

Salomonson, V. V., W. Barnes, and E. J. Masuoka (2006). "Introduction to MODIS and an Overview of Associated Activities". In: *Earth Science Satellite Remote Sensing: Vol. 1: Science and Instruments*. Ed. by J. J. Qu et al. Springer, Berlin, Heidelberg, pp. 12–32. DOI: 10.1007/978-3-540-37293-6_2.

Samuel, A. L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". In: *IBM J. Res. Dev.* 3.3, pp. 210–229. DOI: 10.1147/rd.33.0210.

Sarker, I. H. (2021). "Machine Learning: Algorithms, Real-World Applications and Research Directions". In: *SN COMPUT. SCI.* 2.3, p. 160. DOI: 10.1007/s42979-021-00592-x.

Sassen, K. and Z. Wang (2008). "Classifying clouds around the globe with the CloudSat radar: 1-year of results". In: *Geophys. Res. Lett.* 35.4. DOI: 10.1029/2007GL032591.

Schmetz, J. et al. (2002). "An introduction to Metetsat second generation (MSG)". In: *BAMS* 83.7, pp. 977–992. DOI: 10.1175/1520-0477(2002)083⟨0977:AITMSG⟩2.3.CO;2.

Schumann, U. et al. (2012). "A Parametric Radiative Forcing Model for Contrail Cirrus". In: *J. Appl. Meteorol.* 51.7, pp. 1391–1406. DOI: 10.1175/JAMC-D-11-0242.1.

Seidel, D. J. et al. (2008). "Widening of the tropical belt in a changing climate". In: *Nature Geosci.* 1.1, pp. 21–24. DOI: 10.1038/ngeo.2007.38.

Seneviratne, S. et al. (2012). "Changes in climate extremes and their impacts on the natural physical environment". In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 109–230.

Sherwood, S. C. et al. (2020). "An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence". In: *Reviews of Geophysics* 58.4, e2019RG000678. DOI: https://doi.org/10.1029/2019RG000678.

Sherwood, S. C., S. Bony, and J.-L. Dufresne (2014). "Spread in model climate sensitivity traced to atmospheric convective mixing". In: *Nature* 505.7481, pp. 37–42. DOI: 10.1038/nature12829.

Sherwood, S. C. and A. E. Dessler (2000). "On the control of stratospheric humidity". In: *Geophys. Res. Lett.* 27.16, pp. 2513–2516. DOI: 10.1029/2000GL011438.

Sieglaff, J. M. et al. (2011). "Nowcasting Convective Storm Initiation Using Satellite-Based Box-Averaged Cloud-Top Cooling and Cloud-Type Trends". In: *J. Appl. Meteorol.* 50.1, pp. 110–126. DOI: 10.1175/2010JAMC2496.1.

Singh, A. S. et al. (2017). "The perceived psychological distance of climate change impacts and its influence on support for adaptation policy". In: *Environmental Science & Policy* 73, pp. 93–99. DOI: 10.1016/j.envsci.2017.04.011.

Sønderby, C. K. et al. (2020). *MetNet: A Neural Weather Model for Precipitation Forecasting.* arXiv: 10.48550/ARXIV.2003.12140.

Spekkers, M. et al. (2017). "A comparative survey of the impacts of extreme rainfall in two international case studies". In: *Nat. Hazards Earth Syst. Sci.* 17.8, pp. 1337–1355. DOI: 10.5194/nhess-17-1337-2017.

Srivastava, N. et al. (2014). "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *J. Mach. Learn. Res.* 15.1, pp. 1929–1958. DOI: 10.5555/2627435.2670313.

Stephens, G. L. and C. D. Kummerow (2007). "The Remote Sensing of Clouds and Precipitation from Space: A Review". In: *Journal of the Atmospheric Sciences* 64.11, pp. 3742–3765. DOI: 10.1175/2006JAS2375.1.

Stephens, G. L. et al. (2008). "CloudSat mission: Performance and early science after the first year of operation". In: *J. Geophys. Res. Atmos.* 113.D8. DOI: 10.1029/2008JD009982.

Stephens, G. L. et al. (2024). "Tropical Deep Convection, Cloud Feedbacks and Climate Sensitivity". In: *Surv. Geophys.* 45.6, pp. 1903–1931. DOI: 10.1007/s10712-024-09831-1.

Stevens, B. (2005). "ATMOSPHERIC MOIST CONVECTION". In: *Annual Review of Earth and Planetary Sciences* 33, pp. 605–643. DOI: https://doi.org/10.1146/annurev.earth.33.092203.122658.

Stubenrauch, C. J. et al. (2010). "A 6-year global cloud climatology from the Atmospheric InfraRed Sounder AIRS and a statistical analysis in synergy with CALIPSO and CloudSat". In: *Atmos. Chem. Phys.* 10.15, pp. 7197–7214. DOI: 10.5194/acp-10-7197-2010.

Stubenrauch, C. J., G. Mandorli, and E. Lemaitre (2023). "Convective organization and 3D structure of tropical cloud systems deduced from synergistic A-Train observations and machine learning". In: *Atmos. Chem. Phys.* 23.10, pp. 5867–5884. DOI: 10.5194/acp-23-5867-2023.

Sylolypavan, A. et al. (2023). "The impact of inconsistent human annotations on AI driven clinical decision making". In: *npj Digit. Med.* 6.1, pp. 1–13. DOI: 10.1038/s41746-023-00773-3.

Takahashi, H. et al. (2023). "Revisiting the Land-Ocean Contrasts in Deep Convective Cloud Intensity Using Global Satellite Observations". In: *Geophys. Res. Lett.* 50.5, e2022GL102089. DOI: 10.1029/2022GL102089.

Tan, J. et al. (2015). "Increases in tropical rainfall driven by changes in frequency of organized deep convection". In: *Nature* 519.7544, pp. 451–454. DOI: 10.1038/nature14339.

Thies, B. and J. Bendix (2011). "Satellite based remote sensing of weather and climate: recent achievements and future perspectives". In: *Meteorological Applications* 18.3, pp. 262–295. DOI: 10.1002/met.288.

Tippett, M. K. et al. (2015). "Climate and Hazardous Convective Weather". In: *Curr. Clim. Change Rep.* 1.2, pp. 60–73. DOI: 10.1007/s40641-015-0006-6.

Tompkins, A. M. (2001). "Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Water Vapor". In: *Journal of the Atmospheric Sciences* 58.6, pp. 529–545. DOI: 10.1175/1520-0469(2001)058⟨0529:OOTCIL⟩2.0.CO;2.

Tompkins, A. M. and A. G. Semie (2017). "Organization of tropical convection in low vertical wind shears: Role of updraft entrainment". In: *J. Adv. Model. Earth Syst.* 9.2, pp. 1046–1068. DOI: 10.1002/2016MS000802.

Várnai, T. and A. Marshak (2002). "Observations of Three-Dimensional Radiative Effects that Influence MODIS Cloud Optical Thickness Retrievals". In: *Journal of the Atmospheric Sciences* 59.9, pp. 1607–1618. DOI: 10.1175/1520-0469(2002)059⟨1607:OOTDRE⟩2.0.CO;2.

Vaswani, A. et al. (2017). "Attention is All you Need". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17, pp. 6000–6010. DOI: 10.5555/3295222.3295349.

Westra, S. et al. (2014). "Future changes to the intensity and frequency of short-duration extreme rainfall". In: *Rev. Geophys.* 52.3, pp. 522–555. DOI: 10.1002/2014RG000464.

Wielicki, B. A. et al. (1996). "Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment". In: *BAMS* 77.5, pp. 853–868. DOI: 10.1175/1520-0477(1996)077⟨0853:CATERE⟩2.0.CO;2.

Wilcox, E. M. and V. Ramanathan (2001). "Scale Dependence of the Thermodynamic Forcing of Tropical Monsoon Clouds: Results from TRMM Observations". In: *J. Clim.* 14.7, pp. 1511–1524. DOI: 10.1175/1520-0442(2001)014⟨1511:SDOTTF⟩2.0.CO;2.

Wilcox, E. M., T. Yuan, and H. Song (2023). "Deep convective cloud system size and structure across the global tropics and subtropics". In: *Atmos. Meas. Tech.* 16.21, pp. 5387–5401. DOI: 10.5194/amt-16-5387-2023.

Wilson, J. W. and C. K. Mueller (1993). "Nowcasts of Thunderstorm Initiation and Evolution". In: *WAF* 8.1, pp. 113–131. DOI: 10.1175/1520-0434(1993)008⟨0113:NOTIAE⟩2.0.CO;2.

Windmiller, J. M. and C. Hohenegger (2019). "Convection On the Edge". In: *J. Adv. Model Earth Syst.* 11.12, pp. 3959–3972. DOI: 10.1029/2019MS001820.

Wing, A. A. (2019). "Self-Aggregation of Deep Convection and its Implications for Climate". In: *Curr. Clim. Change Rep.* 5.1, pp. 1–11. DOI: 10.1007/s40641-019-00120-3.

Wing, A. A. et al. (2017). "Convective Self-Aggregation in Numerical Simulations: A Review". In: *Surv. Geophys.* 38.6, pp. 1173–1197. DOI: 10.1007/s10712-017-9408-4.

— (2018). "Convective Self-Aggregation in Numerical Simulations: A Review". In: *Shallow Clouds, Water Vapor, Circulation, and Climate Sensitivity*. Springer, Cham, pp. 1–25. DOI: 10.1007/978-3-319-77273-8_1.

Winker, D. et al. (2017). "Observational Constraints on Cloud Feedbacks: The Role of Active Satellite Sensors". In: *Surv. Geophys.* 38.6, pp. 1483–1508. DOI: 10.1007/s10712-017-9452-0.

Winker, D. M., W. H. Hunt, and M. J. McGill (2007). "Initial performance assessment of CALIOP". In: *Geophysical Research Letters* 34.19, p. L19803. DOI: https://doi.org/10.1029/2007GL030135.

Wodzicki, K. R. and A. D. Rapp (2022). "More Intense, Organized Deep Convection With Shrinking Tropical Ascent Regions". In: *Geophys. Res. Lett.* 49.15, e2022GL098615. DOI: 10.1029/2022GL098615.

Wong, S. and J. Teixeira (2016). "Extreme Convection and Tropical Climate Variability: Scaling of Cold Brightness Temperatures to Sea Surface Temperature". In: *J. Clim.* 29.10, pp. 3893–3905. DOI: 10.1175/JCLI-D-15-0214.1.

Wu, P., X. Dong, and B. Xi (2020). "A Climatology of Marine Boundary Layer Cloud and Drizzle Properties Derived from Ground-Based Observations over the Azores". In: *J. Clim.* 33.23, pp. 10133–10148. DOI: 10.1175/JCLI-D-20-0272.1.

Wu, X., T. Yuan, et al. (2020). "Geographical distribution of extreme deep and intense convective storms on Earth". In: *Atmos. Meas. Tech.* 235, p. 104789. DOI: 10.1016/j.atmosres.2019.104789.

Xie, E. et al. (2021). *SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers*. arXiv: 10.48550/ARXIV.2105.15203.

Young, D. et al. (1998). "Comparison of in situ and satellite-derived cloud properties during SUCCESS". In: *Geophysical Research Letters* 25, pp. 1125–1128. DOI: 10.1029/98GL00116.

Zhang, P. et al. (2019). "General Comparison of FY-4A/AGRI With Other GEO/LEO Instruments and Its Potential and Challenges in Non-meteorological Applications". In: *Front. Earth Sci.* 6. DOI: 10.3389/feart.2018.00224.

Zhao, S. et al. (2023). "The role of satellite remote sensing in mitigating and adapting to global climate change". In: *Science of The Total Environment* 904, p. 166820. DOI: 10.1016/j.scitotenv.2023.166820.

Zipser, E. J. and C. Liu (2021). "Extreme Convection vs. Extreme Rainfall: a Global View". In: *Curr. Clim. Change Rep.* 7.4, pp. 121–130. DOI: 10.1007/s40641-021-00176-0.