UNIVERSITA DEGLI STUDI DI TRIESTE

DOCTORAL THESIS

Understanding Convective Organization and Seeking it Through Observations

---

Author:
Alejandro Casallas Garcia

Supervisor:
Prof. Adrian Mike Tompkins

Ph.D. Coordinator:
Prof. Stefano Maset

Thesis Reviewers:
Prof. Caroline Muller
Prof. Daniel Hernández-Deckers
Prof. Paolina Cerlini

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

Earth System Physics, The Abdus Salam International Centre For Theoretical Physics ICTP
Department of Mathematics and Geosciences
Cycle XXXVI

April 5, 2024

"Even with infinite computing power insufficient understanding limits its application; With infinite computing power we have to understand things that we could in principle compute."

Rudolf Clausius

Abstract

Understanding Convective Organization and Seeking it Through Observations

by Alejandro Casallas Garcia

Idealized simulations of radiative-convective equilibrium (RCE) using cloud-resolving models provide a controlled environment to investigate the influence of diabatic processes on convective clustering. This clustering results in significant drying of the free troposphere and heightened spatial variability in humidity, contributing to a larger increase in outgoing longwave radiation, thus it has the potential of impacting climate sensitivity. However, it's crucial to note that convective clustering in idealized models is influenced by the specific physical parameterizations employed. Furthermore, limited studies have focused on unraveling the mechanisms underlying clustering in observations. Addressing these gaps forms the foundation of this research.

This study pursues two main objectives: firstly, to unravel the sensitivity of convective self-aggregation (SA) to diverse physical parameterizations, and secondly, to comprehend the characteristics of convective organization in observations, shedding light on associated physical mechanisms. Studying organization is of profound significance due to its impact on the radiation budget, climate feedbacks, hydrological cycle, and extreme precipitation patterns, impacting society and climate.

Addressing the first objective, radiative-convective equilibrium simulations explore the effects of 24 parameterization combinations on aggregated and random convective patterns. Key to our analysis is understanding the role of maximum free convection distance (d_{clr}), found crucial for SA, and its modulation by parameterizations. SA predominantly emerges in scenarios with limited convective cores and extensive d_{clr} values (since they are anticorrelated), influenced by sub-grid scale mixing, planetary boundary layer (PBL), and microphysics. Horizontal mixing primarily influences SA by determining the size of convective cores, which is tight to their number and spacing. On the other hand, the influence of microphysics primarily stems from rain evaporation and its subsequent effects on Cold Pools (CPs). Surprisingly, perturbations to ice cloud microphysics had a notably limited effect. Non-local PBL schemes promote SA by enhancing low-level cloud cover, driven by intensified vertical transport within convective cores, and strong entrainment from the boundary layer top, due to stronger moisture gradients. The low cloud radiative forcing drives a circulation that shortens the transport of moist static energy, disallowing moisture to get to the driest regions, favoring SA. Moreover, the heightened low-level cloud cover induces more significant cooling, requiring stronger subsidence to maintain radiative equilibrium. This enhanced subsidence decreases convective core counts, increasing d_{clr}, and favoring SA.

For our second objective, we move beyond idealized models, utilizing state-of-the-art observations, reanalysis data, advanced numerical models, and machine learning techniques. Applying multivariate analysis typically used for SA, we examine convective organization and its impact on total column water vapor variability in three mesoscale sized domains within the tropical western Pacific warm pool region, lying either north of the equator (2-9N or 3-10N), or directly straddling the equator 3S-4N). During periods with limited SST gradients (boreal summer/autumn),

convection tends to be random, with small horizontal humidity gradients. In periods with a weak zonal SST gradient ($> 10^{-3} K km^{-1}$, boreal winter/spring), aggregated convection prevails, exhibiting larger humidity gradients, stronger outgoing longwave radiation, and a drier atmosphere. Intermittent episodes of random convection and smaller humidity gradients disrupt this pattern. A composite analysis of such events associates them with westward propagating convectively coupled Rossby waves in the three regions, suggesting these waves play a key role in mesoscale water vapor variability during boreal winter and spring, potentially influencing long-term variations in convective organization.

To further understand the mechanisms that lead and prevent the development of convection organization. Our study leverages on realistic simulations, underscoring the important role of humidity advection in orchestrating the organization and disorganization of convection, with wind shear playing a dual role in either organizing or disorganizing convection, contingent on its strength and direction. In terms of its characteristics, when convection is organized, the atmosphere is significantly drier compared to the random state. Diabatic feedbacks consistently work to cluster convection, but large-scale dynamics play a more important role in instigating and disrupting organization over the warm pool region, since the large-scale dynamics can export moist gross static energy from moist to dry regions, disallowing organization. This research improve our understanding of the mechanisms that lead to SA in idealized models. Also to bridge the gap between idealized models and real-world tropical atmospheric dynamics, imparting valuable insights into the multifaceted processes governing convective organization and its climatic implications.

Acknowledgements

For me, this represents the culmination of a journey that began in high school and evolved into a goal during my undergraduate studies. I thank Barbara Duse for always being present and wanting to help. I extend my gratitude to Adriana Gomez for her insightful comments on the initial draft of Chapter 3, to Sandrine Bony for her valuable feedback on the results of Chapters 3 and 4, and to Greg Thompson for his comments about microphysics schemes that help in the development of Chapter 3. Special thanks to Sebastian Muller; his work on Lagrangian trackers served as inspiration for creating my own tracker, employed in Chapter 3 and in a side project I led. Graziano Giuliani deserves acknowledgment for his guidance and assistance with Fortran coding on Chapter 3 and discussion on the boundary conditions modification experiments of Chapter 5.

I express appreciation for the contributions of the cdo, numpy, pandas, matplotlib, and xarray development teams; without their work, this thesis would not be feasible. My thanks also extend to all coauthors involved in side projects and collaborations, contributing valuable ideas that enhanced this thesis. Fred Kucharski, Erika Coppola, Cyril Caminade, and Riccardo Farnetti merit appreciation for allowing me to participate in their summer schools, where the methods presented were integral to my thesis. Fred and Erika's discussions and comments played a crucial role in strengthening the results presented in Chapters 4 and 5. Thanks to Pandora Melchose and Susanne Henningsen for all their administrative help, and for their kindness, and also to all the people that have come an go to the ICTP through the years (James, Rene, Abraham, Michie, Alessandro, Luiza, Waeb, Adnan, Natalia, and Chen). I am grateful to the ICTP for awarding me the scholarship and providing the necessary support to complete my Ph.D. Their commitment to assisting individuals from developing countries is truly commendable.

I extend my heartfelt gratitude to Giovanni Biagioli for his unwavering support and availability to discuss my results. His insightful comments were particularly crucial for the introduction and especially in the development of Chapter 3 and strengthening the arguments presented therein. Our collaborative journey has not only been professionally rewarding but has also fostered a meaningful friendship. Special thanks to Caroline Muller for her exceptional guidance and insightful comments, which contributed to outstanding results and inspired new experiments and diagnostics in Chapter 3. Beyond her professional expertise, I appreciate her kindness and support, fostering a positive environment that motivated me to work happily and hard. I extend my deepest gratitude to Adrian Mark Tompkins, my supervisor, for affording me the invaluable opportunity to collaborate with him. I am thankful for the extensive knowledge he imparted in physics, coding, and the principles of robust scientific inquiry. His unwavering support has been essential, extending beyond academic, to even the realm of my family. Moreover, I appreciate his encouragement to engage in various scientific events, exposing me to exceptional research and diverse cultures.

I apologize to the English speakers, but now I will continue in Spanish since many of my friends, previous teachers, and family don't know English, and, in fact, Spanish has an extra flavor that allows me to express gratitude in a more heartfelt way. Para empezar, quiero agradecer a mis profesoras del colegio: Erika Escobar,

Zulu, miss Elvira, miss Luzma, Claudia Cruz, La Chavita, Lina Bastidas, Jenny Pinzon, Francy Gomez y Catalina Montes. Cada una me apoyó en momentos que me marcaron y que me llevaron a ser la persona en la que me convertí, por lo cual espero nunca olvidarlas. A mis profes de la Universidad: Ellie Lopez, Carmen Cecilia, Camila Jimenez, Johanna Vargas, Luis Carlos Belalcazar, Daniel Hernandez, Daniel Molano y Gonzalo Cifuentes, por su apoyo y por mostrarme la belleza de la física y la investigación. A mis estudiantes que siempre han sido una fuente de inspiración. A los peluditos, Simon, Fede, Teo, Luna y ahora Publio, por enseñarme lo que es el amor sin condiciones. A mis amigos, Richi, Alejo, Juani, Andy, Roby, Loki, Laurita, El Calvo, Doly, Sergio V, Carito, Sergio B y Loz por su lealtad, por siempre apoyarme y por regalarme momentos mágicos. Agradezco a Ferro por su amistad, incansable apoyo y colaboración en todos los proyectos en los que participamos, y por siempre responderme preguntas sobre Machine Learning. Un agradecimiento especial a Plats, por ser tan leal y tener tanta confianza en mí; sus decisiones y apoyo fueron invaluables y espero que nuestra amistad prevalezca siempre.

Agradezco mucho a mi tío Luis, a mis tías Ruth y Marlen, y abuela Celia por ayudarme a llegar hasta aquí, a mi tía Alexandra por su confianza. A mi padre Rafael, por ser una persona maravillosa que ama a sus hijos por encima de todas las cosas, y que me enseñó que el trabajo, cualquiera que sea, debe hacerse bien y con toda la energía. A Camilo Argáez por apoyar tanto a mi padre, y a mi hermana Ana por amarlo con tanta fuerza, y por crecer siendo una mujercita tierna, parlanchina y maravillosa. A mi abuelo Armando por estar siempre pendiente de mi madre, y a mi madre, Ximena, por ser un pilar en mi vida, mostrarme que la fuerza del ser va por encima de cualquier contratiempo y dolencia, por amarme con tanta pasión, y por haberme aceptado desde pequeño tal y como soy. A mi abuela, La Cucha, en la que pienso cada vez que suena mi alarma, pues durante muchos años fue ella quien me despertó, siempre con una sonrisa. Su apoyo, su amor, confianza, lealtad y enseñanzas han sido maravillosos, y espero que sigan conmigo por siempre. Es la mejor abuela y espero que nos dure para siempre porque no sé qué haría sin ella. A mi suegro David, por criar a un ser tan maravilloso como mi esposa, Nathalia (Lovy), que llegó a mi vida para hacerme sentir infinitamente feliz, para amarme y mostrarme lo que es querer pasar toda mi vida al lado de alguien. A mi Lovy por ser un apoyo, un pilar, un ser maravilloso en mi vida; gracias a ella, esta experiencia no solo fue enriquecedora, sino una etapa que recordaré por siempre, gracias por mostrarme cuánto se puede amar y ser amado por alguien. Por último, gracias a Trieste, Grignano y al ICTP, que me regalaron años maravillosos que jamás olvidaré.

Contents

List of Figures

List of Tables

For the Marvelous Mysteries of Nature...

Chapter 1

Introduction

1.1 Atmospheric Convection

Atmospheric convection is a crucial process influencing Earth's climate, contributing over 50% of tropical rainfall (e.g., Nesbitt et al. 2000; Betts and Jakob 2002). It plays a vital role in maintaining the planet's energy and hydrological balance by vertically transporting energy and humidity. Defined as any motion driven by buoyancy and governed by the Navier-Stokes equations, atmospheric convection involves complex physical processes like phase changes, cloud microphysics, precipitation, and radiation. It occurs across various spatial and temporal scales, broadly categorized into two main types.

The first type of convection involves non-precipitating cumuli, arising from vertical air currents that may lack the strength for droplets to reach precipitation size. This limitation can be due to insufficient strength to penetrate into deeper atmospheric layers or not reaching temperatures low enough for substantial condensation (Emanuel, 1994). In contrast, the second type includes convection leading to precipitation, occurring when the atmosphere has high instability, which is characterized by strong vertical velocities and elevated humidity levels (e.g., Sherwood et al. 2010). In such cases, ascending air parcels from the surface reach upper atmospheric layers, where low temperatures induce condensation or freezing, leading to the formation of precipitation. The resulting droplets or ice crystals then descend due to their increased weight (Emanuel, 1994; Keil et al., 2008; Sherwood et al., 2010). In this second category is deep convection, further discussed in the following paragraphs.

Deep convection is distinguished by the upward movement of air parcels from the surface to the middle and upper troposphere, where condensation and melting processes transpire (e.g., Sherwood et al. 2010). The development of this convection type necessitates a moist anomaly and is closely linked to regional moisture sources within the atmosphere. To comprehend the nature of deep convection, it is imperative to dive into a multitude of processes integral to this phenomenon, which are outlined below:

(i) An air parcel that is warmer than its surroundings ascends, as it is less dense, exhibiting increased buoyancy. While the parcel ascends (due to sufficient buoyancy), it will expand and cool, becoming saturated, leading to condensation and freezing. This phase releases latent heat, providing additional energy for the parcel to reach the level of neutral buoyancy (LNB). At the LNB, strong stability and

turbulent mixing processes transpire at the cloud boundaries (Emanuel, 1991). Concurrently, the ascending parcel induces a decrease in pressure below the cloud, necessitating the convergence of other parcels to this region due to mass conservation (Holton, 2004). (ii) On the journey to reach the upper troposphere, a substantial proportion of water vapor within the parcel condenses to form water droplets that can precipitate. The remaining portion of water vapor and cloud condensate is detrained at the LNB. (iii) The upward in-cloud convective mass flux is balanced by compensating subsidence in the environment. This downward motion adiabatically warms the surrounding air. The cloud anvil influences the radiation balance through two mechanisms: reducing the emittance of infrared radiation (warming the atmosphere) and increasing cloud albedo (lowering surface temperature) (Hartmann, 2016). (iv) At this juncture, mid-troposphere downdrafts initiate, propelling cold and dry air to the boundary layer, resulting in strong surface gust fronts (cold pools). These gust fronts have the potential to trigger new convective events through both dynamical (Droegemeier and Wilhelmson, 1985; Fuglestvedt and Haerter, 2020) and thermodynamical (Tompkins, 2001b) processes.

Convection, driven by instabilities, seeks to restore environmental stability by heating and moistening the free troposphere. Its impact on crucial processes involves altering the radiation budget through anvil formation, releasing latent heat, and shaping moisture distribution via updrafts and downdrafts. Furthermore, convection plays a vital role in precipitation patterns, dynamically contributing to the hydrological cycle. A comprehensive understanding of atmospheric convection is essential for enhancing models and refining parameterizations. Notably, one characteristic of convection is its capability of organizing into clusters, which has a strong potential to impact climate sensitivity (Bony et al., 2015). Given the challenges faced by General Circulation Models (GCMs) in accurately representing the organization of convection into clusters across scales ranging from 100 km to 10000 km, it becomes imperative to dive deeper into this process and its underlying mechanisms, which is the aim of this research.

1.2 Radiative Convective Equilibrium

Before describing the significance of convective organization, it's essential to describe the Radiative Convective Equilibrium (RCE) framework, which establishes the foundational state for convective organization studies. RCE represents the statistical equilibrium achieved by the atmosphere and surface when lateral energy transport is removed, balancing convective heating and radiative cooling (Wing et al., 2017). Widely used as an idealization of the tropical atmosphere, especially on large scales (Wing and Emanuel, 2014), RCE provides a valuable tool for simplifying the study of convection and its interactions. Convection, typically connecting to the environment by transporting moisture and heat throughout the troposphere, influences the radiation budget through its anvils. Researchers have employed RCE in one-, two-, and three-dimensional models to enhance our understanding of convection (Manabe and Strickler, 1964; Nakajima and Matsuno, 1988; Held et al., 1993; Rennó et al., 1994; Tompkins and Craig, 1998a,b; Muller and Held, 2012; Wing and Emanuel, 2014; Tompkins and Semie, 2017; Muller et al., 2022a). In this context, RCE allows for the examination of the tropical atmosphere in an idealized setup, where

convective latent heating compensates for net radiative cooling in the free troposphere, and large-scale motions are absent.

RCE was initially employed by Manabe and Strickler (1964) in a single-column model (1D) to study the influence of atmospheric absorbers and convection on the thermal structure of the atmosphere. Subsequent researchers expanded its scope to include 2D (e.g., Held et al. 1993; Nakajima and Matsuno 1988; Grabowski and Moncrieff 2001, 2004a) and 3D models (e.g., Tompkins and Craig 1998a; Bretherton et al. 2005; Khairoutdinov and Emanuel 2013; Wing and Emanuel 2014; Hohenegger and Stevens 2018a) to unravel convection characteristics inadequately represented in GCM parameterizations. With technological advancements, modern RCE models span hundreds of kilometers and reach planetary scales in three dimensions (e.g., Arnold and Randall 2015; Bretherton and Khairoutdinov 2015; Müller and Hohenegger 2020a). The key advantage of RCE simulations lies in their ability to comprehend interactions among surface fluxes, the radiation budget (including longwave –LW– and shortwave –SW– components), advection processes, moisture, and sea surface temperature (SST), impacting convection (Wing et al., 2017). RCE facilitates experiment denial and sensitivity studies, aiding in determining the role of various variables in convection representation.

1.3 Convective Organization

Convective clouds exhibit a propensity to cluster. Several factors contribute to this behavior. For instance, convection tends to concentrate in regions with the warmest SSTs. These areas experience lower pressure, leading to convergence, thereby promoting the initiation and persistence of convection in the same location (e.g., Woolnough et al. 2000; Tompkins 2001a; Shamekh et al. 2020b,a; Tompkins and Semie 2021). Another organizing factor is the vertical wind shear, leading to the formation of squall lines (Thorpe et al., 1982; Rotunno et al., 1988; Emanuel, 1994; Weisman and Rotunno, 2004; Helfer and Nuijens, 2021; Abramian et al., 2022a) and promoting new convection. In cases of wind shear, rain may fall away from updrafts, creating a moist environment that does not inhibit the development of nearby convection (e.g., Rotunno et al. 1988; Abramian et al. 2022b).

Another form of organization is self-aggregation (SA), wherein convection organizes into clusters without external forcing. Three decades ago, this noteworthy phenomenon emerged in RCE simulations—one that global models struggle to accurately represent (Tobin et al., 2013). Convective SA, illustrated in Figure 1.1, was first observed by Held et al. (1993) in a two-dimensional model. Despite homogeneous conditions in the model's initialization, they noted that convection spontaneously organized into clusters, moistening proximate regions while inducing dryness in distant areas. Subsequently, Tompkins and Craig (1998a) identified the same phenomenon in a three-dimensional model setup, investigating its sensitivity to radiation. They found that homogenizing heating rates horizontally led to a breakdown of the cluster state into a more random configuration.

Following these first studies, researchers focused on understanding several aspects of convective SA. Muller and Held (2012) explored the sensitivity of SA to domain and grid sizes, advective processes (such as cold pools) feedbacks, and the

role of low clouds. They found that SA strongly depends on domain size and resolution. This sensitivity to domain size was further elucidated by Yanase et al. (2020), who discovered that convection clusters because subsidence induces cooling and drying, generating a circulation that transports moisture from dry to moist regions. In smaller domains, convection is so closely packed that this circulation can't overcome the one produced by cold pools (CPs), which form due to evaporative cooling. Regarding resolution, Biagioli and Tompkins (2023a) showed, in a simple stochastic model similar to that of Craig and Mack (2013) and Windmiller and Craig (2019), that improved resolution increases the number of updraft cores, reducing subsidence regions and hindering aggregation.

Regarding the diabatic feedbacks, Wing and Emanuel (2014) used a Moist Static Energy (MSE) budget, to report that the LW feedback is the most influential, and surface fluxes initially act to aggregate convection but work against it once aggregation is established. Other authors, such as Tompkins (2001c), focused on the water vapor feedback, highlighting its thermodynamic role in clustering convection, showing that the water vapor feedback emerges as a crucial variable, acting as both a precursor and sustainer of convective aggregation (Sherwood et al., 2010).

Studies also explore the impact of SST on convection behavior and its onset time (Bretherton et al., 2005; Hohenegger and Stevens, 2016a; Tompkins and Semie, 2021). Additionally, previous research have focused on how different sub-grid scale parameterizations influence convection behavior (Tompkins and Semie, 2017). These investigations form the foundation for understanding mechanisms that can induce or inhibit clustering, serving as crucial steps toward developing parameterizations that incorporate this phenomenon.

1.3.1 Mechanisms

This section explores processes that contribute to convective SA in non-rotating RCE simulations, for example, surface fluxes, radiation, and moisture feedbacks, play crucial roles in convective SA.

Surface Fluxes

Surface fluxes can promote aggregation through positive feedback, as indicated by prior research (e.g., Tompkins and Craig 1998a; Bretherton et al. 2005; Muller and Held 2012; Wing and Emanuel 2014), although this mechanism is not essential for the occurrence of aggregation. Sensitivity analyses have demonstrated that homogenizing surface fluxes can lead to aggregation, depending on factors like domain size, since larger domains favor clustering (Muller and Held, 2012), radiative feedback (Tompkins and Craig, 1998a), and the imposed background surface fluxes (Holloway and Woolnough, 2016). While surface fluxes can support aggregation, they are insufficient to cluster convection without radiative feedback, at least for current SST temperatures ($\approx$ 301 K to 305 K). An important feature of surface fluxes, is their positive and negative feedback with SA (Wing and Emanuel, 2014). The positive feedback tends to prevail in the initial stages of simulations, before aggregation develops, and transitions to negative feedback when convection is clustered. This surface fluxes feedback pattern is also observed in models with slab oceans (Tompkins and Semie, 2021). Positive feedback arises from stronger winds in moist regions, increasing surface fluxes and enhancing convective organization. In contrast, negative

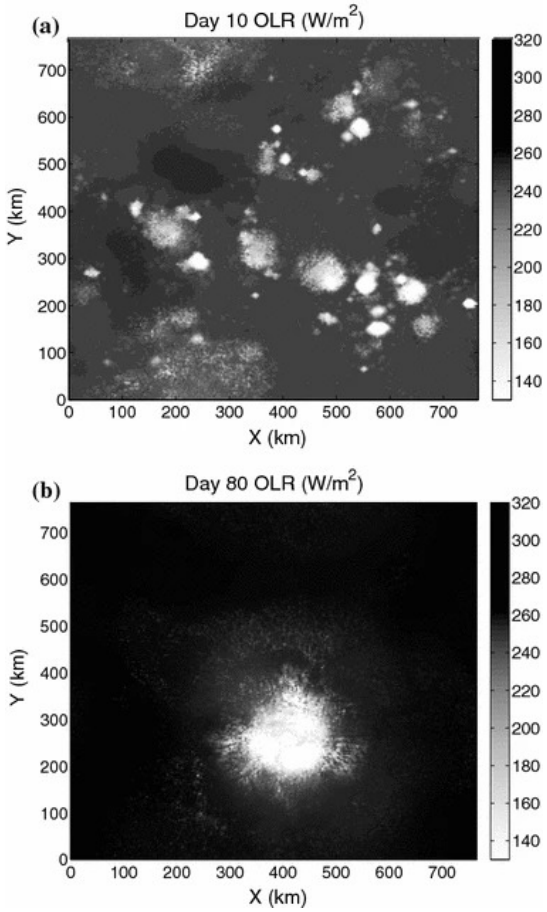

FIGURE 1.1: Snapshot of outgoing long wave radiation (OLR) at (a) day 10 and (b) day 80 of an RCE simulation at 305K. Available in Wing and Emanuel (2014).

feedback results from air-sea enthalpy disequilibrium, leading to reduced surface fluxes in moist regions.

Longwave Radiation

The LW radiation feedback has been extensively studied as it plays a crucial role in the occurrence of aggregation, as demonstrated by various researchers (e.g., Tompkins and Craig 1998a; Bretherton et al. 2005; Muller and Held 2012; Emanuel et al. 2014; Wing and Emanuel 2014; Muller and Bony 2015; Coppin and Bony 2015; Shamekh et al. 2020a). Mechanism denial experiments and the MSE budget have been employed to investigate this feedback. In the early stages, the LW radiation feedback with clouds is the key driver of SA, but gradually the LW-water vapor feedback becomes significant as a dry patch emerges where convection is suppressed, leading to a considerable increase in radiative cooling. This cooling process induces a shallow circulation that can further promote the organization of convection (e.g., Yang 2019). The circulation driven by radiative cooling involves low-level clouds in dry regions transporting MSE upgradient through subsidence and advection from the surface of dry areas to the moist convective zones, enhancing the clustering of convection (Coppin and Bony, 2015; Holloway and Woolnough, 2016). Another outcome of the

LW radiation feedback is the diabatic effect, causing a decrease in MSE in dry regions (due to cooling) and an increase in moist regions (e.g., Yao et al. 2022). The shallow circulation and the reinforcement of the MSE gradient create a strong positive feedback, favoring clustering (Shamekh et al., 2020a).

Unlike surface latent heat flux feedback, LW radiation represents a positive feedback throughout the entire simulation (Wing and Emanuel, 2014). The radiative cooling rates are nevertheless sensitive to the choice of radiation and microphysics schemes (Wing and Cronin, 2016). This sensitivity implies that SA can be influenced by various aspects of the parameterizations employed in models, as demonstrated by Tompkins and Semie (2017) and Cerlini et al. (2023) concerning, for example, sub-grid scale mixing.

Shortwave Radiation

SW radiation, unlike LW radiation, appears to play a secondary role in the aggregation of convection. Some authors (e.g., Muller and Held 2012; Holloway and Woolnough 2016) have suggested that SW radiation may have a negative (albeit weak) feedback on aggregation due to down-gradient transport of MSE induced by heating anomalies driven by deep clouds. Conversely, SW radiation exhibits positive feedback resulting from its diabatic effect and the reduction of SW radiative cooling in clear sky dry regions (Wing and Cronin, 2016). This positive feedback in the dry regions is due to MSE transport, but in the moist regions, it can counteract aggregation due to the effects of deep convection on the upper troposphere. It is important to note that this feedback is not crucial for the occurrence of clustering. The impact of SW radiation feedbacks is stronger in simulations using a reactive slab ocean as a lower boundary condition, rather than spatially fixed SSTs, which is discussed below.

Advective Processes

There is no consensus in the literature regarding the feedback generated by advective processes. Questions persist about whether these processes can induce SA or if they work to sustain and amplify it. Several authors have tackled these questions, identifying two crucial elements in this feedback: a shallow circulation resulting from radiative cooling at the tops of low clouds in dry regions, and changes in MSE variance due to advective processes.

In the first process, the shallow circulation is an indirect effect of radiation. Low-level clouds in dry regions create radiative cooling at their tops, leading to the transport of MSE from dry to moist regions, promoting convection clusters (Muller and Held, 2012; Coppin and Bony, 2015; Muller and Bony, 2015). In the second process, Wing and Emanuel (2014) found that the feedback is positive after aggregation occurs but negative during its initiation. These findings are complemented by Wing and Cronin (2016), who used MSE variance to show that, when convection clusters, MSE is exported from moist to dry regions (negative feedback). However, even in cases where MSE variance produce a negative feedback, a shallow circulation can still enhance convection (Coppin and Bony, 2015). Further insights from Holloway and Woolnough (2016) indicate that this shallow circulation is balanced by other advection terms, suggesting it is produced and enhanced by horizontal heating anomalies due to convection rather than radiative cooling. On the other hand,

Yanase et al. (2020) suggest that advective processes, such as those produced by CPs, may prevent clustering, by acting as a negative feedback.

Water Vapor

Moisture feedbacks are as crucial as radiation feedback in influencing SA. In the absence of radiation feedback, moisture alone can initiate and sustain SA (Craig and Mack, 2013; Coppin and Bony, 2015). Moisture exhibits positive feedback with aggregation, as increased moisture stimulates more convection (Sherwood et al., 2010), and convection produces more moisture in its vicinity, creating a feedback. Tompkins (2001c) highlights that water vapor plays a vital role in determining convection locations, with a dry atmosphere strongly impeding convection due to convective inhibition (CIN) and a moist atmosphere favoring convection due to the convective available potential energy (CAPE). Additionally, convection tends to occur over warmer SSTs and is suppressed over colder SSTs because the air above colder SSTs dries out due to subsidence countering convection. Altering the positions of hot and cold spots does not immediately initiate convection; instead, it triggers an advective moisture process that subsequently leads to convection, emphasizing the vital role of water vapor in convective aggregation.

This crucial feedback mechanism is still not entirely understood, and studies utilizing sensitivity and denial experiments have sought to unveil the underlying physical processes governing this water vapor feedback. For instance, increasing precipitation efficiency enhances convection by boosting instability in the free troposphere (Emanuel et al., 2014). Another key mechanism is that, in a moist environment, an ascending parcel will lose less buoyancy on entrainment mixing as a result of reduced evaporative cooling (Tompkins, 2001c; Mapes and Neale, 2011; Holloway and Woolnough, 2016; Wing et al., 2017; Yang, 2019). The presence of CPs is also integral to the water vapor feedback, as they redistribute humidity. In their absence, aggregation is favored (Muller and Held, 2012; Jeevanjee and Romps, 2013; Muller and Bony, 2015), as moist regions continuously receive moisture due to reduced mixing of dry and moist air, and the evaporative cooling circulation that disfavors aggregation (Yanase et al., 2020) would not develop. While CP sensitivity has been explored by various authors through the removal of rain evaporation, resulting in strong clustering even without radiative feedbacks, changes in rain evaporation could affect convective heating (Yang, 2018), convective downdrafts, and surface fluxes, all crucial for organization (Holloway and Woolnough, 2016; Holloway, 2017). This implies that further research is necessary to comprehend the role of CPs in the development of aggregation.

Sea Surface Temperature

Convection can occur across a broad spectrum of SSTs, ranging from snowball Earth temperatures (Abbot, 2014) to scenarios with an SST of 310 K (Wing and Emanuel 2014; Wing and Cronin 2016). Wing and Emanuel (2014) determined that aggregation does not manifest at temperatures below 300 K but can occur at temperatures as high as 310 K. However, this conclusion was drawn from research employing a fixed SST, necessitating further investigation into this threshold. Other studies have replaced the fixed SST with an interactive slab ocean to disentangle the characteristics and feedback of convective SA with SST variations. Initial slab ocean

experiments (Tompkins, 2001a,c; Bretherton et al., 2005; Reed et al., 2015; Hohenegger and Stevens, 2016a) found that convection clustering onset occurs later in slab oceans compared to fixed SSTs. This delay may arise from different processes in the boundary layer that decelerate the clustering of convection without preventing its occurrence (Tompkins and Semie, 2021).

Shamekh et al. (2020b) demonstrated that organization is responsive to SST hotspots. They introduced a hotspot at the center of the domain, leading to the clustering of convection in that area due to convergence. Further work of Shamekh et al. (2020a) revealed that in an interactive slab ocean, dry patches emerge and are subsequently strengthened by a positive surface pressure anomaly associated with a negative moist anomaly, induced by radiative cooling from subsidence. To establish this, the authors decomposed buoyancy components into temperature and humidity, illustrating that as the dry patches initiate, the negative moisture anomaly surpasses the positive temperature anomaly and sustains the growth of the dry patch due to its close correlation with surface pressure—an insight previously alluded to by Yang (2018) and corroborated by Yao et al. (2022) using the MSE framework.

Another approach to investigate the relationship between SST and convective organization, particularly the aggregation onset, was undertaken by Tompkins and Semie (2021). They employed an ensemble of simulations employing an adaptive Q-flux method designed to prevent any drift in domain-mean temperatures while enabling the surface to respond to diurnal forcing. They utilized an interactive slab ocean model that simulates a vertically uniform temperature with a fixed mixed layer depth, allowing the SST to depend on LH and SH fluxes, as well as radiative fluxes.

They concluded that an interactive ocean delays the onset of organization, and this delay varies according to the ocean layer depth, with smaller ocean depths resulting in larger clustering delays. They also observed that higher temperature gradients between the cloudy and clear sky regions affect the MSE and buoyancy gradients due to LH and SH fluxes, determining the location of convective triggering and clusters. While the LH flux acts as a negative feedback for convective clustering, it is not sufficient to prevent it. An important aspect of Tompkins and Semie (2021)'s research is the demonstration of the relationship between SST anomalies and total column water vapor (TCWV), which open avenues for interesting studies seeking organization in observations, as explored further in Chapters 4 and 5.

1.3.2 RCEMIP

The RCEMIP protocol is described by Wing et al. (2018). The project consists of a set of multiple simulations from different models with consistent configurations to assess their similarities and further understand SA and its consequences for climate. The idea is to perform simulations in different domains (i.e., long channels, large squares, and small squares). The small domains are used as a control since SA is not expected to develop there (Muller and Held, 2012; Yanase et al., 2020; Biagioli and Tompkins, 2023a). To include climate change scenarios and understand how aggregation could be a part of the climate sensitivity, the protocol contemplates different SST temperatures (the models are run with a constant SST).

According to Wing et al. (2020), models align on several SA aspects. As clusters form, the troposphere tends to become drier and warmer, tightly linked to the hydrological cycle. Another consensus emerges around the rise in OLR with increased clustering, indicating heightened climate sensitivity, since more energy is escaping the earth. Regarding clouds, models agree that when aggregation develops, the anvils height ascent and their temperature increases. Models also suggest that higher SST expedite this increase, especially during the clustering phase. Anvil cloud fraction tends to decrease with warming and aggregation, although around 30% of the models lack consensus on this aspect. However, models indicate minimal to negligible changes in mid-level and low-level cloud fractions.

Following the delineation of areas where models align regarding aggregation, it is crucial to address discrepancies, as they bear significance for this thesis. A primary source of disparity lies in the choice of aggregation metrics, where varying results emerge depending on the metric employed (Wing, 2019). Even within a single model, metrics may yield conflicting indications of clustering and randomness, posing a challenge in diagnosing the impact of convection (Biagioli and Tompkins, 2023b). This issue assumes particular importance when characterizing changes in clustering with warming, as no consistent patterns emerge. The spread in outcomes cannot be attributed solely to metric choices; it also stems from the fact that half of the models exhibit increased aggregation with warming, while the remaining models demonstrate a decline (Wing et al., 2020). These findings underscore two critical challenges: the need for improved methods to quantify aggregation and the realization that, even within a consistent and sufficiently large setup conducive to develop clusters (e.g., Muller and Held 2012, Yanase et al. 2020), aggregation could not develop. This suggests sensitivity to factors such as parameterizations, as demonstrated by earlier studies (e.g., Wing and Cronin 2016; Tompkins and Semie 2017; Shi and Fan 2021; Huang and Wu 2022) highlighting the capacity of certain parameters to modify clustering characteristics.

In the realm of climate sensitivity, the RCEMIP project has unveiled crucial insights. Becker and Wing (2020) demonstrated that climate sensitivity is markedly influenced by changes in convective clustering with warming. The capacity of aggregation to either increase or decrease with warming stands as a vital factor in determining climate feedback. The study also reveal that lower climate sensitivities often coincide with an upswing in shallow cloud fraction. These two factors account for 70% to 80% of the spread in model climate sensitivity. Emphasizing the need for precise representation of SA and shallow convection in parameterizations is crucial for more accurate climate sensitivity estimations. Moreover, Becker and Wing (2020) underscored the significance of the moisture memory feedback, particularly in future climate scenarios marked by a saturation deficit. This underscores the imperative of ensuring that parameterizations can faithfully reproduce this feedback. The Becker and Wing (2020) study highlighted that parameterizations, especially in their treatment of radiation, clouds, and moisture, may wield a crucial influence on the emergence and persistence of SA, something we expand upon in Chapter 3.

1.3.3 Organization in Observations

Convective SA, observed in Cloud Resolving Models (CRMs), is crucial for explaining various convective structures like tropical cyclones or the Madden-Julian Oscillation (MJO). For this, one question arises: does this phenomenon occur in real-world

observations or is it just a model artifact? Despite efforts to identify convective organization, studying organization at mesoscales in observational data has been limited (Holloway et al., 2017), with mostly previous work focusing on larger scale convectively coupled waves, such as Rossby and Kelvin waves as well as other forms of convective organization such as the MJO, tropical cyclones and even the Hadley and Walker circulations. Some of these larger scale forms of organization are clearly visible in Figure 1.2, which displays a visible Earth image, revealing diverse convective organization types. This visual evidence suggests convective organization occurs in different natural systems over different spatial scales.

Concerning convective organization, Tobin et al. (2012) and Tobin et al. (2013) observed that the number of clusters in a domain influences domain-wide properties, such as larger OLR and reduced humidity with fewer cores, akin to CRMs. They also noted discrepancies between organization in models and observations. In idealized models, surface fluxes were larger outside convection, a pattern not distinctly observed in real-world data. Tobin et al. (2012) reported increased surface fluxes in both convective and non-convective regions, while Tobin et al. (2013) found no sensitivity of surface fluxes to organization at smaller scales.

From the work of Tobin et al. (2012, 2013), numerous studies have endeavored to quantify convective organization in the atmosphere (Holloway et al., 2017; Beucler et al., 2019). Tobin et al. (2012, 2013) introduced the simple convective aggregation index (SCAI), incorporating parameters such as the number and distance between clusters. While subsequent research has refined and expanded on these indices (see, Coppin and Bony 2015; Wing and Cronin 2016; Tompkins and Semie 2017; White et al. 2018; Popp and Bony 2019; Retsch et al. 2020; Hohenegger and Jakob 2020; Beucler et al. 2020; Brune et al. 2021, among others), it has become evident that the degree of organization is contingent upon the specific index employed (Wing, 2019; Pscheidt et al., 2019; Biagioli and Tompkins, 2023b). This underscores the importance of exploring novel indices and methodologies to quantify organization, particularly in the natural environment where understanding atmospheric phenomena on a large scale is crucial.

Following the comprehensive review by Holloway et al. (2017), a substantial body of literature has focused into quantifying and examining convective organization and their implications for the radiation budget and the hydrological cycle. Convective organization can impact the hydrological cycle by inducing atmospheric drying (Hohenegger and Jakob, 2020). Moreover, it influences extreme precipitation events (Semie and Bony, 2020; Pendergrass, 2020; Bläckberg and Singh, 2022; Angulo-Umana and Kim, 2023) and the broadening of rain belts (Popp and Bony, 2019). The work of Vogel et al. (2022) suggested that SA feedbacks have the potential to produce a net moisture convergence, which would help explain correlations between vertical velocity and relative humidity that are critical for comprehending the trade cumulus feedback. In observational studies of organization and its consequences, Semie and Bony (2020) and Angulo-Umana and Kim (2023) employed the organization index (I_{org}) from Tompkins and Semie (2017). Utilizing the fifth-generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5) product (Hersbach et al., 2020), Angulo-Umana and Kim (2023) revealed that strong clustering leads to more intense precipitation events due to an increase in convective precipitation rate. This underscores the significance of accurately representing organization in parameterized models used for climate change

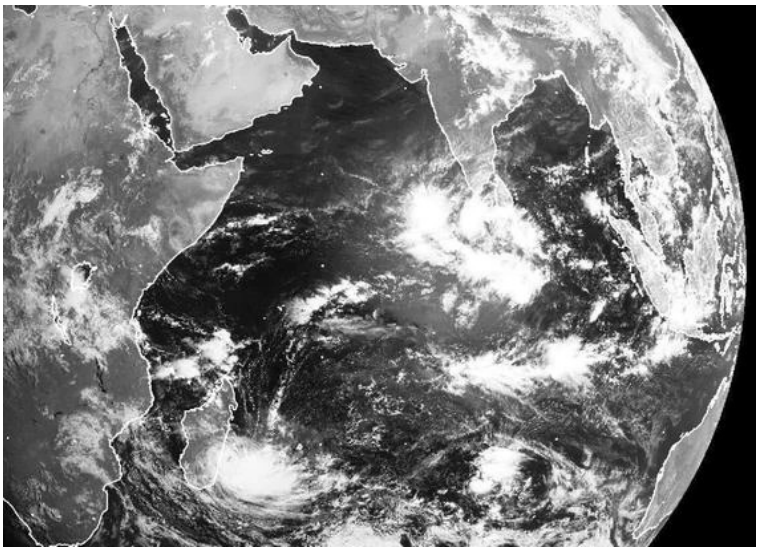

FIGURE 1.2: Visible satellite image of convective organization in long, medium, and large scales. Available in Holloway et al. (2017) and in http://www.sat.dundee.ac.uk/

simulations.

Furthermore, Semie and Bony (2020) discovered that organized convection results in more intense extreme precipitation, both over the ocean and land. The most intense precipitation occurs over land, and also over warm SST regions, with a stronger organization corresponding to increased local precipitation. This emphasizes the importance of accurate representation of convective organization not only over the ocean but also over land (Hohenegger and Stevens, 2018a). Further, Pendergrass (2020) not only establishes a correlation between extreme precipitation likelihood and the degree of organization but also notes that mean precipitation could increase with more clustered convection. Underscoring the need for a better understanding of the mechanisms driving this relationship, as it holds crucial implications for climate change adaptation policies.

Continuing with precipitation-related investigations, Popp and Bony (2019) demonstrated that increased zonal clustering of convection at the equator leads to a meridional widening of the tropical rain belt. The study indicated that this meridional shift in the rain belt is entirely attributable to the zonal organization of convection. Furthermore, the authors observed a broadening of the Intertropical Convergence Zone (ITCZ) when convection exhibits greater zonal organization, signifying a convergence of substantial moisture into convective regions and, likely, a drier subtropical zone. In alignment with this, Beucler et al. (2020), utilizing the moist margin index, and Hohenegger and Jakob (2020), employing a convergence lines index, demonstrate that enhanced organization of the Atlantic ITCZ results in increased tropical moisture and drier subtropics—a phenomenon extensively observed in idealized models that manifest convective self-aggregation (Müller and Hohenegger, 2020b).

The influence of convective organization on humidity has notable implications, as drier atmospheres are less opaque (Galewsky et al., 2023), leading to an anticipated increase in LW clear-sky radiation. Thus LW all-sky radiation should also be influenced by the degree of organization. Investigating these radiative changes,

Bony et al. (2020) found that convection exhibits organization in observations and plays a crucial role in the radiation budget of the large-scale tropics. Their study revealed that convective organization, in conjunction with lower-tropospheric stability, accounts for over 60% of the variance in the net radiative budget of the entire tropics. Consequently, comprehending the mechanisms governing convective organization becomes essential, particularly as it raises concerns about how this phenomenon might evolve in warmer climates and the potential consequences thereof. In light of this, questions arise concerning whether the sensitivity of humidity and the radiative budget to organization remains consistent at the grid-box scales of GCMs, and if the mechanisms orchestrating convection at these scales mirror those observed in CRMs.

A relatively recent and under explored approach involves utilizing realistic simulations to unravel the mechanisms behind convective organization and its impact on the domain mean. Notably, Brune et al. (2018) employed realistic high-resolution simulations to develop an index based on wavelets, enabling the quantification of convective organization and its dynamical properties such as CAPE and wind shear. Through observations and new realistic simulations in the tropical Atlantic, Brune et al. (2020) demonstrated the association between organization, wind shear, and CAPE, revealing that precipitation intensity correlates with TCWV, upper-level divergence, and maximum vertical wind speed. Similarly, Pscheidt et al. (2019) utilized realistic simulations to showcase the model's capability in measuring the degree of organization and capturing certain organizational aspects, such as the correlation between the number of convective objects and precipitable water, consistent with Tobin et al. (2012).

The studies aforementioned, focused on examining the properties of organized events rather than diving into the mechanisms that drive organization and comparing them with idealized studies. Holloway (2017) took a different approach by investigating the similarities between idealized simulations and realistic simulations for 15 organized cases, using an MSE budget. Conducting sensitivity experiments that altered factors such as rain evaporation, surface fluxes, and radiation feedback without changing the boundary conditions, he sought to quantify whether the mechanisms leading to organization in idealized simulations were applicable in more realistic settings. The findings indicated the significance of radiation feedback for organization to develop, with a decrease in organization observed when the radiative feedback was turned off, aligning with results from idealized simulations. The forcings showed positive radiative feedbacks and negative surface fluxes feedback, consistent with idealized studies. This approach not only explores the commonalities between organization in idealized and realistic scenarios but also quantifies potential differences between them.

1.4 Questions to be answered

Here, we outline the key questions that this project seeks to address, providing introductory context for each. It's worth noting that these questions may lead to further inquiries, which, while not explicitly outlined here, will be explored within their respective chapters.

1.4.1 Question 1: What is the sensitivity of self-aggregation to the model subgrid-scale parameterizations?

As discussed in section 1.3.2, convective SA does not consistently occur in domains of the same size, indicating the presence of other influencing factors that require investigation. Wing and Cronin (2016), Tompkins and Semie (2017), and Shi and Fan (2021) demonstrated that the choice of radiative scheme, sub-grid scale mixing, and microphysics scheme significantly impacts the degree of organization in simulations. The sensitivity of convection to these parameters can lead to either highly clustered or randomly distributed patterns. Considering these observations and the insights from RCEMIP, Becker and Wing (2020) emphasized the necessity of a detailed physical examination to comprehend the substantial variations observed across different models. However, the investigation of the sensitivity of SA to parameterization schemes is still relatively underexplored.

This is why Question 1 of this thesis revolves around this topic. The objective is to conduct a series of experiments aimed at comprehending the underlying mechanisms that either promote or hinder SA and how these are impacted by model parameterization choices. Chapter 2 will first introduce the model used for these experiments and relevant parameterization schemes investigated. Subsequently, Chapter 3 will present an analysis of idealized RCE simulations to identify the most important parameters influencing SA. The simulations will encompass various combinations of microphysical schemes, sub-grid scale mixing, and PBL. This comprehensive set of experiments aims to unravel the distinct roles of these parameters in SA. The analysis will show that model disparities are linked to the influence of parameterizations on the number of convective cores and maximum free convective distance (Biagioli and Tompkins, 2023a).

1.4.2 Seeking organization in observations and realistic model simulations

This work endeavors to discern patterns of convective organization in observations and understand their defining characteristics. In this section, we present the two questions that encapsulate our inquiry and underscore their significance.

- **Question 2:** Can we find convective organization in a GCM grid-box sized domain? And if so, what are its impact on the hydrological and radiative budget?

As detailed in Section 1.3.3, numerous endeavors have been devoted to quantifying the spontaneous organization of convection in nature. Many of these efforts utilize various indices, such as the number of convective clusters, providing initial glimpses into the presence of organization in nature. However, further work is essential to refine the quantification of organization and investigate the relevance of the same forcings observed in CRMs to real atmospheric conditions. The challenge in spatially defining indexes arises from the absence of cloud core measurements (Biagioli and Tompkins, 2023b). Holloway et al. (2017) reports that other ways to seek for organization in nature could be to search for some of its characteristics (e.g., water vapor variance). In Chapter 4, we expand on this notion by exploring organization through the SST-TCWV relationship in the Pacific Warm Pool, since in idealized models with an interactive slab ocean, convective clusters align with the warmest SSTs in the moistest regions of the domain (Tompkins and Semie, 2021). Conversely, when this relationship is inverted, leading to the warmest SSTs in the driest domain

zones, convection assumes a random configuration. With this relationship we could explore the mechanims and feedbacks that lead to organization, and quantify the impact of clustering convection in the domain humidity and radiative budget.

- **Question 3:** Can realistic simulations help to understand the mechanisms that organize or disorganize convection in a mesoscale sized domain??

In addressing Question 2, we will show that organization occurs in nature and yields significant consequences for domain mean properties. Revealing that diabatic forcings alone are not always sufficient to induce convective clustering, particularly in boreal summer/autumn and intermittently in boreal winter/spring. In Chapter 5, we dive into both scenarios—organized/random in boreal winter/spring and random in boreal summer/autumn—to enhance our comprehension of the characteristics fostering organization and those impeding it. Our focus extends to events capable of disrupting convection clusters, as these instances may constitute a missing element in convective parameterizations. Leveraging on realistic CRM simulations and sensitivity experiments (Holloway, 2017), involving modifications to boundary conditions and Machine Learning (ML) experiments we aim to understand the mechanisms that drive clustered and random convection, and also the ones that lead to the disbanded of convection clusters.

Each of the three questions delineated here corresponds to a distinct article in preparation for publication. Consequently, the related chapters are self-contained and can be understood independently. Subsequently, a concluding chapter summarizes the thesis and provides concluding remarks, shedding light on the addressed aspects and introducing new questions emerging from our findings. Additionally, Appendix D offers a concise summary of additional publications produced during the Ph.D., elucidating their connection to the methods developed in the thesis along with respective citations.

Chapter 2

Model description

This section will describe the model's main characteristics, such as the governing equations. We also detail some of the related parameterization schemes used to perform the simulations and report the model's important technical features.

2.1 Governing equations

The Weather Research Forecast (WRF) model version 4.2.2 (Skamarock et al., 2019) is used as a cloud resolving model (CRM) both in an idealized and realistic experimental setups. The model has a dynamical core that numerically integrates the non-hydrostatic, compressible Euler equations in their flux form. Euler equations have been generalized based on the vertical coordinate to account for terrain perturbations. To generalize the vertical coordinate, the WRF model follows the hybrid sigma-pressure vertical coordinate, as described in Park et al. (2013), which gives the equation (2.1).

$$p_d = B(\eta)(p_s - p_t) + [\eta - B(\eta)](p_0 - p_t) + p_t \tag{2.1}$$

Where $\eta = (p_d - p_t)(p_s - p_t)^{-1}$ and represents the terrain-following hybrid sigma-pressure vertical coordinate. p_d is the dry hydrostatic pressure, p_s and p_t correspond to the values of p_d in the surface and in the top boundary, respectively. p_0 is the sea-level reference pressure and $B(\eta)$ is the relative weighting between the terrain following the sigma coordinate and the pressure coordinate and have the form of equation (2.2)

$$B(\eta) = c_1 + c_2\eta + c_3\eta^2 + c_4\eta^3 \tag{2.2}$$

In which the values of c depend on pure pressure values of η. This is a new feature in the WRF model, that provides a better simulation results, especially when there is land or different topography features in the domain, as seen in Figure 2.1b of Skamarock et al. (2019).

To define the Euler equations in flux form it is necessary to account not only for the vertical coordinate but also for the vertical coordinate metric (μ_d, equation 2.6 in Skamarock et al. 2019), the wind flux, the potential temperature flux and the mixing ratio of the moisture variables. Variables that are described in equation (2.3)

$$V = \mu_d \mathbf{v} = (U, V, W); \Omega = \mu_d \omega; \Theta_m = \mu_d \theta_m; Q_m = \mu_d q_m \tag{2.3}$$

Where $\mathbf{v} = (u, v, w)$ are the covariant velocities of the wind, ω is the contravariant vertical velocity, Θ_m is the moist potential temperature, and Q_m represents the mixing ratios of each moist variable (e.g. ice, liquid water condensate, water vapor).

The geopotential ($\phi = gz$) is also a prognostic variable, but is not described here since it is not a conserved quantity (Skamarock et al., 2019).

Using the vertical coordinate, the wind, and potential temperature in flux form, we can express the governing Euler equations as from equations (2.4) to (2.10). We can also express the diagnostic equation of dry hydrostatic pressure (equation 2.11) and the diagnostic equation for the total atmospheric pressure (equation 2.12).

$$F_U = \frac{\partial U}{\partial t} + (\nabla \cdot V_u) + \mu_d \alpha \frac{\partial p}{\partial x} + \alpha \rho_d \frac{\partial p}{\partial \eta}\frac{\partial \phi}{\partial x} \tag{2.4}$$

$$F_V = \frac{\partial V}{\partial t} + (\nabla \cdot V_v) + \mu_d \alpha \frac{\partial p}{\partial y} + \alpha \rho_d \frac{\partial p}{\partial \eta}\frac{\partial \phi}{\partial y} \tag{2.5}$$

$$F_W = \frac{\partial W}{\partial t} + (\nabla \cdot V_w) - g[\alpha \rho_d \frac{\partial p}{\partial \eta} - \mu_d] \tag{2.6}$$

$$F_{\theta_m} = \frac{\partial \Theta_m}{\partial t} + (\nabla \cdot V\theta_m) \tag{2.7}$$

$$\frac{\partial \mu_d}{\partial t} + (\nabla \cdot V) = 0 \tag{2.8}$$

$$\frac{\partial \phi}{\partial t} + \mu_d^{-1}[(V \cdot \nabla \phi) - gW] = 0 \tag{2.9}$$

$$F_{Q_m} = \frac{\partial Q_m}{\partial t} + (\nabla \cdot V q_m) \tag{2.10}$$

$$\frac{\partial \phi}{\partial t} = -\alpha_d \mu_d \tag{2.11}$$

$$p = p_0 (\frac{R_d \theta_m}{p_0 \alpha_d})^{\gamma} \tag{2.12}$$

In these equations, g is the gravitational force, α_d is the specific volume of dry air, and α is the inverse of the density accounting for the moist and the dry air. $\gamma = C_p C_v^{-1}$ is the heat capacity ratio of the dry air, ρ_d is the density of dry air, and R_d is the dry air constant. The forcing terms of the model are F_U, F_V, F_W, F_{θ_m}, F_{Q_m} and appear due to the model physics (e.g. cloud microphysics, turbulence). For specific details about these equations and the way to solve them numerically consult Skamarock et al. (2019).

2.2 WRF Technical features

WRF model uses the Third-order Runge-Kutta (RK3) and the high-frequency acoustic modes to integrate the Euler equations in time and prevent numerical instabilities (Wicker and Skamarock, 2002). To implement this integration, the model uses the Arakawa C-grid staggering for every variable and the changes made in the spatial grid (Δx and Δy) resolution that accounts for the sphere projections and the vertical levels. The time step reported in equation (2.13) depends on the Courant number and also on the choice of the advection scheme since higher-order advection schemes require smaller time steps.

$$\Delta t_{max} < \frac{Cr_{theory}}{\sqrt{3}} \frac{\Delta x}{u_{max}} \tag{2.13}$$

In which Δt_{max} is the largest time step that can be selected without numerical instabilities, Cr_{theory} is the Courant number taking from theory (Wicker and Skamarock, 2002) and u_{max} is the maximum velocity expected in the simulation (Skamarock et al., 2019). It is important to mention that for convective permitting resolutions it is better to reduce the time step even more ($\approx 25\%$) to prevent instabilities since the updraft velocities can produce errors in the numerical integration.

The WRF model has 6 different advection order schemes. For the even-order (2^{nd}, 4^{th} and 6^{th}) advection schemes, operators are centered spatially and do not take into account any implicit diffusivity apart from the one that the RK3 integration has. Conversely, odd-order (3^{rd} and 5^{th}) advection schemes are upwind-biased and their discretization is naturally diffusive, which means that the odd-order advection terms are equivalent to the next-order even-order scheme but with a dissipation term of this even-order scheme and a constant similar to the Courant number (Wicker and Skamarock, 2002). In this research, we use the 5^{th} order advection scheme for the horizontal scalar and momentum fields and the 3^{rd} order scheme for the vertical scalar and momentum fields as recommended by Skamarock et al. (2019). Another important feature is the Rayleigh damping that is used to remove or smooth the vertically propagated gravity waves (Klemp et al., 2008; Skamarock et al., 2019) to prevent nonphysical wave reflections in the upper boundaries of the domain.

2.3 Parameterizations

2.3.1 Radiation schemes

The Rapid Radiative Transfer Model for General Circulation Models (RRTMG) scheme, utilized in all conducted runs, represents an enhancement by Iacono et al. (2008) over the previous scheme developed by Mlawer et al. (1997). Employing a k-distribution integration method and gas spectral bands tables, RRTMG represents both LW and SW radiation transfer associated with different gases and accounting for clouds. The scheme assumes a cloud particle effective radii that is consistent with the microphysics scheme. A significant refinement involves addressing uncertainties in the top of the atmosphere (TOA) downward LW flux produced above the model top. The RRTMG scheme has undergone rigorous testing and is widely regarded as one of the most precise in the field (Iacono et al., 2008).

2.3.2 Surface Layer schemes

These surface layer schemes govern the friction velocities and exchange coefficients crucial for computing moisture and surface heat fluxes from both land-ocean and boundary layer surfaces. In scenarios featuring an exclusive ocean surface, the surface layer scheme independently computes surface fields, differing from land surfaces where the scheme serves as input for moisture and heat flux calculations. All schemes adhere to the Monin-Obukhov similarity theory (Monin and Obukhov, 1954), incorporating variations in stability functions to determine surface exchange coefficients for momentum, moisture, and heat, as detailed in equations (2.14) and (2.15) (e.g., Paulson 1970; Webb 1970; Dyer and Hicks 1970).

$$\phi_M(zL^{-1}) = k_v z u_*^{-1} \frac{\partial u}{\partial z} \tag{2.14}$$

$$\phi_H(zL^{-1}) = k_v z \theta_{v*}^{-1} \frac{\partial \theta_v}{\partial z} \tag{2.15}$$

Here, ϕ_M and ϕ_H are the momentum and turbulent heat flux similarity functions, respectively. L is the Monin-Obukhov length scale, k_v is the Von Karman constant, u_* is the friction velocity, θ_{v*} is the scaling temperature and θ_v is the potential temperature. The similarity functions depend on the bulk Richardson number (R_{bulk}) which describes the turbulent state of the atmosphere as in equation (2.16).

$$(z_a L^{-1}) = R_{bulk}(1 - 5R_{bulk})^{-1} \ln(\frac{z_a}{z_0}) \tag{2.16}$$

Where z_a and z_0 are the vertical coordinate inside the boundary layer and in the ground, respectively. On the other hand, the eddy diffusivity exchange coefficients for momentum (K_M) and heat fluxes (K_H) are define in equations (2.17) and (2.18) (Monin and Obukhov, 1954).

$$K_M = \frac{k_v z u_*}{\phi_M(zL^{-1})} \tag{2.17}$$

$$K_H = \frac{k_v z u_*}{\phi_H(zL^{-1})} \tag{2.18}$$

Where K_M and K_H are related to the turbulent Prandtl number (Pr_t), as show in equation (2.19) (Monin and Obukhov, 1954).

$$\frac{K_H}{K_M} = \frac{1}{Pr_t} > 1 \tag{2.19}$$

2.3.3 Sub-Grid Scale Mixing schemes

General Description

GCM models that are not run in the large eddy simulation setup have to use turbulence parameterizations to account for the subgrid atmospheric mixing caused by eddy diffusion processes. For this, the schemes try to represent the turbulent mixing assuming that an adiabatically conserved quantity (θ_f) turbulent flux is related to its own gradient as shown in equation (2.20). Note that the important factor in the turbulence flow is related to the vertical component of the momentum equations as described by many authors (e.g. Holton 2004)

$$\overline{w'\theta'_f} = -k_d \frac{d\overline{\theta_f}}{dz} \tag{2.20}$$

Here, $\overline{w'\theta'_f}$ is the vertical flux, k_d is the eddy diffusion coefficient and $\overline{\theta_f}$ is the horizontal turbulent mean of the advected items.

We have selected three widely used sub-grid scale schemes, consistent with the ones detailed by Tompkins and Semie (2017): Smagorinsky-2D (Smag2), Smagorinsky-3D (Smag3) (Smagorinsky, 1963), and the Turbulent Kinetic Energy (TKE) scheme (Skamarock et al., 2019). Smag2 exclusively operates horizontally, addressing turbulent kinetic energy generated from horizontal shear, with vertical mixing handled

by the PBL scheme. On the other hand, Smag3 calculates mixing using both full 3D wind shear and static stability, allowing it to represent also vertical mixing, and function independently of a PBL scheme if necessary, although mixing is only local. The TKE scheme (Skamarock et al., 2019) employs a 1.5 closure for horizontal and vertical mixing. Its prognostic equation encompasses essential elements such as shear production, buoyancy, transports, and dissipation terms. These schemes have exhibited divergent outcomes in previous studies, reflecting variations in entrainment and horizontal mixing magnitudes Tompkins and Semie (2017). To capture this variability, we incorporate them into our study with different PBL and microphysical schemes. The subsequent sections provide succinct explanations for each of these three schemes.

Smagorinsky 2D

The Smag2 scheme calculates the horizontal eddy viscosity coefficient (K_h). In Skamarock et al. (2019), they defined K_h as in equation (2.21). Notice that the scheme only treats the shear generation of horizontal winds which means that the vertical mixing is done by the PBL scheme.

$$K_h = C_s^2 I_h^2 [0.25(D_{11} - D_{22})^2 + \overline{D_{12}}^{xy}]^{0.5} \tag{2.21}$$

where $0.25(D_{11} - D_{22})^2$ is the tension term and $\overline{D_{12}}^{xy}$ represents the shear strain deformation, as explained in detail by Smagorinsky (1963). C_s is a constant representing the efficiency of eddy transportation. On the other hand, $I_h = (\Delta x \Delta y)^{0.5}$ depends on the horizontal grid size and is the horizontal eddy viscosity length scale. D represents the wind deformation, in which $D_{11} = 2\frac{du}{dx}$, $D_{22} = 2\frac{dv}{dy}$, $D_{12} = \frac{du}{dy} + \frac{dv}{dx}$ being average over x and y grid points.

Smagorinsky 3D

The Smag3 is a generalization of the 2D scheme (Skamarock et al., 2019) described by Smagorinsky (1963). This 3D scheme used the wind shear and the static stability to specify the values of the eddy diffusion coefficient ($K_{h,v}$ following Skamarock et al. (2019) notation) in both the horizontal and vertical resolutions. The scheme calculates the eddy mixing as presented in equation (2.22)

$$K_{h,v} = C_s^2 I_{h,v}^2 max[0, (D^2 - N^2 P_r^{-1})^{0.5}] \tag{2.22}$$

Where $I_{h,v}$ is the horizontal and vertical eddy mixing length scale that can be calculated with an anisotropic or isotropic option, depending on the grid size. N is the Brunt-Vaisala frequency, P_r accounts for the Prandtl number, and D^2 is represented in equation (2.23).

$$D^2 = \frac{1}{2}[D_{11}^2 + D_{22}^2 + D_{33}^2] + (\overline{D_{12}}^{xy})^2 + (\overline{D_{13}}^{x\eta})^2 + (\overline{D_{23}}^{y\eta})^2 \tag{2.23}$$

In which D_{11}, D_{12}, D_{22} are described in Smag2, $D_{33} = 0.5\frac{dw}{dx}$, $D_{13} = \frac{dw}{dx}$ and $D_{23} = \frac{dw}{dy}$. Note that the D_{13} and D_{23} are averaged in x and η and y and η grid points, respectively. It is important to reckon that in our experiments $\Delta z << \Delta x, \Delta y$, so the turbulence is anisotropic and yields that $I_h = (\Delta x \Delta y)^{0.5}$ and $I_v = \Delta z$.

Turbulent Kinetic Energy

The TKE [$e = 0.5(u'^2 + v'^2 + w'^2)$] 1.5 closure calculates the eddy viscosity coefficient as shown in equation (2.24). Notice that as in the Smag3 scheme the TKE scheme represents the horizontal and vertical mixing and in the anisotropic case the I_h is also equal to the Smag3 scheme. On the other hand, C_k is a constant that typically is chosen to be between 0.15 and 0.25.

$$K_{hv} = C_k I_{h,v} e^{0.5} \tag{2.24}$$

The prognostic equation depends on shear production (S in equation 2.25), buoyancy (depicted as ρ_{buo}, and defined in equation 2.26), and dissipation (D_{diss} in equation 2.27). This means that the 3D wind shear vector (S) completely contributes to mixing in addition to the vertical stability and knowing at each time-step the value of e.

$$S = K_h D_{11}^2 + K_h D_{22}^2 + K_v D_{33}^2 + K_h \overline{D_{12}^2}^{xy} + K_v \overline{D_{13}^2}^{x\eta} + K_v \overline{D_{23}^2}^{y\eta} \tag{2.25}$$

where the variables are already shown in the Smag2 and Smag3 in equations (2.21) and (2.23).

$$\rho_{buo} = -K_v N^2 \tag{2.26}$$

where N^2 is the Brunt-Vaisala frequency and the computation is shown in detail by Skamarock et al. (2019).

$$D_{diss} = -\left[1.9C_k + \frac{max(0, 0.93 - 1.9C_k)l}{\Delta s}\right] \frac{e^{\frac{3}{2}}}{l} \tag{2.27}$$

where $\Delta s = (\Delta x \Delta y \Delta z)^{\frac{1}{3}}$, and $l = min(\Delta s, 0.76 e^{0.5} N^{-1})$, and C_k is the same as in equation (2.24). It should be noted that although a 3D Smagorinski scheme could be derived from the TKE scheme by assuming local equilibrium in the momentum variance equations, this is not the case for the WRF implemented schemes due to different varying constants and closure assumptions between the two approaches. As the Smag3 is not the diagnostic equivalent equation of TKE, the domain and time mean mixing from both schemes will differ (Tompkins and Semie, 2017).

2.3.4 Planetary Boundary Layer

The PBL schemes, designed to characterize vertical subgrid-scale fluxes generated by eddy transport throughout the atmospheric column, operate in a one-dimensional framework. These schemes ascertain fluxes within both the well-mixed boundary layer and the stable layer, consequently influencing the evolution of moisture, temperature, and horizontal momentum across the entire atmospheric column (Skamarock et al., 2019).

Yonsei University Scheme - YSU

The YSU scheme, an advancement of Hong and Pan (1996) and widely employed in numerical simulations (Noh et al., 2003; Hu et al., 2010), stands out as a non-local mixing model featuring counter-gradient flux terms in the boundary layer, as outlined in equation (2.28). Notably, the scheme incorporates an explicit treatment of the entrainment layer at the PBL top, a feature validated by Noh et al. (2003) in Large Eddy Simulation (LES) runs. The explicit handling of the PBL top contributes to

realistic moisture representation, distinguishing it from other PBL schemes (Potvin et al., 2020). This, and the non-local mixing capability enables the YSU scheme to better represent the PBL structure, while local vertical mixing is used throughout the rest of the troposphere. The evolution of any prognostic scalar, C, is given by

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial z}[K_c(\frac{\partial C}{\partial z} - \gamma_c) - (\overline{w'c'})_h(zh^{-1})^3] \tag{2.28}$$

where K_c is the eddy diffusivity coefficient, γ_c is a correction to the local gradient, $(\overline{w'c'})_h$ account for the flux in the inversion layer. h is the PBL height and it is defined as the level of minimum flux at the inversion level, and the component $-(\overline{w'c'})_h(zh^{-1})^3$ is the asymptotic entrainment flux term at the inversion layer which is the explicit treatment component of the entrainment processes in the boundary layer.

Bougeault-Lacarrere Scheme - BouLac

The BouLac PBL scheme (Bougeault and Lacarrere, 1989), derived from Therry and Lacarrère (1989), is a 1.5-order local scheme that incorporates a prognostic TKE equation (see equation 8 in their paper). This parameterization captures second-order moments $(\overline{w'X'})$ using an eddy coefficient approximation expressed as $\overline{w'X'} = -K_c\frac{\partial X}{\partial z}$, where X is substituted with either wind components or kinetic energy. The scheme calculates upward and downward length scales for the PBL's top and bottom, selecting the one yielding the smallest scale. Known for its effectiveness in modeling complex terrains and urban scenarios, the BouLac scheme demonstrates precision in representing PBL height on such terrains (Segura et al., 2021). However, local schemes like BouLac may struggle to adequately mix moisture within convective cores (Hu et al., 2010).

2.3.5 Microphysics scheme

Microphysics schemes in the WRF model handle the atmosphere's water, cloud, and precipitation processes. With a plethora of schemes to choose from, including bulk, bin, and spectral (some with double-moment), these parameterizations vary in their methods of calculating moisture variables and can classify moisture into different classes based on ice and mixed-phase processes. The inclusion of mixed-phase schemes becomes crucial in high-resolution simulations where updrafts are resolved, accounting for interactions among water particles in different states and facilitating the production of new substances like hail or graupel (Skamarock et al., 2019).

WRF Single Moment 6-class Scheme - WSM6

The WSM6 microphysics scheme, built upon the foundations laid by Rutledge and Hobbs (1983) and Dudhia (1989), incorporates ice processes from Lin et al. (1983) with enhancements to the graupel class (Hong et al., 2004; Hong and Lim, 2006). Further improvements by Dudhia et al. (2008) address the representation of mixed-phase particle fall speeds for snow and graupel in sedimentation and accretion processes. Comprising five hydro-meteor classes akin to the Lin et al. (1983) scheme, WSM6 captures interaction processes among these hydrometeors.

The model has a new way to represent the size distribution taking into account three size habits (i.e. single column, a single bullet, and rosettes) as represented by equation (2.29)

$$N_I = c(\rho r_I)^d \tag{2.29}$$

Where, N_I is the number of ice (I) distribution, r_I is the ice mixing ratio, c and d are constants that depend on the ice habit and are described in Hong et al. (2004). The terminal velocity is also modified taking into account the ice habits following equation (2.30)

$$V_I = aD_I^b \tag{2.30}$$

Here, V_I is the terminal velocity of ice as in the previous scheme, and a and b are the terminal velocity constants that in this case depend on the ice habit. In addition, the scheme adds a special intercept parameter for snow, it also changes the autoconversion of cloud water to rain and includes the sedimentation of all types of ice crystals.

Goddard Cumulus Ensemble Scheme - GCE

The GCE scheme, conceived by Tao and Simpson (1993) and rooted in the work of Lin et al. (1983) and Rutledge and Hobbs (1984), undergoes substantial modifications, particularly in ice representation by Tao et al. (2014) and Lang et al. (2014). While maintaining the size distribution, slope, and terminal velocity equations of the Lin et al. (1983) scheme, key alterations include adjustments to the intercept parameter N_{0x} and terminal velocity constants a and b (refer to Table 3.1). The scheme introduces several important changes, elucidated below:

- The addition of a 4th ice class
- A snow density mapping is included
- The scheme permits a greater ice supersaturation, improving the evaporation and sublimation processes and allowing a more realistic change in the transition of graupel to snow
- Cloud ice terminal velocity changes that follow Hong et al. (2004)
- The contact nucleation and the saturation adjustment are improved
- Changes in the snow and graupel densities are shown in Table 3.1

The aforementioned changes produce more realistic reflectivities in the entire troposphere, with a more realistic representation below the freezing level. The above is due to the improvements in the hail, snow, and graupel processes (Lang et al., 2014).

Thompson Scheme - Tho

The Thompson scheme (hereafter Tho) is a hybrid scheme based on Lin et al. (1983) which uses single-moment to calculate cloud water, snow, and graupel, and double-moment for rain and cloud ice, it also has important changes in the representation of snow as describe in Thompson et al. (2008), and improved by Iverson et al. (2021).

In addition, the scheme includes aerosols as initiation nuclei for droplets and ice crystals (Thompson and Eidhammer, 2014), which improve the droplets and especially the ice representation of the clouds. Number and velocity are modelled with equations (2.31) and (2.32).

$$N_x(D) = N_{0x} D_x^{\mu} exp(-\lambda_x D_x) \tag{2.31}$$

$$V_{Dx} = a D_x^b exp(-f_x D_x)(\frac{\rho}{\rho_0})^{0.5} \tag{2.32}$$

Where, μ represents the shape parameter, and f_x is a constant, these two variables depend on the hydro-meteor class. The snow class for this scheme has major changes, two of the more important ones being: they are not assumed to be spherical and their size distribution follows equation (2.33) which represents the sum of a gamma and exponential distribution. Other major important details are presented in Table 3.1.

$$N_S(D_S) = \frac{M_4^4}{M_3^3}[k_0 exp(-\frac{M_2}{M_3}\Lambda_0 D_S) + k_1(\frac{M_2}{M3} D_S)^{\mu_S} exp(-\frac{M_2}{M_3}\Lambda_1 D_S)] \tag{2.33}$$

Where k_0, k_1, Λ_0 and Λ_1 are constants, μ_S is the shape parameter of snow and it is also a constant and $M_n = \int D_S^n N_S(D_S) dD_S$ which accounts for the nth moment of the distribution. Another important characteristic that this scheme has is that the threshold for the autoconversion of ice to snow is smaller than in the other schemes.

Morrison Scheme - Mor

Morrison scheme (hereafter Mor) (Morrison and Gettelman, 2008) is a full double-moment bulk microphysics scheme based on the work of Morrison et al. (2005) and Morrison and Pinto (2006). The model predicts the number and mass mixing ratio for six hydro-meteor species, this allows a more robust and realistic treatment of the particle size distributions that are very important for the interactions between hydrometeors (e.g. cloud water, ice). The scheme uses the gamma distribution taking into account the slope and intercept parameters that are calculated from prognostic equations. The prognostic equation for the particle size distribution and mixing ratios are described in equations (2.34) and (2.35).

$$\frac{\partial N_x}{\partial t} + \frac{1}{\rho}\nabla \cdot [\rho \mathbf{v} N_x] = [\sum_i (\frac{\partial N_x}{\partial t})_i] + D_x(N_x) \tag{2.34}$$

$$\frac{\partial r_x}{\partial t} + \frac{1}{\rho}\nabla \cdot [\rho \mathbf{v} r_x] = [\sum_i (\frac{\partial r_x}{\partial t})_i] + D_x(r_x) \tag{2.35}$$

Here, N_x and r_x are the number concentration and mixing ratio, respectively for every hydro-meteor. D_x is the particle diameter and $\mathbf{v}$ is the wind three-dimensional vector. i accounts for evaporation/sublimation, autoconversion, accretion (for cloud droplets and ice by rain and for cloud droplets and ice by snow), heterogeneous and homogeneous freezing, melting, sedimentation, and detrainment. In addition, i for N_x also accounts for nucleation, and in the case of r_x accounts for condensation/deposition. The scheme also calculates the number (V_{N_x}) and mass (V_{r_x}) terminal fall speed with a prognostic approach, which is described in equations (2.36) and (2.37).

$$V_{N_x} = (\frac{\rho}{\rho_0})^{0.54} a\Gamma(1+b+\mu)[\lambda^b \Gamma(\mu+1)]^{-1} \tag{2.36}$$

$$V_{r_x} = (\frac{\rho}{\rho_0})^{0.54} a\Gamma(4+b+\mu)[\lambda^b \Gamma(\mu+4)]^{-1} \tag{2.37}$$

Where, ρ and ρ_0 are the air density and the reference air density, respectively. Γ is the Euler gamma distribution, μ is the shape parameter, and a and b are empirical constants related as in the other schemes (Table 3.1). In this case, the terminal fall speeds are limited by a threshold for rain and snow and there is also a density correction that increases the realism of the parameterization. Further details of these schemes can be found in Morrison and Gettelman (2008).

2.4 Machine Learning

Machine learning (ML) is a field in computer science that uses diverse algorithms to enable computers to learn patterns and make predictions without explicit programming. It involves extracting knowledge from data, allowing models to generalize and adapt (Awad and Khanna, 2015). These models are broadly categorized into supervised learning, where algorithms predict outcomes from labeled datasets, and unsupervised learning, which uncovers patterns in unlabeled data. Reinforcement learning involves agents optimizing a reward-based system through interaction with an environment (Kleine-Deters et al., 2017). In Chapter 4 we use a Random Forest (RF) algorithm, that is part of the supervised learning type of ML, which is why we describe them here in more detail.

2.4.1 Random Forest

RF is an ensemble learning method that operates by constructing a multitude of decision trees during training and outputs the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. RF algorithms are composed by decision trees, bootstrapped sampling, feature randomization, and decision-making process (Liaw and Wiener, 2002).

The foundational work of Breiman (2001) describes the RF elements as follows: **Decision Trees:** RF consists of a collection of decision trees, each trained on a subset of the training data. Decision trees are built by recursively splitting the data based on the most significant features. **Bootstrapped Sampling:** Each tree in the RF is trained on a random subset of the data, known as bootstrapped samples. This involves sampling with replacement from the original dataset. **Feature Randomization:** RF introduces further randomness by considering only a random subset of features at each split during the construction of a tree. This ensures that the trees are diverse and not highly correlated. **Decision-Making Process:** For classification, the mode (most frequent class) of the individual tree predictions is taken as the final output. For regression, the mean of the individual tree predictions is considered.

The prediction from the ensemble of trees for a given instance is represented in equation (2.38), and is described in detail by Breiman (2001).

$$\hat{Y} = \frac{1}{N} \sum_{i=1}^{N} f_i(X) \tag{2.38}$$

Where $\hat{Y}$ is the predicted output, N is the number of trees in the forest, and $f_i(X)$ is the prediction of the ith tree.

Decision Trees

Decision trees, serving as the foundation of RF, operate by recursively partitioning the input space using the data's features. The goal is to create subsets that are as homogeneous as possible concerning the target variable. The process begins at the root node, representing the entire dataset. It selects the feature that yields the best split, maximizing homogeneity. The dataset is then divided into subsets based on the chosen feature (Breiman, 2001). This recursive process continues for each subset, forming a tree structure with nodes as decision points and branches representing outcomes based on feature values. The procedure persists until reaching a stopping criterion, such as a maximum depth or a minimum number of samples in a node. The final nodes or leaves contain the predicted values or class labels, serving as the output (Hastie et al., 2009).

In this context, decision trees employ classification to assess the impurity or disorder of a set of data points (equation 2.39). Additionally, they utilize regression to quantify the variance of the target variable in a set of data points (equation 2.40), as described in detail by Breiman (2001).

$$Gini(t) = 1 - \sum_{i=1}^{c} p_i^2 \tag{2.39}$$

Where $Gini(t)$ is the Gini impurity as a function of each node t, c represents the number of classes and p_i is the proportion of data points belonging to class i in node t.

$$MSEr = \frac{1}{N_t} \sum_{i \in nodet} (y_i - \overline{y}_t)^2 \tag{2.40}$$

Here $MSEr$ is the mean squared error, N_t is the number of data points in node t, y_i represents the target value of the data point i, and $\overline{y}_t$ is the mean target value in node t.

To summarize, in this chapter, we discussed some of the generalities, parameterizations, the characteristics of the WRF model, and a general explanation of the ML model used in this research. This is with the aim to introduce the numerical model we are using in Chapters 3 and 5, and also the ML technique employed in Chapter 5.

Chapter 3

Sensitivity of Self-Aggregation and Key Role of the Free Convection Distance

3.1 Abstract

Past studies have shown that simulations of radiative convective equilibrium with convective permitting models can result in convective self-aggregation (SA), but that SA is sensitive to model resolution, domain sizes and the details of the convective parameterization. Recently, Biagioli and Tompkins (2023a) used a simple stochastic model to derive a dimensionless parameter to predict SA onset, which was based on the derivation of the maximum free convective distance (d_{clr}) expected in the pre-aggregated state. Our goal is to test and further investigate this hypothesis, namely that d_{clr} can predict SA onset, using an ensemble of twenty-four distinct combinations of horizontal mixing, planetary boundary layer (PBL), and microphysical parameterizations. We confirm that the occurrence of aggregated versus random equilibrium states is closely related to d_{clr} in the initial days of the simulations. We thus conclude that the key impact of parameterization schemes is through their control of the number of convective cores and their relative spacing, d_{clr}, which itself is impacted by cold-pool (CP) properties. SA is more likely when the convective core count is small, while CPs intensify local moisture around clouds, generating dry patches through localized drying of the boundary layer, and modify convective counts via gust-front convergence and collisions. Each parameterization scheme emphasizes a different mechanism. Horizontal mixing mainly affects SA through the determination of convective core size and thus spacing, while the sensitivity to the microphysics is mainly through the rain evaporation and subsequent impact on CPs. Instead, perturbations to the ice cloud microphysics have a surprisingly limited effect. Non-local PBL mixing schemes promote SA by enhancing low-level cloud cover, driven by intensified vertical transport within convective cores, and strong entrainment from the boundary layer top, due to stronger moisture gradients. The low cloud radiative forcing drives a circulation that shortens the transport of moist static energy, disallowing moisture to get to the driest regions, favoring SA. Moreover, the heightened low-level cloud cover induces more significant cooling, requiring stronger subsidence to maintain radiative equilibrium. This enhanced subsidence decreases convective core counts, increasing d_{clr}, and favoring SA.

3.2 Introduction

In the early 1990s, Held et al. (1993) introduced convective self-aggregation (SA) through idealized simulations. SA denotes the tendency of convection to spontaneously organize into clusters within simulations of radiative convective equilibrium (RCE), starting from a uniformly distributed state. Subsequent research explored SA within cloud-resolving models (CRMs) conducted by various authors (Nakajima and Matsuno, 1988; Tompkins and Craig, 1998a; Grabowski and Moncrieff, 2004b; Bretherton et al., 2005; Stephens et al., 2008; Muller and Held, 2012; Wing and Emanuel, 2014; Coppin and Bony, 2015; Holloway and Woolnough, 2016; Wing and Cronin, 2016; Yanase et al., 2020; Shamekh et al., 2020a; Beucler et al., 2020; Huang and Wu, 2022). SA typically originates with the birth of dry regions, expanding to inhibit deep convection development and confining it to specific areas in the simulation domain. This process results in reduced domain-mean water vapor and increased domain-mean outgoing longwave radiation (OLR), significantly influencing the hydrological cycle and the energy budget (Mauritsen and Stevens, 2015; Bao and Sherwood, 2019; Wing et al., 2020; Becker and Wing, 2020; Da Silva et al., 2021a).

While the significance of SA is well-acknowledged, uncertainties persist in understanding its initiation and inhibition mechanisms. Past research (see Wing et al. (2017), Wing (2019), and Muller et al. (2022a) for reviews) has highlighted SA's dependence on domain geometry (Wing and Cronin, 2016), size (Yanase et al., 2020), resolution (Muller and Held, 2012), surface properties (Bretherton et al., 2005), and adopted physical parameterizations (Tompkins and Semie, 2017; Shi and Fan, 2021). The Radiative-Convective Model Intercomparison Project (RCEMIP) aimed to improve SA understanding by conducting multiple simulations with different models under standardized settings (Wing et al., 2018, 2020). Convergence observed among models in several SA aspects emphasize the importance of a comprehensive approach to understanding SA. However, the observed differences between models highlight that uncertainties may arise from, for example, variations in model physics.

To address these uncertainties, previous efforts include sensitivity (Bretherton et al., 2005; Wing and Cronin, 2016; Tompkins and Semie, 2017) and mechanisms denial experiments (Jeevanjee and Romps, 2013; Muller and Bony, 2015; Yang, 2019; Haerter et al., 2019). These investigations offer crucial insights into factors like radiative and evaporative cooling feedbacks (Yanase et al., 2020), cold-pools (CPs) (Jeevanjee and Romps, 2013; Yanase et al., 2020), surface temperature gradients (Bretherton et al., 2005; Hohenegger and Stevens, 2016b; Shamekh et al., 2020a; Tompkins and Semie, 2021), and land surface (Hohenegger and Stevens, 2018b), influencing SA emergence or hindrance. Despite their value, mechanisms denial experiments can yield intricate interpretations, as manipulating one mechanism may inadvertently affect others (Holloway et al., 2017). To address the intricacies of mechanisms denial experiments, researchers have turned to simplified models (Craig and Mack, 2013; Windmiller and Craig, 2019; Biagioli and Tompkins, 2023a). Notably, Biagioli and Tompkins (2023a) developed a stochastic reaction-diffusion model capable of replicating both randomly distributed and aggregated states. In their work, they introduce a dimensionless index that sheds light on the influence of various variables—including domain size, resolution, moisture diffusion, convective core number, maximum free convective distance (d_{clr}), subsidence timescales, and convection lifetimes—on SA. Their findings reveal that a reduced number of convective cores

is associated with larger spacing between them, promoting the development of SA. This is attributed to the increased separation of moisture sources, leading to the formation of dry patches.

Here, we contribute to the continuum of research by emphasizing the role of d_{clr} and its implications for SA. Notably, the anti-correlation between d_{clr} and the number of convective cores, influenced by various model physics parameters, is a key focus. Horizontal and vertical mixing (Cohen and Craig, 2004, 2006; Tompkins and Semie, 2017) and entrainment from the boundary layer top (Hu et al., 2010; Xie et al., 2012; Potvin et al., 2020) also play a role in modifying convective core number, d_{clr}, and consequently impacting SA dynamics. CPs influence SA through mechanisms such as moisture redistribution, infusing humidity into the boundary layer, affecting convective core numbers and d_{clr}, inducing localized drying around the CP center, and controlling gust front convergence. Our study dives into the specific impact of different parameterizations within a CRM, including microphysics, sub-grid scale mixing, and planetary boundary layer (PBL) schemes, on d_{clr} and, by extension, SA. The overarching goal is to uncover the sensitivities and interactions governing SA's behavior, revealing its intricate interplay with atmospheric variables.

Given this context, this study posits the question: What is the sensitivity of SA development to the processes that control convective triggering and thus d_{clr}? and the way those processes are parameterized in the model physics?

The next section describes the CRM setup used in the simulations, as well as the diagnostics. Section 3.4 contains the results, starting from an overview (§ 3.4.1), then we focus on the importance of d_{clr} (§ 3.4.2), by evaluating the dry patches evolution and onset. Subsequently, we make a thorough explanation of the mechanisms that control the number of convective cores and d_{clr} for the sub-grid scale mixing (§ 3.4.3), the microphysics (§ 3.4.4), and the PBL (§ 3.4.5) schemes. Concluding remarks are presented in section 3.5.

3.3 Method

Here, we outline the methodology, detailing the model, its configuration, key parameterizations, and the Lagrangian tracker designed for specific diagnostics. Further methodological discussions and related analyses are embedded in the respective sections.

3.3.1 Overall Model Description

We utilized version 4.2 of the WRF model (Skamarock et al., 2019) for all simulations. The model employs a fifth-order horizontal advection scheme and applies Rayleigh damping at the top of the domain to prevent unphysical wave reflection. For radiation parameterization, the RRTMG scheme (Mlawer et al., 1997; Iacono et al., 2008), is employed to handle longwave (LW) and shortwave (SW) radiation. The surface layer scheme follows Monin-Obukhov similarity theory (Monin and Obukhov, 1954), considering stability functions and roughness lengths. We selected three sub-grid scale schemes, named Smagorinsky-3D (Smag3), Smagorinsky-2D (Smag2), and turbulent kinetic energy (TKE). These three schemes are selected due to their documented varied aggregation behaviors in prior studies, attributed to differences in

Scheme	Species	N_{0x}	a_{vx}	b_{vx}	a_{mx}	b_{mx}	ρ_x
WSM6	Ice[a]	-	-	-	-	-	-
	Snow	$min[10^{11}, 2X10^6 exp(-0.12(T_0 - T))]$	11.72	0.41	$\frac{\pi\rho_S}{6}$	3	100
	Graupel	$4X10^6$	330	0.8	$\frac{\pi\rho_G}{6}$	3	500
	Rain	$8X10^6$	841.9	0.8	$\frac{\pi\rho_R}{6}$	3	1000
GCE	Ice[b]	-	-	-	-	-	-
	Snow	$1.6X10^7$	78.63	0.11	$\frac{\pi\rho_S}{6}$	3	100
	Graupel	$4X10^6$	351.2	0.37	$\frac{\pi\rho_G}{6}$	3	400
	Rain	$8X10^6$	2115	0.8	$\frac{\pi\rho_R}{6}$	3	1000
Tho	Ice	Prognostic	1847.5	1	$\frac{\pi\rho_I}{6}$	3	890
	Snow[c]	-	40	0.55	0.069	2	100
	Graupel	$max[10^4, min(\frac{200}{r_G}, 5X10^6)]$	442	0.89	$\frac{\pi\rho_G}{6}$	3	500
	Rain	$(\frac{N_1-N_2}{2})tanh[\frac{(r_{R0}-r_R)}{4r_{R0}}] + \frac{N_1+N_2}{2}$	4854.4	1	$\frac{\pi\rho_r}{6}$	3	1000
Mor	Ice	Prognostic	700	1	$\frac{\pi\rho_I}{6}$	3	500
	Snow	Prognostic	11.72	0.41	$\frac{\pi\rho_S}{6}$	3	100
	Graupel	Prognostic	19.3	0.37	$\frac{\pi\rho_G}{6}$	3	400
	Rain	Prognostic	841.9	0.8	$\frac{\pi\rho_R}{6}$	3	1000

[a] The size distribution is describe in WSM6 section of the SI. The terminal velocity depends on ice mixing ratio (r_I), distribution of ice (D_I), and the habits of the ice crystals.
[b] Number concentration of ice (N_I) depends on the particle distribution and the ice habits. The terminal velocity is a function of D_I.
[c] The size distribution equation is described in the Tho section of the SI. This scheme also assumes non-spherical snowflakes.

TABLE 3.1: Characteristics of the microphysics schemes, and their principal differences: particle size distribution, number intercept parameter (N_{0x}), the velocity-diameter constants (a_{vx} and b_{vx}), the mass-diameter relations constants (a_{mx} and b_{mx}) and the particle density (ρ_x). These values are provided directly from the WRF model modules.

entrainment magnitudes (Tompkins and Semie, 2017).

We select two PBL schemes: the Yonsei University scheme (YSU) and the Bougeault-Lacarrere Scheme (BouLac). YSU, a non-local PBL scheme (Hong and Pan, 1996; Hong et al., 2006) known for its strong entrainment from the boundary layer top (Hu et al., 2010), which may influence SA. On the other hand, BouLac, a 1.5-order local scheme (Therry and Lacarrère, 1989; Bougeault and Lacarrere, 1989), is chosen because local schemes often exhibit insufficient moisture mixing within convective cores.

We conducted simulations employing the four microphysical schemes described in chapter 2: WRF Single Moment 6-Class (WSM6) (Rutledge and Hobbs, 1983; Dudhia, 1989), Goddard Cumulus Ensemble (GCE) (Tao and Simpson, 1993), Thompson (Tho) (Thompson et al., 2008), and Morrison (Mor) schemes(Morrison et al., 2005; Morrison and Pinto, 2006; Morrison and Gettelman, 2008). Table 3.1 outlines the principal characteristics of the microphysics schemes. Chapter 2 describes these parameterizations in detail.

3.3.2 Simulation Setup

The model domain is 510 km^2 in the horizontal with periodic boundary conditions, a fixed sea surface temperature of 301 K as in Tompkins and Semie (2017) and with a 3 seconds time step. The domain size is selected since smaller domains could disallow SA from happening (Muller and Held, 2012; Yanase et al., 2020; Biagioli and Tompkins, 2023a). It also uses a 2 km^2 horizontal resolution to explicitly resolve convection (Tompkins and Craig, 1998a; Bretherton et al., 2005; Tompkins and Semie, 2021), and has a stretched vertical grid (62 levels) as in Tompkins and Semie (2017). The diurnal cycle is active, there are no Coriolis effects, and no mean wind was imposed.

We perform several experiments to study the impact of the sub-grid scale mixing, PBL, and microphysics schemes on convective SA. The experiments last for 45 days with hourly outputs. They consist of the combination of 3 sub-grid scale schemes, with 2 PBL schemes and 4 microphysics schemes, which means that 24 simulations were performed. The nomenclature of the simulations is as follows: First the microphysics scheme used (e.g., Tho, GCE), second the sub-grid scale mixing scheme (i.e., "TKE" for TKE, "SM3" for Smag3 and "SM2" for Smag2), and third the PBL ends the nomenclature, either YSU or BL for the BouLac scheme. For instance, if the simulation uses Thompson microphysics with Smag2 and YSU, the simulation name would be Tho-SM2-YSU, but if the horizontal mixing is Smag3 with BouLac the name would be Tho-SM3-BL.

The initial conditions in our study were adopted from Jordan (1958) and Tompkins and Semie (2017). In addition to our primary experiments (Table 3.2), we conducted six sensitivity simulations. The first, denoted Tho-SM2-BHE, involved a 50% reduction in rain evaporation within the Tho-SM2-BL configuration. In the second, Tho-SM2-YDE, we doubled the rain evaporation within the Tho-SM2-YSU setup. The third, Tho-SM2-YHI, entailed a 50% reduction in the fall speed of ice/snow within the Tho-SM2-YSU configuration. The fourth, Tho-SM2-BDI, involved doubling the fall speed of ice/snow within the Tho-SM2-BL configuration. Finally the fifth experiment starts from the last timestep from the BouLac runs (i.e., Tho-SM2-BL, Tho-TKE-BL, GCE-SM2-BL, and GCE-TKE-BL), but change the PBL scheme to YSU, these runs are named: Tho-SM2-BLY, Tho-TKE-BLY, GCE-SM2-BLY, and GCE-TKE-BLY. Table 3.3 summarizes all sensitivity experiments in this research.

3.3.3 Tracking algorithm

To investigate further the role of CPs in the development of SA, we track and analyze their characteristics using a Lagrangian tracker which is described in this section. The tracker will also be used to follow the evolution of dry patches. The tracker is based on the watershed algorithm following Casallas et al. (2023c), that developed a tracker using the Scikit-image library version 0.16 (Virtanen et al., 2020) in Python 3.6. The algorithm segments the image, with each segment corresponding to an object labeled throughout the entire simulation. The algorithm treats each time step as a "topographic" surface based on a threshold (e.g., height, temperature gradient, total column water vapor – TCWV), where values below (or above) the threshold become the boundaries.

Name	Microphysics	Sub-Grid Scale Mixing	PBL
WSM3YSU	WSM6	Smagorinski 3D	YSU
WSMTYSU	WSM6	TKE	YSU
WSMYSU	WSM6	Smagorinski 2D	YSU
WSM3BL	WSM6	Smagorinski 3D	BouLac
WSMTBL	WSM6	TKE	BouLac
WSMBL	WSM6	Smagorinski 2D	BouLac
GCE3YSU	GCE	Smagorinski 3D	YSU
GCETYSU	GCE	TKE	YSU
GCEYSU	GCE	Smagorinski 2D	YSU
GCE3BL	GCE	Smagorinski 3D	BouLac
GCETBL	GCE	TKE	BouLac
Tho3YSU	Tho	Smagorinski 3D	YSU
ThoTYSU	Tho	TKE	YSU
ThoYSU	Tho	Smagorinski 2D	YSU
Tho3BL	Tho	Smagorinski 3D	BouLac
ThoTBL	Tho	TKE	BouLac
ThoBL	Tho	Smagorinski 2D	BouLac
Mor3YSU	Mor	Smagorinski 3D	YSU
MorTYSU	Mor	TKE	YSU
MorYSU	Mor	Smagorinski 2D	YSU
Mor3BL	Mor	Smagorinski 3D	BouLac
MorTBL	Mor	TKE	BouLac
MorBL	Mor	Smagorinski 2D	BouLac

TABLE 3.2: Overview of the simulations and their sub-grid scale, microphysics and PBL schemes. The sensitivity experiments are also shown with a respective note that describes the changes made to the experiment.

Name	Microphysics	Sub-Grid Scale Mixing	PBL	Notes
GCETBLY	GCE	TKE	BouLac-YSU	Ten days with YSU starting from the last time step of GCETBL
GCEBLY	GCE	Smagorinski 2D	BouLac-YSU	Ten days with YSU starting from the last time step of GCEBL
ThoYDE	Tho	Smagorinski 2D	YSU	Double evaporation of rain
ThoYHI	Tho	Smagorinski 2D	YSU	Half ice/snow fall speed
ThoYTB	Tho	Smagorinski 2D	YSU	25% increase in the PBL height
ThoTBLY	Tho	TKE	BouLac-YSU	Ten days with YSU starting from the last time step of ThoTBL
ThoBHE	Tho	Smagorinski 2D	BouLac	Half evaporation of rain
ThoBDI	Tho	Smagorinski 2D	BouLac	Double ice/snow fall speed
ThoBHB	Tho	Smagorinski 2D	YSU	Half PBL height
ThoBLY	Tho	Smagorinski 2D	BouLac-YSU	Ten days with YSU starting from the last time step of ThoBL

TABLE 3.3: Overview of the sensitivity simulations and their sub-grid scale, microphysics and PBL schemes, with a description of the changes made to the experiment.

The watershed algorithm is employed and tailored for tracking CPs and extracting their characteristics, such as intensity. CP intensity is calculated (Grant and van den Heever, 2016, 2018; Abramian et al., 2022a) by integrating the negative buoyancy as $C^2 = 2\int_0^H(-g\theta_v^{'}/\bar{\theta_v})dz$ (Benjamin, 1968), where C is the CP intensity, H is the height of the CP, g is the gravitational acceleration, $\theta_v^{'}$ is the virtual temperature perturbation and $\bar{\theta_v}$ is the virtual temperature averaged over the domain. An evaluation of the algorithm's efficacy, bench-marked against the conventional criterion of selecting pixels with buoyancy $\leq -0.005\ m^2\ s^{-1}$ (Tompkins, 2001b; Grant and van den Heever, 2016, 2018; Abramian et al., 2022a), reveals a notable similarity (R^2 = 0.93 $\pm$ 0.05 for the vertical velocity W at 100m (W_{100m}) between days 2 and 5 of the simulations). Consequently, we focus on the tracker results, as it demonstrates superior efficacy in capturing aged CP fronts. The CP top is defined as the highest altitude where the buoyancy threshold remains valid (Grant and van den Heever, 2016), constrained to a maximum of 2 km (16th model level) (Grant and van den Heever, 2016, 2018; Abramian et al., 2022a).

The tracker algorithm is also modified and used to follow dry patches to understand their properties and onset. For the CP tracking, the 2 m temperature gradient is calculated, and the 60th percentile of this variable is used as a threshold. To find this threshold we perform a randomized search with the idea of finding a percentile that allows identifying new and old CPs, including their fronts. See Figure A.1a-b-c for an example. Regarding the dry-patch tracking, we used the TCWV field and select 44 mm as a threshold. To select this value, we perform a randomized search to identify the threshold that avoids capturing CP centers but also selects the growing dry zones. See Figure A.1d-e-f for an example.

In terms of temporal tracking, the algorithm is adapted from previous Lagrangian trackers (see Muller et al. 2022b for tracking examples) designed to track the motion of storms. The approach involves following the evolution of each object, assigning one label per object in the first time step. In subsequent time steps, the tracker decides based on the Euclidean distance (of their centers and boundaries) and overlapping of the objects whether the object is new, in which case a new label is created, or if the object identified in the new time step is the same as a previous one, in which case the label is maintained.

3.3.4 Dry to Moist Regions Circulation

Previous studies have highlighted the significance of low clouds in dry regions for the development of SA (Muller and Bony, 2015). Since Section 3.4.5 will elucidate the importance of low-level clouds, we further analyze their role by computing the stream function Ψ as introduced by Bretherton et al. (2005). Ψ serves as a useful metric for quantifying the circulation between dry and moist regions. It is computed by ordering the columns based on TCWV-percentile and calculating the corresponding mass flux for each column, as outlined in Equation (3.3.4).

$$\Psi(i,z) = \Psi(i-1,z) + \sum_{\int TCWV \in (\int TCWV_{i-1}, \int TCWV_i)} w(z)\overline{\rho}(z) \tag{3.1}$$

With $\Psi(0,z) = 0$ for every z, where $\overline{\rho}$ is the reference density profile used in the anelastic governing equations, w represents the vertical velocity, and i denotes the TCWV column index. $\Psi(i,z)$, as explained by Muller and Held (2012), signifies

the total vertical mass flux across all columns where $\int TCWV \leq \int TCWV_i$. This circulation must be interpreted independently of physical space due to the TCWV ordering but provides valuable insights into the exchange between dry and moist regions (Shamekh et al., 2020a; Yanase et al., 2020).

3.4 Results

In this section, we present an overview of the study's results, highlighting key variables that influence the development of convective SA or maintain a random configuration in simulation setups. Our analysis underscores the sensitivity of SA to sub-grid scale mixing, PBL, and microphysics schemes.

3.4.1 Overview of the Main Results

Our findings reveal a sensitivity of SA to the employed parameterizations, as depicted in Figure 3.1, showcasing the TCWV across all 24 experiments conducted. Simulations employing the Smag3 scheme exhibit rapid aggregation, in contrast to Smag2 and TKE, where SA emerges primarily when YSU is employed as the PBL scheme. This suggests a significant role for sub-grid scale mixing and the PBL in triggering SA. Figure 3.2 provides a concise summary, presenting the SA index calculated as the mean between the last 5 days of each experiment of the difference of the ninetieth (90th) and tenth (10th) percentiles (hereafter NMTP) of TCWV (Bretherton et al., 2005; Müller and Hohenegger, 2020b). The colors in the figure represent the horizontal mixing coefficient (K_h).

Figure 3.2 highlights key insights. Firstly, the significant impact of horizontal mixing, since regardless of variations in the PBL and microphysics configurations all the Smag3 simulations develop organization. This underscores the essential role of Smag3 in influencing SA. Secondly, surprisingly the microphysics have a fairly limited impact on SA, simulations using the Mor scheme exhibit SA, while those employing the WSM scheme consistently display random organization. This underscores the substantial influence of microphysics as an independent factor impacting SA, regardless of the chosen PBL scheme. Thirdly, simulations employing nonlocal PBL schemes (YSU) tend to promote SA, whereas those utilizing local schemes (BouLac) tend to inhibit SA—except in cases where microphysics appears to play a dominant role in driving SA.

We hypothesize that the sensitivity of SA to various parameterizations can be clarified by considering d_{clr} (refer to $d_{max,clr}$ in Biagioli and Tompkins (2023a) Figure 8) given its established inverse correlation with the number of convective cores (Biagioli and Tompkins, 2023a) before SA develops, which is why we adopt their framework. These cores are intricately governed by multiple physical mechanisms embedded within the parameterizations. To test this hypothesis, we examine the TCWV-NMTP (mean of the last 5 simulations days) as a function of the mean d_{clr} calculated between days 2 and 5 (the first 48h are for model spin-up), a period preceding the onset of SA. Additionally, the size of the markers on the plot corresponds to the mean CP intensity between days 2 and 5 (Figure 3.3). Other periods (i.e., 2-4, 2-7, and 2-10 days) were used to calculate the mean d_{clr} and CP intensity, and the results yield the same conclusions.

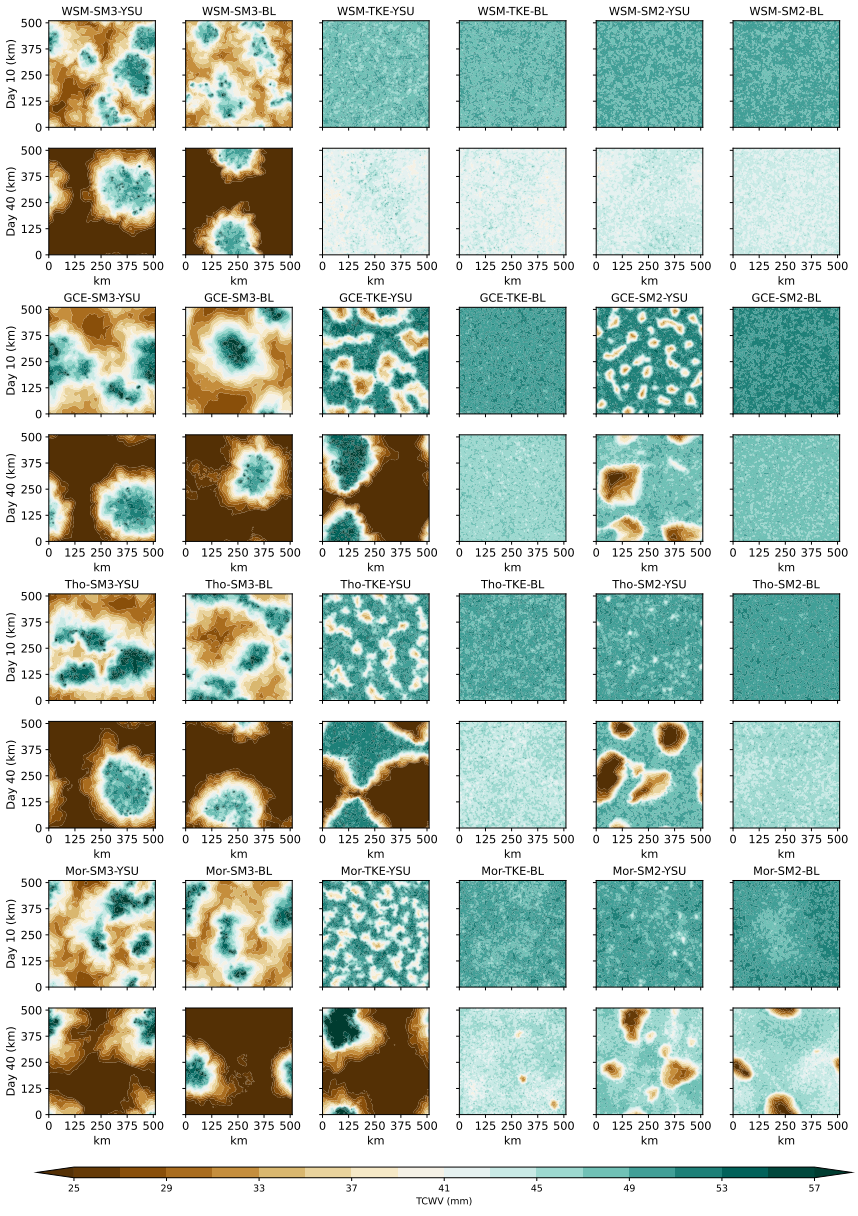

FIGURE 3.1: Snapshot of TCWV of all the 24 experiments performed (see Table 3.2). The columns represent the different model configuration. Each microphysics scheme is represented by two rows with the first depicting a snapshot of day 10, and the second showing a snapshot of day 40.

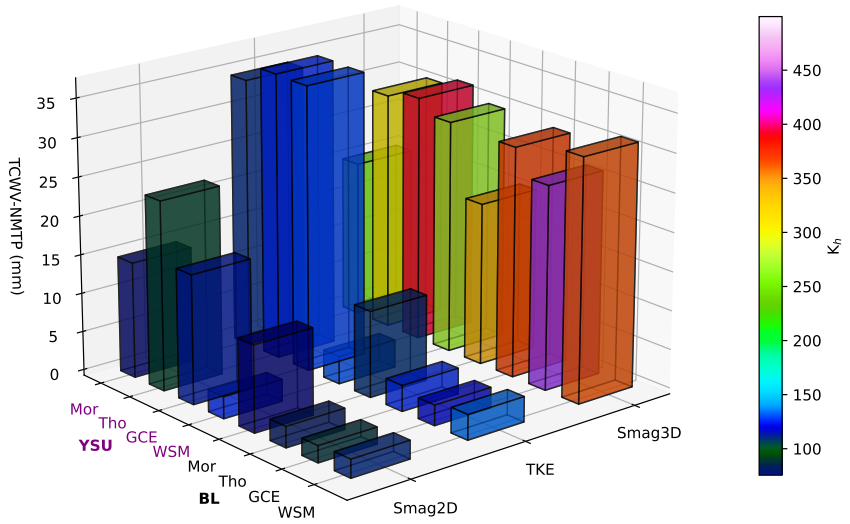

FIGURE 3.2: The height of the bars represents the mean TCWV-NMTP (calculated as percentile 90 - percentile 10 of TCWV and used as SA metric) between days 40 to 45 of each simulation. The colors show the mean horizontal eddy viscosity K_h inside the convective cores during the entire simulation. The x-axis shows the Sub-Grid Scale mixing scheme. The y-axis includes the microphysical parameterization and the PBL schemes. Notice that, the purple font represents the simulations that use YSU and the black font the experiments using BouLac.

Figure 3.3 reveals that simulations developing SA (blue-green colors) consistently exhibit higher values of d_{clr} in the pre-onset phase, irrespective of the specific parameterizations employed. This observation resonates with the theoretical framework of Biagioli and Tompkins (2023a), positing that greater spacing between cores generate drier regions between the sources, setting the stage for the emergence of SA. Simulations employing the Smag3 horizontal mixing scheme exhibit the largest d_{clr}. A comparison between GCE-SM2-YSU, Tho-SM2-YSU, and Tho-TKE-BL simulations, with comparable d_{clr} values, highlights the contrasting effect of CP intensity on SA. While Tho-TKE-BL's more intense CPs hinder organization, this is not the case for the other two simulations.

3.4.2 The Importance of d_{clr}

Here, we focus on studying the relevance of d_{clr} by analyzing dry patch evolution using a Lagrangian tracker approach (detailed in § 3.3) and conducting confirming experiments.

Dry Patch Growth

We first focus on the growth and maintenance of dry patches. We employ the Lagrangian tracker described in the methods section, focusing on Smag2 simulations

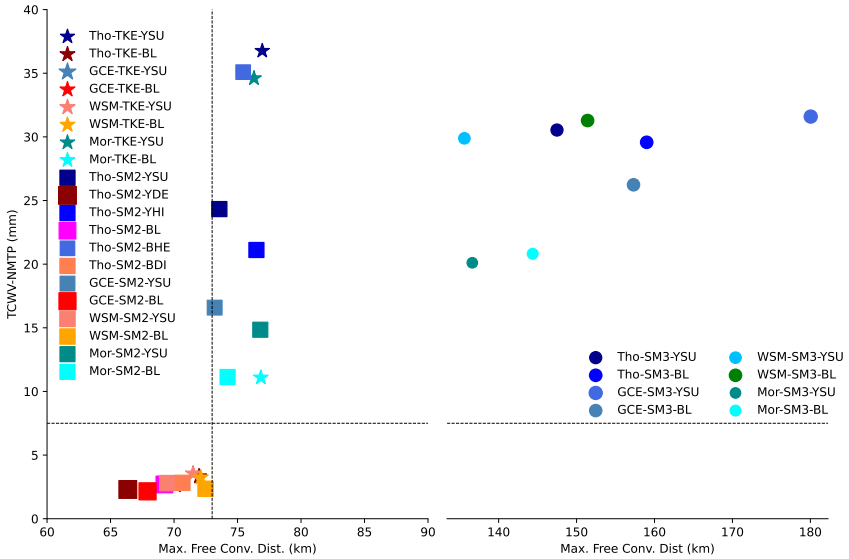

FIGURE 3.3: This figure shows the mean TCWV-NMTP (calculated as in Figure 3.2) between day 40 and day 45 for each simulation as a function of the d_{clr}, which is averaged from day 2 to day 5 for all the experiments. The circles represent the Smag3 experiments, stars account for the TKE runs and the squares show the Smag2 simulations. The blue-green colors represent the experiments that develop SA, and red-yellow colors account for runs with random convection. The size of the markers depends on the CPs intensity averaged from day 2 to day 5, so larger markers imply stronger CPs. Notice that for plotting purposes the x-axis is divided into two zones.

(the results are consistent for simulations using TKE and Smag3, not shown). Importantly, dry patches do not emerge immediately at the beginning of the simulation. As our interest lies in understanding the atmospheric conditions preceding and during dry patch development, we identify the atmospheric behavior before the dry patch onset. To achieve this, we implement a back-propagation technique: Once a dry patch reaches an area of 16 km^2, we select its center and create a buffer with a radius of 20 km^2, centered around the core of the dry patch (varying buffer sizes yield similar results). We then save the coordinates of the pixels within this buffer for subsequent analysis of the atmospheric properties in the locations where dry patches form before their initial appearance.

At the initial stages of the simulation, a marginally positive pressure anomaly triggers its development (Figures 3.4, A.2, A.3, and A.4). This anomaly is prominently observed in regions characterized by large d_{clr} values, which are typically drier areas. This positive pressure anomaly gradually strengthens (Yang, 2018; Yao et al., 2022) and induces a divergent flow, as detailed by Shamekh et al. (2020a), facilitating the export of moist static energy (MSE) up-gradient. This implies that dry regions tend to become drier, while moist regions become moister, thereby reinforcing the high-pressure anomaly and its associated divergent flow. This feedback mechanism, in turn, contributes to the expansion of the dry zone. As illustrated in Figure 3.4d, the initial signs of slightly positive anomalies emerge around day 2, after which the anomaly intensifies and extends throughout the entire boundary layer.

Here, we employ the framework developed by Shamekh et al. (2020a) to gain insights into the emergence of dry patches and the concurrent high-pressure anomaly. However, we extend the analysis to encompass the atmospheric characteristics at the beginning of the simulations, not exclusively when the dry patches manifest. Assuming hydrostatic balance, the surface pressure anomaly is related to the column density anomaly, itself related to the virtual temperature anomaly. Shamekh et al. (2020a) decompose the virtual potential temperature ($\theta_v^{'}/\overline{\theta_v}$, where the quote ($'$) represent a perturbation for each pixel, and the over line ($\overline{\ }$) is the domain mean) anomaly into two distinctive components within the boundary layer: the moist component [$0.61q_v^{'}/(1+0.61\overline{q_v})$, where q_v is the water vapor mixing ratio] and the temperature component ($T^{'}/\overline{T}$). The pressure anomaly is intricately linked to density anomalies within the boundary layer in tropical regions where the Coriolis parameter is relatively small, and mid and upper tropospheric density anomalies remain negligible (Sobel et al., 2001).

During the initial 20 days of the simulation, the surface moisture anomaly (Figure 3.4b) exhibits a negative trend (drying), becoming more pronounced over time. The surface temperature anomaly (Figure 3.4c) displays a positive tendency (stronger in magnitude compare to the moist anomaly) which grows slower in magnitude compare to the drying anomaly. This dynamic implies that the $\theta_v^{'}/\overline{\theta_v}$ anomaly initiates as a positive anomaly at the surface (Figure 3.4a). However, as the surface negative moist anomaly intensifies, a surface negative $\theta_v^{'}/\overline{\theta_v}$ anomaly materializes around day $\approx$7 and persists. This inversion of the $\theta_v^{'}/\overline{\theta_v}$ anomaly from positive to negative occurs due to the rapid growth of the drying surface anomaly, outpacing the increase of the positive surface temperature anomaly. This phenomenon could be attributed to the divergent flow and the consequential export of MSE from dry to moist regions. The transition to a surface negative $\theta_v^{'}/\overline{\theta_v}$ anomaly favors SA and leads to an accelerated rise in the pressure anomaly, as evident from day ≈ 7 in Figure 3.4d. Notably, these findings align with the results obtained from experiments utilizing alternative microphysics schemes, as depicted in Figures A.2, A.3, and A.4 in Appendix A.

According to Shamekh et al. (2020a), the growth of a drying anomaly and the subsequent development of a high-pressure anomaly in dry areas are primarily driven by radiative cooling (Q_{rad}), since enhanced cooling, through subsidence, would cool and dry the local environment. Figure 3.5 (see Figures A.5, A.6, and A.7 for other microphysics results) illustrates the temporal progression of Q_{rad}, including all its components. The net Q_{rad} (Figure 3.5g) exhibits strong negative values at the surface, within the altitude range of 2 km to 5 km, and from 6 km to 10 km. These regions align well with the heights of high-pressure anomalies depicted in Figure 3.4d.

The cooling at the surface is a consequence of the combined effects of both the LW cloud radiative effect –CRE– (Figure 3.5a) and LW clear-sky radiative cooling (Figure 3.5b). This combined cooling induces a surface divergent flow, directing air from the dry patches toward the moister zones. These findings are consistent with previous research (Yang, 2019; Naumann et al., 2019; Shamekh et al., 2020a; Muller et al., 2022a), as well as with the conclusions drawn by Muller and Bony (2015) regarding the significance of radiative cooling from clouds. On the other hand, the net Q_{rad} between altitudes of 2 km and 5 km is primarily driven by the LW clear-sky component, regardless of the warming produced by the SW CRE component (Figure 3.5e). The radiative cooling help to generate a moisture anomaly through subsidence (Sobel et al., 2001; Shamekh et al., 2020a), contributing to the formation

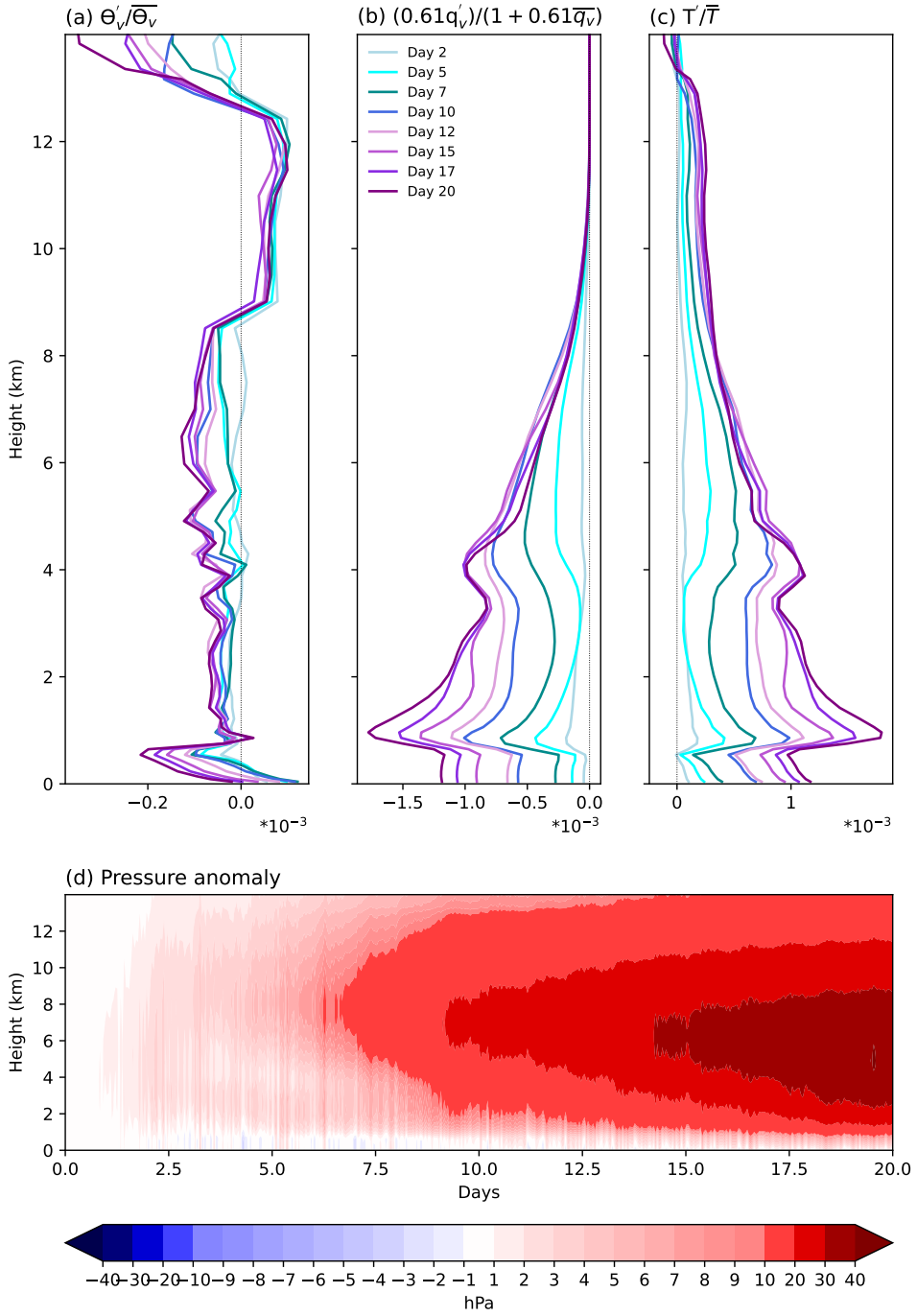

FIGURE 3.4: The plot shows daily dry patch composite means of the virtual potential temperature anomalies (a), and the contributions to this variable from the qv anomaly (b) and the temperature anomaly (c) of the Tho-SM2-YSU experiment. The results are similar for the other runs that develop organization. (d) Hourly composite of pressure anomaly from the dry patches as they evolve with time.

of a high-pressure anomaly at this height (Figure 3.4d), which initiates around day 3.

As for the net Q_{rad} produced between altitudes of 6 km to 11 km, it is mainly driven by the LW CRE component, although the SW clear-sky component plays a smaller role. The LW CRE component exhibits a cooling effect at this height, increasing with altitude due to cloud-related processes and warming at the PBL tops, with its most significant values observed between 8 km and 10 km, coinciding with the location of high-pressure anomaly (Figure 3.4d) responsible for initiating a divergent flow at the same altitude around day 5. This divergent flow leads to the export of MSE from dry to moist regions. These three divergent branches described here (surface; 2 km-5 km; 6 km-10 km) resemble the circulations reported by Muller and Held (2012), Shamekh et al. (2020a), and Cerlini et al. (2023) when convection is clustered due to the upgradient MSE transport. It is important to notice that the results described here also hold for the other experiments, as shown in Figures A.5, A.6, and A.7 in Appendix A.

Fundamentally, in regions with ample d_{clr}, subsidence prevails, inducing a negative (i.e., more cooling) LW Q_{rad} anomaly at three distinct altitudes. This increase cooling effect is further reinforced by a negative SW clear-sky Q_{rad} anomaly in the mid-troposphere. These combined factors result in the establishment of a divergent flow, associated with a positive pressure anomaly that transfers MSE from dry to moist regions, driving a divergent feedback. This confluence of a d_{clr} region, large subsidence, radiative cooling, moisture anomaly, and pressure anomaly collectively contributes to the formation of a dry patch that leads to SA. This section analysis of dry patch onset and evolution reveals that a CP originating at the periphery of a large d_{clr} region serves as a seed to initiate the dry patch. This insight is essential, as it allows us to conduct target experiments involving CPs situated in different regions to validate the role of d_{clr} and the ability of a CP to trigger SA. This is the topic of next section.

Dry Patch Onset

We argue that a CP can act as a seed for a dry patch when produced under specific conditions (i.e., large d_{clr} and in an already dry zone). To test this hypothesis, we introduce artificial CPs into the domain and monitor their evolution to identify instances where they generate dry patches that lead to SA.

Regarding its initial conditions and parameterizations, the experiment starts from the last output hour (day 45) of the Tho-SM2-BL experiment, which does not exhibit SA. It runs for 10 days with an hourly output and a 3-second time step. This duration is sufficient for the initiation of one or multiple dry patches (Figure 3.1). To ensure the development of SA and maintain similarity to Tho-SM2-BL, the Tho scheme is used in combination with Smag2 and YSU as the PBL, as this configuration demonstrates SA development (Figures 3.3 and 3.2). The domain size, resolution, LW and SW radiation, and surface layer parameterizations remain consistent with all other experiments (see § 3.3 for details). To establish a control, we conduct experiments with the same initial conditions and parameterizations but without any CP additions (hereafter CP-control) to confirm that CPs indeed contribute to the development of dry patches.

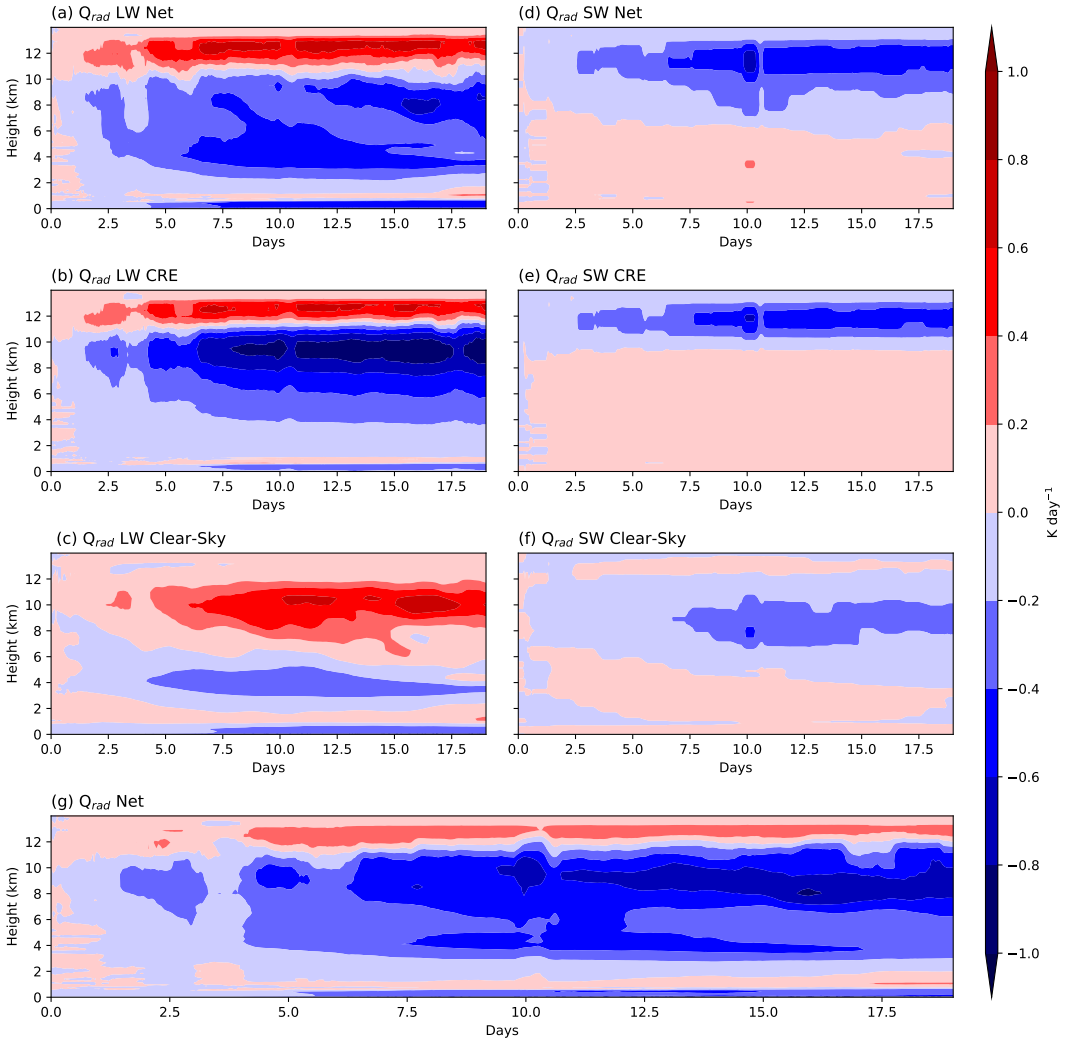

FIGURE 3.5: The plot shows dry patch composite mean evolution of the Q_{rad} (a) LW Net (b) LW CRE, (c) LW Clear-Sky, (d) SW Net, (e) SW CRE, (f) SW Clear-Sky, and (f) the full Q_{rad} cooling for Tho-SM2-YSU. All fields are subject to a 24h-running mean.

In terms of the CP characteristics, we based on the work of Grant and van den Heever (2018). We introduce multiple cold bubbles in the domain. Each bubble has a $\theta_v^{'}$ of -5 K with a radius of 2 km^2 and a height of 2 km (16th model level). To make a more realistic CP we also include artificial rain by adding a perturbation of 1.5 $\times 10^{-5}$ kg kg^{-1} to the rain mixing ratio at $\approx$ 2.2 km of altitude (17th model level). This rain mixing ratio is selected since it is the mean value of the domain at that height. Experiments with other bubble radius/height, different artificial rain additions and perturbations lead to the same conclusions, except for perturbations $<$ -1.5K that are not enough to develop the dry patch. As the idea is to probe that the d_{clr} is important, as well as the dryness of that area, we include CPs in different domain regions depending on the dryness and on the free convective area as shown in Table 3.4.

The CP-control simulation (not shown) exhibits SA, with dry patches emerging from day 1. These patches are located to the right of the domain and are in close

Experiments ID	Moist percentile	Free convection area size (km^2)
Zone 1	30th	276
Zone 2	35th	624
Zone 3	50th	176
Zone 4	65th	64
Zone 5	70th	16

TABLE 3.4: Characteristics of the location of the artificial cold bubble. Notice that the cold bubbles are included in places with different amounts of moisture. The bubbles are also located at the center of the free convection area. To have more robust results, the CPs are also located at the edges of the free convection area, but no significant differences were found.

proximity, leading to their merger by day 5 and the formation of a sustained large dry patch until the simulation's end. For this reason, the artificial CPs are introduced at the center-left part of the domain, where no dry patches are observed in the CP-control. To account for the stochastic nature of SA development highlighted by Tompkins and Semie (2021), three additional control runs are conducted to confirm that the patches consistently emerge on the right side of the domain (not shown). This was expected since the patches grew in the largest free convection areas of the domain, which as described in the previous sections are the areas more favorable for SA to develop.

In the experiment with added CPs, only two CPs contribute to the development of sustained dry patches throughout the entire simulation (Figure 3.6). These two CPs were the ones placed in the driest regions with the largest d_{clr}. Conversely, the other CPs vanish before the fourth day of the simulation. Specifically, the CPs introduced in zones 4 and 5, situated in moister areas with small d_{clr}, generate a dry anomaly that dissipates within the first simulation day, aligning with expectations. The CP in zone 3 produces a dry patch that persists until day four, but the appearance of convection in its vicinity leads to humidity mixing and the eventual disappearance of the dry zone. In contrast, the CPs in zones 1 and 2 successfully generate growing dry patches, as evident in Figure 3.6. The dry patch originating from zone 1 experiences a significant expansion around day ≈ 7 due to its merger with another dry zone that develops on the right side of the domain, where a dry patch forms in the CP-control run.

The selective growth and persistence of dry patches initiated by specific CPs can be attributed to the fact that pre-existing dry zones are located in the largest d_{clr} regions. In the presence of a CP, the dry anomaly in the zone intensifies as the inherently dry core of the CP augments the existing negative moisture anomaly. The moist surge from the CP expands toward the boundaries rather than the core of the dry zone. This phenomenon may result from the development of a relatively weak high-pressure system and variations in air density between the dry CP center and the moist peripheries of the patch (Ross et al., 2004). Consequently, the influx of moisture into the dry zone remains modest, and the dry CP center accentuates the dryness of an already dry region, amplifying the nascent positive pressure anomaly and initiating the divergent feedback. In summary, the experiments presented here highlight that a CP can initiate a dry patch (serving as seeds), but this occurs effectively only when the CP develops in an already dry region featuring a sufficiently

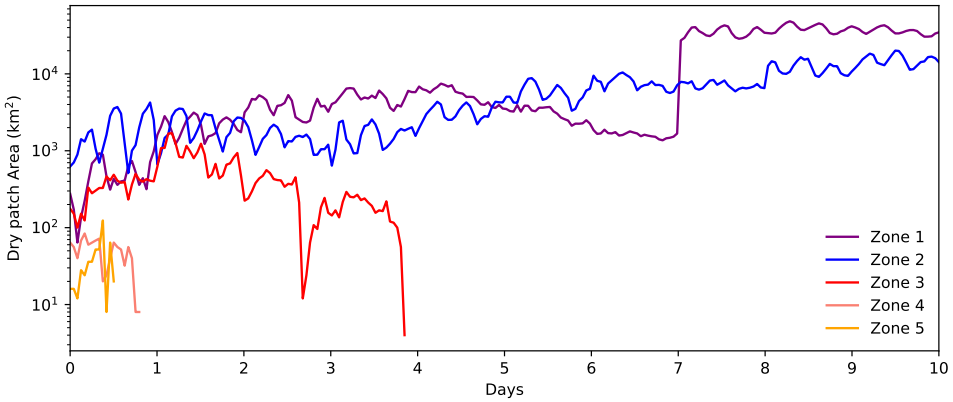

FIGURE 3.6: Evolution from day 0 to day 10 of the zones described in Table 3.4 where the artificial CPs were introduced. Notice that the disappearance of the line implies that the dry zone was moistened. The blue-purple colors represent the experiments that develop SA, and red-yellow colors account for runs with random convection.

large d_{clr}.

3.4.3 Sensitivity to Sub-Grid scale mixing

Figure 3.3 suggests a robust sensitivity to the sub-grid scale mixing scheme, with Smag3 simulations consistently developing SA, regardless of the chosen PBL or MP schemes. This trend is confirmed by Figure 3.2, illustrating that the K_h within convective cores (identified where vertical velocity W $\geq$ 1 m s^{-1} at 750 hPa, following Tompkins and Semie (2017)) is one order of magnitude larger in Smag3 simulations compared to TKE or Smag2 schemes. This aligns with Tompkins and Semie (2017) and Cerlini et al. (2023), emphasizing that high K_h within convection induces significant entrainment, crucial for SA development. Entrainment can modify convective mass-flux/heating, influencing SA (Bretherton et al., 2005; Stephens et al., 2008; Tompkins and Semie, 2017; Yang, 2019). Additionally, it affects the number of convective cores and, consequently, d_{clr}.

The findings of Tompkins and Semie (2017) suggest entrainment's potential to induce atmospheric dryness anomalies, triggering SA. Our reinterpretation underscores the intricate links among entrainment, the number of convective cores, and d_{clr}. In this context, Figure 3.7 elucidates these connections during the initial simulation phase (days 2 to 5), focusing on SA onset. The analysis excludes the first 48 hours for model spin-up, and after day 5, some simulations (e.g., Smag3 experiments) exhibit significant SA (Figure 3.1). Alternative periods (2-4, 2-7, and 2-10 days) yield consistent conclusions.

Smag3 simulations exhibit pronounced entrainment through enhanced K_h within convective cores (Figure 3.7a), resulting in a notable reduction in core count (Figure 3.7c). The larger core areas (around 17 $\pm$ 1.5 km^2) in Smag3, three to four times larger than others (around 3.5 $\pm$ 1 km^2), can be linked to the requirement for convective mass flux equilibrium in RCE scenarios (Cohen and Craig, 2004, 2006). This leads to a reduced core count necessary to offset Q_{rad}, thereby contributing to a decrease in

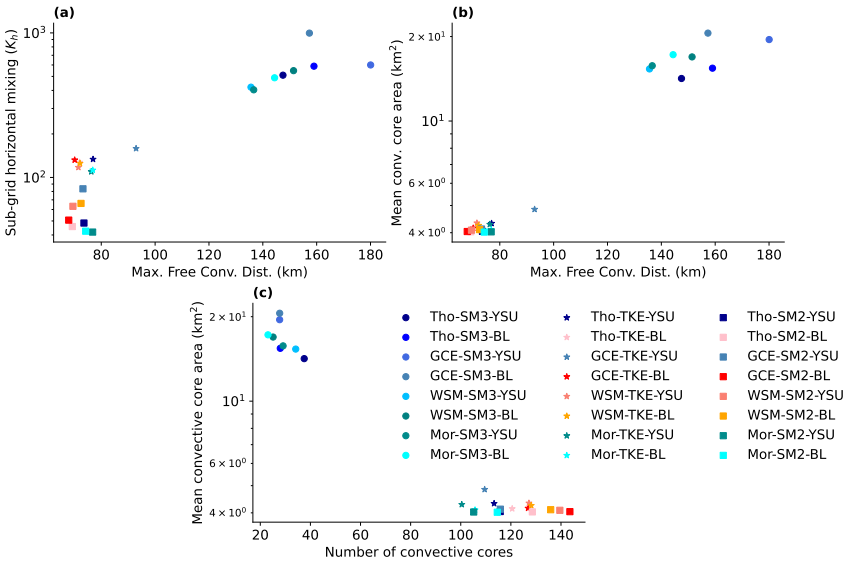

FIGURE 3.7: (a) Horizontal mixing (K_h) mean (from day 2 to 5) inside convective cores as a function of mean d_{clr} (from day 2 to 5), notice that the y-axis has a log scale, (b) mean convective core area as a function of mean d_{clr} (from day 2 to 5) and (c) mean convective core area (from day 2 to 5) as a function of mean number of convective cores (from day 2 to 5). Each plot includes all 24 simulations, including Smag3 (circles), Smag2 (squares), and TKE (stars) simulations. Blue-green colors indicate runs that produce SA, while red-yellow colors represent runs without SA.

core number, since the vertically integrated Q_{rad} remains consistent across all experiments before SA attains full development (not shown). In this context, the observed increment in mean convective core area translates to a diminished core count and, concomitantly, an increase in d_{clr} due to their inverse relationship (Figure 3.7b-c). Thus, the discernible association between strong entrainment, reducing core counts and increasing d_{clr}, offers a coherent explanation for the Smag3 simulations results.

3.4.4 Sensitivity to the Microphysics

Since both Smag2 and TKE schemes exhibit SA under identical PBL and microphysics configurations (Figure 3.2), for simplicity here we focus on the simulations that use Smag2 to assess the sensitivity to microphysics (similar results are obtained with the simulations that use the TKE scheme, as shown by Figure A.8 in Appendix A). There could be many ways in which microphysics schemes control the number of cores (Weverberg et al., 2013; Han et al., 2013) and *ergo* d_{clr}, for example with the autoconversion of rain (not shown), or the ice/snow fall speed (Figure 3.3). Surprisingly sensitivity experiments increasing (Tho-SM2-BDI) and decreasing (Tho-SM2-YHI) the ice/snow fall speed on Tho-SM2-YSU and Tho-SM2-BL simulations do not change the SA behavior (Figure 3.3). The fall speed experiment only delayed the aggregation onset compare to Tho-SM2-YSU simulation, consistent with Bretherton et al. (2005).

Another way in which microphysics can modify the number of convective cores is by changing the evaporation of rain, which is known to impact SA since it modifies the CPs intensity. In fact, comparing the Tho-SM2-YSU and WSM-SM2-BL runs

from Figure 3.3, we see that although Tho-SM2-YSU has a larger d_{clr} compared to WSM-SM2-BL, the values are not significantly different, but their aggregation differ. One explanation for this could be their CP intensities, that can prevent SA by redistributing humidity (Jeevanjee and Romps, 2013; Wing and Emanuel, 2014; Muller and Bony, 2015; Holloway and Woolnough, 2016; Yanase et al., 2020; Muller et al., 2022a). In our example, the experiment with weaker and smaller CPs (Tho-SM2-YSU) favor the development of SA compared with the one with stronger and larger CPs (WSM-SM2-BL) because weaker CP intensities are less efficient reallocating moisture, favoring the development of dry enough zones that could onset SA, and also because they are more space filling, disfavoring SA.

From the aforementioned, it is clear that the natural path to follow is to analyze differences in the CP by modifying the evaporation of rain and also by analyzing their intensity (see § 3.3 for details). The analysis of CP intensity (Figures 3.8a-b-d and A.8) reveals a clear trend: simulations that disallow the onset of SA tend to feature stronger CP intensities and a larger number of large CPs (Figure 3.8e), consistent with our earlier discussions. This observation supports the argument put forth by Yanase et al. (2020), suggesting that the intensification and enlargement of CPs can trigger an evaporative-driven negative feedback. Such feedback mechanisms have the potential to offset the positive feedback driven by radiatively induced near-surface horizontal divergence, leading to a redistribution of humidity and ultimately preventing the emergence of dry patches.

In experiments with SA, CPs exhibit lesser strength and smaller size, accompanied by heightened convergence at their associated gust fronts (Figure 3.8). This reduced convergence (W_{100}) at CP gust fronts (Figure 3.8c) leads to a lower convective triggering rate (Tompkins, 2001b; Stevens, 2005; Sherwood et al., 2009; Dawson et al., 2010; Li et al., 2015; Fuglestvedt and Haerter, 2020; Casallas et al., 2023b), and subsequently to a decreased number of convective cores (resulting in larger d_{clr}) that favors the development of SA. This underscores the dual role of CPs in both moisture redistribution and their potential influence on convective triggering.

Fuglestvedt and Haerter (2020) highlighted the role of CP fronts in fostering a conducive environment for convection initiation by generating moist patches through convergence at their gusts (Schlemmer and Hohenegger, 2014). When these moist gust fronts collide, they trigger new convection (Tompkins, 2001b,c). This underscores the significance of both convergence and the area/intensity of CPs in influencing convective triggering. Moreover, these gust fronts induce potent and sustained vertical velocities at the surface, fostering the initiation and intensification of convection (Böing et al., 2012). These insights shed light on why experiments using Mor's scheme consistently result in SA, while those employing the WSM scheme tend to remain random. Mor's scheme generates weaker and smaller CPs with limited convergence at their gust fronts, whereas the WSM scheme produces stronger and larger CPs with substantial convergence. This disparity leads to two critical outcomes: less efficient moisture redistribution (Figure 3.8d), favoring the development of dry regions in Mor's scheme and the opposite in the WSM scheme, and varying convective core numbers driven by CP size/convergence, altering d_{clr} to favor or hinder the development of SA.

These findings are supported by the evaporation of rain experiments. Tho-SM2-YDE, featuring increased CP area/intensity and gust front convergence, does not

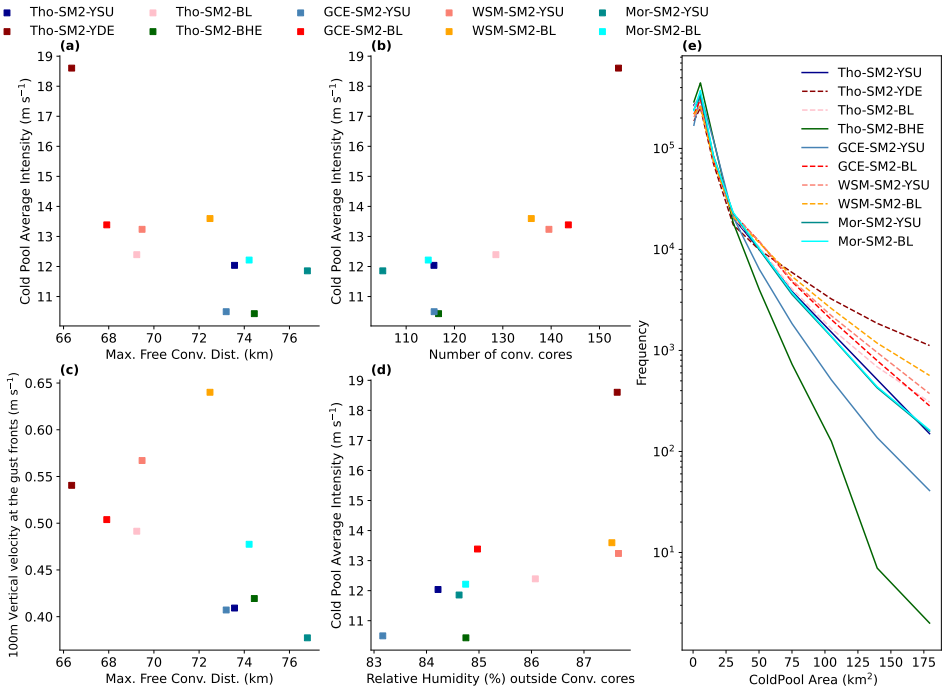

FIGURE 3.8: Cold pool average intensity (from day 2 to 5) as a function of (a) mean d_{clr}, (b) number of convective cores and (d) boundary layer relative humidity outside convective cores (from day 2 to 5). (c) Mean vertical velocity at 100m (W_{100}) at the gust fronts of the Cold-Pools as a function of mean d_{clr} (from day 2 to 5). (d) Frequency of Cold-Pool area between day 2 to 5. Continuous lines depicted aggregated runs, meanwhile dashed lines show runs with random convection. Blue-green colors indicate runs that produce SA, meanwhile red-yellow colors represent runs without SA.

lead to SA (unlike Tho-SM2-YSU that develops SA). Conversely, Tho-SM2-BHE, with decreased CP area/intensity and gust front convergence, exhibits SA (unlike Tho-SM2-BL that does not develop SA). These results must be taken with caution since modifying rain evaporation impacts CPs, but also latent heating profiles, tropospheric moisture budgets, and downdraft patterns (Holloway et al., 2017). It also induces changes in gust front convergence, core counts, and in d_{clr} (Figure 3.8), which are essential for SA development. This underscores the need for new experiments or simple models to isolate CPs' role in SA development or prevention.

3.4.5 Sensitivity to the PBL

Experiments conducted with the Mor and WSM microphysical schemes reveal no sensitivity to PBL configurations. For this, this section focuses on the Tho and GCE microphysics, uncovering a sensitivity of the SA phenomenon to PBL schemes (Figure 3.9a-b). Simulations employing the YSU scheme (non-local) exhibit larger low-level cloud cover, fewer cores, and increased core distances compared to those using the BouLac scheme (local).

To ascertain whether non-local schemes indeed yield greater low-level cloud cover, we conducted four perturbation experiments designed to provide a more deterministic validation of our findings. These experiments, initially configured with

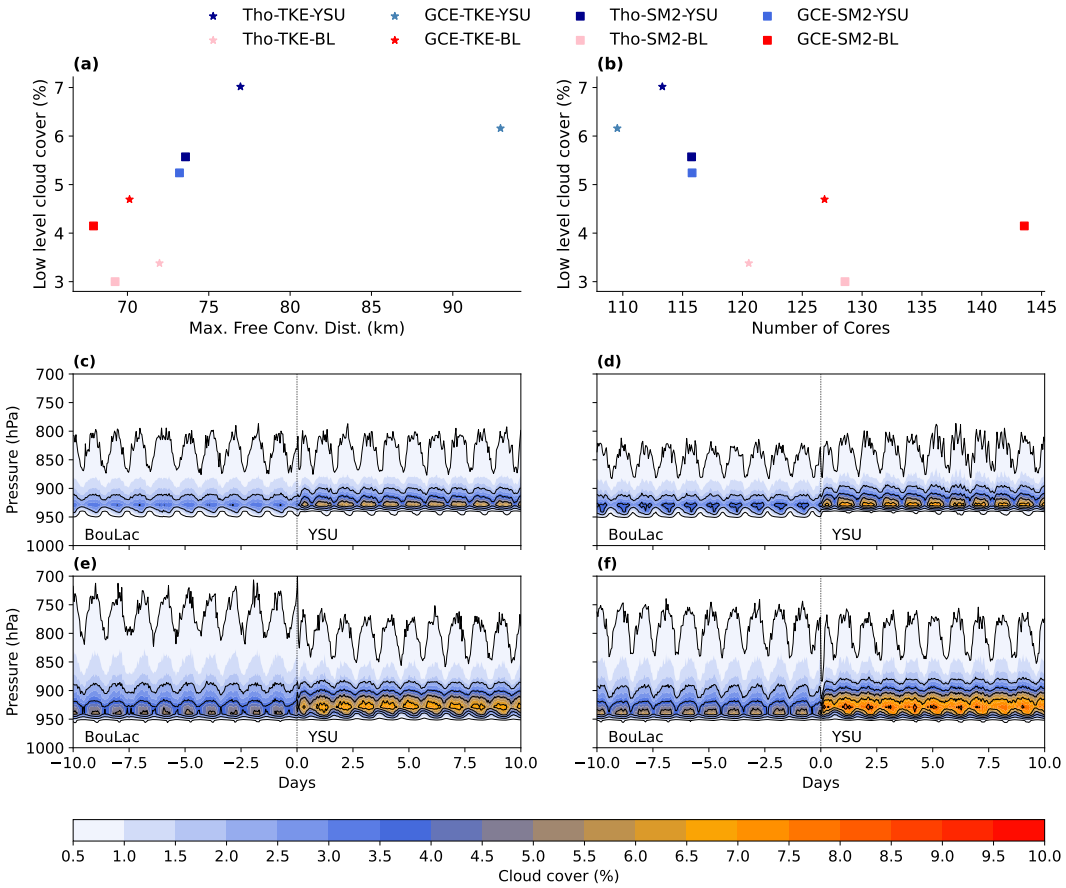

FIGURE 3.9: Low level cloud cover mean between day 2 to 5, plotted as function of (a) mean d_{clr} and (b) mean number of convection cores between day 2 to 5. Blue-green colors indicate runs that produce SA, meanwhile red-yellow colors represent runs without SA. Low cloud cover evolution for (c) Tho-SM2-BLY, (d) Tho-TKE-BLY, (e) GCE-SM2-BLY, and (f) GCE-TKE-BLY. Notice that on the x-axis from -10 to 0, the simulation is the last 10 days of the simulations using BouLac, and then the values are from the simulation that start from the last timestep of the BouLac runs but that use YSU.

BouLac scheme, utilized the final state of these runs as initial conditions of a random state (details in § 3.3). Subsequent 10-day simulations with a transition to the YSU scheme are performed, and reveal two consistent trends: (i) the onset of a dry patch typically occurs around day 6 (not shown). (ii) An increase in low-level cloud cover is observed between 950 to 850 hPa across all simulations, regardless of sub-grid scale or microphysics scheme (Figure 3.9c-d-e-f), thus indicating a direct relationship between changes in low clouds and the choice of PBL scheme.

Non-local schemes generate larger low-cloud cover probably due to two main reasons. Firstly, they feature stronger vertical mixing within the boundary layer compared to local schemes (Hu et al., 2010; Xie et al., 2012, 2013). This enhanced mixing promotes the ascent of moist air parcels, facilitating low-level cloud formation. Secondly, non-local schemes induce stronger entrainment from the top of the boundary layer, driven by larger moisture gradients (Figure A.9), thereby promoting enhanced low-cloud formation, consistent with Hu et al. (2010) results.

Based on our findings, we propose that the observed increase in low-level cloud cover instigate a circulation pattern, facilitating the transfer of MSE from drier to moister regions, as initially suggested by Muller and Bony (2015). To quantify the circulation between dry to moist regions circulation, we use the stream function Ψ (see method section for details). Figure 3.10 show a consistent net circulation from dry to moist regions, resembling the circulation reported by previous studies (Bretherton et al., 2005; Coppin and Bony, 2015; Muller and Romps, 2018). This circulation is primarily induced by anomalies produced by low clouds (Figure A.10), consistent with the LW radiative cooling anomalies (Figure A.11), although the clear sky component also exerts an influence, meanwhile the SW radiative cooling (Figure A.12) has a minor contribution.

This aforementioned low-level circulation implies that the spatial variability of low-level cooling strongly favors the development of SA (Muller and Bony, 2015; Shamekh et al., 2020a). This is because low-level cooling (due to low-clouds) in dry regions promotes subsidence, which, in turn, triggers a low-level flow from dry to moist areas, resulting in an upgradient transport of MSE and, consequently, favoring SA. This is evident in Figure 3.10, where the circulation of the aggregated runs fails to extend to the driest columns, especially between $\approx$ 850 and 900 hPa (as indicated, as an example, by the bold dashed line). Instead, it remains predominantly within intermediate humid regions, promoting dryness in the dry areas while simultaneously moistening the already humid regions (Muller and Bony, 2015), reminiscent to the idea that redistribution of moisture from dry to moist regions is an essential aspect for SA development (Jeevanjee and Romps, 2013). This finding is supported by a mechanism denial experiment we performed (not shown), in which the contribution of low clouds to radiative cooling is eliminated, and SA do not develop, consistent with previous studies that have conducted similar experiments (Muller and Held, 2012; Muller and Bony, 2015).

The LW clear-sky component (Figure A.11) also influences the generation of a low-level circulation that exports MSE upgradient. This occurs because non-local schemes indirectly incorporate horizontal moisture advection within the boundary layer over broader spatial scales (Potvin et al., 2020; Segura et al., 2021), unlike local schemes that only consider neighboring pixels (Hu et al., 2010). This efficient large-scale moisture transport also promotes an increase in moisture gradients within the boundary layer, leading to an enhancement in low-level circulation. Additionally, heightened moisture gradients can facilitate the development of dry regions and, subsequently of SA, aligning with the findings of Yang (2018) and Yao et al. (2022).

These result arise the following question: what is the relation between low-clouds and convective core counts, depicted in Figure 3.9a-b? Simulations that aggregate has larger low level cloud amounts, producing stronger cooling (Figure A.13), as this cooling need to be balance by subsidence, the subsidence strength increase. This means that there are less zones prompt for the triggering of convection, disfavoring core counts, subsequently increasing d_{clr}, and enhancing the development of SA, which is consistent with the simple model of Biagioli and Tompkins (2023a). To show this, we compute the convective inhibition (CIN) for each pixel (Ladwig, 2017) in the simulations discussed in this section, since CIN provides insights into regions where convection may be hindered or triggered (Emanuel, 1994; Holton, 2004).

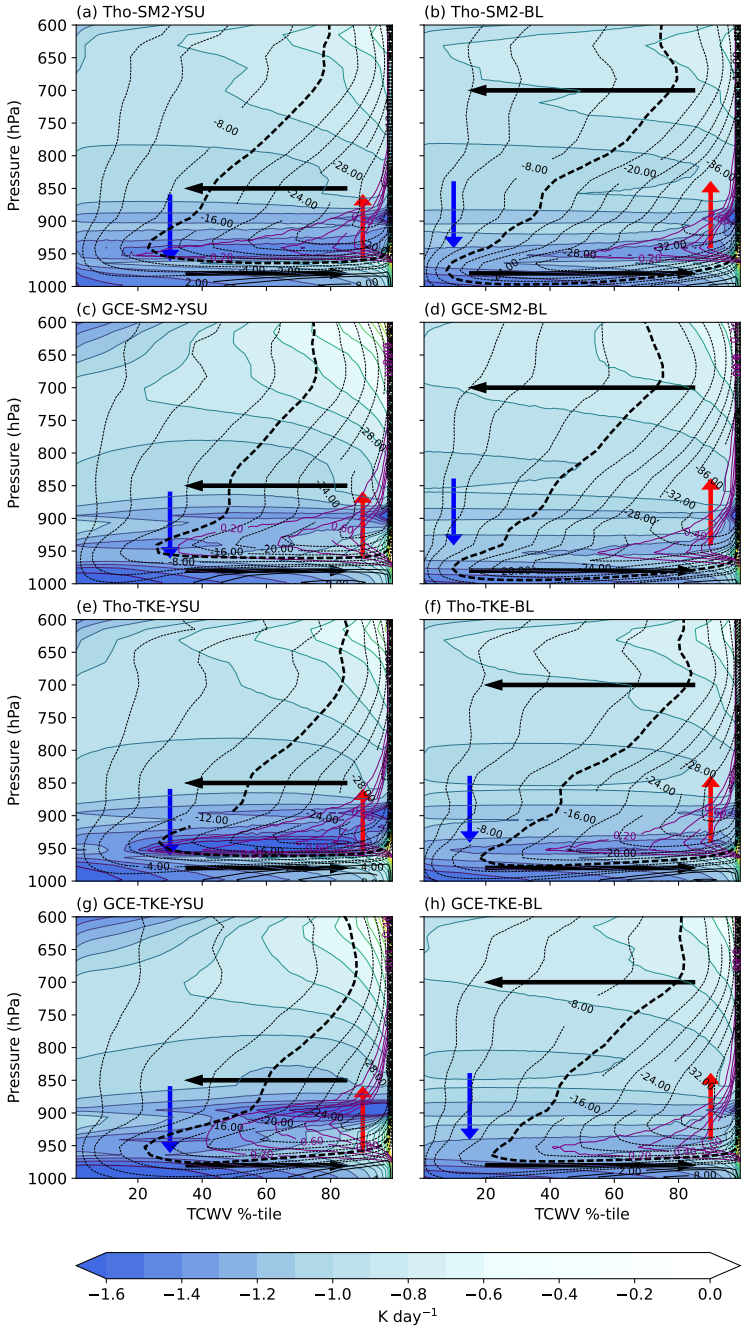

FIGURE 3.10: Net radiative cooling (colors) between day 2 to 5 ordered on TCWV-%tile. For (a) Tho-SM2-YSU, (b) Tho-SM-BL, (c) GCE-SM2-YSU, (d) GCE-SM2-BL, (e) Tho-TKE-YSU, (f) Tho-TKE-BL, (g) GCE-TKE-YSU, and (h) GCE-TKE-BL. The left column runs produce SA, while the runs on the right column produce random convection. Black contours indicate the stream function Ψ as a function of TCWV-%tile and height (units 10^{-2} kg m^{-2} s^{-1}). The bold dashed line represents the -12 line, illustrating, how the circulation persists within intermediate humid and the moistest regions. Purple contours indicate the liquid water mixing ratio (units 10^{-5} kg kg^{-1}). The arrows schematically show the subsidence due to radiative cooling (blue) and rising motion caused by warming (red), along with the low-level and mid-level (solid black) flows induced.

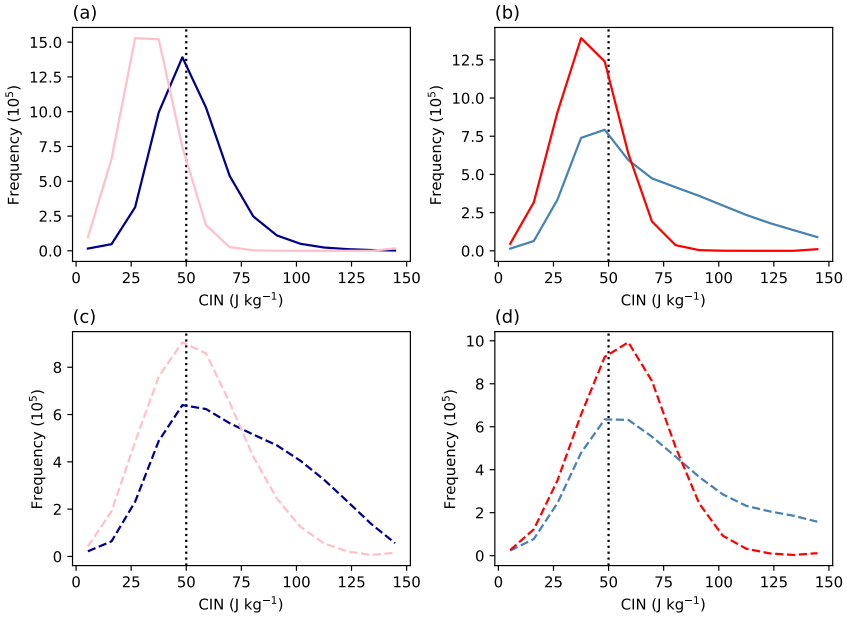

FIGURE 3.11: PDFs for Convective Inhibition (CIN) for (a) Tho-SM2-YSU and Tho-SM2-BL, (b) GCE-SM2-YSU and GCE-SM2-BL, (c) Tho-TKE-YSU and Tho-TKE-BL, and (d) GCE-TKE-YSU and GCE-TKE-BL from day 2 to day 5. The vertical dotted line indicates the threshold identified by Behrendt et al. (2011) beyond which deep convection is strongly disfavor. Notice that blue colors depict aggregated runs, while red colors represent runs without SA.

Figure 3.11 illustrates that simulations where SA occurs exhibit fewer regions with weak CIN values and more regions with stronger CIN compared to simulations without SA development. This is highlighted by the vertical dotted line, representing the threshold distinguishing between favorable and unfavorable conditions for deep convection initiation (Behrendt et al., 2011). In aggregated runs, convection is less likely to be triggered, not only because there are fewer regions conducive to new convection, but also due to larger areas of strong CIN, which actively inhibit convection initiation. Consequently, runs with SA display fewer convective cores and a larger d_{clr}.

3.5 Chapter Summary and Conclusions

The aim of this study was to understand the mechanisms influencing the occurrence or prevention of convective self-aggregation (SA) across diverse idealized modeling scenarios. Using the WRF model, we explored 24 different combinations of parameterizations, including sub-grid scale mixing, planetary boundary layer (PBL), and microphysics, without modifying domain characteristics. Our analysis and experiments highlight the crucial role played by the maximum free convective distance (d_{clr}) in determining whether simulations undergo a transition to an SA state or maintain a random configuration. This parameter, d_{clr}, reveals the nuanced sensitivity of SA to perturbations in sub-grid scale mixing, PBL, and microphysics.

In essence, the propensity for convective SA intensifies with an increased d_{clr}, especially when cold pools (CPs) are characterized by limited size, intensity, and

gust front convergence. These combined factors are significant due to several phenomena: **Convective Triggering:** Reduced CP intensity, size, and gust front convergence result in decreased convective triggering. This is attributed to fewer CP collisions and a reduction in moisture convergence at their fronts, ultimately leading to a lower core count and an increase in d_{clr}—favoring aggregation. **Distinct Moisture Redistribution:** Larger d_{clr} and weaker CP intensity and size foster the emergence of dry regions. The spatial separation of moisture sources (Biagioli and Tompkins, 2023a) and less efficient moisture redistribution, characterized by reduced effectiveness from distant sources, collaborate to create and reinforce these dry zones. **Divergent Feedback:** The dry patches intensify through radiative cooling and subsidence. Subsidence accentuates the cooling and drying of non-convective regions, resulting in dry anomalies and positive pressure anomalies (Yao et al., 2022). This induces a divergent flow that further dries the region, creating a feedback mechanism, named here divergent feedback (Shamekh et al., 2020a).

Alternatively, the interconnected variables influencing d_{clr} highlight its important role in understanding SA development. These variables exhibit intricate interplays and modulations:

- **Sub-Grid Scale Mixing:** The horizontal mixing imparts entrainment, amplifying transported mass within clouds to yield large convective cores, albeit in fewer numbers. A diminished number of cores escalates d_{clr}, culminating in the onset and amplification of dry patches and leading to SA.

- **Microphysics:** The evaporation of rain entails multifaceted repercussions. Beyond affecting CP intensity/size, it imparts changes to convective core count, and d_{clr}. Weaker evaporation translates to less efficient moisture redistribution and weaker CP gust fronts convergence, resulting in less conducive conditions for convection triggering (Tompkins, 2001b; Bretherton and Blossey, 2017; Fuglestvedt and Haerter, 2020), thus modulating convective core number, d_{clr}, and subsequently SA.

- **PBL:** Non-local schemes generate larger low-cloud cover, even in dry regions, due to strong entrainment from the top Hu et al. (2010). These low clouds heightens radiative cooling, fostering a circulation that maintain moisture in the moist regions, favoring SA. Additionally, the pronounced radiative cooling prompts stronger subsidence, reducing convective triggering, diminishing core counts, increasing d_{clr}, and promoting SA (Muller and Bony, 2015; Coppin and Bony, 2015).

Our investigation has yielded valuable insights, yet has also shown unexplored avenues that merit investigation. We outline several questions warranting attention for future research:

1. **Microphysical Factors Governing Convective Cores:** Beyond rain evaporation, an unaddressed query pertains to the microphysical parameters steering the number of convective cores.

2. **Unraveling Cold Pool Impact on SA:** The intricate role of CPs in SA prompts further exploration: (i) Understanding their influence on moisture redistribution; (ii) Investigating their ability to modify convective core number and d_{clr}, particularly through a space-filling framework, affecting SA development probability; (iii) Examining how CP intensity, size, and moisture redistribution

collectively influence SA emergence; (iv) Modifying rain evaporation induces changes in CP characteristics, requiring further experiments or simple models for a comprehensive understanding of their role in SA.

3. **Seeking Mechanisms in Observations:** Observational campaigns can enhance our understanding of the presented results. In the context of emerging dry regions associated with CPs, conducting surface moisture measurements in CP-prone regions can differentiate between dry anomalies caused by CPs and location-specific variations. This, along with examining dry anomaly lifetimes and clustering patterns, can provide insights into whether CPs lead to dry anomalies and favorable conditions for SA in nature, as suggested by the presented results.

In summation, these questions underscore the avenues yet to be traversed in comprehending the intricacies of SA. Addressing these inquiries would allow to enhance our understanding of the phenomenon and also to illuminate novel facets of its governing mechanisms.

Chapter 4

Dynamical and Thermodynamical Drivers of Humidity Variability in the Tropical Western Pacific Warm Pool

4.1 Abstract

Idealized model simulations reveal spontaneous aggregation of tropical convection on the beta-mesoscale due to radiative-cloud feedbacks, influencing humidity variance and potentially impacting climate sensitivity. Our multivariate analysis of tropical Western Pacific observations challenges idealized models, showing that despite radiative and surface fluxes promoting convection clustering in boreal summer/autumn, a random configuration with homogeneously mixed humidity prevails due to homogeneous sea surface temperature (SST). Conversely, in boreal winter/spring, with a weak meridional SST gradient, organized convection occurs over warm SST regions, introducing substantial meridional humidity gradients. Periodic episodes of counter-gradient convection and limited humidity variance are linked to westward propagating convectively coupled Rossby waves. This underscores SST gradients and equatorial wave dynamics as key drivers of convective organization and humidity variability in the Pacific warm pool.

4.2 Introduction

Water vapor, a key greenhouse gas, plays a major role in the tropical energy budget. Climate models project a positive feedback, anticipating an increase in specific humidity under the assumption of constant relative humidity (Colman and Soden, 2021). However, spatial variations in water vapor, influenced by convection patterns, may impact this response. In the tropics, convective precipitation correlates exponentially with column humidity (Bretherton et al., 2004; Rushley et al., 2018). The aftermath of convection is a saturated free troposphere, while drying occurs in non-convective areas due to compensating subsidence (Sun and Lindzen, 1993). This knowledge has facilitated the development of simple models, grounded in column saturation over high sea surface temperatures (SSTs) or areas characterized by upward motions balanced by advection, and compensatory subsidence elsewhere (Sherwood, 1996; Pierrehumbert, 1998). These models have been instrumental in replicating the large-scale water vapor distribution in the tropics, helping us explain the climatological relationship between total column water vapor (TCWV) and SSTs

(Kanemaru and Masunaga, 2013).

While these broad, tropics-scale models can explain the mean climatology of the water vapor distribution and radiation budget, recent evidence suggests that the details of the arrangement of convective clouds, particularly the degree of spatial "clustering" over scales of a few hundred kilometers (the beta and alpha mesoscale), also play a crucial role in understanding temporal variations in the tropical energy budget. Deep convective activity in the tropics migrates around the interior and boundaries of a moist region with TCWV exceeding 48 kg m^{-2}, separated from the dry subsidence regions by sharp horizontal moisture gradients (Mapes et al., 2018). This, combined with low-level static stability that determines low cloud amount, a metric of the deep convective arrangement, I_{org} (Tompkins and Semie, 2017), has been shown to explain nearly all of the variability in the net tropical radiation budget over monthly to annual time scales (Bony et al., 2020). This result is particularly intriguing, as a scale analysis of I_{org} has demonstrated that it measures convective clustering almost exclusively over the beta mesoscale (20-200 km) (Biagioli and Tompkins, 2023b). Thus, mesoscale variations in the arrangement of convection appear to determine the moisture leaving the moist region to the adjacent subsidence zones, impacting the energy budget of the drier "radiator fins" of the tropical atmosphere (Pierrehumbert, 1995).

Understanding what determines the organization of convection and associated patterns of mesoscale humidity variability is crucial for gaining insight into drivers of variations in the tropical radiative budget, and it may also be essential for assessing tropical climate sensitivity. Experiments with cloud-resolving models (CRMs) show that convection can spontaneously cluster, leading to very dry atmospheres, even in the absence of heterogeneous large-scale dynamical forcing, due to spatial heterogeneity in diabatic heating, such as long-wave (LW) cloud forcing (Tompkins and Craig, 1998a; Bretherton et al., 2005; Muller and Held, 2012; Wing and Emanuel, 2014). If this organization of convection is sensitive to the tropical mean SST, it could act as an additional climate feedback, which is as yet relatively unexplored and poorly understood (Emanuel et al., 2014). One concern is that if the spatial scales of the clustering are important on the alpha (200-1000 km) and beta-mesoscale, then it might be missing or poorly represented in global climate models that employ classical parameterization schemes. Unfortunately, initial attempts to assess this using idealized model intercomparisons have failed to show any consensus between models (Wing et al., 2020).

One hindrance to progress is the difficulty of assessing convective organization with present-generation observations. Various attempts to measure convective organization have been made, using a variety of "measures" of convective organization, which in a recent comprehensive review by Biagioli and Tompkins (2023b), were divided into direct and indirect metrics. Direct measures of organization attempt to identify the location of convective updrafts and the relative spatial organization of these. This can be done using ground-based scanning precipitation or Doppler radar (Radtke et al., 2022), but it is restricted to specific limited spatial areas where the radar is located. The 94 GHz polar-orbiting CloudSat radar, on the other hand, was not a scanning instrument and thus could not discern spatial organization. The lack of Doppler radar in space and the relatively coarse resolution of remotely sensed precipitation observations mean that space-bound measures of organization must thus attempt to use cloud locations or top-of-atmosphere (TOA) outgoing longwave

radiation (OLR) as proxies for convective location. While this can discern organizational modes of convection on large scales (>1000 km), such as the Madden Julian Oscillation, eastward propagating moist-Kelvin waves, and westward propagating Rossby waves (Wheeler and Kiladis, 1999), cloud overlap complicates this over the mesoscale in visible and infrared imagery (Weger et al., 1992; Tobin et al., 2012), despite the application of filters to try to mitigate such effects and identify overshooting tops (Bony et al., 2020).

Complementary to this approach, indirect metrics attempt to identify convective clustering through its impact on the environment, such as clustering increasing spatial humidity or moist static energy variance (Holloway et al., 2017). As noted by Biagioli and Tompkins (2023b), the issue with these univariate signatures is that they are influenced by various factors in addition to convective organization, such as wind shear or lower boundary temperature, which complicates the comparison of scenes with different lower boundary conditions. However, progress can be made by adopting other indirect metrics from idealized modeling studies, notably identifying bi- or multivariate relationships between SST and TCWV variance in different phases of clustering. To date, limited use of such bivariate metrics has been made in past observational studies (Holloway et al., 2017). Here, we aim to further understand the controls of humidity variability on the mesoscale and its association with deep convective organization in state-of-the-art observations in the tropical western Pacific, applying a novel but simple multivariate analysis technique adopted from idealized modeling studies (Tompkins and Semie, 2021).

4.3 Data and Methods

The analysis uses the latest generation satellite retrievals for column water vapor (from MIMIC-TPWV2), and SST (from Himawari-8 and NOAA OISST), supplemented by ERA-5 reanalysis.

4.3.1 Satellite datasets

We used the 3-hourly TOA OLR from the Gridded Satellite (GridSat) B1 dataset (Knapp et al., 2011). This dataset has a temporal uncertainty of less than 0.1 K per decade and has a horizontal resolution of 1.0°. Hourly SST (level 3) and Cloud Top Height (CTop; level 2) were downloaded from the Himawari-8 Collection version 1.2. The Himawari-8 is a geostationary meteorological satellite operated by the Japan Meteorological Agency (JMA) that was launched on October 2014 and became operational on July 2015 (Bessho et al., 2016). The SST has a horizontal resolution of 2 km and has been previously evaluated by other authors (e.g., Kurihara et al. 2016, 2021). The CTop was used to calculate the cloud fraction and has a horizontal resolution of 5 km. This product is regarded as one of the most precise cloud products (e.g., Ishida and Nakajima 2009; Ishida et al. 2011). Additionally, we incorporate daily SST data from NOAA Optimum Interpolation SST (OISST) version 2.1 with a spatial resolution of 0.25° (Huang et al., 2021). This dataset combines ship measurements, satellite imagery, and buoy data to create a blended product that provides an accurate representation of SST. For comprehensive information on the dataset and its validation, the reader is refer to Huang et al. (2021).

The Morphed Integrated Microwave Imagery CIMSS for Total Precipitable Water (hereafter MIMIC) product developed by the Cooperative Institute for Meteorological Satellite Studies (CIMSS) was used in this study. It is an experimental product based on the morphological compositing of data from several polar-orbiting satellites (Wimmers and Velden, 2011). Its mean average error is between 0.5-2 mm over the ocean (Wimmers and Velden, 2011), and its horizontal resolution is 0.25°.

For precipitation, we used data from the Global Precipitation Measurement (GPM) Core Observatory spacecraft, which has an advanced dual-frequency precipitation radar and a state of the art microwave imager. Due to these sensors, the instrument is capable of sensing light rain and falling snow (Hou et al., 2008, 2014), making it more reliable. The data used has a 30-minutes temporal resolution, a horizontal resolution of 5 km, and a bias of less than 50% at 1 mm h^{-1} and of 25% at 10 mm h^{-1} (Hou et al., 2013; Skofronick-Jackson et al., 2018). It is important to mention that all the datasets used in this study were downloaded for the time period of 1 January 2017 to 31 December 2021.

4.3.2 Meteorological Reanalyses

Reanalyses are one of the best estimates of global atmospheric conditions available (Hersbach et al., 2020). Here we used the ERA5 reanalysis data, which is the fifth generation atmospheric reanalysis product from the European Centre for Medium-Range Weather Forecasts (ECMWF). It provides us with hourly data for wind components, air temperature, specific humidity, relative humidity (RH), and specific ice and liquid water content. We use data from the surface until the 750 hPa level with a vertical resolution of 25 hPa and then with a 50 hPa resolution until 100 hPa level. We also used single-level hourly data for the SST, latent heat (LH) and sensible heat (SH) fluxes, 10 meter u and v components of the wind, 2 meter air temperature, TCWV, both components of the vertically integrated water vapor flux (WVF), cloud ice water path (CIWP), cloud liquid water path (CLWP), all-sky TOA OLR and clear-sky TOA OLR, mean surface upward LW clear (LWCUpw) and all sky (LWUpw) radiation, all-sky and clear-sky shortwave (SW) radiation.

Hourly data were downloaded for the time period of 1 January 2017 to 31 December 2021 with a horizontal resolution of 0.25°. Note that in order to reduce the uncertainty (Chen et al., 2022), all the ERA5 reanalysis datasets were compared with their respective (available) satellite product counterparts. Noteworthy the fact that all map plots were constructed using the Cartopy Python package (Met-Office, 2015).

4.3.3 Identifying Organize and Reversal Regimes

To identify convective organization in these regions, we employ simple multivariate analysis metrics that have been applied to idealized CRMs (Wing et al., 2020). Using experiments with a slab ocean coupled to a high-resolution atmospheric CRM, several characteristics can be discerned in random and clustered convective states. Beginning with homogeneous conditions, convection in these simulations is initially randomly arranged, with the exception of scales below 20 km due to the action of cold pools (Tompkins and Semie, 2017).

Cloud-resolving model studies have sought to comprehend the multivariate relationships between convective activity, water vapor, and SST by analyzing their

evolution in water vapor space or energy space, following the approach introduced by Bretherton et al. (2005) and Wing and Emanuel (2014). In these studies, the mesoscale-sized, two-dimensional longitude-latitude simulation domain is reordered into a single vector arranged from driest to moistest columns (Wing and Emanuel, 2014; Tompkins and Semie, 2021) or, similarly, in terms of moist static energy (Bretherton et al., 2005). Subsequently, other variables can be binned as a function of the TCWV.

In situations characterized by random convection, spatial variations in SST and water vapor are limited, and SST shows a spatial anti-correlation with water vapor, as SW radiation is intensified in clear sky moist regions. Conversely, following the onset of convective aggregation, the spatial variance of humidity significantly increases, manifesting dry areas alongside convectively moist regions (Shamekh et al., 2020a; Tompkins and Semie, 2021). The presence of these dry areas allows the enhanced OLR and LH flux to outweigh the increased incoming SW in the surface energy budget, leading to cooler SSTs. In summary, the correlation between SST and TCWV is negative, with warmer SSTs over the moisture-conducting regions in states of random convection, but becomes positives in states of clustered convection.

We apply this method to analyze observations over the target regions. At each time step, we determine the linear regression relationship using SST anomalies (from Himawari-8) as a function of the sorted TCWV field (from MIMIC-TPWV2) and calculate the slope ($\frac{dSST}{dTCWVf}$ with units of Km^2kg^{-1}). Composites of "reversal events" during the boreal winter/spring period are created by applying a threshold to the SST-humidity linear relationship. Random "reversal" states are identified when the relationship is less than half of the standard deviation for a minimum of 24 consecutive hours. We conducted sensitivity tests, varying the standard deviation threshold (0.25*std and 0.75*std) while maintaining the same temporal requirement. Additionally, we tested the sensitivity of the temporal threshold by keeping the standard deviation threshold at 0.5*std but changing the temporal threshold to 12 and 36 consecutive hours. None of these tests yielded significant differences in the conclusions or arguments presented in this study.

The sensitivity of this fit was tested by excluding pixels with active precipitation to address increased uncertainty in the Himawari-8 product in the presence of clouds. However, no significant changes were observed. The slope was also calculated using ERA5 SST and TCWV data, yielding largely similar results (mean squared error = 9.39e-6 Km^2kg^{-1}, with a correlation of 0.88 for 1440 data points).

To identify differences between organized and reversal states and analyze their implications, we calculated the moisture, wind, cloud, and radiative characteristics of these states using daily means from ERA5 data. Additionally, to mitigate uncertainty in the procedure and analysis using the slope's quartiles, calculations were repeated using satellite products. Results from these datasets continue to support our arguments and conclusions.

Since we use a very simple statistical relationship to capture the association between SST and TCWV. We decide to apply a multivariate rotational empirical orthogonal function (REOF) analysis for robustness, to do it, we utilize the xeofs library version 2.2.4 (Rieger and Levang, 2023) in Python 3.8. The results indicate that

the combined sum of the first three principal components (PCs) effectively reproduces the $\frac{dSST}{dTCWVf}$ calculation (Figure B.1 in Appendix B), showcasing a correlation of 0.67 for the entire dataset and 0.87 for all boreal winter/spring months.

The initial three PCs possess the capability to replicate the $\frac{dSST}{dTCWVf}$, as each one encapsulates information regarding the distinct positions of the highest TCWV values in relation to SST. Consequently, the cumulative sum of these positions inherently reproduces the slope calculation (Figure B.2 in Appendix B). Utilizing the aggregated sum of the first three PCs for conducting all statistical analyses presented in this study yields results consistent with those obtained using the slope (not shown), underscoring the robustness of our findings and conclusions.

4.3.4 Organization Feedbacks and Latent Heat Flux Decomposition

An important aspect of this research involves quantifying the radiative and surface flux feedbacks that contribute to organizing convection in idealized model setups. Building on the methodologies established in previous studies, both idealized and real setups (e.g., Wing and Emanuel 2014; Holloway and Woolnough 2016; Holloway 2017; Tompkins and Semie 2021), we computed the LW, SW, SH flux, and LH flux feedbacks for the entire study period and for the three designated study zones using ERA5 data.

In addition to this calculation, we decomposed the LH flux contributions by following and modifying the surface flux decomposition proposed by Tompkins and Semie (2021), in which the authors avoided the execution of a complex surface flux scheme in offline mode and simplified the procedure by using a machine learning (ML) approach to approximate the scheme results. The X-matrix (input) of the ML technique correspond to the 10 m wind speed, $\Delta q = q_{sat}(SST) - q_2$, and $\Delta T = T_2 - SST$, where q is the water vapor mixing ration, T is the air temperature, 2 is the measurement height, and the Y-matrix (output of the ML technique) is the LH flux. Both the input and output data (>100000 data points) are from the ERA5 reanalysis.

In this study, we employed a random forest (RF) algorithm for the ML analysis. The selection of this ML technique involved a randomized search aimed at determining the most effective ML technique and hyper-parameters, with an implemented early stoppage method to prevent over-fitting (Casallas et al., 2023a). While other ML techniques were explored as suggested by McGovern et al. (2019), such as a neural network, support vector machine, and a decision tree algorithm, no significant sensitivity was observed (not shown). Consequently, we present only the RF results, given its superior precision. The training process utilized 80% of the data, and the remaining 20% was reserved for evaluating the RF model under both random and clustered conditions. The validation results demonstrated an r^2 of 0.93, a root mean square error (RMSE) of 0.28 W m^{-2}, and a mean bias of 0.12 W m^{-2}. Once the model was trained, we calculated the impact of each variable (wind, Δq, and ΔT) following the procedure outlined by Tompkins and Semie (2021). This involved using the value of one variable in combination with the area-mean values of the other two variables, ensuring that the RF was employed only within the range of the training data.

4.3.5 Offline radiation calculations

To assess the radiative differences between organized and reversal states, we conducted offline calculations using the rapid radiative transfer model (RRTM) (Mlawer et al., 1997). The ERA5 data served as input, and the climlab code developed by Rose (2020) facilitated the decomposition of LW total and clear-sky flux anomalies. Ninety random instances of organized and reversal states were selected, and for each state, we manipulated SST, specific humidity, temperature, and cloud properties (i.e., cloud fraction, ice and liquid water cloud mixing ratio) in the RRTM model. For instance, to isolate the SST contribution, we replaced the SST value of a reversal episode among the 90 organized episodes while keeping all other variables constant (i.e., specific humidity, cloud properties, and air temperature), and then we quantify its contribution to the total change. This approach ensured that changes in the RRTM were solely attributed to SST variations. The methodology drew inspiration from the radiative anomaly decomposition by Bony et al. (2020).

4.3.6 Equatorial Rossby Wave Analysis

We first use a space-time bandpass filtering following the methodology outlined by Wheeler and Kiladis (1999). Initially, the OLR data from NOAA (spanning from January 1, 2002, to December 31, 2023) is detrended and temporally tapered by a split cosine bell (Gehne et al., 2022). Subsequently, a two-dimensional Fast Fourier Transform (FFT) (Brigham and Morrow, 1967) is applied for the wave ($n = 1$) to filter Equatorial Rossby (ER) Waves (Wheeler and Kiladis, 1999; Kiladis et al., 2006; Gehne et al., 2022).

Utilizing the aforementioned filtered OLR data, we adopt the method proposed by Gehne et al. (2022), which builds upon the work of Gottschalck et al. (2010). This approach involves computing empirical orthogonal functions (EOFs) (Rieger and Levang, 2023) in the Pacific region (20S to 20N latitude and 120E to 100W longitude). These EOFs serve as a basis for estimating ER wave activity within our study regions. We then project the filtered OLR data onto the spatial structures (EOFs) specific to each of the three study regions (for further details, refer to Gehne et al. (2022) Appendix A), thereby deriving the wave activity within each respective area.

4.4 Results

4.4.1 Seasonal Variation of Water Vapour

Convection organization is evaluated through a multivariate analysis of atmospheric retrievals, including TCWV, SST, cloud, and rainfall. The analysis focuses on mesoscale regions of approximately 10^6 km^2 in the tropical western Pacific (Figure 4.1), a size consistent with domains commonly employed in idealized studies on convective aggregation. The primary region under study spans from 2N to 9N and from 135E to 145E, strategically located in the Western Pacific warm pool, away from direct influences of the maritime continent. Additional analyses are conducted in secondary regions to the east. One located at 3N to 10N and 147E to 157E and the other region straddling the equator (between 3S and 4N, and from 156E to 166E) to verify the robustness of our findings. Details on the analysis of these secondary regions can be found in the Appendix B.

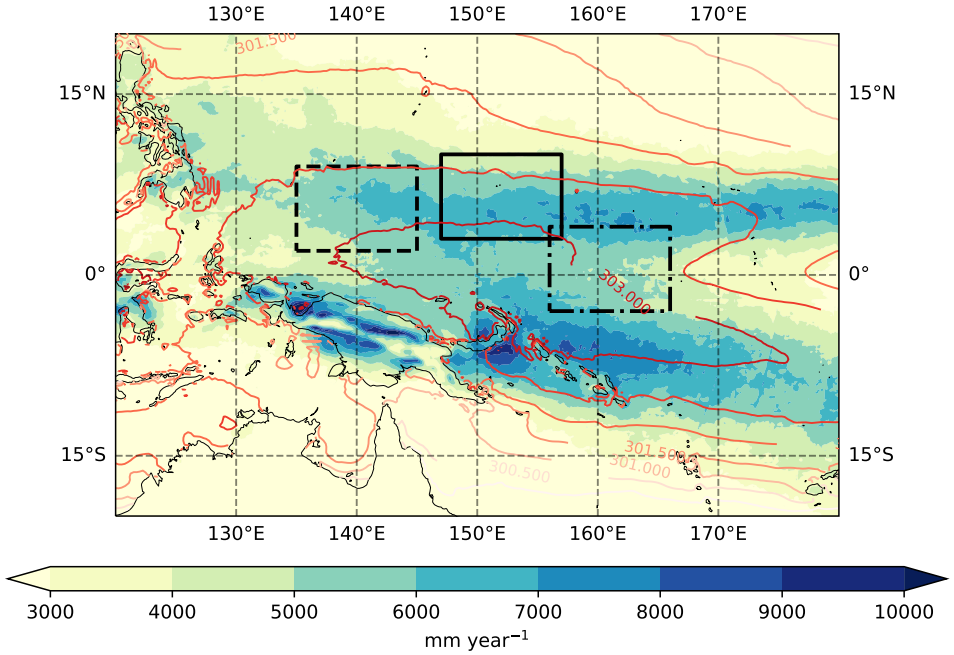

FIGURE 4.1: Annual mean distribution of Precipitation (colors) and SST (contours) for the study areas. The SST contours start from 301.75 and increase every 0.25K. The dashed rectangles delimits the three study regions.

Regarding the SST annual cycle, a subtle north-south gradient of approximately 1 K spans the 7-degree latitude range during the boreal winter and spring (Figure B.3 in Appendix B includes the seasonal evolution of the SST, TCWV, and wind in the regions of interest). Conversely, in the boreal summer months, mean SST gradients are virtually nonexistent (Figure 4.2). Convection responds to these changes in SST, with maximum rainfall occurring uniformly across the domain during the summer (JJA) months, slightly diminishing in the boreal autumn. As the SST cools by approximately 1 degree in the boreal winter and the main warm pool shifts south of the equator, rainfall also decreases, exhibiting a local peak at 6N. This precipitation gradient is most prominent in spring, surpassing a factor of two. While the primary precipitation maximum is south of the equator during this period, it falls outside the domain.

These shifts in the distribution of convection significantly impact the spatial variability of water vapor, a crucial variable for monitoring the onset of convective aggregation in model studies (Bretherton et al., 2005). From June to November, the mean zonal gradient remains minimal throughout the troposphere, aligning with the uniform distribution of rainfall and indicative of randomly scattered deep convection. In contrast, during the winter and spring months, a robust north-south gradient of relative humidity prevails above the boundary layer, with mid-tropospheric relative humidity averaging as low as 40%. This alignment with a peaked precipitation profile suggests organized convection. In summary, spatial variance is at its minimum during the boreal summer months and peaks in the MAM months. While large-scale ocean dynamics establish the north-south gradient of convection in winter/spring, our energy balance calculations will reveal that the radiation budgets tend to amplify this pattern. The reader is also referred to Figure B.3 in Appendix B, which include maps of the seasonal evolution of water vapor, wind, and SST in the

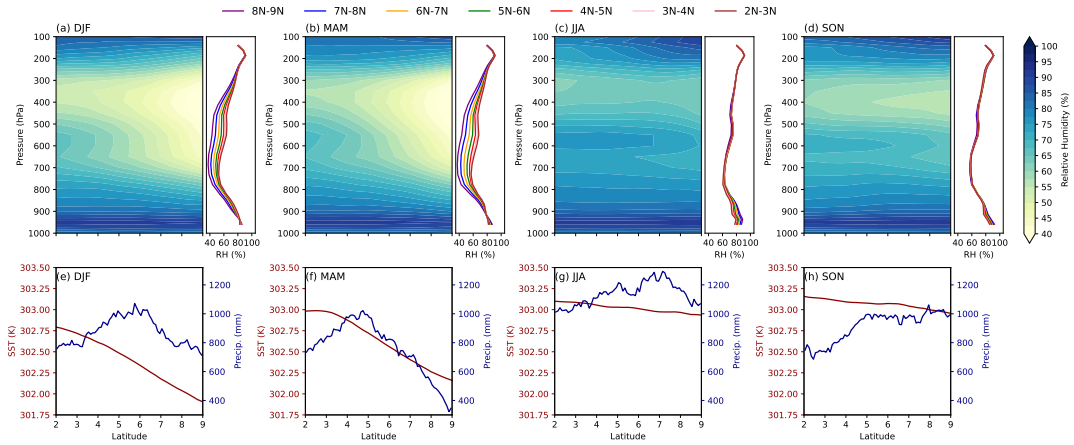

FIGURE 4.2: Mean seasonal variations in (a-b-c-d) relative humidity as a function of latitude and height extracted from ERA5 data. It is important to observe that the right panel illustrates the mean relative humidity as function of height for each latitudinal degree within the domain (indicated by colors and explain in the legend). Additionally, the mean seasonal variations in (e-f-g-h) SST obtained from Himawari with a 24 hour running mean to give a smoothed version, depicted in red and corresponding to the left y-axis, are presented alongside rainfall data derived from GPM, shown in blue and associated with the right y-axis as a function of latitude.

three regions.

4.4.2 Subseasonal Variations of Water Vapour and Organization Regimes

Having established the key differences in meridional seasonal means, we turn our attention to the evolution of sub-seasonal variability in water vapor and its correlation with convection and SST (refer to methods for details). We focus on two distinct periods, one in boreal summer and the other in spring, chosen for their representation of the extremes in background humidity gradients.

During boreal summer, when our focal region is positioned within the main warm pool region, we validate that mesoscale SST gradients remain limited, even over shorter time spans (Figure 4.3). Temporal alterations in the distribution of TCWV are minimal, with the 1st percentile of TCWV rarely surpassing the 48 kg m^{-2} threshold that demarcates the boundary between deep convective and non-convective regions in the deep tropics in the present climate (Mapes et al., 2018). In this regime, the driest regions often coincide with the SST warm perturbations within the domain, where SST gradients are small (Figure 4.3c). Convection consistently aligns with the moistest columns, a pattern consistent with prior analyses (shading in Figure 4.3a) (Bretherton et al., 2004; Holloway et al., 2012; Rushley et al., 2018). Generally, periods characterized by the highest mean rainfall (Figure 4.3d) correspond to the largest maximum humidity values within the domain.

This scenario, characterized by constrained humidity and SST spatial variability and an anti-correlation between SST and TCWV, closely mirrors the characteristics of randomly organized convection observed in idealized CRM experiments. In such

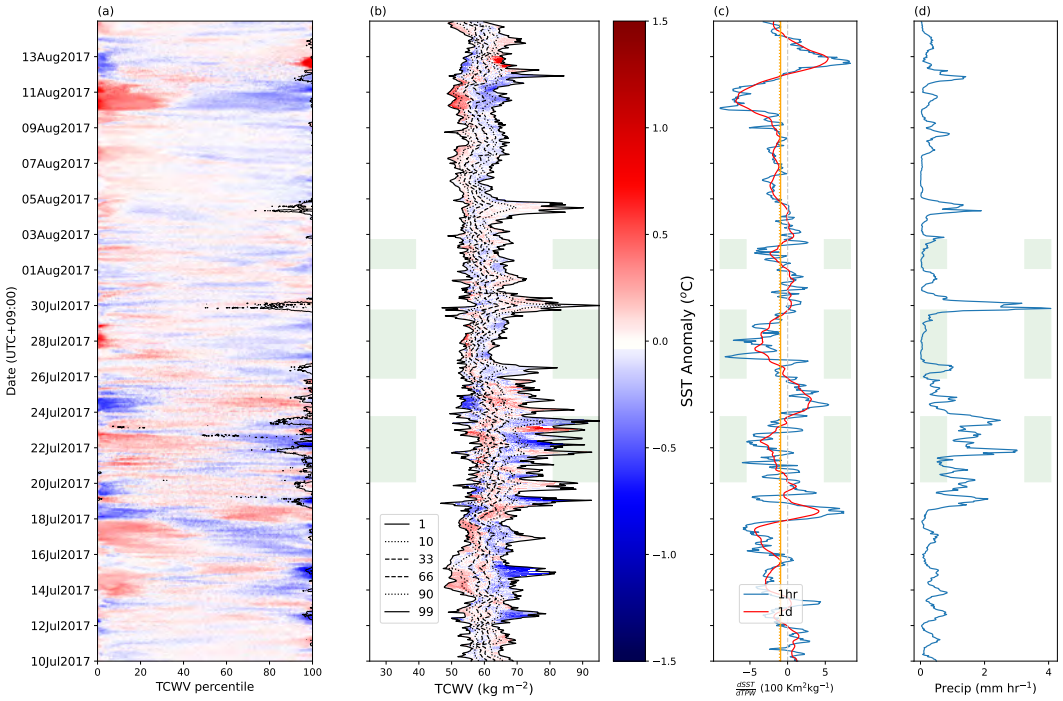

FIGURE 4.3: For the 2N-9N, 135E-145E study region during 2017-06-20 to 2017-08-10. (a) Hovmöller plot of NOAAOI SST anomaly (depicted by colors) against MIMIC TCWV retrieval percentile, featuring overlaid black contours indicating areas of GPM IMERG precipitation at 5 mm hr^{-1}. (b) Hovmöller plot of MIMIC SST anomaly (represented by colors) relative to absolute TCWV, clipped between the 1st and 99th percentiles to eliminate anomalous extremes, with percentile values depicted as contours. (c) Time series of the SST-TCWV regression. (d) Domain mean GPM IMERG precipitation rate. Green shading in panels b-d highlights reversal events, signifying instances where the SST-TCWV regression shifts from positive to negative for a duration of at least 24 hours.

experiments, convection is distributed across the entire domain, ensuring local moisture sources that prevent any significant drying of specific regions (Figure 4.5a-b).

The situation undergoes a significant transformation in the boreal spring, as illustrated by the example in Figure 4.4 during April/May 2017 (see Figure B.4 in Appendix B). Two distinct regimes become evident in this analysis. For the majority of the period, humidity variance surpasses that observed in summer, with the driest region of the domain registering values below 30 kg m^{-2}. In these instances, the relationship with SST reveals that the coldest SSTs align with the locations of the driest columns, while, consistent with the summer pattern, convection and rainfall are linked to the moistest columns. SST variance also exhibits an increase. In contrast to the summer scenario, this pattern, characterized by heightened SST and TCWV variance and a positive correlation between the two fields, resembles a situation of aggregated deep convection, where convection is confined to a specific part of the domain.

Although the aggregated state is the predominant condition during the depicted spring period, there are distinct multi-day "events", two of which are highlighted

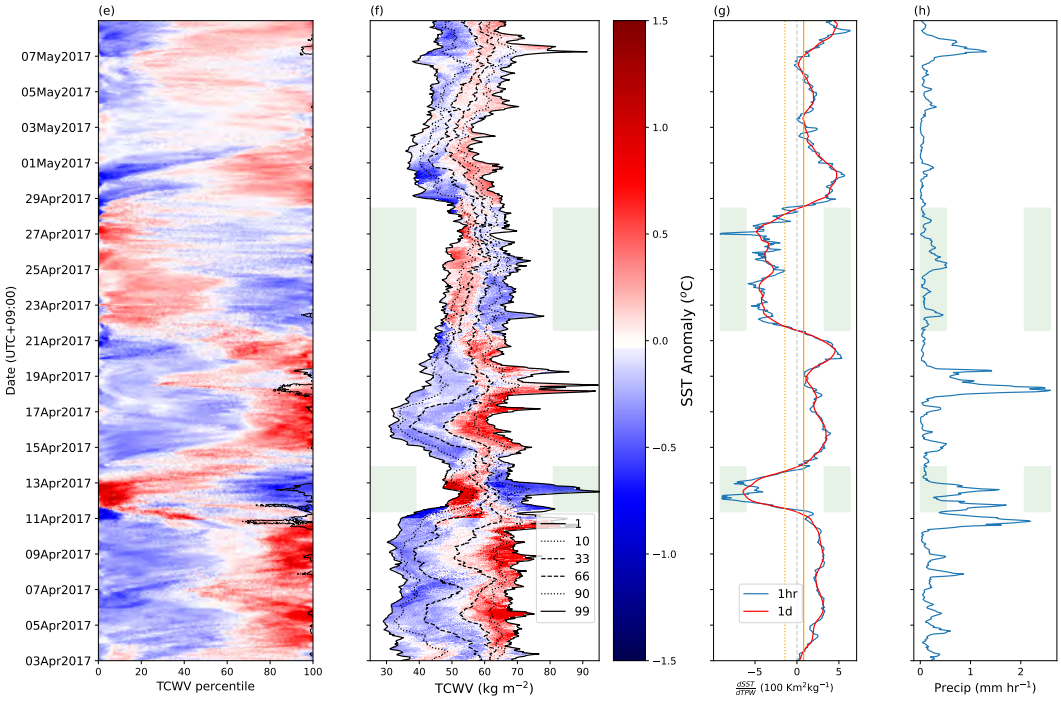

FIGURE 4.4: For the 2N-9N, 135E-145E study region during 2017-04-01 to 2017-05-22. (a) Hovmöller plot of NOAAOI SST anomaly (depicted by colors) against MIMIC TCWV retrieval percentile, featuring overlaid black contours indicating areas of GPM IMERG precipitation at 5 mm hr^{-1}. (b) Hovmöller plot of MIMIC SST anomaly (represented by colors) relative to absolute TCWV, clipped between the 1st and 99th percentiles to eliminate anomalous extremes, with percentile values depicted as contours. (c) Time series of the SST-TCWV regression. (d) Domain mean GPM IMERG precipitation rate. Green shading in panels b-d highlights reversal events, signifying instances where the SST-TCWV regression shifts from positive to negative for a duration of at least 24 hours.

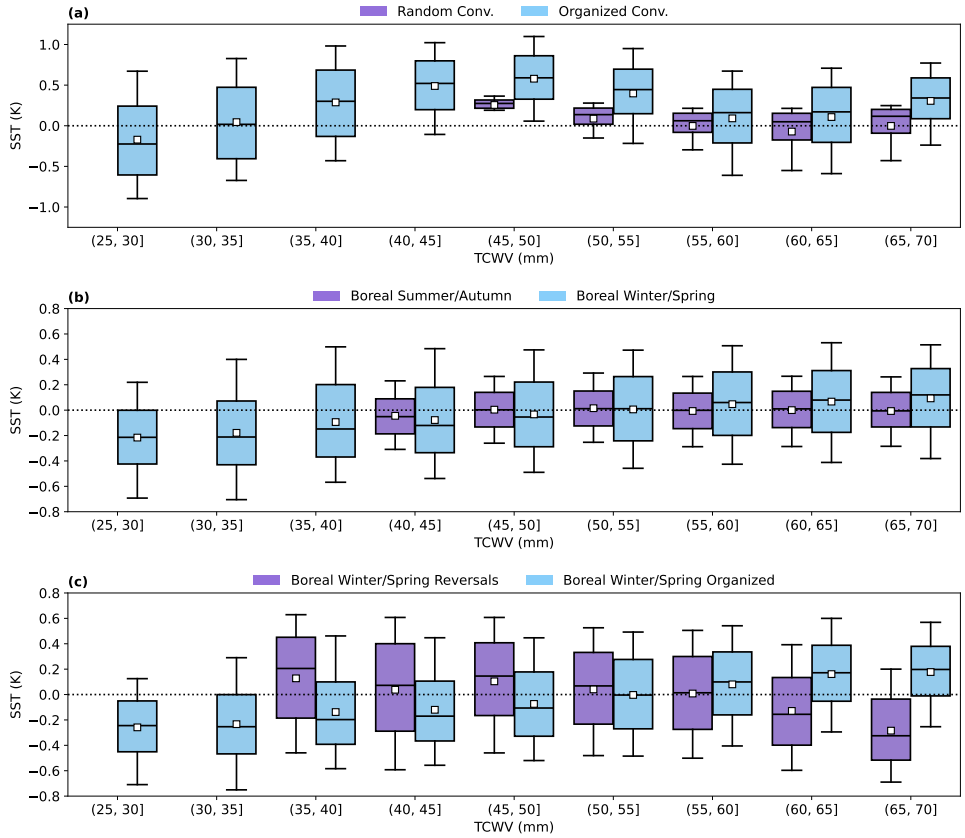

FIGURE 4.5: Boxplot representations of SST anomalies are categorized based on TCWV for: (a) a random and an organized state from results of the slab ocean model of Tompkins and Semie (2021), (b) boreal summer/autumn and winter/spring periods, and (c) boreal winter/spring reversals and organized regimes, using Himawari and MIMIC datasets. Consistent results are observed with NOAA OISST and ERA5 datasets (not-shown). In each boxplot, the lower end represents the 25th percentile, the upper end shows the 75th percentile, and the middle line indicates the median (50th percentile). The whiskers extend to the 10th and 90th percentiles. The white square inside the box represent the mean of the data.

with green shading. These events witness substantial changes in humidity distributions, marked by a rapid reduction in spatial humidity variability—a key metric widely employed for assessing convective aggregation. During these events, the driest regions in the domain undergo a swift moistening of at least 20 kg m^{-2}, while the moistest regions associated with convection remain relatively stable. This outcome aligns with expectations, as these values are consistent with an approximately saturated moist adiabatic process, with the lower boundary defined by the warmest SST values. At the same time the SST-humidity relationship undergoes a reversal, with the warmest SSTs being situated over the driest regions of the domain. These occurrences, termed "reversals," persist for several days in this example before returning to the typical aggregated-like pattern. These reversal events exhibit characteristics reminiscent of the random convective situations prevalent during the summer period.

An examination of Hovmöller plots for other months (not shown) indicates that

convection in this specific region is randomly distributed when SST gradients are limited from June to November. However, during the other months, characterized by a weak zonal SST gradient and intermittent episodes of randomly organized convection, the SST-humidity relationship becomes inverted or reversed.

For a more systematic analysis spanning the entire 6-year period during the boreal winter/spring months, Figure 4.6 presents the joint probability density function of the standard deviation of *TCWV* (Figure 4.6a-c) and the 5th percentile of *TCWV* (indicative of the driest areas in the domain, Figure 4.6b-d), both in comparison with the SST-TCWV regression relationship. The joint PDF reveals that during periods of reversals when the regression relationship is 0.5*std smaller than the mean, σ(TCWV) consistently exhibits low values, signifying homogeneous convection. This is further emphasized by the remarkably high minimum value within the domain. The vertical lines in the plot indicate the slope threshold used to distinguish between random and clustered convection. The horizontal line represents a TCWV value of 48 kg m^{-2}, a threshold identified by Mapes et al. (2018) to discern convecting from non-convecting regions in the deep tropics. For nearly all scenes characterized by a noticeable random/reversal regime, very few points exhibit drier conditions than this threshold within the domain. Conversely, all scenes featuring clustered convection display a positive slope, with the majority of points situated in the drier regions within the domain.

Hence, during the boreal winter-spring months (see Figure 4.6a-b), scenes characterized by a reversed SST-TCWV regression—indicating that the warmest SSTs are beneath the driest columns—are consistently associated with homogeneously arranged convection and uniformly distributed water vapor. In contrast, when convection is clustered, nearly all points fall below the 48 kg m^{-2} threshold and exhibit positive slope values. The limited number of points below 48 kg m^{-2} and with a reverse/random slope likely corresponds to transition periods, during which convection shifts from a random to an organized state. This observation reinforces the notion that convection is primarily situated above the warmest SSTs, with clear skies prevailing over the driest and coolest regions.

In the boreal summer/autumn months (Figure 4.6c-d), TCWV variance diminishes significantly compared to the boreal winter/spring months, aligning with a regime of random convection observed in idealized models (Müller and Hohenegger, 2020b; Biagioli and Tompkins, 2023a). Regarding the 48 kg m^{-2} threshold line, nearly no points fall below this threshold, especially when considering the random part of Figure 4.6d. This indicates that the warmest SST regions align with the driest areas of the domain, consistent with random conditions in idealized models (Shamekh et al., 2020a; Tompkins and Semie, 2021).

4.4.3 Impacts of reversal events

To delineate and assess the reversal events, quantify their frequency, evaluate their impact on the energy budget, and identify their underlying causes, we employed the criterion of a SST-TCWV relationship persisting above/below half standard deviation of the mean for more than 24 consecutive hours (see methods). Our analysis identified a total of 44 "reversal" events for the 2N-9N region, 46 events for the 3N-10N area, and 41 events on the equatorial box, during the boreal winter/spring

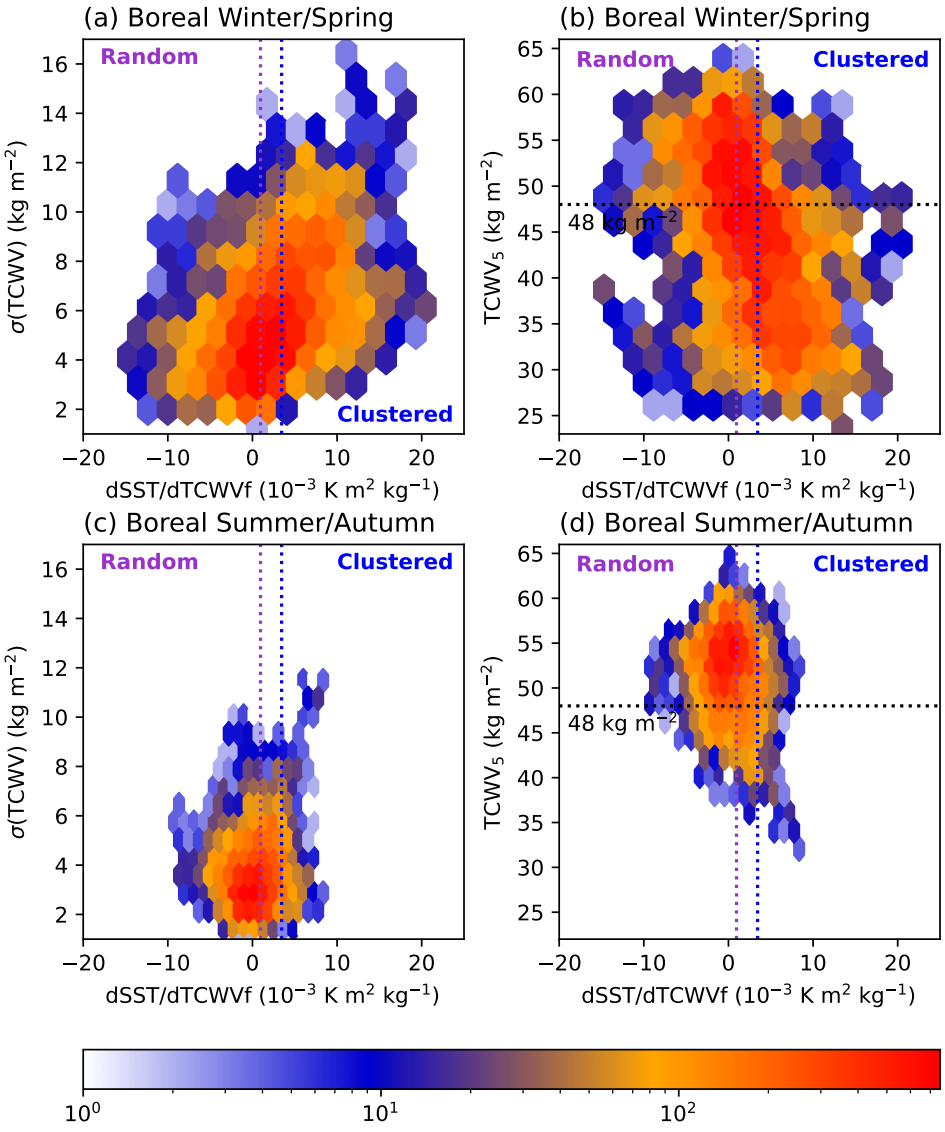

FIGURE 4.6: Joint Probability Density Function (PDF) plots for (a-b) Boreal Winter/Spring and (c-d) Summer/Autumn, illustrating (a-c) $\delta(TCWV)$ and $\frac{dSST}{dTCWVf}$, and (b-d) TCWV 5th percentile and $\frac{dSST}{dTCWVf}$. Hexagons colors represent the distribution of data points in each bin. The purple vertical dashed line denotes the upper boundary threshold for considering a random event, while the blue vertical dashed line indicates the lower boundary threshold for clustered convection. The black horizontal dashed line represents the 48 kg m^{-2} TCWV threshold distinguishing convective and non-convective regions, as identified by Mapes et al. (2018).

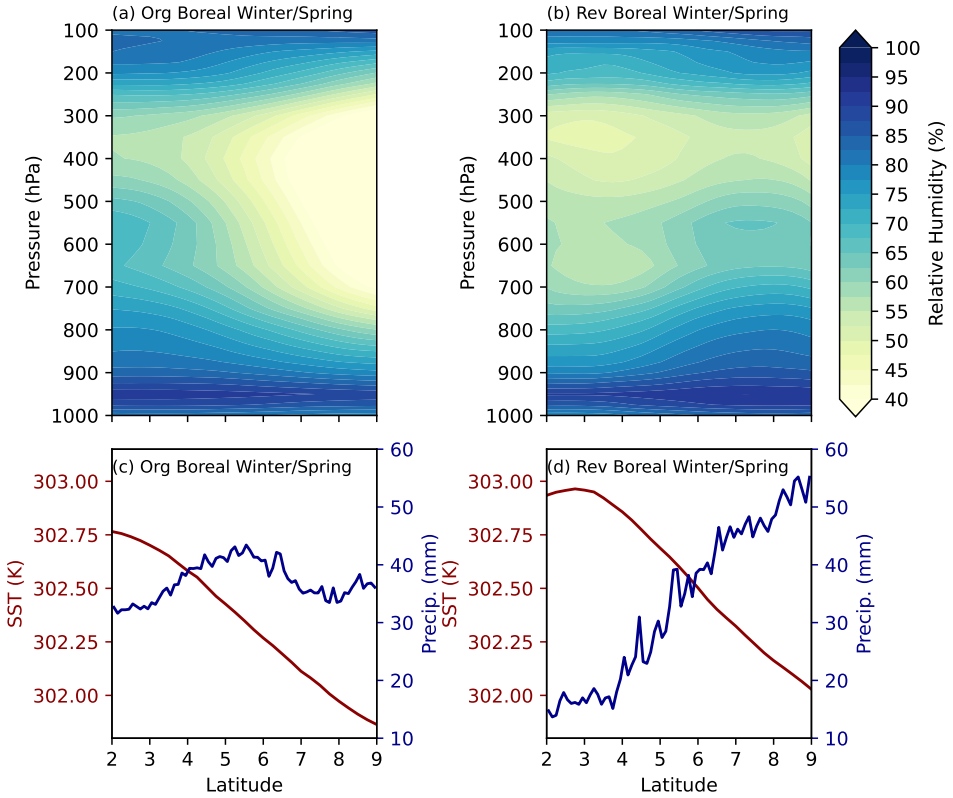

FIGURE 4.7: (a-b) ERA5 relative humidity as function of latitude and height and (c-d) Himawari SST (left y-axis) and GPM precipitation (right y-axis) as a function of latitude, for the (a-c) organized and (b-d) reversal regimes of all boreal winter/spring seasons.

period spanning from 2017 to 2022. Averaging over these reversal and clustered periods allows us to elucidate the impact of these events and investigate their origin.

The reversal events notably influence the meridional distribution of RH within the domain, as illustrated by dividing the boreal winter/spring period into organized and reversal phases (Figure 4.7). Throughout the majority of the period when the SST-TCWV correlation does not indicate a reversal, a pronounced north-south gradient of RH prevails (see Figure B.5a in Appendix B for a composite map of organized events in boreal winter-spring months). In this scenario, free tropospheric RH descends to as low as 40% above the boundary layer, extending northward from 6N. This low humidity is attributed to the absence of deep convection-induced moistening, occurring at latitudes with the coolest SST, albeit with a meridional SST gradient of less than 1 K across the 770 km domain. Precipitation peaks just north of the equator at 5N in this period. Instead, during reversal episodes (Figure 4.7b-d), the precipitation peak shifts further north to 8N, and the water vapor variance is constrained, resulting in a generally moister domain. The robust anticorrelation between SST and precipitation becomes evident during these reversal events (see Figures B.5b in Appendix B for a composite map of reversals and Figure B.6 for an example of a reversal episode in the boreal winter-spring months).

The uniform humidity distribution (Figure 4.8d-h) significantly influences the

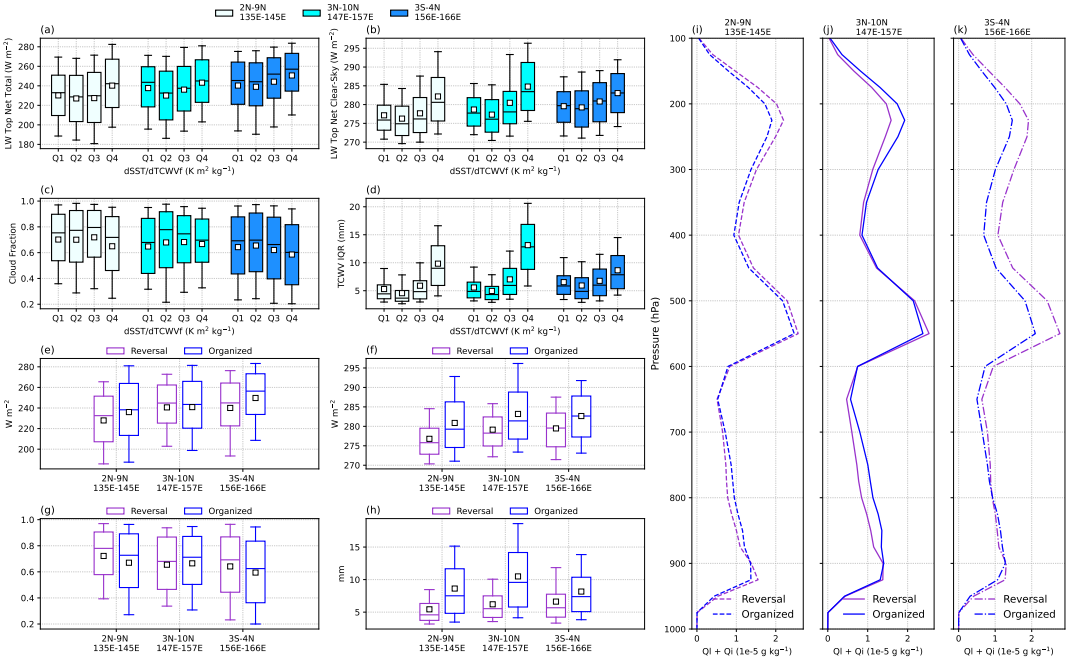

FIGURE 4.8: (a) All-sky TOA all-sky OLR, (b) TOA OLR, (c) cloud fraction, and (d) Interquartile Range (IQR) of TCWV as a function of $\frac{dSST}{dTCWVf}$ quartile for each study region (represented by box color). For reversal and organized days, (e) all-sky TOA OLR, (f) clear-sky TOA OLR, (g) cloud fraction, (h) IQR of TCWV, liquid water, and ice mixing ratio are shown for (i) 2N-9N and 135E-145E, (j) 3N-10 and 147E-157E, and (k) 3S-4N and 156E-166E. Each boxplot displays the 25th percentile at the lower end, the 75th percentile at the upper end, and the median (50th percentile) as the middle line. Whiskers extend to the 10th and 90th percentiles, with a white square inside the box representing the mean of the data.

energy balance, causing a distinct reduction in clear-sky TOA OLR during the reversal period (Figure 4.8b-f). This reduction is primarily due to increased atmospheric opacity, as confirmed by RRTM model calculations in Figure B.7 (in Appendix B), emphasizing humidity's primary role in clear-sky TOA OLR changes. Conversely, all-sky TOA OLR (Figure 4.8a-e) also shows a statistically significant decrease, mainly driven by cloud properties, as shown in Figure 4.9, highlighting clouds' substantial contribution to TOA OLR differences. Both clear-sky and all-sky OLR experience a marked decrease exceeding 10 W m^{-2} (refer to Figure B.8 in Appendix B for region-specific significant differences). This consistent pattern holds for the 2N-9N and equatorial regions. In the 3N-10N area, clear-sky LW increases notably, while all-sky LW shows subtle changes due to a stable cloud fraction in this region. Despite this, the observed decrease in LW is partly attributed to reversal periods promoting a more homogeneous and moist domain, thereby reducing OLR. Although statistically significant, this impact is less pronounced compared to the contrast between aggregated and non-aggregated states in CRMs.

The prominence of cloud properties as the primary contributors to significant differences in all-sky TOA OLR during clustered convection raises the question: What specific cloud characteristics drive these changes? Unraveling this complexity is challenging due to the strongly non-linear nature of cloud properties (Bony et al.,

2020), preventing a straightforward decomposition of their contributions. Given that the RRTM model utilizes cloud fraction, liquid water, and ice mixing ratios inside the cloud, we can gain insights into the importance of each element by examining their behavior.

Figure 4.8i-j-k illustrates the sum of liquid water and ice mixing ratios as a function of height for both reversal and organized regimes. Notably, clouds appear thicker during reversal events at both high and low levels, with the difference being less pronounced in low clouds. This suggests that the contribution of liquid water and ice mixing ratios is not of zero order, as they act to decrease the all-sky TOA OLR difference. Consequently, we posit that cloud fraction (Figure 4.8c-g) most probably from low clouds, since high clouds seem thick, stands out as the most crucial contributor to all-sky TOA OLR differences. This conclusion is supported by the observation that the two regions exhibiting a significant increase in OLR experience a decrease in cloud fraction, while the region with no change in cloud fraction also demonstrates a minimal and statistically insignificant alteration in TOA OLR.

In summary, our findings suggest that the increase in all-sky TOA OLR observed during days with organized convection is primarily linked to changes in cloud fraction most probably from low clouds, rather than variations in the liquid and ice mixing ratios within the clouds.

4.4.4 Understanding Reversal Events

What is driving the reversals? In idealized simulations of RCE, where all boundary conditions are uniform, studies have elucidated the essential role of diabatic feedbacks in fostering organization. Specifically, cloud-LW interactions take precedence, primarily influencing the pre-onset stage, with an additional contribution from surface LH flux feedback (Muller and Held, 2012; Wing and Emanuel, 2014; Tompkins and Semie, 2021). In line with this approach, we organize ERA5 reanalysis flux anomalies based on TCWV, adopting a technique akin to that employed by Wing and Emanuel (2014) and Holloway and Woolnough (2016). Positive flux anomalies in the moistest columns signify a propensity to augment moist static energy variance and propel aggregation. We segment this diabatic feedback analysis between boreal winter/spring and summer/autumn periods (Figure 4.10a-b). Further, we dissect the winter/spring period into organized and reversal periods (Figure 4.10c-d).

The LW cloud and LW clear-sky feedbacks play crucial roles in organizing convection (Wing et al., 2017), with these factors generating the strongest feedbacks (Tompkins and Craig, 1998a; Stephens et al., 2008; Holloway and Woolnough, 2016; Holloway et al., 2017). However, in the progression of reversals, the LW clear-sky feedback counters organization by causing the less cloudy regions to move towards the warmest SSTs. This movement induces regional heating, creating a negative feedback with the organization process (refer to Figures B.9 and B.10 in Appendix B). Simultaneously, anomalies in SW cloud heating remain negative during both periods, as high clouds attenuate incoming solar radiation.

The LH flux heating anomalies observed here align closely with findings from previous idealized and observational research (Wing et al., 2017; Holloway et al.,

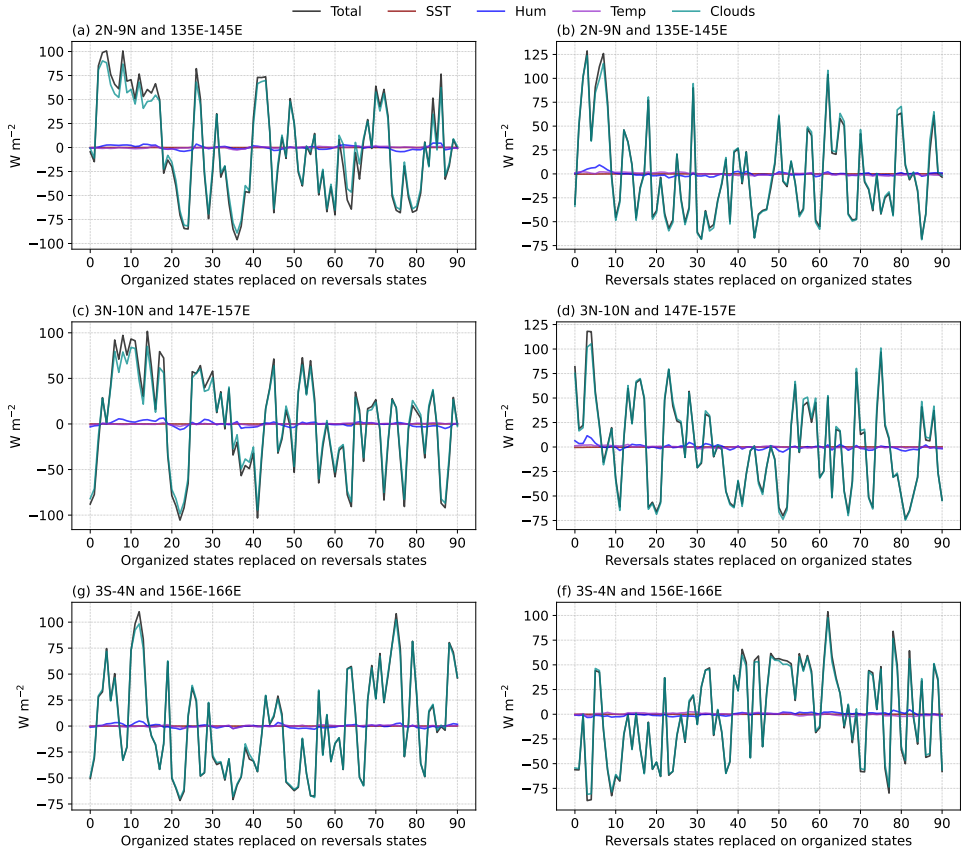

FIGURE 4.9: Contributions to all-sky TOA OLR from SST, specific humidity (Hum), temperature, and cloud components (liquid water, ice, and cloud fraction) for (a-c-e) organized days replaced with reversal events and for (b-d-f) reversal days replaced with organized events. The depicted regions are (a-b) 2N-9N and 135E-145E, (c-d) 3N-10 and 147E-157E, and (e-f) 3S-4N and 156E-166E. The x-axis represents the days on which the replacement was performed.

2017). Surface fluxes play a reinforcing role in organized convection when in its random state and disfavor it when already organized. We dissect the LH flux anomalies into contributions from humidity, stability, and wind (Figure 4.11). Tompkins and Semie (2021) found a LH flux positive feedback during random convection primarily stemming from temperature and, to a lesser extent, wind speed, while the negative feedback arises from humidity structures. Our results align with theirs regarding the sign of LH flux feedbacks. However, a notable distinction emerges: the wind-induced surface heat exchange (WISHE) feedback overwhelmingly governs both negative and positive feedbacks, overshadowing the influence of temperature and humidity structures, which consistently exhibit a different sign compared to the wind speed contribution (Figure 4.11g-h-i). Under organized conditions, slow winds prevail in the convective regions (Wing and Emanuel, 2014). In contrast, during reversal episodes, strong wind speeds characterize the moistest regions within the domain (Figure 4.11g-h-i). This observation aligns with the findings of Retsch et al. (2022), who emphasize the significant role of wind in organizing convection based on observational evidence.

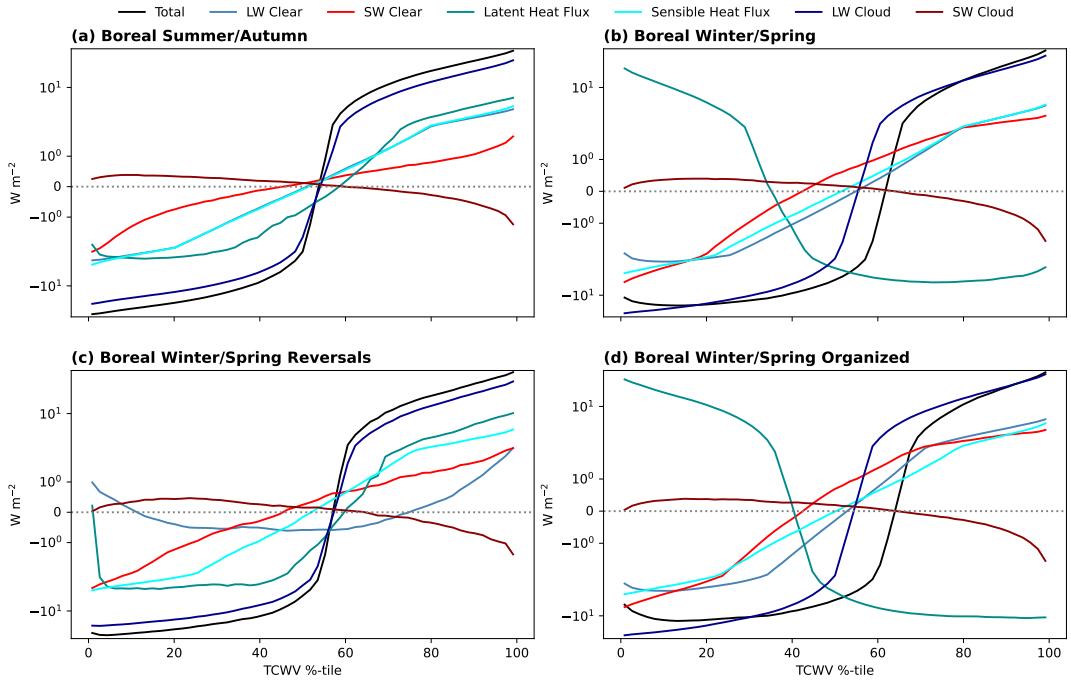

FIGURE 4.10: Mean anomalies of total atmospheric convergence (i.e., diabatic and diabatic feedbacks), ordered by TCWV for (a) boreal summer/autumn, (b) boreal winter/spring, and for boreal winter/spring (c) reversals and (d) organized episodes. Notice that warming anomalies in the atmosphere are represented by positive values.

As diabatic forcing consistently acts to aggregate convection in these observations, the query arises as to how and why the aggregated convection transitions into a state of random convection characterized by a narrowed humidity distribution in the winter/spring seasons under the influence of a zonal SST gradient. To elucidate this, we construct composites of 44 reversal events in the 2N-9N region, 46 in the 3N-10N area, and 41 in the equatorial box, occurring between 2017 and 2021. The zero hour is identified as the time when the SST-humidity negative correlation is at its minimum within each event (Figure 4.12). The composite, focusing on the main target region at 2N-9N and 135E-145E, reveals a westward-propagating mode of convection initiating to the east of the target region around 160E (see Figures B.11 and B.12 in Appendix B for the same composite analysis but accounting for the other two study regions, yielding similar conclusions). A symmetric convective perturbation develops to the north and south of the equator at approximately 10N and 10S, with convection relatively suppressed on the equator (Yoneyama, 2003). The convective signal propagates westward at a speed of ≈ 4.5 m s^{-1}, accelerating as it reaches the western Pacific (to ≈ 5.5 m s^{-1}), consistent with the findings of Mayta et al. (2022). The propagation speeds, vorticity, OLR anomaly structure, and wavelength align with the characteristics of equatorial convectively coupled Rossby waves (Kiladis and Wheeler, 1995; Kiladis et al., 2009). This conclusion is consistent with Mayta et al. (2022), who demonstrated that westward-propagating Rossby waves also explain a significant portion of cloud fraction variability in this region. These waves represent a moisture mode excited by diabatic feedbacks (Figure 4.11), including wind-induced surface flux enhancement, consistent with Chen (2022). Additionally, alongside constructing the composite using the index, we performed a multivariate

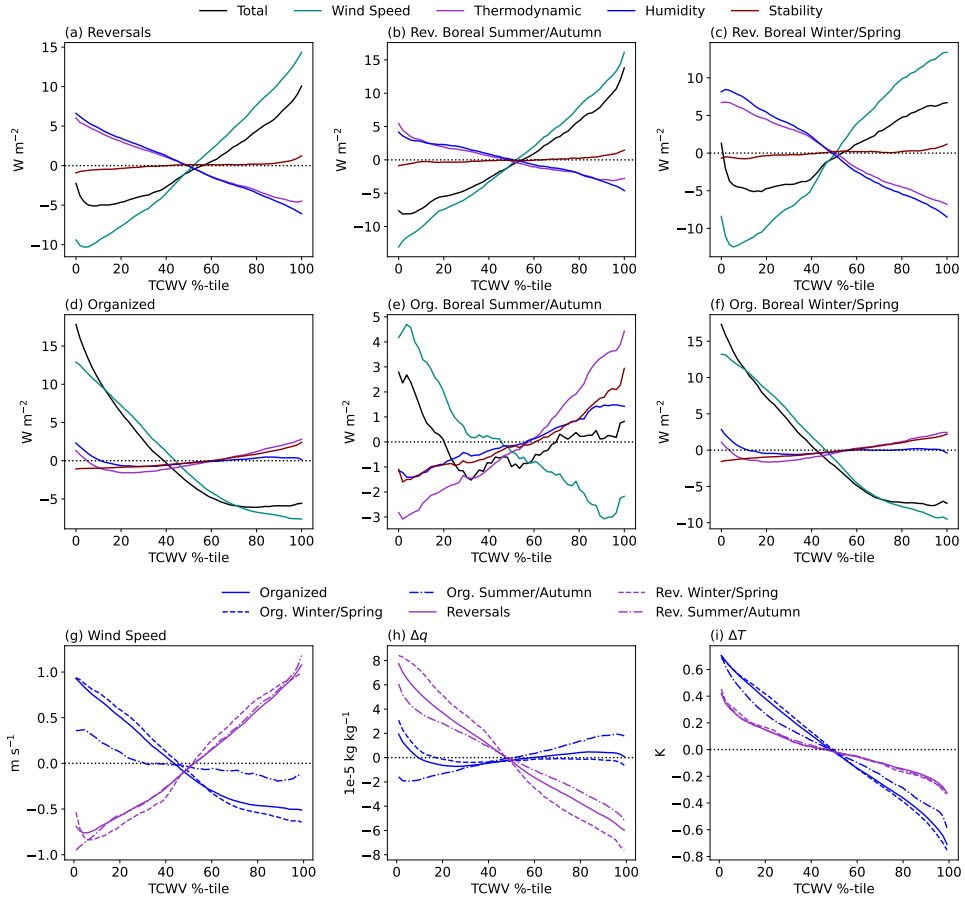

FIGURE 4.11: Contributions to LH surface flux perturbations in the reversal regime are depicted for (a) the entire dataset, and separately for (b) boreal summer-autumn and (c) winter-spring. Similarly, contributions are illustrated for the organized regime, encompassing (c) the entire dataset, and for boreal (b) summer-autumn and (c) winter-spring. These contributions emanate from (g) wind speed perturbations, (h) humidity (Δq), and (i) stability (ΔT) differences. Thermodynamic perturbations refer to the combined contribution of ΔT and Δq terms. The calculations are executed using a random forest (RF) algorithm fitted to the ERA5 dataset (refer to Section 4.3 for details).

REOF analysis of SST, and TCWV, yielding the same conclusions (see Figure B.13 in Appendix B for the 2N-9N region).

Another crucial characteristic of the westward-propagating wave is its ability to induce convergence, altering wind direction and generating a southerly flow (Figure B.14 in Appendix B). This phenomenon has been consistently observed in prior studies focused on warm pool wave dynamics (Takayabu and Nitta, 1993; Takayabu, 1994; Kiladis and Wheeler, 1995). The southerly flow facilitates the transport of moisture from the center of the warm pool into the domain, leading to a displacement of the majority of convection to the north and the initiation of convection to the south of the domain. This process results in homogeneous TCWV conditions and the occurrence of random convection, irrespective of the strong SST gradients. This underscores the significance of wind dynamics in maintaining organization (Retsch et al., 2022).

The key determinant of mesoscale humidity variability in these tropical warm pool study regions seems to be large-scale wave dynamics. Consequently, a comprehensive understanding emerges, where convection during the winter/spring period typically exists in a clustered state characterized by robust meridional gradients of free tropospheric humidity, owing to the presence of a weak north-south SST gradient. However, this clustered state is intermittently disrupted by the passage of westward-propagating Rossby wave as shown in Figure 4.13a-c-e. These induce a more homogeneous convection pattern that generally opposes the SST gradient, and is characterized by a southerly flow that imports humidity within the domain resulting in very uniform distributions of TCWV.

We hypothesize that the aforementioned Equatorial Rossby waves instigate convective anomalies at 10N and 10S, leading to peak convective activity over the coolest SSTs in the domain. This counter-gradient convective activity could have a relation with the TCWV variance of the entire Pacific [$\sigma(TCWV)$]. If there is indeed a positive relation between convective activity and $\sigma(TCWV)$ across the entire Pacific (as the preliminary analysis of Figure 4.13b-d-f suggests), it could imply an intensified export of humidity to the subsiding branches of the Hadley/Walker circulations. Furthermore, we propose that the observed correlation between mesoscale organization and mid-tropospheric humidity, highlighted by Bony et al. (2020), may indicate variability in Rossby wave activity in the tropical western Pacific region.

In a preliminary analysis, we seem to see that the TCWV variance in the area between 3S-4N and 156E-166E shows a correlation with the OLR of the whole tropics. This implies that the controls of $\sigma(TCWV)$ can account for a substantial portion of the annual and seasonal variance in TOA OLR in the Pacific. In the region where the Hadley/Walker circulations are centered, the correlation is notably strong. However, in the western region (between 2N-9N and 135E-145E), although the correlation remains high in the south where convection typically develops, the rest of the region does not exhibit a statistically significant correlation. A more detailed analysis is necessary to understand this correlations, and this will be addressed in the forthcoming paper based on this chapter.

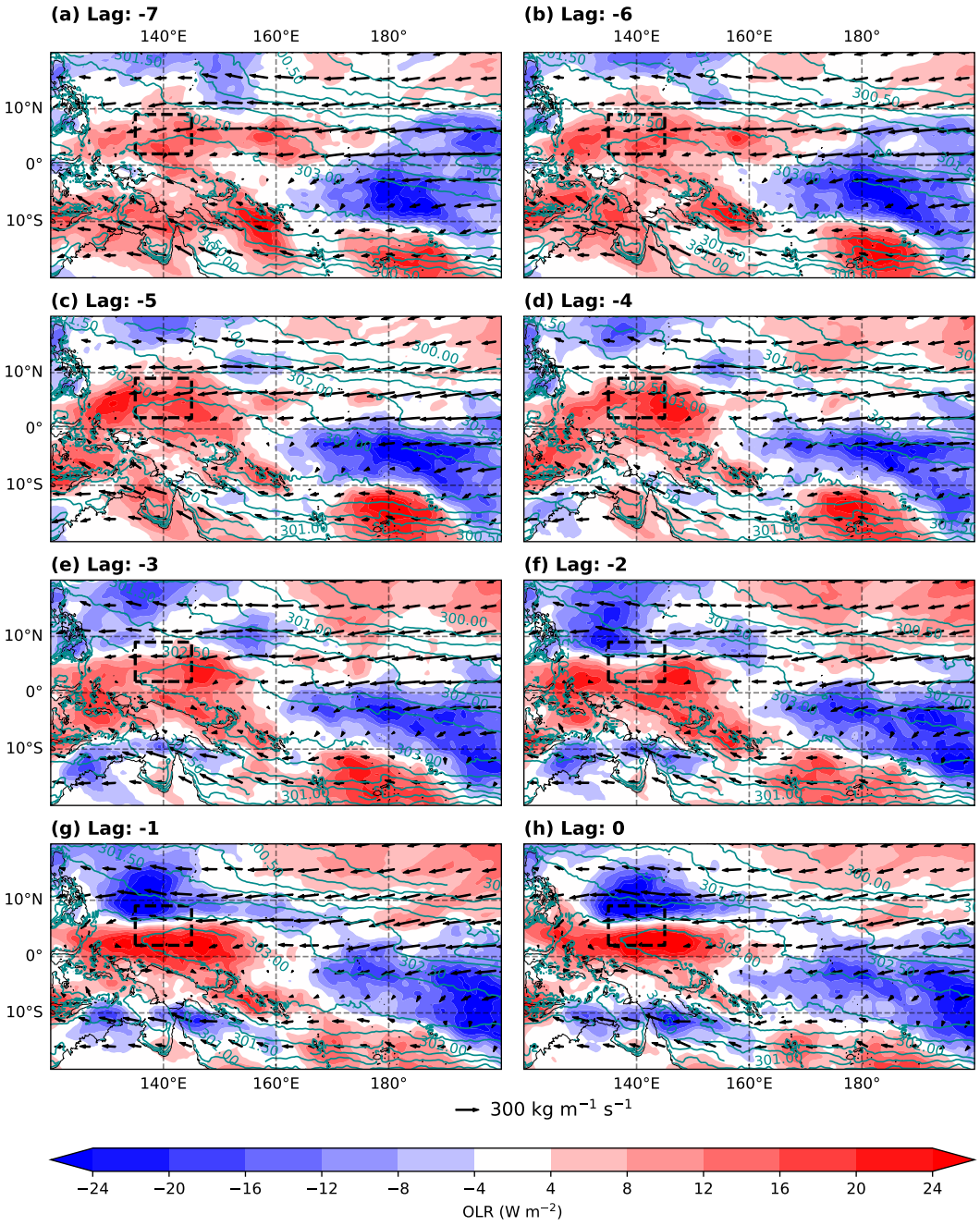

FIGURE 4.12: Composite of 44 reversal states on boreal winter-spring months and their related lags. The colors represent the OLR anomaly for each of the lags. The contours are the SST and the arrows size represent the WVF magnitude and their orientation shows the WVF direction. Notice that each panel indicate the lag in number of days. The dashed rectangle represents the study area between 2N-9N and 135E-145E.

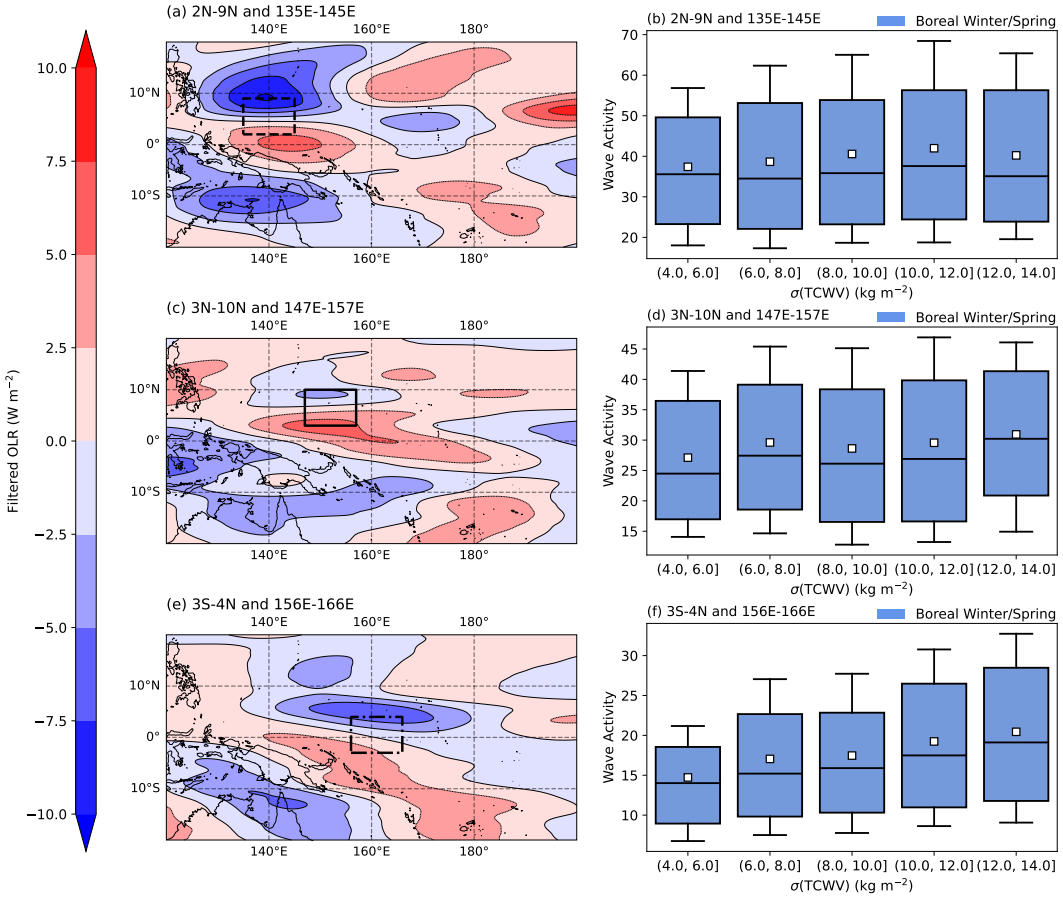

FIGURE 4.13: Equatorial Rossby wave filtered (and deseasonalized) OLR composites of reversal states during boreal winter/spring months. The composites depict (a) 44 events in the 2N-9N and 135E-145E region, (c) 46 episodes in the 3N-10N and 147E-157E zone, and (e) 41 events in the 3S-4N and 156E-166E area. Wave activity for the (b) 2N-9N and 135E-145E, (d) 3N-10N and 147E-157E, and (f) 3S-4N and 156E-166E regions as a function of TCWV variance for the entire Pacific during boreal winter/spring.

4.5 Chapter Summary and Conclusions

Water vapor, a crucial greenhouse gas and regulator of the tropical energy budget, is well-understood in terms of the mean activity of deep convection. However, recent findings by Bony et al. (2020) suggest that variations in convection distribution at scales less than 1000 km, specifically the alpha and beta meso-scales, play a significant role in shaping interannual variations in the water vapor distribution. These variations, coupled with changes in low clouds, appear to have a profound impact on the year-to-year fluctuations in the TOA net radiative budget.

If convection is more clustered at mesoscale, the local atmosphere is drier, which in turn impacts the mid-tropospheric humidity in the descending branches of the Walker and Hadley circulations, increasing the net infra-red flux to space averaged over the tropics and subtropics. Understanding the first order controls of convection organization and water vapor variability in the ITCZ is crucial. To address this, we employed a simple multivariate analysis commonly used in idealized cloud-resolving studies to assess the impacts of convective organization. Idealized studies

reveal that in a clustered state, the TCWV distribution widens due to the presence of dry columns, and SST and water vapor amounts exhibit a positive correlation. Conversely, in a more randomly organized convection, the water vapor distribution is narrow, with high mean values, and SST is negatively correlated with TCWV. As a case study, we examine these relationships in a domain north of the equator in the warm pool region of the tropical western Pacific.

Focusing on mesoscale domains of approximately O(1e6) km^2 in the warm pool region, our analysis indicates that during summer/autumn months, when spatial SST gradients are small, the multivariate signature aligns strongly with randomly organized convection. This occurs despite diabatic forcings consistently acting to cluster convection and spatial humidity gradients remaining small. In contrast, during boreal winter/spring months when the warmest SSTs shift southward, and our study regions experience a weak north-south SST gradient, the analysis suggests a predominantly aggregated convection state with notable mesoscale humidity variability. However, this aggregated state is intermittently interrupted by episodes where convection transitions to a random state, characterized by a moister atmosphere and reduced humidity gradients.

Using the SST-TCWV correlation as an index for the clustered state, we constructed a composite of random events over our study periods. Our analysis of diabatic forcings revealed consistent behaviors in radiative-cloud and moisture feedbacks, which act to cluster convection, while the surface LH feedback promotes clustering when convection is in a random state. However, the latter feedback reverses its role once convection becomes aggregated, aligning with findings from idealized models.

Despite radiative feedbacks favoring convection aggregation, sporadic breakup of clustering during the boreal winter/spring period, in spite of the weak meridional SST gradient, is attributed to westward propagating, convectively coupled equatorial Rossby waves. These waves initiate convective anomalies at 10N and 10S, resulting in peak convective activity occurring over the coolest SSTs in the domain. This counter-gradient convective activity results in a more uniform humidity distribution, leading to a notable impact on TOA OLR, with a decrease of approximately 10 W m^{-2} attributed to reduced humidity and cloud cover. In regions where cloud cover remains relatively stable, the changes in TOA OLR are not statistically significant. The broader zone of convective activity, with TCWV exceeding 48 kg m^{-2} from the equator to 9N, suggests increased export of humidity to the subsiding branches of the Hadley/Walker circulations. The observed connection between mesoscale organization, quantified by the I_{org} index, and mid-tropospheric humidity as identified in Bony et al. (2020), may indicate variations in westward propagating wave activity within the tropical western Pacific. A preliminary examination supports the notion that the TOA OLR in the Pacific could be predominantly influenced by wave dynamics and the arrangement of convection.

Chapter 5

From Observations to a Realistic Simulations: the Mechanisms that Develop and Prevent Organization in Nature

5.1 Abstract

This study examines the spatial organization of convection and humidity in a 1000 km x 700 km region of the Pacific Warm Pool lying north of the equator. In boreal winter/spring, clustered convection prevails, and westward propagating waves disrupt convection clusters, impacting wind patterns and large-scale moisture structures. In contrast, boreal summer/autumn experiences fully random convection despite conducive diabatic feedbacks. This research explores the factors influencing organized and random convection, particularly during boreal winter/spring breakup episodes. Utilizing advanced satellite observations and reanalysis data, we evaluate variables that contribute to or hinder clustering. Cloud-resolving models and machine learning aid in conducting sensitivity experiments by manipulating large-scale conditions. Our findings highlight moisture advection as the critical force steering convection states, determining either randomness or organization. Breakup episodes, characterized by southerly flows importing moisture, redirect deep convection northward, inducing new convection in the southern region, promoting random convection. Organized conditions involve a weak northerly flow, retaining convection in the southern region. While large-scale humidity primarily contributes to clustering, it fails to prevent breakup episodes, emphasizing the crucial role of the wind. Low-level meridional winds emerge as significant contributors to strong breakup episodes, even capable of inducing prolonged episodes through continual moisture import from the south. Alternatively, consistently imposed southward winds extend clustered convection periods, yet breakup episodes persist, highlighting the role of large-scale humidity structures. Surprisingly, zonal wind exhibits a feeble role in clustering, random convection, and breakup episodes. This investigation sheds light on the interplay among large-scale moisture, wind, and convection dynamics in shaping convection states.

5.2 Introduction

The phenomenon of convective self-aggregation has been a subject of extensive research since its initial discovery by Held et al. (1993) in radiative convective equilibrium (RCE) simulations. A multitude of studies have focused into understanding

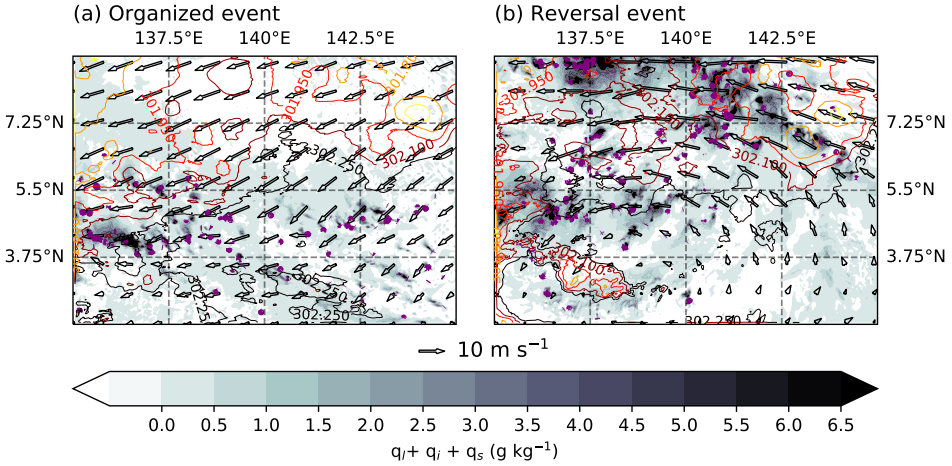

FIGURE 5.1: **Graphical Abstract:** Snapshots of organized and reversal states. Clouds, depicted in colors, are accompanied by SST contours, wind arrows, and convection cores marked by purple points (Vertical velocity > 1 m s^{-1}). **Organized:** Clouds are arranged in a line at the southern edge of the domain, aligning with the warmest SST. Northeasterly winds prevail, facilitating convection confinement to the southern domain. **Reversal:** Clouds are disperse across the entire domain, with a notable concentration in colder regions. This dispersion is linked to a moisture mode triggered by a Rossby wave, fostering a southerly flow that transports moisture from the south into the domain. This influx of moisture initiates new convection in the southern region, marked by the presence of low clouds.

the underlying mechanisms driving clustering (Nakajima and Matsuno, 1988; Held et al., 1993; Muller and Held, 2012; Wing and Emanuel, 2014; Muller and Bony, 2015; Coppin and Bony, 2015; Tompkins and Semie, 2017) and its broader implications for climate (Emanuel et al., 2014; Coppin and Bony, 2015; Holloway and Woolnough, 2016; Wing and Cronin, 2016; Wing et al., 2020; Becker and Wing, 2020), the hydrological cycle, particularly in terms of changes in domain humidity (Bretherton et al., 2005; Müller and Hohenegger, 2020b; Biagioli and Tompkins, 2023a,b), and extreme precipitation (e.g., Muller 2013; Pendergrass 2020; Da Silva et al. 2021b). Recent studies have expanded the scope by incorporating sea surface temperature (SST) gradients and interactive slab ocean models (e.g., Hohenegger and Stevens 2016b; Shamekh et al. 2020a,c; Tompkins and Semie 2021) to investigate how aggregation develops in a more realistic setup. These studies found that SST gradients favor clustering, while a slab ocean delays the onset of organization due to a stronger negative feedback from latent heat (LH) flux compared to non-slab ocean simulations. However, the longwave (LW) and shortwave (SW) feedbacks act similarly to simulations with constant SSTs, with the LW cloud and clear-sky components and the SW clear-sky feedback strongly promoting organization, while the SW cloud feedback opposes it (see Wing et al. 2017; Wing 2019; Muller et al. 2022a for comprehensive reviews).

A number of investigations focus on the impact of convective organization, revealing a consensus on its significance for the hydrological cycle and thus climate (Wing et al., 2017; Holloway et al., 2017; Becker and Wing, 2020). Recognizing its importance in RCE simulations, various observational studies aimed to ascertain

whether organization has comparable effects in nature as seen in idealized models (see the review by Holloway et al. 2017). Pioneering efforts by Tobin et al. (2012, 2013) quantified organization's influence on mean properties, uncovering increased outgoing LW radiation (OLR) and decreased humidity with fewer cores, consistent with idealized modeling studies. Their conclusions align with other studies (e.g., Brune et al. 2018; Popp and Bony 2019; Brune et al. 2020; Pendergrass 2020; Beucler et al. 2020; Retsch et al. 2020; Bläckberg and Singh 2022; Shamekh et al. 2023; Vogel et al. 2022), indicating that LW feedbacks favor organization, whereas LH flux supports it before the onset of clustered convection but acts against it when the convection is already organized, albeit with diminished strength compared to the modeling outcomes (Beucler et al., 2019).

The impact of convective organization on humidity also has implications for atmospheric opacity, with drier atmospheres expected to be less opaque, leading to an increase in LW clear-sky radiation. This radiative effect of organization was investigated by Bony et al. (2020), revealing its significance in the radiation budget of the large-scale tropics, consistent also with the work of Angulo-Umana and Kim (2023). Bony et al. (2020) showed that convective organization and lower-tropospheric stability jointly account for over 60% of the variance in the net radiative budget across the tropics. This shows alignment between observation and idealized simulations, and suggests that organization can modify the radiative budget, influencing climate sensitivity. This raises questions about whether the sensitivity of humidity and the radiative budget to organization remains consistent at GCM grid-box scales. Additionally, it prompts inquiries into whether the mechanisms orchestrating convection at GCM scales align with those observed in CRMs.

To study this question, Chapter 4 (hereafter Chap4) focused on alpha-mesoscale (O(1000km) sized) regions within the Pacific Warm Pool (Figure 5.2), chosen for weak SST, precipitation and humidity gradients, and proximity to the study area of Tobin et al. (2012, 2013). Chap4 revealed that, despite strong diabatic feedbacks favoring clustering throughout the year, convection maintains a random configuration in boreal summer/autumn. In contrast, in boreal winter/spring, weak meridonal SST gradients induce organized convection, resulting in drier mean conditions, higher OLR, and diminished cloud cover. During this season, this situation is intermittently broken up in multiday episodes where the mesoscale humidity is much more well-mixed spatially, indicative of random convection, which were demonstated to be due to westward-propagating waves. This raises questions concerning the mechanisms leading to the well-mixed humidity when the wave traverses the domain—subjects of interest in this study.

The aim of this study is to further examine these reversal events, using both dynamical models and machine learning (ML) to understand the sensitivity to the larger scale properties in which the domain is embedded. To begin, section 5.3 provides a brief observational analysis, building upon Chap4 results, outlining research objectives and hypotheses. In Section 5.4, we detail the datasets, realistic modeling setup, and parameterizations used, contributing to a more comprehensive understanding of organization, wave dynamics, and adding to the literature employing realistic setups for such studies (e.g., Holloway et al. 2017; Brune et al. 2018; Pscheidt et al. 2019; Brune et al. 2020). This section also covers diagnostics and the design of the ML model. Section 5.5 analyzes the role of large-scale moisture advection in producing both clustering and random convection configurations. We explore the

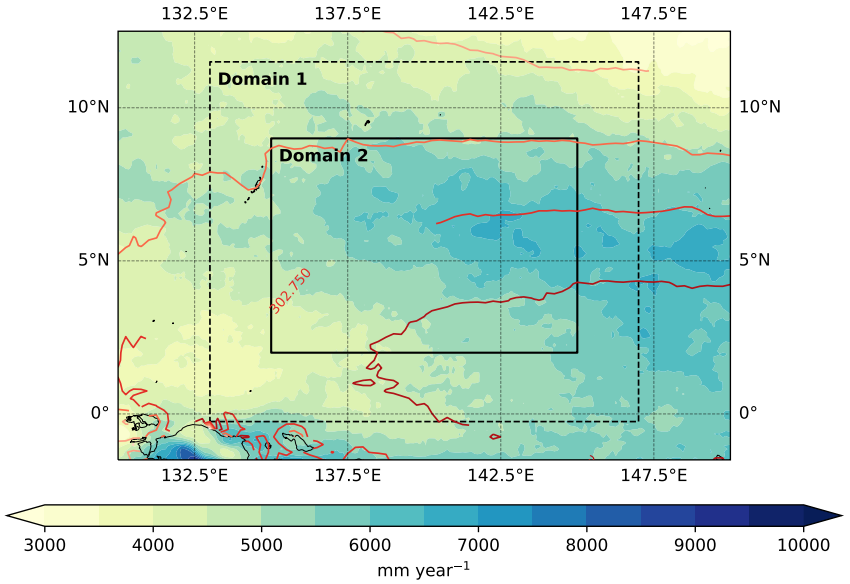

FIGURE 5.2: Annual mean distribution of Precipitation (colors) and SST (contours) for the study area. The SST contours start from 301.75 and increase every 0.25K. The dashed rectangle delimits the outer domain (Lon: 132.25 to 147.25; Lat: -0.25 to 12) and the continuous rectangle delimits the inner second domain (Lon: 135 to 145; Lat: 2 to 9), which is the one we used for all the calculations of this research.

influence of moisture and wind, including their components, on clustering development or prevention (§ 5.5.1), with a comprehensive evaluation of wind role in organization (§ 5.5.2). Section 5.6 presents a comprehensive overview of results from all simulations conducted using the ML model while concluding remarks are provided in Section 5.7.

5.3 Brief Observational Analysis

Our analysis focuses on a specific region within the Warm Pool studied in Chap4 represented by Domain 2 in Figure 5.2. This region undergoes a transition from the warm pool's boundaries in boreal winter/spring to its center in boreal summer/autumn, leading to a SST gradient during winter/spring that diminish to almost zero during the summer/autumn period (indicated in Figure 5.3c-f). Similar behavior is observed with the total column water vapor (TCWV) (Figure 5.4a-b). While this region was not strongly connected to the Pacific-wide energy budget, its local humidity variance and energy budget was clearly impacted the most by westward propagating convectively coupled Rossby waves in the boreal winter/spring periods, and it is the impact of these waves what controls the variance of TCWV which is strongly correlated to the OLR of the entire Pacific. The precipitation climatology is fairly spatially uniform in this zone, with stronger values at the south-easter region within the domain (Figure 5.2). The juxtaposition of uniform precipitation and varying SST and TCWV conditions in this area makes it conducive for the study of convective organization, aligning with idealized studies and the regions selected by Tobin et al. (2012, 2013).

Prior investigations (Bony et al., 2020; Semie and Bony, 2020; Angulo-Umana and Kim, 2023; Galewsky et al., 2023) have attempted to understand convective organization, but challenges arise due to the difficulty of directly measuring convective core locations (see Biagioli and Tompkins, 2023b, for a comprehensive review of some of the issues with various approaches of gauging convective organization). Hence, Chap4 approached the issue by seeking consequences or "fingerprints" of clustering rather than direct measurements (Holloway et al., 2017; Biagioli and Tompkins, 2023b). This drew inspiration from the RCE simualtions using a slab ocean model of Shamekh et al. (2020a) and Tompkins and Semie (2021), who utilized Hovmöller plots to correlate SST with TCWV percentiles to reveal that the transition from random to organized convection shifts the warmest SSTs from drier to moister regions, thus inverting the SST-TCWV correlation from negative to positive.

To quantify organization in observations, Chap4 also ordered SST according to the TCWV in the target mesocale domain and computed the slope ($\frac{dSST}{dTPWf}$) by fitting a linear regression through positive SST anomalies against this sorted TCWV-%tile field at each time step (Figure 5.3). This regression allowed us to distinguish different states of convection organization, with the spatial variance of TCWV very large when the regression relationship was positive in boreal winter/spring, reminiscent of clustered convective states, while variance was much smaller and the mean TCWV much larger when the regression relation was reversed, similar RCE simulations of randomly organized convection. By establishing thresholds based on the slope's standard deviation (std) and on temporal considerations, organized and reversal states were identified. The purple (reversal) and blue (organized) shades in Figure 5.3b-e illustrate these states, while typical organized and reversal patterns are visualized in Figure 5.4c and 5.4d, respectively.

Under the defined metric (Figure 5.3g), boreal summer/autumn (JJA-SON) exhibits random convection, while boreal winter/spring (DJF-MAM) tends towards organization. Importantly, Chap4 shows that, despite diabatic feedbacks favoring organized convection, it can still manifest as a random configuration, potentially due to minimal SST gradients and suppression of TCWV heterogeneities by large-scale dynamics. This aligns with idealized experiments homogenizing humidity (Tompkins, 2001c). Conversely, boreal winter/spring displays organized convection due to robust diabatic feedbacks, stronger SST gradients, and substantial TCWV heterogeneities.

In Chap4 we find that organized convection coincides with drier atmospheric conditions (Angulo-Umana and Kim, 2023; Galewsky et al., 2023), significantly increased outgoing LW clear-sky radiation (Bony et al., 2020), and moderately higher outgoing LW all-sky radiation (Bony et al., 2020), influenced to a lesser extent by cloud fraction, which, although reduced, remains relatively stable. Intriguingly, Chap4's investigations reveal that during the boreal winter/spring, convection can transition from organized to random states due to moisture advection by westward propagating waves (Gonzalez and Jiang, 2019), evident in Figure 5.4d displaying an equatorial Rossby wave pattern consistent with Kiladis and Wheeler (1995) and Fuchs-Stone et al. (2019).

Figure 5.5 offers a composite of 44 reversals identified during the MAM season. Each panel portrays a daily lag, originating from 7 days prior to the onset of reversal events. Color-coded regions represent OLR anomalies, while contour lines

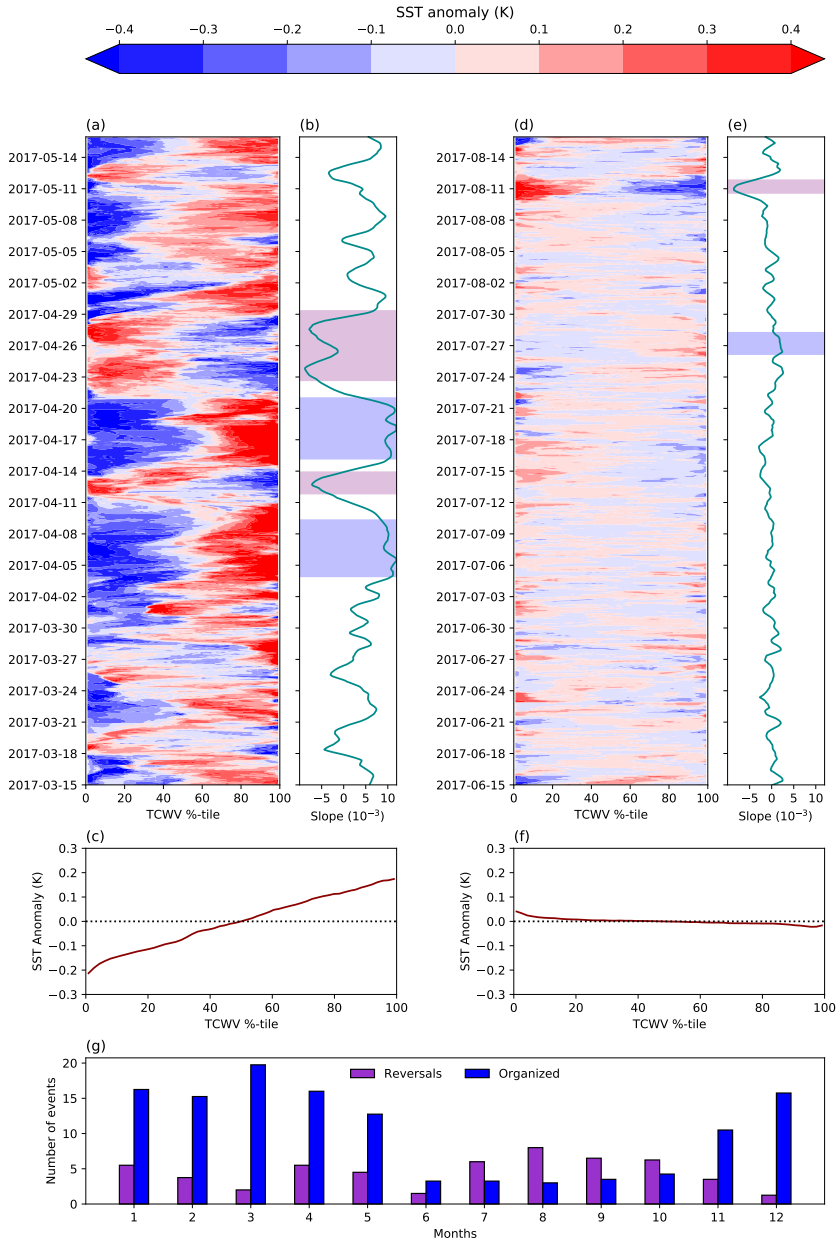

FIGURE 5.3: Hourly mean spatial anomaly of SST arranged as a function of TCWV percentile (ERA5 data, see section 5.4 for details about the data) for (a) MAM and (d) JJA season and their temporal average (c) and (f), respectively. Columns are arranged in order of TCWV-%tile and then an average of 81 columns is applied (no block averaging). Slope ($\frac{dSST}{dfTCWV}$) for (a) MAM and (e) JJA, the blue shading show examples of organized convection (i.e., 0.5*std larger than Slope mean), meanwhile the purple shading show examples of reversal events (i.e., 0.5*std smaller than Slope mean). Notice that for the JJA season the convection stays random since the SST tend to be homogeneous. (g) Annual average number of organized (blue) and reversals (purple) events per month from 2016 to 2020.

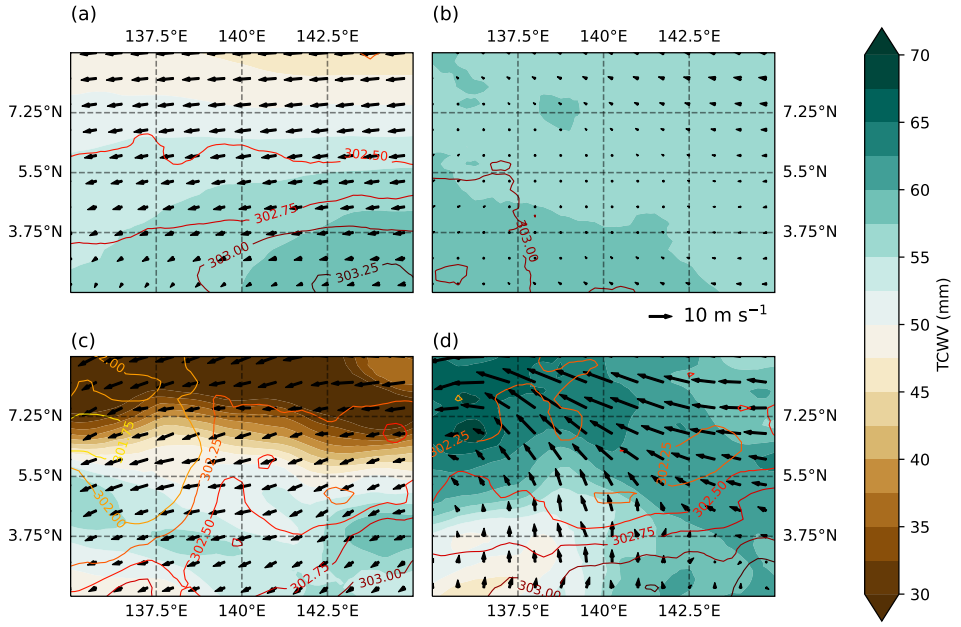

FIGURE 5.4: TCWV (colors), SST (contours) and wind vector (arrows) mean from ERA5 data for the (a) MAM and (b) JJA seasons. (c) Snapshot of organized convection on 2017-04-12 and (d) snapshot of random convection on 2017-03-21.

illustrate the magnitude of Vertically Integrated Water Vapor Flux (WVF), accompanied by directional arrows. Notably, four days prior to the moisture mode's arrival in our study area (as depicted in Figure 5.5h), the moisture mode associated with an Equatorial Rossby wave had already begun to manifest. This manifestation was characterized by a symmetric anomaly in OLR extending across the equator. The wave's initiation occurs 7 days ahead of its passage through the study zone, approximately spanning Lon: 150E-160E and Lat: 8N-12N. The onset of this anomaly can be attributed to moisture flux convergence, as illustrated in Figure C.1 in Appendix C. Interestingly, it is not primarily driven by the direction of moisture flux in this area but rather by speed convergence, resulting in a net inflow of moisture into the region, as depicted in Figures 5.5 and C.2 in Appendix C. This imbalance augments water vapor content (and convergence), fostering a robust moisture anomaly. As it propagates westward, the anomaly eventually disrupts convective clusters positioned in the southern portion of the study area. This impact is not only attributed to the moisture mode but also stems from a concurrent southern advection (Takayabu and Nitta, 1993; Takayabu, 1994) of moisture prompted by the convergence generated by the advancing wave.

In this context, our aim is to identify the key factors driving both organized and random convection episodes in the boreal winter/spring, focusing on cluster breakup episodes associated with westward propagating Equatorial Rossby waves. In particular we want to understand what drives the homogenization of the humidity in these wave passages and the importance of humidity advection. We also wish to investigate what sustains convection in a random configuration during boreal summer despite the diabatic forcing which acts to induce convective spatial clustering.

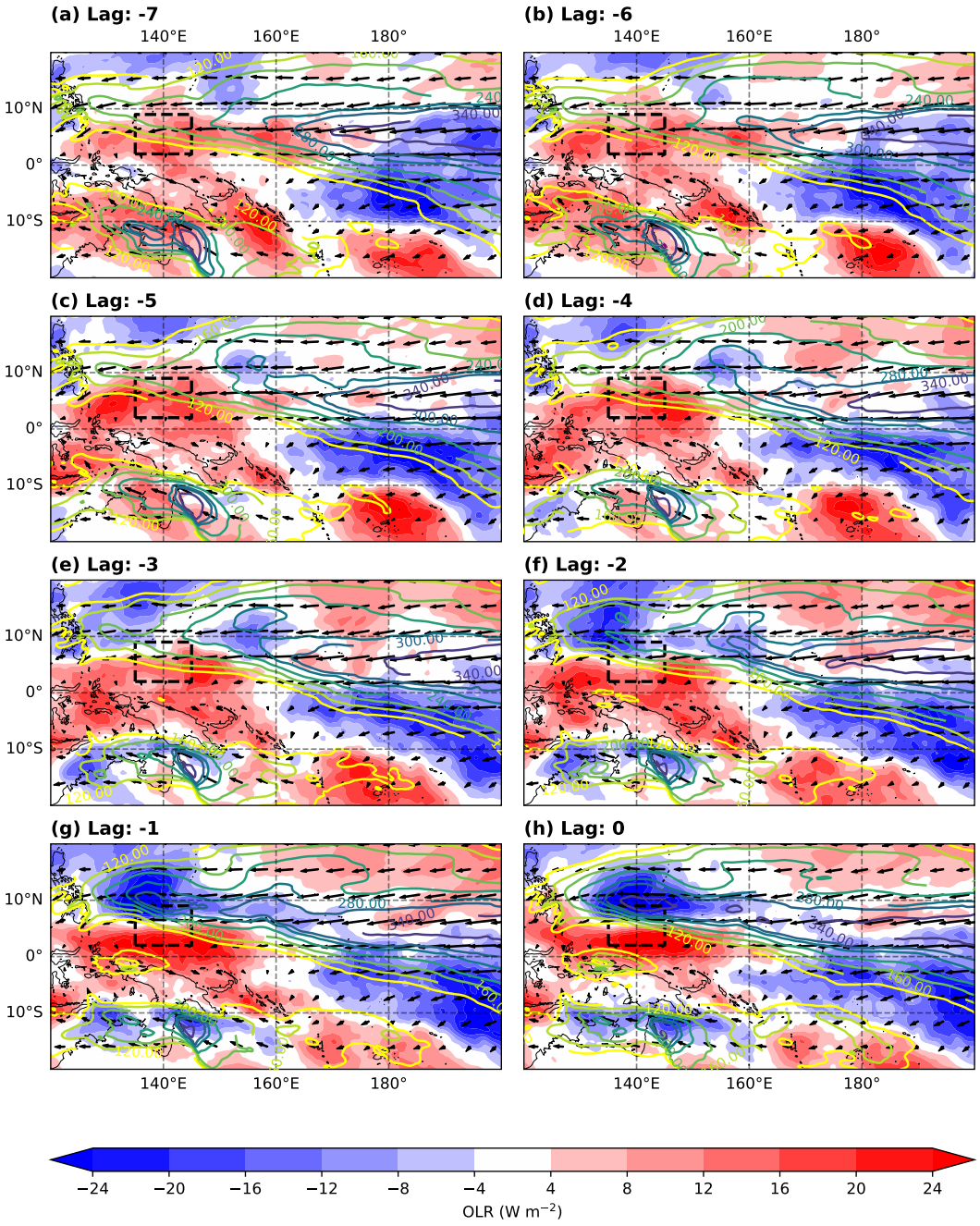

FIGURE 5.5: Composite of 44 reversal states on MAM seasons and their related lags. The colors represent the OLR anomaly for each of the lags. The contours are the magnitude and the arrows the direction of the WVF. Notice that each panel indicate the lag in number of days. The dashed rectangle represents the study area, whose close-up is presented in Figure 5.2 as Domain 2.

In light of the insights from Chap4 and previous studies that have highlighted the role of southerly flow in enhancing the moisture mode by transporting humidity from the south as waves propagate through the warm pool (Takayabu and Nitta, 1993; Takayabu, 1994; Gonzalez and Jiang, 2019), we anticipate that moisture advection will play a crucial role in determining both organized and reversal states. In Chap4, organized episodes in boreal winter/spring feature a strong easterly zonal wind and a weak northerly meridional component, with convection in the southern part of the domain over the warmest SSTs. In contrast, random convection episodes are projected to involve a weakened zonal wind and a southerly flow. Additionally, the meridional wind is expected to contribute significantly to breakups due to the southerly flow resulting from convergence initiated by the propagating wave (Figure 5.4d). During these random episodes, we also anticipate a net outflow of Moist Gross Static Energy (MGSE) (Bretherton et al., 2005; Raymond et al., 2009; Muller and Held, 2012), countering the diabatic forces that Chap4 reported to consistently promote convective clustering during boreal summer/autumn when neither SST nor TCWV heterogeneities develop.

5.4 Datasets and Method

5.4.1 Datasets

Regarding observational datasets used for section 5.3, we used the datasets already described in detail in Chap4. For this, here we briefly mention them. We employ the Global Precipitation Measurement Core Observatory spacecraft to retrieve precipitation data. The TCWV data is acquired on an hourly basis from the morphed integrated microwave imagery for the total precipitable water (MIMIC-TPW) product, version 2, from the Cooperative Institute for Meteorological Satellite Studies. Our analysis also incorporates 3-hourly Top Of Atmosphere (TOA) OLR data from the Gridded Satellite (Gridsat) B1 dataset (Knapp et al., 2011), featuring a horizontal resolution of 1.0°. We use hourly SST data (level 3) from the Himawari-8 Collection version 1.2 (Bessho et al., 2016).

Observations are supplemented by ERA5, the fifth-generation atmospheric reanalysis product developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach et al., 2020). We use hourly output of wind components (zonal, meridional, and vertical), air temperature, and specific humidity as a function of height as well as surface variables of SST, LH and sensible heat (SH) fluxes, TCWV, both components of the vertically integrated water vapor flux (WVF), vertically integrated moisture divergence (VMD), and TOA net all-sky SW and LW radiation, including both clear-sky and all-sky components. It's worth noting that all datasets were acquired for the period spanning from 2016-10-01 to 2019-12-31, a range dictated by satellite data availability.

5.4.2 Model Description, Simulation Setup and Validation

In this study, we employed the Weather Research and Forecasting (WRF) model version 4.2 (Skamarock et al., 2019) for simulations. We implemented two nested domains (depicted in Figure 5.2), mitigating potential performance reduction associated with a single domain (Hernandez-Deckers et al., 2022). The model configuration included 32 vertical levels, with a model top at 20.5 km. Horizontal resolutions

were set at 6 km^2 and 2 km^2 for **Domain1** (Longitude: 133E-147E; Latitude: 0.25S-11.25N) and **Domain2** (Longitude: 135E-145E; Latitude: 2N-9N).

Microphysical processes were simulated utilizing the Thompson microphysics scheme (Thompson et al., 2008; Thompson and Eidhammer, 2014; Iverson et al., 2021), while radiative transfer was modeled with the RRTMG model (Mlawer et al., 1997; Iacono et al., 2008), encompassing both LW and SW radiation. The YSU scheme (Hong and Pan, 1996; Hong et al., 2006) was employed for the boundary layer, Smagorinsky 2D (Smagorinsky, 1963) for sub-grid scale horizontal mixing, and the Monin-Obukhov similarity theory (Monin and Obukhov, 1954) for the surface layer. Oceanic Mixed Layer Depth (MLD) and SST schemes are detailed by Pollard et al. (1973) and Skamarock et al. (2019). These parameterizations are selected since they are recommended by the WRF developers team (Powers et al., 2017).

We produced hourly outputs for two distinct periods: 2017-03-15 to 2017-05-15 (encompassing organized and reversal states, as observed in Figure 5.3a-b-g), and 2017-06-15 to 2017-08-15 (encompassing mostly a random state during the JJA season, as evident in Figure 5.3d-e-g). The model's initial and boundary conditions were derived from 0.25° GFS analysis (NOAA, 2015) at 6-hour intervals, featuring 32 vertical levels up to a model top of 20.5 km. For initial conditions, SST values were homogenized, set as the mean domain value for the first simulation hour using ERA5 data. The MLD commenced at an initial depth of 60 m, consistent with typical conditions in the Pacific Warm-Pool (Kara et al., 2003). Although we explored various MLD depths (9 m, 20 m, 30 m) (Tompkins and Semie, 2021), constant MLD, and MLDs without wind dependence, but no indication of significant sensitivity to this parameter was found.

We verified the WRF model's capability to accurately replicate both the MAM (2017-06-15 to 2017-08-15) and JJA (2017-06-15 to 2017-08-15) periods, as well as its proficiency in reproducing organized and reversal states (Figure C.3 in Appendix C). The model exhibited (r) correlation coefficients of 0.69 and 0.52 for the MAM and JJA Slopes, respectively, along with mean square errors of $12.05e^{-6}$ and $14.022e^{-6}$ for the corresponding seasons. Notably, the model faithfully captured all convective aspects pertinent to this study, encompassing organized (e.g., 2017-03-21) and reversal (e.g., 2017-04-12) states, as indicated in Figure C.3 in Appendix C. Nevertheless, WRF simulate weaker SST gradients compare to ERA5 data,due to this the SST-TCWV relation during organized episodes is less strong (April, 4th to 9th 2017), although, they can still be consider organized episodes, since they are larger than 0.5*std of the mean SST-TCWV relation calculated from WRF output.

The model and observations exhibited remarkable concordance in the spatial distribution of moisture, SST, and wind, evident in the comparison between Figure 5.4 and Figure 5.6. This agreement extended to capturing both the overarching conditions of MAM and JJA (Figures 5.4a-b and 5.6a-b) and the specifics of organized situations (Figures 5.4c and 5.6c) characterized by elevated SST and TCWV in the southern domain, accompanied by a southwestward flow. The model also adeptly reproduced reversal instances (Figures 5.4d and 5.6d), marked by heightened TCWV in the northern domain, while higher SSTs remained situated in the south, alongside a southeasterly wind—a pattern consistent with observations. However, as a limitation, the model exhibited a tendency to slightly underestimate SST values, resulting in weaker SST anomalies. Nevertheless, these findings provide a robust foundation

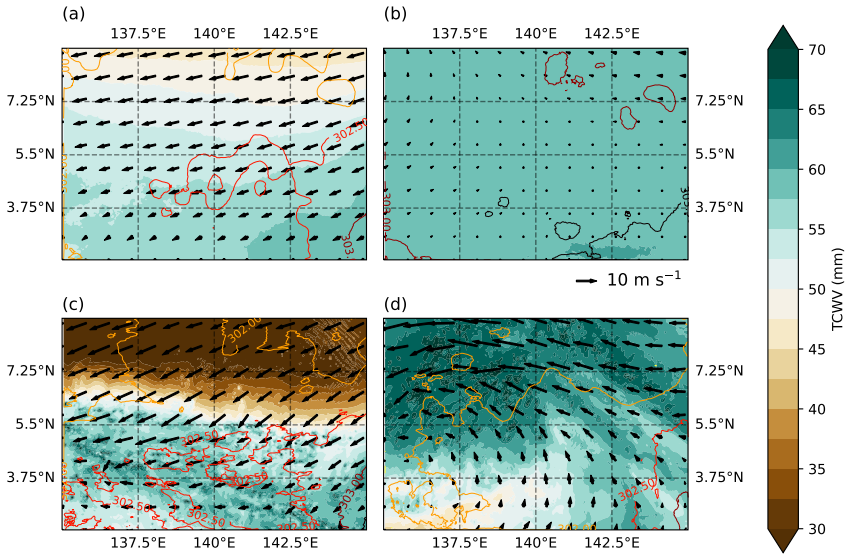

FIGURE 5.6: TCWV (colors), SST (contours) and wind vector (arrows) mean from WRF model for the (a) MAM and (b) JJA seasons. (c) Snapshot of organized convection on 2017-04-12 and (d) snapshot of a reversal event on 2017-03-21

for investigating the underlying mechanisms governing the emergence, persistence, and disruption of organization. In pursuit of this, we introduced modifications to the boundary conditions of the MAM simulation (serving as our control, since it includes organized and reversal conditions) in three distinct experiments that respectively subsistute boundary value (i) moisture advection ($\vec{u}q$), (ii) moisture (q), and (iii) wind components (u and v) for the entire period with values taken from a single fixed day. Two experiments are conducted, with the fixed day in the first chosen to correspond to a day subject to a highly organized convective state (2017-03-20 18:00:00), whereas the second experiment adopts a day in a "reversal" state where domain humidity was very homogeneous (2017-04-12 00:00:00). Table 5.1 (in Appendix C) contains a summary of all the sensitivity experiments performed using the WRF model.

5.4.3 Machine Learning Model

To complement our analysis, we trained a Random Forest (RF) algorithm (from the SKlearn package V.1.0.1 of Python 3.9.7) to perform sensitivity tests including changing the low/high wind shear, this RF idea is borrowed from Casallas et al. (2023a). We explored the use of neural networks, support vector machine, and Decision Trees (not shown) as advised by McGovern et al. (2019) to reduce uncertainties. However, our findings demonstrated that the choice of ML technique did not significantly impact the results, underscoring their robustness. As such, here, we will solely focus on the RF method, as it yielded the most favorable validation results.

Feature Selection, Input/Output Structure and Validation

The ML model is designed to emulate the dynamics and variations of the SST-TCWV regression, drawing on key variables from the ERA5 dataset: wind vector (27 levels), TCWV, SST, and adiabatic and diabatic fluxes (LH, SH, LW, and SW). This selection

Simulation Name	Boundary Conditions Modification
Moisture Advection Reversals (MAR)	Replace the water vapor and the wind boundary conditions with those of the strong reversal event of 2017-04-12 00:00:00
Moisture Advection Organized (MAO)	Replace the water vapor and the wind boundary conditions with those of the strong organized event of 2017-03-20 18:00:00
Moisture Reversals (MoiRev)	Replace the water vapor boundary conditions with those of the strong reversal event of 2017-04-12 00:00:00
Moisture Organized (MoiOrg)	Replace the water vapor boundary conditions with those of the strong organized event of 2017-03-20 18:00:00
Wind Reversals (WinRev)	Replace the wind boundary conditions with those of the strong reversal event of 2017-04-12 00:00:00
Wind Organized (WinOrg)	Replace the wind boundary conditions with those of the strong organized event of 2017-03-20 18:00:00
Zonal Wind Reversals (URev)	Replace the zonal wind boundary conditions with those of the strong reversal event of 2017-04-12 00:00:00
Zonal Wind Organized (UOrg)	Replace the zonal wind boundary conditions with those of the strong organized event of 2017-03-20 18:00:00
Meridional Wind Reversals (VRev)	Replace the meridional wind boundary conditions with those of the strong reversal event of 2017-04-12 00:00:00
Meridional Wind Organized (VOrg)	Replace the meridional wind boundary conditions with those of the strong organized event of 2017-03-20 18:00:00

TABLE 5.1: Overview of the simulations performed with WRF by changing the large-scale conditions through the boundary conditions.

is based on the role these variables play in both the organization and breakdown of convection, as elucidated by Chap4. SST and TCWV are chosen for their potential influence on organizational processes, while the wind vector is significant in modifying moisture advection, potentially crucial for both organization and reversal events. Vertical velocity provides insights into convective core characteristics, and surface and radiation fluxes offer a glimpse into local diabatic feedbacks, along with convection patterns.

We applied the variance inflation factor (VIF) methodology (Ghahremanloo et al., 2021) to determine an optimal and less complex combination of variables (Behrens et al., 2022; Grundner et al., 2022). This selection criterion ensured that we identified variables contributing to the understanding mechanisms that lead to reversal and organization states, while also maintaining multicollinearity below a threshold of five (Kline, 2015). This consideration is crucial, as prior studies have highlighted the adverse impact of multicollinearity on ML model accuracy (e.g., Wei et al. 2019). It's

pertinent to note that all variables were converted to domain-mean values to align their dimensions with the slope.

After the VIF analysis, six variables were selected as inputs for the RF algorithm: SST, TCWV, LW radiation, vertical velocity at varying pressure levels (850-500-100hPa), zonal wind at different pressure levels (1000-850-650-450-200-150-100hPa), and meridional wind at distinct pressure levels (1000-850-650-500-400-200-150-100hPa). These variables were chosen due to their significance in influencing the slope and their low collinearity. The input matrix consists of hourly data points for each selected variable, amounting to approximately 35000 data points per variable. The output vector also has hourly time resolution and is composed of the slope values.

For training, 90% of the data was used, while the remaining 10% was allocated for validation. It's worth noting that within this 10%, the period from 2017-03-15 to 2017-05-15 was included, which aligns with the time frame used in our WRF experiments. During training, a randomized search method was employed to identify the optimal trees depth. The validation outcomes (Figure C.4 in Appendix C) showcased great performance, with an R^2 of 0.97, a mean squared error of $0.1022e^{-7}$ Km^2kg^{-1}, and an r value of 0.99. These results affirm the model's aptitude to capture the intricate relationship between input variables and the slope (Figure C.4c in Appendix C). This high precision empowers the model to provide insights into the impact of variables that are challenging to manipulate in WRF due to potential disruptions to mass or momentum conservation (e.g., wind shear).

5.5 The Role of the Moisture Advection and Diabatic Forcings

We conducted sensitivity experiments altering the moist and wind boundary conditions in our simulations. These conditions were substituted with values corresponding to either a reversal or an organized event during the MAM season, which encompasses both convection states.

The outcomes of the boundary condition experiments are illustrated in Figure 5.7. Initially, we focus on explaining the moisture advection (MAR for reversal conditions and MAO for organized conditions) experiments depicted in Figure 5.7a-b. These experiments, as expected, exhibit strong resemblance to the reversal and organized scenarios presented in Figures 5.4c-d and 5.6c-d, in terms of TCWV, SST, and wind distribution. Figure 5.8 presents the SST-TCWV relationship for MAM and JJA seasons and also for the MAR and MAO scenarios, depicting the state of convection and allowing the comparison between experiments. In the context of a reversal scenario, the highest TCWV values are situated to the north, while the warmest SSTs are found in the southern region. A southerly flow emerges, conveying moisture from the south. This SST-TCWV relationship is depicted in Figure 5.8b, indicating its persistence across the entire simulation, unlike the sporadic cases seen in Figure 5.8a. In the organized setting (Figures 5.7b and 5.8c), peak TCWV values occur in the southern domain where the warmest SSTs reside. The wind direction is southwestward, characterized by a strong zonal component. This implies a narrower warm pool with drier northern boundaries, akin to the findings of Hohenegger and Jakob (2020) and Beucler et al. (2020) regarding the Atlantic ITCZ.

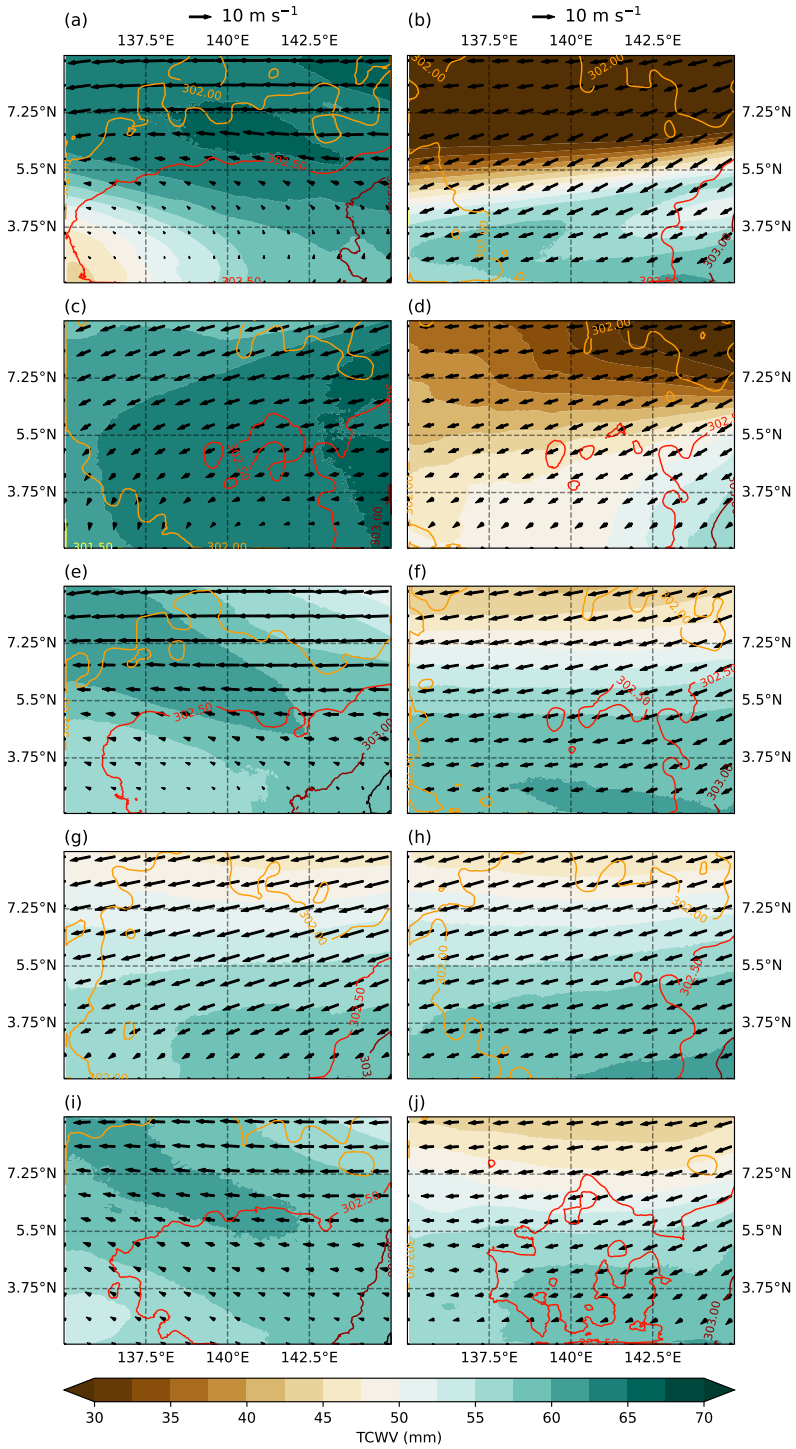

FIGURE 5.7: Mean TCWV (colors), SST (contours), and wind vectors (arrows) from the WRF model for each sensitivity experiment. The left (right) column displays experiments in which the reversal (organized) conditions are replaced in the boundary layer. Panels (a) and (b) depict the moisture advection experiments, while (c) and (d) illustrate the moisture boundaries experiment. Conversely, (e) and (f) showcase experiments that replace the winds, whereas panels (g) and (h) display experiments that only modify the zonal wind, similar to panels (i) and (j), which represent experiments modifying the meridional wind.

These four configurations offer a valuable perspective to explore the interplay between local feedbacks and large-scale influences in shaping convective organization. Following the approach of Tompkins and Semie (2021), based on previous work (Bretherton et al., 2005; Muller and Held, 2012; Wing and Emanuel, 2014), we calculate the diabatic feedbacks. Figures C.5 to C.8 (in Appendix C) present the forcing for the LW and SW Cloud, LW and SW clear-sky, and LH and SH flux feedbacks for the MAM, JJA, and moist advection experiments over the entire period. In contrast, Figure 5.9 illustrates the cumulative diabatic and adiabatic feedbacks derived for each experiment.

Consistent with idealized modeling studies, in MAM the LW cloud and LW clear-sky feedbacks act to organize convection (Holloway and Woolnough, 2016), exhibiting the strongest feedbacks (Tompkins and Craig, 1998a; Stephens et al., 2008). However, during reversal development, the LW clear-sky feedback opposes organization by moving less cloudy regions toward the warmest SSTs, inducing regional heating and creating a negative feedback with organization (Figure C.5 in Appendix C). This negative feedback intensifies in the MAR experiment (Figure C.7 in Appendix C), where the warmest regions persistently locate in dry, clear-sky zones, intensifying SST warming and magnifying the feedback. The largest cooling rates of LW clear-sky occur at intermediate water vapor levels, aligning with Tompkins and Semie (2021), while mild warming appears in the moistest regions. Simultaneously, SW cloud heating anomalies remain negative across all four simulations, as high clouds dampen incoming solar radiation. Interestingly, the SW clear-sky exhibits a positive feedback in the MAO experiment (Figure C.8 in Appendix C) and during the MAM season due to heightened SW absorption in the moistest regions (Wing and Emanuel, 2014). This effect weakens when convection becomes random, as observed in the MAR experiment and during the JJA season (Figure C.6 in Appendix C).

The LH and SH flux heating anomalies align with prior idealized and observational studies (for comprehensive reviews, consult Wing et al. 2017, Holloway et al. 2017, Muller et al. 2022a). Surface fluxes reinforce organized convection in its random state and disrupt it when already organized, consistent with our study's results and observations from Chap4. Similar to Tompkins and Semie (2021), we decompose LH flux anomalies into contributions from humidity, stability, and wind. Their research showed that the positive feedback when convection is random primarily comes from stability and, to a lesser extent, wind speed, while the negative feedback arises from humidity differences between the surface and boundary layer. Correspondingly, the Chap4 observational results demonstrate that LH flux feedbacks have the same sign as in Tompkins and Semie (2021), with wind speed strongly dictating the feedback's sign and strength. Our modeling outcomes align with the results of Chap4, illustrating that wind speed overwhelmingly governs both negative and positive feedbacks, overshadowing the influence of temperature and humidity structures. In brief, in scenarios of random convection, the wind-induced surface heat exchange (WISHE) feedback is characterized by the enhancement of surface winds in intensely convecting regions with convective gustiness (Wing and Emanuel, 2014). Conversely, as convection becomes organized, winds typically diminish in the convective regions while strengthening in the dry zones, giving rise to a negative feedback linked to organizational patterns.

Figure 5.9 illustrates the average total heating anomalies for the entire simulation

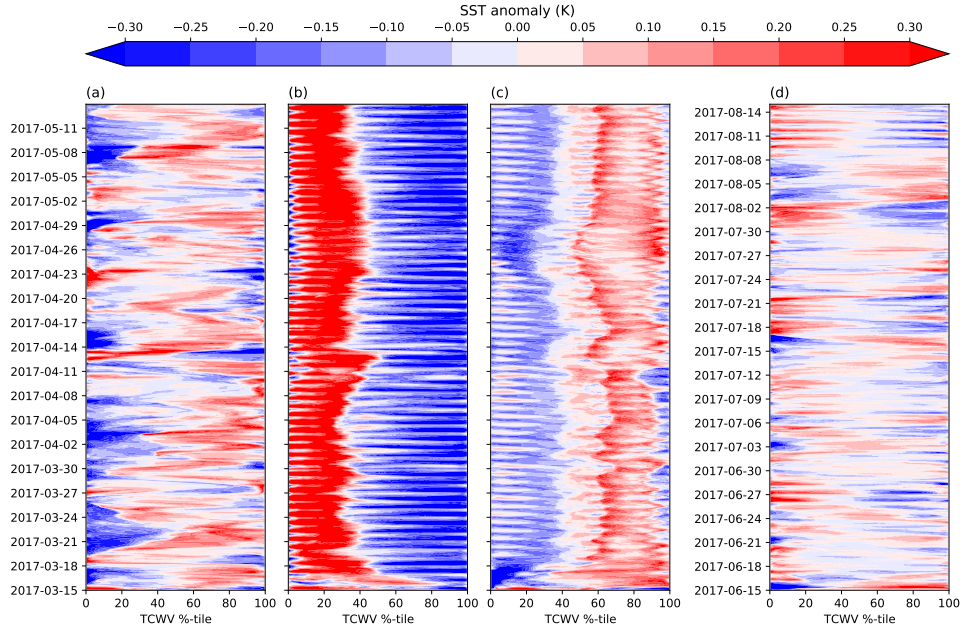

FIGURE 5.8: Hourly mean spatial anomaly of SST arranged as a function of TCWV percentile for (a) MAM, (b) the MAM simulations but with reversals Moisture Advection conditions, (c) as (b) but with organized Moisture Adv. conditions and for (d) JJA. Columns are arranged in order and then an average of 1656 columns is applied (no block averaging)

period, encompassing experiments representing organized (Figure 5.9a) and reversal conditions (Figure 5.9b), including both MAM and JJA scenarios. All experiments indicate that local diabatic forcing contribute to organizing convection, yet their potency is not always sufficient to cluster convection. This suggests that convection should export MGSE to counterbalance these feedbacks (e.g., Raymond et al. 2009; Muller and Held 2012). To study this, we utilize the stream function calculation, widely employed in diverse studies (e.g., Bretherton et al. 2005; Muller and Bony 2015; Shamekh et al. 2020a; Tompkins and Semie 2021). Figure 5.10 illustrates the net radiative cooling (Q_{rad}), liquid and ice mixing ratios, and stream function, sorted by TCWV percentiles. In MAM and MAO experiments, the highest net Q_{rad} is observed between 950hPa to 850 hPa, particularly in intermediate moisture regions (60-80%-tile) compared to JJA and MAR experiments, in which the largest net Q_{rad} is located in the driest zones. This suggests increased subsidence regions when convection is organized, consistent with previous idealized studies (e.g., Coppin and Bony 2015; Müller and Hohenegger 2020b). The MAM and MAO experiments present a stronger cooling in the mid-troposphere (500hPa to 300hPa), especially in intermediate humidity regions (40-60%-tile), akin to the results of Tompkins and Semie (2021).

Importantly, the MAO case exhibits a lower cloud fraction compared to the other experiments. This aspect gains significance due to insights from idealized studies (Wing and Emanuel, 2014), reporting that organized configurations prompt heightened OLR. This aligns with the observed cloud-associated OLR surge in our results (not shown), corroborated by findings from the same region in Chap4. However, it is noteworthy that the MAO case displays a more pronounced manifestation of this behavior compared to observations and the MAM simulation, possibly attributed

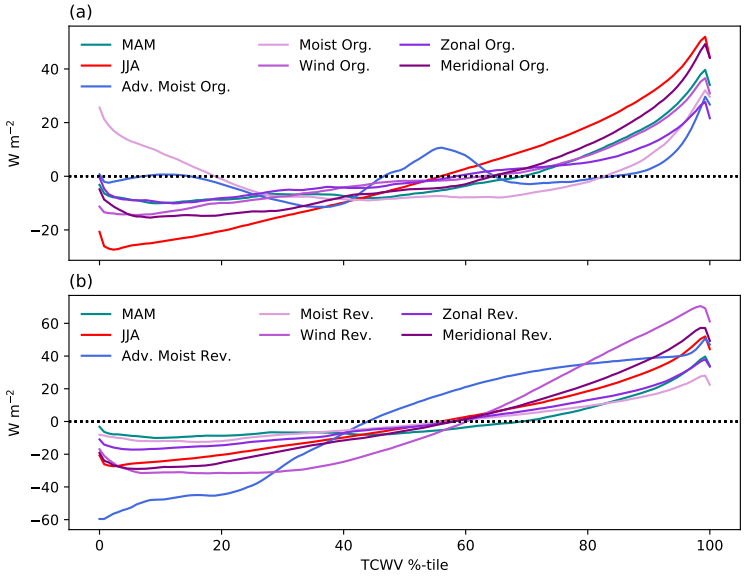

FIGURE 5.9: Total atmospheric convergence anomaly mean (positive is warming anomaly of atmosphere) ordered according to TCWV for (a) organized and (b) reversals conditions experiments. Both panels include the simulations for MAM and JJA seasons for comparison.

to the assistance provided by the large-scale atmospheric conditions (of the experiment) in maintaining organized convection.

Comparing the circulation patterns between the MAM and JJA seasons (Figure 5.10a and c), several observations stand out. Firstly, in MAM, stronger low-level Q_{rad} leads to greater surface up-gradient MGSE export, particularly in moister locations. Subsidence is pronounced between the 0-60 TCWV %-tile range in MAM, whereas in JJA, the largest subsidence occurs at the 20th TCWV %-tile. The expansion of the largest Q_{rad} to intermediate moisture regions in MAM restricts the strongest circulation, hindering moisture transport to the driest areas and favoring organization (Jeevanjee and Romps, 2013; Muller and Bony, 2015). In contrast, during JJA, the circulation effectively exports MGSE to the driest zones, countering local diabatic feedbacks. This export appears to rely more on dynamic and large-scale motions, which homogenize the water vapor field, reducing inhomogeneities. These findings are consistent with our sensitivity experiments and align with idealized studies that investigate homogenization of the water vapor field (Tompkins, 2001c; Yang, 2019).

Similar to the MAM season, the MAO experiment (Figure 5.10b) exhibits a contraction of the circulation, particularly notable at 900hPa and 500hPa, compared to the JJA or MAR experiment (Figure 5.10d). This contraction is likely due to the expanded area of strong Q_{rad} (1000hPa to 800hPa) between 0-60 TCWV %-tiles, possibly influenced by the SST gradients present during these experiments, this cooling also leads to stronger subsidence, and to a larger MGSE export at the surface (Bretherton et al., 2005; Muller and Held, 2012; Muller and Bony, 2015; Shamekh et al., 2020a). Notably, additional radiative cooling occurs at 300-400hPa in intermediate moisture regions (40-60 TCWV%-tile), creating a secondary circulation that

further inhibits moisture transport to the driest areas, encouraging organized conditions, consistent with the results of Tompkins and Semie (2021). In the MAR experiment, the most intense cooling is concentrated in the driest regions at 300-400hPa, facilitating moisture export to these areas and discouraging organized convection. This phenomenon is also observed in the JJA season but is more pronounced in the MAR experiment.

In synthesis, the diabatic feedbacks alone lack the potency to induce convection organization without the presence of a SST gradient, as they can be counteracted by MGSE export, potentially propelled by dynamic large-scale processes that constrain moisture (and SST) heterogeneities. In contrast, when convection is organized, the circulation responsible for transporting moisture from moist to dry regions contracts, which is a result from expanded areas of strong radiative cooling and increased subsidence. These coupled with diabatic feedbacks working synergistically to organize convection, further fueled by the strong influence of large-scale dynamics on developing SST and TCWV heterogeneities, strongly favor clustering.

5.5.1 Analyzing Moisture Structures and Wind

This section studies the role of the moisture structures and the wind, including an evaluation of the wind components. A summary of the results is presented in Figure 5.11, and Figures C.9 (reversals), and C.10 (organized) in Appendix C, illustrating the slopes for the control experiment and the experiments utilizing reversed and organized boundary layer conditions for each of the variables under investigation. For instance, Figure 5.11a displays the slopes for the MAR and MAO experiments explained in the preceding section.

The Role of the Humidity

Results from the experiments that manipulate humidity boundary conditions, labeled as MoiRev for reversal conditions and MoiOrg for organized conditions, reveal intriguing findings. The MoiRev experiment (Figure 5.7c) closely resembles the humidity distribution during the JJA season (Figures 5.4b and 5.6b), characterized by its homogeneity. This is corroborated by the slope calculation (Figure 5.11b and C.9b), showing minimal variability and reflecting essentially random conditions. This indicates that large-scale humidity alone can lead to random convection, even in the presence of strong diabatic feedbacks (Figure 5.9b and C.11), which exhibit a similar sign and magnitude to the JJA and MAR experiments, but are not strong enough to cluster convection.

However, it's worth noting that the MoiRev experiment also generates short periods of "organized" convection as measured by the TCWV variance (e.g., 2017-04-16 to 2017-04-17) and exhibits very intense reversal events in which convection is advected to the northern part of the domain. These conditions are reminiscent of the westward propagating wave, in which a wave strongly disrupts convection during MAM season and the control. This suggests that, while moist structures can contribute to random conditions, the role of wind is crucial, particularly in events that disrupt convection.

Concerning the MoiOrg experiment, it exhibits elevated TCWV values predominantly concentrated in a specific region of the domain (Figure 5.7d) for the majority

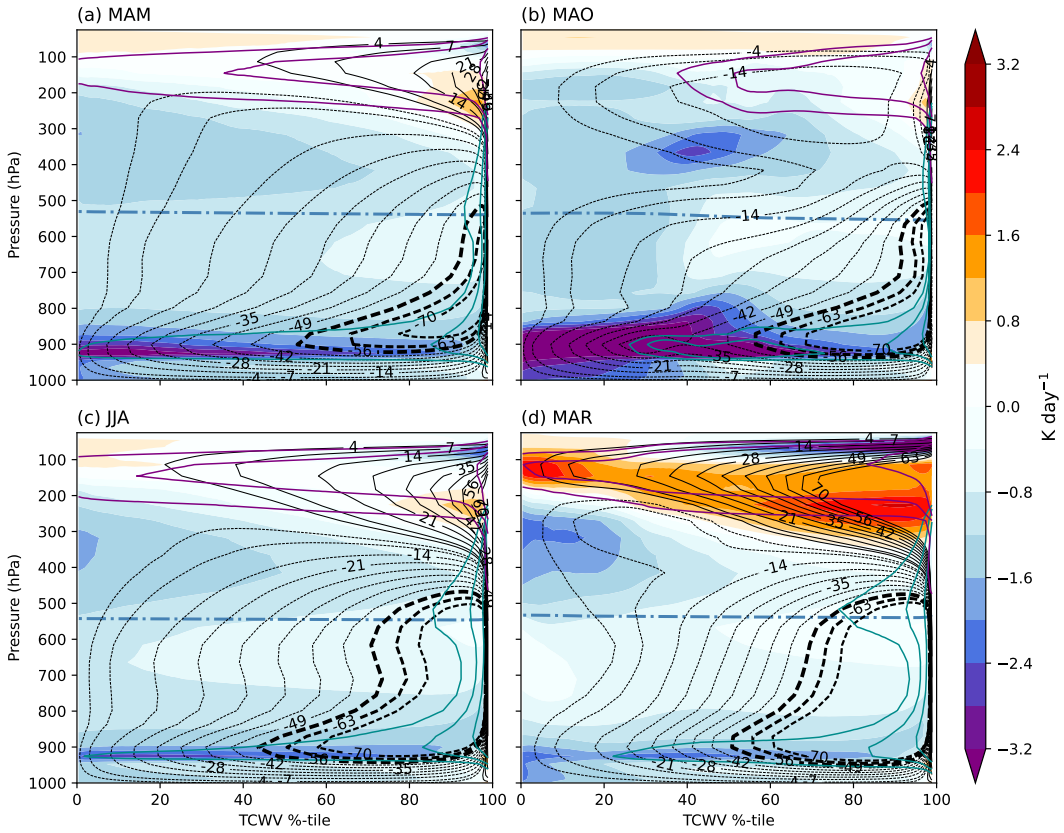

FIGURE 5.10: Average (first to last day) of net radiative cooling (color shading), liquid water mixing ratio (dark-cyan contours marking 10^{-5}, 3.10^{-4}, and 10^{-4} kg kg^{-1}) and ice water mixing ratio (purple contours marking 10^{-6}, 3.10^{-5}, and 10^{-5} kg kg^{-1}) as a function of the TCWV percentile for (a) MAM season, (b) the MAM simulations but with organized Moisture Adv. conditions, (c) JJA season and (d) as (b) but with reversals Moisture Adv. conditions. Black dashed contours show the mass flux streamlines as defined by Bretherton et al. (2005). The bold dashed lines represent the represent the -56, -63, and -70 in units of 10^{-2} kg m^{-2} s^{-1} to ilustrate that the strongest circulations persist within intermediate humid and the moistest regions.

of the simulation, indicating that moist structures can lead to an increase in clustering (Tompkins, 2001c). This insight is further supported by the slopes shown in Figures 5.11b and C.10b in Appendix C, which closely resemble those of the MAO experiment (Figure 5.11a). The clustering effect arises not only from the large-scale moisture structures but also from the diabatic feedback mechanisms that contribute to organizing convection (Figures 5.9a and C.10 in Appendix C) and are similar to the MAO experiment.

It's worth noting that, despite the large-scale humidity structures favoring organization, convection clusters can still be disrupted, as observed during specific periods, such as April 13-14 and April 18-24, 2017 (Figures 5.11b and C.10b). This highlights the significant role of wind in disrupting convection by advecting moisture from the south to the domain during breakup events. These events occur in the control and both the MoiRev and MoiOrg scenarios (Figure 5.6d). Additionally,

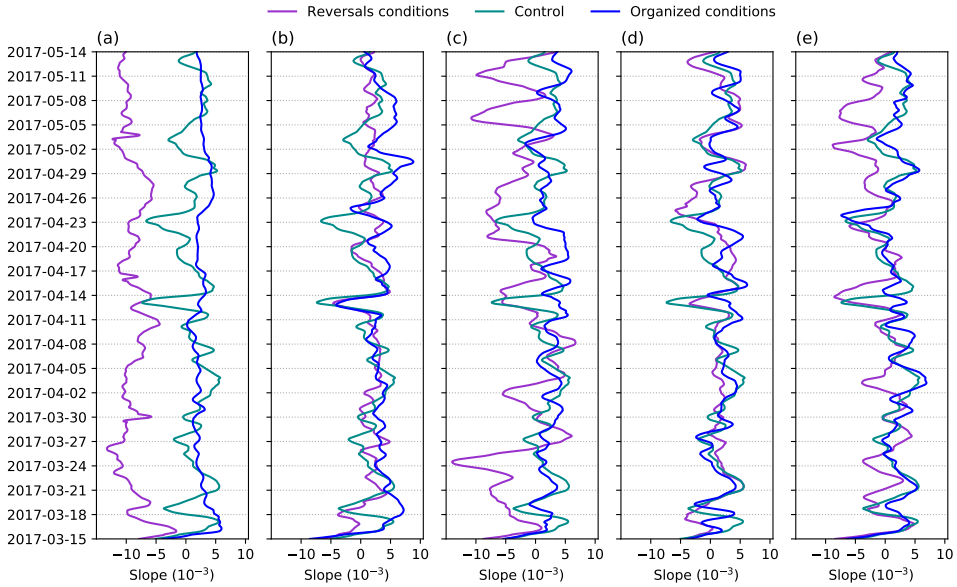

FIGURE 5.11: Slope calculated from the WRF output for (a) Moist Advection, (b) Humidity, (c) Wind, (d) Zonal wind component, and (e) Meridional wind component experiments. The darkcyan line represent the control (MAM season) run, meanwhile the purple(blue) line represent the reversals(organized) conditions.

the LW clear-sky feedback (Figure C.12b in Appendix C), although similar to that of the MAO experiment, is negative for more extended periods, particularly after a reversal event, where the highest SSTs are situated in the clear sky regions, which explain the negative feedback. This observation suggests that the wind can play a significant role in the reorganization of convection after a wave crosses the domain and breaks up convection clusters.

The Role of the Wind

In this section we modify the wind boundary conditions with reversals (WinRev) and organized (WinOrg) conditions (Figure 5.7e-f). One particularly noteworthy observation is the remarkable similarity between the outcomes of the WinRev and WinOrg experiments and those of the MAR and MAO experiments, respectively. In the case of the WinRev experiment, deep convection is shifted to the north-west of the domain, while in the southern region, new convective events are triggered and maintained as low-level clouds (Figure 5.7e). This configuration indicates random convection configuration, consistent with the findings presented in Figures 5.11c and C.9c. This pattern arises because a southeasterly flow persists in this experiment, resulting in the advection of moisture from the southern region of the domain (which during this season is very moist, due to the location of the warm pool), akin to the results of Takayabu (1994) and Gonzalez and Jiang (2019). This influx of moisture causes the bulk of convection to shift towards the northern part of the domain while new convection is triggered in the southern region. This random configuration bears a resemblance to the findings reported in Chap4 when a westward-propagating wave crossed the domain, generating a southerly flow due to the convergence induced by the passage of the wave (Kiladis and Wheeler, 1995; Haertel and Johnson, 1998). This is in contrast to the control experiment, where clear skies mostly prevail

in the northern part of the domain, while the majority of deep clouds are concentrated in the southern region, but it coincides with the control breakup events.

Furthermore, the diabatic feedbacks in the WinRev experiment act to organize convection (Figure 5.9b). However, these feedbacks are not strong enough to induce clustering. When analyzing individual fluxes (Figure C.13 in Appendix C), it is seen that all of them contribute to organizing convection. While LH primarily promotes clustering, aligning with the results obtained in Chap4 for the same region. The SW cloud on the other hand, acts to oppose organization, due to high cloud effects on incoming solar radiation. It is also important to mention that despite convection being random almost the entire period, there are some days in which it clusters for a small amount of time (e.g., April 8th, April 18th of 2017) probably due to the role of the large-scale moisture, which can strongly act to produce, and maintain convective clusters.

On the other hand, the WinOrg experiment exhibits similarities to the control experiment in terms of the mean spatial distribution of TCWV, with deep convection primarily located in the southern region of the domain, while the northern region remains predominantly cloud-free (Figure 5.7f). However, the WinOrg experiment features more organized convection compared to the control experiment (Figures 5.11c and C.10c) and experiences weaker breakup events, as observed during the event on April 14th, 2017. It's worth noting that the event that occurs between April 23rd and April 29th, 2017, in the control experiment, is also present in the WinOrg experiment, as also happens in the MoiOrg experiment. This demonstrates that while the wind can reduce the intensity of breakup episodes, large-scale humidity structures can independently induce randomness in convection. In fact, the MoiOrg experiment is capable of sustaining clustered convection for longer duration's but does not diminish the magnitude of breakup events. These findings suggest that both wind and moisture play crucial roles in generating breakup episodes and organization, with moisture structures being more efficient at maintaining clustering, and the wind being more important for the intensity of the breakup episodes.

When considering the diabatic feedbacks (Figures 5.9a and C.14) for the WinOrg experiment, they collectively act to favor clustering, with the exception of the SW cloud feedback, and of the LH flux, which consistently opposes organization as convection is already organized. In summary, all the aforementioned findings highlight the crucial role of wind in the development of breakup events and clustering. For this, the next logical step is to investigate the specific contributions of the wind components in terms of clustering convection and the development of breakup events or random convection configurations.

In evaluating the influence of the zonal wind, experiments involving the replacement of the simulation's boundary conditions with the zonal component of the wind from a reversal (URev) and an organized (UOrg) event exhibit notable similarities to the control experiments. This resemblance is apparent both in the spatial distribution of the TCWV (Figures 5.6a and 5.7g-h) and in the slope calculations (Figures 5.11d). The URev experiment (Figures 5.11d and C.9) demonstrates its capability to generate breakup events, often stronger than those in the control scenario. This suggests that, while the zonal wind may not be a dominant factor, it can have a modest impact on the breakup of convection, especially in terms of allowing convection to cluster after such events. In fact, these events tend to last longer than in the control

case, exemplified by the period from April 23rd to 29th, 2017. Regarding diabatic feedbacks, they closely resemble those of the control scenario, promoting organization (Figure 5.9a), while the LH and SW cloud feedbacks oppose clustering (Figure C.15 in Appendix C).

Regarding the UOrg experiment, it closely resembles the control experiment. Generally, convection tends to be located in the southern region of the domain, with clear skies prevailing in the northern region (Figures 5.11d and C.10). However, in the UOrg experiment, convection exhibits a higher degree of organization (Figure 5.11) compared to the control scenario, along with milder breakup events. This suggests that the zonal wind may contribute to the organization of convection and potentially plays a modest role in the occurrence of breakup events, although it does not exert the most substantial influence. In terms of diabatic feedbacks (Figure 5.9a), they collectively act to encourage organization, with the exception of the LH flux and SW cloud feedback, which act in opposition to organization, consistent with the control simulation (Figure C.16 in Appendix C), and with previous experiments using idealized and realistic modeling setups (e.g., Holloway 2017; Shamekh et al. 2020a), and observations (e.g., Holloway et al. 2017).

The experiments that replaced the meridional wind boundary conditions with reversal (VRev) and organized (VOrg) events offer intriguing insights. The VRev experiment exhibits similarities to the MAR and WinRev experiments in terms of the spatial distribution of TCWV (Figure 5.7a-e-i). In this scenario, deep convection predominantly occurs in the northwest region of the domain, with low-level clouds and new deep convection events developing in the southern region of the domain. This behavior is similar to the one observed when a breakup event occurs, as a southerly flow develops due to the convergence produced by a westward-propagating wave (Takayabu and Nitta, 1993; Takayabu, 1994; Kiladis and Wheeler, 1995). An examination of the slope (Figures 5.11e and C.9e) reveals that convection tends to be random for the majority of the simulation's duration under continuous reversal conditions. In some instances, convection briefly clusters but is quickly broken up. This behavior is likely attributed to the strong influence of large-scale moist structures that promote convection clustering, as suggested by the MoiOrg experiment (Figures 5.11b and C.10b). Additionally, diabatic fluxes favor organization (Figure 5.9b), including the LH flux feedback, which mostly acts to organize convection, while the SW cloud feedback opposes clustering (Figure C.17), mirroring the conditions observed in the MAR and WinRev experiments. These results suggest that the meridional wind on its own can lead to random convection configurations, although in certain cases, large-scale moisture structures may induce short-lasting clusters.

In the case of the VOrg experiment, the spatial distribution of TCWV closely resembles that of the control run (Figures 5.6a and 5.7j), with deep convection primarily developing in the southern part of the domain. The slope analysis (Figures 5.11e and C.10e) indicates that the breakup episodes are less intense but still occur, and convection is not as organized as observed in the MoiOrg experiment. This suggests that the meridional wind can contribute to organizing convection, but the large-scale moisture structures play a more significant role. The moisture structures on their own can also lead to the development of breakup episodes, as demonstrated by the event occurring between April 20th to 23rd, 2017 on both the Vrev and the Vorg experiments. Regarding the diabatic fluxes (Figures 5.9a and C.18), they collectively promote convection clustering, with the LH flux and the SW cloud feedback

opposing clustering, consistent with the control run and in line with previous studies (e.g., Wing and Emanuel 2014; Wing et al. 2017).

In summary, as discussed in § 5.5.1, large-scale humidity, when considered in isolation, does not possess the capability to impede reversals when organized conditions are enforced. However, it is remarkable that moisture structures alone can induce strong clustering of convection, surpassing the organization strength achieved by the wind or in the control experiments. Large-scale moisture structures can also engender random convection throughout the simulation when reversal conditions are imposed at the model boundaries. In some instances, this randomness transitions into convection clusters, which are rapidly disorganized. Notably, breakup episodes, akin to those observed in the control run, can develop, highlighting the critical role of the wind in explaining these episodes.

Conversely, the wind plays a more substantial role in initiating convection breakup episodes. It can diminish the intensity of reversals within the organized conditions experiment. Moreover, when reversal conditions are applied to the wind, convection exhibits a random state throughout the simulation, with a few instances of cluster formation quickly dispersed. This random convection results from moisture advection from the southern boundary into the domain. Regarding specific wind components, the zonal wind's influence on disrupting convection is minimal, weakly dampening reversals under organized conditions and having a similarly weak effect on strengthening breakup episodes when reversal conditions are imposed. In contrast, the meridional wind exerts a countervailing influence by transporting moisture from the southern region into the domain when reversal conditions are imposed, setting the stage for prolonged reversal events. However, due to the presence of large-scale moisture structures, convection can briefly cluster. Additionally, when organized conditions are imposed, the meridional wind contributes to convection organization by restraining the magnitude of reversals and maintaining a relatively more organized state compared to the control experiment, although it does not reach the level of organization observed in the experiment altering large-scale humidity.

5.5.2 Analyzing Wind Shear and Low-High Level Wind

Artificially modifying wind shear in WRF without affecting mass and momentum conservation is challenging. To address this, we harnessed the machine learning RF algorithm (as detailed in § 3.3) to emulate the SST-TCWV relationship (regression slope) using the variables identified in the previous sections as important for the regression. Despite the model's performance (R^2 = 0.97 in RF MAM simulation, see Figure C.4 in Appendix C), it's essential to validate its ability to replicate WRF's sensitivity experiments, including moisture advection, humidity, and wind (Figure 5.12). As the model relies on domain-mean values for training, direct modification of boundary conditions is not feasible. Instead, we replace domain-mean values from organized/reversal conditions for corresponding WRF simulation days to facilitate comparison. Notably, diverse ML techniques may yield different outcomes (McGovern et al., 2019), but our tests with alternative methods such as neural networks, support vector machines, and decision tree algorithms yielded consistent results with the RF algorithm, the latter demonstrating superior precision (not shown). For conciseness, we present results obtained using the RF algorithm.

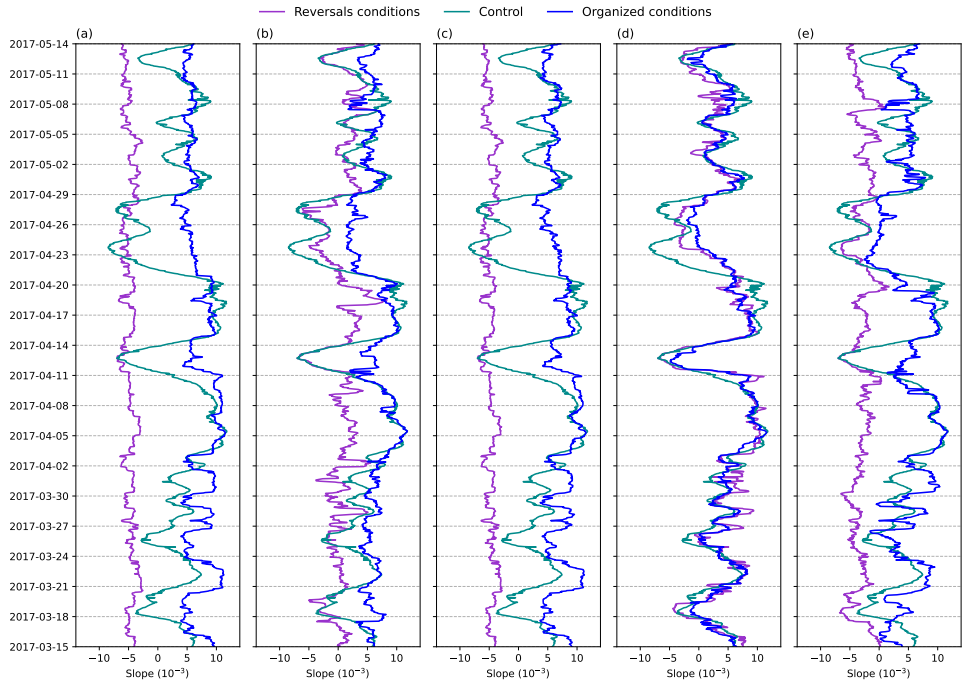

FIGURE 5.12: Random Forest simulations for the MAM season in the period simulated by WRF, for (a) Moisture Advection, (b) Moisture, (c) Wind, (d) Zonal wind and (e) Meridional Wind. Dark-cyan lines represent the control, the purple line represent the experiments that include reversals conditions and the blue line the ones with organized conditions.

Results underscore the RF's capability in emulating sensitivity experiments conducted within WRF (Figure 5.12). It effectively highlights the significance of moisture advection, humidity, and the meridional wind component while revealing that the influence of the zonal wind on clustering and random convection is minimal, although it can dampen reversal events. The RF model reliably reproduces the outcomes of experiments with altered humidity, but with less strength. In WRF, the introduction of humidity from reversal conditions results in random convection and breakup episodes. The RF experiment with reversal conditions replicates this behavior, displaying random states through the simulation, and also producing strong breakup-like episodes, notably on April 13th and between April 23rd and 27th, 2017.

In WRF, introducing organized moisture conditions significantly enhances convection organization, yielding weaker reversals compared to the control. The RF experiment with organized moisture conditions mirrors this pattern, displaying strong organization throughout the simulation with milder breakup episodes (compared to the control) occurring on the same dates. The RF model effectively mirrors WRF's results concerning the wind and its components, although it exaggerates the wind's role in the organized experiments, resulting in a significant weakening of breakup events that in WRF experiments is attributed to a combination of both the large-scale moisture structures and the wind. The RF model replicates the control simulation's slope evolution, and closely aligns with WRF sensitivity experiment outcomes (Figure 5.12). It characterizes the interplay between the slope and input variables, highlighting the essential role of the wind. We thus conclude that RF can serve as a tool

for a deeper exploration of the impact of surface wind and wind shear on either promoting or impeding convective organization.

In this context, we employed the RF model to conduct a series of experiments, wherein we altered the wind conditions by substituting them with data from a reversal event (2017-04-12 00:00:00) and an organized state (2017-03-20 18:00:00). An overview of these sensitivity experiments is provided in Table 5.2, encompassing variations such as the individual removal of the meridional and zonal wind shear, both separately and in combination, as well as the isolated removal of low and high tropospheric wind shear. These experiments were designed to unravel whether the surface or low level wind or the wind shear variables exert a more dominant influence on the emergence of breakup events or the formation of convection clusters.

Figure 5.13 offers a summary of the conducted experiments in which we eliminated meridional shear by replacing surface meridional wind with data either from a reversal period (MedSheRev) or subsequently from an organized (MedSheOrg) event throughout the troposphere (Figure 5.13a) . The MedSheRev experiment indicated that the absence of shear results in persistent random convection throughout the simulation, featuring breakup events of reduced intensity. This implies that meridional shear plays a minimal role in the development of breakup events compared to surface meridional wind. Conversely, the MedSheOrg experiment showed slightly more organized convection compared to the control, accompanied by a significant reduction in breakup episode strength. This emphasizes that surface meridional wind can mitigate breakup events even in the absence of shear. A comparison between the MedSheOrg and VOrg experiments reveals remarkable similarities, suggesting that shear does not significantly contribute to convection organization, and surface wind plays a more crucial role in promoting clustering.

Concerning the impact of low meridional shear in reversal (MLowSheRev) and organized (MLowSheOrg) conditions, the results align with those from the MedSheRev and MedSheOrg experiments, respectively. In the MLowSheRev experiment (Figure 5.13b), low tropospheric winds from the south persist, indicating continuous moisture import from the region to the south of the focus domain, where the warm pool center is located. This results in random convection conditions, even without shear. Conversely, the MLowSheOrg experiment (Figure 5.13b) demonstrates that substituting low-level wind with the surface wind of the organized state leads to slightly more organized convection and a weakening of breakup episodes. This highlights the role of low-level wind, especially at the surface, in establishing a southerly flow that imports moisture into the domain. Consistent with previous studies (e.g., Takayabu and Nitta 1993; Kiladis and Wheeler 1995), that showed how westward-propagating waves crossing the domain induce convergence, creating a southerly flow that further moistens the wave. This, in turn, advects deep convection to the northern region of the domain and triggers new convection in the southern region (see Chap4 for details).

The results indicate that high meridional shear (Figure 5.13c), observed in both reversal (MHigSheRev) and organized (MHigSheOrg) conditions, does not have a substantial impact on organization or the breakup of convection clusters. The outcomes closely resemble those of the control experiment, with the only notable distinction being slightly weaker breakup episodes. This suggests that while high-level wind does not play a significant role, it can still influence the strength of breakup

Simulation Name	Mean Wind Modification Modification
No Meridional Shear Reversals (MedSheRev)	Replace each meridional wind level of the entire simulation with the surface meridional wind from the strong reversal event of 2017-04-12 00:00:00
No Meridional Shear Organized (MedSheOrg)	Replace each meridional wind level of the entire simulation with the surface meridional wind from the strong organized event of 2017-03-20 18:00:00
No Meridional Low Tropospheric Shear Reversals (MLowSheRev)	Replace the meridional wind levels from 1000hPa to 650hPa of the entire simulation with the surface meridional wind from the strong reversal event of 2017-04-12 00:00:00
No Meridional Low Tropospheric Shear Organized (MLowSheOrg)	Replace the meridional wind levels from 1000hPa to 650hPa of the entire simulation with the surface meridional wind from the strong organized event of 2017-03-20 18:00:00
No Meridional High Tropospheric Shear Reversals (MHigSheRev)	Replace the meridional wind levels from 500hPa to 100hPa of the entire simulation with the 500hPa meridional wind from the strong reversal event of 2017-04-12 00:00:00
No Meridional High Tropospheric Shear Reversals (MHigSheOrg)	Replace the meridional wind levels from 500hPa to 100hPa of the entire simulation with the 500hPa meridional wind from the strong organized event of 2017-03-20 18:00:00
No Zonal Shear Reversals (ZonSheRev)	Replace each zonal wind level of the entire simulation with the surface zonal wind from the strong reversal event of 2017-04-12 00:00:00
No Zonal Shear Organized (ZonSheOrg)	Replace each zonal wind level of the entire simulation with the surface zonal wind from the strong organized event of 2017-03-20 18:00:00
No Zonal Low Tropospheric Shear Reversals (ZLowSheRev)	Replace the zonal wind levels from 1000hPa to 650hPa of the entire simulation with the surface zonal wind from the strong reversal event of 2017-04-12 00:00:00
No Zonal Low Tropospheric Shear Organized (ZLowSheOrg)	Replace the zonal wind levels from 1000hPa to 650hPa of the entire simulation with the surface zonal wind from the strong organized event of 2017-03-20 18:00:00
No Zonal High Tropospheric Shear Reversals (ZHigSheRev)	Replace the zonal wind levels from 450hPa to 100hPa of the entire simulation with the 450hPa meridional wind from the strong reversal event of 2017-04-12 00:00:00
No Zonal High Tropospheric Shear organized (ZHigSheOrg)	Replace the zonal wind levels from 450hPa to 100hPa of the entire simulation with the 450hPa meridional wind from the strong organized event of 2017-03-20 18:00:00
No Shear Reversals (NoSheRev)	Replace every zonal and meridional wind throughout the entire simulation with their respective surface zonal and meridional winds recorded during the strong reversal event on 2017-04-12 00:00:00
No Shear Organized (NoSheOrg)	Replace every zonal and meridional wind throughout the entire simulation with their respective surface zonal and meridional winds recorded during the strong organized event on 2017-03-20 18:00:00

TABLE 5.2: Overview of all the wind shear experiments performed using the Random Forest model.

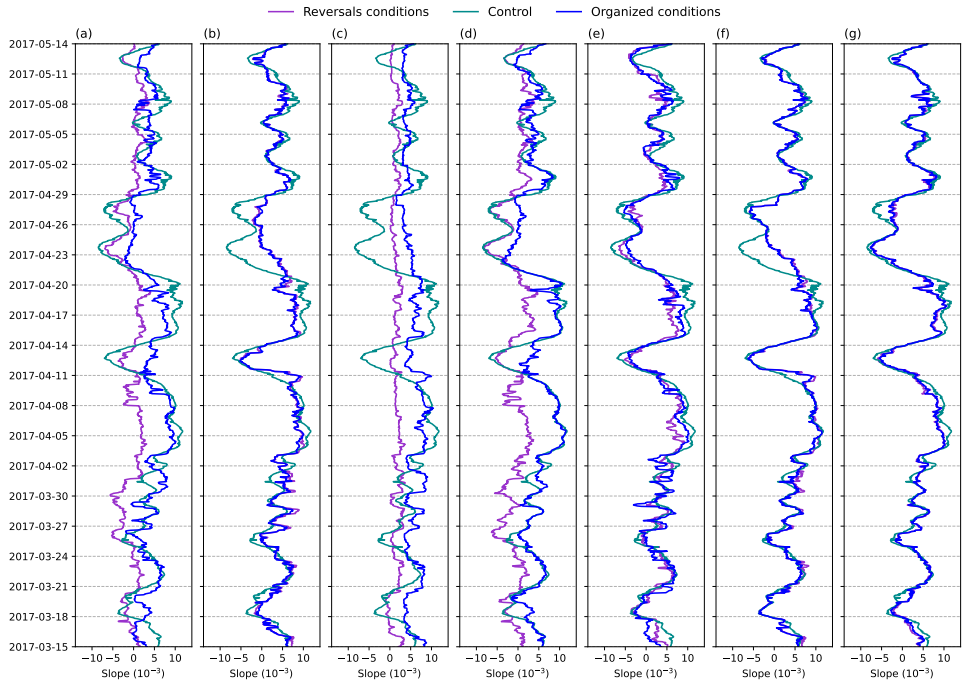

FIGURE 5.13: RF simulations for the MAM season in the period simulated by WRF, for (a) No Meridional-Shear. No Meridional (b) Low and (c) High Tropospheric Shear, (d) No Zonal-Shear. No Zonal (e) Low and (f) High Tropospheric Shear, and (g) No Shear. Dark-cyan lines represent the control, the purple line represent the experiments that include reversals conditions and the blue line the ones with organized conditions.

episodes, likely through a not very efficient transport of moisture from the south. Another important element is that the MHigRev produce weaker organization, meaning that although not significantly, high tropospheric shear could lead to less strong clustering.

Concerning the experiments aimed at eliminating zonal shear by substituting all zonal wind levels with the surface zonal wind values from both reversal (ZonSheRev) and organized (ZonSheOrg) events throughout the troposphere (Figure 5.13d), the results affirm the minimal influence of the zonal component on convection clusters, their prevention, or disbandment. Both the ZonSheRev and the ZonSheOrg, closely resemble the control conditions, with the only distinguishing feature being weaker and shorter breakup episodes, similar to the URev and UOrg experiments (Figure 5.12d). This weakening of the breakup episodes is more influenced by the low-tropospheric wind (ZLowSheRev and ZLowSheOrg), as illustrated in Figure 5.13e, rather than the high-tropospheric wind (ZHigSheRev and ZHigSheOrg), which demonstrates a minor impact primarily on the April 23rd to 29th breakup episode (Figure 5.13f). It is worth noting that the no zonal shear experiments (Figure 5.13d-e-f) occasionally exhibits reduced clustering, as evident in the period from April 15th to 20th, 2017, suggesting that shear can sometimes act to weakly decrease the degree of organization.

The final set of wind modification experiments involved the replacement of all

zonal and meridional wind levels with their respective surface values from a reversal (NoSheRev) and an organized (NoSheOrg) event, as presented in Figure 5.13g. Both experiments exhibit strong similarities to the WinRev and WinOrg experiments (Figure 5.12c), respectively. This suggests that surface (and low-level) wind plays the most significant role in shaping convective clusters, random convection, and breakup episodes. However, it's important to note some nuances. In the case of the NoSheRev experiment, it results in random convection conditions throughout the entire period, albeit with less intensity compared to the WinRev experiment. This implies that shear, most likely from low levels, can contribute to the promotion of random conditions and breakup events. The NoSheOrg experiment produces more organized convection than the control but less than the WinOrg experiment, especially in specific periods (e.g., April 14th to 20th, 2017), during which convection is less organized than in the control. This indicates that shear possibly from lower levels, can play a more prominent role in favoring convection organization under certain conditions.

In summary, the RF model accurately reproduces WRF boundary condition experiments, enabling an in-depth exploration of the role of wind in shaping convection patterns. Zonal wind's impact is constrained, mainly linked to low-level winds. On the contrary, meridional wind significantly influences convection clusters, breakup events, and random convection. Both surface and low-level meridional winds play an essential role in inducing intense breakup events or creating entirely random conditions. This is driven by the southerly flow associated with reversal episodes, transporting moisture into the domain and equalizing moisture disparities. Consequently, the majority of convection shifts to the northern part of the domain, and new convection is triggered in the southern region. In the absence of this southerly flow, clustered convection prevails with minimal reversal events, likely due to the presence of large-scale moisture structures and the absence of a southerly wind flow.

5.6 Overview of the Machine Learning Model Results

Figure 5.14 illustrates the Slope IQR (Inter-Quartile Range) as a function of the maximum slope across all RF experiments, including those which closely resemble to the WRF experiments. The arrows on the graph serve as indicators: a larger IQR corresponds to a more pronounced reversal/breakup episode, signifying a greater shift in the SST-TCWV regression relationship (slope). Additionally, the arrows demonstrate that the level of organization during a simulation period is contingent on the maximum slope; a higher maximum slope results in stronger clustering.

Our study, as summarized in Figure 5.14, underscores the crucial role of large-scale moisture advection in both organizing and maintaining random convection patterns. This process sustains convective clusters in the southern region within the domain (organized conditions) and facilitates their northward migration (reversals conditions), triggering new convection in the south and resulting in random convection. Key insights from Figure 5.14 include: (i) Zonal wind weakly contributes to maintaining clustering and inducing breakup episodes. (ii) The meridional wind significantly influences organized and random states, with a weak and southerly flow favoring organization, and a more substantial meridional wind from the south inducing breakup convection and random conditions. (iii) Moisture structures play

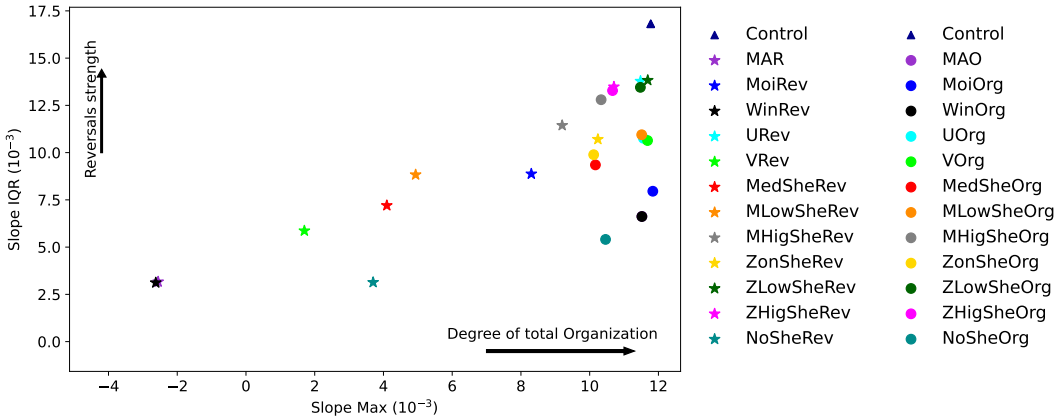

FIGURE 5.14: Slope IQR (75-25%-tile) as a function of Maximum Slope. The Slope IQR represent the strength of the reversal event, and the Maximum Slope (Slope Max) shows the degree of organization, as the plotted arrows indicate. The circles represent simulations that used organized conditions, the stars the ones that use reversal conditions, and the triangle is the control. The colors of the stars and the circles show the variable that is selected to replace with the organized/reversal condition. This means that if the simulation name in the legend is Meridional wind, this is the variable that was replaced with the organized/reversals conditions

a vital role in both organized and random conditions. When organized conditions are applied, they lead to strong clustering and weakened breakup episodes. Conversely, under reversal conditions, convection remains random, and breakup episodes occur concurrently with the control, emphasizing the importance of wind conditions, aligning with previous work (e.g., Takayabu and Nitta 1993; Takayabu 1994; Kiladis and Wheeler 1995; Haertel and Johnson 1998; Gonzalez and Jiang 2019) that emphasize the role of large-scale moisture dynamics in determining convection organization.

Interestingly, wind shear plays a role, though not as pronounced, in both convection clustering and breakup events. This aligns with findings from Helfer et al. (2020) and Helfer and Nuijens (2021) in trade-wind regions, indicating that shear can create conditions favorable for organization, depending on its direction and strength. Zonal shear, encompassing low-level and high-level shear, shows minimal impact compared to the control, consistent with the zonal wind experiments discussed earlier. On the other hand, experiments eliminating meridional shear reveal substantial effects on both organized and random conditions. Eliminating shear by substituting all meridional wind levels with surface wind from a reversal episode results in primarily random convection, principally driven by low-level winds creating a southerly flow (Takayabu and Nitta, 1993; Haertel and Johnson, 1998). Eliminating meridional shear by replacing all wind levels with surface wind from an organized event yields highly organized convection, notably reducing breakup episodes, similar to the meridional wind experiment. This reaffirms previous findings emphasizing the role of surface meridional wind (and low-level meridional winds) in influencing both the promotion and disruption of convection. Notably, despite weakened breakup episodes in the absence of meridional shear, they persist, potentially influenced by moisture structures, an important factor in reversal episodes as well.

Other significant factors warrant consideration. Runs exhibiting the strongest reversals also demonstrate strong clustering. The relationship between SST-TCWV undergoes the most significant change when convection transitions from a highly organized configuration to a random one, typically due to a westward-propagating wave. This phenomenon is particularly evident in the control experiment, characterized by significant clustering, and experiences disruptions from westward-propagating waves, leading to pronounced changes in the SST-TCWV relationship and strong reversal events. Conversely, runs with maximum slope > 8 and slope IQR > 12 exhibit strong clustering and reversals but feature less intense waves due to modifications in key variables, diminishing the wave strength, and subsequently the reversal's magnitude.

The MoiOrg experiment demonstrates the highest degree of clustering, surpassing even the control, MAO, or WinOrg experiments, highlighting moisture's essential role (Tompkins, 2001c). Nonetheless, MoiOrg still produces weak reversal episodes, stronger than MAO and WinOrg, indicating wind's importance in breakup event development. On the other hand, in the VRev, MLowSheRev, MedSheRev, and NoSheRev experiments, convection stays mostly random, but with episodes of weak clustering, which is why those runs lays in intermediate levels of clustering and reversal strength.

5.7 Chapter Summary and Conclusions

Numerous studies highlight the significant impact of convective organization on climate sensitivity and the hydrological cycle. In Chap4 we introduce an index based on the SST-TCWV relationship, revealing that in the Pacific Warm Pool, boreal-winter convection tends to organize into clusters, disrupted by westward-propagating waves. This clustering results in a drier domain, allowing more LW radiation to escape the atmosphere, confirming the role of organization into influencing climate feedbacks and the hydrological cycle. In boreal-summer convection maintains a random configuration despite strong diabatic and adiabatic feedbacks favoring clustering.

This study aims to unravel the mechanisms driving clustered convection, breakup events during boreal/winter, and the processes leading to random convection in the Pacific Warm Pool. Through observational analysis akin to Chap4, we identify moisture advection as an essential variable determining the spatial arrangement of convection. Utilizing the WRF model, we simulate the MAM and JJA seasons, designating MAM as the control. We manipulate the large-scale moisture advection in MAM by replacing it with the moisture advection conditions of a reversal (MAR) and an organized (MAO) state. These experiments demonstrate that moisture advection can induce either fully random or clustered convection throughout the MAM season. In contrast, the JJA simulation exhibits random convection despite potent diabatic feedbacks.

Employing a stream function, we observe a moist gross static energy (MGSE) export from moist to dry regions during JJA, counteracting diabatic feedbacks. This effect is more pronounced in the MAR experiment, where strong diabatic feedbacks are compensated by MGSE export facilitated by large-scale moisture structures. In the MAM season, MGSE is imported into moist regions, likely due to large-scale

dynamics, coupled with favorable diabatic feedbacks, and including SST gradients, and moisture disparities, that promote organized convection. This multiple aspects that favor clustering are further strengthens in the MAO experiment.

Having established that moisture structures influence both clustered and random convection, we then attempted to comprehend the roles of large-scale moisture structures and the wind, including its components. Substituting large-scale moisture fields at the domain boundaries with those from a single day with organized conditions results in strong clustering throughout the entire simulation. This substitution also mitigates breakup episodes but does not entirely prevent them, underscoring the crucial role of winds. Replacing moisture with reversal conditions induces random scenarios with breakup-like episodes. These episodes disperse all the cores, interspersed with brief clustered periods influenced by wind conditions.

In terms of wind influence, organized conditions result in strong organization (though not as pronounced as in the moist boundary conditions experiment), coupled with weaker breakup episodes than in the control, underscoring its role in shaping reversal conditions. This also emphasizes that moisture boundary conditions alone can induce breakup episodes, since the wind perturbations are not able to fully prevent them. Reversal wind conditions exhibit a weak zonal wind component and a southerly meridional component, creating a sustained influx of moisture to the domain. This displaces deep convection northward while initiating new convection in the southern region within the domain. This scenario is reminiscent of Chap4 conditions, where a westward-propagating wave induces a similar southerly flow, driven by convergence produce by the wave.

The wind results hint that the zonal component makes a minimal contribution to generating organized or random convection, as demonstrated by our experiments imposing the zonal wind of both reversal and organized states. In contrast, the meridional wind exhibits a behavior similar to the wind experiments with reversals conditions, results produce an almost entirely random convection configuration throughout the entire MAM season. This is attributed to the mentioned moisture influx from the south, where the center of the warm pool is located. The experiment also briefly produces cluster conditions that quickly disband, revealing that moisture boundary conditions can drive organization even in the presence of disfavoring meridional wind, albeit less strongly (and shortly). Meridional wind from organized conditions can lead to stronger organization than normal, with weak breakup episodes, underscoring that moisture structures alone can induce breakup episodes, although less prominently than in combination with the meridional wind. In essence, moisture boundary conditions exhibit greater efficiency in transporting moisture from dry to moist regions, whereas meridional wind boundary conditions emerge as the primary driver exporting MGSE from humid to dry regions, particularly during reversal episodes.

Having elucidated the wind conditions influencing random and clustered convection, we further investigate the roles of surface winds, low-level winds, high-level winds, and wind shear. This exploration involves a new set of experiments employing a Random Forest model with high precision in replicating the SST-TCWV relationship (slope) and WRF sensitivity experiments. Unlike the WRF model, this

approach allows modifications to large-scale wind conditions without inducing momentum imbalances. Results indicate that the zonal wind component's weak contribution to both organized and reversal states emanates from surface and low-level winds. A parallel finding emerges for the meridional component, where the impact is attributed to surface and low-level meridional winds, even in the absence of wind shear at low levels. Wind shear, on the other hand, exhibits a dual role, capable of both promoting and hindering clustering, aligning with findings from other studies (e.g., Helfer et al. 2020; Helfer and Nuijens 2021).

The findings presented in this chapter underscores the predominant role of large-scale dynamics in clustering convection, disrupting it, and also producing random convection. Spatial diabatic feedbacks appear less influential in determining when clustering occurs but rather amplify it through the impact of dynamics. Large-scale dynamics can induce a MGSE export from the moist to the dry regions, offsetting the diabatic positive feedbacks.

Chapter 6

Summary and Conclusions

Deep convection organization impacts the radiative budget and hydrological cycle. One crucial facet of deep convection, with the potential to influence climate sensitivity and alter the hydrological cycle by modifying storm intensity and atmospheric humidity, is its organization. Examples of convection organization include mesoscale convective systems such as squall lines and tropical cyclones, and larger scale coupled waves including the Madden-Julian Oscillation. Organized convection contributes to an increase in precipitation extremes. Clustered convection increase clear-sky OLR as a result of diminishing moisture relative to a state where convection is random. Depending on the reduction of the cloud fraction resulting from organized convection, the all-sky OLR can also be affected, introducing potential climate sensitivities that remain incompletely understood. Consequently, the scientific community has employed idealized models, observations, and realistic simulations to enhance our comprehension of the mechanisms driving organization and its implications for the atmospheric mean state. One intriguing phenomenon observed in idealized convection permitting modelling studies within the radiative convective equilibrium framework is convective self aggregation (SA). In SA, convection clusters despite originating from homogeneous initial conditions and in the absence of external forcings at the domain boundaries, such as large-scale forcing, waves, or SST heterogeneities. Extensive research has revealed that convective SA can occur in different models and domain configurations; however, it may be impeded in smaller domains. In particular, previous research has indicated that the occurrence of SA can be very sensitive to the choice of specific combinations of parameterization schemes employed. This is why the focus of the first objective of this thesis is to comprehend the factors driving the sensitivity of SA to the model parameterizations.

Due to the convective SA impact on the domain mean in idealized models, researchers have turned to observations and realistic simulations to discern its consequences in the atmosphere, with the specific objective of identify SA's effects on the radiative budget and the hydrological cycle. Authors have shown that SA can influence the tropical atmosphere's radiative budget, affecting OLR, and that organized convection results in intensified daily precipitation with associated drying in the sub-tropics. However, studying organization in observations poses challenges in quantification, with no consensus on methodology despite numerous proposed indices. Addressing this gap, the second objective of this thesis focuses on developing a fingerprint index to quantify organization based on its consequences on the humidity distributions and relationship between humidity and SST, rather than the direct position of convection. The index is employed to analyze the impacts of organization on the domain mean cloud fraction, OLR and humidity in three GCM-like grid boxes situated in the Pacific Warm Pool. The study also investigates the mechanisms driving organized and random convection, drawing comparisons with

findings from idealized modeling studies. Building on the insights from this objective, the third objective of this thesis focused on understanding the mechanisms fostering or impeding organization in the warm pool. It also seeks to uncover why organized convection can undergo breakup episodes that could also modify the domain's mean properties.

In the following, the conclusions of the three objectives outlined in this paper, starting from the general question we sought to address, a question which is also described in the introduction:

Question 1: What is the sensitivity of self-aggregation to the model subgrid-scale parameterizations?

Chapter 3 aimed to comprehend the mechanisms influencing the occurrence or prevention of convective SA in diverse idealized modeling scenarios. Using the WRF model, we explored 24 different combinations of parameterizations, including sub-grid scale mixing, PBL, and microphysics, without altering domain characteristics. Our analysis emphasizes the critical role of factors that impact the convective spacing in the initial phase of near random convection, as exemplified by the maximum free convective distance (d_{clr}) metric in determining SA transition or maintenance of a random configuration.

In essence, convective SA propensity increases with higher d_{clr}, particularly when cold pools (CPs) exhibit limited size, intensity, and gust front convergence. This is significant since weaker/smaller and less convergent CPs lead to fewer collisions, decreased moisture convergence (and moisture redistribution), leading to a lower core count, and to an increased d_{clr}—favoring aggregation. Conversely, larger d_{clr} and weaker/smaller CPs elevate the likelihood of a dry region emerging, which can act as a starting point for the aggregation of convection. When a dry region develops, subsidence intensifies the cooling and desiccation of non-convective regions, resulting in dry anomalies and positive pressure anomalies. This sets off a divergent flow that further amplifies regional drying, establishing a feedback mechanism.

Interconnected variables influencing d_{clr} underscore its crucial role in SA development, showcasing intricate an interplay: (i) Horizontal mixing (sub-grid scale) induces entrainment, amplifying transported mass within clouds to yield large convective cores, escalating d_{clr} and leading to SA. (ii) Rain evaporation (microphysics) impacts CP intensity/size, convective core count, and d_{clr}. Weaker evaporation hampers moisture redistribution and CP gust fronts, influencing convective core number, d_{clr}, and favoring SA. (iii) Non-local schemes induce heightened low-cloud cover, even in dry regions, by drawing in air from the upper atmosphere. This increased low-cloud cover enhances radiative cooling, which reinforces a circulation pattern that retains moisture in moist regions, thereby favoring self-aggregation. Moreover, the intensified radiative cooling stimulates stronger subsidence, which suppresses convective initiation, reduces convective core frequency, expands the d_{clr}, and ultimately promotes SA.

In general, the sub-grid scale mixing scheme stands out as the primary driver of organized convection, regardless of the chosen PBL or microphysics scheme, making it the most influential factor for convective SA. Microphysics schemes can also play a significant role (although not as the sub-grid schemes), with certain schemes either

facilitating or inhibiting SA. PBL schemes, although still relevant, are of secondary importance. Notably, in cases where the Thompson/GCE microphysics scheme is paired with TKE/Smag2, the locality of the PBL emerges as the key determinant for SA, with non-local schemes promoting SA and local schemes inhibiting it.

Question 2: Can we find convective organization in a GCM grid-box sized domain? And if so, what are its impact on the hydrological and radiative budget?

Chapter 4 dived into the intricate relationship between convection organization, water vapor variability, and their impact on the tropical energy budget. While the mean activity of deep convection has been well-understood, recent insights have highlighted the crucial role of mesoscale convection distribution, particularly at alpha and beta meso scales, in shaping interannual variations in water vapor.

We employed a multivariate analysis, commonly used in idealized cloud-resolving studies, to unravel the impacts of convective organization. Findings revealed that in a clustered state, the distribution of TCWV widens due to the presence of dry columns, with positive correlations between SST and water vapor. Conversely, in a more randomly organized convection, the water vapor distribution narrows, featuring high mean values and a negative SST-TCWV correlation.

Focusing on the northern of the equator study regions of the mesoscale domain located in the warm pool region of the tropical western Pacific, the analysis uncovered distinct behaviors during different seasons. In summer/autumn, when spatial SST gradients are small, convection tends to align with a randomly organized pattern, despite diabatic forcings favoring clustering. In contrast, during boreal winter/spring, with a weak north-south SST gradient, a predominantly aggregated convection state emerges, intermittently interrupted by episodes of randomly organized convection. These interruptions lead to a moister atmosphere and reduced humidity gradients.

Utilizing the SST-TCWV correlation as an index for the clustered state, three composites of 44 (2N-9N box), 46 (3N-10N box), and 41 (3S-4N box) random events during the study period were constructed. Diabatic forcings, including radiative-cloud and moisture feedbacks, consistently acted to cluster convection. However, the surface LH flux feedback, typically favoring clustering in a random state, reversed its role when convection becomes aggregated. This feedback is dependent on wind speed rather than moisture structures, diverging from the results observed in idealized models. Sporadic breakup of clustering during boreal winter/spring was attributed to westward propagating waves, inducing a counter-gradient convective activity that resulted in a more homogeneous humidity distribution. This phenomenon significantly impacted TOA OLR, with an approximate significant decrease of 10 W m^{-2} due to cloud cover changes, in the 3N-10N box, the all-sky OLR change was not significance due to stable cloud cover in both organized and random conditions.

Question 3: Can realistic simulations help to understand the mechanisms that organize or disorganize convection in a mesoscale sized domain?

Building on Chapter 4, Chapter 5 focus on understanding the drivers of clustered

convection, breakup events in boreal winter/spring, and the processes behind random convection in the Pacific Warm Pool during boreal summer/autumn. Focusing on the MAM and JJA seasons, with MAM as the control, we manipulate large-scale moisture advection by introducing reversal (MAR) and organized (MAO) state conditions. These experiments revealed that diabatic feedback always favor convection organization. However, they also showed that moisture advection can induce either fully random or clustered convection in MAM. The JJA simulation exhibits random convection, counteracting potent diabatic feedbacks through moist gross static energy (MGSE) export from moist to dry regions, akin to the MAR experiment. In contrast, the MAM and MAO experiments showcase MGSE export from dry to moist regions.

Having established the influence of moisture structures on both clustered and random convection, we then focused on its components. Substituting large-scale moisture structures with those sampled in conditions with organized convection promotes strong clustering throughout the simulation, mitigating breakup episodes but not entirely preventing them—highlighting the role of the wind. Replacing moisture with reversal conditions induces random scenarios with breakup-like episodes throughout the entire simulation. In terms of large-scale wind, organized conditions lead to strong organization (though less pronounced than in the moist structures experiment) and weaker breakup episodes, underscoring its role in shaping reversal conditions. Reversal wind conditions feature a weak zonal wind component and a southerly meridional component, fostering a sustained influx of moisture into the domain. This displaces deep convection northward while initiating new convection in the southern region.

The wind component results reveal a minimal contribution from the zonal component to organized or random convection, with a predominant influence from low-level winds. Conversely, reversal conditions in meridional winds, especially at low levels, result in almost entirely random convection during the MAM season. Briefly clustered conditions can arise, but quickly dissipate, indicating that moisture structures can drive organization despite a less favorable meridional wind, albeit less strongly and briefly. Meridional winds, particularly from low levels, under organized conditions foster stronger organization than usual, accompanied by weak breakup episodes, highlighting that moisture structures alone can induce breakup, albeit less prominently. Wind shear plays a dual role, capable of both promoting and hindering clustering depending on its strength and direction.

The findings presented in this research underscore the predominant role of large-scale dynamics in clustering convection, disrupting it, and also producing random convection. Spatial diabatic feedbacks appear less influential in fostering clustering, as large-scale dynamics can induce MGSE export from moist to dry regions, offsetting the diabatic positive feedbacks. The results also indicate that breaking up convection clusters requires a strong perturbation, suggesting the potential development of hysteresis, akin to idealized models.

Future Research

Our research raised new questions that warrant deeper exploration. In Chapter 3, we analyze the role of rain evaporation, nevertheless, an unexplored aspect revolves around the microphysical parameters that dictate the number of convective

cores. The intricate involvement of cold pools in SA prompts investigation on various fronts, including the use of simplified models. This exploration aims to unravel the moisture-redistribution impact of cold pools, probe into their potential to reconfigure both convective core number and d_{clr}, and comprehend how factors such as cold pool intensity, size, gust front convergence, and moisture redistribution collectively shape the emergence of SA. Additionally, our findings suggest that convection lifetimes may play a role in influencing SA, as enduring clouds have the potential to moisten a specific location, creating favorable conditions for convection and thereby fostering SA.

Chapters 4 and 5 also highlight potential research avenues. An interesting hypothesis we propose involves the long-term variations in the convective organization index, linked to the tropics-wide energy budget. We speculate that these variations may be influenced by the frequency of westward-propagating waves that disrupt convection clusters in the Pacific Warm Pool, potentially due to their ability to redistribute moisture across the tropics. Exploring the impact of intensified convective organization on daily extreme precipitation events in the Pacific Warm Pool and assessing the potential influence of breakup episodes on these events would be an intriguing avenue of research. Another compelling area of study involves investigating how convective organization might evolve in the future, particularly under warmer SSTs in the Pacific Warm Pool. Understanding whether these changes could contribute to climate sensitivity or modify the overall organizational impacts on the domain mean presents a valuable research opportunity. To address this, conducting pseudo-climate change realistic simulations spanning 4 or 6 years in the warm pool and comparing the levels of organization and their effects during these periods with the present climate could offer valuable insights.

Furthermore, the broader implications stemming from the findings in Chapter 4 propose a potential connection between mesoscale organization, quantified by the I_{org} index, and mid-tropospheric humidity, as elucidated by Bony et al. (2020), with the variability in the activity of westward-propagating waves in the tropical western Pacific region. Grasping these complex dynamics is paramount for unraveling the primary determinants of convection organization and water vapor variability in the warm pool, holding far-reaching significance for the tropical energy budget.

In summation, these questions underscore the avenues yet to be traversed in comprehending the intricacies of convective organization (including SA). Addressing these inquiries promises would allow to enhance our understanding of the phenomenon and also to illuminate novel facets of its governing mechanisms.

Side Projects

Utilizing the insights gained during my Ph.D., I had the opportunity to lead and co-author several side projects, some of which resulted in publications. While these publications may not directly align with the main theme of my thesis, they were facilitated by the methods and diagnostics developed for this thesis. In Appendix D, a concise summary of these publications is provided, including a brief explanation of their connection to the methods developed in the thesis and respective citation.

Bibliography

Abbot, D., 2014: Resolved snowball earth clouds. *Journal of Climate*, **27 (12)**, 4391–4402, doi:10.1175/JCLI-D-13-00738.1.

Abramian, S., C. Muller, and C. Risi, 2022a: Shear-convection interactions and orientation of tropical squall lines. *Geophysical Research Letters*, **49 (1)**, e2021GL095 184, doi:https://doi.org/10.1029/2021GL095184.

Abramian, S., C. Muller, and C. Risi, 2022b: Shear-Convection Interactions and Orientation of Tropical Squall Lines. *Geophysical Research Letters*, **49 (1)**, e2021GL095 184, doi:https://doi.org/10.1029/2021GL095184.

Angulo-Umana, P., and D. Kim, 2023: Mesoscale convective clustering enhances tropical precipitation. *Science Advances*, **9 (2)**, eabo5317, doi:10.1126/sciadv.abo5317.

Arnold, N. P., and D. A. Randall, 2015: Global-scale convective aggregation: Implications for the madden-julian oscillation. *Journal of Advances in Modeling Earth Systems*, **7 (4)**, 1499–1518, doi:https://doi.org/10.1002/2015MS000498.

Awad, M., and R. Khanna, 2015: *Machine Learning*, 1–18. Apress, Berkeley, CA, doi:10.1007/978-1-4302-5990-9_1.

Bao, J., and S. C. Sherwood, 2019: The Role of Convective Self-Aggregation in Extreme Instantaneous Versus Daily Precipitation. *Journal of Advances in Modeling Earth Systems*, **11 (1)**, 19–33, doi:https://doi.org/10.1029/2018MS001503.

Becker, T., and A. A. Wing, 2020: Understanding the Extreme Spread in Climate Sensitivity within the Radiative-Convective Equilibrium Model Intercomparison Project. *Journal of Advances in Modeling Earth Systems*, **12 (10)**, e2020MS002 165, doi:https://doi.org/10.1029/2020MS002165.

Behrendt, A., and coautors, 2011: Observation of convection initiation processes with a suite of state-of-the-art research instruments during cops iop 8b. *Quarterly Journal of the Royal Meteorological Society*, **137 (S1)**, 81–100, doi:https://doi.org/10.1002/qj.758.

Behrens, G., T. Beucler, P. Gentine, F. Iglesias-Suarez, M. Pritchard, and V. Eyring, 2022: Non-linear dimensionality reduction with a variational encoder decoder to understand convective processes in climate models. *Journal of Advances in Modeling Earth Systems*, **14 (8)**, e2022MS003 130, doi:https://doi.org/10.1029/2022MS003130.

Benjamin, T. B., 1968: Gravity currents and related phenomena. *Journal of Fluid Mechanics*, **31 (2)**, 209–248, doi:10.1017/S0022112068000133.

Bessho, K., and coautors, 2016: An Introduction to Himawari-8/9; Japans New-Generation Geostationary Meteorological Satellites. *Journal of the Meteorological Society of Japan. Ser. II*, **94 (2)**, 151–183, doi:10.2151/jmsj.2016-009.

Betts, A., and C. Jakob, 2002: Study of diurnal convective precipitation over Amazonia using single column mode. *J. Geophys. Res*, **D23 (107)**, 4732, doi:10.1029/2002JD002264.

Beucler, T., T. H. Abbott, T. W. Cronin, and M. S. Pritchard, 2019: Comparing Convective Self-Aggregation in Idealized Models to Observed Moist Static Energy Variability Near the Equator. *Geophysical Research Letters*, **46 (17-18)**, 10 589–10 598, doi:https://doi.org/10.1029/2019GL084130.

Beucler, T., D. Leutwyler, and J. M. Windmiller, 2020: Quantifying convective aggregation using the tropical moist margin's length. *Journal of Advances in Modeling Earth Systems*, **12 (10)**, e2020MS002 092, doi:https://doi.org/10.1029/2020MS002092.

Biagioli, G., and A. Tompkins, 2023a: A Dimensionless Parameter for Predicting Convective Self-Aggregation Onset in a Stochastic Reaction-Diffusion Model of Tropical Radiative-Convective Equilibrium. *Journal of the Advances in Modeling Earth Systems*, **15 (5)**, e2022MS003 231, doi:10.1029/2022MS003231.

Biagioli, G., and A. M. Tompkins, 2023b: Measuring convective organization. *Journal of the Atmospheric Sciences*, doi:https://doi.org/10.1175/JAS-D-23-0103.1.

Bläckberg, C. P. O., and M. S. Singh, 2022: Increased large-scale convective aggregation in cmip5 projections: Implications for tropical precipitation extremes. *Geophysical Research Letters*, **49 (9)**, e2021GL097 295, doi:https://doi.org/10.1029/2021GL097295.

Böing, S. J., H. J. J. Jonker, A. P. Siebesma, and W. W. Grabowski, 2012: Influence of the Subcloud Layer on the Development of a Deep Convective Ensemble. *J. Atmos. Sci.*, **69 (9)**, 2682–2698, doi:10.1175/JAS-D-11-0317.1.

Bony, S., A. Semie, R. Kramer, B. Soden, A. Tompkins, and K. Emanuel, 2020: Observed Modulation of the Tropical Radiation Budget by Deep Convective Organization and Lower-Tropospheric Stability. *AGU Advances*, **1 (3)**, e2019AV000 155, doi:https://doi.org/10.1029/2019AV000155.

Bony, S., and coautors, 2015: Clouds, circulation and climate sensitivity. *Nature Geoscience*, **8 (4)**, 261–268, doi:10.1038/ngeo2398.

Bougeault, P., and P. Lacarrere, 1989: Parameterization of Orography-Induced Turbulence in a Mesobeta–Scale Model. *Monthly Weather Review*, **117 (8)**, 1872–1890, doi:10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.

Breiman, L., 2001: Random forests. *Machine learning*, **45 (1)**, 5–32.

Bretherton, C. S., and P. N. Blossey, 2017: Understanding mesoscale aggregation of shallow cumulus convection using large-eddy simulation. *Journal of Advances in Modeling Earth Systems*, **9 (8)**, 2798–2821, doi:https://doi.org/10.1002/2017MS000981.

Bretherton, C. S., and M. F. Khairoutdinov, 2015: Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *Journal of Advances in Modeling Earth Systems*, **7 (4)**, 1765–1787, doi:https://doi.org/10.1002/2015MS000499.

Bretherton, C. S., J. R. McCaa, and H. Grenier, 2004: A New Parameterization for Shallow Cumulus Convection and Its Application to Marine Subtropical Cloud-Topped Boundary Layers. Part I: Description and 1D Results. *Mon. Wea. Rev.*, **132 (4)**, 864–882, doi:10.1175/1520-0493(2004)132<0864:ANPFSC>2.0.CO;2.

Bretherton, S., P. Blossey, and M. Khairoutdinov, 2005: An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST. *Journal of Atmospheric Sciences*, **62 (12)**, 4273–4292, doi:10.1175/JAS3614.1.

Brigham, E. O., and R. E. Morrow, 1967: The fast fourier transform. *IEEE Spectrum*, **4 (12)**, 63–70, doi:10.1109/MSPEC.1967.5217220.

Brune, S., S. Buschow, and P. Friederichs, 2020: Observations and high-resolution simulations of convective precipitation organization over the tropical atlantic. *Quarterly Journal of the Royal Meteorological Society*, **146 (729)**, 1545–1563, doi:https://doi.org/10.1002/qj.3751.

Brune, S., S. Buschow, and P. Friederichs, 2021: The local wavelet-based organization index – quantification, localization and classification of convective organization from radar and satellite data. *Quarterly Journal of the Royal Meteorological Society*, **147 (736)**, 1853–1872, doi:https://doi.org/10.1002/qj.3998.

Brune, S., F. Kapp, and P. Friederichs, 2018: A wavelet-based analysis of convective organization in icon large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, **144 (717)**, 2812–2829, doi:https://doi.org/10.1002/qj.3409.

Casallas, A., M. Castillo-Camacho, E. Sanchez, Y. Gonzalez, N. Celis, J. Mendez-Espinosa, L. Belalcazar, and C. Ferro, 2023a: Surface, satellite ozone variations in northern south america during low anthropogenic emission conditions: a machine learning approach. *Air Quality, Atmosphere, and Health*, **16**, 745–764, doi:https://doi.org/10.1007/s11869-023-01303-6.

Casallas, A., M. P. Castillo-Camacho, M. A. Guevara-Luna, Y. González, E. Sanchez, and L. C. Belalcazar, 2022: Spatio-temporal analysis of pm2.5 and policies in northwestern south america. *Science of The Total Environment*, **852**, 158 504, doi:https://doi.org/10.1016/j.scitotenv.2022.158504.

Casallas, A., D. Hernandez-Deckers, and M.-P. H, 2023b: Understanding convective storms in a tropical, high-altitude location with in-situ meteorological observations and GPS-derived water vapor. *Atmosfera*, **36 (2)**, 225–238, doi:https://doi.org/10.20937/ATM.53051.

Casallas, A., and coautors, 2023c: Air pollution analysis in northwestern south america: A new lagrangian framework. *Science of the Total Environment*, **905**.

Cerlini, P. B., M. Saraceni, and L. Silvestri, 2023: Competing effect of radiative and moisture feedback in convective aggregation states in two crms. *Journal of Advances in Modeling Earth Systems*, **15 (2)**, e2022MS003 323, doi:https://doi.org/10.1029/2022MS003323.

Chen, G., 2022: A Model of the Convectively Coupled Equatorial Rossby Wave over the Indo-Pacific Warm Pool. *Journal of the Atmospheric Sciences*, **79 (9)**, 2267–2283, doi:10.1175/JAS-D-21-0326.1.

Chen, Y., and coautors, 2022: Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover. *Nature Geoscience*, **15**, 609–614, doi:10.1038/s41561-022-00991-6.

Cohen, B. G., and G. C. Craig, 2004: The response time of a convective cloud ensemble to a change in forcing. *Quarterly Journal of the Royal Meteorological Society*, **130 (598)**, 933–944, doi:https://doi.org/10.1256/qj.02.218.

Cohen, B. G., and G. C. Craig, 2006: Fluctuations in an Equilibrium Convective Ensemble. Part II: Numerical Experiments. *Journal of the Atmospheric Sciences*, **63 (8)**, 2005 – 2015, doi:https://doi.org/10.1175/JAS3710.1.

Colman, R., and B. J. Soden, 2021: Water vapor and lapse rate feedbacks in the climate system. *Reviews of Modern Physics*, **93 (4)**, 045 002, doi:10.1103/RevModPhys.93.045002.

Colón-González, F., M. Odhiambo-Sewe, A. Tompkins, H. Sjödin, A. Casallas, J. Rocklöv, C. Caminade, and R. Lowe, 2021: Projecting the risk of mosquito-borne diseases in a warmer and more populated world: a multi-model, multi-scenario intercomparison modelling study. *The Lancet Planetary Health*, **5 (7)**, e404–e414, doi:https://doi.org/10.1016/S2542-5196(21)00132-7, URL https://www.sciencedirect.com/science/article/pii/S2542519621001327.

Coppin, D., and S. Bony, 2015: Physical mechanisms controlling the initiation of convective self-aggregation in a General Circulation Model. *Journal of Advances in Modeling Earth Systems*, **7 (4)**, 2060–2078, doi:https://doi.org/10.1002/2015MS000571.

Craig, G., and J. Mack, 2013: A coarsening model for self-organization of tropical convection. *Journal of Geophysical Research: Atmospheres*, **118 (16)**, 8761–8769, doi:https://doi.org/10.1002/jgrd.50674.

Da Silva, N. A., C. Muller, S. Shamekh, and B. Fildier, 2021a: Significant Amplification of Instantaneous Extreme Precipitation With Convective Self-Aggregation. *Journal of Advances in Modeling Earth Systems*, **13 (11)**, e2021MS002 607, doi:https://doi.org/10.1029/2021MS002607.

Da Silva, N. A., C. Muller, S. Shamekh, and B. Fildier, 2021b: Significant amplification of instantaneous extreme precipitation with convective self-aggregation. *Journal of Advances in Modeling Earth Systems*, **13 (11)**, e2021MS002 607, doi:https://doi.org/10.1029/2021MS002607.

Dawson, D. T., M. Xue, J. A. Milbrandt, and M. K. Yau, 2010: Comparison of Evaporation and Cold Pool Development between Single-Moment and Multimoment Bulk Microphysics Schemes in Idealized Simulations of Tornadic Thunderstorms. *Monthly Weather Review*, **138 (4)**, 1152 – 1171, doi:https://doi.org/10.1175/2009MWR2956.1.

Droegemeier, K., and R. Wilhelmson, 1985: Three-Dimensional Numerical Modeling of Convection Produced by Interacting Thunderstorm Outflows. Part I: Control Simulation and Low-Level Moisture Variations. *Journal of Atmospheric Sciences*, **42 (22)**, 2381–2403, doi:10.1175/1520-0469(1985)042<2381:TDNMOC>2.0.CO;2.

Dudhia, J., 1989: Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of*

Atmospheric Sciences, **46 (20)**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Dudhia, J., S.-Y. Hong, and K.-S. Lim, 2008: A new method for representing mixed-phase particle fall speeds in bulk microphysics parameterizations. *Journal of the Meteorological Society of Japan*, **86 (1)**, 33–44.

Dyer, A., and B. Hicks, 1970: Flux-gradient relationships in the constant flux layer. *Quarterly Journal of the Royal Meteorological Society*, **96 (410)**, 715–721, doi:https://doi.org/10.1002/qj.49709641012.

Emanuel, K., 1991: A Scheme for Representing Cumulus Convection in Large-Scale Models. *Journal of Atmospheric Sciences*, **48 (21)**, 2313–2329, doi:10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2.

Emanuel, K., A. Wing, and E. Vincent, 2014: Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, **6 (1)**, 75–90, doi:https://doi.org/10.1002/2013MS000270.

Emanuel, K. A., 1994: *Atmospheric convection*. Oxford University Press, Inc., New York, 580 pp.

Fuchs-Stone, Z., D. J. Raymond, and S. Sentić, 2019: A simple model of convectively coupled equatorial rossby waves. *Journal of Advances in Modeling Earth Systems*, **11 (1)**, 173–184, doi:https://doi.org/10.1029/2018MS001433.

Fuglestvedt, H., and J. Haerter, 2020: Cold Pools as Conveyor Belts of Moisture. *Geophysical Research Letters*, **47 (12)**, e2020GL087 319, doi:https://doi.org/10.1029/2020GL087319.

Galewsky, J., M. Schneider, C. Diekmann, A. Semie, S. Bony, C. Risi, K. Emanuel, and H. Brogniez, 2023: The influence of convective aggregation on the stable isotopic composition of water vapor. *AGU Advances*, **4 (5)**, e2023AV000 877, doi:https://doi.org/10.1029/2023AV000877.

Gehne, M., B. Wolding, J. Dias, and G. N. Kiladis, 2022: Diagnostics of tropical variability for numerical weather forecasts. *Weather and Forecasting*, **37 (9)**, 1661 – 1680, doi:10.1175/WAF-D-21-0204.1, URL https://journals.ametsoc.org/view/journals/wefo/37/9/WAF-D-21-0204.1.xml.

Ghahremanloo, M., Y. Lops, Y. Choi, and B. Yeganeh, 2021: Deep learning estimation of daily ground-level NO2 concentrations from remote sensing data. *Journal of Geophysical Research: Atmospheres*, **126**, e2021JD034 925, doi:https://doi.org/10.1029/2021JD034925.

Gonzalez, A. O., and X. Jiang, 2019: Distinct propagation characteristics of intraseasonal variability over the tropical west pacific. *Journal of Geophysical Research: Atmospheres*, **124 (10)**, 5332–5351, doi:https://doi.org/10.1029/2018JD029884.

Gottschalck, J., and coautors, 2010: A framework for assessing operational madden–julian oscillation forecasts: A clivar mjo working group project. *Bulletin of the American Meteorological Society*, **91 (9)**, 1247 – 1258, doi:10.1175/2010BAMS2816.1, URL https://journals.ametsoc.org/view/journals/bams/91/9/2010bams2816_1.xml.

Grabowski, W., and M. Moncrieff, 2001: Large-scale organization of tropical convection in two-dimensional explicit numerical simulations. *Quarterly Journal of the Royal Meteorological Society*, **127 (572)**, 445–468, doi:https://doi.org/10.1002/qj.49712757211.

Grabowski, W., and M. Moncrieff, 2004a: Moisture–convection feedback in the tropics. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3081–3104, doi:https://doi.org/10.1256/qj.03.135.

Grabowski, W. W., and M. W. Moncrieff, 2004b: Moisture–convection feedback in the tropics. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3081–3104, doi:https://doi.org/10.1256/qj.03.135.

Grant, L. D., and S. C. van den Heever, 2016: Cold pool dissipation. *Journal of Geophysical Research: Atmospheres*, **121 (3)**, 1138–1155, doi:https://doi.org/10.1002/2015JD023813.

Grant, L. D., and S. C. van den Heever, 2018: Cold pool-land surface interactions in a dry continental environment. *Journal of Advances in Modeling Earth Systems*, **10 (7)**, 1513–1526, doi:https://doi.org/10.1029/2018MS001323.

Grundner, A., T. Beucler, P. Gentine, F. Iglesias-Suarez, M. A. Giorgetta, and V. Eyring, 2022: Deep learning based cloud cover parameterization for icon. *Journal of Advances in Modeling Earth Systems*, **14 (12)**, e2021MS002 959, doi:https://doi.org/10.1029/2021MS002959.

Haertel, P. T., and R. H. Johnson, 1998: Two-day disturbances in the equatorial western pacific. *Quarterly Journal of the Royal Meteorological Society*, **124 (546)**, 615–636, doi:https://doi.org/10.1002/qj.49712454611.

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

Han, M., S. A. Braun, T. Matsui, and C. R. Williams, 2013: Evaluation of cloud microphysics schemes in simulations of a winter storm using radar and radiometer measurements. *Journal of Geophysical Research: Atmospheres*, **118 (3)**, 1401–1419, doi:https://doi.org/10.1002/jgrd.50115.

Hartmann, D., 2016: Chapter 3 - Atmospheric Radiative Transfer and Climate. *Global Physical Climatology (Second Edition)*, D. L. Hartmann, Ed., second edition ed., Elsevier, Boston, 49–94, doi:https://doi.org/10.1016/B978-0-12-328531-7.00003-7.

Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning*. Springer.

Held, I., R. Hemler, and V. Ramaswamy, 1993: Radiative-convective equilibrium with explicit two-dimensional moist convection. *Journal of Atmospheric Sciences*, **50 (23)**, 3909–3927, doi:10.1175/1520-0469(1993)050<3909:RCEWET>2.0.CO;2.

Helfer, K., and L. Nuijens, 2021: The morphology of simulated trade-wind convection and cold pools under wind shear. *Journal of Geophysical Research: Atmospheres*, **126 (20)**, e2021JD035 148, doi:https://doi.org/10.1029/2021JD035148.

Helfer, K., L. Nuijens, S. de Roode, and A. Siebesma, 2020: How wind shear affects trade-wind cumulus convection. *Journal of Advances in Modeling Earth Systems*, **12 (12)**, e2020MS002 183, doi:https://doi.org/10.1029/2020MS002183.

Hernandez-Deckers, D., T. Matsui, and A. M. Fridlind, 2022: Updraft dynamics and microphysics: on the added value of the cumulus thermal reference frame in simulations of aerosol–deep convection interactions. *Atmospheric Chemistry and Physics*, **22 (2)**, 711–24, doi:10.5194/acp-22-711-2022.

Hersbach, H., and coautors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi:https://doi.org/10.1002/qj.3803.

Hohenegger, C., and C. Jakob, 2020: A relationship between itcz organization and subtropical humidity. *Geophysical Research Letters*, **47 (16)**, e2020GL088 515, doi:https://doi.org/10.1029/2020GL088515.

Hohenegger, C., and B. Stevens, 2016a: Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *Journal of Advances in Modeling Earth Systems*, **8 (3)**, 1468–1482, doi:https://doi.org/10.1002/2016MS000666.

Hohenegger, C., and B. Stevens, 2016b: Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *Journal of Advances in Modeling Earth Systems*, **8 (3)**, 1468–1482, doi:https://doi.org/10.1002/2016MS000666.

Hohenegger, C., and B. Stevens, 2018a: The role of the permanent wilting point in controlling the spatial distribution of precipitation. *Proceedings of the National Academy of Sciences*, **115 (22)**, 5692–5697, doi:10.1073/pnas.1718842115.

Hohenegger, C., and B. Stevens, 2018b: The role of the permanent wilting point in controlling the spatial distribution of precipitation. *Proceedings of the National Academy of Sciences*, **115 (22)**, 5692–5697, doi:10.1073/pnas.1718842115.

Holloway, C., A. Wing, S. Bony, C. Muller, H. Masunaga, T. L'Ecuyer, D. Turner, and P. Zuidema, 2017: Observing Convective Aggregation. *Surveys in Geophysics*, **38**, 1199–1236, doi:10.1007/s10712-017-9419-1.

Holloway, C., and S. Woolnough, 2016: The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations. *Journal of Advances in Modeling Earth Systems*, **8 (1)**, 166–195, doi:https://doi.org/10.1002/2015MS000511.

Holloway, C. E., 2017: Convective aggregation in realistic convective-scale simulations. *Journal of Advances in Modeling Earth Systems*, **9 (2)**, 1450–1472, doi:https://doi.org/10.1002/2017MS000980.

Holloway, C. E., S. J. Woolnough, and G. M. S. Lister, 2012: Precipitation distributions for explicit versus parametrized convection in a large-domain high-resolution tropical case study. *Quarterly Journal of the Royal Meteorological Society*, **138 (668)**, 1692–1708, doi:10.1002/qj.1903.

Holton, J., 2004: *An Introduction to Dynamic Meteorology*, Vol. 4. Elsevier Academic Press, USA, 535 pp.

Hong, S., and J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *Asia-pacific Journal of Atmospheric Sciences*, **42**, 129–151.

Hong, S., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, **134 (9)**, 2318 – 2341, doi:10.1175/MWR3199.1.

Hong, S., and H. Pan, 1996: Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model. *Monthly Weather Review*, **124 (10)**, 2322 – 2339, doi:10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Monthly Weather Review*, **132 (1)**, 103 – 120, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

Hou, A. Y., G. Skofronick-Jackson, C. D. Kummerow, and J. M. Shepherd, 2008: *Global precipitation measurement*, 131–169. Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-540-77655-0_6.

Hou, A. Y., G. Skofronick-Jackson, W. A. Petersen, E. F. Stocher, J. Kaye, and R. Kakar, 2013: NASA GPM Science Implementation Plan, NASA Goddard Space Flight Center. NASA, URL https://gpm.nasa.gov/resources/documents/gpm-science-implementation-plan.

Hou, A. Y., and coautors, 2014: The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, **95 (5)**, 701–722, doi:10.1175/BAMS-D-13-00164.1.

Hu, X., J. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. *Journal of Applied Meteorology and Climatology*, **49 (9)**, 1831 – 1844, doi:10.1175/2010JAMC2432.1.

Huang, B., C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H.-M. Zhang, 2021: Improvements of the daily optimum interpolation sea surface temperature (doisst) version 2.1. *Journal of Climate*, **34 (8)**, 2923 – 2939, doi:https://doi.org/10.1175/JCLI-D-20-0166.1, URL https://journals.ametsoc.org/view/journals/clim/34/8/JCLI-D-20-0166.1.xml.

Huang, J.-D., and C.-M. Wu, 2022: A Framework to Evaluate Convective Aggregation: Examples With Different Microphysics Schemes. *Journal of Geophysical Research: Atmospheres*, **127 (5)**, e2021JD035 886, doi:https://doi.org/10.1029/2021JD035886.

Iacono, M., J. Delamere, E. Mlawer, M. Shephard, S. Clough, and D. W. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, **113 (D13)**, doi:https://doi.org/10.1029/2008JD009944.

Ishida, H., and T. Y. Nakajima, 2009: Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. *Journal of Geophysical Research: Atmospheres*, **114 (D7)**, doi:https://doi.org/10.1029/2008JD010710.

Ishida, H., T. Nakjima, T. Yokota, N. Kikuchi, and H. Watanabe, 2011: Investigation of GOSAT TANSO-CAI Cloud Screening Ability through an Intersatellite Comparison. *Journal of Applied Meteorology and Climatology*, **50 (7)**, 1571 – 1586, doi: 10.1175/2011JAMC2672.1.

Iverson, E., G. Thompson, and B. Nygaard, 2021: Improvements to melting snow behavior in a bulk microphysics scheme. *Atmospheric Research*, **253**, 105 471, doi: 10.1016/j.atmosres.2021.1054711.

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

Jordan, C., 1958: MEAN SOUNDINGS FOR THE WEST INDIES AREA. *Journal of Atmospheric Sciences*, **15 (1)**, 91 – 97, doi:10.1175/1520-0469(1958)015<0091:MSFTWI>2.0.CO;2.

Kanemaru, K., and H. Masunaga, 2013: A Satellite Study of the Relationship between Sea Surface Temperature and Column Water Vapor over Tropical and Subtropical Oceans. *Journal of Climate*, **26 (12)**, 4204–4218, doi:10.1175/JCLI-D-12-00307.1.

Kara, A. B., P. A. Rochford, and H. E. Hurlburt, 2003: Mixed layer depth variability over the global ocean. *Journal of Geophysical Research: Oceans*, **108 (C3)**, doi:https://doi.org/10.1029/2000JC000736.

Keil, C., A. Rpnack, G. Craig, and U. Schumann, 2008: Sensitivity of quantitative precipitation forecast to height dependent changes in humidity. *Geophys. Res. Lett.*, **35**.

Khairoutdinov, M., and K. Emanuel, 2013: Rotating radiative-convective equilibrium simulated by a cloud-resolving model. *Journal of Advances in Modeling Earth Systems*, **5 (4)**, 816–825, doi:https://doi.org/10.1002/2013MS000253.

Kiladis, G. N., C. D. Thorncroft, and N. M. J. Hall, 2006: Three-dimensional structure and dynamics of african easterly waves. part i: Observations. *Journal of the Atmospheric Sciences*, **63 (9)**, 2212 – 2230, doi:10.1175/JAS3741.1, URL https://journals.ametsoc.org/view/journals/atsc/63/9/jas3741.1.xml.

Kiladis, G. N., and M. Wheeler, 1995: Horizontal and vertical structure of observed tropospheric equatorial rossby waves. *Journal of Geophysical Research: Atmospheres*, **100 (D11)**, 22 981–22 997, doi:https://doi.org/10.1029/95JD02415.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Reviews of Geophysics*, **47 (2)**, doi:10.1029/2008RG000266.

Kleine-Deters, J., R. Zalakeviciute, M. Gonzalez, and Y. Rybarczyk, 2017: Modeling pm2.5 urban pollution using machine learning and selected meteorological parameters. *Journal of Electrical and Computer Engineering*, **2017**, 5106 045, doi: 10.1155/2017/5106045.

Klemp, J., J. Dudhia, and A. Hassiotis, 2008: An Upper Gravity-Wave Absorbing Layer for NWP Applications. *Monthly Weather Review*, **136 (10)**, 3987 – 4004, doi: 10.1175/2008MWR2596.1.

Kline, R., 2015: *Principles and practice of structural equation modeling*, Vol. 1. Guilford publications, Inc, 370 Seventh Avenue, Suite 1200, New York, NY 10001, 493 pp.

Knapp, K., and coautors, 2011: Globally Gridded Satellite Observations for Climate Studies. *Bulletin of the American Meteorological Society*, **92 (7)**, 893–907, doi:https://doi.org/10.1175/2011BAMS3039.1.

Kurihara, Y., H. Murakami, and M. Kachi, 2016: Sea surface temperature from the new Japanese geostationary meteorological Himawari-8 satellite. *Geophysical Research Letters*, **43 (3)**, 1234–1240, doi:https://doi.org/10.1002/2015GL067159.

Kurihara, Y., H. Murakami, K. Ogata, and M. Kachi, 2021: A quasi-physical sea surface temperature method for the split-window data from the second-generation global imager (sgli) onboard the global change observation mission-climate (gcom-c) satellite. *Remote Sensing of Environment*, **257**, 112 347, doi:https://doi.org/10.1016/j.rse.2021.112347.

Ladwig, W., 2017: wrf-python (Version 1.3.4.1). UCAR/NCAR, Boulder, Colorado, URL https://doi.org/10.5065/D6W094P1, Software.

Lang, S., W. Tao, J. Chern, D. Wu, and X. Li, 2014: Benefits of a Fourth Ice Class in the Simulated Radar Reflectivities of Convective Systems Using a Bulk Microphysics Scheme. *Journal of the Atmospheric Sciences*, **71 (10)**, 3583 – 3612, doi:10.1175/JAS-D-13-0330.1.

Li, Z., P. Zuidema, P. Zhu, and H. Morrison, 2015: The Sensitivity of Simulated Shallow Cumulus Convection and Cold Pools to Microphysics. *Journal of the Atmospheric Sciences*, **72 (9)**, 3340 – 3355, doi:https://doi.org/10.1175/JAS-D-14-0099.1.

Liaw, A., and M. Wiener, 2002: Classification and regression by randomforest. *R news*, **2 (3)**, 18–22.

Lin, Y., R. Farley, and H. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *Journal of Applied Meteorology and Climatology*, **22 (6)**, 1065 – 1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Manabe, S., and R. Strickler, 1964: Thermal Equilibrium of the Atmosphere with a Convective Adjustment. *Journal of Atmospheric Sciences*, **21 (4)**, 361–385, doi:10.1175/1520-0469(1964)021<0361:TEOTAW>2.0.CO;2.

Mapes, B., and R. Neale, 2011: Parameterizing convective organization to escape the entrainment dilemma. *Advances in Modeling Earth Systems 3*, **10**, 1942–2466.

Mapes, B. E., E. S. Chung, W. M. Hannah, H. Masunaga, A. J. Wimmers, and C. S. Velden, 2018: The Meandering Margin of the Meteorological Moist Tropics. *Geophysical Research Letters*, **45 (2)**, 1177–1184, doi:10.1002/2017GL076440.

Mauritsen, T., and B. Stevens, 2015: Missing iris effect as a possible cause of muted hydrological change and high climate sensitivity in models. *Nature Geoscience*, **8**, 346–351, doi:https://doi.org/10.1038/ngeo2414.

Mayta, V. C., Á. F. Adames, and F. Ahmed, 2022: Westward-Propagating Moisture Mode Over the Tropical Western Hemisphere. *Geophysical Research Letters*, **49 (6)**, e2022GL097 799, doi:10.1029/2022GL097799.

McGovern, A., R. Lagerquist, D. J. Gagne, G. E. Jergensen, K. L. Elmore, C. R. Homeyer, and T. Smith, 2019: Making the black box more transparent: Understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, **100 (11)**, 2175 – 2199, doi:https://doi.org/10.1175/BAMS-D-18-0195.1.

Met-Office, 2015: Cartopy: a Cartographic Python Library with Matplotlib Interface, Met Office Exeter, UK. MetOffice, URL http://scitools.org.uk/cartopy/docs/latest.

Mlawer, E., S. Taubman, P. Brown, M. Iacono, and S. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, **102 (D14)**, 16 663–16 682, doi:https://doi.org/10.1029/97JD00237.

Monin, A., and A. Obukhov, 1954: Basic Laws of Turbulent Mixing in the Surface Layer of the Atmosphere. *Tr. Akad. Nauk SSSR Geophiz. Inst.*, **24 (151)**, 163 – 187.

Morrison, H., J. Curry, and V. Khvorostyanov, 2005: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description. *Journal of the Atmospheric Sciences*, **62 (6)**, 1665 – 1677, doi:10.1175/JAS3446.1.

Morrison, H., and A. Gettelman, 2008: A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests. *Journal of Climate*, **21 (15)**, 3642 – 3659, doi:10.1175/2008JCLI2105.1.

Morrison, H., and J. Pinto, 2006: Intercomparison of Bulk Cloud Microphysics Schemes in Mesoscale Simulations of Springtime Arctic Mixed-Phase Stratiform Clouds. *Monthly Weather Review*, **134 (7)**, 1880 – 1900, doi:10.1175/MWR3154.1.

Muller, C., 2013: Impact of convective organization on the response of tropical precipitation extremes to warming. *Journal of Climate*, **26 (14)**, 5028 – 5043, doi:https://doi.org/10.1175/JCLI-D-12-00655.1.

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

Muller, C., and I. Held, 2012: Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences*, **69 (8)**, 2551 – 2565, doi:10.1175/JAS-D-11-0257.1.

Muller, C., and D. M. Romps, 2018: Acceleration of tropical cyclogenesis by self-aggregation feedbacks. *Proceedings of the National Academy of Sciences*, **115 (12)**, 2930–2935, doi:10.1073/pnas.1719967115.

Muller, C., and coautors, 2022a: Spontaneous aggregation of convective storms. *Annual Review of Fluid Mechanics*, **54 (1)**, 133–157, doi:10.1146/annurev-fluid-022421-011319.

Muller, S., and coautors, 2022b: Evaluation of alpine-mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms. *Climate Dynamics*, doi:10.1007/s00382-022-06555-z.

Müller, S., and C. Hohenegger, 2020a: Self-Aggregation of Convection in Spatially Varying Sea Surface Temperatures. *Journal of Advances in Modeling Earth Systems*, **12 (1)**, e2019MS001 698, doi:https://doi.org/10.1029/2019MS001698.

Müller, S. K., and C. Hohenegger, 2020b: Self-aggregation of convection in spatially varying sea surface temperatures. *Journal of Advances in Modeling Earth Systems*, **12 (1)**, e2019MS001 698, doi:https://doi.org/10.1029/2019MS001698.

Nakajima, K., and T. Matsuno, 1988: Numerical experiments concerning the origin of cloud clusters in the tropical atmosphere. *Journal of Meteorological Society of Japan*, **66**, 309–329.

Naumann, A., B. Stevens, and C. Hohenegger, 2019: A Moist Conceptual Model for the Boundary Layer Structure and Radiatively Driven Shallow Circulations in the Trades. *Journal of the Atmospheric Sciences*, **76 (5)**, 1289–1306, doi:10.1175/JAS-D-18-0226.1.

Nesbitt, S., E. Zipser, and D. Cecil, 2000: A Census of Precipitation Features in the Tropics Using TRMM: Radar, Ice Scattering, and Lightning Observations. *Journal of Climate*, **13 (23)**, 4087–4106, doi:10.1175/1520-0442(2000)013<4087:ACOPFI>2.0.CO;2.

NOAA, 2015: Ncep gfs 0.25 degree global forecast grids historical archive. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder CO, URL https://doi.org/10.5065/D65D8PWK.

Noh, Y., W. Cheon, S. Hong, and S. Raasch, 2003: Improvement of the K-profile Model for the Planetary Boundary Layer based on Large Eddy Simulation Data. *Boundary-Layer Meteorology*, **107**, 401 – 427, doi:10.1023/A:1022146015946.

Park, S., W. Skamarock, J. Klemp, L. Fowler, and M. Duda, 2013: Evaluation of global atmospheric solvers using extensions of the jablonowski and williamson baroclinic wave test case. *Monthly Weather Review"*, **141 (9)**, 3116 – 3129, doi:10.1175/MWR-D-12-00096.1.

Paulson, C., 1970: The Mathematical Representation of Wind Speed and Temperature Profiles in the Unstable Atmospheric Surface Layer. *Journal of Applied Meteorology and Climatology*, **9 (6)**, 857 – 861, doi:10.1175/1520-0450(1970)009<0857:TMROWS>2.0.CO;2.

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

Pierrehumbert, R. T., 1995: Thermostats, Radiator Fins, and the Local Runaway Greenhouse. *Journal of the Atmospheric Sciences*, **52 (10)**, 1784–1806, doi:10.1175/1520-0469(1995)052<1784:TRFATL>2.0.CO;2.

Pierrehumbert, R. T., 1998: Lateral mixing as a source of subtropical water vapor. *Geophysical Research Letters*, **25 (2)**, 151–154, doi:10.1029/97GL03563.

Pollard, R., P. Rhines, and R. Thompson, 1973: The deepening of the wind-mixed layer. *Geophysical Fluid Dynamics*, **4 (4)**, 381–404, doi:10.1080/03091927208236105.

Popp, M., and S. Bony, 2019: Stronger zonal convective clustering associated with a wider tropical rain belt. *Nature Communications*, **10**, 4261, doi:https://doi.org/10.1038/s41467-019-12167-9.

Potvin, C., and coautors, 2020: Assessing systematic impacts of pbl schemes on storm evolution in the noaa warn-on-forecast system. *Monthly Weather Review*, **148 (6)**, 2567 – 2590, doi:10.1175/MWR-D-19-0389.1.

Powers, J. G., and coautors, 2017: The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions. *Bulletin of the American Meteorological Society*, **98 (8)**, 1717 – 1737, doi:10.1175/BAMS-D-15-00308.1.

Pscheidt, I., F. Senf, R. Heinze, H. Deneke, S. Trömel, and C. Hohenegger, 2019: How organized is deep convection over Germany? *Quarterly Journal of the Royal Meteorological Society*, **145 (723)**, 2366–2384, doi:https://doi.org/10.1002/qj.3552.

Radtke, J., A. K. Naumann, M. Hagen, and F. Ament, 2022: The relationship between precipitation and its spatial pattern in the trades observed during EUREC4A. *Quarterly Journal of the Royal Meteorological Society*, **148 (745)**, 1913–1928, doi:10.1002/qj.4284.

Raymond, D. J., S. L. Sessions, A. H. Sobel, and Z. Fuchs, 2009: The Mechanics of Gross Moist Stability. *Journal of Advances in Modeling Earth Systems*, **1 (3)**, doi:https://doi.org/10.3894/JAMES.2009.1.9.

Reed, K., B. Medeiros, J. Bacmeister, and P. Lauritzen, 2015: Global Radiative–Convective Equilibrium in the Community Atmosphere Model, Version 5. *Journal of the Atmospheric Sciences*, **72 (5)**, 2183 – 2197, doi:10.1175/JAS-D-14-0268.1.

Rennó, N., K. Emanuel, and P. Stone, 1994: Radiative-convective model with an explicit hydrologic cycle: 1. formulation and sensitivity to model parameters. *Journal of Geophysical Research: Atmospheres*, **99 (D7)**, 14 429–14 441, doi:https://doi.org/10.1029/94JD00020.

Retsch, M. H., C. Jakob, and M. Singh, 2020: Assessing convective organization in tropical radar observations. *Journal of Geophysical Research: Atmospheres*, **125 (7)**, e2019JD031 801, doi:https://doi.org/10.1029/2019JD031801.

Retsch, M. H., C. Jakob, and M. S. Singh, 2022: Identifying Relations Between Deep Convection and the Large-Scale Atmosphere Using Explainable Artificial Intelligence. *Journal of Geophysical Research: Atmospheres*, **127 (3)**, e2021JD035 388, doi:10.1029/2021JD035388.

Rieger, N., and S. J. Levang, 2023: xeofs v. 2.2.4: Comprehensive eof analysis in python with xarray: A versatile, multidimensional, and scalable tool for advanced climate data analysis. CRM Centre for Mathematical Research, Universitat Autònoma de Barcelona, URL https://github.com/xarray-contrib/xeofs, doi:10.5281/zenodo.6323011.

Rose, B., 2020: Insolation, the climate laboratory. University at Albany, accessed: 2023-09-30, https://brian-rose.github.io/ClimateLaboratoryBook/courseware/insolation.html.

Ross, A. N., A. M. Tompkins, and D. J. Parker, 2004: Simple models of the role of surface fluxes in convective cold pool evolution. *Journal of the Atmospheric Sciences*, **61 (13)**, 1582 – 1595, doi:https://doi.org/10.1175/1520-0469(2004)061<1582:

SMOTRO>2.0.CO;2, URL https://journals.ametsoc.org/view/journals/atsc/61/13/1520-0469_2004_061_1582_smotro_2.0.co_2.xml.

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A Theory for Strong, Long-Lived Squall Lines. *Journal of Atmospheric Sciences*, **45 (3)**, 463 – 485, doi:https://doi.org/10.1175/1520-0469(1988)045<0463:ATFSLL>2.0.CO;2.

Rushley, S. S., D. Kim, C. S. Bretherton, and M.-S. Ahn, 2018: Reexamining the Nonlinear Moisture-Precipitation Relationship Over the Tropical Oceans. *Geophysical Research Letters*, **45 (2)**, 1133–1140, doi:10.1002/2017GL076296.

Rutledge, S., and P. Hobbs, 1983: The Mesoscale and Microscale Structure and Organization of Clouds and Precipitation in Midlatitude Cyclones. VIII: A Model for the "Seeder-Feeder" Process in Warm-Frontal Rainbands. *Journal of Atmospheric Sciences*, **40 (5)**, 1185 – 1206, doi:10.1175/1520-0469(1983)040<1185:TMAMSA>2.0.CO;2.

Rutledge, S., and P. Hobbs, 1984: The Mesoscale and Microscale Structure and Organization of Clouds and Precipitation in Midlatitude Cyclones. XII: A Diagnostic Modeling Study of Precipitation Development in Narrow Cold-Frontal Rainbands. *Journal of Atmospheric Sciences*, **41 (20)**, 2949 – 2972, doi:10.1175/1520-0469(1984)041<2949:TMAMSA>2.0.CO;2.

Schlemmer, L., and C. Hohenegger, 2014: The Formation of Wider and Deeper Clouds as a Result of Cold-Pool Dynamics. *Journal of the Atmospheric Sciences*, **71 (8)**, 2842–2858, doi:10.1175/JAS-D-13-0170.1.

Segura, R., A. Badia, S. Ventura, J. Gilabert, A. Martilli, and GaraVillalba, 2021: Sensitivity study of PBL schemes and soil initialization using the WRF-BEP-BEM model over a Mediterranean coastal city. *Urban Climate*, **39**, 100 982, doi:10.1016/j.uclim.2021.100982.

Semie, A. G., and S. Bony, 2020: Relationship Between Precipitation Extremes and Convective Organization Inferred From Satellite Observations. *Geophysical Research Letters*, **47 (9)**, e2019GL086 927, doi:https://doi.org/10.1029/2019GL086927.

Shamekh, S., K. D. Lamb, Y. Huang, and P. Gentine, 2023: Implicit learning of convective organization explains precipitation stochasticity. *Proceedings of the National Academy of Sciences*, **120 (20)**, e2216158 120, doi:10.1073/pnas.2216158120.

Shamekh, S., C. Muller, J. Duvel, and F. D'Andrea, 2020a: Self-Aggregation of Convective Clouds With Interactive Sea Surface Temperature. *Journal of Advances in Modeling Earth Systems*, **12 (11)**, e2020MS002 164, doi:https://doi.org/10.1029/2020MS002164.

Shamekh, S., C. Muller, J.-P. Duvel, and F. D'Andrea, 2020b: How Do Ocean Warm Anomalies Favor the Aggregation of Deep Convective Clouds? *Journal of the Atmospheric Sciences*, **77 (11)**, 3733 – 3745, doi:https://doi.org/10.1175/JAS-D-18-0369.1.

Shamekh, S., C. Muller, J.-P. Duvel, and F. D'Andrea, 2020c: How do ocean warm anomalies favor the aggregation of deep convective clouds? *Journal of the Atmospheric Sciences*, **77 (11)**, 3733 – 3745, doi:https://doi.org/10.1175/JAS-D-18-0369.1.

Sherwood, Roca, Werckwerth, and Andronova, 2009: Tropospheric water vapor, convection and climate: A critical review. *Geophysica*, **48**, RG2001, doi:10.1029/2009RG000301.

Sherwood, S. C., 1996: Maintenance of the free-tropospheric tropical water vapor distribution .1. Clear regime budget. *J. Climate*, **9**, 2903–2918.

Sherwood, S. C., R. Roca, T. M. Weckwerth, and N. G. Andronova, 2010: Tropospheric water vapor, convection, and climate. *Rev. Geophys.*, **48 (2)**, RG2001, doi:10.1029/2009RG000301.

Shi, X., and Y. Fan, 2021: Modulation of the Bifurcation in Radiative-Convective Equilibrium by Gray-Zone Cloud and Turbulence Parameterizations. *Journal of Advances in Modeling Earth Systems*, **13 (10)**, e2021MS002 632, doi:https://doi.org/10.1029/2021MS002632.

Skamarock, W., and coautors, 2019: A Description of the Advanced Research WRF Version 4. Tech. rep., NCAR Technical Note NCAR/TN-556+STR. doi:10.5065/1dfh-6p97.

Skofronick-Jackson, G., D. Kirschbaum, W. Petersen, G. Huffman, C. Kidd, E. Stocker, and R. Kakar, 2018: The global precipitation measurement (gpm) mission's scientific achievements and societal contributions: reviewing four years of advanced rain and snow observations. *Quarterly Journal of the Royal Meteorological Society*, **144 (S1)**, 27–48, doi:https://doi.org/10.1002/qj.3313.

Smagorinsky, J., 1963: GENERAL CIRCULATION EXPERIMENTS WITH THE PRIMITIVE EQUATIONS: I. THE BASIC EXPERIMENT. *Monthly Weather Review*, **91 (3)**, 99 – 164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Sobel, A., J. Nilsson, and L. Polvani, 2001: The Weak Temperature Gradient Approximation and Balanced Tropical Moisture Waves. *Journal of the Atmospheric Sciences*, **58 (23)**, 3650 – 3665, doi:10.1175/1520-0469(2001)058<3650:TWTGAA>2.0.CO;2.

Sokhi, R. S., and coautors, 2021: A global observational analysis to understand changes in air quality during exceptionally low anthropogenic emission conditions. *Environment International*, **157**, 106 818, doi:https://doi.org/10.1016/j.envint.2021.106818.

Stephens, G., S. van den Heever, and L. Pakula, 2008: Radiative–Convective Feedbacks in Idealized States of Radiative–Convective Equilibrium. *Journal of the Atmospheric Sciences*, **65 (12)**, 3899–3916, doi:10.1175/2008JAS2524.1.

Stevens, B., 2005: Atmospheric moist convection. *Annu. Rev. Earth Planet. Sci.*, **33**, 605–643.

Sun, D.-Z., and R. S. Lindzen, 1993: Distribution of tropical tropospheric water vapor. *J. Atmos. Sci.*, **50**, 1643–1660.

Takayabu, Y. N., 1994: Large-scale cloud disturbances associated with equatorial waves. *Journal of the Meteorological Society of Japan. Ser. II*, **72 (3)**, 433–449, doi:10.2151/jmsj1965.72.3_433.

Takayabu, Y. N., and T. Nitta, 1993: 3-5 day-period disturbances coupled with convection over the tropical pacific ocean. *Journal of the Meteorological Society of Japan. Ser. II*, **71 (2)**, 221–246, doi:10.2151/jmsj1965.71.2_221.

Tao, W., and J. Simpson, 1993: Goddard Cumulus Ensemble Model. Part I: Model Description. *Terrestrial, Atmospheric and Oceanic sciences journal*, **4**, 35–75, doi:10.1007/s00703-001-0594-7.

Tao, W., and coautors, 2014: The Goddard Cumulus Ensemble model (GCE): Improvements and applications for studying precipitation processes. *Atmospheric Research*, **143**, 392–424, doi:https://doi.org/10.1016/j.atmosres.2014.03.005.

Therry, G., and P. Lacarrère, 1989: Improving the Eddy Kinetic Energy model for planetary boundary layer description. *Monthly Weather Review*, **25**, 63–88, doi:10.1007/BF00122098.

Thompson, G., and T. Eidhammer, 2014: A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. *Journal of the Atmospheric Sciences*, **71 (10)**, 3636 – 3658, doi:10.1175/JAS-D-13-0305.1.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Wea. Rev.*, **136 (12)**, 5095–5115, doi:10.1175/2008MWR2387.1.

Thorpe, A. J., M. J. Miller, and M. W. Moncrieff, 1982: Two-dimensional convection in non-constant shear: A model of mid-latitude squall lines. *Quarterly Journal of the Royal Meteorological Society*, **108 (458)**, 739–762, doi:https://doi.org/10.1002/qj.49710845802.

Tobin, I., S. Bony, C. Holloway, J. Grandpeix, G. Sèze, D. Coppin, S. Woolnough, and R. Roca, 2013: Does convective aggregation need to be represented in cumulus parameterizations? *Journal of Advances in Modeling Earth Systems*, **5 (4)**, 692–703, doi:https://doi.org/10.1002/jame.20047.

Tobin, I., S. Bony, and R. Roca, 2012: Observational Evidence for Relationships between the Degree of Aggregation of Deep Convection, Water Vapor, Surface Fluxes, and Radiation. *Journal of Climate*, **25 (20)**, 6885–6904, doi:10.1175/JCLI-D-11-00258.1.

Tompkins, A., 2001a: On the Relationship between Tropical Convection and Sea Surface Temperature. *Journal of Climate*, **14 (5)**, 633 – 637, doi:https://doi.org/10.1175/1520-0442(2001)014<0633:OTRBTC>2.0.CO;2.

Tompkins, A., 2001b: Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Cold Pools. *Journal of Atmospheric Sciences*, **58 (13)**, 1650–1672, doi:10.1175/1520-0469(2001)058<1650:OOTCIL>2.0.CO;2.

Tompkins, A., 2001c: Organization of Tropical Convection in Low Vertical Wind Shears: The Role of Water Vapor. *Journal of the Atmospheric Sciences*, **58 (6)**, 529 – 545, doi:10.1175/1520-0469(2001)058<0529:OOTCIL>2.0.CO;2.

Tompkins, A., and G. Craig, 1998a: Radiative–convective equilibrium in a three-dimensional cloud-ensemble model. *Quarterly Journal of the Royal Meteorological Society*, **124 (550)**, 2073–2097, doi:https://doi.org/10.1002/qj.49712455013.

Tompkins, A., and G. Craig, 1998b: Time-scales of adjustment to radiative-convective equilibrium in the tropical atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **124 (552)**, 2693–2713, doi:https://doi.org/10.1002/qj.49712455208.

Tompkins, A., and A. Semie, 2017: Organization of tropical convection in low vertical wind shears: Role of updraft entrainment. *Journal of Advances in Modeling Earth Systems*, **9 (2)**, 1046–1068, doi:https://doi.org/10.1002/2016MS000802.

Tompkins, A., and A. Semie, 2021: Impact of a mixed ocean layer and the diurnal cycle on convective aggregation. *Journal of Advances in Modeling Earth Systems*, doi:https://doi.org/10.1029/2020MS002186.

Virtanen, P., and coautors, 2020: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, **17**, 261–272, doi:10.1038/s41592-019-0686-2.

Vogel, R., A. Albright, J. Vial, G. George, B. Stevens, and B. Sandrine, 2022: Strong cloud–circulation coupling explains weak trade cumulus feedback. *Nature*, **612**, 696–700, doi:https://doi.org/10.1038/s41586-022-05364-y.

Webb, E., 1970: Profile relationships: The log-linear range, and extension to strong stability. *Quarterly Journal of the Royal Meteorological Society*, **96 (407)**, 67–90, doi:https://doi.org/10.1002/qj.49709640708.

Weger, R. C., J. Lee, T. Zhu, and R. M. Welch, 1992: Clustering, randomness and regularity in cloud fields: 1. Theoretical considerations. *J. Geophys. Res.*, **97**, 20 519–20 536.

Wei, J., W. Huang, Z. Li, W. Xue, Y. Peng, L. Sun, and M. Cribb, 2019: Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach. *Remote Sensing of Environment*, **231**, 111 221, doi:https://doi.org/10.1016/j.rse.2019.111221.

Weisman, M. L., and R. Rotunno, 2004: "a theory for strong long-lived squall lines" revisited. *Journal of the Atmospheric Sciences*, **61 (4)**, 361 – 382, doi:https://doi.org/10.1175/1520-0469(2004)061<0361:ATFSLS>2.0.CO;2, URL https://journals.ametsoc.org/view/journals/atsc/61/4/1520-0469_2004_061_0361_atfsls_2.0.co_2.xml.

Weverberg, K. V., and coautors, 2013: The Role of Cloud Microphysics Parameterization in the Simulation of Mesoscale Convective System Clouds and Precipitation in the Tropical Western Pacific. *Journal of the Atmospheric Sciences*, **70 (4)**, 1104 – 1128, doi:https://doi.org/10.1175/JAS-D-12-0104.1.

Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *Journal of the Atmospheric Sciences*, **56 (3)**, 374 – 399, doi:https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2.

White, B., A. Buchanan, C. Birch, P. Stier, and K. Pearson, 2018: 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 the Atmospheric Sciences*, **75 (2)**, 425 – 450, doi:https://doi.org/10.1175/JAS-D-16-0307.1.

Wicker, L., and W. Skamarock, 2002: Time-Splitting Methods for Elastic Models Using Forward Time Schemes. *Monthly Weather Review*, **130 (8)**, 2088 – 2097, doi:10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2.

Wimmers, A. J., and C. S. Velden, 2011: Seamless advective blending of total precipitable water retrievals from polar-orbiting satellites. *Journal of Applied Meteorology and Climatology*, **50 (5)**, 1024 – 1036, doi:10.1175/2010JAMC2589.1.

Windmiller, J., and G. Craig, 2019: Universality in the Spatial Evolution of Self-Aggregation of Tropical Convection. *Journal of the Atmospheric Sciences*, **76 (6)**, 1677–1696, doi:10.1175/JAS-D-18-0129.1.

Wing, A., 2019: Self-aggregation of deep convection and its implications for climate. *Current Climate Change Reports*, **5**, 793–813, doi:10.1007/s40641-019-00120-3.

Wing, A., and T. Cronin, 2016: Self-aggregation of convection in long channel geometry. *Quarterly Journal of the Royal Meteorological Society*, **142 (694)**, 1–15, doi:https://doi.org/10.1002/qj.2628.

Wing, A., and K. Emanuel, 2014: Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations. *Journal of Advances in Modeling Earth Systems*, **6 (1)**, 59–74, doi:https://doi.org/10.1002/2013MS000269.

Wing, A., K. Emanuel, C. Holloway, and C. Muller, 2017: Convective Self-Aggregation in Numerical Simulations: A Review. *Surveys in Geophysics*, **38**, 1173–1197, doi:10.1007/s10712-017-9408-4.

Wing, A. A., K. A. Reed, M. Satoh, B. Stevens, S. Bony, and T. Ohno, 2018: Radiative–convective equilibrium model intercomparison project. *Geoscientific Model Development*, **11 (2)**, 793–813, doi:10.5194/gmd-11-793-2018.

Wing, A. A., and coautors, 2020: Clouds and Convective Self-Aggregation in a Multimodel Ensemble of Radiative-Convective Equilibrium Simulations. *Journal of Advances in Modeling Earth Systems*, **12 (9)**, e2020MS002 138, doi:https://doi.org/10.1029/2020MS002138.

Woolnough, S., J. Slingo, and B. Hoskins, 2000: The Relationship between Convection and Sea Surface Temperature on Intraseasonal Timescales. *Journal of Climate*, **13 (12)**, 2086 – 2104, doi:https://doi.org/10.1175/1520-0442(2000)013<2086:TRBCAS>2.0.CO;2.

Xie, B., J. C. H. Fung, A. Chan, and A. Lau, 2012: Evaluation of nonlocal and local planetary boundary layer schemes in the wrf model. *Journal of Geophysical Research: Atmospheres*, **117 (D12)**, doi:https://doi.org/10.1029/2011JD017080.

Xie, B., J. C. R. Hunt, D. J. Carruthers, J. C. H. Fung, and J. F. Barlow, 2013: Structure of the planetary boundary layer over southeast england: Modeling and measurements. *Journal of Geophysical Research: Atmospheres*, **118 (14)**, 7799–7818, doi:https://doi.org/10.1002/jgrd.50621.

Yanase, T., S. Nishizawa, H. Miura, T. Takemi, and H. Tomita, 2020: New Critical Length for the Onset of Self-Aggregation of Moist Convection. *Geophysical Research Letters*, **47 (16)**, e2020GL088 763, doi:https://doi.org/10.1029/2020GL088763.

Yang, D., 2018: Boundary layer diabatic processes, the virtual effect, and convective self-aggregation. *Journal of Advances in Modeling Earth Systems*, **10 (9)**, 2163–2176.

Yang, D., 2019: Convective Heating Leads to Self-Aggregation by Generating Available Potential Energy. *Geophysical Research Letters*, **46 (17-18)**, 10 687–10 696, doi:https://doi.org/10.1029/2019GL083805.

Yao, L., D. Yang, and Z.-M. Tan, 2022: A Vertically Resolved MSE Framework Highlights the Role of the Boundary Layer in Convective Self-Aggregation. *Journal of the Atmospheric Sciences*, **79 (6)**, 1615 – 1631, doi:10.1175/JAS-D-20-0254.1.

Yoneyama, K., 2003: Moisture Variability over the Tropical Western Pacific Ocean. *Journal of the Meteorological Society of Japan*, **81 (2)**, 317–337, doi:10.2151/jmsj.81.317.