

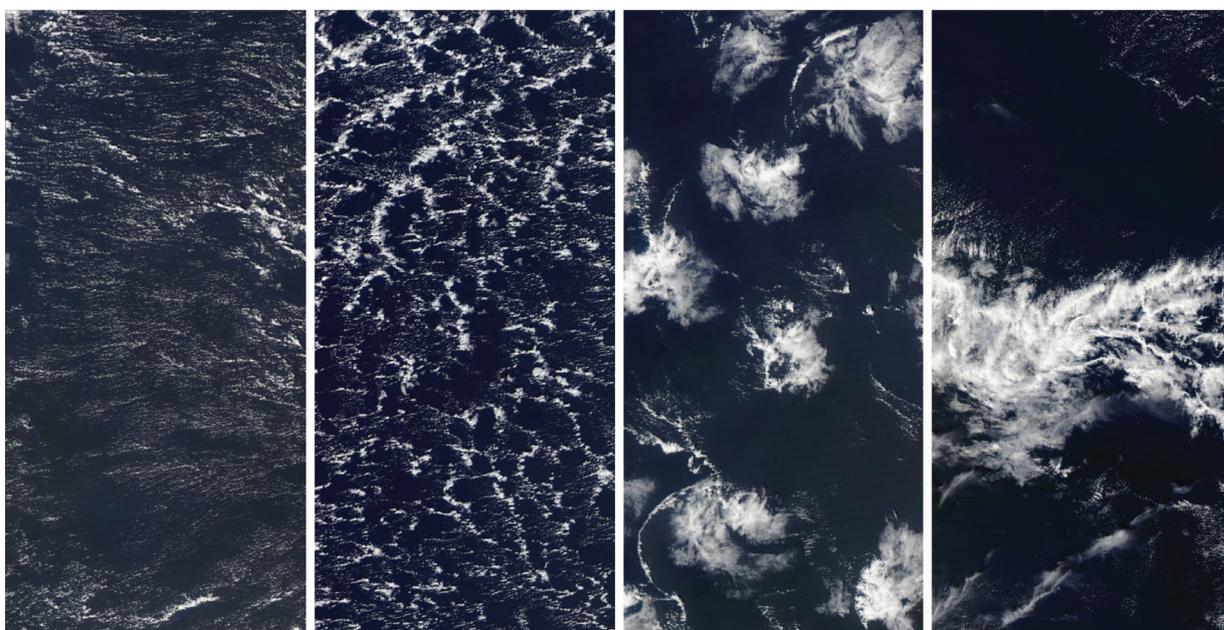


MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE



International Max Planck Research School on Earth System Modelling

Meso-Scale Patterns of Shallow Convection in the Trades



Hauke Schulz

Hamburg 2021

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0

Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de

www.mpimet.mpg.de

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including PhD theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Layout

*Bettina Diallo and Norbert P. Noreiks
Communication*

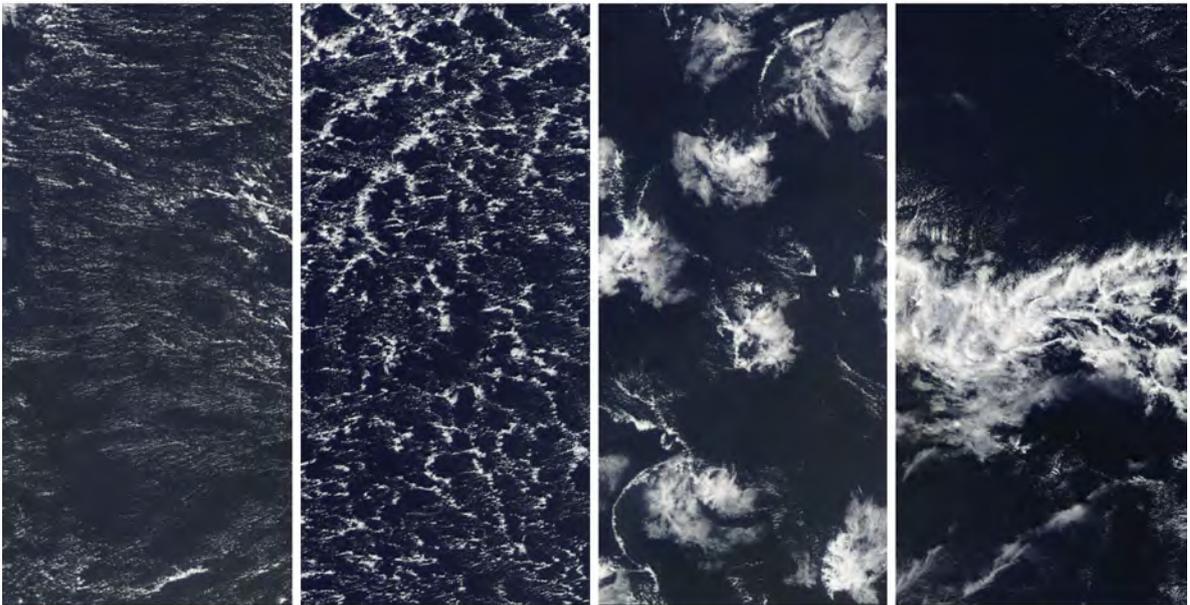
Copyright

Photos below: ©MPI-M

*Photos on the back from left to right:
Christian Klepp, Jochem Marotzke,
Christian Klepp, Clotilde Dubois,
Christian Klepp, Katsumasa Tanaka*



Meso-Scale Patterns of Shallow Convection in the Trades



Hauke Schulz

Hamburg 2021

Hauke Schulz

aus Hamburg, Deutschland

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 3. Dezember 2021

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Bjorn Stevens
Prof. Dr. Stephan Bühler

Vorsitzender des Promotionsausschusses:

Prof. Dr. Hermann Held

Dekan der MIN-Fakultät:

Prof. Dr. Heinrich Graener

Cover graphic: satellite images of the four meso-scale patterns Sugar, Gravel, Flowers and Fish (left to right) in natural colors. Tiles are 300km x 644km in size. Images are from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS).

Typeset based on the classicthesis template developed by André Miede, available at:
<https://bitbucket.org/amiede/classicthesis/>

Hauke Schulz

Meso-Scale Patterns of Shallow Convection
in the Trades

ABSTRACT

How will marine low-level cloudiness change in a warming climate? To answer this question a better process understanding of low-level cloudiness is needed. This dissertation uses a multitude of observations and large-eddy simulations to explore how meso-scale patterns of shallow convection relate to this challenging question. This study focuses on the downwind trades and its meso-scale patterns that only recently raised interest based on the work of Stevens et al. (2020) who supplemented the traditional classes of meso-scale patterns of the upstream trades. These new classes are named based on their visual impression *Sugar*, *Gravel*, *Flowers* and *Fish*. Here they are further investigated in terms of their climatic relevance, physical characteristics, atmospheric environment and emergence.

The core of these investigations consists of deep neural networks that have been trained to identify these patterns in satellite images. At the same time, the deep neural networks proved to be a valuable tool to derive a common perception of subjectively defined classes that do not have a ground truth. Although the crowd-sourced labels were therefore very noisy, the neural networks ranked among the highest in inter-annotator agreements.

The classification of the neural network reveals that the patterns are common to the trades beyond the winter season in the western North Atlantic and can represent more than 40% of the observed variability depending on season and region. This variability expresses itself not only in changes of the visual appearance but also physically in the cloud cover. A linear relationship between the cloud cover and the cloud radiative effect makes the processes leading to the patterns relevant for the climate.

The underlying physical processes of each meso-scale pattern are related to the air-mass origin with an influence of diurnal variations that are potentially modulating the large-scale factors. One large-scale factor that is most distinct among the patterns is wind speed. Other factors are only related to a particular pattern but can be a necessity for the pattern to form. *Fish* for example is associated with anomalously strong convergence. *Sugar* favors warmer surface temperatures. Both the forcing of *Fish* and *Sugar* are related to air-masses intruding from outside the trades, leaving *Gravel* and *Flowers* be the only native trade-wind patterns.

Large-eddy simulations reveal that they are in general capable of replicating the observed variability in meso-scale cloud patterns. However, they are unable to match the observed vertical distribution of cloudiness in both their absolute values and their variability in particular for *Flowers* and *Fish*. Nevertheless, the distribution of moisture and the presence of meso-scale circulations indicates that the responsible processes for the formation of the different patterns are captured and the simulations are a valuable tool to complement the observations to gain a better process understanding.

Based on the relationships between large-scale forcing and mesoscale patterns found in this dissertation, conditions preferred by patterns with higher cloud amount and more negative cloud radiative effect are expected to occur less frequently in a warming climate.

ZUSAMMENFASSUNG

Wie wird sich die niedrige Bewölkung über den Meeren in einem sich erwärmenden Klima verändern? Um diese Frage zu beantworten, ist ein besseres Verständnis der flachen Bewölkung und dessen Prozesse erforderlich. In dieser Dissertation wird anhand einer Vielzahl von Beobachtungen und Grobstruktursimulationen untersucht, wie mesoskalige Muster der flachen Konvektion mit dieser schwierigen Frage zusammenhängen. Diese Studie konzentriert sich auf die mesoskaligen Muster die in den westlichen Passatwindregionen aufkommen und erst kürzlich durch die Arbeit von Stevens u. a. (2020) beschrieben wurden. Sie sind somit ergänzend zu den bisherigen Klassen der mesoskaligen Muster die vorzüglich in den östlichen Passatwindregionen auftreten. Diese neuen Klassen werden aufgrund ihres visuellen Eindrucks *Sugar*, *Gravel*, *Flowers* und *Fish* genannt. Hier werden sie im Hinblick auf ihre klimatische Relevanz, ihre physikalischen Eigenschaften, ihre atmosphärische Umgebung und ihre Entstehung weiter untersucht.

Den Kern dieser Untersuchungen bilden tiefe neuronale Netze, die darauf trainiert wurden, diese Muster in Satellitenbildern zu erkennen. Gleichzeitig erwiesen sich die tiefen neuronalen Netze als wertvolles Instrument zur Ableitung eines gemeinsamen Verständnisses von subjektiv definierten Klassen, für die es keine Grundwahrheit gibt. Obwohl die von einer Vielzahl von Personen vergebenen Klassifizierungen sehr veräusert waren, erreichten die neuronalen Netze eine der höchsten Übereinstimmungen zwischen den Beobachtern.

Die Klassifizierung des neuronalen Netzes zeigt, dass die Muster über die Wintersaison hinaus im westlichen Nordatlantik verbreitet sind und je nach Jahreszeit und Region mehr als 40% der beobachteten Variabilität ausmachen können. Diese Variabilität äußert sich nicht nur in Veränderungen des visuellen Erscheinungsbildes, sondern auch physikalisch unter anderem in der Wolkenbedeckung. Eine lineare Beziehung zwischen der Wolkenbedeckung und der Strahlungswirkung der Wolken macht die Prozesse, die zu den Mustern führen, für das Klima relevant.

Die den einzelnen mesoskaligen Mustern zugrundeliegenden physikalischen Prozesse hängen mit dem Ursprung der Luftmasse zusammen, wobei tageszeitliche Schwankungen eine Rolle spielen, die möglicherweise die großräumigen Einflussgrößen modulieren. Eine großräumige Einflussgröße, die sich am deutlichsten von Muster zu Muster unterscheidet, ist die Windgeschwindigkeit. Andere Größen stehen nur mit einem bestimmten Muster in Beziehung, können aber für dessen Entstehung notwendig sein. *Fish* zum Beispiel ist mit einer anomal starken Konvergenz verbunden. *Sugar* bevorzugt wärmere Oberflächentemperaturen. Sowohl *Fish* als auch *Sugar* werden von Luftmassen beeinflusst, die von außerhalb der Passatwindregionen kommen, sodass *Gravel* und *Flowers* die einzigen ursprünglichen Passatwindmuster sind.

Grobstruktursimulationen zeigen, dass sie im Allgemeinen in der Lage sind, die beobachtete Variabilität in mesoskaligen Wolkenmustern zu replizieren. Sie sind jedoch nicht in der Lage, die beobachtete vertikale Verteilung der Bewölkung sowohl in Bezug auf die absoluten Werte als auch auf deren Variabilität, insbesondere für *Flowers* und *Fish*,

nachzuzahlen. Dennoch deuten die Verteilung der Feuchtigkeit und das Vorhandensein von mesoskaligen Zirkulationen darauf hin, dass die für die Bildung der verschiedenen Muster verantwortlichen Prozesse erfasst werden. Die Simulationen bilden somit ein wertvolles Instrument zur Ergänzung der vorhandenen Beobachtungen, um ein besseres Verständnis der Prozesse zu erlangen.

Auf der Grundlage der in dieser Dissertation gefundenen Beziehungen zwischen großskaligem Antrieb und mesoskaligen Mustern wird erwartet, dass Bedingungen, die von Mustern mit höherem Bedeckungsgrad und negativerem Wolkenstrahlungseffekt bevorzugt werden, in einem sich erwärmenden Klima weniger häufig auftreten.

ACKNOWLEDGMENTS

The dissertation is the culmination of the PhD, however, it lacks the ability to tell the story of people, possibilities and decisions that were influential and led not only to this dissertation but also my personal development. Here I'd like to acknowledge at least a few people.

Bjorn Stevens, who gave me the freedom to develop my own ideas, follow my interests, taught me to think beyond, but also to stay focused deserves a lot of credit. He supported me in any way possible and it is his enthusiasm that sparked also mine and gets me excited about the research ahead. Calling him my supervisor is a clear understatement.

I'm very grateful to Stephan Bühler and Victor Brovkin who also let me follow my own interests, were great mentors and made the panel meetings a joyful event.

I thank Rob Wood and Chris Bretherton for hosting me for three months at the University of Washington and giving me the opportunity to discuss my work with very bright students and colleagues on the other side of the Atlantic. A research stay that has a lasting effect owing to many new friends and colleagues. In particular I want to thank Jim, his wife and Adam for their generousness; Isabel, Ryan and Sam are thanked for enlightening scientific discussions in the office, pub or at sea and the great collaborations that followed so far.

The Tropical Cloud Observatory group has been a long-term companion and I am happy for the common evolution and all the technical insights and experiences I could gain along the way, not to mention the planning of and participation in the EUREC⁴A field campaign.

This dissertation would not have been possible without the whole atmosphere department at the Max-Planck Institute for Meteorology and a group of scientists around Sandrine Bony at the LMD giving me their trust in my work and classifying satellite images for a whole workday. I would still be sitting and classifying these images, if it weren't for you.

In times when office colleagues were the norm, it has been a joy to share an office with Laura. Geet has been more helpful than he probably imagined and it has been a great pleasure to get to know him. Ideas outside the box were best developed with him and Tobias.

Further, I would like to thank Geet, Irene and Yakelyn for their valuable feedback on an earlier draft of the dissertation.

Last but not least, I am grateful for the financial support of the International Max Planck Research School on Earth System Modelling and the Max Planck Gesellschaft and even more so that this support let me follow my own ideas.

PREFACE

The research on trade-wind cloudiness has been very active during my time as a PhD student, especially because of the activities my research group, the Tropical Observation Group, organized and participated in. This gave me the great opportunity to develop my own ideas and also join forces with other scientists at the institute but also internationally. This led ultimately to several publications that are listed on page [xi](#).

Here, a selection of seven studies that is explicitly marked on page [xi](#) is chosen for a focus on the meso-scale patterns of shallow convection in the trades. These studies form the basis of this cumulative dissertation and their essence and motivation is presented in a unifying essay and complemented by an overall conclusion and outlook. These studies are added as appendices. My contributions to each of these publications is stated at the beginning of each study in the appendix.

PRE-PUBLISHED OR INTENDED FOR PUBLICATION WORK
RELATED TO THIS DISSERTATION

MANUSCRIPTS FOCUSED ON IN THIS DISSERTATION AND INCLUDED IN THE AP-
PENDIX

- Bony, Sandrine, **Schulz, Hauke**, ... et al. (2020). Sugar, Gravel, Fish, and Flowers: Dependence of Mesoscale Patterns of Trade-Wind Clouds on Environmental Conditions. *Geophysical Research Letters* 47.7, e2019GL085988. DOI: [10.1029/2019GL085988](https://doi.org/10.1029/2019GL085988).
- Rasp, Stephan, **Schulz, Hauke**, ... et al. (Nov. 2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. *Bulletin of the American Meteorological Society* 101.11, E1980–E1995. DOI: [10.1175/BAMS-D-19-0324.1](https://doi.org/10.1175/BAMS-D-19-0324.1).
- Stephan, Claudia Christine et al. (Feb. 2021). Ship- and Island-Based Atmospheric Soundings from the 2020 EUREC⁴A Field Campaign. *Earth System Science Data* 13.2, pp. 491–514. DOI: [10.5194/essd-13-491-2021](https://doi.org/10.5194/essd-13-491-2021).
- Vial, Jessica, ..., **Schulz, Hauke** et al. (2021). On the Daily Cycle of Mesoscale Cloud Organization in the Winter Trades. *Quarterly Journal of the Royal Meteorological Society* 147.738, pp. 2850–2873. DOI: [10.1002/qj.4103](https://doi.org/10.1002/qj.4103).
- Schulz, Hauke** (in preparation). Meso-scale patterns of shallow convection during EUREC⁴A.
- Schulz, Hauke**, ... et al. (2021). Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades. *Journal of Geophysical Research: Atmospheres* 126.17, e2021JD034575. DOI: [10.1029/2021JD034575](https://doi.org/10.1029/2021JD034575).
- Schulz, Hauke**, Stevens, Bjorn et al. (in preparation). Assessment of meso-scale patterns of shallow convection in realistically forced large-eddy simulations.

ADDITIONAL MANUSCRIPTS I CONTRIBUTED TO DURING MY TIME AS A PHD STU-
DENT AND HELPED TO GAIN FURTHER INSIGHTS INTO SHALLOW CONVECTION IN
A BROADER CONTEXT

- George, Geet, ..., **Schulz, Hauke**, ... et al. (2021). JOANNE : Joint Dropsonde Observations of the Atmosphere in Tropical North atlAntic Meso-Scale Environments. *Earth System Science Data Discussions*, pp. 1–33. DOI: [10.5194/essd-2021-162](https://doi.org/10.5194/essd-2021-162).
- Stevens, Bjorn, ..., **Schulz, Hauke**, ... et al. (2020). Sugar, Gravel, Fish and Flowers: Mesoscale Cloud Patterns in the Trade Winds. *Quarterly Journal of the Royal Meteorological Society* 146.726, pp. 141–152. DOI: [10.1002/qj.3662](https://doi.org/10.1002/qj.3662).
- Stevens, Bjorn et al. (Aug. 25, 2021). EUREC⁴A. *Earth System Science Data* 13.8, pp. 4067–4119. DOI: [10.5194/essd-13-4067-2021](https://doi.org/10.5194/essd-13-4067-2021).
- Vogel, Raphaela, ..., **Schulz, Hauke**, ... et al. (in review). A Climatology of Trade Cumulus Cold Pools and Their Link to Mesoscale Cloud Organization, p. 28.

CONTENTS

I UNIFYING ESSAY

1	BACKGROUND	3
1.1	Importance of low-level clouds	3
1.2	Meso-scale patterns of shallow convection in the trades	3
2	IDENTIFICATION	7
2.1	Design/Workflow	7
2.2	Evaluation	9
3	IMPORTANCE OF MESO-SCALE SHALLOW CONVECTION IN THE TRADES	11
3.1	Temporal and spatial distribution	11
3.2	Physical impact	12
4	PHYSICAL CHARACTERIZATION	15
4.1	Cloudiness	15
4.2	Cloud-scale and large-scale drivers	16
5	TOWARDS A PROCESS UNDERSTANDING WITH LARGE-EDDY SIMULATIONS	21
6	CONCLUSION	25
7	OUTLOOK	27

II APPENDICES

A	COMBINING CROWDSOURCING AND DEEP LEARNING TO EXPLORE THE MESOSCALE ORGANIZATION OF SHALLOW CONVECTION	31
B	SUGAR, GRAVEL, FISH AND FLOWERS: DEPENDENCE OF MESOSCALE PATTERNS OF TRADE-WIND CLOUDS ON ENVIRONMENTAL CONDITIONS	51
C	CHARACTERIZATION AND EVOLUTION OF ORGANIZED SHALLOW CONVECTION IN THE DOWNSTREAM NORTH ATLANTIC TRADES	63
D	ON THE DAILY CYCLE OF MESOSCALE CLOUD ORGANIZATION IN THE WINTER TRADES	91
E	ASSESSMENT OF MESO-SCALE PATTERNS OF SHALLOW CONVECTION IN REALISTICALLY FORCED LARGE-EDDY SIMULATIONS	121
F	SHIP- AND ISLAND-BASED ATMOSPHERIC SOUNDINGS FROM THE 2020 EUREC ⁴ A FIELD CAMPAIGN	143
G	MESO-SCALE PATTERNS OF SHALLOW CONVECTION DURING EUREC ⁴ A	177
	BIBLIOGRAPHY	197

ACRONYMS

ABI	Advanced Baseline Imager
BCO	Barbados Cloud Observatory
CERES	Clouds and the Earth's Radiant Energy System
CNN	Convolutional Neural Network
CRE	Cloud Radiative Effect
ECS	Equilibrium climate sensitivity
ERA-Interim	ECMWF Reanalysis Interim
ERA5	ECMWF Reanalysis v5
EUREC ⁴ A	Elucidating the Role of Cloud–Circulation Coupling in Climate
GOES	Geostationary Operational Environmental Satellite
ICON	ICOsahedral Nonhydrostatic
ITCZ	Inter-Tropical Convergence Zone
LES	Large Eddy Simulation
LCL	Lifting Condensation Level
LTS	lower tropospheric stability
MODIS	Moderate Resolution Imaging Spectroradiometer
NN	neural network
PAMTRA	Passive and Active Microwave TRANSfer package
SST	sea surface temperature
TOA	top of the atmosphere

Part I

UNIFYING ESSAY

BACKGROUND

*In order to understand geometric shapes,
I believe that one must see them*

— Benoit Mandelbrot (Mandelbrot, 2010)

1.1 IMPORTANCE OF LOW-LEVEL CLOUDS

Clouds are omnipresent and cover more than 70% of the earth at any time (Siebesma et al., 2020). They occur at all latitudes and altitudes and play a primary role in the planet's radiative energy budget. On the annual average, they reduce the absorbed energy by about -18 W m^{-2} at the top of the atmosphere (TOA) in the current climate (Loeb et al., 2018). Depending on factors like the cloud type and latitude of occurrence this so called Cloud Radiative Effect (CRE) differs. High, thin clouds for example are less efficient in reflecting sunlight than low, in particular optically thick clouds. In fact, thin high clouds trap more energy than they reflect, causing a positive CRE. In contrast, the shortwave cloud radiative effect - the ability to reflect solar radiation - surpasses the longwave cloud radiative effect - the ability to trap outgoing longwave radiation - for low clouds. This negative CRE paired with the sheer abundance of this cloud type make them largely contributing to the current climates' net CRE and important for future climate scenarios.

In particular over the oceans, marine low-level clouds contribute largely to a mean cooling effect of clouds. How this low-level cloudiness changes in a warming climate is however one of the leading questions in climate science. The uncertainty in tropical low-cloud feedback to a doubling of CO_2 largely contributes to the uncertainty of the estimated equilibrium climate sensitivity (ECS; the equilibrium surface temperature response to a doubling of atmospheric CO_2) (Vial et al., 2013). Meehl et al. (2020) emphasized that Earth system model experiments alone are not sufficient to reduce the uncertainty in ECS as their estimates have stagnated over the last three decades. Sherwood et al. (2020) show that additional lines of evidence are needed to reduce the uncertainty by e.g. paleoclimate evidence and process understanding. With the low-level clouds carrying most of this uncertainty, its processes need to be better understood.

A process that has attracted little attention in the past is the meso-scale patterning of shallow convection into clusters, in particular in the downwind trades.

1.2 MESO-SCALE PATTERNS OF SHALLOW CONVECTION IN THE TRADES

The mesoscale and trade-wind convection share a common issue. They were both for a long time hardly investigated. It wasn't until Ligda (1951), who studied precipitating echoes in radar observations that he could neither attribute to the micro-scale nor to the

synoptic scale, coined the term "meso-scale". Just two years earlier, Joanne Simpson, who did pioneering work on trade-wind convection, was derided about studying trade-wind convection in general due to its assumed pettiness by senior researchers (Fleming, 2020). She thankfully continued her pioneering work and among other insights recognized the structure present in tropical clouds despite being limited to aircraft observations that provide far less of an overview on the cloud patterning than satellite images enabled later (Malkus and Riehl, 1964).

With the advent of the satellite era the patterning of clouds on many different scales became visible. Especially, the dominant widespread stratiform layers in the upstream trades at the western coasts of the continents attracted considerable attention (e.g. Agee (1987)). Because these stratiform clouds have a strong negative net radiative effect which is sensitive to slight changes in their characteristics (Wood, 2012), these clouds and their patterning are still a very active field of research. However, the traditional classes of this patterning, namely open and closed mesoscale cellular convection, are occurring predominately in the mid-latitudes over cold ocean water upwelling at the western sides of the continents (Muhlbauer et al., 2014).

Recently, Stevens et al. (2020) tied up to this history in an attempt to close a long standing gap: do recognisable meso-scale patterns exist besides the ones in the upwind trades, also in the downwind trades? The curiosity was led by the question regarding whether shallow convection can alter the net energy balance at the TOA by means of organization forms, like in the case of deep convection (Tobin et al., 2012). While Stevens et al. (2020) did not answer this question, they did identify and agreed upon four patterns of shallow convection that seem to be reoccurring in the western North Atlantic winter trades. Fig. 1.1 shows an example cloud scene for each of the identified patterns.

These patterns were chosen subjectively purely by their appearance and named accordingly: *Sugar*, *Gravel*, *Fish*, *Flowers*.

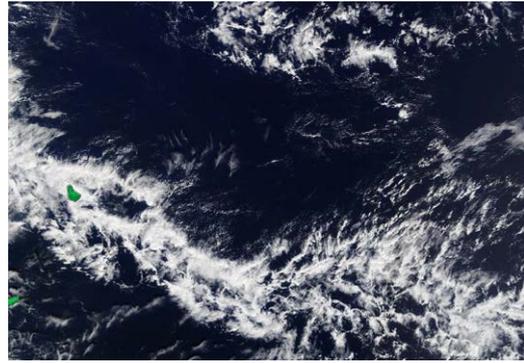
- *Sugar* is named after the very fine dusting of cloud patches that are often arranged by the large-scale flow.
- *Gravel* describes a cloud field that is characterised by cloud patches arranged along arc-like or hexagonal structures at the meso- β scale (20 km to 100 km) whose enclosures are often cloud free.
- *Fish* is the pattern containing the largest cloud patches and its organization looks similar to a skeleton of a Fish. It is organized on the meso- α scale (200 km to 2000 km) as per Orlanski (1975). Its longitudinal cloud patch and the large clear-sky areas enclosing it are both characteristic.
- *Flowers* are describing a collection of rather circular cloud patches on the meso- β scale (20 km to 200 km) which are quasi-regularly distributed and well separated from each other by clear-sky regions.

The proof that patterns of shallow convection exist in the downwind trades, in particular in the North Atlantic trades, has been made by Stevens et al. (2020). Because their patterns were purely based on the visual impression these cloud scenes made to the 12 authors studying them, the following questions arise naturally:

Sugar



Fish



Gravel



Flowers

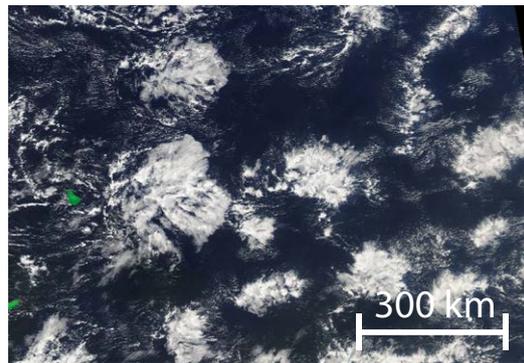


Figure 1.1: Meso-scale patterns of shallow convection in the trades as defined by Stevens et al. (2020) and observed in Moderate Resolution Imaging Spectroradiometer images of true color channel composite. Green overlay indicates landmasses with Barbados in the western part of the image.

1. Are these meso-scale patterns of shallow convection important for the climate system?
2. Are the patterns besides their visual appearance also physically distinct?
3. How do the patterns relate to cloud-scale and large-scale forcing?
4. Can large-eddy simulations help to gain a process understanding of the pattern morphology?

This dissertation will address these questions by means of observations and large-eddy simulations. First of all, the question of whether these patterns can even be efficiently detected will be answered in the next section.

IDENTIFICATION

To study the newly defined patterns of meso-scale organization of Stevens et al. (2020), a method to detect these patterns consistently and quickly, preferably on global scale was needed. The subjective definition of these patterns can be described quite well to other humans during a training session with visuals, but is hard to formulate. Supervised neural network architectures can directly infer the rules from human visual classifications themselves without any abstraction. They are in this case advantageous to a traditional rule-based algorithms. It should be noted though that several rule-based algorithms have been developed to identify organization also in the field of meteorology (Beucler et al., 2020; Tobin et al., 2012; Tompkins and Semie, 2017). However, they are often oversimplifying and are dependent on cloud size, cloud number or specific to a certain application. The struggle to find a rule-based algorithm that separates cloud scenes is also reflected in Janssens et al. (2021), where 21 metrics have been tested to separate the observed variance of shallow cloud patterns in the trades.

Study B also relies on two of these metrics, the mean object size and the organization index (I_{org}) to identify the different patterns. Although this combination describes well the variability of shallow convection, it is not able to cover the different scales of patterns noted in Sec. 1. It is restricted to the description of a fixed domain that might however contain different forms of meso-scale patterns. For brevity this classification approach is not further described here, but the reader is pointed to B.2.1. Here the focus is on an object detection algorithm, which is able to classify subsections of a domain.

Image recognition with the help of deep neural networks (in the following just neural networks) has developed quickly in recent years, but usage in geophysics was however still limited. Study A show-cases how such an algorithm can be applied to spatial geophysical problems that are often connected with visual classifications of e.g. extreme weather patterns or land-use detection (Reichstein et al., 2019). It also lays out how the workflow including the creation of labels can look like. In particular, it shows how to detect the subjectively chosen patterns of shallow convection in the trades. Here the main concepts are outlined.

2.1 DESIGN/WORKFLOW

Study A presents two deep neural networks: one object detection algorithm and one segmentation algorithm. Here, the focus is on the object detection algorithm as its classification format is closest to the human classifications that are explained in the following section and is also used in some variant in the following physical analysis studies C, D and E.

For the training of a supervised neural network, classifications of the subjects are needed. Because the task to classify these particular cloud patterns was new, only the classifications from Stevens et al. (2020) existed. However, one of the conclusions of Stevens et al. (2020) is that the defined meso-scale patterns have different characteristic

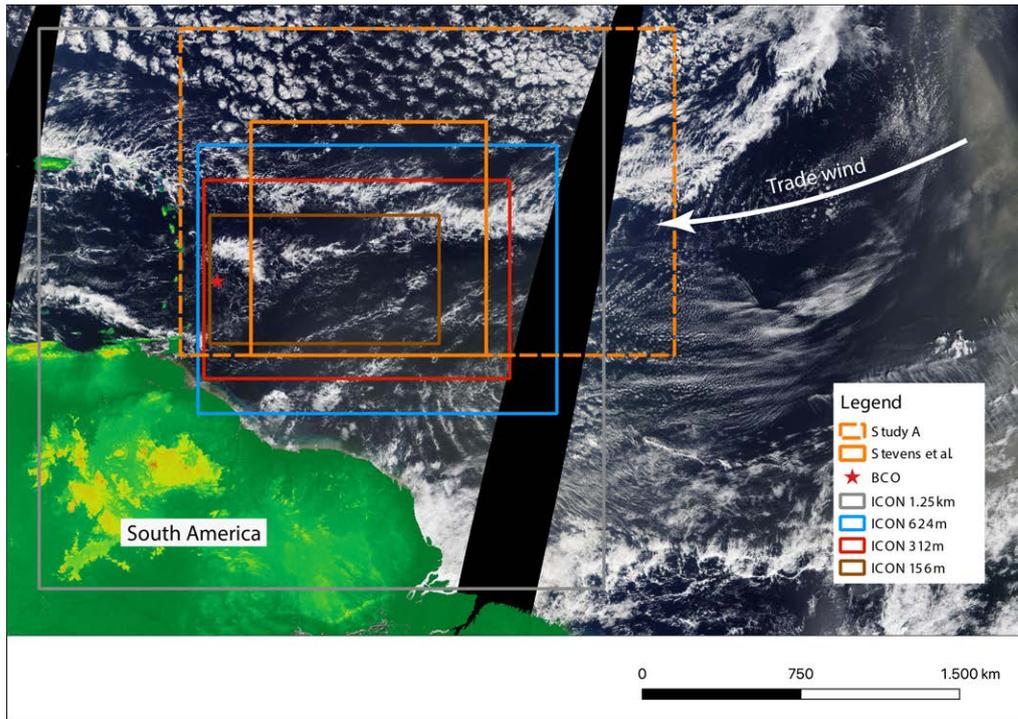


Figure 2.1: Overview of domains used in this work and Stevens et al. (2020). MODIS image from 20th January 2020 in the background overlaid with elevation map of landmasses.

area-fractions. *Gravel* was found most often dominating the study area, while *Sugar* has been noted by the authors to occur frequently but often did not dominate in the 10 deg x 10 deg domain they studied. It was therefore less often classified as they only classified entire satellite images. To better reflect the variability of patterns in the classifications, smaller sub-regions need to be classified. The example satellite image in Fig. 2.1 illustrates how this rather small domain of Stevens et al. (2020) can contain several meso-scale patterns. While the northern part would be labeled as *Fish*, the southern part contains cloud patches that can be attributed to *Sugar*.

To consider these findings and increase the amount of labels to sufficiently train a neural network, especially in case of noisy classes, a new workflow has been designed.

2.1.1 Labels

While Stevens et al. (2020) have classified an entire domain of 10 deg x 10 deg (see Fig. 2.1), study *A* increases the domain size to 21 deg x 14 deg and allows to classify sub-regions. These labels were created in form of rectangular bounding boxes. In two internal crowd-sourcing events at the Max-Planck-Institut für Meteorologie and the Laboratoire de Météorologie Dynamique with a total of 67 participating scientists, almost 50.000 of such labels were created based on 10.000 satellite images. They were captured by the MODIS instrument aboard the satellites AQUA and TERRA. This quick growth in labels was possible thanks to the scientists taught to identify the cloud patterns, who were highly motivated as their own interest was often linked to clouds in the trades. For many this was also a great opportunity to get a better feeling for these clouds as they normally

focus on model simulations. Second, besides the region-season used in Stevens et al. (2020) additional regions and seasons were added based on their similarity to the first. For the similarity analysis large-scale cloud controlling factors were used and increased the chance that images would contain the patterns. This analysis is further described in A.11.1. Third, an online platform common for crowd-sourcing labels enabled a straight forward classification. The participants were automatically served a random cloud scene and were asked to draw bounding boxes around the patterns. An example of these bounding boxes can be seen in Fig. A.3. Last, the bounding boxes approach allowed to increase the domain size to 21×14 deg as shown by the dashed orange rectangle in Fig. 2.1 and the marking of several classes at one sight. Bounding boxes were chosen over precise outlines to optimize for labeling speed rather than precision which in any case is hard to achieve for these patterns with undefined boundaries.

2.1.2 Neural network

The deep neural networks used in studies A, C, D and E are Convolutional Neural Network (CNN) RetinaNets. They are used with the ResNet-50 backbone that has been pre-trained with a variety of different images and classes (Lin et al., 2018) making the training process for the classification of new classes, here the classes of meso-scale patterns of shallow convection, computationally less expensive. General features that are useful for many object detection tasks like edge detection are already learned in a pre-trained model and will only adjust slightly to the new task. In addition, the Keras implementation enables an easy integration into the python framework many scientists are familiar with.

While the neural network trained in study A is based on the visible images that have been labeled, a second neural network is trained with the same labels but with infrared images as input. These infrared images are captured by MODIS simultaneously to the visual ones such that they cover the identical cloud scene. This infrared neural network (NN) extends the application to night time images and also makes it rather independent of the solar cycle and satellite viewing angle. The neural network trained with infrared images could therefore be applied to a different imager, the Advanced Baseline Imager (ABI) installed on Geostationary Operational Environmental Satellite (GOES)-16. Because GOES-16 is a geostationary satellite, its sensors capture the research area at a higher temporal resolution as compared to the MODIS imager flown on the satellites AQUA and TERRA, which are polar-orbiting and capturing the area only twice a day each. The comparison of the different NNs and image sources show a good agreement to each other (C.9.1).

2.2 EVALUATION

The challenge of the development of an algorithm that should detect subjective classes compared to those where consent exists is the lack of a ground-truth. The perception of what a human or a car defines are universal and the quality of labels can be validated against a ground-truth. In case of patterns of meso-scale shallow convection however, even trained scientists might disagree on classifications as there are no clear boundaries

of the patterns and nature's variability is much larger than the canonical examples shown in Fig. 1.1.

The NN in this case is rather trained to detect the joint consent of the scientific community and as a consequence produces less noisy labels. This reduction of noise becomes evident in the inter-comparison of the labels created by the NN and humans. The NN ranks at the top with one of the largest agreements among all participants.

The neural network has therefore proven to have learned the skills to detect the newly defined patterns of shallow convection in the trades and built a community consent.

IMPORTANCE OF MESO-SCALE SHALLOW CONVECTION IN THE TRADES

The patterns were chosen purely subjectively by the visual impression they left on trained scientists who defined the most recognizable ones in an effort described in Stevens et al. (2020). This appearance based approach has been used already by Howard (1803) and led to the distinctions of cloud types as seen from the ground and known to people also outside the field. We know today that these visual distinctions have also a physical meaning and are tightly connected to weather conditions. Whether the defined patterns of shallow convection are meaningful as well and relevant to the trade-wind region are valid questions. In this section they are addressed by answering the following subquestions:

1. Do the patterns occur beyond the original study area and season and can they be generalized to the broader global trade-wind regimes?
2. Do the patterns represent the variability in the trades or are they a marginal phenomenon?
3. Is the variability in visual manifestations based on physical differences?

3.1 TEMPORAL AND SPATIAL DISTRIBUTION

Stevens et al. (2020) focused on a 10deg × 10deg region just east of Barbados in the western North Atlantic trade-wind regime (Fig. 2.1). With their approach they attributed about 37% of the cloud scenes in the winter time to one of the patterns. This can however be seen as a lower bound, because their classification technique was very conservative and only allowed clear and widespread patterns. Study C derives a similar percentage of about 40%, but also quantifies that an additional 36% can be attributed to days with a mixture of patterns. Only 20% could not be attributed to any of the four classes. (Tab. C.2)

While 37% of occurrence already state the local importance of these patterns for the northern hemispheric winter, the question remains, whether these patterns also describe large parts of the variability in cloudiness in other regions of the globe, specifically other maritime trade-wind regions. The manual classifications used for the training of the NNs give already a hint that other regions with a similar atmospheric environment enclose these patterns as well (not shown).

However, the trained neural network is able to extend these classifications to the global scale. As Fig. 3.1 reveals, the four patterns are indeed present in all major ocean basins, in particular in the respective trade-wind regimes. Individual patterns can dominate the cloud structures with a frequency of occurrence of up to 50% and highlight the importance they might have.

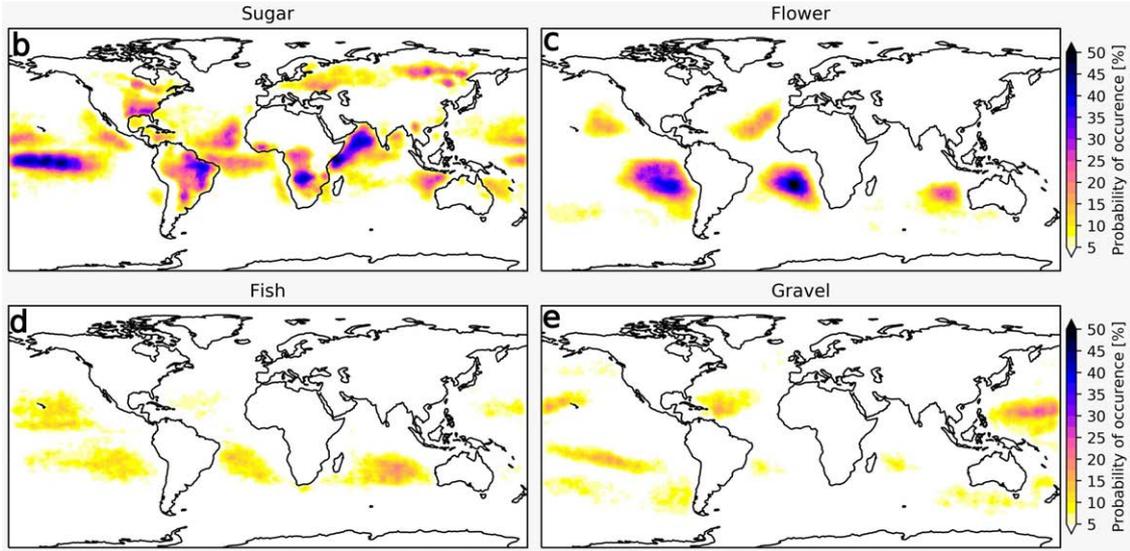


Figure 3.1: Global distribution of meso-scale shallow convection patterns in 2017 based on the neural network classifications

The patterns are not equally distributed and favour different geographical locations which hint that different atmospheric environments might be necessary to cause these patterns. This hypothesis is further strengthened by analysing the seasonal dependence of these patterns. Fig. 3.2 shows this dependence for the North Atlantic.

During the winter season (DJF) initially chosen by Stevens et al. (2020) all patterns are present with similar frequencies of occurrences albeit with slight differences in their geographical distribution. *Gravel* and *Fish* are mostly confined to the downwind trades and mainly occur only in the winter season. *Sugar* is showing no preferred season, but is enfaming the Inter-Tropical Convergence Zone (ITCZ) and is following its seasonal migration. By contrast, *Flowers* expose a strong seasonal cycle with a peak in spring in the upstream trades, which is traditionally associated with stratiform cloudiness.

3.2 PHYSICAL IMPACT

The frequency of occurrence alone however does not make the patterns relevant for the climate system. A measure that is commonly used to identify the importance of clouds is the CRE. It is a measure of how clouds affect the radiative energy budget at the TOA and estimated by comparing the TOA-fluxes of clear-sky regions ($F_{\text{net,clear}}$) to all regions ($F_{\text{net,allsky}}$).

$$\text{CRE} = F_{\text{net,clear}} - F_{\text{net,allsky}} \quad (3.1)$$

From a negative CRE follows that cloudy regions have a cooling effect on the planet relative to clear-sky regions. In contrast, a positive CRE would describe situations when clouds warm the planet compared to clear-sky regions.

Study B presents such an investigation. Based on satellite products captured by the satellite-born instruments MODIS and CERES (Clouds and the Earth's Radiant Energy

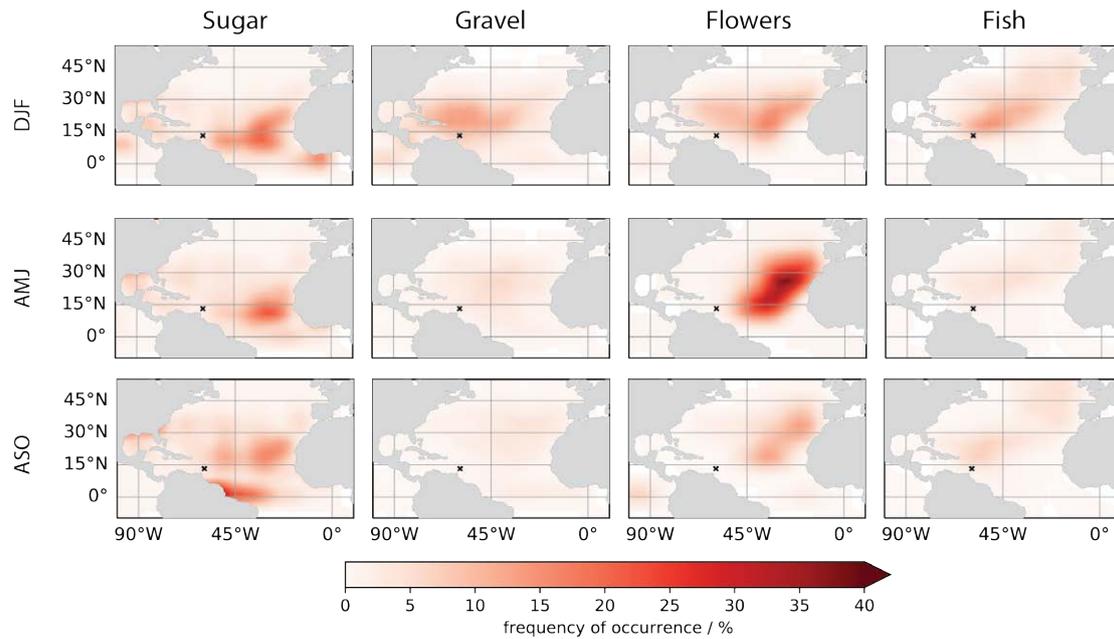


Figure 3.2: Seasonal distribution of meso-scale shallow convection patterns in the North Atlantic in the (December-January-February [DJF]), transitional- (April-May-June [AMJ]) and wet- (August-September-October [ASO]) season

System) the low cloud amount and the net CRE have been estimated. Grouped by the detected patterns¹ distinct CRE become visible (Fig. 3.3b).

Depending on the pattern, the net CRE differs. While it is negative for all patterns, *Flowers* and *Fish* are able to reflect more sunlight than *Sugar* and *Gravel*. This is primarily caused by the strong dependence of CRE on marine low-level cloud amount. Both quantities have a linear relationship as seen in Fig. 3.3a. *Flowers* with their highest cloud amount compared to the other patterns have also the highest cooling effect, while *Sugar* is only half effective.

As a consequence, this investigation of the CRE reveals that these patterns of shallow convection indeed matter for the Earth's system energy balance. The CRE alone says however little about the impacts of these cloud patterns in a changing climate. The frequency of occurrence, the area covered and finally the characteristics of the patterns themselves might change. A thorough understanding of the mechanisms that lead to one pattern or the other is therefore important to gain. The following chapter will build the basis for this understanding by characterising the patterns from a physical perspective that will help to judge in Chapter 5 whether simulations capture the characteristics and can be used to gain further process understanding.

¹ Please note, the patterns are in this case distinguished by the metrics I_{org} and mean object size (S). More details can be found in study B.2.1

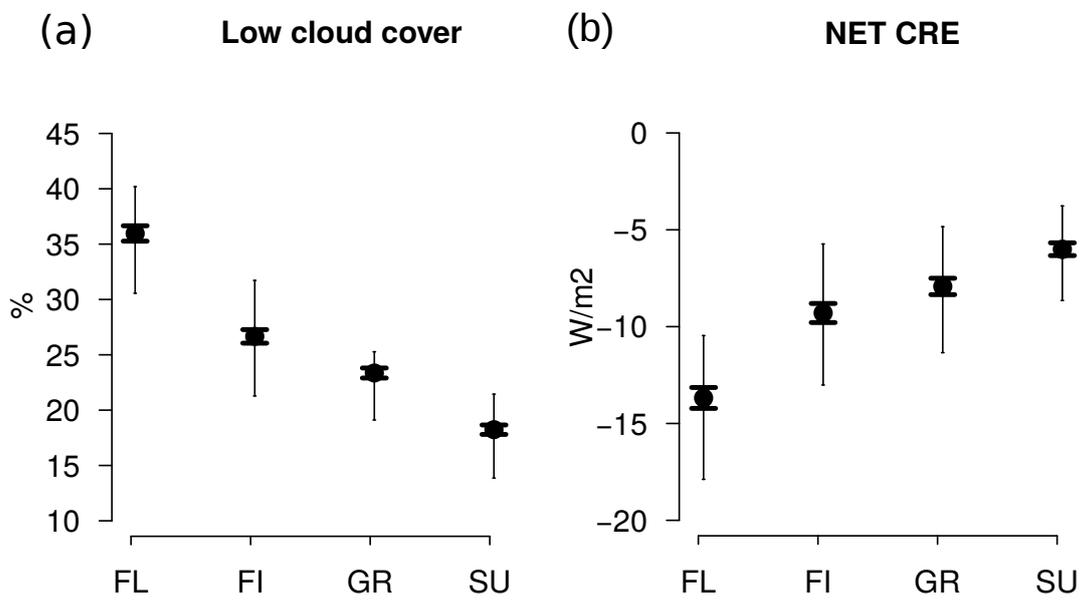


Figure 3.3: Low cloud cover and net CRE for the different cloud patterns: Flowers (FL), Fish (FI), Gravel (GR) and Sugar (SU)

OBSERVED CHARACTERISTICS OF MESO-SCALE CLOUD PATTERNS

4.1 CLOUDINESS

The characterization of cloudiness follows naturally from the visual definition of the patterns. The highly reflective cloud patches above the dark ocean surface attract the viewer's attention. The investigation of cloudiness beyond the spatial distribution of cloud patches larger than a few kilometers is however challenging from space. Clouds in the trades are often smaller than the resolution of satellite imagers and cannot be resolved. Study C combines the satellite classification of the meso-scale patterns with ground-based measurements from the Barbados Cloud Observatory (BCO) that are able to resolve individual clouds but also their environment to overcome this challenge. With profiling instruments like lidar and radar it is further possible to capture thermo-dynamic profiles including the vertical cloud distribution with great detail.

The BCO (Stevens et al., 2016) is located at the most windward tip of the Caribbean island Barbados (Fig. 2.1), capturing the undisturbed maritime trade-winds during the northern hemispheric winter. It is equipped besides standard meteorological sensors for pressure, temperature, humidity and wind among others with a microwave rain radar and a cloud radar that are used extensively in this chapter. Additional radiosonde launches to the ones at the nearby airport are done during site visits or field campaigns (e.g. study F). The technical details of the instrument-specific analysis are left out for brevity, but the interested reader is pointed to the respective section in study C.

The meso-scale contextualisation is common to all measurements. With the help of the satellite observations the measurements from BCO are put into the meso-scale context. Differences between the spatial and temporal perspective that the combination of satellite observations with observations from a fixed ground-based station implies, are solved in three steps. First, the neural network is applied to infrared images captured by the ABI-instrument onboard the geo-stationary satellite GOES-16. This enables the identification of patterns every ten minutes both day and night. Second, the classifications are evaluated at the point of interest e.g. the BCO. Last the ground-based observations are attributed to the most dominant meso-scale pattern within a given time window (study C) or the closest classification (study D).

The satellite measured cloud cover in Fig. 3.3a can now be decomposed into its vertical distribution. Two characteristics become visible. First, as Fig. 4.1 reveals the bimodality of the echo fraction. One mode is observed at the Lifting Condensation Level (LCL) at about 800 m while the other one is closer to 2 km just below the typical trade-inversion height.

The cloudiness at the LCL is fairly constant and differences can be attributed to precipitation that is included in the measure of echo fraction. The differences in cloud cover that are observed in Fig. 3.3a are therefore primarily caused by the cloudiness aloft. This is in agreement with Nuijens et al., 2015 who found that the variability in

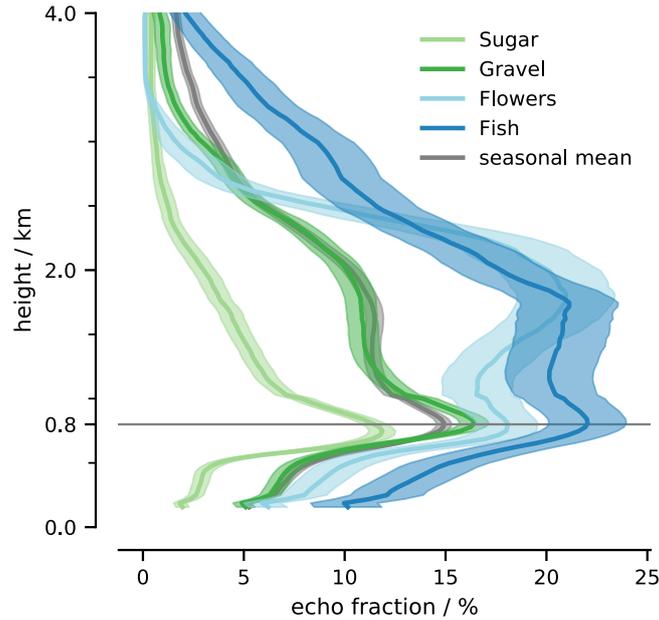


Figure 4.1: Echo fraction measured at the BCO grouped by the four meso-scale patterns of shallow convection. The wintertime (DJF) mean is included for comparison.

trade-wind cloudiness is not caused by cloudiness at the LCL but by cloudiness aloft. Now, this variability can be attributed to different meso-scale patterns.

Besides the fairly constant cloud fraction of cumulus humilis across patterns, also cloud geometries like cloud base height and cloud top height of traditional cloud types did not show a relationship to the apparent meso-scale pattern. A stratiform cloud for instance had the same cloud thickness when observed during *Flowers* or *Fish*. Only the cloud length of stratiform layers increased from *Sugar* via *Gravel* to *Fish* and *Flowers*. Although in case of *Sugar* and *Gravel* stratiform clouds are very rarely observed.

The intra-pattern variability found between 1.5 km and 2.5 km is the reason for the differences in overall cloud cover and therefore the differences in CRE.

4.2 CLOUD-SCALE AND LARGE-SCALE DRIVERS

Precipitation is one process that can play a major role in organizing convection through the creation of cold pools. Although each individual precipitation event is rather local, in bulk they are able to alter the form of organization as they actively force new convection when the cold pools spread and eventually collide. In the most idealized case hexagonal structures build up very similar to those seen in open mesoscale cellular convection and *Gravel*. The meso-scale appearance can therefore be affected by cloud-scale events but also by the large-scale through e.g. changes of the large-scale environment. These influences manifest themselves in the inversion strength, the subsidence rate or in averaged quantities of wind speed, temperature and pressure measured at the surface.

A quantification of these properties is therefore very valuable to further understand the differences across patterns and also to provide guidance to evaluate and eventually improve climate simulations.

Study C examines these properties by the means of observational data from the BCO as outlined in the previous section and ECMWF Reanalysis v5 (ERA5) reanalysis data. While the examination of each individual property is left for the reader to read up in study C the synergy is illustrated in Fig. 4.2.

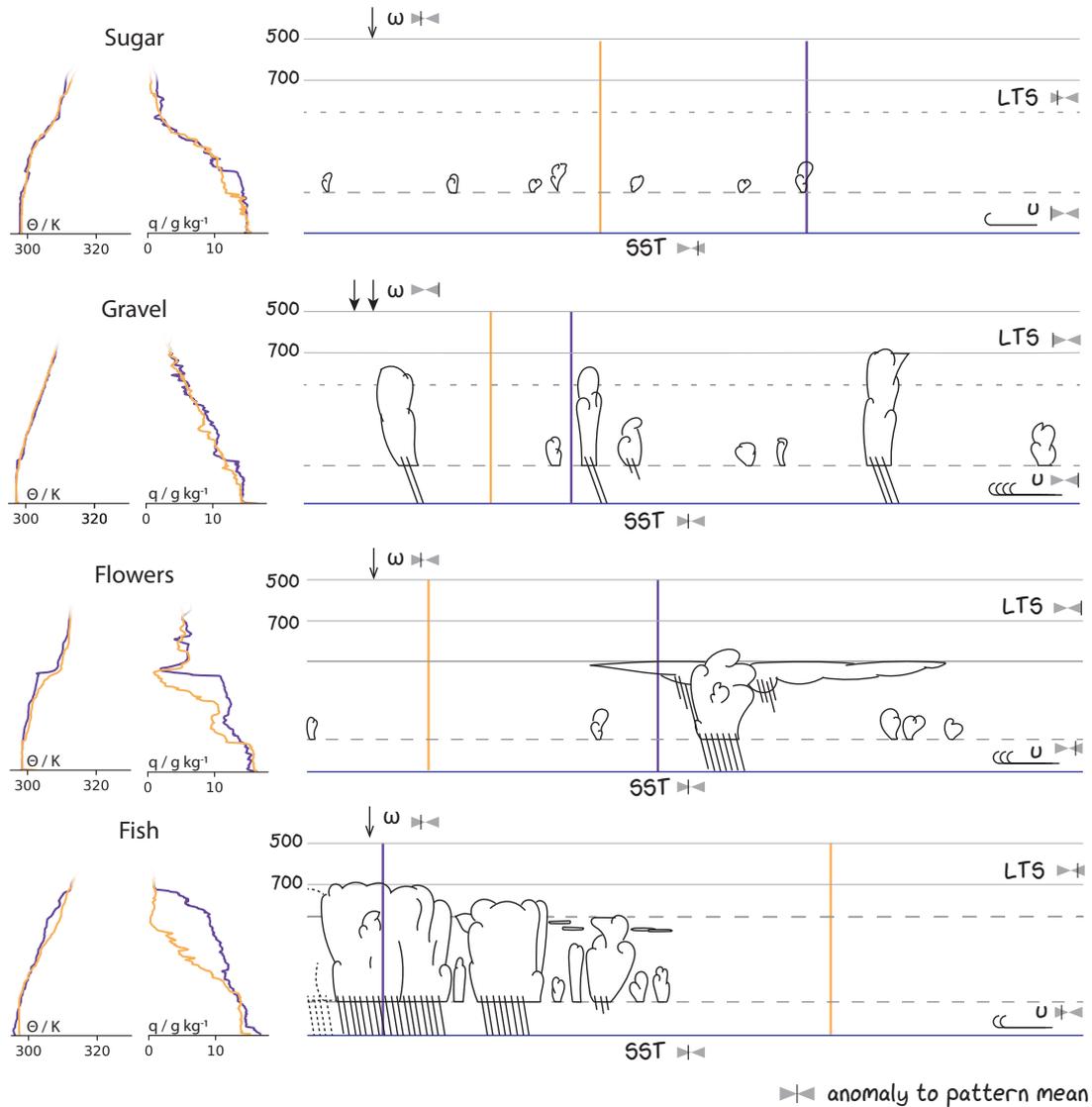


Figure 4.2: Schematic of the four meso-scale cloud patterns with their associated large-scale forcing (right; determined by subsidence ω , horizontal wind u , sea surface temperature SST) and the thermodynamic profiles (left; characterized by specific humidity q and potential temperature Θ). Vertical lines indicate the contrasting positions of the thermodynamic profiles, purple being in the moist part and orange in the dry area. Profiles origin from EUREC⁴A soundings (study F)

The large-scale forcing has been studied in both studies C and B. Despite the different approaches, study B relying on ECMWF Reanalysis Interim (ERA-Interim) output and C primarily on in-situ measurements, wind speed has been identified to separate best the different patterns. The correlation of other large-scale factors peaks only for individual patterns. *Flowers* are observed during high lower tropospheric stability (LTS), *Gravel*

occurs during strong subsidence, *Sugar* is identified when surface temperatures are higher than average and *Fish* occurs during cases of strong convergence.

4.2.1 Evolution of air masses

Understanding the large-scale environment and its evolution plays an important part in understanding the processes leading to the cloud patterns but even more so in estimating how the occurrence of patterns might change in a warming climate depending on changing large-scale environments.

The evolution of air masses and their origin are addressed in study C by calculating the back-trajectories for all patterns identified in the downstream North Atlantic trades. The trajectories show that on average, the air masses have different geographical origins. Despite the identification of *Fish* and *Sugar* in the downwind trades, their air masses origin outside the trades. Only *Gravel* and *Flowers* are actually patterns that are native to the trades, while the others are caused by intruding air masses. *Fish* originates from dissipating cold fronts that intrude as shear lines into the trades from the north and give the *Fish* pattern its characteristic latitudinal band-like structure. In contrast, *Sugar* air masses have a southern origin, which is consistent with the earlier finding that *Sugar* preferably occurs in the vicinity of the ITCZ.

4.2.2 Diurnal cycle

Besides the advection of air masses, the diurnal cycle modulates parts of the large-scale forcing contributing to a detectable diurnal cycle in the pattern frequency and cloudiness as highlighted by Fig. 4.3. The change in pattern-cloud fraction - the characteristic cloud fraction of a particular pattern over the course of a day - only plays a minor role.

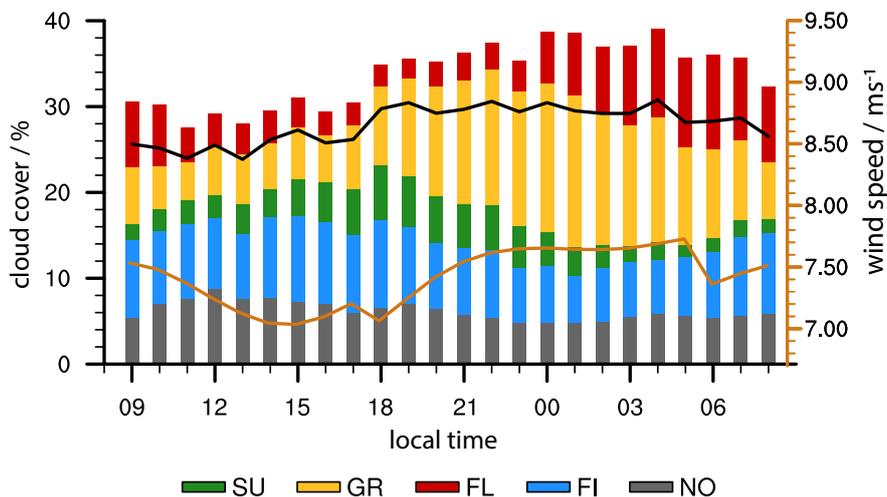


Figure 4.3: Daily cycle of total cloud cover measured at the BCO (black line) and its breakdown by the four patterns and an additional "NO" clear pattern category (colored bars). Stacked bars exceed the total cloud cover because overlaps of NN-classifications are counted towards each classification.

It becomes visible again that *Fish* is a constantly forced pattern as it shows only a very muted diurnal cycle. *Sugar*, *Gravel* and *Flowers* are much more distributed in time. Their peak in cloud cover contribution shifts from *Sugar* around sunset, to *Gravel* at midnight and *Flowers* at dawn.

The strong relationship of patterns with wind speed as seen in the previous section holds also on the hourly timescale. During comparably calm conditions *Sugar* peaks in its relative contribution to total cloud cover. When the wind increases after sunset, so does the occurrence of *Gravel*. The *Flowers* peak when the wind drops again.

This small excerpt of study D underpins which significance the sub-daily and therefore the meso-scale in its temporal sense has. Hence, processes on this scale need to be understood and represented in simulations along with those on the cloud- and large-scale.

These relations between the air mass origin, the large-scale forcing and the patterns let the patterns with the larger cloud amounts, *Flowers* and *Fish*, become less likely to occur. The widening of the tropics (Seidel et al., 2008) and the poleward shift of the extra-tropical storm tracks (Ulbrich et al., 2008; Yin, 2005) in a warming climate would reduce the conditions favoured by them and consequently reduce the average cloud radiative effect observed in the downwind trades.

TOWARDS A PROCESS UNDERSTANDING WITH LARGE-EDDY SIMULATIONS

The previous sections have shown how distinct the patterns on the meso-scale are besides their visual impression. They have also shown how manifold these differences are and that distinctions are observed across several scales, namely the large-scale, the meso-scale and the cloud-scale.

The precise processes that lead to the different patterns are not yet well understood. A field campaign that is helping to reveal and study these processes is the EUREC⁴A field campaign because of its sampling strategy to measure the atmosphere across different scales and because it was conducted in the region where the meso-scale patterns have been observed for the first time: the downwind North Atlantic trades close to Barbados (Stevens et al., 2021). However, even such a dense network of measurement platforms can profit from additional simulations that can fill gaps and as such can help to gain an understanding of the processes that lead to the different cloud formations.

Study E investigates if Large Eddy Simulations (LESs) are able to capture the meso-scale variability and hence can serve as a tool to study the physical mechanisms behind them. LESs are currently the best compromise in terms of domain size and the resolution for a problem of meso-scale extent. With a horizontal grid-spacing of only a few hundred meters and a domain size of over $(500 \text{ km})^2$ they resolve most of the scales that have shown to relate with the meso-scale patterns of shallow convection.

Whether simulations consequently replicate these patterns and are able to mimic the observed variability is evaluated by means of the quantification done in the earlier sections. For this purpose, study E conducted one continuous LES run with a grid-spacing of 624 m (ICON-624m) and two additional nests with grid-spacings of 312 m (ICON-312m) and 156 m (ICON-156m) covering 37 days in January-February 2020. ICON-312m is however only run for a short amount of time as a proof of concept for future investigations.

A general indicator of whether the patterns are represented or not is the visual inspection of the spatial distribution of cloud patches, just like done in the earlier cloud classifications. In this case, the cloud scenes are however not captured by an actual satellite sensor but are simulated from the model output by a satellite forward operator.

This forward operation is shown for the infrared in Fig. 5.1. It depicts the synthetic satellite images for the dates of the canonical pattern examples of Fig. 1.1. Indeed, the visualization demonstrates that most of the scenes match the principle structure of the patterns. This is further confirmed by the classifications of the neural network that has been run for comparison on both the actual satellite images and the synthetic ones of the simulations for the studied period.

It is however noticeable that the simulation tends to produce more *Sugar*- and *Gravel*-like cloud patches at the cost of *Fish* and especially *Flowers*. In Fig. 5.1 *Flowers* are hardly distinguishable from the *Sugar* scene independent of the LES's resolution (see Fig. E.14 for comparison).

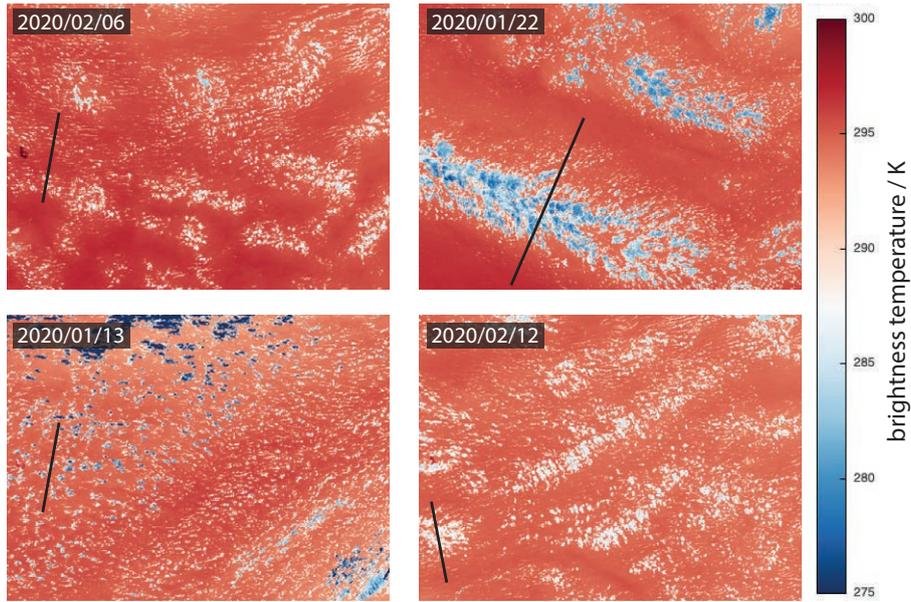


Figure 5.1: Overview of simulated satellite images of ICON625 matching the cloud scenes shown in Fig. 1.1. Different to Fig. 1.1 the infrared channel of ABI is shown. Cross-sections along the black line are shown in Fig. 5.3.

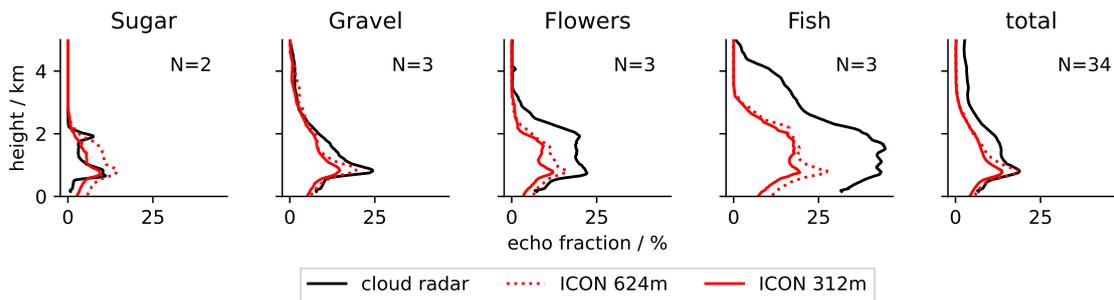


Figure 5.2: Cloud fraction variability grouped by meso-scale pattern identified in study G.

Looking at the vertical distribution of cloudiness, differences between *Sugar* and *Flowers* do become visible mainly due to different amounts of cloudiness aloft the lifting condensation level. Compared to the observations, the variability of the vertical cloud distribution is however substantially reduced throughout the cloud layer for all patterns. Especially the variability of cloudiness above the lifting condensation level is not captured (Fig. 5.2). The *Flowers* with their characteristic stratiform layers are therefore less well represented.

The large deviations in the echo fraction profile shown for the *Fish* are caused by two superimposed factors. First, the *Fish* is not sustained in the simulations for the three days it is continuously seen in the observations. While the pattern is passing over the site in both cases as seen in the supplemental animation ([10.5281/zenodo.5553825](https://zenodo.org/record/5553825)), the *Fish* in the simulations get suppressed and eventually gives rise to *Sugar* and *Gravel* which have much smaller cloud fractions. The second reason for the disagreement is the difference in the vertical extend. The *Fish* pattern is much shallower in the simulations and is capped at about 3 km while in the observations cloudiness is present up to 5 km.

In agreement with the simulated suppression, the inversion strength is much stronger in the simulations hindering the convection to grow deeper. As a consequence less precipitation is formed.

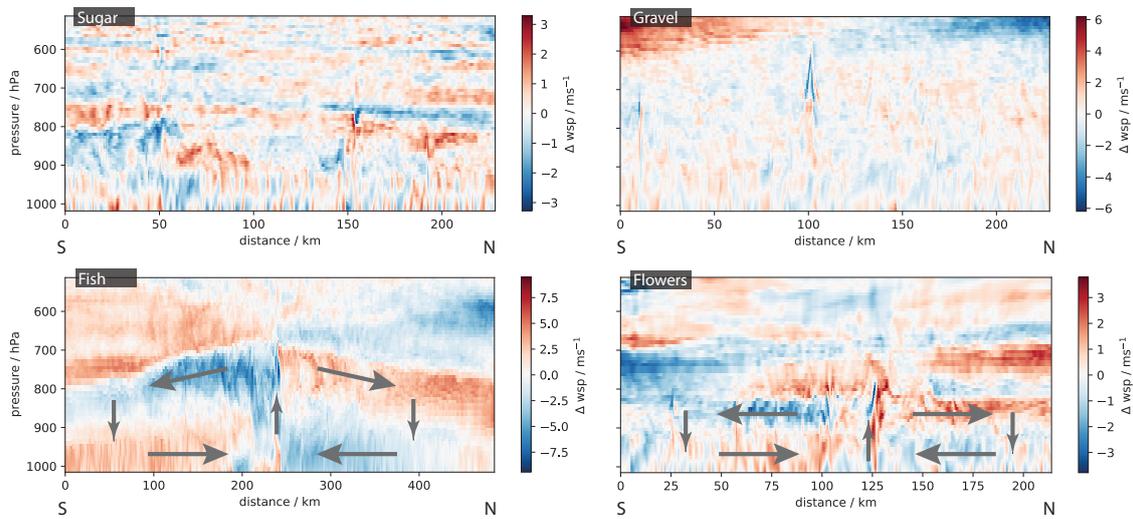


Figure 5.3: Boundary-layer circulations inferred from anomalous wind speeds along the cross-sections indicated in Fig. 5.1. The cross-sections are illustrated from south to north. Arrows indicate circulations. Note the different scales of the colorbar.

Despite these differences, the main processes leading to the different patterns seem to be captured. Fig. 5.3 shows that along the cross-sections marked in Fig. 5.1 a meso-scale circulation is present in *Fish* and *Flowers*. This strengthens the hypothesis that analog to the organization of deep convection a shallow circulation is also evident in patterns of shallow convection. This circulation transports moisture upgradient and increasingly suppresses convection in the expanding dry regions (Bretherton C. S. and Blossey P. N., 2018). It should be noted that these circulations are evident without the correct representation of the stratiform cloudiness. The cloud-top cooling is therefore an unlikely driver of the circulation. The efficient cooling in the dry regions seems more likely to play a role and is the subject of a future study.

Sugar and *Gravel* do not show a circulation on the same scale. This is expected because *Sugar* does not show any signs of self-organization and *Gravel* has much finer granularity than *Fish* and *Flowers*. However, *Gravel*, which visually reminds of colliding cold pools, frequently reveals those in the simulation emphasizing their probable importance.

All in all, it can be concluded that the LESs are a promising tool to investigate the physical processes leading to meso-scale patterns of shallow convection in further detail.

CONCLUSION

For the first time characteristic meso-scale patterns of shallow convection of the downwind trades are systematically studied. This dissertation bases its investigation on the recently observed and defined patterns by Stevens et al. (2020). Because these patterns were solely defined by their visual impression and their usefulness was unknown, this work starts with automatizing the classification of these patterns and continues with subsequent studies on their importance and physical characteristics. Finally, it is evaluated whether large-eddy simulations can be used to fill in observations and gain insights into the processes at play.

The key findings to the overarching questions posed in the introduction are summarized as follows:

0. **Can subjectively defined cloud patterns be efficiently detected?**

The trained deep neural network is able to condense the noisy labels made by humans to a common perception and can successfully be applied to satellite images in a fraction of the time a human would need. The temporal and spatial coverage of the analysis could therefore be extended.

1. **Are the four meso-scale patterns of shallow convection important for the climate system?**

The patterns of shallow convection observed in the downwind North Atlantic winter trades are generic to all global oceanic trade-wind regimes and can reflect more than 40% of the observed variability in cloudiness depending on the location and season. In particular cases, a single pattern can occur up to 50% of the time in a season. Their widespread occurrence paired with their different CRE makes them an important phenomenon that is able to influence the earth's energy budget.

2. **Are the patterns besides their visual appearance also physically distinct?**

The most dominant physical characteristic that separates the patterns from each other is their stratiform cloud amount. The differences observed in CRE can be to first-order attributed to their variability. Geometric cloud properties of cloud genes were independent of the meso-scale context and did not change across patterns with the exception of cloud size that is increased for *Flowers* and *Fish*.

3. **How do the patterns relate to cloud-scale and large-scale forcing?**

Precipitation may play a role in triggering new convection through cold-pools in all meso-scale contexts except for *Sugar*. Changes in the large-scale forcing are playing a major role on the meso-scale both in terms of daily variations, but also on sub-daily timescales. Especially the average wind speed shows the highest correlation with patterns across analyses. Other factors stand out only for particular patterns: divergence in case of *Fish*, subsidence for *Gravel*, surface temperature for *Sugar* and LTS in case of *Flowers*. With the knowledge gained above and the anticipated widening of the tropics (Seidel et al., 2008) and poleward shift of the extra-tropical

storm tracks in a warming climate (Ulbrich et al., 2008; Yin, 2005), the patterns with higher cloud fractions will occur less often with reduced cloud radiative effects.

4. **Can large-eddy simulations help to gain a process understanding of the pattern morphology?**

Large-eddy simulations are presented as a valuable tool to fill in observations and to study the processes that lead to the different meso-scale cloud patterns. However, even at horizontal grid-spacings of about 300 m the meso-scale structure is not matched in all cases. In particular the *Flowers* and their associated variability in stratiform cloudiness are hardly captured. Nevertheless, the distribution of moisture and the presence of meso-scale circulations hint that most of the processes that lead to the patterns are represented. In case of *Flowers* a higher vertical resolution in particular at the inversion strength is hypothesized to result in the characteristic stratiform layers. All in all, the LESs present themselves as a promising additional tool to gain further process understanding of the meso-scale patterning of shallow convection and hence may help to narrow down the tropical low-level cloud feedback to a warming climate.

The answers to these questions wouldn't have been possible without the efforts made in collecting the data. While most of the data is described in the respective data section of the studies and is left out here for brevity, two studies should explicitly be mentioned here because the comprehensiveness of their data collection and processing justified individual manuscripts. First, the radiosonde data used in study C partly originates from the soundings launched during EUREC⁴A. Its measurement strategy, launch procedures and quality control are elaborated in study F. Second, pattern classifications covering the EUREC⁴A period have been in high demand by the scientific community to investigate the manifold observations in the meso-scale context. The manual classifications made during an online event are described in study G and used in E.

OUTLOOK

This unifying essay illustrates the whole process of detecting meso-scale cloud patterns in the trades, identifying their importance, characterizing them and their air masses and finally point out the potential processes that might form and maintain them. The peculiarity of the four classes is their ability to break down nature's complexity into separate - physically distinct, as shown in this work - categories, giving us the chance to focus on the processes individually. Detailed studies can now investigate further processes based on these classifications, for example Vogel et al. (2021) focus on the characterization of cold pools that seem to play an important role in organizing at least one of the patterns, Gravel, as shown by the conducted simulations.

This study focuses on the current climate. The large-scale forcing and air mass trajectories might however change in a warming climate leading to different distributions of these patterns. Changes might also occur on the meso-scale itself.

Because the simulations are able to reproduce the large-scale forcing as shown in this dissertation, the model might be useful to study how the large-scale forcing changes in response to a doubling of CO₂, but also how changes in the occurrence of meso-scale patterns might influence the larger scales. Fermepin and Bony (2014) show for example that a change in tropical low clouds CRE affects the atmospheric overturning circulation and also reinforces precipitation and surface wind speed, all factors that influence the meso-scale patterns.

The meso-scale patterns themselves might also react to a warming and change their characteristics. At least for the patterns *Sugar*, *Gravel* and *Fish* this might be tested by rerunning the simulations of study E with an increased SST and analyse the intra-pattern effects. Will the spacing between single cloud patches change because the strength of the meso-scale circulations alters? Vogel et al. (2020) report that with increased sea surface temperature (SST) the cloudiness will deepen, precipitate more and result in a reduction of the inversion strength. Consequently the pattern type could be altered locally without changes to the large-scale circulation.

The question of how these patterns relate to each other and whether they are partly a continuation of the upstream studied forms of meso-scale organization, mainly open and closed mesoscale cellular convection, is currently investigated by combining the different datasets and classification methods.

This particular dataset and especially the manual classifications that made the development of the neural networks possible can also spark new ideas and foster research in other fields like computer vision. How can subjective classes without a ground-truth be best utilized? An online hosted competition on the platform kaggle (https://www.kaggle.com/c/understanding_cloud_organization) to find the most precise neural network raised large interest in the neural network community. With over 1500 teams participating, it shows the large potential that inter-disciplinary research offers. The winning solutions were all based on ensembles of neural networks which

made the use for the studies presented unpractical due to the larger computational efforts.

Future will tell, if *Sugar*, *Gravel*, *Flowers* and *Fish* will become similarly known as the cloud classifications by Howard (1803) that are now known beyond the scientific community. As shown here, they have their physical justification. In addition, they eased the communication during the EUREC⁴A field campaign tremendously by allowing to describe very complex cloud systems by a single word.

*Any intelligent fool can make things bigger, more complex, and more violent.
It takes a touch of genius — and a lot of courage to move in the opposite direction.*

— Ernst F. Schumacher (Schumacher, 1973)

Part II

APPENDICES

COMBINING CROWDSOURCING AND DEEP LEARNING TO EXPLORE THE MESOSCALE ORGANIZATION OF SHALLOW CONVECTION

The work in this appendix has been published as:

Rasp, Stephan, **Schulz, Hauke**, Bony, Sandrine, and Stevens, Bjorn (Nov. 2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. *Bulletin of the American Meteorological Society* 101.11, E1980–E1995. doi: [10.1175/BAMS-D-19-0324.1](https://doi.org/10.1175/BAMS-D-19-0324.1)

The contributions of the authors to this publication are as follows:

SR and HS designed the data collection experiment and analysed the data. SB and BS supervised the study and made the data collection an institute wide event. SR drafted the manuscript with contributions of all co-authors. All authors proof-read several iterations and accepted the final manuscript.

Combining crowd-sourcing and deep learning to explore the meso-scale organization of shallow convection

Stephan Rasp¹, Hauke Schulz², Sandrine Bony³, Bjorn Stevens²

¹Department of Informatics, Technical University of Munich

²Max Planck Institute for Meteorology

³Sorbonne Université, LMD/IPSL, CNRS

ABSTRACT

Humans excel at detecting interesting patterns in images, for example those taken from satellites. This kind of anecdotal evidence can lead to the discovery of new phenomena. However, it is often difficult to gather enough data of subjective features for significant analysis. This paper presents an example of how two tools that have recently become accessible to a wide range of researchers, crowd-sourcing and deep learning, can be combined to explore satellite imagery at scale. In particular, the focus is on the organization of shallow cumulus convection in the trade wind regions. Shallow clouds play a large role in the Earth's radiation balance yet are poorly represented in climate models. For this project four subjective patterns of organization were defined: Sugar, Flower, Fish and Gravel. On cloud labeling days at two institutes, 67 scientists screened 10,000 satellite images on a crowd-sourcing platform and classified almost 50,000 mesoscale cloud clusters. This dataset is then used as a training dataset for deep learning algorithms that make it possible to automate the pattern detection and create global climatologies of the four patterns. Analysis of the geographical distribution and large-scale environmental conditions indicates that the four patterns have some overlap with established modes of organization, such as open and closed cellular convection, but also differ in important ways. The results and dataset from this project suggests promising research questions. Further, this study illustrates that crowd-sourcing and deep learning complement each other well for the exploration of image datasets. (Capsule Summary) Crowd-sourcing and deep learning are combined to explore the meso-scale organization of shallow clouds in the subtropics.

A.1 INTRODUCTION

A quick glance at an image, be it taken from a satellite or produced from model output, is often sufficient for a scientist to identify features of interest. Similarly arranged features across many images form the basis for identifying patterns. This human ability to identify patterns holds true also in situations where the features, let alone the patterns they build, are difficult to describe objectively—a situation which frustrates the development of explicit and objective methods of pattern identification. In these situations, machine learning techniques, particularly deep learning (see Sidebar 1), have demonstrated their ability to mimic the human capacity for identifying patterns, also from satellite cloud imagery (e.g., Wood and Hartmann, 2006). However, the application and assessment of such techniques is often limited by the tedious task of obtaining sufficient training

data, so much so that (in cloud studies at least) these approaches have not been widely adopted, let alone assessed.

Recently, Stevens et al., 2020 described a collective cloud classification activity by a team of 13 scientists supported by the International Space Science Institute (ISSI). This ISSI team aimed to identify mesoscale cloud patterns in visible satellite imagery taken over a trade-wind region east of Barbados. Organization, or clustering, of clouds has been shown to have important implications for climate in the case of deep convection (Tobin et al., 2012), which raises the question to what extent this is the case in shallow clouds. The ISSI-team's hand-labeling effort resulted in around 900 subjectively classified images. An initial application of machine learning to these images (by the first author) proved promising but also highlighted the need for more training data in order to obtain robust and interpretable results.

Based on these first insights the authors organized a crowd-sourced project (see Sidebar 2) that would allow us to collect a substantially larger set of labeled images. This activity was designed to provide a better foundation for the application of machine learning to the classification of patterns of shallow clouds, as well as to explore methodological questions raised when attempting to marry crowd-sourcing with machine learning to address problems in climate and atmospheric science. Specifically we sought to answer four questions:

- Q1 How should a community-driven labeling exercise be set up to ensure a) a good user experience for participants and b) the usefulness of the gathered data for subsequent analysis?
- Q2 Can a diverse set of scientist identify the subjective modes of cloud organization established by the ISSI team with satisfactory agreement to warrant further scientific analysis?
- Q3 Can a deep learning algorithm learn to classify images as well as trained scientists?
- Q4 To the extent that a machine can be trained to classify large numbers of images, what can be learned from applying this algorithm to global data?

In this paper, we present our findings. They suggest that, for suitable problems, the combination of crowd-sourcing and deep learning allow scientists to analyze data on a scale beyond what would be possible with traditional methods. Though our main findings will be of particular interest to researchers interested in the mesoscale organization of shallow clouds, the methods used to obtain them may be of more general interest, and are presented with this in mind.

We begin by describing how the cloud patterns (or classes) we sought to classify were defined, followed by a summary of the crowd-sourcing project. Then the results from the human data is presented before we explain how deep learning is used to extend the analysis. Finally, we summarize our findings as pertains to the above stated research questions, from which inferences of potential relevance to future studies are drawn.

A.2 SUGAR, FLOWERS, FISH AND GRAVEL

Mesoscale patterning of shallow cumulus is a common feature in satellite imagery. However, organization on these scales is largely ignored in modeling studies of clouds

and climate. This applies to process studies with large-eddy simulations e.g., Bretherton, 2015; Rieck et al., 2012 as well as general circulation models, be it in traditional or super-parameterizations (Arakawa and Schubert, 1974; Parishani et al., 2018).

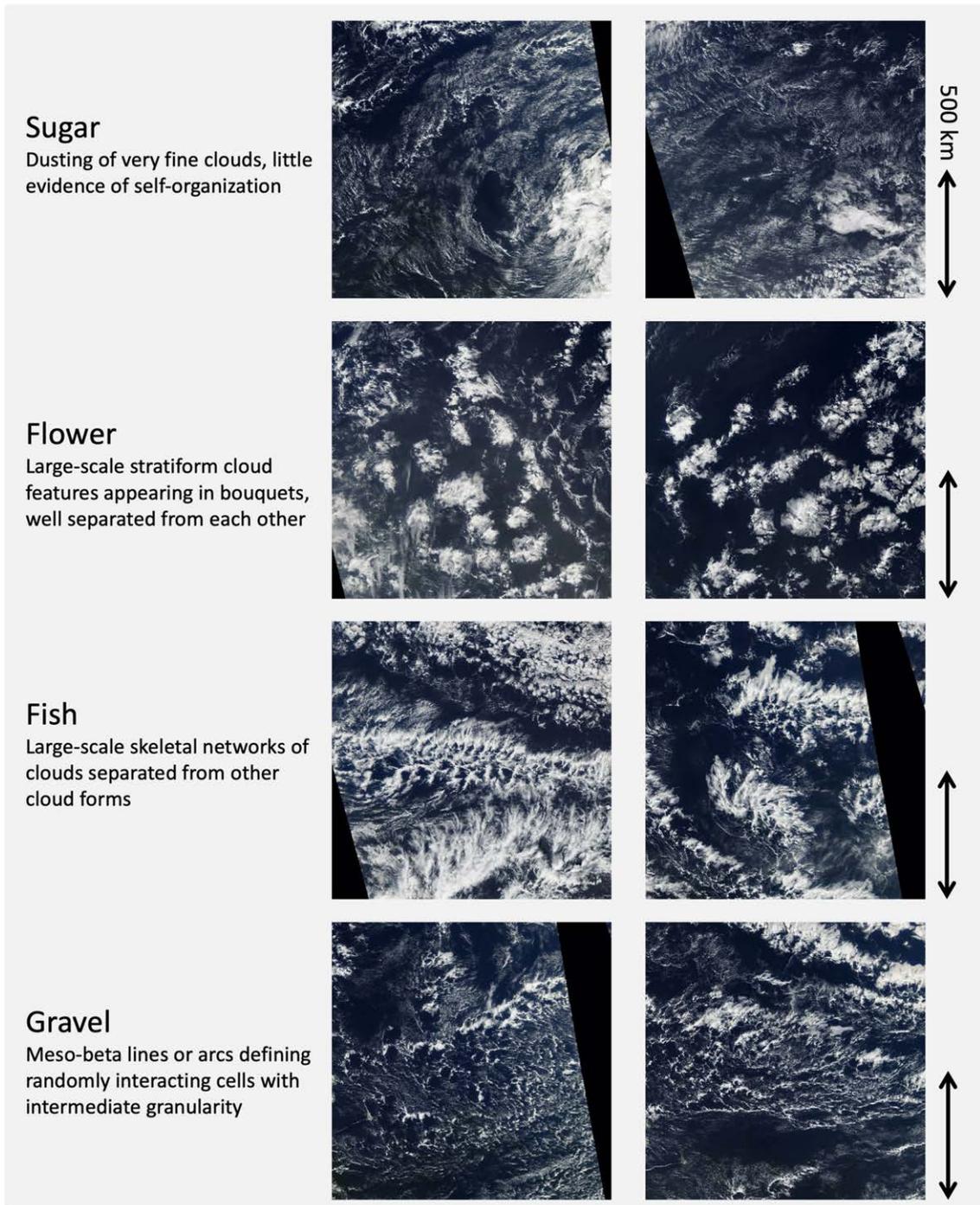


Figure A.1: Canonical examples of the four cloud organization patterns as selected by the ISSI team.

The prevalence of mesoscale patterning in satellite cloud imagery led the ISSI team (Stevens et al., 2020) to identify four cloud patterns that frequent the lower trades of the North Atlantic. They named these patterns Sugar, Flower, Fish and Gravel (Fig. A.1). The

choice of new and evocative names was motivated by the judgement that the patterns were different than those that have been previously described, for instance in studies of stratocumulus or cold-air outbreaks. Support for this judgement is provided by an application of the neural network from Wood and Hartmann, 2006 and Muhlbauer et al., 2014, which was trained to distinguish between “No Mesoscale Cellular Convection (MCC)”, “Closed MCC”, “Open MCC” and “Cellular, but disorganized”. When applied to the scenes classified by the ISSI team the algorithm mostly resulted in the “disorganized” classification (Personal communication with I. L. McCoy). Despite the lack of a simple link between the patterns classified by the ISSI team and patterns previously described in the literature, below we point out previously identified patterns that may be related to the four patterns used here.

“Sugar” describes wide-spread areas of very fine cumulus clouds. Overall these fields are not very reflective, do not have large pockets of cloud-free regions and, ideally, exhibit little evidence of meso-scale organization. Often, though, they are embedded within the larger-scale flow which gives them some structure. In strong flow, Sugar can form thin “veins”, or feathers, which have previously described as dendritic clouds (Nicholls and Young, 2007).

“Flowers” are areas with isotropic cloud structures, each ranging from 50 to 200 km in diameter, with similarly wide cloud-free regions in-between. This pattern overlaps to some degree with canonical closed-cell MCC. Flowers, however, are often less densely packed than typical closed cells, which only have narrow cloud-free regions at the edges, and they are identified well outside of regions where stratocumulus are found (Norris, 1998). One hypothesis is that they are successors of more closely packed closed-cell MCC which are in the process of breaking up.

“Fish” are elongated, skeletal structures that sometimes span up to 1,000 km, mostly longitudinally. As noted by Stevens et al., 2020, these features appear similar to what Garay et al., 2004 called actinoform clouds. They presented examples of these particularly well structured cloud forms taken from all ocean basins, near but typically downwind of regions where stratocumulus maximize. To the extent Fish are variants of the actinoform clouds found by Garay et al., they may be more common than previously thought.

Finally, “Gravel” describes fields of granular features marked by arcs or rings. The typical scale of these arcs is around 20 km. We suspect that these patterns are driven by cold pools caused by raining cumulus clouds (Rauber et al., 2007). In this regard, Gravel is fundamentally different from open-cell MCC, which has larger cells that are driven by overturning circulations in the boundary layer. However, the line between these two mechanisms can blur at times.

It is also interesting to compare our subjectively chosen labels to those of Denby, 2020 who used an unsupervised learning algorithm to automatically detect different types of cloud organization (their Fig. 2). Some of their patterns bear resemblance to our classes, e.g. “Sugar” seems to most closely correspond to their patterns A and B, “Gravel” to G and H. However, their automatically detected classes appear less striking to the human eye.

A.3 CROWD-SOURCED LABELS

To obtain a large pool of labeled images from the community, an accessible user interface is needed. Zooniverse¹ is an open web platform that enables researchers to organize and present research questions in ways that enable contributions from the broader public (see also Sidebar 2). Its flexibility in serving and presenting images, choosing between different labeling tasks and its ability to monitor and organize the information associated with the labeling activities made Zooniverse very well suited for our task.

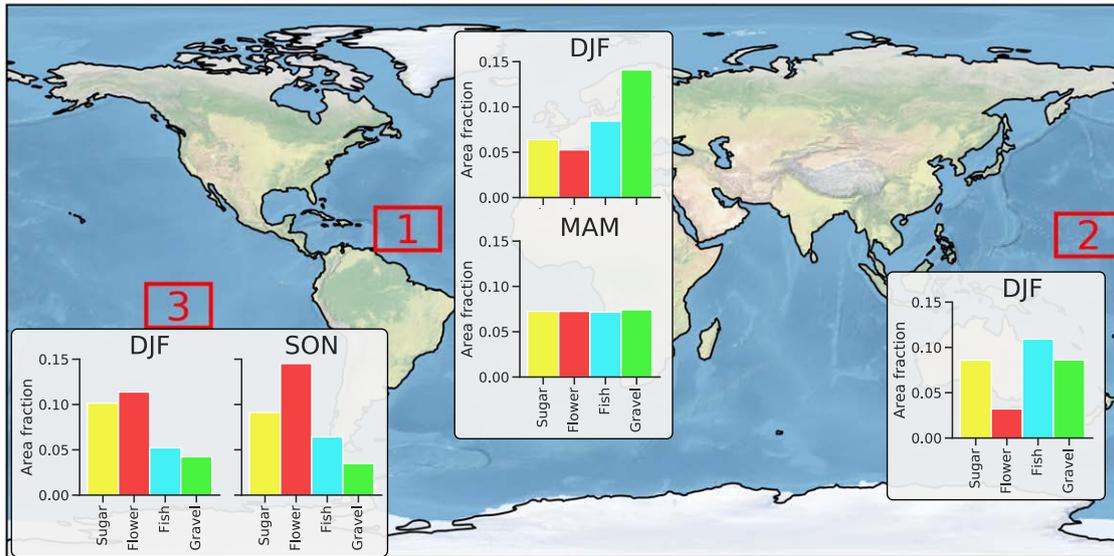


Figure A.2: World map showing the three regions selected for the Zooniverse project. Bar charts are showing which fraction of the image area was classified into one of the four regions by the human labelers. Note that the areas do not add up to one. The remaining fraction was not classified.

For our project we downloaded roughly 10 000 $14^\circ \times 21^\circ$ (lat-lon) Terra and Aqua MODIS visible images from NASA Worldview. To select the regions and seasons, we started with the Boreal winter (DJF) east of Barbados as a reference. Barbados is home to the Barbados Cloud Observatory (Stevens et al., 2016). The clouds in its vicinity were not only the focus of the ISSI-teams study, but have more generally come to serve as a laboratory for studies of shallow clouds and climate (Bony et al., 2017; Medeiros and Nuijens, 2016; Stevens et al., 2020, 2016). To obtain more images and sample a greater diversity of clouds we subsequently added images from two further regions in the Pacific, which were chosen based on their climatological similarity to the original study region upwind of Barbados (Fig. A.2; see Supplement for details). Images were downloaded for an eleven year period from 2007 to 2017.

Stevens et al., 2020 speculated that their protocol of assigning a single label to the entire $10^\circ \times 10^\circ$ image resulted in considerable ambiguity and disagreement between labelers. In an attempt to minimize this issue we presented participants with slightly larger images and experimented with ways to allow the labeling of multiple, and possibly overlapping sub-regions. This was accomplished by allowing users to draw rectangles

¹ <https://www.zooniverse.org/projects/raspstephan/sugar-flower-fish-or-gravel>

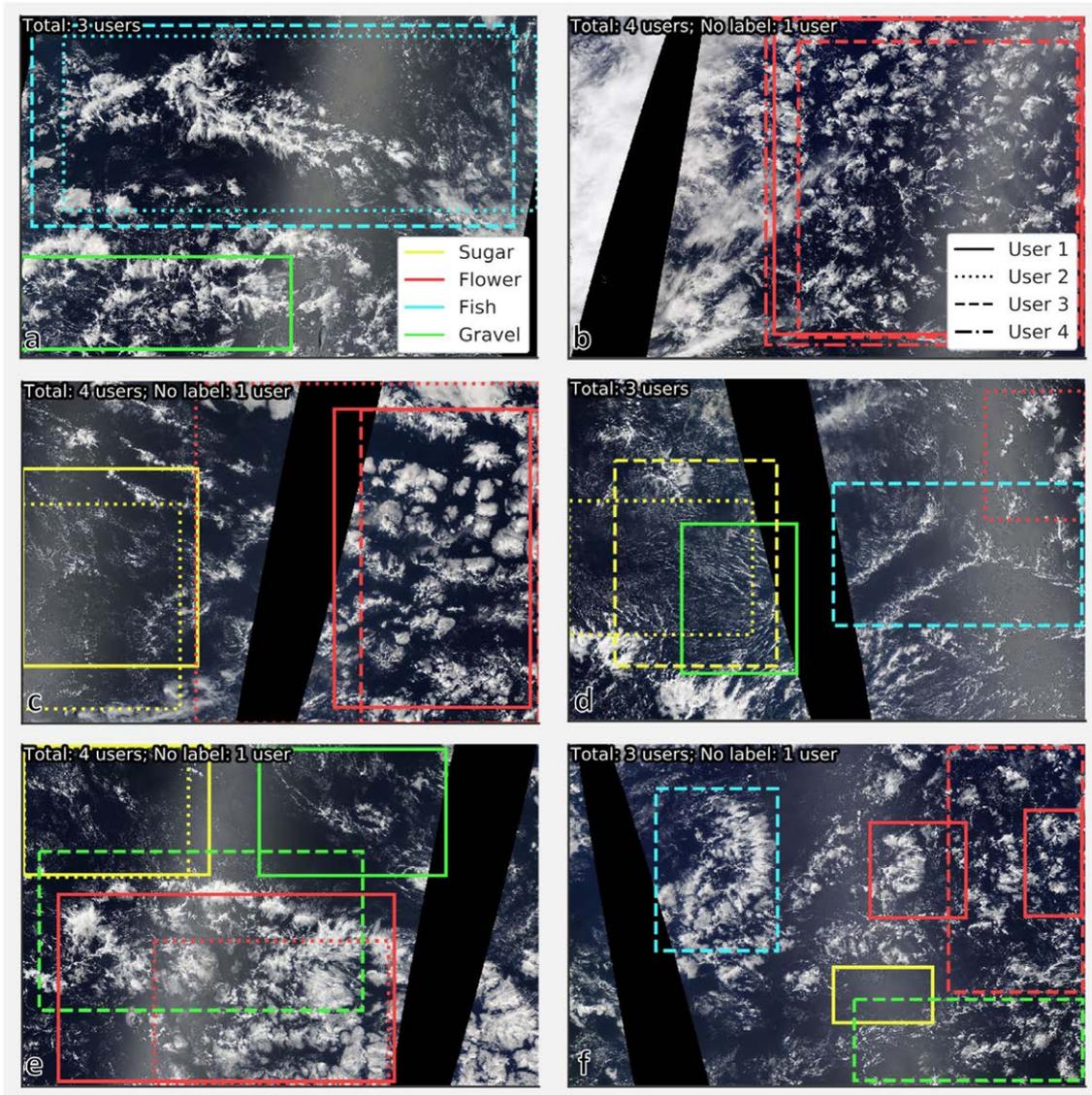


Figure A.3: Six example images showing annotations drawn by human labelers. Different line styles correspond to different users. In addition the IoU values for each image and class are shown in the table.

around regions where they judged one of the four cloud patterns to dominate (see Fig. A.3 for examples). Participants had the possibility to draw any number of boxes, including none, with the caveat that the box would cover at least 10% of the image. We arrived at this setup after experimenting with other options, such as labeling subsections of a predefined grid, or allowing users to label regions that they defined using polygons with an arbitrary number of sides. We opted for the rectangles to increase labeling speed and improve the user experience. Our thinking was that it would be better to have less accurate but more plentiful data, and that given the vague boundaries of the cloud structures, it was anyway doubtful that a more accurate labeling tool would add much information. As we will show later, this thinking paid off for the machine learning models we trained.

The Zooniverse interface was further configured to serve participants an image randomly drawn from our library of 10 000 images. After being classified by four different users, images were retired, i.e. removed from the image library. In addition, no user was shown the same image twice. With the interface in place, cloud classification days were set up at the Max Planck Institute for Meteorology in Hamburg, Germany on Nov 2nd and at the Laboratoire de Météorologie Dynamique in Paris, France on Nov 29th, 2018. After a brief instruction at the start of the day and a warm-up on the training dataset, 67 participants, most of them researchers, from the two institutes, labeled images for an entire day. The activity yielded roughly 30 000 classified images, i.e., each image was classified about three times on average. Because an image could have sub-regions with different classifications, the number of labels was somewhat larger with 49 000. On average, participants needed around 30 s to classify one image, amounting to approximately 250 h of concentrated human labor. There was however considerable differences among users, as the interquartile range in classification times ranged from 20 s to 38 s. Overall, the four patterns occupied similarly large areas but notable differences occurred depending on the geographic region and season (Fig. A.2).

A.4 INFERENCES FROM HUMAN LABELS

Given the subjective nature of labels assigned by visual inspection, our first research question was to what extent the human labelers agreed with each other. In the initial classification exercise of 900 images reported in Stevens et al., 2019 a majority of scientists agreed on one pattern in 37 % of the cases, significantly more than random. In this project, in addition to choosing the category of the clouds, participants also had to choose the location. To explore the agreement we started by looking at many examples, six of which are reproduced in Fig. A.3. Many more can be found at Rasp, 2020b. The most notable conclusions from this visual inspection are that users agreed to a high degree on features that closely resemble the canonical examples of the four classes but also that there was a lot of disagreement otherwise. Take Fig. A.3a where two out of three participant agreed on the presence of Fish in the top half of the image; or Fig. A.3b where three out of four participants recognized a region of Flowers. On the other hand Fig. A.3d shows an example of an image with plenty of ambiguity. Also note that users applied different methodologies when labelling, some labeling a single large region, other many small regions. Overall, we came to the conclusions that, while certainly noisy, clear examples of what was defined as Sugar, Flower, Fish and Gravel could be robustly detected.

Next, we aimed to quantify the agreement. To our knowledge there is no standard way of evaluating subjective labels from multiple users. The most commonly used metric for comparing a label prediction with a ground truth is the Intersect over Union (IoU) score, also called the Jaccard index. Given two sets, A and B , it is defined as the ratio of their intersection to their union, i.e., $I = A \cap B$ divided by $U = A \cup B$. An IoU score of one indicates perfect overlap, while zero indicates no overlap. We adapted the IoU score to this task by first iterating over every image and then computing the intersect I_{image} and union U_{image} for every user-user combination for this image. To compute the final "Mean IoU between humans" we computed the sum of the intersect and union over all images: $I = \sum_{\text{image}} I_{\text{image}}$ and $U = \sum_{\text{image}} U_{\text{image}}$. This was done for each cloud class separately. We also computed an IoU score for the "Not classified" area. Finally, the "All

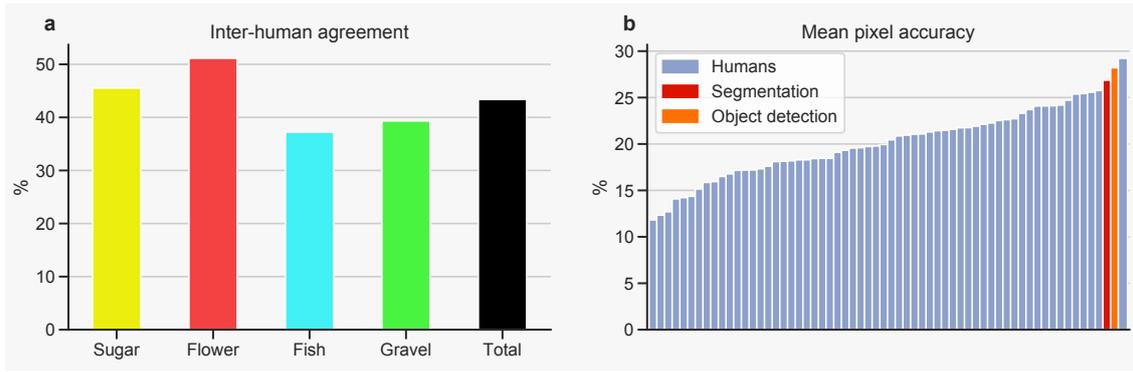


Figure A.4: (a) Mean IoU between humans. Dashed line represents random IoU; see text for details. (b) Mean IoU for each human participant and the two deep learning algorithms for a validation dataset.

classes" IoU was computed by additionally taking the sum of I and U over all classes. The results for the inter-human mean IoU are shown in Fig. A.4a.

At first glance, IoU values of around 0.2 seem low. In classical computer vision tasks such values would certainly indicate low agreement. However, as mentioned above, this dataset is different from classical object detection tasks in that there are more than two labelers for most images and there are many cases in which one or more participants did not label an image. In fact, the primary reason for the low mean IoU score are zero values, which arise from some users detecting a feature while others did not. Take Fig. A.3b as an example. Here, three of four users agreed to a high degree of accuracy on the location of Flowers but the last user did not submit a label. This results in three "no label"- "label" comparisons and three "label"- "label" comparisons. Even with perfect agreement between the three Flower labelers, the mean IoU would only be 0.5. In reality it is 0.44 for this example. These "no label"- "label" pairs with $\text{IoU} = 0$ make up 63% of all user-user comparisons (see Supplement Fig. S2). Omitting these gives a mean IoU of 0.43. To get a feel for what this value means consider the two Sugar rectangles in Fig A.3d, which have an IoU of 0.46. The table at the bottom of Fig. A.3 shows the mean IoU values for each of the example images. These numbers suggest that even for images where one would visually detect a high degree of agreement between the users, the IoU numbers are quite low. For this reason, the actual values should not be compared to other tasks where the IoU is used. Rather, for this paper they simply serve the purpose of comparing different classes and methods. To further illustrate this point we computed the IoU score for many randomly drawn labels from the same number and size distribution as the human labels, which gives an IoU of only 0.04. What the numbers do show is that there are noticeable differences between the four patters. People agreed most on Flowers while Fish proved more controversial. With regards to Q2, we came to the conclusion that, despite the noise in the labels, there was sufficient consensus between the participants on clear features to warrant further analysis, especially since as we will see the noise will largely disappear in the statistical average.

Another question that the new methodology of labeling allows us to answer is whether or not the patterns tend to span larger or smaller areas. Based on the Zooniverse labels, Flower boxes tended to be largest, covering around 25% (around 900,000 km²) of the image. Fish and Gravel were somewhat smaller with a box size of around 20% (around

720,000 km²). Sugar spanned regions smaller yet with boxes only taking up 15% (around 540,000 km²) of the image on average. Because the initial classification by Stevens et al., 2020 required labelers to identify the entire scene, the relative infrequency with which they detected Sugar is likely due in part to the infrequency with which it covers large areas.

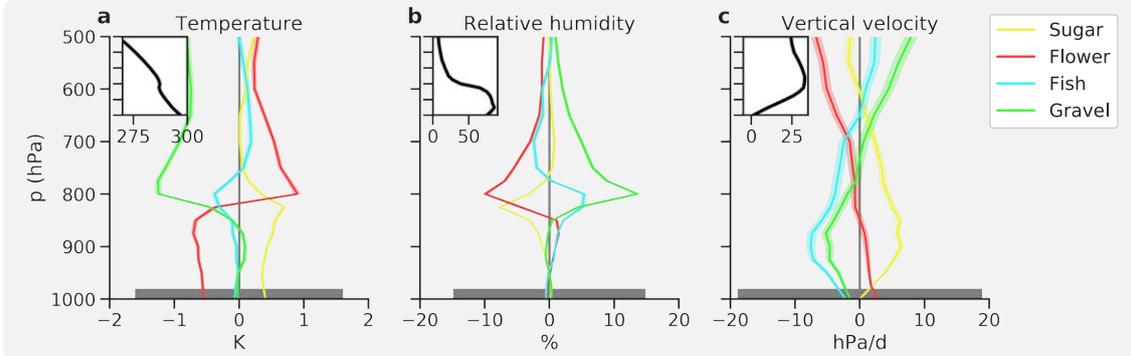


Figure A.5: Median of large-scale environmental conditions corresponding to the four patterns as identified by the human labelers. Figures show deviations of (a) temperature, (b) specific humidity and (c) vertical velocity (shown in $\omega = dp/dt$) relative to the climatological mean which is shown in the figure insets. The shading about the lines shows the standard error, and hence the statistical difference between the mean conditions associated with any particular pattern. The bar at the base of the figure shows the average inter-quartile spread (for the level where this spread maximizes, around 800 hPa) in the thermodynamic state associated with each pattern, indicating that the conditions associated with any given pattern can vary considerably.

Further we can ask whether the four patterns, which were purely chosen based on their visual appearance on satellite imagery, actually correspond to physically meaningful cloud regimes. To investigate this, we created composites of the large-scale conditions from ERA-Interim reanalyses² corresponding to each pattern (Fig. A.5). To the extent the ERA-Interim accurately represents the meteorological conditions in the region, the composites suggest that Sugar, Flower, Fish and Gravel appear in climatologically distinct environments. This is supported by the standard error being smaller than the difference between the patterns. The standard error is a measure for how well the mean conditions of a given pattern can be estimated and is defined as σ/\sqrt{N} , where σ is the standard deviation and N is the sample size. At the same time, there is variability between individual profiles within a composite, as shown by the inter-quartile range. Hence, while the compositing suggests that the occurrence of a particular pattern is associated with significant changes in the large-scale environmental conditions, this is clearly not the only factor at play, and things like airmass history are likely also important.

Flowers tend to be associated with a relatively dry and cold boundary layer with a very strong inversion (note that Fig. A.5 shows deviations from the climatological mean). Sugar on the other hand appears in warm and humid boundary layers with strong downward motion maximizing near the cloud base. For Fish and particularly Gravel, on the other hand, the inversion and downward motion is rather weak. The fact that Flowers and Gravel are essentially opposites in terms of their environmental profiles

² <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim>

suggests that they are not simply manifestations of closed and open-cellular convection, which often transition smoothly into one another in similar large-scale environments (Muhlbauer et al., 2014).

A.5 APPLICATION OF DEEP LEARNING

While the 10 000 images labeled on Zooniverse already provide a useful dataset for further analysis, they only cover a small fraction of the globe for a small fraction of time. Only 0.6% of the data available during the selected eleven year period were labeled. In this section, we explore whether deep learning (see Sidebar 1) can help to automate the detection of the four organization patterns and if so what can be learned from it.

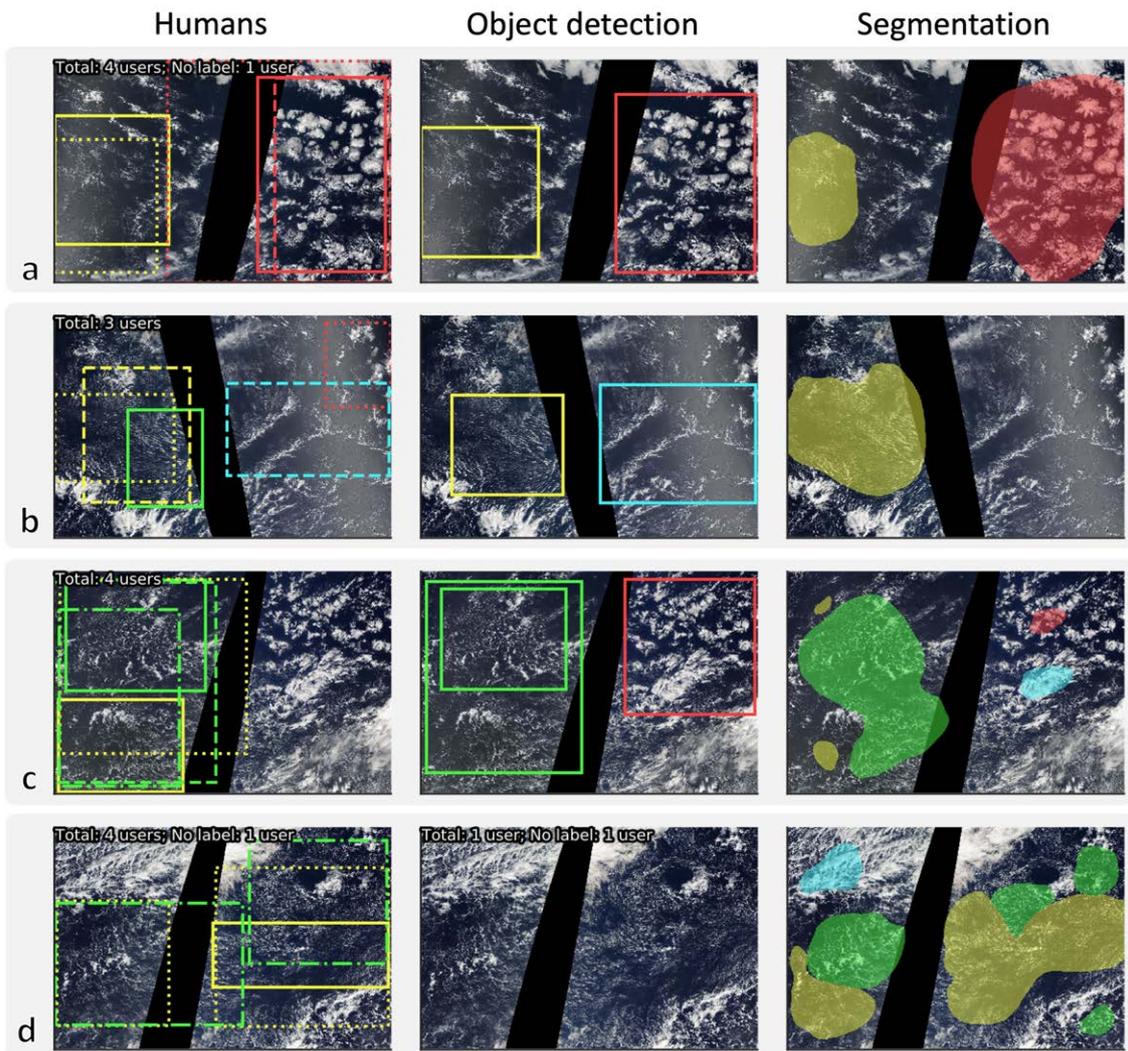


Figure A.6: Human and machine learning predictions for four images from the validation set. Note that images a) and b) are also shown in Fig. A.3.

The pattern recognition task can be framed as one of two machine learning problems: object detection and semantic segmentation. Object detection algorithms draw boxes around features of interest, essentially mirroring what the human labelers were doing.

In contrast, segmentation algorithms classify every pixel of the image. Fig. A.6 shows examples of these two approaches for images from a validation dataset that was not used during training (see Rasp, 2020a for more, randomly chosen examples). Details about the neural network architectures and preprocessing steps can be found in the Supplement. Both types of algorithm accurately detect the most obvious patterns in the image and agree well with human labels. Neither algorithm is perfect, however. The object detection algorithm sometimes misses features, as is visible in Fig. A.6d. The segmentation algorithm, on the other hand tends to produce relatively small patches (Fig. A.6c and d) because, other than humans and the object detection algorithm, in which the range of possible box sizes is an adjustable parameter, it has not been given instructions to only label larger patches. An interesting and advantageous feature of the segmentation algorithm is that, despite all training labels being rectangular, it appears to focus on the actual, underlying shape of the patterns, as visible by the rounded outlines of the predicted shapes. This suggests that despite the uncertainty in the human dataset, the deep learning algorithms are able to filter out a significant portion of this noise and manage to distill the underlying human consensus.

To quantitatively compare the deep learning algorithms against the human labelers, we compute the mean IoU for each human individually as well as for the two algorithms (Fig. A.4b). Both algorithms show a large agreement with the human labels for a random validation dataset. The fact that the scores are higher than the mean inter-human IoU directly reflects the fact that the algorithms tend to produce less noisy predictions. Further analysis shows that the algorithms inherit some biases from the human training labels. The frequency and accuracy of the predicted labels is higher for patterns with a higher inter-human agreement, most notably flowers (Supplemental Fig. 3), which could slightly bias the deep learning predictions.

The main advantage of deep learning algorithms is that they are very fast at inference, one second per image compared to the 30 seconds a human needed on average, and they are more scalable. This allows us to apply the algorithm to the entire globe (Fig. A.7a; see Supplement for details). A healthy skepticism is warranted when applying machine learning algorithms outside of their training regime (Rasp et al., 2018; Scher and Messori, 2019). A visual inspection of the global maps (see Rasp, 2020c for more examples), however, suggests that the algorithm's predictions are reasonable and physically interpretable as discussed below. Naturally, over land the predictions have to be assessed with greater care because no land was present in the training dataset. Nevertheless, Fig. A.7a suggests that the algorithm even appears to correctly identify shallow cumuli over the tropical landmasses as sugar.

To obtain global climatologies of Sugar, Flower, Fish and Gravel we ran the algorithm on daily global images for the entire year of 2017 (Fig. A.7b–e). The resulting heatmaps reveal coherent hotspots for the four cloud patterns. The spatial distribution of these hotspots helps answer some further questions raised by the ISSI team's study. For instance, the heat maps indicate that organization is most common over the ocean. Only Sugar – the one pattern characterized by its lack of mesoscale organization – was identified over land (but keeping in mind the potential bias of the algorithm). Our results also indicate that Sugar, followed by Flower, are the most common forms of organization globally. This indicates a bias arising from the ISSI team's focus on a single study region, as large areas of Sugar are relatively rare near Barbados. A prevalence of Sugar in the

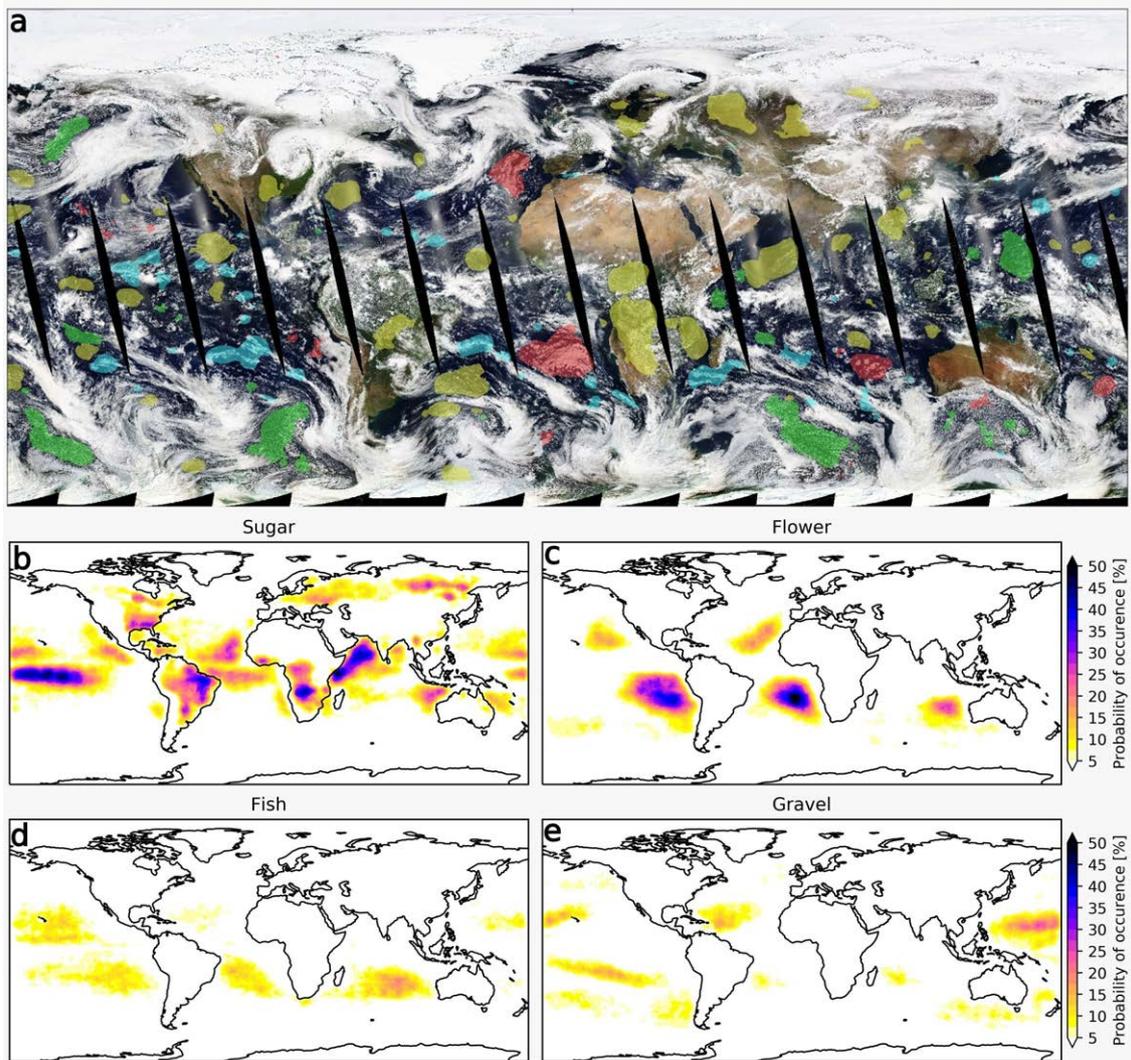


Figure A.7: (a) Global predictions of the image segmentation algorithm for May 1 2017. The colors are the same as in the previous figures. See Rasp, 2020c for more examples. (b–e) Heatmaps of the four patterns for the year 2017.

trades adjacent to the deep tropics, and regions such as the Arabian sea, is consistent with its coincidence in association with strong low-level subsidence and a somewhat drier cloud layer (as seen by the large-scale composites, Fig. A.5) indicating that it might be most favored in regions where convection is suppressed by strong subsidence from neighboring regions of active convection, or strong-land sea circulations.

Flowers prevail slightly downstream of the main stratocumulus regions. Composites of the environmental conditions in which they form show them to be, on average, associated with large scale environmental conditions characterized by more pronounced lower tropospheric stability, and a somewhat drier free troposphere (Fig. A.5). This lends credence to the idea that they are manifestations of closed-cell MCCs. Whereas the climatology of closed-cell MCC by Muhlbauer et al., 2014, their Fig. 5 shows similar hotspots in the subtropics, it also has strong maxima across the mid and high-latitude oceans. The absence of such hotspots in our classification of Flowers could indicate a bias

of our algorithm towards the regions it was trained on. However, it could also suggest that “Flowers” differ from typical closed-cellular convection in their scale and spacing.

Further downstream in the trade regions, Flowers make way to Gravel and Fish. These two patterns are more geographically intertwined, which is in agreement with the similarity of the environmental profiles in Fig. A.6. Interestingly, Gravel seems to be relatively confined to the Barbados region, the west of Hawaii, and the southern tropical Pacific near regions – like the South Pacific Convergence Zone – of climatological convergence (Fig. A.2). Hence the prevalence of Gravel in the more limited classification activity of Stevens et al., 2020, is not representative of the trade-wind regions more broadly. There is also some coincidence of Gravel hot-spots with regions of open-cell MCC regions as highlighted by the classification by Muhlbauer et al., 2014, specifically around Hawaii and in the Southern sub-tropical Atlantic, but as with Flowers their MCC algorithm picks up many more open cells in higher latitudes. This, again, suggests that there may be a fundamental difference between the classes, something we already suspected based on their physical driving mechanisms, i.e. cold pools versus boundary layer circulations. Fish, appears linked to stronger synoptic upward motion (Fig. A.5c), which the image snapshot from 1 May 2017 (Fig. A.6a) suggests is associated with synoptic convergence lines, often connected to trailing mid-latitude fronts.

Globally, the patterns are coherent, with hot-spots for a given pattern appearing in a few spatially extensive and plausibly similar meteorological regimes. This coherence supports the hypothesis that the subjective patterns are associated with meaningful and distinct physical processes. Though the combination of crowd-sourced labels and deep learning helped answering many of the questions raised at the outset of this study it also raises some new ones, for instance whether important cloud regimes are missing from our classification. Unsupervised classification algorithms like the one deployed by Denby, 2020 can be a good starting point to explore this question.

A.6 THE FOUR QUESTIONS

In this paper, we described a project to combine crowd-sourcing, to detect and label four subjectively defined patterns of mesoscale shallow cloud organization from satellite images, with deep learning. The design and execution of the project raised a number of questions, four of which have been highlighted in this paper, and the answers to which we present as follows.

The first question (Q1) was concerned with how best to configure a crowd-sourcing activity. We found that speed and ease of use for the participants is paramount. Drawing crude rectangles on the screen only took tens of seconds for each image, whereas more detailed shapes such as polygons would have taken significantly longer. Further, the quickness of drawing boxes on an image meant that less of an attention span was required from the participants. (Some even reported to have had fun.) For our task, which involves judgements with inherent uncertainty, the added noise introduced by crude labels turned out to be insignificant in the statistical average, as shown by the “consensus” found by the deep learning algorithm. Based on our experience, quantity trumps quality. This might, of course, be different for tasks where object boundaries are more clearly defined.

Our second question, Q₂, asked whether sufficient agreement exists between the human labelers to warrant scientific use of the labels. We believe that this is indeed the case. As discussed in the section titled "Inferences from human labels" there is a significant amount of disagreement between the participants, particularly because many cloud formations did not fit one of the four classes exactly. However, more importantly there was significant agreement on patterns that closely matched the canonical examples of "Sugar", "Flower", "Fish" and "Gravel". Taking a statistical average – training a deep learning model can be viewed as doing just that – removes some of the ambiguity from the labels and crystallizes the human consensus. Of course, the four classes chosen are not a complete description of all modes of organization, and others could have been defined. But the fact that the results are compatible with physical understanding suggest that the four classes do indeed capture important modes of cloud organization in the sub-tropics.

Q₃ asked whether deep learning can be used to build an automated labeling system. The answer is a resounding yes. Both deep learning algorithms used in this paper, show high agreement scores. Further, visual analysis of the deep learning predictions suggest that these are less noisy than the human predictions. In other words, the deep learning models have learned to disregard the noise of the human labels and instead extract the common underlying pattern behind points of agreement, i.e., the essence of the proposed patterns. In addition, the deep learning models are both many orders of magnitude faster than humans at labeling images, and less costly and difficult to maintain.

The application of deep learning enabled us to classify a significantly larger geographical and temporal set of data. This allowed us to look at global patterns of "Sugar", "Flower", "Fish" and "Gravel" thereby addressing our fourth research question (Q₄). Here our main finding is that heat-maps of pattern occurrence are distributed in a geographically coherent way across all the major ocean basins, and sample significantly different meteorological conditions. Heat maps for two of the patterns (Flower and Gravel) show some overlap with closed-cell mesoscale cellular convection, but only over portions of the sub-tropics. As a rule the regions where patterns are identified (particularly for Fish and Flower) are not in regions familiar from past work on cloud classification.

A.7 INFERENCES AND OUTLOOK

The coherence of the heat maps for individual patterns suggests the presence of physical drivers underpinning their occurrence; drivers that may change as the climate changes. Using the same classification categories but a different way of classifying the images, Bony et al., 2020 showed that differences in cloud radiative properties are associated with different forms of organization. Our study thus lends weight to the idea that quantifying the radiative effects of shallow convection, and potential changes with warming, may require an understanding of, or at least ability to represent, the processes responsible for the mesoscale organization of fields of shallow clouds. This might seem to be a daunting task. However, if the occurrence of different modes of organization can be reliably linked to large-scale conditions, reanalysis data or historical climate model simulations could help reconstruct cloud fields. This could offer clues as to how meso-scale organization, and hence cloudiness, has changed in the past, and may change in the future.

This example helps highlight how crowd-sourced and deep-learned data-sets create new ways to study factors influencing shallow clouds and their radiative properties, and hopefully stimulates ideas for adapting the approach to other problems. The growing accessibility of these new research methodologies makes their application all the more attractive. Platforms like Zooniverse make it easy to set up a labeling interface free of charge. Plus, even if we did not do so, it is also possible to make the interface open to the public. Deep learning has also become much more accessible. Easy-to-use Python libraries³ with pre-trained models for many applications in computer vision as well as accessible online courses⁴ make it possible even for non-computer scientists to apply state-of-the-art deep learning techniques.

Our study also illustrates how crowd-sourcing and deep-learning effectively complement one another, also for problems in climate science. Deep learning algorithms typically need thousands of samples for training. These are not readily available for most problems in the geosciences. A key lesson from our project is that, even for the ambiguously defined images that characterize many problems in atmospheric and climate science, it is feasible to create sufficient training data with a moderate amount of effort. We found that 5000 labels (i.e. a 6th of what was collected here) were enough to obtain similarly good results to the ones shown here. This translates to a day of labeling for around 15 people.

This means that combining crowd-sourcing and deep learning is a promising approach for many questions in atmospheric science where features are easily – albeit not unambiguously – detectable by eye but hard to quantify using traditional algorithms. In our case, the combination of the two tools allowed us to generate global heatmaps, something that would have been impossible with traditional methods. Potential examples of similarly suited problems in the geosciences are detecting atmospheric rivers and tropical cyclones in satellite and model output⁵, classifying ice and snow particles images obtained from cloud probe imagery, or even large-scale weather regimes.

DATA AND CODE AVAILABILITY. All data and code are available at <https://github.com/raspstephan/sugar-flower-fish-or-gravel>.

A.8 ACKNOWLEDGMENTS

First and foremost, we would like to thank all the participants of the cloud labeling days. Special thanks go to Ann-Kristin Naumann and Julia Windmiller for initiating this collaboration and to Katherine Fodor for suggesting Zooniverse. SR acknowledges funding from the German Research Foundation Project SFB/TRR 165 “Waves to Weather”. This paper arises from the activity of an International Space Science Institute (ISSI) International Team researching “The Role of Shallow Circulations in Organising Convection and Cloudiness in the Tropics”. Additional support was provided by the European Research Council (ERC) project EUREC4A (Grant Agreement 694768) of the European

³ Keras (Chollet and Others, 2015) and fastai (<https://docs.fast.ai/>) were used for this study, see Supplement.

⁴ <https://course.fast.ai/>, <https://deeplearning.ai>

⁵ See <https://www.nersc.gov/research-and-development/data-analytics/big-data-center/climatenet/> for a similar project.

Union's Horizon 2020 Research and Innovation Programme and by the Max Planck Society. We acknowledge the use of imagery from NASA Worldview, part of the NASA Earth Observing System Data and Information System (EOSDIS).

A.9 SIDEBAR 1: CROWD-SOURCING

Crowd-sourcing describes projects where a task is collaboratively solved by a group of people. This can be a small research group or a large group of internet users. One of the first examples of crowd-sourcing in the natural sciences is Galaxy Zoo⁶, a project that has citizen scientists classify different galaxy types and has produced 60 peer-reviewed publications so far. An early meteorological example focused on estimating hurricane intensity (Hennon et al., 2015). Recent climate projects on the crowd-sourcing platform Zooniverse^{7 8} asked volunteers to transcribe old, hand-written weather records. Thanks to the collaboration of many individuals such projects produce a wealth of data that would be unattainable for a single scientist. Note that for this paper we understand the term crowd-sourcing to indicate active labor by the participants rather than providing data through personal sensors or cameras. For a broader review of citizen science and crowd-sourcing studies in the geosciences, see Zheng et al., 2018.

A.10 SIDEBAR 2: DEEP LEARNING FOR VISION TASKS IN THE GEOSCIENCES

Deep learning describes a branch of artificial intelligence based on multi-layered artificial neural networks (Nielsen, 2015). In recent years, this data-driven approach has revolutionized the field of computer vision which up to 2012 was to a large extent based on hard-coded feature engineering (LeCun et al., 2015). More specifically, the success of deep learning in vision tasks is based on convolutional neural networks which exploit the translational invariance of natural images (i.e. a dog is a dog whether it is in the top right or bottom left of the image) to greatly reduce the number of unknown parameters to be fitted. Deep neural networks also have many potential applications in the Earth sciences, particularly where already existing deep learning techniques can be transferred to geoscientific problems (Reichstein et al., 2019). A perfect example of this is the detection of features in images, the topic of this study. One obstacle is that deep learning requires a large number, typically several thousands, of hand-labeled training samples. For Earth science problems, these are usually not available. For this reason, previous studies that used deep neural networks to detect atmospheric features relied on training data created by traditional, rule-based algorithms (Hong et al., 2017; Kurth et al., 2018; Liu et al., 2016; Mudigonda et al., 2017; Racah et al., 2016). A notable exception is the aforementioned study by Wood and Hartmann, 2006. They hand-labeled 1000 images of shallow clouds and used a neural network to classify them into four cloud types, making it a predecessor to our study.

6 <https://www.zooniverse.org/projects/zookeeper/galaxy-zoo>

7 <https://www.zooniverse.org/projects/edh/weather-rescue>

8 <https://www.zooniverse.org/projects/dreeweepsouth/southern-weather-discovery>

A.11 SUPPLEMENTAL MATERIAL

A.11.1 Region selection criteria

The regions were selected ahead of the classification days according to a similarity analysis of atmospheric conditions that resemble the conditions encountered during the DJF season east of Barbados where these patterns were first found (Stevens et al., 2020).

Because the mesoscale organization of shallow cumulus is a relatively new research topic, the meteorological conditions influencing it are primarily an educated guess. Lower tropospheric stability (LTS), surface wind speed (FF) and total integrated column water vapour (TCWV) are three parameters one could naively imagine to describe the meteorological setting to a sufficient degree. Starting with the inter-annual seasonal mean of these atmospheric properties at the region east of Barbados, we searched for climatologically similar regions and seasons within a 120° -wide latitudinal belt (60°N to 60°S) around the globe. We used a k-means clustering with eight clusters to find similar patterns within our search perimeter. As input to the algorithms we used the climatological means of LTS, FF10 and TCWV for each of the four seasons. The eight clusters explain more than 90% of the variance in the dataset and provide large enough regions to fit 21° longitude by 14° latitude boxes reasonably well.

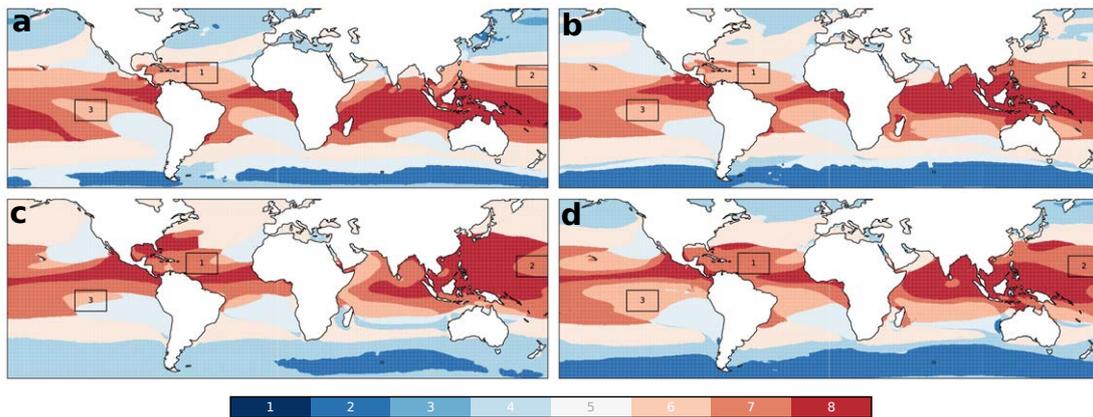


Figure A.8: Cluster analysis of LTS, FF10, TCWV separated by season (DJF, MAM, JJA, SON). The colors identify the 8 clusters as a result of the k-means algorithm. For a better visual impression the clusters are sorted by cluster mean column integrated moisture with cluster 1 being the driest. Black boxes indicate regions chosen for human-classifications.

Fig. A.8 shows the clusters for the four seasons. Our analysis indicates that the meteorological conditions over the Northwestern Atlantic change with season. This is not surprising due to the migration of the ITCZ, but it illustrates that we shouldn't expect to see the same cloud patterns or at least the same distribution throughout the year. The final choice of seasons and regions was made to match the climate of region 1 in DJF (Table A.1)

Table A.1: Selected domains used for human-classification of cloud patterns.

Domain	Bounds	Seasons used
1	-61°E -40°E; 10°N 24°N	DJF, MAM
2	159°E 180°E; 8°N 22°N	DJF
3	-135°E -114°E; -1°N -15°N	DJF, SON

A.11.2 Deep learning models

Two deep learning models are used, one for object detection and one for semantic segmentation. For object detection, an algorithm called Retinanet (Lin et al., 2018) is used. Here we used the following implementation in Keras (Chollet and Others, 2015): <https://github.com/fizyr/keras-retinanet>, which uses a Resnet50 (He et al., 2015) backbone. The original images had a resolution of 2100 by 1400 pixels. For Retinanet the images were downscaled to 1050 by 700 pixels. This is necessary to fit the batch (batch size = 4) into GPU RAM.

For semantic segmentation, we first converted each human classification, i.e. all boxes by one user for an image, to a mask. Sometimes boxes for different patterns overlap. In this case, the mask is chosen to represent the value of the smaller box. Overall, the amount of overlapping boxes is small, however, so that the resulting error is most likely negligible. To create a segmentation model, we used the fastai Python library v1⁹. The network architecture has a U-Net (Ronneberger et al., 2015) structure with a Resnet50 backbone. For the segmentation model the images were downscaled to 700 by 466 pixels (batch size = 6).

To create the prediction masks, first a Gaussian filter with a half-width of 10 pixels was applied to smooth the predicted field. Then, for each pixel the highest probability for each of the four patterns was used, if this probability exceeded 30%. This last step counteracts the tendency to predict background, which is by far the most common class in the training set.

A.11.3 Global heatmaps

To create the heatmaps, the segmentation algorithms was used. Predictions were created for a 21° longitude by 14° latitude region at a time, with a windows sliding in 10.5° and 7° increments over the globe. The highest pattern probability for the overlapping images was then taken to create the global mask. This was necessary because the algorithm tends to predict background at the edges of the image, a consequence of the human labelers not drawing boxes that extend all the way to the edge of the image. The climatology was created from one year of MODIS-AQUA data.

⁹ <https://docs.fast.ai/>

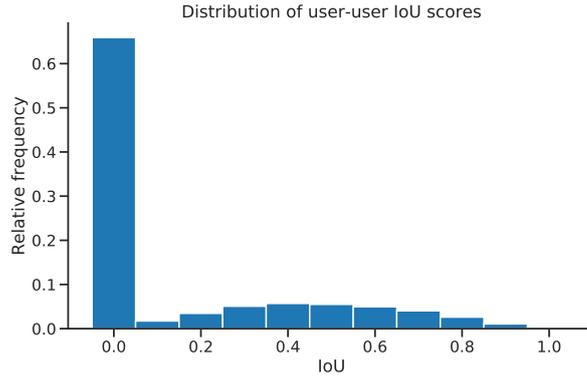


Figure A.9: Histogram of IoU values for each user-user comparison for each class. Zero values mostly indicate cases where one user labeled a feature of a given class and the other user did not.

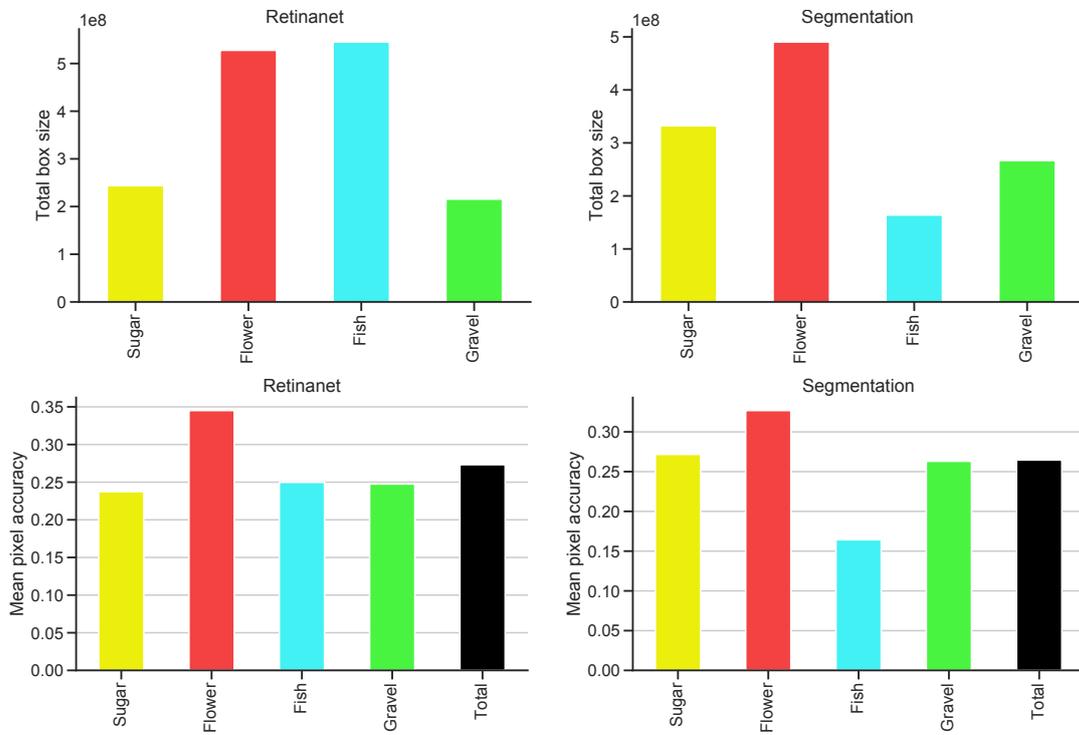


Figure A.10: (Top row) Total size of classifications for the two deep learning algorithms for a random validation dataset. (Bottom row) Mean pixel accuracy (= mean IoU) for the two algorithms stratified by pattern, also for a random validation set.

SUGAR, GRAVEL, FISH AND FLOWERS: DEPENDENCE OF
MESOSCALE PATTERNS OF TRADE-WIND CLOUDS ON
ENVIRONMENTAL CONDITIONS

The work in this appendix has been published as:

Bony, Sandrine, **Schulz, Hauke**, Vial, Jessica, and Stevens, Bjorn (2020). Sugar, Gravel, Fish, and Flowers: Dependence of Mesoscale Patterns of Trade-Wind Clouds on Environmental Conditions. *Geophysical Research Letters* 47.7, e2019GL085988. doi: [10.1029/2019GL085988](https://doi.org/10.1029/2019GL085988)

The contributions of the authors to this publication are as follows:

SB designed the study and prepared the manuscript with contributions of all authors. HS processed the satellite data and build the foundation for the separation of auxiliary data by pattern by calculating the organization metrics. SB conducted with HS most of the analysis with contributions and feedback from JV and BS. All authors contributed in the editing process and provided feedback.

Sugar, Gravel, Fish and Flowers: Dependence of Mesoscale Patterns of Trade-wind Clouds on Environmental Conditions

Sandrine Bony¹, Hauke Schulz², Jessica Vial¹ and Bjorn Stevens²

¹LMD/IPSL, CNRS, Sorbonne University, Paris, France

²Max Planck Institute for Meteorology, Hamburg, Germany

ABSTRACT

Trade-wind clouds exhibit a large diversity of spatial organizations at the mesoscale. Over the tropical western Atlantic, a recent study has visually identified four prominent mesoscale patterns of shallow convection, referred to as Flowers, Fish, Gravel and Sugar. We show that these four patterns can be identified objectively from satellite observations by analyzing the spatial distribution of infrared brightness temperatures. By applying this analysis to 19 years of data, we examine relationships between cloud patterns and large-scale environmental conditions. This investigation reveals that on daily and interannual timescales, the near-surface wind speed and the strength of the lower-tropospheric stability discriminate the occurrence of the different organization patterns. These results, combined with the tight relationship between cloud patterns, low-level cloud amount and cloud-radiative effects, suggest that the mesoscale organization of shallow clouds might change under global warming. The role of shallow convective organization in determining low-cloud feedback should thus be investigated.

PLAIN LANGUAGE SUMMARY

Satellite imagery shows that clouds in the trade-wind regions exhibit a large diversity of patterns. Over the tropical Atlantic close to Barbados, the population of low-level clouds can organize in different ways, adopting patterns evocatively referred to as: 'flowers', 'fish', 'gravel', 'sugar'. This study shows that these different patterns, originally identified subjectively, can be recognized more objectively from space measurements of infrared radiation. It also shows that the relative occurrence of these different patterns relates to the strength of the trade winds near the ocean surface, and to the stability of the lower atmosphere. Finally, it shows that each pattern is associated with a different cloud amount, and thus impacts the radiative cooling of the Earth differently. These results suggest that under global warming, the change in environmental conditions might perturb the frequency of different patterns, which might affect the Earth's radiative response to warming in a way that has not been previously considered.

B.1 INTRODUCTION

Shallow cumuli are ubiquitous over the World Ocean, and therefore their sensitivity to a change in environmental conditions has the potential to greatly influence Earth's radiation balance and climate sensitivity. Actually, the response of trade-wind cumuli to

warming constitutes a primary source of uncertainty in climate model estimates of cloud feedbacks (Bony and Dufresne, 2005; Medeiros et al., 2015; Vial et al., 2013). During the last decade, much progress has been made in understanding the mechanisms through which trade-wind cumuli could respond to warming (Bretherton, 2015; Klein et al., 2017; Rieck et al., 2012; Vogel et al., 2016). However, despite having long recognized that shallow convective clouds are patterned – or organized – on the mesoscale in many different ways (Agee, 1987; Malkus and Riehl, 1964), the role that this organization might play in cloud feedbacks remains largely unexplored (Nuijens and Siebesma, 2019; Vial et al., 2017). Thus it remains an open question as to whether the mesoscale organization of clouds, which is left out of most parameterizations and the many large-eddy simulations over small domains, influences how shallow convective clouds respond to warming.

To help answer this question, here we first explore whether the observed organization of shallow convection can be linked to variability in large-scale environmental conditions, and whether the different patterns of organization imprint themselves differently on the radiation budget. We do so by using observations over a region of the North-Atlantic trades (windward of Barbados) where the shallow clouds are known to be representative of the broader trades (Medeiros and Nuijens, 2016), and where the mesoscale organization of shallow clouds has been well characterized (Stevens et al., 2020). During boreal winter, this region is associated with sea surface temperatures (SSTs) of 26 °C to 28 °C, a moderate large-scale subsidence in the free troposphere (about 25 hPa d⁻¹ to 30 hPa d⁻¹), and a predominance of shallow clouds (Stevens et al., 2016). In this region, the prominent patterns of organization do not correspond to the classical and well-characterized open and closed patterns of mesoscale cellular convection found over colder oceans (McCoy et al., 2017; Wood and Hartmann, 2006). Rather, shallow clouds in the trades appear organized in a less regular fashion, on scales ranging from 20 km to 2000 km. By inspecting ten years of satellite imagery, Stevens et al. (2020) identified four recurrent patterns that they labeled ‘Flowers’, ‘Fish’, ‘Gravel’ and ‘Sugar’. In their classification, *Sugar* consists of a dusting of very fine scale clouds with small vertical extension, *Gravel*, of clouds organized along lines or arcs defining cells with intermediate granularity, sometimes looking like cold pools. *Fish*, were so named due to the appearance of a fishbone-like skeletal network of clouds separated by well defined cloud-free areas, and *Flowers* denoted the presence of larger, seemingly more stratiform, cloud structures in the form of very large but dispersed closed cells. Examples of these four patterns are provided in Figure B.1a.

Considerable daily and interannual variability in the appearance of the patterning (Stevens et al., 2020) offers an opportunity to investigate its co-variability with large-scale meteorological conditions. Even so, the relatively weak variability of the large-scale environment (e.g., 90% of SST variations are weaker than 2 K), requires a record longer than ten winter seasons already classified. To access a longer record, we first attempt to identify the mesoscale patterns using an objective methodology, and use this for exploring the co-variability amongst patterns, their radiative effects, and the large-scale environment in which they form. We do so by first demonstrating (section B.2) that the four prominent patterns of cloud mesoscale organization pointed out by Stevens et al. (2020) project well onto a simple characterization of the spatial variability of infrared brightness temperatures measured from satellite. Then, in section B.3, we investigate relationships between the four cloud patterns and the large-scale environment in which

they form. Finally, a summary of the main findings, and a discussion of their implications for understanding low-cloud feedbacks, are presented in section B.4.

B.2 CLASSIFICATION OF MESOSCALE ORGANIZATION PATTERNS

We follow Stevens et al. (2020) by analyzing shallow clouds over the tropical Atlantic ocean East of Barbados (48W-58W, 10N-20N) during the boreal winter (DJF, 1 Dec. to 28 Feb.), for the period from Dec. 2000 through Feb. 2019. Our analysis is, however, based on 3-hourly infrared (11 μm) brightness temperature (T_b), gridded (0.07°) data from the GridSat-B1 dataset (Knapp et al., 2011). The calibration uncertainty of T_b is less than 0.5 K and its stability better than 0.1 K/decade. To avoid situations obscured by occasional cirrus associated with deep convection over South America or within the ITCZ, we restrict our analysis to images for which the 25th percentile of T_b is higher than 285 K. Pixels for which $280 \text{ K} \leq T_b \leq 290 \text{ K}$ are associated with the presence of marine low-cloud objects. This definition is purposefully conservative to exclude some of the thinnest cloud features that may correspond to evaporating cloud fragments. The threshold of 290 K corresponds to the temperature of cloud tops around 1 km, near the penetration depth of the most buoyant surface parcels (Stull, 1988; Vogel et al., 2019), and somewhat above the hazy layer of cumulus debris near cloud base. Finally, cloudy areas (or cloud objects) are defined through a nearest neighbor segmentation (e.g. Tobin et al. (2012)). Each object is subsequently identified by its centroid, and area.

B.2.1 Organization metrics

The population of cloud objects within the $10^\circ \times 10^\circ$ area is characterized through a few metrics. Those include the total number of cloud objects N within the domain, the total fractional area A of the domain covered by shallow clouds, and a clustering measure, I_{org} , defined by Tompkins and Semie (2017) based on earlier work by Weger et al. (1992). I_{org} compares the distribution of the nearest-neighbor distances among the centroids of objects to that expected for a random distribution of objects. $I_{\text{org}} = 0.5$ corresponds to randomly distributed centroids, while I_{org} values significantly lower than 0.5 correspond to regular distributions, values higher than 0.5 correspond to ‘clustered’ or ‘organized’ distributions (Figure S1). These metrics are calculated for each 3-hourly satellite image, and then daily-mean values are computed.

The visual inspection of the day-to-day variability of cloud organizations suggests that at first order, the diversity of patterns can be characterized by only two metrics: the mean object size, $S = \frac{A}{N} \times 10^4$, which distinguishes patterns associated with a predominance of small or large cloud objects, and I_{org} . Over the period 2000-2019, S and I_{org} exhibit a large variability with fairly continuous distributions (Figure B.1b). By selecting situations that fall in the upper or lower terciles of both the S and I_{org} distributions, we define four classes, or quadrants, that we refer to as A, B, C, and D, and which, as we show below, match well with the four cloud patterns identified by Stevens et al. (2020).

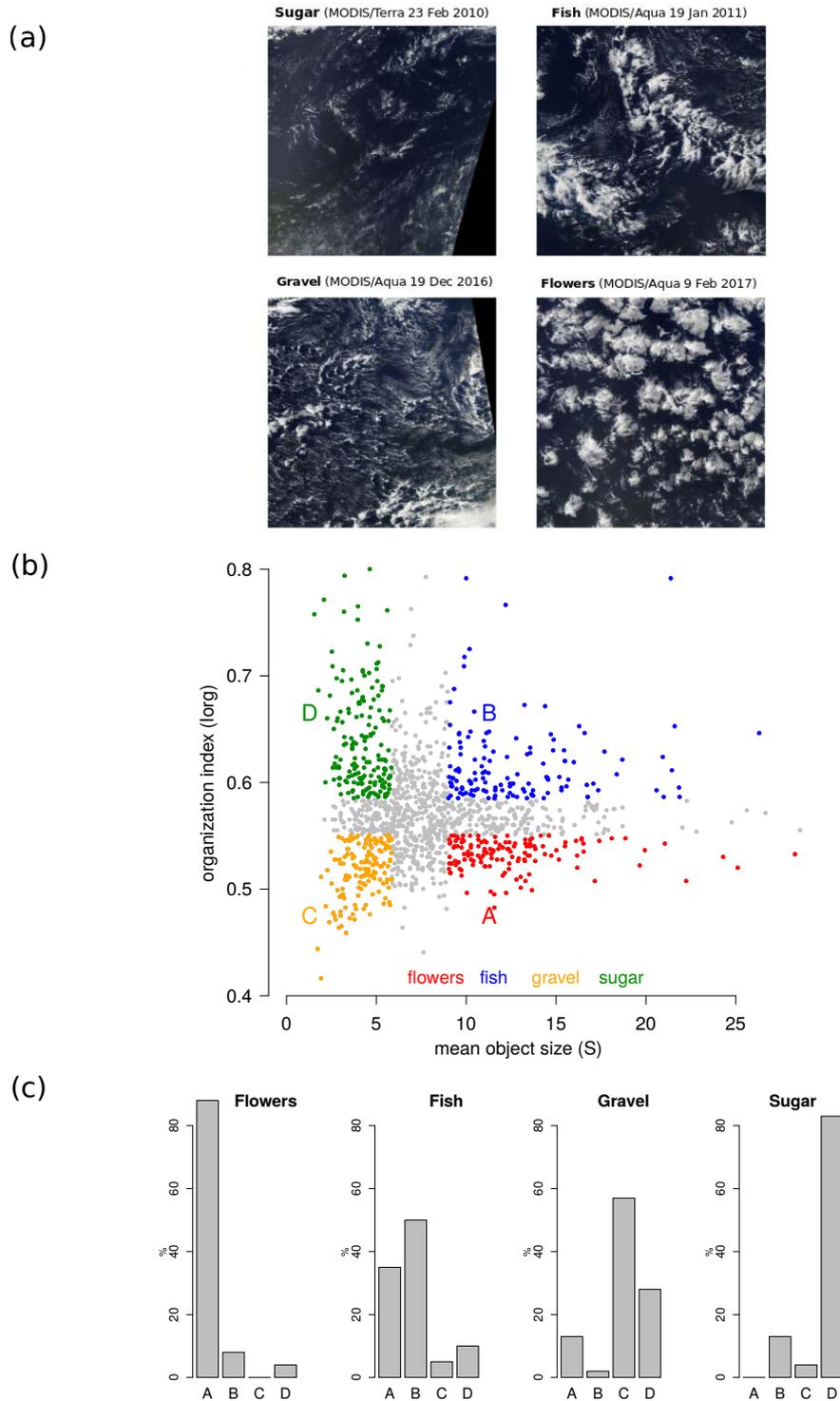


Figure B.1: (a) Illustration of the four prominent cloud patterns of shallow convective organization pointed out by Stevens et al. (2020) over the tropical western Atlantic near Barbados. The four satellite images (48W-58W;10N-20N) are derived from MODIS imagery. (b) Characterization of the shallow convective organization using infrared geostationary satellite data through two metrics: a convective organization index (I_{org}) and the mean object size (S). The lower and upper terciles of I_{org} and S define four classes of mesoscale organization (quadrants A, B, C and D). (c) Relative occurrence of the four cloud patterns defined by Stevens et al. (2020) in each quadrant of the (S , I_{org}) distribution.

B.2.2 *Cloud patterns*

Of the 900 images considered by Stevens et al. (2020), 815 were classified by at least one person as being dominated by one of the four patterns, and 337 were classified robustly (consistent classification by at least four people). To show that the cloud patterns are well delineated in the (S, I_{org}) space, we consider all the robustly classified images falling into one of the four A-B-C-D quadrants (154 images), and ask how frequently each pattern (Flowers, Fish, Gravel and Sugar) fall into each quadrant. Figure B.1c shows that the four quadrants of (S, I_{org}) discriminate among the patterns reasonably well. ‘Flowers’ occurs predominantly in quadrant A, ‘Fish’ in B, ‘Gravel’ in C, and ‘Sugar’ in quadrant D. This is consistent with the visual impression that the Sugar and Gravel patterns are mostly associated with small-scale cloud features while the Flowers and Fish patterns are associated with more extended cloud features. According to the I_{org} index, Flowers and Gravel patterns are associated with a close-to-random distribution of cloud features, while the Fish and Sugar patterns are associated with more clustered cloud objects.

The way in which the patterns distribute themselves in the (S, I_{org}) space is largely intuitive, the one exception being the association of Sugar with large values of I_{org} . As explained above, the cloud objects selected by the chosen brightness temperature thresholds do not correspond to the entire cloud population that exceeds the lifting condensation level, but only to the population of clouds whose top reaches the 290 K isotherm (about 1 km altitude). Sugar situations are characterized by the predominance of very fine scale clouds of very small vertical extent (Stevens et al., 2020). The rare active clouds that reach the 290 K isotherm often appear as isolated, so that their spatial distribution within the $10^\circ \times 10^\circ$ area is characterized by a large I_{org} . Whereas Flower and Sugar are clearly separated classes, Fish and Gravel patterns show some overlap with Flowers and Sugar respectively, an ambiguity that Stevens et al. (2020) also found in the visual classification.

Given the satisfactory correspondence between the visually identified patterns and the four (S, I_{org}) quadrants, in the following we use the objective labeling to associate scenes distributed in quadrants A, B, C, D with ‘Flowers’, ‘Fish’, ‘Gravel’ and ‘Sugar’, respectively. Adopting this methodology allows us to use the full GridSat-B1 record to diagnose the daily occurrence of the four cloud patterns and their co-variability with environmental conditions.

B.2.3 *Robustness of the classification*

We test the robustness of the classification by repeating it using higher resolution MODIS (1 km) channel 31, and GOES-16 (2 km) channel 13 brightness temperatures (Tables S1 and S2, Figures S2 and S3). MODIS provides twice daily data for the 2000-2019 period. From GOES-16, we use 3-hourly data for the last two winter seasons. The higher-resolution data changes the number of cloud objects, the mean cloud object size, and the absolute value of I_{org} (Figure S4). However, the day-to-day variability of the I_{org} , A and S metrics correlates well among the different datasets (Table S3). Despite their very different spatial resolution (8 km vs 2 km), GridSat and GOES-16 classifications correlate best. For the case of I_{org} , the geostationary data (GOES, GridSat) correlate less well with

MODIS, suggesting that I_{org} may vary more with temporal sampling than it does with resolution.

The daily time series of mesoscale patterns (A, B, C, D labels) determined from the upper and lower terciles of the MODIS data (Figure S2) correlate well with those identified using GridSat data (Table S3). The robustness of the classification provides further justification for our association of the upper and lower terciles of the (S , I_{org}) distributions (the four quadrants) with ‘Flowers’, ‘Fish’, ‘Gravel’ and ‘Sugar’.

B.3 DEPENDENCE OF PATTERNS ON THE LARGE-SCALE ENVIRONMENT

To explore how large-scale environmental conditions vary among the quadrants (patterns), we use 6-hourly reanalyses of meteorological data as provided by the ERA-interim product (Dee et al., 2011) for each DJF season from 2000 to 2019 and for several environmental variables: the SST, the near-surface wind speed V_s , the zonal and meridional components of the surface wind u_s and v_s , the zonal wind shear between 700 hPa and the surface, the large-scale vertical velocity at 700 hPa, the Lower Tropospheric Stability (LTS, defined as $\theta_{700} - \theta_{1000}$, where θ is the potential temperature, Klein and Hartmann (1993)), and the Estimated Inversion Strength (EIS, Wood and Bretherton (2006)), defined as $EIS = LTS - \Gamma_m^{850}(z_{700} - \text{LCL})$ where Γ_m^{850} is the moist-adiabatic potential temperature gradient at 850 hPa, z_{700} is the height of the 700 hPa level, and LCL is the height of the Lifting Condensation Level assuming a surface relative humidity of 80%. We also use layered free tropospheric relative humidity data from the Megha-Tropiques satellite (Sivira et al., 2015). Each of these variables is computed as a daily-mean average over the domain.

B.3.1 Day-to-day variability

To test whether different environmental conditions are associated with different patterns, a quadrant composite of each daily-mean environmental variable is constructed. Most of the environmental variables considered do not differ significantly, or differ only marginally, from one pattern to another (Figure S5). However a few variables, namely V_s and EIS (equivalently LTS which correlates nearly perfectly (0.99) with EIS, but we adopt EIS because it generalizes to warmer climates more readily), were discriminating (Figure B.2).

The analysis shows that ‘Flowers’ are associated with relatively cold SSTs, strong surface winds and greater stability. ‘Fish’ pattern were found over more moderate SSTs, weaker winds and strong stability. ‘Gravel’ was likewise associated with moderate SSTs but strong surface winds and low stability. ‘Sugar’ prevailed over the warmest SSTs, when surface winds were weak and stability was low. It thus appears that EIS (or LTS, Figure S6) best discriminates the patterns with small vs large S : the patterns associated with large cloud objects (Flowers and Fish) predominantly occur in situations with a more stable lower troposphere. This is consistent with the expectation that larger stratiform cloud fields are to be expected in situations with enhanced stability (Klein and Hartmann, 1993; Wood and Bretherton, 2006). V_s best discriminates the type of convective organization (I_{org}): random to more regular organizations of cloud centroids (Flowers and Gravel, both associated with low I_{org} values) tend to occur when the trade-winds

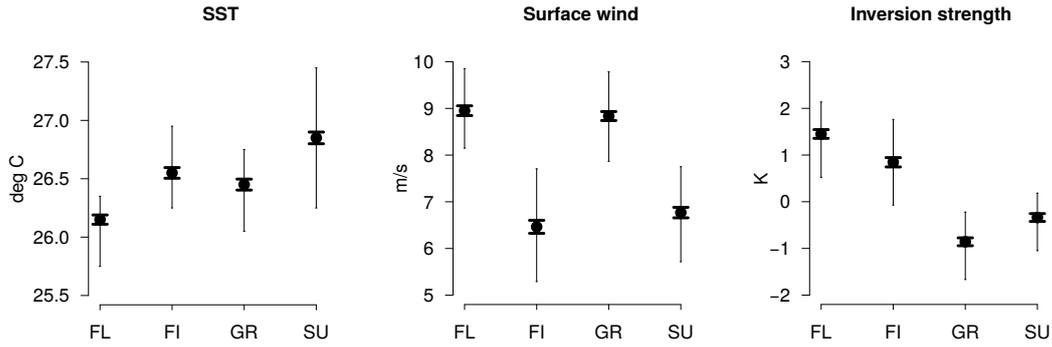


Figure B.2: Large-scale environmental conditions (daily-mean SST, V_s and EIS) composited over the 2000-2019 period as a function of the mesoscale cloud patterns (FL=Flowers; FI=Fish; GR=Gravel; SU=Sugar) inferred from GridSat data. Black markers indicate the mean of the distribution, thin vertical bars the range between the 25th and 75th percentile values, and thick lines \pm the standard error on the mean.

are strong ($V_s \geq 8 \text{ m s}^{-1}$), while the more 'organized' distributions (Sugar and Fish) tend to occur when the trades are weaker. Overall, the cloud patterns that correspond to the most contrasted S and I_{org} metrics, namely the Flowers and Sugar patterns, are those that occur in the most contrasted environments. Repeating this analysis using the MODIS classification (Figure B.3) leads to similar conclusions.

B.3.2 Interannual variability

The analysis of daily variability was extended to explore interannual variability by comparing year-to-year variations of DJF means, each DJF mean being computed by filtering out the days obscured by upper-level clouds or without cloud pattern classification. Once again, the variability of I_{org} and S derived from GridSat-B1 and MODIS datasets are consistent with each other (Figures B.4a-b). The interannual relationships between these metrics and environmental conditions are also consistent with those found at the daily timescale: I_{org} anomalies exhibit a negative correlation with V_s anomalies, and S anomalies exhibit a positive correlation with EIS anomalies (Figures B.4c-d, Table S4). On the other hand, the correlation between these metrics and SST (whose time evolution is shown in Figure S7) is not significant at the interannual timescale.

For each season, the relative prominence of the four patterns is consistent with the V_s and stability anomalies of that season (Figure B.4e). For instance, the 2009-2010 DJF season which was characterized by an anomalously weak V_s (Figure B.4c) and an anomalously strong stability (Figure B.4d) was associated with a predominance of 'Fish'. In contrast, the 2013-2014 DJF was associated with a very strong surface wind and a predominance of 'Gravel' and 'Flowers' patterns, while the 2017-2018 DJF was associated with weak stability and was mostly associated with 'Gravel' and 'Sugar'. The association between cloud patterns and large-scale environmental conditions (as characterized by V_s and EIS) pointed out at the daily timescale is thus able to also explain

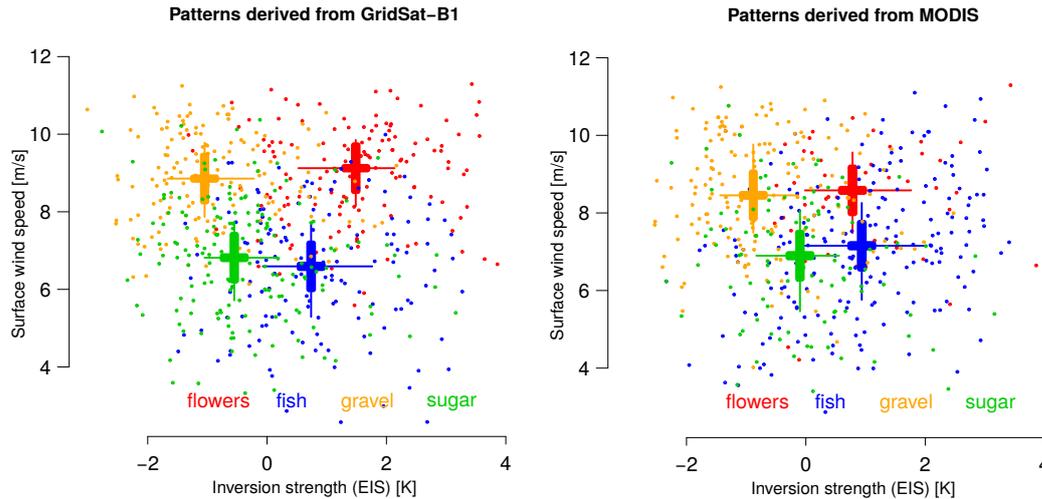


Figure B.3: Scatter plot of daily-mean values of EIS and near-surface wind V_s over 2000-2019. The mesoscale cloud patterns classified as Flowers, Fish, Gravel or Sugar using (left) GridSat or (right) MODIS observations are indicated in colors. Also reported is the mean (EIS, V_s) value computed over the whole period for each cloud pattern. Thin bars indicate the 25th and 75th percentiles of the distributions, and thick bars indicate \pm the standard error on the mean.

the year-to-year variations of the spatial organization metrics and the predominance of a specific mesoscale cloud pattern (Figure S8).

B.4 SUMMARY AND DISCUSSION

Stevens et al. (2020) showed, based on a visual and thus subjective classification, that the tropical western Atlantic during boreal winter is associated with four prominent mesoscale patterns of shallow convection. The present study shows that these patterns can be objectively identified based on the size and degree of clustering of segmented cloud objects as identified from infrared brightness temperatures. The classification is largely insensitive to the spatial resolution of the brightness temperature data: GridSat data with a resolution of 8 km and MODIS data with a resolution of 1 km lead to very similar classifications.

The analysis of daily and interannual variations shows that the relative occurrence of the different cloud patterns correlates strongly with two environmental factors: the strength of the near-surface wind speed and the strength of the lower-tropospheric stability (Figures B.3 and S2). Flowers tend to occur in windy ($V_s > 8 \text{ m.s}^{-1}$) and stable environments ($EIS > 0.5 \text{ K}$), while Sugar tends to occur in calm ($V_s < 8 \text{ m.s}^{-1}$) and unstable environments ($EIS < 0.5 \text{ K}$). Fish appears to prefer calm and stable environments, while Gravel tends to occur in windy and unstable environments. These relationships beg a physical explanation. For this purpose, data from the forthcoming 2020 EUREC⁴A (*Elucidating the role of cloud-circulation coupling in climate*) field campaign should be well suited (Bony et al., 2017). With its large complement of air and sea-going vessels in the same study region as examined here, EUREC⁴A will not only quantify the

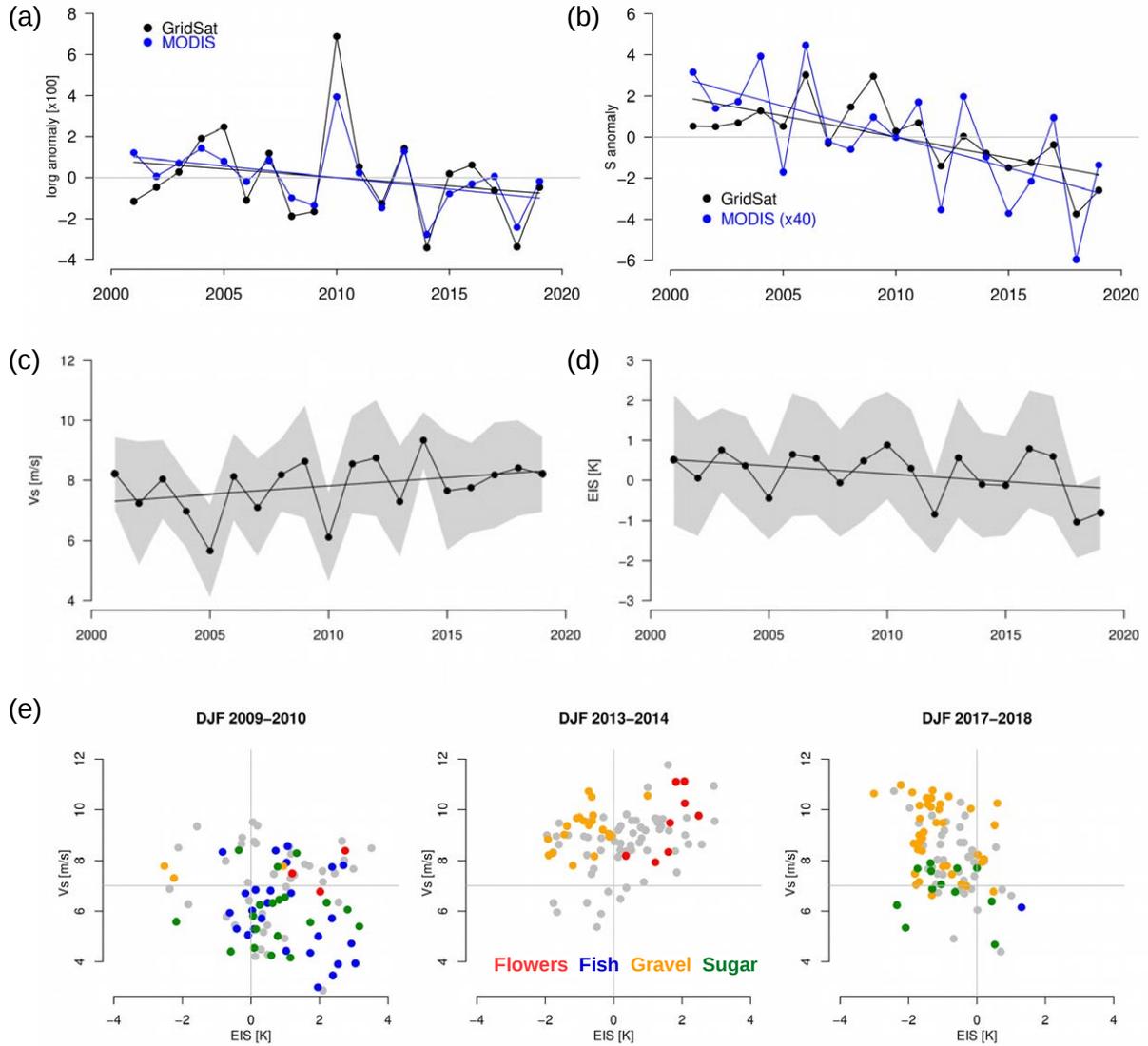


Figure B.4: Interannual anomalies of (a) the organization index (I_{org}) and (b) the mean cloud object size S computed from GridSat or MODIS observations over the period 2000–2019 during DJF (the correlation between GridSat and MODIS timeseries is 0.90 for I_{org} and 0.78 for S). Interannual evolution of (c) V_s and (d) EIS derived from ERA interim for the same period. Note that in (a–d), the year is defined by the Jan–Feb months of the DJF season (e.g. 2010 corresponds to Dec 2009 – Feb 2010). The shading represents \pm one standard deviation of daily-mean values around the DJF mean. (e) Examples, for a few DJF seasons, of the daily cloud patterns identified from GridSat data represented as a function of the daily (EIS, V_s) conditions of that season (the grey lines are just visual guides).

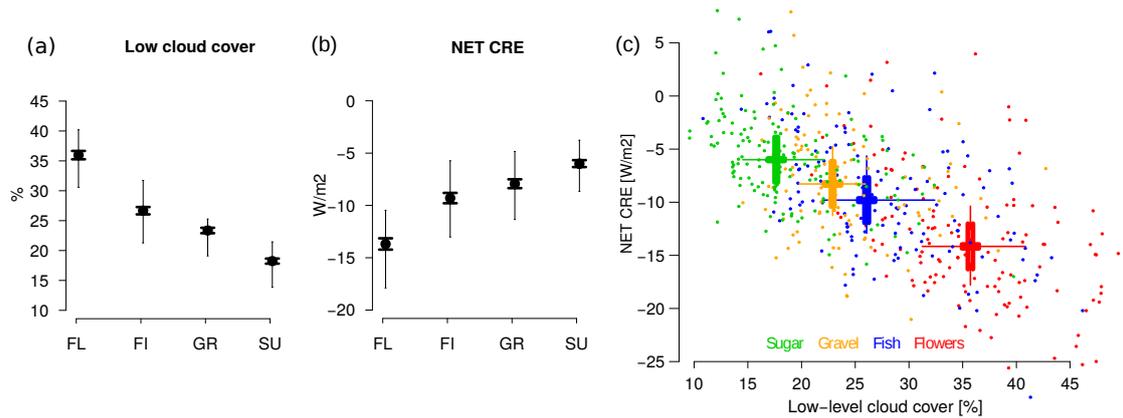


Figure B.5: (a) and (b) Same as Figure B.2 but for the low-cloud amount derived from MODIS cloud products and the NET cloud-radiative effect derived from CERES observations. (c) Same as Figure B.3 but for daily-mean values of NET CRE and low-level cloud amount.

relationship between cloud patterns and large-scale variables, but also the circulation systems that connect the two. These measurements should thus also help determine how much the mesoscale organization of shallow convection has to be considered if one wants to understand and predict the response of shallow clouds to changes in environmental conditions.

A closely related question is whether the mesoscale organization of shallow convection matters for cloud-radiative effects. To shed light on this issue, we used daily estimates of top-of-atmosphere radiative fluxes and cloud products from the CERES (Clouds and the Earth's Radiant Energy System) geostationary enhanced temporally interpolated dataset (Wielicki et al., 1996), along with low-cloud amount retrievals provided as part of the same data set for the period 2001-2017. Low-level cloud amount varies by a factor of two across the different patterns, and the net CRE associated with 'Flowers' is about double that of the Sugar pattern (Figure B.5). At first order, the CRE (dominated by its shortwave component) varies linearly with the low-cloud amount (Klein and Hartmann, 1993), so that radiative differences across the patterns are related to differences in the low-cloud amount. However, unlike what has been found for other types of mesoscale organizations of marine low-clouds (McCoy et al., 2017), for a given low-cloud amount we do not notice significant radiative difference among patterns. It suggests that over the western tropical Atlantic, changes in the mesoscale organization of trade-wind cumuli primarily affect the top-of-atmosphere radiation budget through associated changes in the low-level cloud amount.

Could this tight relationship between convective organization and low-cloud amount, or CRE, imply that changes in cloud organization have the potential to influence cloud-radiative feedbacks? The large-scale environment in which the trade-wind cumuli form might change under global warming. Climate models predict EIS increases over the western tropical Atlantic as the planet warms (Qu et al., 2015). On the other hand, the change in V_s remains uncertain, partly because the geographical pattern of surface warming can act against the anticipated slow-down of the large-scale circulation (Ma

et al., 2016). Indeed, in climate change experiments run with the IPSL climate model (Dufresne et al., 2013), EIS always increases with global warming over the tropical western Atlantic (by 0.1 to 0.7 K K^{-1} depending on the type of experiment and model version), whereas, V_s does not change in a robust fashion. Assuming that V_s and EIS remain the main controlling factors of the mesoscale organization of shallow clouds in a perturbed climate, these projections would suggest a more frequent occurrence of Fish or Flower at the expense of Sugar or Gravel with global warming, and thus a larger cloud fraction. This is in conflict with the prevailing idea, based on models which do not account for mesoscale organization, that low-cloud amount will reduce in response to rising SST (Klein et al., 2017). In our analysis, SST does not appear to be a strong controlling factor of the cloud mesoscale organization on daily and interannual timescales (Table S2), but it remains an open question whether it could play a bigger role in climate change. In either case, better understanding the extent to which the mesoscale patterning of clouds affects their response to warming appears relevant to establishing confidence in how clouds respond to warming as a whole.

Future investigations of this issue using numerical models that predict explicitly these different cloud patterns and are able to reproduce the relationships discussed in this paper should help determine how much the cloud organization is sensitive to SST, and how much it could affect the magnitude and even maybe the sign of the change in low-cloud amount. This should fill an important gap in our understanding and our assessment of low-cloud feedbacks under climate change.

B.5 ACKNOWLEDGMENTS

We gratefully acknowledge the International Institute for Space Studies (ISSI) science team which was instrumental in initiating this study, and two anonymous reviewers for their thoughtful input. The GridSat-B1 primary data are available at <http://www.ncdc.noaa.gov/gridsat/gridsat-index.php?name=data>. The MODIS/Terra and Aqua Calibrated Radiances 5Min L1B Swath 1km and the MODIS/Terra and Aqua Geolocation Fields 5Min L1A Swath 1km datasets were acquired from the Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (doi.org/10.5067/MODIS/MOD021KM.061, doi.org/10.5067/MODIS/MYD021KM.061, doi.org/10.5067/MODIS/MOD03.061, doi.org/10.5067/MODIS/MYD03.061). GOES-16 data are available from the GOES-R Calibration Working Group: NOAA GOES-R Series Advanced Baseline Imager Level 1b Radiances, NOAA NCEI ([doi:10.7289/V5BV7DSR](https://doi.org/10.7289/V5BV7DSR), accessed 04/19/2019). The Synoptic Radiative Fluxes and Clouds dataset (SYN1deg-Day, Edition 4A, [doi:10.5067/Terra+Aqua/CERES/SYN1degDAY_{_}L3.004A](https://doi.org/10.5067/Terra+Aqua/CERES/SYN1degDAY_{_}L3.004A)) is made available by the NASA CERES group. ERA-interim reanalyses were downloaded from <http://climserv.ipsl.polytechnique.fr/fr/les-donnees/era-interim.html>. Megha-Tropiques retrievals of layered relative humidity (version 2 of the daily product L2B-RH gridded at 1×1 deg for Dec 2011 - Feb 2018) were downloaded from <http://www.icare.univ-lille1.fr/mt/products>. This project has received funding from the European Union's Horizon 2020 research and innovation programme (European Research Council grant EUREC⁴A No 694768, CONSTRAIN project No 820829), and from the Max Planck Society. Data are available upon request.

CHARACTERIZATION AND EVOLUTION OF ORGANIZED
SHALLOW CONVECTION IN THE DOWNSTREAM NORTH
ATLANTIC TRADES

The work in this appendix has been published as:

Schulz, Hauke, Eastman, Ryan, and Stevens, Bjorn (2021). Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades. *Journal of Geophysical Research: Atmospheres* 126.17, e2021JD034575. doi: [10.1029/2021JD034575](https://doi.org/10.1029/2021JD034575)

The contributions of the authors to this publication are as follows:

The conceptual design, research focus, analysis and writing have been planned and conducted by HS. ERA-5 trajectory data has been calculated by RE. All authors proof-read several drafts and approved the final manuscript. BS supervised the study.

Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades

Hauke Schulz¹, Ryan Eastman², Bjorn Stevens¹

¹Max Planck Institute for Meteorology

²University of Washington

ABSTRACT

Four previously identified patterns of meso-scale cloud organization in the trades – called *Sugar*, *Gravel*, *Flowers* and *Fish*– are studied using long-term records of ground-based measurements, satellite observations and reanalyses. A deep neural network trained to detect these patterns is applied to satellite imagery to identify periods during which a particular pattern is observed over the Barbados Cloud Observatory. Surface-based remote sensing at the observatory is composited and shows that the patterns can be distinguished by differences in cloud geometry. Variations in total cloudiness among the patterns are dominated by variations in cloud-top cloudiness. Cloud amount near cloud base varies little. Each pattern is associated with a distinct atmospheric environment whose characteristics are traced back to origins that are not solely within the trades. *Sugar* air-masses are characterized by weak winds and of tropical origin. *Fish* are driven by convergence lines originating from synoptical disturbances. *Gravel* and *Flowers* are most native to the trades, but distinguish themselves with slightly stronger winds and stronger subsidence in the first case and greater stability in the latter. The patterns with the higher cloud amounts and more negative cloud-radiative effects, *Flowers* and *Fish*, are selected by conditions expected to occur less frequently with greenhouse warming.

C.1 INTRODUCTION

The organization of deep convection has long been recognized to influence the global distribution of moisture and, as a consequence, the climate. Shallow convection, as is common in the trades, is usually not thought of as being organized. Rather, in the mind's eye of many researchers, trade-wind clouds are randomly distributed, have little vertical development and have generally been assumed to play little role in the climate system. Over the past twenty years however, the out-sized role of maritime shallow clouds on Earth's radiation budget – and discrepancies in how models predict their changes with warming (Bony and Dufresne, 2005) – have made a determination of processes controlling their coverage a central focus of climate science. During this period, observational studies such as RICO (Rauber et al., 2007) and the emergence of satellite imagery with spatial resolution on the hectometer scale began emphasizing how shallow clouds in the trades adopt different forms of organization, often in association with precipitation development and the formation of cold pools (Seifert and Heus, 2013; Seifert et al., 2015; Zuidema et al., 2011). More recently, several approaches to characterize these forms of organizations have been developed (Denby, 2020; Janssens et al., 2021; Stevens et al., 2020). Stevens et al. (2020) identified that most of the large-scale patterns of shallow convection can be categorized into four categories, which they called *Sugar*,

Gravel, Flowers and *Fish*. Although based solely on the subjective visual inspection of visible satellite imagery in the vicinity of the Barbados Cloud Observatory, these patterns have varied net cloud radiative effects and thus may influence Earth’s climate sensitivity (Bony et al., 2020).

In contrast to Denby (2020) and Janssens et al. (2021) who retrieve a continuum of classes, the limitation to the few categories of Stevens et al. (2020) helps to break down natures’ complexity into manageable pieces, arguably the building blocks of shallow convection in the downstream trades. Understanding what factors help select these patterns may help understand factors controlling cloudiness as a whole.

Motivated by the potential impact on climate sensitivity and the striking differences in the visual appearance of the four patterns identified by Stevens et al. (2020), we are interested in better understanding the basic features of these four patterns, and the factors that influence their emergence. Specifically:

1. How do the four patterns differ in terms of the observed cloud geometry e.g., cloud fraction, cloud size and cloud base height?
2. Are the different patterns associated with different large-scale environments?
3. To what extent are these large-scale differences reflective of different air mass origins?

To answer these questions we collocate the four cloud patterns of meso-scale organization as automatically detected in satellite measurements with observations made at the Barbados Cloud Observatory. This contextualization of the high-resolution ground-based cloud measurements within the meso-scale patterning enables deeper insights about their characteristics than possible using satellite measurements alone.

The methods adopted and the data used are described in Section C.2. A characterization of the cloud patterns, with a focus on the cloud-geometric properties and how they fit with our preconceptions as derived from the satellite images is given in Section C.3. In Section C.4 we analyze the meteorological conditions under which the patterns occur and the extent to which they can be distinguished. The effects of air-mass origin are analyzed with the aid of back-trajectories in Section C.5. We conclude with Section C.6.

C.2 DATA AND METHODS

This study uses several data sources as shown in Tab. C.1.

C.2.1 *Pattern detection of shallow convection*

To detect the four patterns of shallow convection, we use the Keras RetinaNet (Gaiser et al., 2018). We trained this deep neural network (in the following just *neural network*), as described in Rasp et al. (2020) with only a difference in the input datas’ radiance. While we used the same 49 000 manually created labels that were performed based on 10 years of visible imagery captured by the Moderate Resolution Imaging Spectroradiometer (MODIS) instruments aboard the satellites AQUA and TERRA, here we use the simultaneously captured brightness temperatures.

Table C.1: Primary data

Platform	Instrument	Variables	Location	Sampling
GOES-16	ABI	brightness temperature (channel 13)	10°N-24°N; 61°W-40°W	2km; 30min
AQUA	MODIS	- brightness temperature (ch. 31) - corrected reflectance (ch. 1,3,4)	10°S-55°N; 100°W-10°E	daily (daytime overpass)
ERA5	n/a	ω , div, T, u, v, SST	0°N-60°N; 70°W-10°E	0.25° x 0.25°; 3h
BCO	- Ka-Band radar - Raman lidar - MRRR - weather sensor - microwave radiometer	- reflectivity profile - mixing ratio profile (only Fig. C.2) - rain rate - surface p, T, RH, u, v - LWP, IWV	13.16°N; 59.43°W	- 10s; 30m - 120s; 60m - 60s - 10s - 1s
Grandley Adams Airport	- soundings	- p, T, RH, u, v	13.07°N; 59.50°W	12h-24h

To capture the inter-pattern variability at a fixed site, like the Barbados Cloud Observatory, it needs to be assumed, that the patterns persist for at least the time it takes to advect over the site. In case of the patterns studied here, the characteristic length scale of a typical pattern can easily exceed 90 km, or about 3 h, for an advection speed of 8 m s^{-1} . However, 3 h is about the time between AQUA and TERRA overpasses, where we have noticed differences in the distribution of pattern types. Using infrared data allows us to overcome this issue by also retrieving useful classifications at night and increasing the number of classifications per day, even more so by applying the neural network to infrared satellite imagery taken by the Advanced Baseline Imager (ABI) aboard GOES-16. Although the training of a separate neural network directly on the ABI data would have been preferred, it was not possible due to missing overlap with the manual classifications. Nevertheless, the channels of MODIS (channel 31; $10.78 \mu\text{m}$ to $11.28 \mu\text{m}$) and ABI (channel 13; $10.18 \mu\text{m}$ to $10.48 \mu\text{m}$) used here are relatively close to each other and in a wavelength range where absorption due to water vapor is relatively constant and overall minor. Differences are therefore expected to be minimal. A comparison of the overall performance of the visual classifications used in Rasp et al. (2020) to the infrared classifications used in this study is given in the supplemental material.

While the GOES-16 Advanced Baseline Imager can capture images every minute for pre-selected regions, here we extract the region of interest ($10\text{N}, 24\text{N}$; $-61\text{E}, -40\text{E}$) from the full-disk scenes and restrict ourselves to a temporal resolution of 30 min and the nominal spatial resolution of 2 km. Because the GOES-16 satellite only recently started its operation, we restrict our analysis to the three boreal winter seasons 2017/2018 (JFM), 2018/2019 (NDJFM) and 2019/2020 (NDJFM).

To attribute one of the four patterns to the observations made at the Barbados Cloud Observatory, each classification of the neural network is evaluated at the location of the observatory. Because the RetinaNet is an object-detection algorithm, classifications can overlap or contain cloud formations that are less characteristic of the main identified pattern. As we wish to include only clear and long-lasting patterns, while still remaining a statistically robust sample size, the complete time-series is split into 6 h time windows which are associated with a specific pattern in cases where a pattern dominates this time window e.g., is detected for at least half of the time (3 h). This results in 42 % of the 6 h windows being associated with one of the four patterns, while another 37 % could not be attributed clearly to any category, only in 21 % of the cases was no pattern detected for 3 h. Because the patterns are not equally likely to occur, the number of detected cases differs as indicated in Tab. C.2. Time windows that do not contain a long-lasting pattern were grouped together as *Others*.

Table C.2: Number of time windows that contain robustly identified patterns in the boreal winters of 2017/2018 (JFM), 2018/2019 (NDJFM) and 2019/2020 (NDJFM)

pattern	# of 6h windows	% of total	% of robust patterns
Sugar	145	9	22
Gravel	305	19	46
Flowers	77	5	12
Fish	141	9	21
Others	846	58	
<i>mixed</i>	567	36	N/A
<i>no pattern</i>	337	21	

For the detection of the seasonal cycle of the patterns and the trajectory analysis we do not need to have high temporal sampling. Hence, we use the daytime MODIS AQUA overpass to identify pattern which allows us to extend the record to 11 years (2010-2020).

c.2.2 Back-trajectories

To analyze the origin of the patterns and the evolution of the air mass in which they are found, we calculate back-trajectories following the framework of Eastman and Wood (2016). Vertical winds are assumed to be negligible compared to the horizontal components, such that the trajectories are followed near the top of the sub-cloud (boundary) layer and kept constant at the initial height of 925 hPa. These boundary-layer trajectories are calculated using winds from the 5th European Center Reanalysis of meteorological observations (ERA5) (Hersbach et al., 2020) on a 6-hourly time step.

The back-trajectories are initialized at the center of each classification within the domain 10°N - 24°N and 61°W - 40°W and most closely to the AQUA overpass time. The analysis covers the winter months (November through March) of 2010-2020. Trajectories are calculated for an 84 h period and atmospheric properties along the trajectory are extracted at each 12-hour time-step coinciding with an AQUA overpass. Reanalysis variables are taken from a 1×1 degree latitude-longitude grid, with averages produced for all boxes with centers that fall within 100 km of trajectory sampling points.

c.2.3 Ground-based measurements

We use surface observations from the Barbados Cloud Observatory (BCO). The only long-term cloud observatory in the broader trades.

c.2.3.1 Instrumentation

The BCO uses advanced remote sensing instrumentation to measure the undisturbed marine subtropical atmosphere (Medeiros and Nuijens, 2016; Stevens et al., 2016). In this study, we use simultaneous measurements from the CORAL Ka-band cloud radar and Raman lidar to characterize clouds and their thermodynamic environment, especially the surrounding humidity structure of the clouds. These advanced remote-sensing measurements are complemented by 752 soundings of the nearby Grantley-Adams airport to improve the statistics above clouds, which can quickly attenuate the lidar signal and make a retrieval inside and above clouds impossible. Radiosondes are launched once or twice a day, usually an hour or so before their 0 UTC and 12 UTC report times. The closest of these soundings has been attributed to each of the 6 h time windows.

To detect only hydro-meteors with the cloud radar and no sea-salt aerosols, we apply a threshold of -50 dBZ as used in Klingebiel et al. (2019).

Integral measurements of liquid water path (LWP) and total integrated water vapor (IWV) are retrieved from microwave radiometer measurements following Löhnert and Crewell (2003) and Steinke et al. (2015).

In addition to the standard surface meteorological measurements from a Vaisala WXT-520, we use the rain rate measurements from a micro-rain radar (MRR). Due to its larger sampling area compared to the also available acoustic rain sensor, it detects more

reliable light (0.01 mm h^{-1} (Peters et al., 2002)) and/or short rain events. However, this comes at the cost of measuring the rain rate above the surface (325 m) rather than at the surface – which will be larger than at the surface, disproportionately so for light rain.

These datasets are available for 90% of the 6 h time windows and are equally distributed among the patterns. The only exception is the radiometer data with only about 60% availability.

c.2.3.2 *Cloud entity classification*

The identifications of meso-scale patterns of shallow convection are supplemented with cloud-type classifications derived from the BCO measurements.

Cloud-geometric properties of single cloud entities are retrieved based on the segmentation of the radar reflectivity. Individual clouds are identified by testing the connectivity of radar retrievals in height and time. Since a main part of this study focuses on stratiform layers, we use a running window of 100 s in time and only direct connections in the vertical to account for the fact that the stratiform layers can be so thin that they are not continuously detected by the radar.

Similar to Lamer et al. (2015) we classify individual clouds by their cloud-base heights (CBH). Stratiform layers are defined as clouds that have a frequent CBH above 1 km up to 2.5 km. As shall be seen later (Fig. C.3) the echo fraction minimizes at 1 km and therefore a threshold of 1 km separates best the stratiform cloudiness from the cloud layer below. This layer of clouds with CBHs below 1 km we classified as originating from the cumulus gene. An example of the radar reflectivity and the derived cloud-type classifications is shown in Fig. C.1. It illustrates that also a mixture of cumulus with an attached stratiform layer may exist. These cases are actually classified as “StSc+Cu” in the case the stratiform layer lasts for at least 20% of the time it takes for the cumulus cloud-entity to pass over the observatory.

Based on the single cloud entities, geometric properties like stratiform extent and mean thickness of stratiform layers are calculated and associated with each entity.

C.3 SURFACE BASED CHARACTERIZATION OF CLOUDINESS AND PRECIPITATION

The four patterns identified by Stevens et al. (2020) – *Sugar*, *Gravel*, *Flowers*, *Fish*– are purely defined by their visual impression from space, predominantly the spatial distribution of cloudiness. The cloudiness is therefore the physical quantity closest to the definition of these patterns. Among the physical differences that may accompany these patterns, aspects of cloudiness that go beyond the spatial arrangement of reflectivity as seen from above will be important to characterize, especially in so far as it influences the cloud radiative effect.

An overview of these patterns and the ground-based observations linked to them is shown in Fig. C.2. Differences in cloudiness are readily apparent, and conform to what has been previously noted in the literature. *Sugar* is identified with a fine dusting of clouds, *Gravel* with cloud features arranged around arc-like structures. *Flowers* and *Fish* are composed of elements that are yet larger in scale and show a clearer separation between cloudy and clear-sky areas. For *Flowers* the clouds show a more isotropic

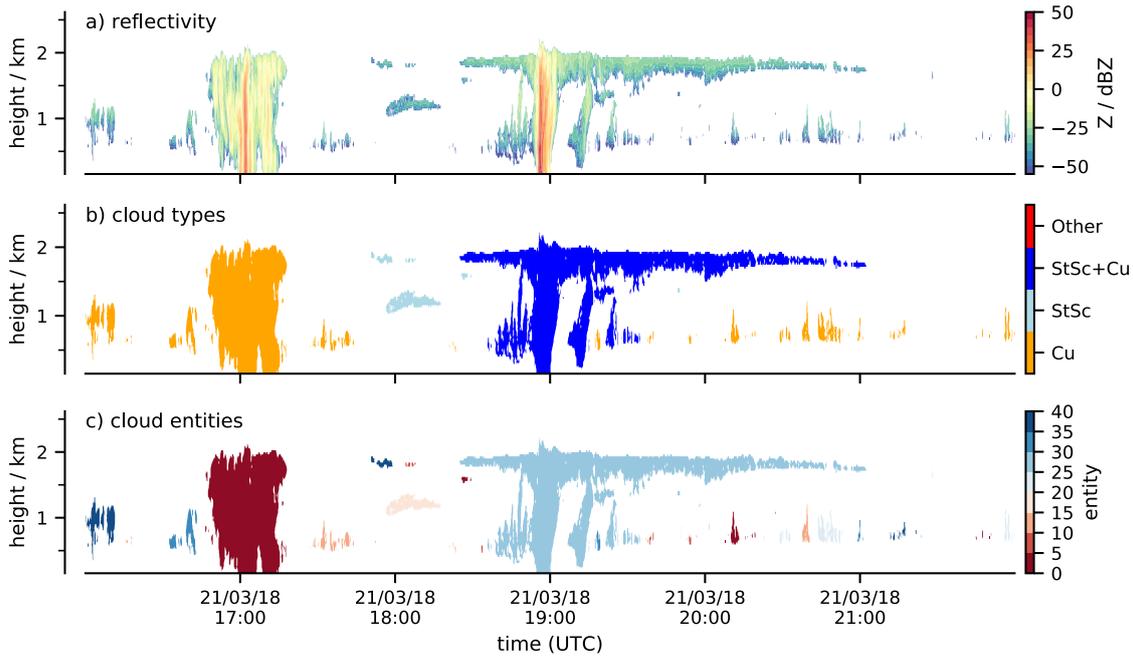


Figure C.1: Example of cloud-type classification based on radar reflectivity

distribution, while they are usually elongated – roughly West to East – in the case of *Fish*.

From these illustrations, which add to the examples shown in Stevens et al. (2020), it is natural to develop preconceptions about differences in the three-dimensional structure of the boundary layer associated with the patterns. For example, *Gravel* is generally thought to be associated with precipitation due to the visible cold-pool signature in the cloud field, and *Flowers* are thought to be composed of stratiform clouds with suppressed convection around them. Assessing whether such preconceptions are supported by the data, is one of the goals of this section.

We first focus on the characteristics of the cloudiness in terms of their geometric properties. Thereafter we analyze the precipitation signatures of the patterns, as they might help to gain a process understanding on how these different patterns form.

C.3.1 Cloudiness

The cloud cover at the Barbados Cloud Observatory is shaped by the ubiquitous appearance of cumulus humilis – i.e., cumulus clouds of very limited vertical extent. Cumulus humilis are not the only cloud type measured at the site. Even in the northern hemispheric winter, when the Intertropical Convergence Zone is furthest away from Barbados and the region experiences strong subsidence, the measured cloud fraction is not solely caused by non-precipitating cumulus humilis (Riehl, 1954). This is demonstrated by an analysis of the mean radar echo fraction profile (a combination of cloud- and rain-fraction) shown in Fig. C.3. Echoes are detected extending to depths above 3 km.

Looking at the mean echo fraction profiles of *Sugar*, *Gravel*, *Flowers* and *Fish* and also the overall wintertime mean echo fraction, suggests that all but *Fish* are some form of

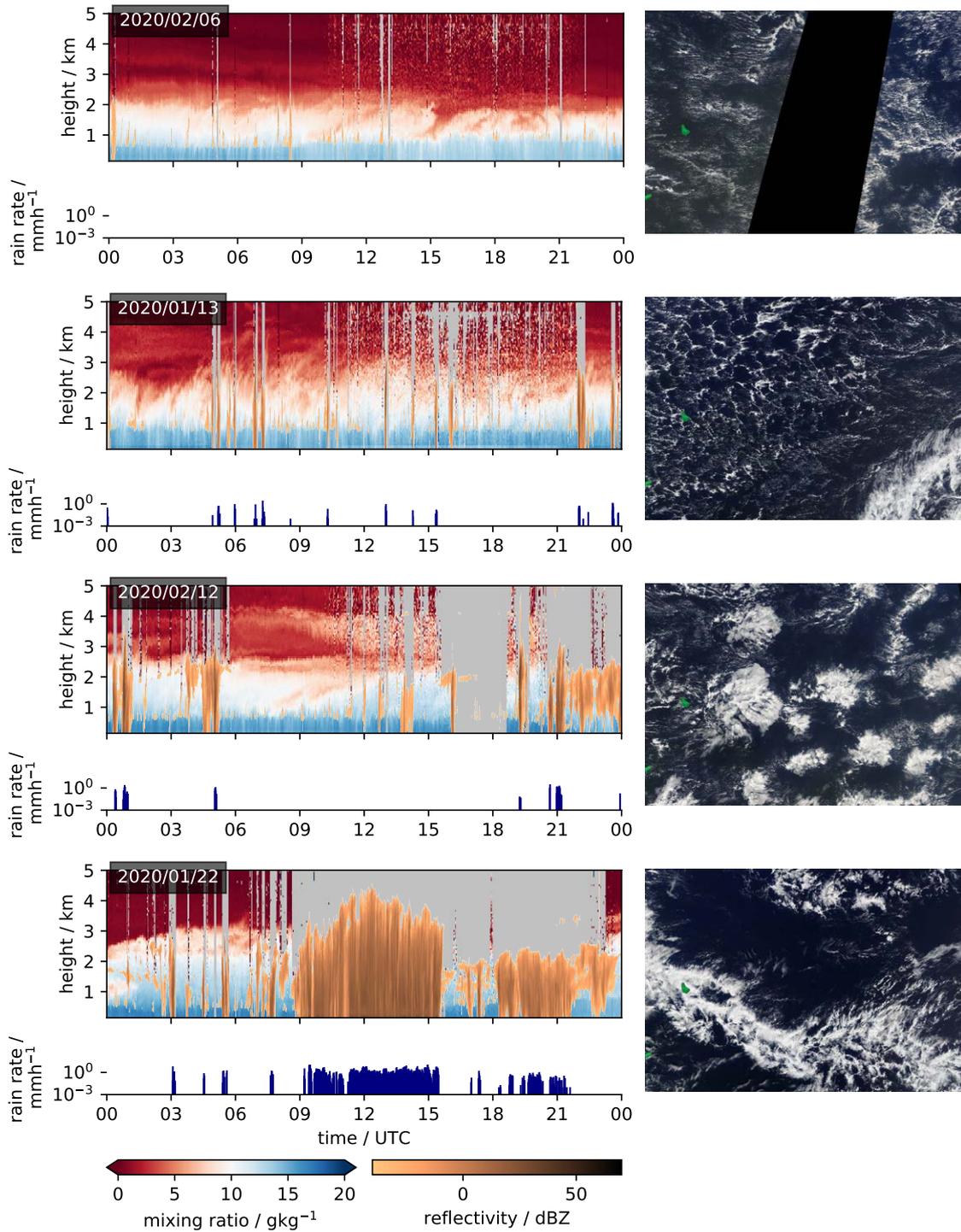


Figure C.2: Time-series of each cloud pattern as identified by scientists participating in the EUREC⁴A campaign (top to bottom: *Sugar*, *Gravel*, *Flowers*, *Fish*). Water vapor measurements from the Raman lidar overlaid by radar reflectivity shown in upper panels, while rain rates measured at 325 m are shown in the lower panels. The according MODIS images from the TERRA satellite overpass are shown on the right. Missing values are colored grey.

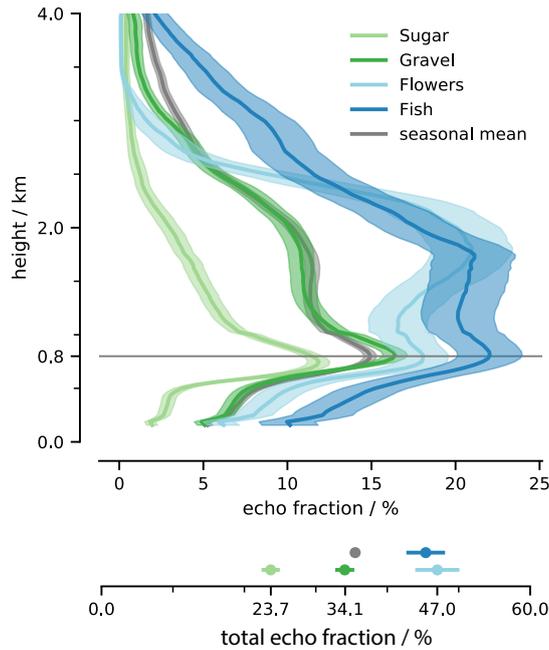


Figure C.3: Echo-fraction measured at the BCO and grouped by detected pattern indicating the combination of cloud- and rain-fraction (upper panel). The height integral, total echo fraction, is shown in the bottom panel. The overall mean of the analyzed winter seasons is shown in grey with the height of maximum echo fraction shown as horizontal line. Shading and whiskers indicate standard error of mean.

shallow convection, with very small echo fractions (less than 3% at 4 km) extending much above 2.5 km.

The largest inter-pattern variability in echo fraction is found in the layer between 1.5 km to 2.5 km. At 1.7 km echo fractions vary between 5% to 20% and explain a large part of the differences that give rise to the differences seen in the satellite imagery (e.g., Fig. C.2). For instance *Flowers*, with its cloudy patches of high reflectivity paired with the sheet-like structure anticipate a strong stratiform component in the cloud fraction compared to *Sugar* and *Gravel*, as is evident in the echo-fraction profiles. The overall echo fraction of *Flowers* (0.47) is therefore much more influenced by the stratiform cloud component as compared to *Sugar* (0.24) and *Gravel* (0.34). *Fish* has high echo-fractions throughout the cloud layer, but are less obviously dominated by a stratiform component as compared to simply more cloudiness, which often extends much more deeply through the lower troposphere.

Common to all patterns is the similarity in difference between the echo fraction at the surface and 700 m, which we interpret as the cloud base cloud fraction. The inter-pattern variations in echo-fraction at 700 m that do exist in Fig. C.3 can largely be attributed to rain events, i.e., differences below 500 m – which are a signature of precipitation – are similar to those at 700 m. The lack of variability of cloud amount at the cloud-base height was emphasized by Nuijens et al. (2014). That *Flowers* would have a similar echo fraction at cloud-base as *Sugar* when neglecting the rain contribution to the echo fraction was not something we would have guessed from the satellite imagery. It shows that an

abundance of clouds near cloud-base under the cloud shield compensates for an absence of shallow-cloudiness in the cloud-free part of the *Flowers* pattern.

Looking at the cloudiness of the patterns as a whole, we recognize that the echo fraction of *Gravel* has the strongest similarity to the seasonal mean echo fraction, which is the average of all 6 h windows independent of any pattern. This is consistent with *Gravel* being the most common pattern detected in this study (about 19% of all regarded time windows and 45% of the windows with any dominant pattern). Further, it also suggests that a large portion of the more uncertain and mixed time-windows contains cloudiness similar to the *Gravel* pattern. *Sugar*, in contrast, occurs rather seldom with 9%. This might seem to contradict Rasp et al. (2020) who found that *Sugar* is actually more often identified than *Gravel*. However, similar to Stevens et al. (2020), who were looking for dominating patterns on a fixed domain, we look for dominating patterns within a fixed time-period. Both methods register only patterns that are persistent for a long time or cover a large area, both of which de-emphasize *Sugar*. The cloud pattern with randomly distributed clouds of little vertical extent occurs frequently, but is often not dominant and thus not picked out by our analysis.

Fig. C.4 confirms that the differences in echo fraction at the lifting condensation level and below are indeed caused by different contributions at the higher end of the reflectivity spectrum (> 0 dBZ) which is indicative of precipitating hydro-meteors. Much more similar across patterns is that the vast majority of hydro-meteors are found at the lower end of the reflectivity spectrum. While the reflectivities below -50 dBZ close to the surface are characteristic for hygroscopically grown sea-salt particles Klingebiel et al. (2019), with increasing height and reflectivity (towards -15 dBZ) at about 2 km the imprint of non-precipitating cumulus humilis (Lonitz et al., 2015) is found.

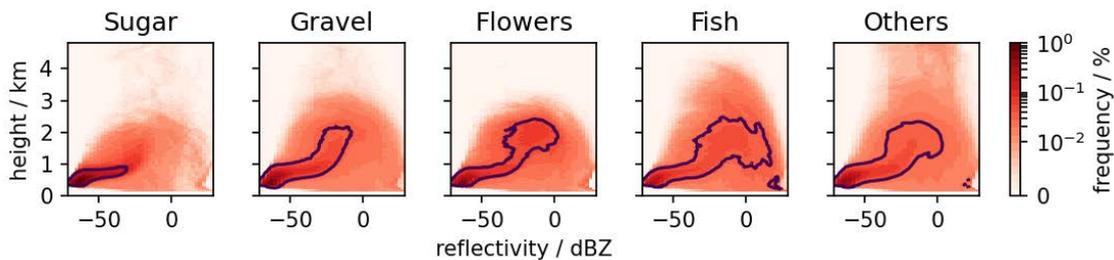


Figure C.4: Contoured frequency by altitude diagram (CFAD) for the four patterns of shallow convection and less clear patterns gathered in the group *Others*. The colors indicate the frequency of occurrence of a reflectivity-height tuple within a specific pattern. The tuples explaining 50% of all values are contoured in black. Hydrometeor-free profiles are excluded.

Despite a similar cloud cover for both *Flowers* and *Fish* of about 0.5, Fig. C.4 indicates different relationships between the cumulus and the stratiform cloud layer. Whereas *Flowers* show a second distinct reflectivity maximum at about 2.2 km and near -5 dBZ, the distribution is more monomodal for *Fish*. This two-layer structure suggests that *Flowers* are only sporadically connected by higher cumulus convection whereas for *Fish*, clouds aloft appear to be deeper and as a more continuous extension of clouds near cloud base. More like the more active and deeper distribution of *Gravel*. The deeper

echoes for *Fish* are also evident in a stronger precipitation feature (i.e., 25 dBZ near surface mode).

Looking at single entities of the clouds detected within the classified 6 h analysis periods as described in Sec. C.2.3.2, we found no evidence that these different couplings of the stratiform cloud layer to the underlying cumulus convection influences the geometric properties of these layers. Independent of the meso-scale organization, cumulus coupled stratiform layers (*StSc+Cu*) have a CBH between 980 m and 1050 m and a cloud top height between 1240 m to 1370 m. More broadly, any of the individual cloud components differentiated here: stratus (*StSc*), cumulus (*Cu*) and the connection of both/stratocumulus cumulonimbus (*StSc+Cu*) show similar geometric properties independent of the pattern.

The cloud base height (CBH) of cumulus, which to a first approximation is the same as the lifting condensation level of near surface air, is at about 650 m to 700 m with cloud tops about 160 m higher. Stratus are about 130 m thick on average, with an average CBH between 1600 m to 1850 m, which is somewhat higher than for stratiform layers that are connected to a convective core during the time of observation.

The only differences that do exist, are the lack of stratiform layers in case of *Sugar* and the increase in size of the stratiform components (*StSc;StSc+Cu*) from *Sugar* (3 km; 10 km) via *Gravel* (5 km; 27 km) to *Fish* (7 km; 63 km) and *Flowers* (11 km; 55 km). The 95th percentile is given in brackets as this can be assumed to better capture the characteristic length of the stratiform cloud decks by excluding very small entities and entities whose path length is much smaller than the actual characteristic length as the observatory rarely samples the clouds at their characteristic cross section. The translation from cloud entity length in time to space has been done by using the wind speed at cloud height measured by the nearest sounding.

C.3.2 Rainfall

To assess, how important precipitation might be for different patterns, we characterize its frequency and strength in the following. From the example time-series shown in Fig. C.2 and the results from the previous section, we expect a clear separation of the precipitation characteristics among the patterns: from the lack of rain during the occurrence of *Sugar*, to frequent showers in the case of *Gravel*, to yet stronger rain events for *Fish*.

To test this expectation, we analyze at the precipitation measurements from the BCO within the same 6-hour time windows used in the section above. First, we quantify how many analysis windows contain any rain event. With the exception of *Sugar*, in more than 50% of the identified cases, rain is present. For *Sugar* precipitation can be detected in only 35% of the cases.

This absence of rain events in case of *Sugar* is even more evident in the quantification of the mean near-surface rainfall (Fig. C.5a). Rain amounts are similar for *Flowers* and *Gravel*, consistent with the frequency of near surface echoes evident in Figs. C.3 and C.4, nearly twice as large for *Fish*. We also quantify rain intensity by averaging the maximum rain-rates within each analysis window for each of the patterns. Among the precipitating patterns rain intensities do not differ as substantially. In all of these cases the precipitation is intense (approach 10 cm a day), and well above the threshold (1 mm h^{-1} to 2 mm h^{-1})

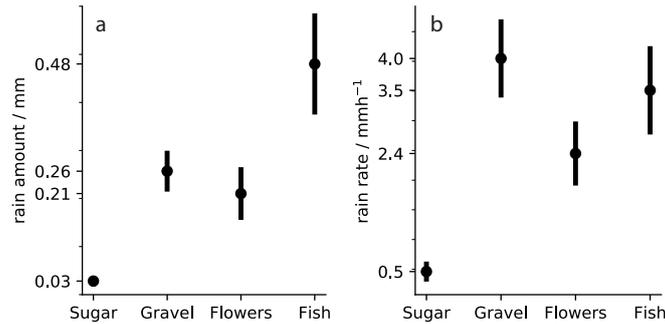


Figure C.5: Rain statistics of each pattern averaged over a 6-hour period. The average rain amount (left) and the average maximum rain rate of each window (right) are shown with their standard error.

that past studies have associated with the formation of cold-pools (Barnes and Garstang, 1982; Drager and Heever, 2017).

By applying the threshold of 1 mm h^{-1} to the maximum rain events, the number of cases with significant rainfall decreases to 12% in case of *Sugar* and about 35% for the other patterns. 35% might not seem to be a lot, but it has to be kept in mind that these patterns are of meso-scale extent and even a 6 h-period cannot capture the complete variability. This is especially the case for the *Fish* pattern, where a 6 h-period might only capture the clear-sky part of the *Fish* and therefore the importance of precipitation may be underestimated.

Our data do not contain sufficient samples to evaluate to what extent the spatio-temporal characteristics of precipitation differs among the patterns. However, by analyzing the precipitation signature in all 138 6-hourly windows of *Fish*, there is evidence of a bimodal distribution of rain events, with a second mode consisting of extended periods of precipitation (like the one shown in Fig. C.2) that is not evident for either the case of *Flowers* or *Gravel*.

Overall, precipitation events of significant strength occur during *Gravel*, *Flowers* and *Fish* periods and suggest that precipitation plays a role in the patterning process, or at least in the persistence of these patterns. There is no hint that precipitation is important for *Sugar*.

C.4 METEOROLOGICAL ENVIRONMENT

In the previous section we characterized similarities and differences in cloud- and precipitation-signatures among the four patterns. To the extent the patterns are forced, this forcing might be evident in the local meteorological setting. In this section, we address this possibility and investigate the meteorological settings, first at the surface and then within the free troposphere, for the different patterns.

C.4.1 Surface measurements

Near surface (5 m and 25 m above mean sea level) meteorological measurements at the BCO are composited by pattern in Fig. C.6. Common to all variables shown is a distinguished value for at least one of the patterns.

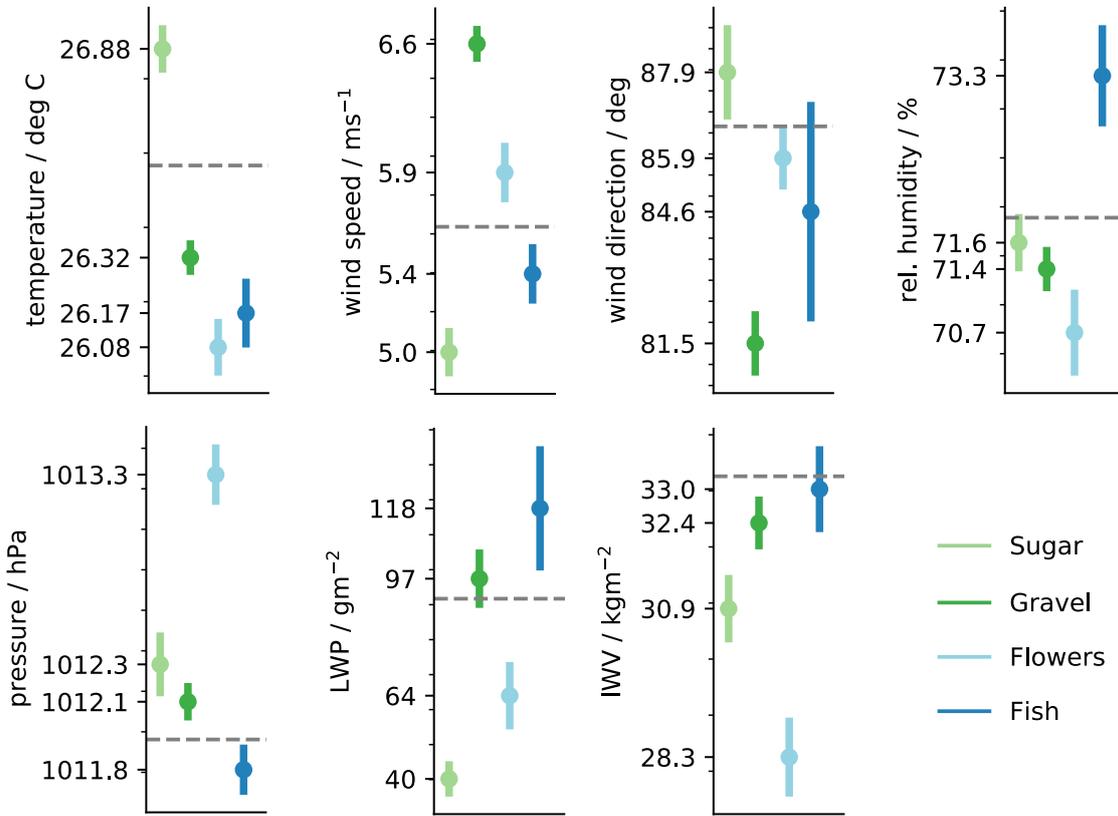


Figure C.6: Surface meteorology measured at the BCO during the observation of the four patterns. The seasonal mean of the observed time-period independent of any pattern is drawn as grey line.

Sugar distinguishes itself from other patterns by virtue of its mean temperature and low wind-speed. Given that annual cycle of surface temperatures is just over 2K this 0.6 K difference is large. While *Sugar* is associated with unseasonably warm conditions, the contrast with the other patterns is due in equal part to them being unseasonably cool, which is consistent with *Sugar* arising during periods with little northerly contribution to the mean flow. *Gravel* is distinguished by the surface winds being unseasonably strong and northerly. *Flowers* are found on the days that are coolest, when surface winds are strong, but not so strong as for *Gravel*, and when the surface pressure is unseasonably high (1013.3 hPa). In contrast *Fish* which are also associated with extensive cloud coverage (Fig. C.3) are found on days with unseasonably low pressure (1011.8 hPa), high humidity and relatively low but quite variable (in terms of direction) winds, consistent with more disturbed conditions and extended periods of precipitation.

Amongst all analyzed surface observations, wind speed is the best proxy for a specific pattern. The lowest mean wind is measured during *Sugar* situations with 5 m s^{-1} . For the other patterns the mean wind speed increases by an increment of 0.5 m s^{-1} from *Fish* to

Flowers to *Gravel*. The finding that *Flowers* and *Gravel* occur in conditions of higher winds is consistent with what was found by Bony et al. (2020), but further discriminates among all patterns rather than two groupings (e.g., *Flowers* and *Gravel* as high wind-speed and *Sugar* and *Fish* as low wind-speed patterns). This suggests that there may be processes that are not captured by the reanalysis, especially so as we gain similar results to Bony et al. (2020) when compositing the ERA5 surface data (Fig. C.18)

C.4.2 Vertical structure

The previous analysis is extended in the vertical through a composite analysis of the Grantley Adams radiosonde data. The pattern mean-soundings, and their associated uncertainty estimates, are presented in Fig. C.7. Composites are made of the equivalent potential temperature, potential temperature difference ($\Theta - \bar{\Theta}$, where $\bar{\Theta}$ is the mean sounding across all patterns), relative humidity and wind speed.

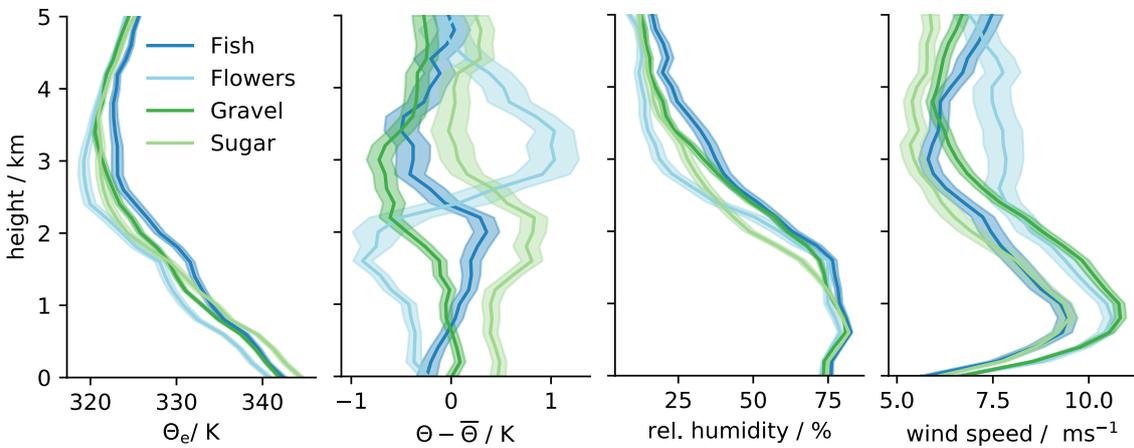


Figure C.7: Average profiles of equivalent potential temperature, potential temperature difference to the overall pattern mean, relative humidity and wind speed from soundings at the Grantley Adams Airport (standard error is shaded).

Surface temperature differences measured at the BCO are also evident in the soundings, and extend through the depth of the moist (lower 3 km) layer. *Flowers* distinguish themselves not only by lower surface temperatures, but also by a much stronger stratification atop the humid layer, showing a strong inversion at about 2.5 km. *Sugar* appears associated with a much shallower cloud layer, also capped by an inversion. The apparent instability (decrease in $\Theta - \bar{\Theta}$ with height) for the other patterns simply indicates that they are less stable on average. The lower-tropospheric stability (LTS) is $16.2 \pm 0.3\text{K}$ in case of *Flowers* and nearly 2 K lower for *Fish* ($14.7 \pm 0.2\text{K}$), *Sugar* ($14.4 \pm 0.2\text{K}$) and *Gravel* ($14.1 \pm 0.1\text{K}$). However, in case of *Sugar*, the value of Θ at 700 hPa (which is used to construct LTS) may miss the shallow stable layer that appears to cap the convective development of this pattern.

The relative humidity profile is strongly coupled to the convective activity and hence the echo fraction (e.g., Fig. C.3) As we have shown in the last section, *Sugar* is mostly characterized by cloudiness at cloud-base height with few clouds reaching up to 1.8 km. In agreement, the according moisture profile shows a shallower layer compared to the

other patterns that more regularly reach the inversion height and distribute moisture. Likewise *Fish*, with echos reaching more deeply through the lower troposphere is also considerably moister than the other patterns above 3 km. These humidity profiles also help explain differences in θ_e , particularly in the upper cloud layer and lower free-troposphere. For example as seen by contrasting *Fish* and *Gravel*.

Based on measurements made during RICO, Nuijens et al. (2009) analyzed differences in θ_e similar to those shown in Fig. C.7. Consistent with their findings these profiles seem to co-vary consistently with surface wind speeds. Stronger surface winds for *Gravel* and *Flowers* are mostly confined to the moist layer for *Gravel*, but extend through the lower troposphere for *Flowers*. These winds are one component of what is often thought of as an externally imposed large-scale forcing, to which the boundary layer thermodynamic profiles relatively quickly equilibrate.

C.4.3 Large-scale forcing

We use ERA5 data in a $20^\circ \times 20^\circ$ domain centered around each ABI classification to examine how the large-scale subsidence (ω_{500}) varies as a function of pattern. Those domains are afterwards averaged to one composite that shows the strength of subsidence at the center of each pattern, but also in its surrounding.

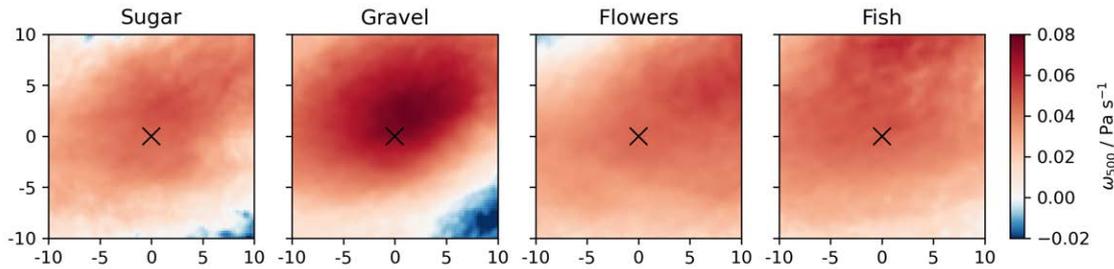


Figure C.8: Distribution of subsidence strength ω_{500} relative to identified pattern centers composed by 20×20 degree domains around each identified pattern. Pattern centers are marked with a cross.

Fig. C.8 reveals that all patterns occur during times of subsidence and that this subsidence is in most cases also similar to the typical subsidence rate of 0.05 hPa s^{-1} in the Atlantic trade-wind regime (Holland and Rasmusson, 1973). However, it also shows that some variability in the large-scale forcing exists and stronger subsidence is, contrary to expectation, not occurring during *Sugar* and *Flowers* cases, but rather during *Gravel* cases (Tab. C.3).

Table C.3: Large-scale forcing averaged by pattern from fixed-location sounding data (snd) and ERA5 data from pattern center

Pattern	LTS _{snd}	LTS _{ERA5}	$\omega_{500, \text{ERA5}}$	freq. of convergence
Sugar	14.3 K	14.8 K	0.046 Pa s^{-1}	28%
Gravel	14.0 K	14.4 K	0.072 Pa s^{-1}	38%
Flowers	16.2 K	16.6 K	0.046 Pa s^{-1}	34%
Fish	14.6 K	16.0 K	0.048 Pa s^{-1}	59%

In the subtropics, particularly in association with stratocumulus, subsidence co-varies positively with LTS. On shorter time scales and deeper in the tropics, other factors may play a role. In particular the temperatures above the cloud layer are tightly coupled to moisture, so as to homogenize the density temperature on isobaric surfaces. This partly explains the stronger temperature inversion for *Flowers*. It also means that boundary layer variability may play a more important role in determining the LTS, consistent with near-surface temperature differences as illustrated in Fig. C.6.

C.5 ARE THE FOUR PATTERNS INDICATIVE OF SPECIFIC AIR MASSES?

Rasp et al. (2020) showed that globally, the four patterns predominate in the dry tropics, regions often associated with the trade winds. The analysis in the previous section identified subtle differences in the environments in which the four patterns form. This raises the question as to the origin of these environmental differences, i.e., to what extent they arise from subtle variations within the trades, or what one might alternatively think of as disturbances to or departures from canonical trade-wind conditions. We explore this question by analyzing the seasonal cycle of the four patterns within our North Atlantic study region as well as the air-mass histories of the different patterns by compositing reanalysis data along back-trajectories constructed from that same data.

c.5.1 Seasonality

Considering just the region of the downstream trades, taken to be the tropical North Atlantic west of 45°W, most patterns predominate in the boreal winter trades as shown by Fig. C.9. *Fish* and *Gravel* seem only to occur in this region in conditions (DJF) when the trades are well developed. *Flowers* are also present in boreal spring and early summer (AMJ). *Sugar* shows very little seasonality. Rather, and consistent with the analysis by Rasp et al. (2020), it appears associated with suppressed conditions bordering the ITCZ whose seasonal migration it follows. Based on this we hesitate to call *Sugar* a trade-wind cloud pattern.

Flowers are even more common in the ‘upper’ trades (east of 45°W), even more so in the April-June period, (e.g., Fig. C.9). Such a distribution is consistent with an affinity for conditions that favor stratocumulus. This distribution is in agreement with the analysis in the previous section, which showed that *Flowers* favor conditions of higher lower tropospheric stability, and lower surface temperatures, as compared to the other patterns. This supports the idea that *Flowers* are the downstream manifestation of the familiar, but much smaller, closed cellular stratocumulus (Stevens et al., 2020); alternatively, it may be indicative of a failing of the neural network in the upper trades because it has not been trained to distinguish between the very similar looking *Flowers* and closed-cells.

c.5.2 Lagrangian evolution of air masses by meso-scale organization

Here we use the back-trajectories, initialized at the center of the MODIS AQUA classifications following the boundary layer winds at 925 hPa for 84 h, to investigate possible reasons for the environmental differences associated with each pattern as described in Section C.2.

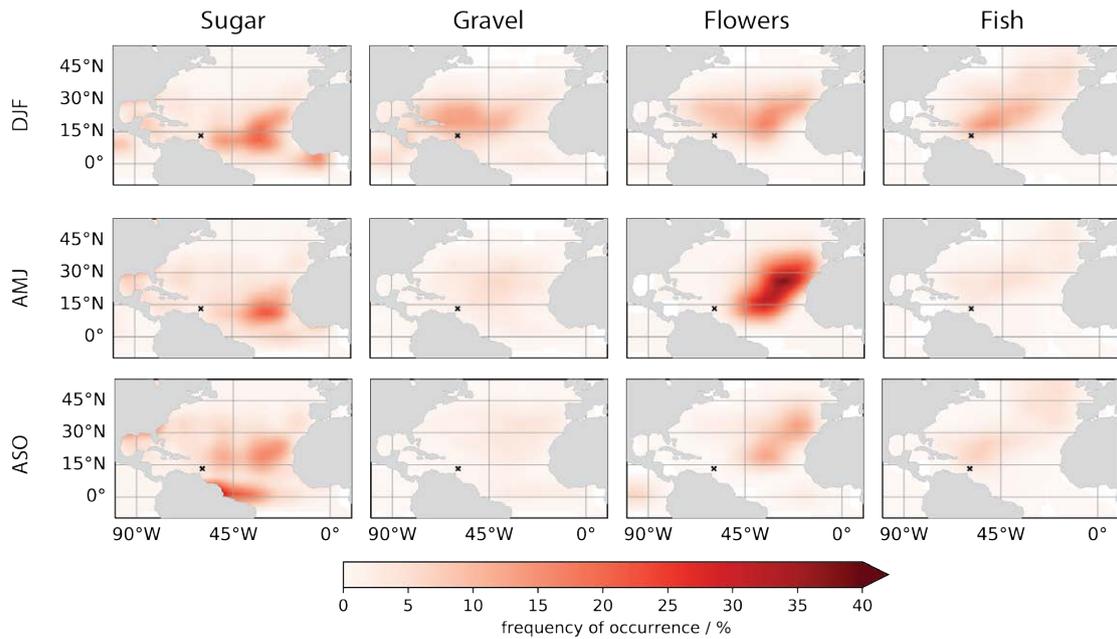


Figure C.9: Seasonal distribution of patterns in the North Atlantic in the dry- (DJF), transitional- (AMJ) and wet- (ASO) season (top to bottom) detected in infrared imagery (AQUA MODIS 2010-2020).

Fig. C.10 shows that the back-trajectories are consistent with the steadiness that characterizes the winter trades, with the trajectories aligning well along the general flow of the trades as they come to their point of initialization. They do however differentiate themselves as one follows their history back in time. Most notably *Fish* which originates far to the west of the other trajectories. A more tropical influence on *Sugar* is also consistent with its back-trajectories which start furthest south.

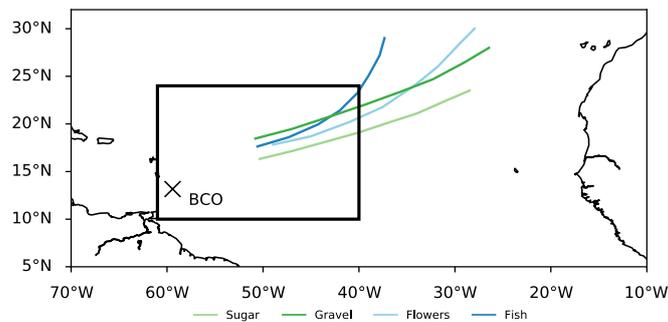


Figure C.10: Mean back-trajectories for the different patterns initialized at the center of individual classifications within the indicated black box at 925 hPa and calculated for 84 h.

Compositing the large-scale conditions on the trajectories shows that many of the environmental differences previously documented are apparent well in advance (and upstream) of where the pattern was eventually identified (Fig. C.11). *Sugar* has warmer sea-surface temperatures, weaker winds and a relatively moist free-troposphere along its entire back-trajectory, consistent with a more tropical influence. *Flowers* evolve over cold ocean temperatures throughout the trajectory paired with persistently high LTS (despite

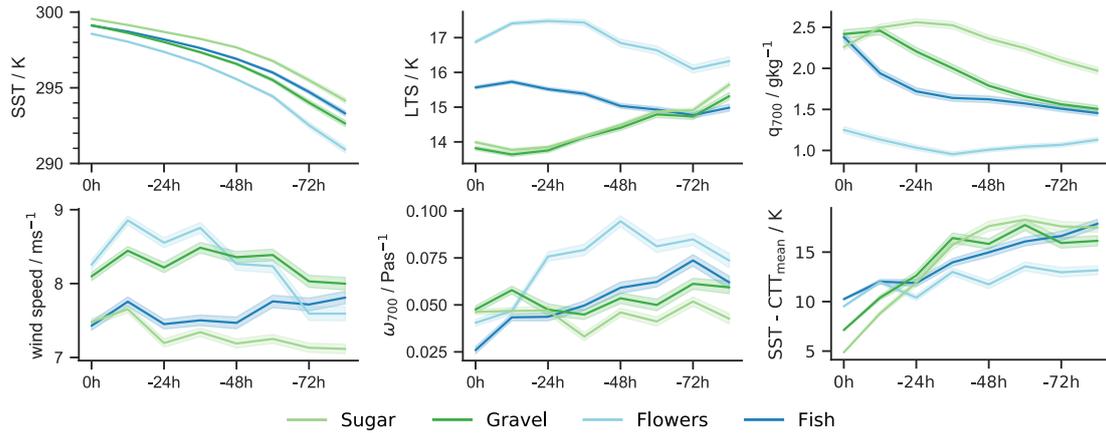


Figure C.11: Environmental conditions along the back-trajectory of air-masses before time of pattern detection. All values are ERA5 reanalysis properties, except the cloud top height estimate where the cloud top temperature (CTT) is sourced from MODIS AQUA observations. Shading indicates standard error.

rising SSTs), a dry free troposphere and stronger low-level winds. And differences in LTS among the patterns are robust and in place already 48 h earlier.

The time-evolution of different fields is also indicative of dynamic influences. For instance, for *Flowers* an acceleration of the low-level winds between -24 h to -84 h may be driving the strong subsidence at 700 hPa, which in turn would support the already anomalously dry free-troposphere to dry further and increase the LTS. This pattern preceding process may drive the differences between *Flowers* from *Gravel* with the slight slackening of the winds and the decrease of the subsidence nearer the time and place where the pattern is identified, playing less of a role. In contrast, for *Fish* a strong temporal evolution within the last 24 h, as manifest through a moistening of the lower troposphere, might be indicative of a dynamic disturbance. *Sugar* seems less representative of a sudden stalling in the winds in association with *local* suppression, if anything recovering from more suppressed conditions and weaker winds upstream.

The time-evolution of cloud top height, estimated as the difference between the ERA5 sea surface temperature and the mean cloud top temperature sourced from MODIS within 100 km around the trajectory sampling point, can further be an indication of different lifetimes of the patterns. *Sugar* and *Gravel* seem to set up only shortly before the detection (-36 h) when the cloud top height dropped quickly, which would be indicative of a shorter lifetime. In contrast, *Flowers* and *Fish* might have persisted longer because the cloud top height evolves only little. It has to be kept in mind that this estimation of cloud depth is sensitive to high clouds that are filtered by the nature of the classification at the time of detection, but likely influence the CTTs further upstream. The changes of cloud depth might therefore appear magnified.

The diurnality that is pronounced in a number of fields (wind speed, ω_{700} , SST-CTT) is explored in more depth by Vial et al. (2021).

C.5.3 Extra-tropical disturbances

Although we focus on the northern hemispheric winter season where the trades are well formed, disturbances to the trade-wind mean flow are well documented (Bunker et al., 1949; Riehl, 1945). The earlier literature identifies two types of disturbances, one associated with anomalously deep easterlies (which might be associated with active deep convection), the other associated with extra-tropical intrusions in the form of trailing cold-fronts from extra-tropical cyclones.

The older literature conceptualized the later (extra-tropical intrusions) as tropical incursions of the ‘polar front’. Especially in the boreal winter, when the Intertropical Convergence Zone is further south and the Azores high is less well established, frontal disturbances can extend equator-ward. By the time they reach the subtropics, their temperature signature is muted and they become most pronounced in the form of a shear line that separates the light easterlies from the stronger north-easterlies (Riehl, 1945).

Such a frontal passage can be seen in the surface analysis charts e.g., in association with a deepening cyclone over the mid-Atlantic (near 45°N and 45°W) on 25th December 2018. Through the course of six days the cold front, initially supported by the outflow of cold-continental air (a cold air outbreak) from the east-coast of North-America, occludes upon reaching the tropics as far south as Barbados (see Fig. C.12).

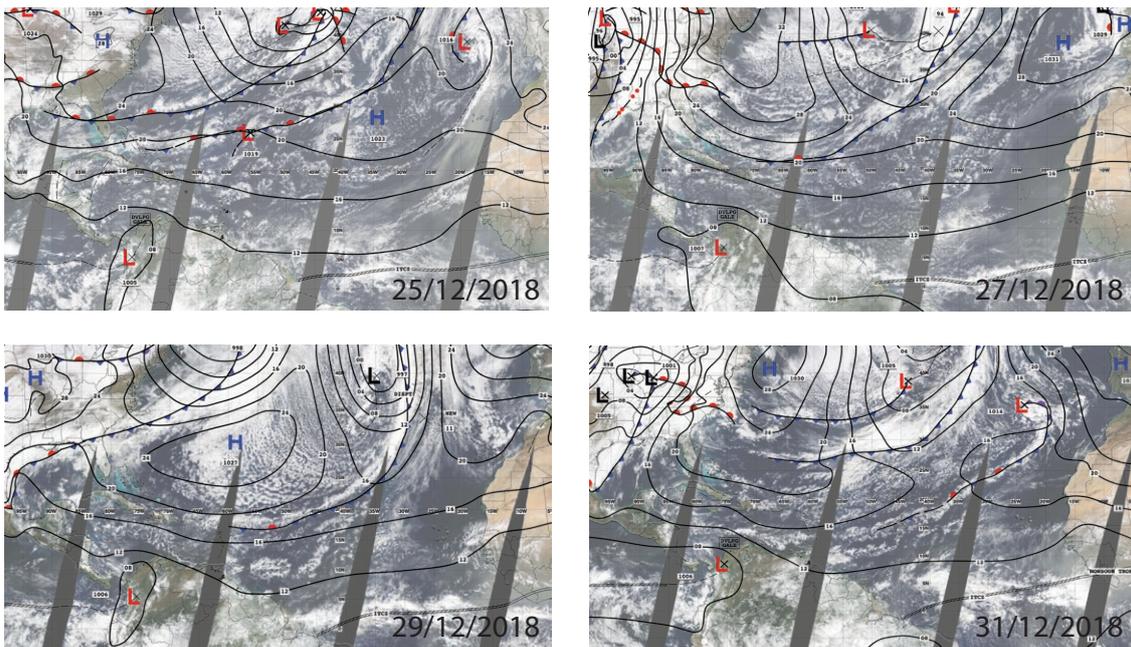


Figure C.12: Cold air outbreak between 25th of December 2018 and 31th December 2018. Surface analysis charts from the National Hurricane Center are adapted and overlaid on Terra MODIS images

By comparing the surface analysis chart with the satellite image, we recognise that the front is visible as a band of convection that we would classify as *Fish*. In the former cold sector, just north of the front, one can also notice on the 29th December cloud structures north of the front similar to *Flowers* (more pronounced to the west) or perhaps *Gravel*.

We repeat our composite analyses done to create Fig. C.8 with surface convergence to test, whether the frontal character is typical for *Fish* and whether other patterns can be related to the fronts as well. We find a signal of strong convergence ($> 1 \times 10^{-6} \text{ s}^{-1}$) connected with *Fish* about 60% of the identified cases (Fig. C.13). A clear signature for the other patterns is less pronounced, although the absence of convergence for *Sugar* is consistent with it being more locally suppressed. For *Fish*, the pattern of convergence extends zonally in a way that supports the hypothesis of *Fish* arising in association with disturbances associated with trailing cold-fronts or shear-lines from extra-tropical intrusions.

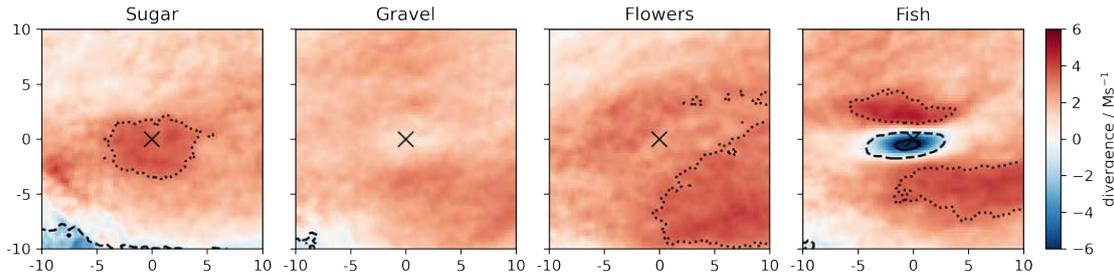


Figure C.13: Distribution of divergence at 950 hPa relative to identified pattern centers composited by 20x20 degree domains around each identified pattern. Contours indicate frequency of events with convergence larger $1 \times 10^{-6} \text{ s}^{-1}$ (30%: dotted, 45%: dashed, 60%: solid). Pattern centers are marked with a cross.

C.6 DISCUSSION AND CONCLUSION

Cloud- and environmental properties associated with four patterns of meso-scale organization in the lower trades (50°W to 60°W) of the North Atlantic are examined. The four patterns follow the *Sugar*, *Gravel*, *Fish*, *Flowers* taxonomy of Stevens et al. (2020) and are identified using a neural-network applied to high-resolution infra-red imagery from the GOES-16 and AQUA satellites.

We conditioned surface observations, back-trajectories, and reanalysis data on the identification of different patterns to answer three questions: One, do the four patterns show physical differences in the cloud geometry as seen by surface-based remote sensing? Two, can differences in the large-scale environment associated with different patterns be discerned? And, three can we identify the origins of discernible environmental differences among the patterns.

Fig. C.14 summarizes these results and illustrates that the four patterns differ in more than just their satellite presentation. Cloud coverage and its vertical distribution differ and differences in the environment of different patterns are discernible. The thermodynamic profiles in Fig. C.14 show inter-pattern differences, but also intra-pattern differences as measured by radiosondes at points whose position relative to other features within a pattern is schematized.

Many preconceptions from earlier studies, either inferred from snapshots (Stevens et al., 2020) or from compositing reanalysis data on values of a cloud-clustering index that correlate with different patterns (Bony et al., 2020), are supported by our analysis. As an example, *Flowers*, and to some extent *Fish*, have a stratiform component detectable

from surface-based remote sensing. In the latter this is less distinctly a capping stratiform, or stratocumulus layer, as it is associated with more cloudiness throughout the cloud layer. Compared to the mean conditions, or the other patterns, LTS is higher (0.5 K to 1.0 K) for *Fish* and (2 K) for *Flowers*.

Non-precipitating cloud coverage at the lifting condensation level, as emphasized by Nuijens et al. (2014) for the entirety of trade-wind cloudiness, also holds across the four patterns. This came as a surprise given that *Flowers* and *Fish* are characterized in part by their cloud free areas. Differences in cloud-base echo fraction largely reflect differences in precipitation, suggesting that to the extent environmental conditions demand an increase in the mass flux out of the sub-cloud layer, for instance as shown by George et al. (2020), this is largely associated with the development of deeper clouds and precipitation.

Similar to what was found by Bony et al. (2020), near surface winds identify *Flowers* and *Gravel* with strong near-surface winds, and *Fish* and *Sugar* with light winds. Our analysis, further discriminates within these two groups, with *Sugar*, *Fish*, *Flowers* and *Gravel* each being separated by a roughly 0.5 m s^{-1} increase in surface wind speeds. Precipitation increases with near surface winds, as previously noted for measurements during RICO (Nuijens et al., 2009), with *Fish* being an outlier whose large rain rates are associated with extra-tropical disturbances and anomalous low-level convergence.

Seasonal variations and back-trajectories provide further insight into the origin of differences in the environments of the different patterns. The view of trade-wind clouds as cumulus humilis, and hence non-precipitating with little vertical extent, as popularized by studies based on data from BOMEX (Siebesma and Cuijpers, 1995) and most closely associated with *Sugar* suggests that these are at least in the form of large-scale homogeneous areas rather uncharacteristic of the trades. *Sugar* is found to favor more suppressed conditions, uncharacteristically (for the trades) weak winds, and proximity to deeper convection in the ITCZ.

As a historical note, the third author recalls that when the large-eddy simulation community began focusing on shallow trade-wind convection through simulations of conditions derived from BOMEX data (Siebesma et al., 2003), Bruce Albrecht admonished us that less suppressed and more stratiform capped conditions – as for instance seen and simulated in association with the Atlantic Trade-Wind Experiment Stevens et al., 2001, ATEX and which we might today call *Flowers*– were more characteristic of the trades. We find confirmation for his point of view, twenty years later, in our data. Given the association of *Fish* with shear lines from remnant extra-tropical cold fronts intruding deep into the sub-tropics, only *Gravel* is left to add to *Flowers* as an archetypical form of trade-wind convection. *Fish* and *Sugar* are intruders.

Gravel and *Flowers* differ substantially in their cloud amounts (as seen here) and their cloud radiative effects, as shown by Bony et al. (2020). We attribute this to *Flowers* forming in conditions of weaker winds, and a drier and warmer free troposphere. These differences to *Gravel* would support a more pronounced capping inversion, and stronger boundary layer cooling. Based on back-trajectories we hypothesize that these conditions arise from an acceleration of the trades and stronger subsidence in the upstream flow along *Flowers* back-trajectories. This hypothesis lends itself well to tests with LES, and may even be evident at the somewhat coarser resolution now being simulated by a new generation of global storm-resolving models (Satoh et al., 2019).

Independent of the formation mechanism, understanding of the conditions favoring one or the other pattern may help anticipate to what extent climate change, by virtue of changes in wind-speeds, or the frequency of extra-tropical disturbances, or changes in the opacity and stability of the free troposphere, will affect the frequency of occurrence for different patterns, and thus cloud-radiative effects in the lower trades. The widening of the tropics (Seidel et al., 2008) and the poleward shift of the extra-tropical storm-tracks (Ulbrich et al., 2008; Yin, 2005) would, following our analysis, disfavor *Fish* and *Flowers* patterns in favor of *Sugar* with a much smaller cloud fraction, and less pronounced cloud radiative effects.

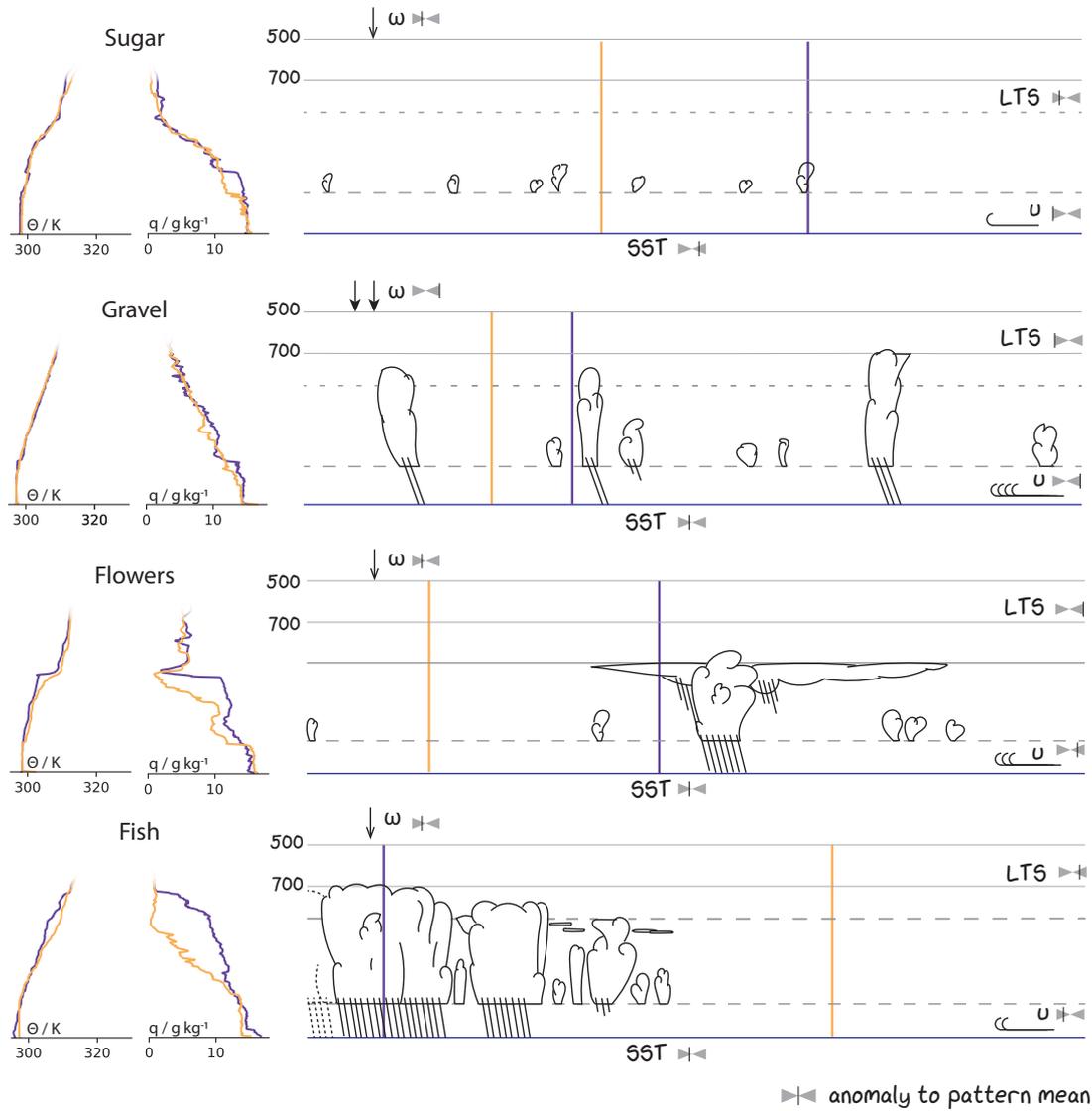


Figure C.14: Illustration of the cloud field during the four patterns of meso-scale organization and the associated large-scale forcing (right) including the thermodynamic profiles (left). The anomaly in forcing to the pattern mean is indicated by grey sliders. Vertical lines indicate the contrasting positions of the thermodynamic profiles, purple being in the moist part and orange in the dry area. Thermodynamic profiles are based on soundings during the EUREC⁴A field campaign (Stephan et al., 2020b)

C.7 OPEN RESEARCH

Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the authors' work can be obtained from <https://doi.org/10.5281/zenodo.4767674>. The ERA5 datasets used in this study (Hersbach, H. et al., 2018) have been provided by the Climate Data Store. GOES-16 Advanced Baseline Imager Level 1b radiances are available at <https://doi.org/10.7289/V5BV7DSR> and were converted with (Raspaud et al., 2019) to brightness temperatures. MODIS imagery originates from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS).

C.8 ACKNOWLEDGMENTS

We thank the members of the Tropical Cloud Observation group for maintaining the Barbados Cloud Observatory and fruitful discussions. We would also like to thank Jiawei Bao for helpful comments on an earlier version of this manuscript. We thank Minghua Zhang for his editorial contributions and three anonymous reviewers for their constructive feedback.

C.9 SUPPORTING INFORMATION

C.9.1 Agreement between neural network classifications

To compare our results with the earlier study of Rasp et al. (2020) we apply their neural network to visible MODIS-AQUA images of the time-period where all classification datasets overlap: 2017/2018 (JFM), 2018/2019 (NDJFM) and 2019/2020 (NDJFM). An overview of these different classification datasets is shown in Fig. C.15 for four arbitrarily chosen days.

A more quantitative way to compare the classifications, is to use the Intersection over Union (I) score, which is commonly used as a measure of agreement, with $I = 0$ being no agreement and $I = 1$ if identical. To account for the overlap of labels, we define the mean pattern specific Intersection over Union between two classification datasets (c_1, c_2) as

$$I_{c_1, c_2, p} = \frac{1}{S} \sum_{s=1}^S \frac{L_{c_1, s, p} \cap L_{c_2, s, p}}{L_{c_1, s, p} \cup L_{c_2, s, p}}, \quad \text{with } L_{c_x, s, p} = \bigcup_{i=1}^n l_{c_x, s, p, i}. \quad (\text{C.1})$$

First, the union of individual labels l for a certain pattern p of a particular dataset (c_x) and satellite scene s is calculated to remove the intra-pattern overlap that is common for object-detection algorithms and can be seen here e.g. on January 12, 2020, where two classifications of Gravel overlap in the visible classification dataset (Fig. C.15). Second, pairs of these unions ($L_{c_1, s, p}, L_{c_2, s, p}$) from two datasets are then used to calculate the Intersection over Union for a specific scene and pattern. $I_{c_1, c_2, p}$ is finally the pattern averaged Intersection over Union.

For each label and each combination of the classification datasets the agreement is shown in Fig. C.16. In cases where both approaches lead to a classification, I is between 0.5 and 0.7 (grey bars). Taking all labels into account, also those that are missing in one of

the classification datasets, reduces the agreement to the range of 0.3 to 0.6. This is mostly due to the number of scenes that remained unclassified in the MODIS_{VIS} dataset. While this agreement might seem to be low, it has to be kept in mind that these patterns are not as well defined as objects traditionally identified with neural networks. This has also been shown by (Rasp et al., 2020) where the inter-human agreement is on average below 0.3. As expected, the agreement between the infrared datasets is in generally higher as any comparison with the visible dataset. Part of the differences might be due to different observation times and the slightly different wavelengths.

A general agreement between the classifications can also be seen in Fig. C.17. The mean area fraction a pattern classification covers is relatively constant with classifications of *Sugar* being on average the smallest.

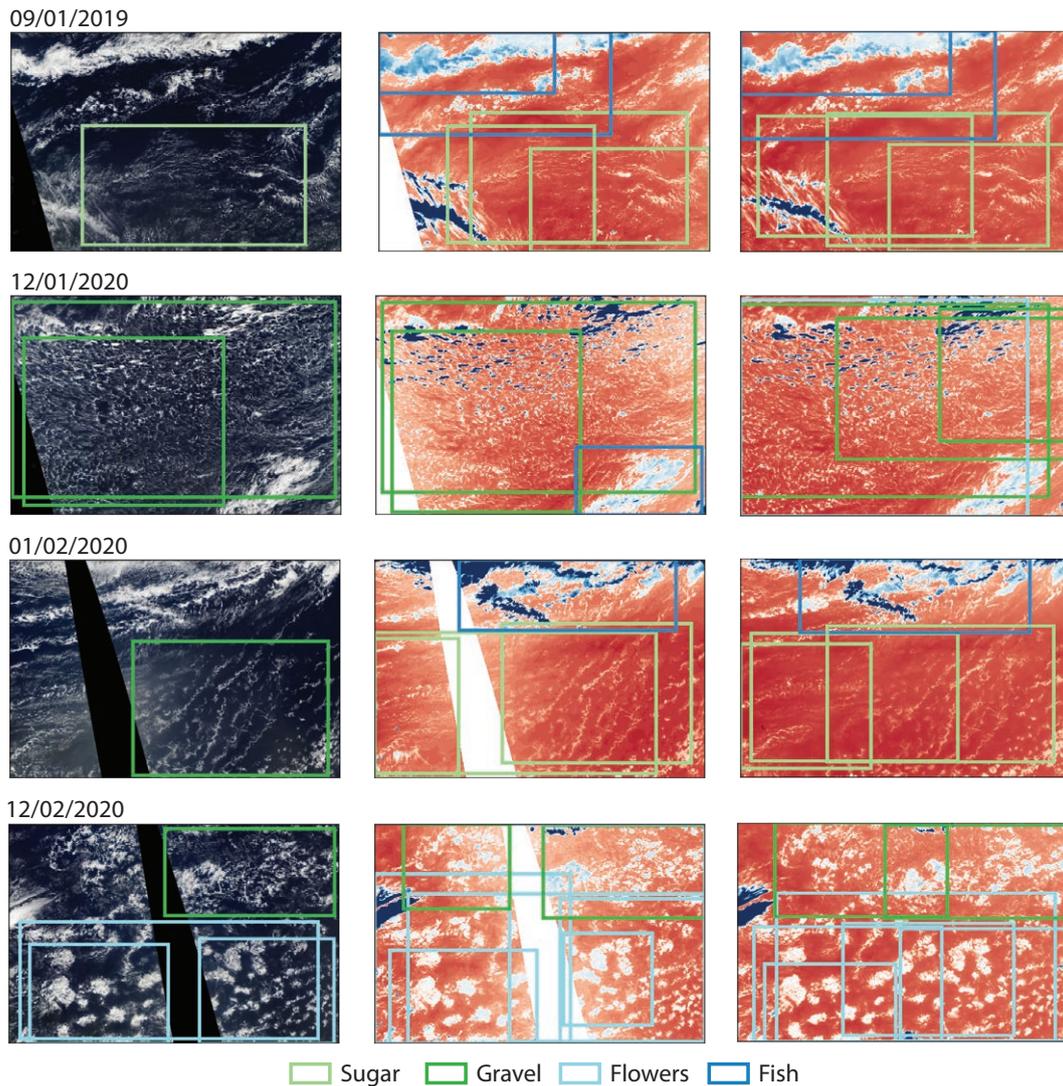


Figure C.15: Example of labels for the different classification datasets (left to right): MODIS_{VIS} , MODIS_{IR} , ABI_{IR} .

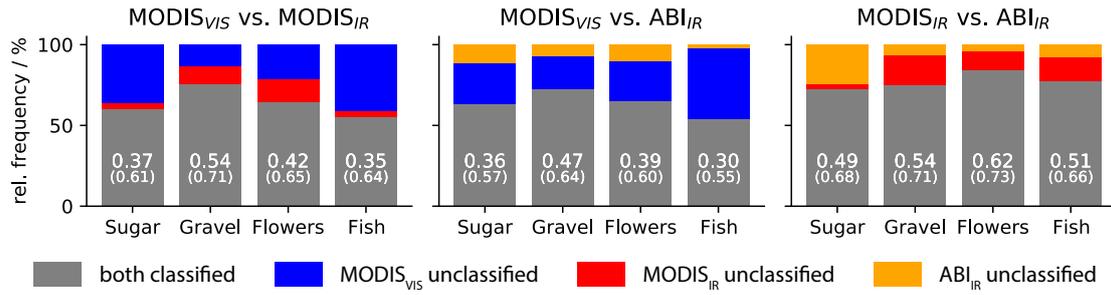


Figure C.16: Agreement between the classification datasets based on the Intersection over Union (I) score. Colored bars quantify the influence of missing labels ($I = 0$) on the averaged I . The I excluding partially unclassified scenes is given in brackets.

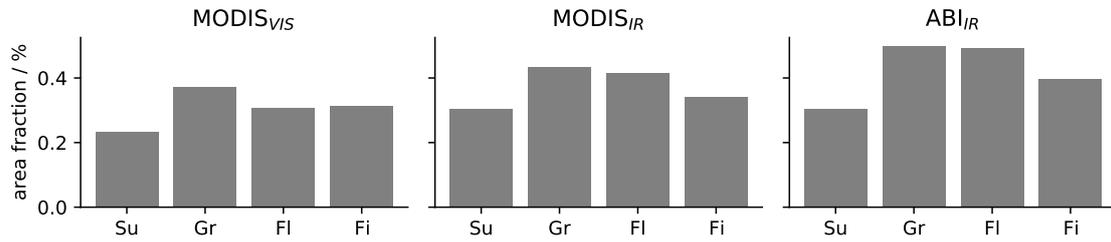


Figure C.17: Comparison of mean area fractions between the different classifications for the DJF season.

c.9.2 ERA5 surface meteorology

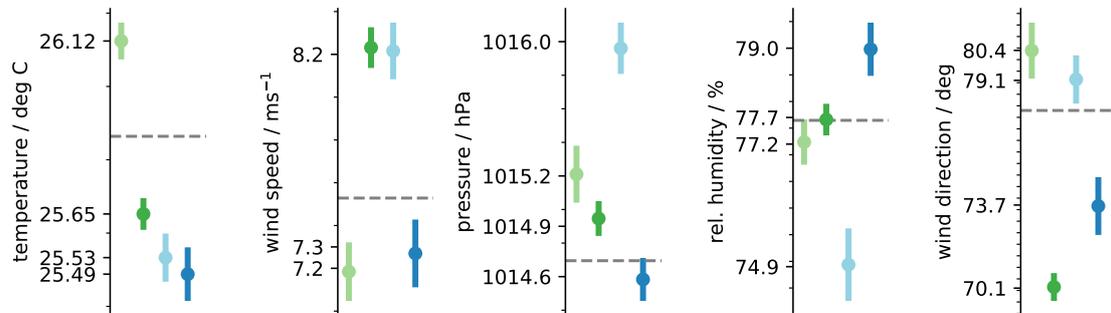


Figure C.18: Like Fig. 6 but with ERA5 data slightly east of Barbados (13.25N, -59.25 E)

ON THE DAILY CYCLE OF MESOSCALE CLOUD ORGANIZATION IN THE WINTER TRADES

The work in this appendix has been published as:

Vial, Jessica, Vogel, Raphaela, and **Schulz, Hauke** (2021). On the Daily Cycle of Mesoscale Cloud Organization in the Winter Trades. *Quarterly Journal of the Royal Meteorological Society* 147:738, pp. 2850–2873. DOI: [10.1002/qj.4103](https://doi.org/10.1002/qj.4103)

The contributions of the authors to this publication are as follows:

The idea together with a proof of concept of detecting the diurnal cycle of patterns of shallow convection with a neural network, the neural network design itself and its application have been performed by HS. All authors contributed to the analysis. JV shaped the paper and wrote with input from all co-authors the manuscript. All authors iterated on the manuscript and approved the final manuscript.

On the daily cycle of mesoscale cloud organization in the winter trades

Jessica Vial¹, Raphaela Vogel¹, Hauke Schulz²

¹LMD/IPSL, Sorbonne Université, CNRS, Paris, France

²Max Planck Institute for Meteorology, Hamburg, Germany

ABSTRACT

How spatial organization of clouds at the mesoscale contributes to the daily cycle of shallow cumulus clouds and precipitation is here explored, for the first time, using three years of high-frequency satellite- and ground-based observations. We focus on the four prominent patterns of cloud organization – Sugar, Gravel, Flowers and Fish – which were found recently to characterize well the variability of the North Atlantic winter trades. Our analysis is based on a simple framework to disentangle the parts of the daily cycle of trade-wind cloudiness that are due to changes in (i) the occurrence frequency of patterns and (ii) cloud cover for a given pattern. Our investigation reveals that the contribution of mesoscale organization to the daily cycle in cloudiness is largely mediated by the frequency of pattern occurrence. All forms of mesoscale organization exhibit a pronounced daily cycle in their frequency of occurrence, with distinct 24-hour phasing. The patterns Fish and Sugar can be viewed as *daytime* patterns, with a frequency peak around noon for Fish and towards sunset for Sugar. The patterns Gravel and Flowers appear instead as *nighttime* patterns, with a peak occurrence around midnight for Gravel and before sunrise for Flowers. The cloud cover for a given pattern, however, always maximizes at nighttime (between 00 and 03 local times), regardless of the specific pattern. Analyses of the role of large-scale environmental conditions shows that the near-surface wind speed can explain a large part of the diurnal variability in pattern frequency and cloudiness.

D.1 INTRODUCTION

As one of the most fundamental modes of tropical climate variability, the daily cycle has been thoroughly studied for many cloud types except, surprisingly, for shallow cumuli in the trade-wind regime. Indeed, the daily cycle of trade-wind cumuli has only been recently described in some details by Vial et al., 2019, forty years after it was first documented (Brill and Albrecht, 1982; Nitta and Esbensen, 1974). In typical conditions of the North-Atlantic winter trades, cloudiness overall maximizes at the end of the night and is minimum in the afternoon. This daily cycle reflects the evolution of two distinct cloud populations: (i) a daytime population of non-precipitating small cumuli, which peaks around sunset and has weak vertical extents of only a few hundred meters above the Lifting Condensation Level (LCL), and (ii) a nighttime population of deeper precipitating clouds, peaking just before sunrise, which is often accompanied by a stratiform cloud shield spreading below the trade inversion (Vial et al., 2019). In storm-resolving and large-eddy simulations run over large domains, these two cloud populations seem to be associated with distinct spatial organizations at the mesoscale (see Figure 4 in Vial

et al., 2019). The daytime population exhibits a multitude of small cumuli scattered in a regular pattern over hundreds (or even thousands) of kilometers, while the nighttime field tends to show fewer and larger cloud clusters, sometimes organized along extensive lines or arcs reminiscent of gust fronts accompanying rainfall-generated cold pools (Ruppert and Johnson, 2016; Zuidema et al., 2011). Whether this visual impression from the simulations is realistic and systematic, and whether this type of spatial organization plays a role in the daily cycle of trade-wind clouds and convection, however, remain open questions.

It has long been recognized that shallow convection can organize spatially into various patterns; the most classical forms being cloud streets, closed or open cells in cold air outbreaks or in subtropical upwelling areas (Atkinson and Zhang, 1996; McCoy et al., 2017; Wood and Hartmann, 2006). In the North-Atlantic trades, different environmental conditions (e.g., warmer sea surface temperature, weaker subsidence) give rise to other forms of cloud organization, which have been recently discovered and characterized in satellite- and ground-based observations (Bony et al., 2020; Schulz et al., 2021; Stevens et al., 2020). Each organization pattern features a specific type of clouds and a specific spatial layout of the cloud field on scales from 20 km to 2000 km. The patterns range from isolated, shallow, non-precipitating cumulus clouds (*Sugar* organization), to precipitating cumuli forming along lines or arcs defining gust fronts (*Gravel* organization), and to organized structures of deeper precipitating cumuli with a stratiform cloud layer at their top that can extend up to hundreds of kilometres and that are separated by large and well defined cloud-free areas (*Fish* and *Flowers* organizations). These four patterns are illustrated in Figure D.1.

The fact that the daily cycle and mesoscale organization of trade-wind cumuli have only recently been discovered (or revived) clearly shows that there is a significant gap in our understanding of the dynamics of trade-wind shallow convection and clouds. This is of particular concern as the coupling between marine shallow trade-wind clouds and circulation is known to play a central role in the uncertainty of the tropical cloud feedback and climate sensitivity estimated by models (Bony et al., 2015; Brient and Bony, 2013; Rieck et al., 2012; Sherwood et al., 2014; Tomassini et al., 2015; Vial et al., 2016, 2017; Vogel et al., 2016). To help fill this knowledge gap, and finally, improve our ability to predict the Earth's climate response to warming, in this article, we build upon the work of Vial et al. (2019) to investigate more thoroughly the link between the daily cycle and the spatial organization of trade-wind shallow cumulus convection. To our knowledge, this study is the first to question the role of spatial organization in the daily cycle of shallow cumulus convection. Specifically, we explore whether the observed occurrence of the four aforementioned patterns of organization exhibits variability on the daily timescale, and whether the different patterns of organization influence differently the daily cycle of trade-wind cloudiness and precipitation.

Our analysis combines satellite- and ground-based remote sensing observations, as well as in-situ surface measurements in the North-Atlantic trade-wind region, windward of Barbados, which is known to be representative of the trade-wind regime in other ocean basins (Medeiros and Nuijens, 2016; Rasp et al., 2020; Stevens et al., 2016). We

focus on the boreal winter season, when the Intertropical Convergence Zone is at its southernmost position, and thus when shallow cumuli predominate. After a description of the observational datasets and analysis framework (Section D.2), we document the daily cycle in the occurrence frequency of mesoscale patterns (Section D.3), and investigate the different ways in which the daily cycle of trade-wind cloudiness and precipitation depend on the mesoscale patterns of organization (Section D.4). Finally, we explore the role of the large-scale environment in the variability of pattern occurrence and cloudiness at the daily timescale (Section D.5). Our conclusions are presented in Section D.6.

D.2 OBSERVATIONAL DATASETS AND METHODOLOGY

We use satellite- and ground-based remote sensing and in-situ observations over the tropical Atlantic Ocean around Barbados during the boreal winter months (DJFM) from 1 January 2018 to 31 March 2020. The different datasets and their use are described in the following subsections.

D.2.1 GOES-16 satellite data

The Geostationary Operational Environmental Satellite (GOES)-16 is the current satellite in the GOES-East location (centered at 75.2°W), providing data since December 2017. We use 30-minute infrared ($13\ \mu\text{m}$) brightness temperature (T_b) at a spatial resolution of 2 km from the Advanced Baseline Imager (ABI) Level 1b data product (GOES-R Calibration Working Group and GOES-R Series Program, 2017).

D.2.1.1 Cloud organization classification

Two different approaches are employed here to classify mesoscale patterns of shallow cloud organization using GOES-16 ABI data.

The first method, developed in Bony et al., 2020 and referred to here as the I_{org}/S method, characterizes the organization of a marine shallow cloud population within a fixed $10^\circ \times 10^\circ$ domain east of Barbados (48°W - 58°W , 10°N - 20°N) based on the mean size (S) and the clustering (I_{org}) of segmented cloud objects, which correspond to pixels for which $280\ \text{K} \leq T_b \leq 290\ \text{K}$. The index I_{org} was defined by Tompkins and Semie, 2017 such that $I_{org} < 0.5$ corresponds to a regularly distributed cloud population, $I_{org} = 0.5$ to a random distribution and $I_{org} > 0.5$ to a clustered distribution. The lower and upper terciles of S and I_{org} distributions are then used to classify the mesoscale patterns into four categories (Figures D.1a and D.10): Sugar is classified as high I_{org} and low S ; Gravel as low I_{org} and low S ; Fish as high I_{org} and high S ; and Flowers as low I_{org} and high S . The unclassified pattern (also referred to as the No category in the figures) is defined as all the times when I_{org} and S fall in the intermediate terciles. Refer to Bony et al., 2020 for more details.

The second method, called *NN* method, is based on the neural network (NN) algorithm originally developed and trained with visible satellite images in Rasp et al., 2020, and adapted to infrared images in Schulz et al., 2021. Rectangles of the four cloud organization classes (Sugar, Gravel, Flowers, Fish) are detected over a $14^\circ \times 21^\circ$ domain

including Barbados (9.3°N-23.3°N, 45°W-66°W), and the No category is considered as the remaining part of the domain where none of the four patterns is detected (see illustration in Figure D.1). Any number of pattern rectangles of various sizes can be detected at a given time (with a minimum rectangle size of about 10% of the domain area), with potential overlaps between them. When overlaps occur between several rectangles of the same pattern (e.g., 3 overlapping Gravel rectangles), we merge them into one polygon with the pattern area being the union of all overlapping rectangles; we thus only count the overlapping area once. Overlaps can also occur between different labels (e.g., 1 Fish overlaps with 1 Gravel rectangle) due to ambiguous forms of organization, or connectivity among patterns (Rasp et al., 2020; Stevens et al., 2020). This could either result from a weak machine learning prediction (which is affected by the quality of the human labels), or it could have a physical explanation (for instance, the overlaps could occur during transitions between patterns). Figure D.1b shows an example of an overlap between Gravel and Fish (on the top-right corner of the domain), since the Fish here appears as a network of Gravel-like cold-pool structures – which can be seen particularly well in the visible image from MODerate-resolution Imaging Spectroradiometer (MODIS) at the top. In these situations (i.e., of overlaps between different labels), we simply count the total area of all rectangles without removing any overlap. The total area of patterns (including the unclassified pattern) is thus greater than the domain area (we discuss this further in Section D.2.1.2).

These two classification methods are quite different in nature. I_{org}/S is based on geometrical and statistical properties of the cloud field, and as such could be considered as the most objective of the two approaches. However, since I_{org} and S are continuous measures and the patterns can only be robustly identified at the extremes (here, subjectively chosen first and third terciles) of the paired (I_{org} , S) distributions, the inner paired tercile is by default unclassified and marks the regime of transitional or unclear patterns. This transitional regime represents 5/9 of the paired distributions ($\sim 55\%$ of the time), which thus potentially constitute an important methodological bias in the interpretation of our results (cf. Section D.3). The NN method does not suffer from this issue, but given the subjective categories, which human labelers sometimes did not agree on, it can sometimes yield ambiguous classifications as well (cf. Figures D.1b and D.9). Stevens et al. (2020) and Bony et al. (2020) have shown that especially the patterns Fish and Flowers, on one hand, and Gravel and Sugar, on the other hand, can be confused. Examples of these two ambiguities are shown in Figure D.9.

Another difference between NN and I_{org}/S is that the former can detect several cloud patterns within a domain and therefore does not have to classify a complete fixed domain like I_{org}/S . To compare the two methods, we can ask to which NN-detected predominant pattern does the domain-scale pattern identification with I_{org}/S correspond. To address this question, we compute the relative occurrence and area of NN-detected patterns overlapping the $10^\circ \times 10^\circ$ classification domain at times when I_{org}/S detects a specific pattern. Figure D.2 shows that the two methods result in fairly consistent classifications. That is, when I_{org}/S detects a specific pattern, about half (or more) of the $10^\circ \times 10^\circ$ domain is covered with the NN-detected patterns of the same category. For instance, in 80% of Sugar cases detected by I_{org}/S , NN detects predominantly Sugar patterns with an average coverage of 50% of the I_{org}/S domain. In the case of Gravel, nearly all

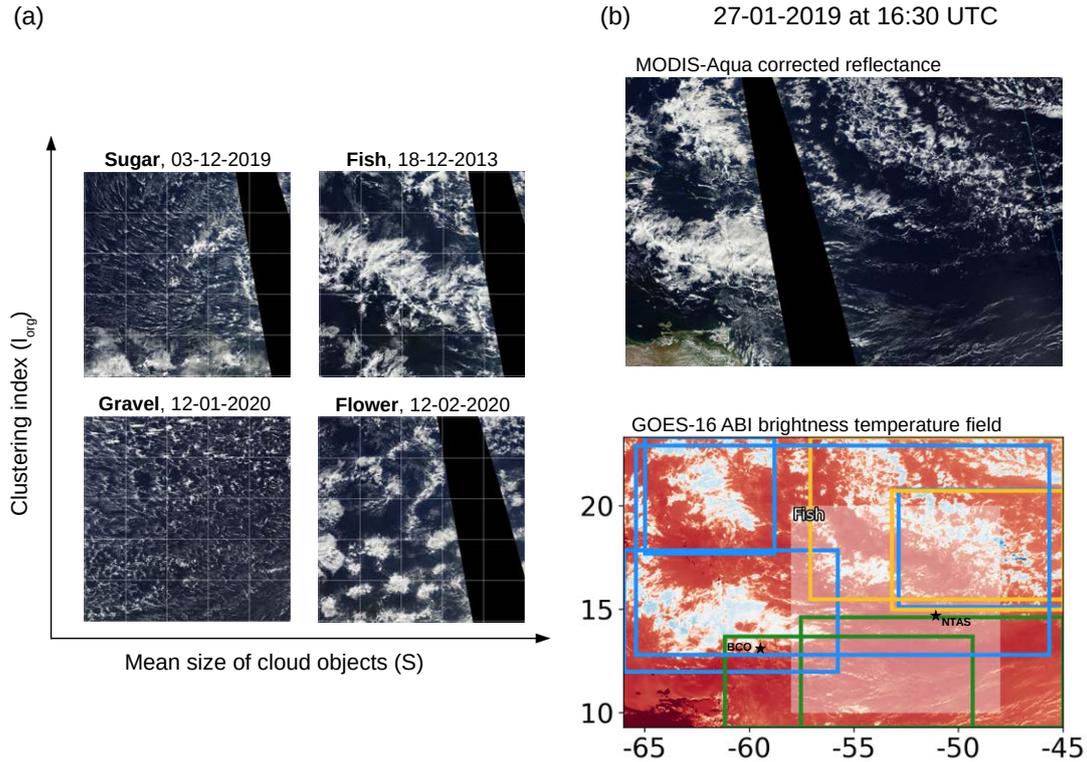


Figure D.1: (a) MODIS-Aqua satellite images from NASA Worldview, illustrating the four prominent mesoscale cloud organization patterns of the North Atlantic trade-wind region over the $10^\circ \times 10^\circ$ I_{org}/S classification domain. The images are positioned in the four quadrants defined by the lower and upper terciles of the S and I_{org} distributions. (b, bottom panel) Illustration of NN-detected pattern rectangles using the GOES-16 ABI T_b field over the $14^\circ \times 21^\circ$ classification domain (Gravel in yellow, Fish in blue and Sugar in Green) and the corresponding I_{org}/S classification (Fish) over the smaller domain (in grey). The top panel is the corresponding visible MODIS image from NASA Worldview. The location of the two sites, BCO and NTAS, is highlighted in the lower panel (b). Note that the NN classification domain is positioned such that BCO and NTAS are equally distant from the lateral edges of the domain, and the I_{org}/S domain is chosen as in Bony et al. (2020), upwind of Barbados.

I_{org}/S classifications correspond to NN-detected Gravel patterns covering on average 80% of the I_{org}/S domain. The Flowers pattern is the one for which the correspondence between the I_{org}/S and NN identifications is the least clear, because of a relatively high occurrence of NN-detected Gravel (Figure D.2, third panel). The exact reason for this ambiguity has not clearly been identified, but it might be related to the thresholding by terciles chosen in Bony et al., 2020 and applied here as well. As shown in Figure D.10, the distributions of I_{org} and S are skewed toward high I_{org} and low S values, which disfavor the detection of Flowers situations (as also evidenced in the Figures D.2 and D.3a by the relatively small sample size of I_{org}/S -detected Flowers), and in particular, Flowers situations with large mean cloud size S . This results in the Gravel and Flower situations to be quite close from each other in the I_{org}/S space, and thus perhaps less easily distinguishable with this method. Note that the smallest ambiguity in the third

panel of Figure D.2 occurs for NN-detected Sugar, which is precisely the furthest from the Flowers pattern in the I_{org}/S space.

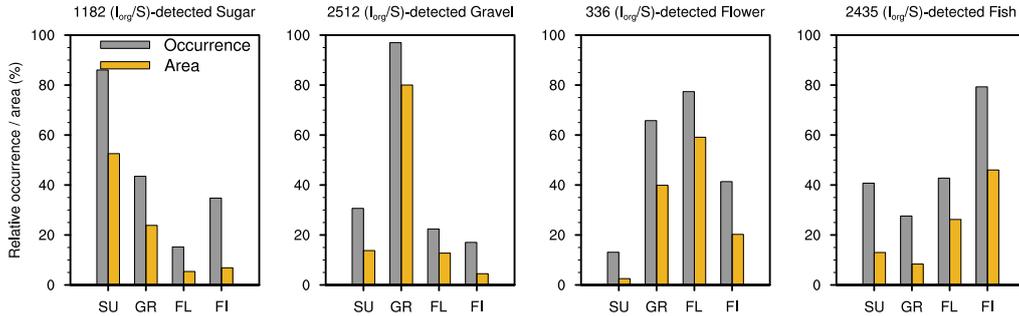


Figure D.2: Relative occurrence (grey bars) and area (yellow bars) of NN-detected patterns overlapping the $10^\circ \times 10^\circ$ classification area at times when I_{org}/S detects a pattern. The frequencies of occurrence are computed with respect to the number of I_{org}/S -detected patterns (as indicated at the top of each panel), and the areas are computed over the overlapping part of the NN-detected rectangles with the $10^\circ \times 10^\circ$ domain and normalized by the $10^\circ \times 10^\circ$ area. The sum of occurrences and areas will thus be greater than 100%.

The correspondence in pattern detection and classification between the two approaches is thus fairly satisfactory. Nevertheless, as discussed above, each of these two classification methods has its own limitations, and thus both methods will be used in our study in order to more robustly assess the daily cycle of mesoscale cloud organization and its influence on the daily cycle of cloudiness. It is also worth-mentioning that the NN approach can offer two additional advantages with respect to I_{org}/S : (i) a more accurate characterization of the atmospheric or surface properties of the patterns (since it does not have to classify an entire domain) and (ii) the possibility to select those patterns that are located over instrumented sites to characterize the patterns in greater details using the ground-based measurements.

D.2.1.2 Pattern cloud covers

The shallow cloud cover (CC) is computed from the GOES-16 ABI T_b mask as defined earlier – i.e., CC is 1 in pixels for which $280 \text{ K} \leq T_b \leq 290 \text{ K}$, and 0 otherwise.

When using the I_{org}/S classification, the cloud cover is averaged over the $10^\circ \times 10^\circ$ domain at each timestep and is assigned to one of the four cloud patterns or to the unclassified category. In doing so, the domain-mean CC averaged over the entire period (DJFM 2018-2020) can be expressed as: $\overline{CC} = \sum_k (CC_k \times F_k)$ where k refers to pattern labels (SU, GR, FL, FI, NO), CC_k is the CC of a given pattern k (also referred to as ‘pattern-related cloud cover’) and F_k is the frequency of occurrence of pattern k at a given time of the 24-hour day, such that $\sum_k F_k = 100\%$ across the daily cycle. The product $CC_k \times F_k$ is the effective contribution of pattern k to the total cloud cover.

The analysis, based on this simple decomposition allows us to both (i) quantify the relative contribution of mesoscale patterns to the daily cycle in total cloud cover, and (ii) disentangle the part of the daily cycle that is due to changes in pattern occurrence frequency and the part due to changes in cloudiness for a given pattern. For reference, we present in Table D.1 the daily characteristics of these two main quantities (F_k and CC_k)

Table D.1: Observed daily characteristics of pattern frequency of occurrence (F_k) and cloud cover (CC_k), as derived from I_{org}/S and NN over their respective computational domain and (with NN only) at the location of the instrumented sites BCO and NTAS—diurnal mean (in %), amplitude (referred to as amp , in %) and local time of daily maximum (the $phase$, rounded to the nearest hour)—as derived from the first harmonics.

	Sugar			Gravel			Flowers			Fish			No		
	mean	amp	phase	mean	amp	phase	mean	amp	phase	mean	amp	phase	mean	amp	phase
Frequency of occurrence (F_k)															
I_{org}/S	7.4	8.8	14	15.7	13.8	00	2.1	2.4	03	15.2	8.8	12	59.5	4.2	05
NN	12.8	6.4	17	36.0	14.3	00	17.2	11.0	06	19.4	4.1	12	32.7	6.5	14
NN-BCO	16.2	8.4	18	37.1	17.5	00	15.7	12.0	06	21.6	4.6	12	27.3	8.4	13
NN-NTAS	21.5	11.0	16	46.3	17.6	00	24.2	15.6	05	24.8	5.6	12	14.4	4.3	14
GOES-16 ABI cloud cover (CC_k)															
I_{org}/S	7.1	1.1	07	13.2	1.2	03	26.9	3.5	02	25.9	5.5	03	19.4	4.0	03
NN	6.5	1.3	01	16.8	2.4	02	26.7	4.3	00	28.2	3.3	00	19.9	3.8	01
NN-BCO	5.8	1.1	02	15.4	2.5	02	24.8	4.0	02	26.2	3.4	01	-	-	-
NN-NTAS	6.8	1.8	01	16.6	2.6	02	25.8	4.8	01	27.1	3.7	00	-	-	-
BCO radar cloud cover (CC_k)															
NN-BCO	20.3	2.3	23	26.3	3.1	23	33.9	6.3	20	36.7	2.7	19	26.7	5.1	01

derived from the different methods, geographical locations and observational datasets (as described below).

We follow the same approach when using the NN classification, but because multiple patterns with different sizes can be detected at one timestep over the domain, the frequency of pattern occurrence (F_k) becomes dependent on both the area of patterns (a_k) relative to the domain area (a_{dom})—i.e., a_k/a_{dom} —and on their temporal frequency of occurrence (f_k), such that $F_k = f_k \times a_k/a_{dom}$, where $a_{dom} = \sum_k a_k$ and thus $\sum_k F_k = \sum_k f_k = 100\%$.

Note that while in the formulation above, the domain area corresponds to the sum of all NN-pattern areas (including unclassified patterns), in practice, here, a_{dom} tends to overestimate the actual domain size due to the overlaps between rectangles of different patterns. The difference between a_{dom} and the actual domain size is somewhat proportional to the number of detected patterns, ranging between 15% in the afternoon and 30% at nighttime (Figure D.11). While this difference seems significant, tests on the accuracy of pattern detection have shown that the interpretation of our results is not sensitive to these overlaps (Figures D.11 and D.12). Our findings also remain consistent when we discard multiple pattern occurrences at one timestep and location (Section D.2.2 and Figure D.13).

In addition to the “domain-mean” pattern-related cloud covers, we also consider the GOES-16 ABI cloud cover for those pattern rectangles that overlap specific locations on the domain (i.e., the location of the instrumented sites described in the following section). In doing this, we similarly weight the pattern cloud cover by its spatial coverage, except for the ‘No’ category (which is not a distinct class, and thus does not have a delineated area around the site location).

D.2.2 *Ground-based remote sensing and in-situ data at BCO and NTAS*

Following Vial et al. (2019), we use ground-based radar, ceilometer, and Micro Rain Radar (MRR) measurements from the Barbados Cloud Observatory (BCO), which is located at the most windward tip of Barbados at 59.48°W, 13.15°N and samples undisturbed trade-wind conditions (see Nuijens et al., 2014 and Stevens et al., 2016 for a detailed description of the BCO and its instrumentation). The cloud and rain statistics are aggregated into 5-min averages. Periods with a radar signal between 4 km and 8 km, including the hour before and after, are discarded to limit our analysis to shallow convection.

The mean rain rate is derived from MRR data at 325 meters above ground (the lowermost level with reliable data). The MRR is also used to compute a rain flag, which is set to 1 when rain rates greater than 0.05 mm/h are measured in at least five range gates in the lowest 3 km (following Nuijens et al., 2014).

The vertical distributions of hydrometeors (i.e. cloud and rain droplets) and clouds are derived from a 35.5 GHz (Ka-Band) Doppler cloud radar. The hydrometeor mask is derived using a threshold of -50 dBZ on the equivalent radar reflectivity Z_e (a 10 dBZ lower threshold than used in Vial et al., 2019 to increase the sensitivity to smaller clouds). Cloud fraction profiles are obtained from the hydrometeor mask by discarding periods of rain: when the ceilometer does not detect a cloud base (due to strong rain) or when the MRR rain flag is 1. Moreover, radar signals below the ceilometer-detected cloud-base height are set to 0 in the cloud fraction profiles. Periods when neither the ceilometer nor the MRR are running are also discarded.

The (rain-corrected) radar cloud profiles are also used to derive the total cloud cover, as well as the contribution to the total cloud cover from three distinct categories of clouds: (i) shallow cumuli with cloud base (CB) below 1 km and cloud top (CT) below 1.3 km, (ii) deeper cumuli with CB < 1 km and CT > 1.3 km, and (iii) clouds aloft with CB > 1 km. This decomposition is slightly different than the more commonly used decomposition from Nuijens et al., 2014, as the use of radar rather than ceilometer data allows us to further decompose the ‘LCL’ cloud category into two sub-categories according to the cloud-top height. The decomposition used here is similar to the one applied to model data in Vial et al., 2019 (see details in their Appendix A).

Note that the rain-correction applied will lead to an underestimation of total cloud cover, as periods of rain are usually also periods of cloudiness. Because rain is most frequent during nighttime (Nuijens et al., 2009; Vial et al., 2019) and because the rain frequency also depends on the patterns (i.e. fish is the rainiest, followed by Flowers and Gravel Schulz et al., 2021), we tested that this underestimation does not bias our results by comparing the daily cycles of the total cloud cover and the total hydrometeor cover. The hydrometeor cover includes both cloud and rain droplets and thus overestimates cloud cover (being on average ~10% larger for Sugar, ~15% larger for Gravel and Flow-

ers, and $\sim 20\%$ larger for Fish). However, we find the daily cycles of total cloud cover and total hydrometeor cover to be very similar for all patterns (compare the black and grey curves on the top-right panels of Figure D.7), and thus conclude that the underestimation of cloud cover due to the rain-correction does not bias our results.

We also use measurements of sea surface temperature (SST) and near-surface wind speed from the Northwest Tropical Atlantic Station (NTAS) open ocean surface mooring at 51.02°W , 14.82°N . NTAS measurements have a temporal sampling rate of 1 minute. Wind data are collected at about 3 meters above sea level and SSTs at 1-meter depth. To directly compare the buoy observations of wind speed with ERA5 estimates (section D.2.3), we adjust the 3-meter wind to conform to the reference height of 10 meters, using the simple power-law wind profile ($u_2 = u_1(z_2/z_1)^{0.11}$ – where u_2 is the wind speed at the reference 10-m height (z_2), and u_1 the wind speed measured at height z_1 ($= 3.4$ meters)), which was shown to be a good approximation for use over the ocean, where near-neutral stability conditions prevail (Hsu et al., 1994). This adjustment implies an increase of the wind speed of about 1 m/s between 3 and 10 meters.

In a similar way as described in Section D.2.1, we construct pattern-related composites for the BCO and NTAS sites, but we only select the NN-detected patterns overlapping the location of these instrumented sites. The compositing is instantaneous in the sense that we average all measurements within ± 15 minutes around the classification time.

As explained above, several pattern rectangles can be detected at one location, potentially introducing a bias in the composite when different patterns occur at the same time. That said, we show in Figure D.13 that discarding the timesteps when multiple labels occur at BCO does not affect the results, but does reduce significantly the sample size of our composites. We therefore keep all detected patterns at the site locations to construct the pattern-related composites.

D.2.3 The large-scale environment from ERA5 reanalysis

The ERA5 reanalysis is based on the Integrated Forecasting System (IFS) Cy41r2, operational since 2016. It provides hourly estimates of atmospheric variables, at a horizontal resolution of 31 km (0.25° or TL639) and 137 vertical levels from the surface to 0.01 hPa (Hersbach et al., 2020).

We here use hourly output for the 10-meter wind speed and the lower tropospheric stability (LTS, defined in Klein and Hartmann, 1993 as $\theta_{700} - \theta_{1000}$, where θ is the potential temperature in Kelvin) over the NN classification domain (9.3°N - 23.3°N , 45°W - 66°W), in order to explore how the daily cycles of cloudiness and mesoscale patterns relate to these large-scale environmental factors that are known to play a role in the variability of trade-wind cloudiness and organization at longer timescales (Bony et al., 2020; Brueck et al., 2015; Nuijens et al., 2015).

Following the approach described in Section D.2.1, we construct the pattern-related composites for the near-surface wind speed (U_k) and LTS (LTS_k) and sample their daily cycle. The relationship between the large-scale environment (including U and LTS) and the mesoscale patterns of organization was first explored in observations at the day-to-day and inter-annual timescales using the Iorg/S pattern classification method (Bony

et al., 2020). Here, we extend the analysis by considering the daily timescale and patterns detected with the NN approach. Note also that using the I_{org}/S method, combined with hourly estimates of the environmental conditions, the sample size is too small to see a robust signal on the daily timescale.

D.3 DAILY CHANGES IN THE OCCURRENCE OF MESOSCALE PATTERNS OF ORGANIZATION

The first question we address here is whether the observed occurrence of mesoscale patterns of shallow convection varies at the daily timescale.

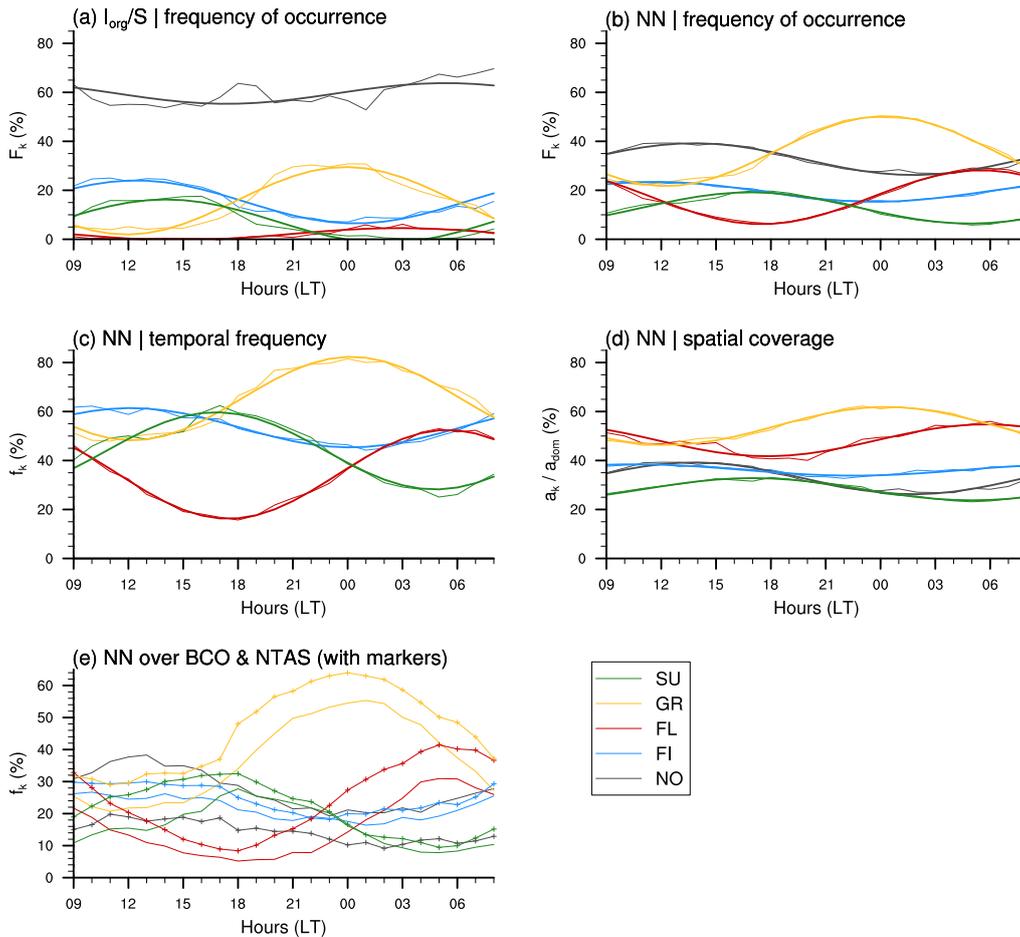


Figure D.3: Top panels: daily cycle in the occurrence frequency of mesoscale patterns for Sugar (green), Gravel (yellow), Flowers (red), Fish (blue) and No category (grey), with the first harmonics superimposed in thicker lines. Panel (a) shows the $10^\circ \times 10^\circ$ domain-scale I_{org}/S pattern frequency and panel (b) shows the occurrence frequency of NN-detected patterns within the $14^\circ \times 21^\circ$ domain. Middle panels: contributions to the NN occurrence frequency owing to (c) the temporal frequency and (d) spatial coverage. Bottom panel (e): temporal frequency of NN-detected patterns over BCO (solid line) and NTAS (line with markers). Note that at a single point, the temporal frequency corresponds to the frequency of occurrence.

Figure D.3 (a,b) shows that whatever the method we use to classify mesoscale organization (I_{org}/S or NN), the frequency of occurrence of all forms of organization exhibits a pronounced daily cycle with distinct 24-hour phasing. The patterns Fish and Sugar preferentially occur during daytime, and as such can be viewed as daytime patterns, while Gravel and Flowers appear rather as nighttime patterns. More specifically, the frequency of Fish patterns increases in the morning and reaches a maximum at 12 local time (LT); the peak in the Sugar pattern frequency is shifted towards the afternoon, at 14LT for I_{org}/S and 17LT for NN (Table D.1); Gravel increases during the afternoon and peaks at midnight; and the Flowers population grows soon after sunset until reaching a peak at the end of the night between 03LT and 06LT depending on the method.

The amplitude of these daily cycles is substantial, with a minimum of 20% relative to the daily mean for the NN-detected Fish, between 35% and 60% for NN-detected Gravel, Sugar and Flowers, and more than 100% daily variation for the I_{org}/S -detected Sugar and Flowers patterns (Table D.1).

When using the NN approach, the frequency of pattern occurrence depends both on the temporal frequency (Figure D.3c) and spatial coverage of patterns (Figure D.3d). These two attributes of pattern occurrence exhibit very similar diurnal phasings, which means that when the NN detects more of a given pattern, it also extends over a larger area and vice versa. However, the daily variability in the occurrence of patterns (Figure D.3b) appears to be more strongly driven by their temporal frequencies – whose variations range between 10% and 45% with respect to daily means – than their spatial coverage, which varies between 5% and 15% (relative to daily means) among the different patterns.

Gravel appears to be the most frequent pattern and on average covers larger areas of the domain (more than 50% of the domain area on average). The Flowers pattern also covers about half of the domain on average, but with a very pronounced daily cycle in its temporal frequency – it is the least frequently observed pattern around sunset, as well as one of the most frequent patterns at night. Sugar and Fish are the least spatially extended patterns (about 30% and 40% of the domain, respectively), but they can be frequently observed, especially during daytime (about 60% of the time at 12LT and 17LT, respectively).

Similar daily phasings as on the entire domain are also found when looking at the surface sites BCO and NTAS independently. However, the comparison between BCO and NTAS (which lies 8° east of BCO) reveals a geographical dependency in pattern occurrence, especially regarding the daily means and amplitudes of the daily cycles in the temporal frequency of patterns (Figure D.3e). Overall, the frequency of detected patterns is systematically higher at NTAS than at BCO, while the frequency of unclassified patterns is greater at BCO. Note that larger differences were found, especially for the unclassified category, by applying the NN algorithm on a 5°-eastward shifted domain (not shown). This was due to a lower pattern detection at BCO, which lay closer to the edge of the domain. With the present domain, however, BCO and NTAS are equally distant from the edge of the domain (about 6°), thus we expect the differences between the two sites to be real.

In particular, the differences for the Gravel and the Flowers patterns, which are the most important throughout the 24-hour daily cycle, might reflect an east-west gradient in the frequency of occurrence of these patterns, with a greater occurrence on the east

due to stronger easterlies (not shown). This is consistent with the findings in Bony et al. (2020) who show that the Gravel and Flowers patterns mostly occur in conditions of stronger near-surface wind speed. The frequencies of the Fish patterns are fairly similar at BCO and NTAS, which can be expected given the large-scale characteristic of this pattern that, moreover, is most often oriented along the east-west direction (Schulz et al., 2021; see also in Figure D.9). Finally, we note also that the daily amplitude of the Sugar pattern is higher at NTAS than at BCO due to a much higher occurrence frequency during daytime.

D.4 DEPENDENCE OF THE DAILY CYCLE OF TRADE-WIND CUMULI AND PRECIPITATION ON MESOSCALE PATTERNS OF ORGANIZATION

D.4.1 Mesoscale pattern signatures on the daily cycle of GOES-16 ABI cloud cover

Figure D.4 shows the averaged daily cycle of GOES-16 ABI cloud cover associated with the different patterns (CC_k) detected over the entire classification domains (panels a and b) and over the sites at BCO (c) and NTAS (d). The similarity between all four panels is salient, and suggests that the daily cycle in cloudiness is overall independent of the classification method, of the patterns themselves and of the geographical location.

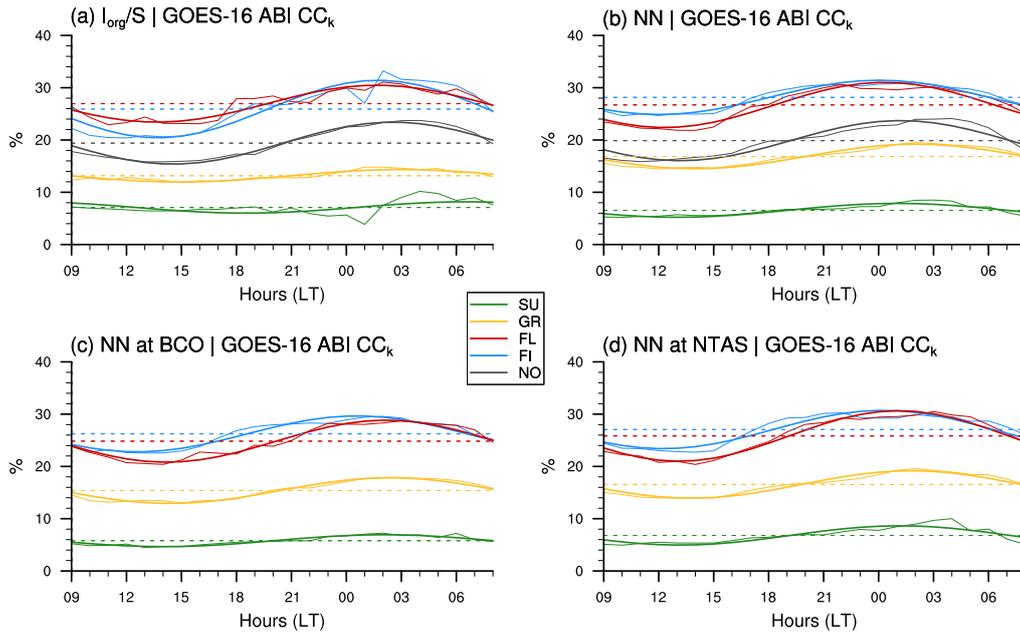


Figure D.4: Daily cycles and first harmonics of patterns cloud cover (CC_k) derived from GOES-16 ABI using (a) I_{org}/S classified patterns, (b) NN-detected patterns over the entire domain, (c) NN-detected patterns over BCO and (d) NN-detected patterns over NTAS. The dash lines represent the daily-mean cloud covers. Note that the NO category does not appear in panels (c–d) as there is no delineated area over which we can compute a cloud cover. The color code is the same as in Figure D.3.

The weak dependency of the daily phase of cloud cover on NN patterns is particularly striking, especially so for the large domain (Figure D.4b): cloudiness always minimizes

in the afternoon (between 13LT and 16LT) and maximizes at nighttime (between 00LT and 04LT), consistent with Vial et al. (2019) and the 3-year winter-time climatological daily cycle in cloud cover calculated here (cf. black lines in Figure D.6). The exception is for Sugar identified with I_{org}/S , which shows a phase shift of about 5 hours, although the robustness of this daily cycle might be questioned given the small number of Sugar cases identified with I_{org}/S at nighttime (Figure D.3a).

From a more quantitative point of view, the daily mean and amplitude in cloud cover are, to some extent, dependent on pattern, and more precisely, on the mean size of cloud objects (S): patterns with small cloud entities (Sugar and Gravel), tend to have a smaller daily mean and amplitude in cloud cover than the Flowers and Fish patterns, which have much larger cloud structures and a larger fraction of stratiform cloudiness near the inversion that is particularly sensitive to the daily cycle (Vial et al., 2019). Nevertheless, the daily variability in cloud cover for a given pattern remains small compared to the differences in daily-mean cloud cover between the patterns. The cloud cover varies by at least a factor of two across the different patterns (see also Bony et al., 2020), while the daily variations range between 10% and 20% relative to daily means.

Note that Bony et al. (2020) found a higher cloud cover (from MODIS cloud products) for the Flowers pattern than for the Fish pattern identified with I_{org}/S , which is consistent with our daytime estimates (Figure D.4a), given that only daytime measurements of MODIS over the North Atlantic trade-wind region (mid-morning for Terra and early afternoon for Aqua) were used. Nevertheless, it should also be noted that our GOES-16 ABI estimates of cloud cover are overall lower than MODIS estimates (cf. Bony et al., 2020), presumably because the T_b cloud mask and lower resolution of GOES-16 ABI prevent from detecting the smallest clouds (see also the discussion at the end of Section D.4.2).

Overall, our results suggest that the mesoscale patterns of cloud organization constitute a fairly robust constraint on cloud cover, and that the dependence of the daily cycle on mesoscale organization is essentially due to the daily changes in pattern frequency of occurrence. This also means that knowing the daily variation in pattern occurrence and the mean cloud cover for a given pattern, we can recover to a large degree the daily cycle in the effective cloud cover of the different patterns (Figure D.5).

D.4.2 Contribution of mesoscale patterns to the daily cycle of cloud cover

Combining the daily cycles in F_k (Section D.3) and CC_k (Section D.4.1) into the product $CC_k \times F_k$ allows us to quantify more explicitly the relative contribution of mesoscale patterns to the daily cycle in total cloud cover. We present this “effective pattern cloud cover decomposition” in Figure D.6 for GOES-16 ABI cloudiness and detected patterns over the entire classification domains (panels a and b), as well as for the radar cloud cover at BCO and the NN-patterns overlapping the BCO site (panel c). Recall that when using the NN-patterns, the sum of all contributions is greater than the total cloud cover (black lines) because of the overlaps between multiple label occurrences (Section D.2.1.1).

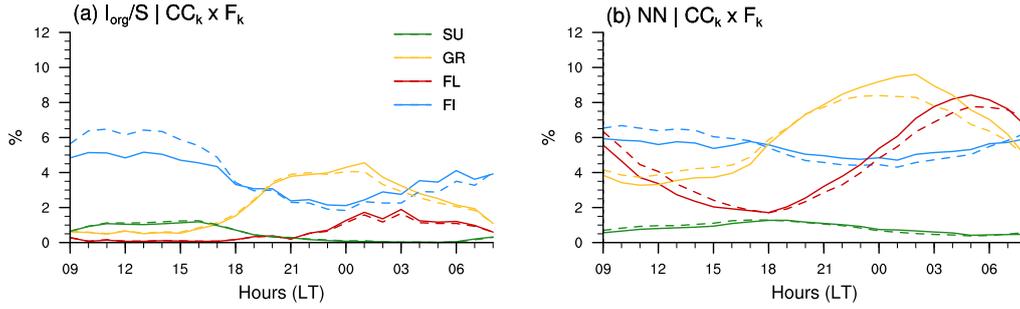


Figure D.5: Daily cycle in the effective pattern cloud covers ($CC_k \times F_k$, solid lines) from GOES-16 ABI using (a) I_{org}/S classified patterns and (b) NN-detected patterns over the entire domain. The dashed lines represent the product $CC_k \times F_k$ but with CC_k fixed to the daily-mean. The color code is the same as in Figure D.3.

Overall, the contribution to total cloudiness from the four defined patterns is greater at night – when the total cloud cover is maximum – than during the day. However, the extent to which the patterns explain the total cloudiness strongly depends on the classification method. When using I_{org}/S , about 60% of total cloudiness is explained by the No category throughout the 24-hour cycle (Figure D.6a). This percentage reflects the frequency of occurrence of the No category (Figure D.3a) that is set by the classification criteria of patterns (Section D.2.1.1). With the NN method, only about 30% of the cloud cover is due to the No category, and the daily cycle for this contribution remains weak (Figures D.6b,c) owing to the opposite diurnal phasing of F_{NO} and CC_{NO} (compare for instance Figures D.3b and D.4b).

The contribution of the No category to the overall daily cycle in cloud cover, however, deserves some more discussion. It is notable that adding the percentage of cloud cover due to the overlaps between NN patterns (i.e., the difference between the sum of all contributions and the actual total cloud cover in Figure D.6b) and the percentage of $CC_{NO} \times F_{NO}$ (the grey area in Figure D.6b), leads to a similar contribution of *unclear* patterns as that of the No category for I_{org}/S . This further supports the coherence between these two intrinsically different methods. Moreover, it suggests that using the NN method, we can unravel the contribution to total cloud cover due to forms of cloud organization that are somehow related to the four predefined patterns (i.e., the contribution from overlaps) and that due to organization forms that are not related to the predefined patterns (i.e., contribution from the No category). Therefore, we argue here that the most likely contribution of unclassified forms of organization to total cloud cover is about 30% (the percentage given by the NN method) – and thus that the largest extent of cloud cover can be explained by these four forms of mesoscale cloud patterns, as follows.

The Fish pattern is the most important contribution to the afternoon cloudiness regardless of the classification method (Figure D.6a,b). This is partly because this pattern occurs more frequently during daylight hours (especially when using I_{org}/S), but also because the other patterns are less frequent and therefore contribute relatively little to cloudiness at this time. Actually, when using the NN method, the effective contribution of the Fish pattern to the daily cycle in cloud cover tends to be quite small (Figure D.6b,c), because of opposing phases between F_{FI} and CC_{FI} (Figure D.5b) and relatively weak daily amplitudes in F_{FI} (Table D.1 and Figure D.3). Note that the Fish pattern is

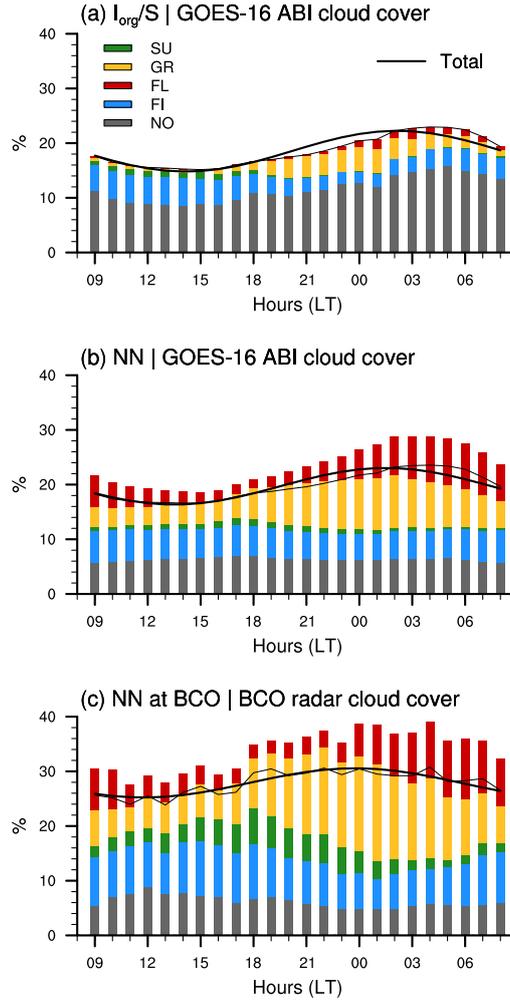


Figure D.6: Daily cycle of total cloud cover (black thin line, with the first harmonics superimposed in thicker line) and relative contributions of the different mesoscale patterns of cloud organization ($CC_k \times F_k$): Sugar (green), Gravel (yellow), Flowers (red), Fish (blue) and No classification (grey). In panels (a,b) the pattern cloud covers are derived with GOES-16 ABI T_b cloud mask on (a) the $10^\circ \times 10^\circ$ I_{org}/S domain and (b) the NN-detected patterns within the entire $14^\circ \times 21^\circ$ classification domain. In panel (c): same as in (a–b) but the cloud covers are diagnosed from the radar at BCO using NN-detected patterns overlapping the BCO site.

often associated with a synoptic disturbance that persists for several days, continuously forced by a convergence line (Aemisegger et al., 2021; Schulz et al., 2021). This may thus explain the small daily cycle in the occurrence frequency of the Fish patterns.

When using the NN method, Gravel and Flowers are the dominant contributions at nighttime, both over the entire classification domain and at BCO. Gravel explains about 45% of the total cloud cover around midnight, while the Flowers contribution maximizes just before sunrise with values ranging between 30% (at BCO) and 40% (over the large domain). A similar 24-hour phasing is observed for these two patterns when using I_{org}/S , although their contribution to total cloudiness remains small (Figure D.6a).

The case of the Sugar pattern is interesting because it can occur quite frequently (Figures D.3c,e) but its spatial extent is relatively small (Figure D.3d). Consequently, its

contribution to total cloudiness appears much larger at BCO (Figure D.6c) than over the large domain (Figure D.6b). Moreover, the daily phases of F_{SU} and CC_{SU} are opposed (Figures D.3b and D.4b), which reduces the effective contribution of this pattern to the daily cycle of total cloudiness. With I_{org}/S , Sugar is the pattern that contributes the least to the total cloud cover and its daily cycle.

Overall, these results reveal that unclassified and Fish patterns are the most important for daytime cloudiness, while Gravel and Flowers contribute most to nighttime cloudiness. The contribution from the Sugar pattern, although never dominant, maximizes around sunset (with NN) and can be more important when viewed locally than at large-scale.

Finally, we draw attention to the three different cloud cover estimates in Figure D.6 (black lines). We note, in particular, that the cloud cover from the radar at BCO is about 3% to 10% larger than the satellite-based estimates over the large domain. This difference is even larger when using the ceilometer-based cloud cover (not shown). This difference, which has also been reported in previous studies comparing BCO data with other satellite-based products (Nuijens et al., 2015; Vial et al., 2019), is likely due to different capabilities of the instruments to measure low-level clouds – the BCO radar or ceilometer being much more sensitive to low-level cumuli than the GOES-16 ABI infrared channel. Moreover, the T_b cloud mask is defined such as to exclude some of the shallowest clouds, in particular those with a cloud-top height below 1 km (Bony et al., 2020). So we might expect the difference in the cloud cover estimates (between GOES-16 ABI and the BCO radar or ceilometer) to be particularly pronounced for the Sugar pattern that essentially consists of clouds with little vertical extent above the LCL (Schulz et al., 2021). Although this difference is indeed slightly larger for the Sugar pattern, it remains, nevertheless, of the same order of magnitude regardless of the pattern (compare the pattern-related CC_k 's in Table D.1). This could be explained based on findings from Schulz et al., 2021 showing that the cloudiness near the LCL does not vary substantially from pattern to pattern, and therefore the difference between the BCO and satellite-based estimates should also remain relatively similar from pattern to pattern.

D.4.3 Mesoscale pattern signatures on the daily cycle of clouds and precipitation at BCO

We here take advantage of the BCO dataset (Section D.2.2) to characterize further the different cloud and precipitation properties associated with each pattern of organization. Schulz et al. (2021) provides a detailed description of the mean structure of clouds and of the convective boundary layer for each pattern. Here, we focus more on the daily evolution of precipitation and of the vertical distribution of cloudiness (Figure D.7). We show that, even if the daily cycle of the overall cloud cover (CC_k) is very similar among the different patterns (Section D.4.1), each organization pattern appears with its own daily cycle of shallow convection:

- The Sugar pattern essentially consists of small non-precipitating cumuli with a cloud-base height close to the LCL (below 1 km) and weak vertical extent during daylight hours. Near sunset, the overall cloud cover starts to increase due to a blooming of slightly deeper clouds reaching the upper cloud layer (above 1.3 km) between 19 LT and 02LT. The proportion of shallower clouds (i.e., $CT < 1.3$ km) is

dominant and remains fairly constant throughout the day. However, despite this overall cloud deepening at night, the precipitation rates measured at the surface remain low throughout the day.

- For the Gravel pattern, clouds overall reach higher levels in the cloud layer, with about $2/3$ of the cloud population having their top above 1.3 km. The daily cycle of the total cloud cover is mainly driven by the population of thicker clouds ($CB < 1$ km and $CT > 1.3$ km), which increases in the afternoon and maximizes at 00LT. The cloud cover from both the very shallow clouds ($CB < 1$ km and $CT < 1.3$ km) and clouds aloft ($CB > 1$ km) remains roughly constant throughout the day. Interestingly, the precipitation peak is delayed by about 6 hours with respect to the maximum in cloud cover. Seifert and Heus, 2013 noted a similar feature in large-eddy simulations of shallow convection with cold-pool organization (cf. their figure 2). One reason could be that there is, first, a reduction of rainfall due to evaporation below cloud-base and that, later, the generated cold-pools created a moister environment allowing more rainfall at the surface. Further investigation is needed to verify the robustness of this time shift between cloudiness and precipitation and to provide an explanation for it.
- The Flowers pattern has to be interpreted with caution as the number of detected patterns is small over BCO, especially between 18LT and 21LT (Figure D.3), at times when the peak in cloudiness is observed (Figure D.7). Nevertheless, we find that about 80% of the cloud cover is explained by clouds with cloud-top height above 1.3 km, and the cloud fraction near the inversion tends to be more pronounced at late night hours. The daily cycle in precipitation seems weak and local maxima are not always correlated with peaks in cloudiness (for instance at 12 LT) – a feature that could also be explained by the presence of cold pools as for the Gravel case.
- For the Fish pattern, the relative contributions of shallow and deeper clouds are similar as for the Flowers pattern. The daily variability in CC exhibits two local maxima, around sunset and at early morning hours, but with an overall tendency (given by the first harmonic) for a daily maximum in cloudiness around sunset. The vertical cloud fraction profiles reveal that the nighttime inversion and clouds can reach higher levels than during the day. The precipitation daily cycle given by the first harmonic is similar as for the Gravel pattern, albeit with more pronounced variability.
- The unclassified cloud scenes (the 'No' category) might be the most variable on the daily timescale. The cloud profile is fairly similar to that of the Sugar pattern during the day, whereas at night it appears as a mixture between the Gravel and the Fish patterns (with both a strong inversion near 2 km and a significant cloud fraction above). Despite the nighttime cloud deepening, surface rainfall remains quite low throughout the day. However, there is small difference in the amplitude of the daily cycle between the total hydrometeor cover and the total cloud cover (of about 3% – between the grey and black curves in the top-right panel), which suggests a nighttime enhancement of clouds with precipitation at higher levels.

One notable feature worth mentioning is the different 24-hour phasing in pattern-related cloud cover at BCO whether the cloud cover is derived using GOES-16 ABI

(Figure D.4c) or BCO remote sensing instruments (Figure D.7). When using the BCO measurements, the peak in cloud cover is systematically earlier in time compared to when GOES-16 ABI retrievals are used (with a phase shift of several hours—between 2h and 7h—depending on the pattern, see Table D.1). The time difference in the solar forcing between the western and eastern boundary of the NN classification domain is at most $1\text{h}24'$ (4 minutes for every degree longitude), which cannot explain the aforementioned time shifts. One reason explaining these differences could, however, be related to sampling: a small number of detected patterns at BCO (e.g., Flowers at 18LT) and/or a too small temporal averaging for a given pattern to capture the averaged properties of the pattern at a given time. For example, given the large size of Flowers (~ 100 km or more), they can take several hours to cross entirely over BCO (see Figure 7 in Stevens et al., 2020). Complementary tests of the influence of spatial/temporal scales on our results support this explanation (Figure D.14).

D.5 THE ROLE OF ENVIRONMENTAL FACTORS

Two related questions that can be asked now are what controls the daily variability of pattern frequency, on the one hand, and the constancy in the nighttime peak of cloudiness regardless of the pattern, on the other hand. We shed light on these questions by diagnosing the daily cycle of some of the environmental factors that are known to be either determinant for pattern occurrence (Bony et al., 2020) and/or controlling factors of winter trade cloudiness at timescales longer than a day (Brueck et al., 2015; Nuijens et al., 2015). In those studies, the near-surface wind speed and LTS appear to be the most influential factors on the day-to-day and interannual timescales. We thus consider those two variables, as well as the SST, which is known to be an important ingredient for the daily cycle of convection when the near-surface wind speed is weak (Bellenger et al., 2010; Ruppert and Johnson, 2016).

The results presented in Figure D.8 show that these three variables exhibit a daily cycle, with distinct phasings and amplitudes depending on the variable itself, on the pattern, and on the dataset that is considered.

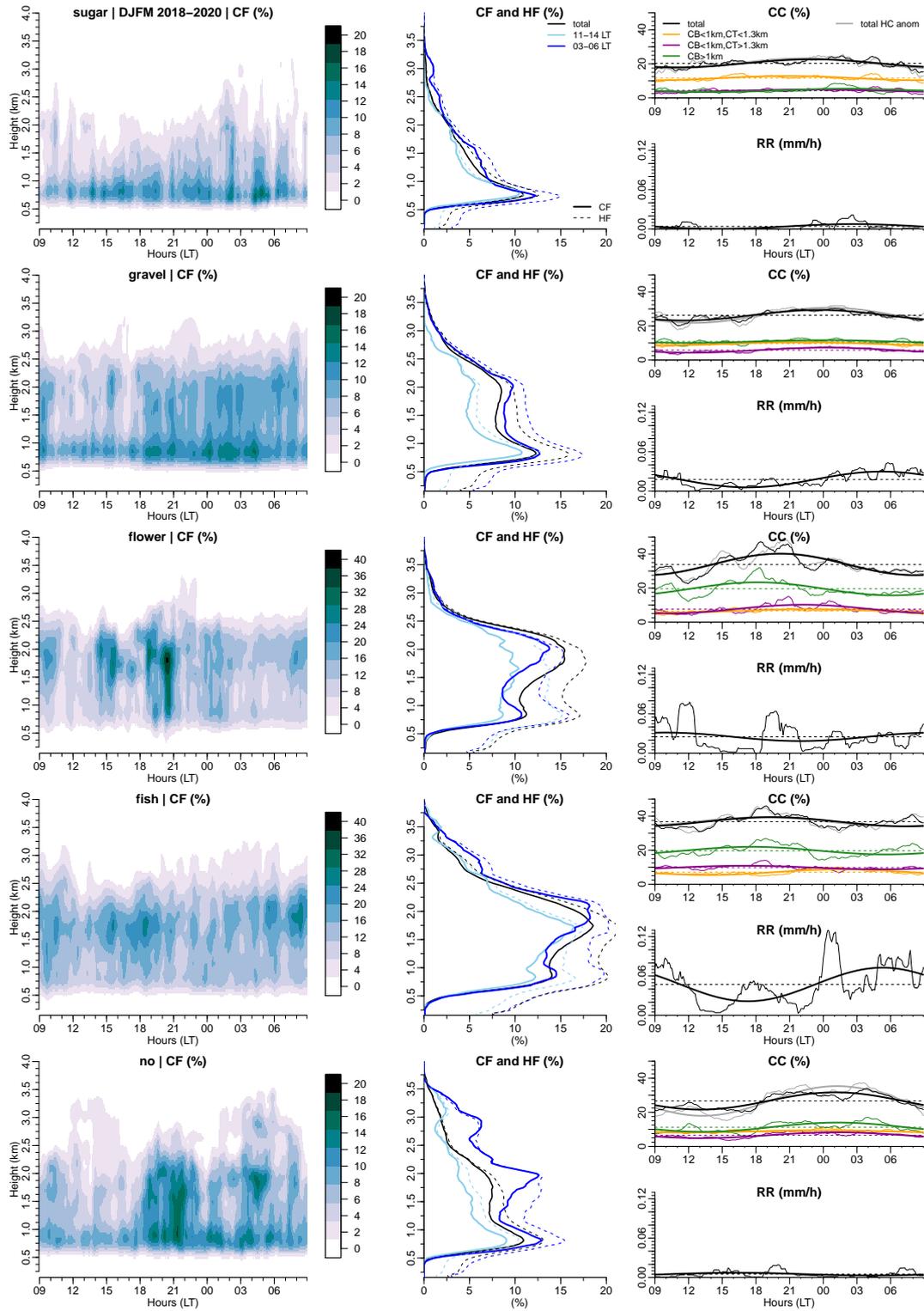


Figure D.7: Pattern-related daily cycle of cloudiness and precipitation from the Barbados Cloud Observatory for (top-to-bottom) Sugar, Gravel, Flowers, Fish and No patterns. Left panels: rain-corrected cloud fraction derived from the radar; Middle panels: profiles at selected times for the rain-corrected (cloud) and hydrometeor (cloud and rain droplets) fraction derived from the radar; Right panels: (top) radar-derived cloud covers and (bottom) rain rate derived from the Micro Rain Radar (MRR). In the right panels, the thin solid lines represent the actual data, the thicker lines are the first harmonics, and the thin dotted horizontal lines are the daily means. Also shown in the top right panels by the grey curve, is the daily cycle in total hydrometeor cover (HC), with the difference in the daily means between HC and CC removed (for ease of readability in the figure). The daily means in HC are 31.5% for Sugar, 40.6% for Gravel, 49.8% for Flowers, 57.4% for Fish and 46.3% for No.

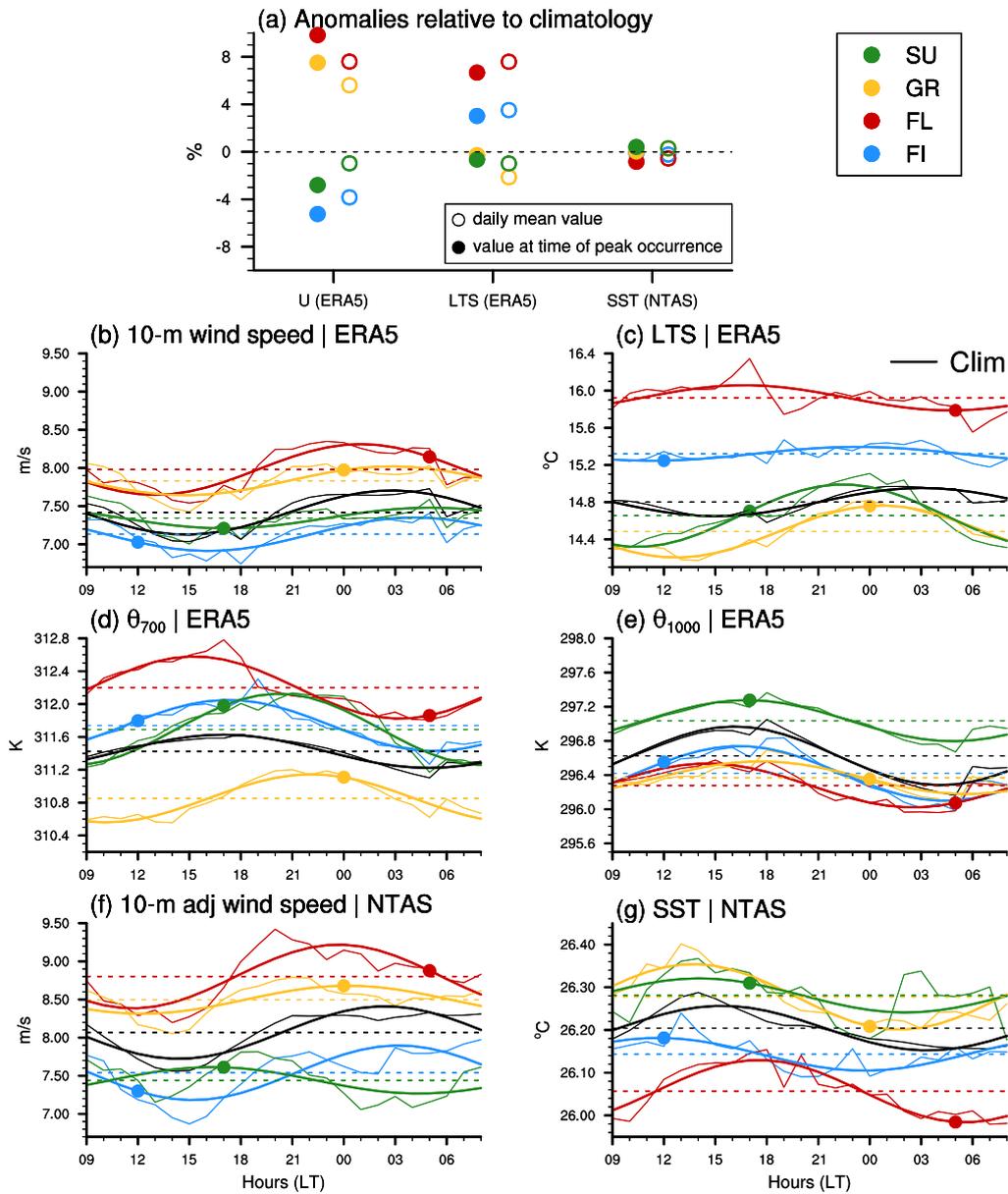


Figure D.8: (a) Fractional anomalies (with respect to climatological mean) of pattern-related large-scale 10-m wind speed (U) and LTS from ERA5 and local SST from NTAS measurements. The empty circles represent the anomalies based on the daily-mean U, LTS and SST, whereas the filled circles represent the anomalies at the time of the peak occurrence of the specific patterns (which is indicated by the filled markers in panels b–g). (b–e) Pattern-related daily cycle of U, LTS, θ_{700} and θ_{1000} sampled over the large NN classification domain when using ERA5 data. (f–g) Pattern-related daily cycle of U and SST sampled over the NTAS site and using NTAS measurements (d, e). The first harmonics is shown by the thicker line. Note that the 3-meter wind speed measured at NTAS has been adjusted to conform to the reference 10-m height section D.2.2). The color code for the patterns is as indicated in the legend, and the black line represents the climatological mean (DJFM 2018-2020).

D.5.1 *Relation to the daily cycle in pattern frequency*

For daily means in large-scale wind speed and LTS, our findings are consistent with Bony et al. (2020): the large-scale environment tends to be less stable with weak winds for Sugar, less stable with strong winds for Gravel, more stable with strong winds for Flowers, more stable with weak winds for Fish (Figure D.8a, empty circles). Here, we show that this holds at any time of the day, and that the daily cycle in wind speed amplifies the relationship between wind speed and pattern occurrence (Figure D.8a, filled circles). The large-scale wind tends to be stronger at night regardless of the pattern (Figure D.8b), and thus it discriminates the occurrence of the organization patterns in the same way at sub-daily, daily and inter-annual timescales: Gravel and Flowers occur mostly at night when the wind is stronger, while Fish and Sugar occur mostly during daytime when the wind is weaker. It is, however, worth noting that this relationship can be different depending on whether the wind speed is diagnosed at large-scale (here, with ERA5 over the large NN classification domain in Figure D.8b) or locally (e.g., at NTAS in Figure D.8f). Indeed, the overall increased variability in the pattern-related daily cycles in wind speed at NTAS could explain the difference for the Sugar pattern (compare the thin lines between Figures 8b and 8f).

The daily cycle in LTS can be different from one pattern to another (Figure D.8c), and the relationship between LTS and pattern occurrence at daily timescale is opposite to that found at longer timescales: at times of maximum occurrence of Fish and Flowers patterns the environment is the least stable of the day, and at times of maximum occurrence of Gravel and Sugar patterns the environment tends to be more stable compared to the pattern-related daily-mean LTS (Figure D.8a).

The SSTs tend to be colder for the Flowers pattern and warmer for the Sugar pattern (as in Bony et al., 2020), but overall this variable does not significantly explain the variability in pattern occurrence on the day-to-day or daily timescales (Figure D.8a).

Therefore, the near-surface wind speed is here the factor that explains best the daily variability in pattern occurrence; it discriminates the daytime from nighttime patterns.

D.5.2 *Relation to the constancy of the nighttime peak of cloudiness*

As mentioned earlier, the large-scale wind is overall stronger at night whatever the pattern (Figure D.8b), and thus correlates quite well with the daily cycle in pattern-related cloudiness (Figure D.4). This co-variability between trade-wind cloudiness and near-surface wind speed has already been discussed in the context of slowly-varying observations (Brueck et al., 2015; Nuijens et al., 2015) or in equilibrated large-eddy simulations (Nuijens and Stevens, 2012). Here, we show that it happens on the daily timescale whatever the organization pattern in place, and could therefore constitute a basic ingredient of trade-wind convection: as the winds reinforce, surface evaporation increases, providing the moisture that is needed for the clouds to grow deeper, which then helps increase the overall cloud cover. Surface winds, in turn, can be enhanced at night as radiative cooling destabilizes the boundary layer and strengthens the momentum transport by the shallow convection (Hourdin et al., 2015; Schlemmer et al., 2017).

The domain-mean LTS has a small daily cycle with a maximum at nighttime – at times when the cloud cover is maximum (compare black line in Figures D.8b and Figures D.6b).

This result is somewhat expected given that large nighttime cloudiness is mainly related to the spreading of a stratiform cloud layer below the trade inversion (Vial et al., 2019), and that stratiform cloudiness is more likely to occur under stronger stability (Wood and Bretherton, 2006). Given that strong LTS is favored by weak θ_{1000} and/or strong θ_{700} and that both θ_{1000} and θ_{700} exhibit a daily cycle with a minimum at nighttime, the nighttime maximum in LTS then primarily owes to the minimum in (near-)surface warming.

Note, however, that the daily cycle in LTS is different from one pattern to another; this is related to a θ_{700} -dependency, as the daily cycle in θ_{1000} is fairly similar for all patterns (Figures D.8d,e). The Flowers pattern in particular is associated with less stable conditions during the night, owing to a large decrease in θ_{700} . From these results, we hypothesize that while more stable conditions can be more favorable to stratiform cloudiness at night, once the Flowers stratiform cloud layer are present they might produce locally less stable conditions, presumably through enhanced radiative cooling at cloud-top (see for instance in Figure 6 of Albright et al., 2021). Case studies of the field campaign EUREC⁴A (Elucidating the role of clouds-circulation coupling in climate, Stevens et al., 2021) that took place windward of Barbados in January-February 2020 could provide new opportunities to further investigate how the cloud patterns impact the local environment.

Due to the sustained easterlies in this season, the daily cycle in SST is not expected to be strong and hence also not expected to play a major role in the daily cycle of trade-wind cumuli (Brill and Albrecht, 1982; Vial et al., 2019). Indeed, as shown in Figure D.8e, the daily amplitudes of SSTs remain small, ranging between 0.1°C and 0.2°C across the different patterns. Moreover, there is an overall tendency for higher SSTs during the day, at times when the cloud cover is minimum, revealing that the nocturnal increase in cloud cover is not forced by surface warming (a somewhat obvious fact).

The dependence of the SST daily cycles on the mesoscale patterns of organization and associated wind speed does not appear straightforward. Whereas there is an anti-correlation between the daily cycles of the wind speed and of the SST for the 3-winter climatology (black curves in Figures D.8d,e), consistently with previous observational analyses over this area on long time scales (e.g., Xie, 2004), more diverse relationships are found at the daily timescale for the individual patterns. We note, for instance, a time shift of about 6 hours between the daily maximum in wind speed and the daily minimum in SST for the patterns Sugar and Flowers. Moreover, there does not seem to be a linear relationship between the daily-mean wind speed and SST among the different patterns. These results therefore suggest that different factors (e.g., precipitation, upper ocean eddies) might affect the air-sea coupling on short time scales and depending on the mesoscale pattern of cloud organization.

D.6 CONCLUSIONS AND DISCUSSIONS

High-frequency geostationary satellite observations over the tropical Atlantic ocean and ground-based remote sensing measurements from the Barbados Cloud Observatory (BCO) are used to explore how the daily cycle of cloudiness in the winter trades depends on the spatial organization of shallow convection. We focus on the four prominent patterns of cloud organization of this region – Sugar, Gravel, Flowers and Fish – that have been characterized recently (Stevens et al., 2020). We apply two existing classification

methods on 30-minute infrared brightness temperatures from GOES-16 ABI to sample the daily cycle of these four forms of organization: one based on a neural network (referred to as NN) and one based on the mean size and degree of clustering of segmented cloud objects (referred to as I_{org}/S). A fifth category is also considered for unclassified mesoscale cloud scenes. Although these two classification methods are quite different in nature, they both yield qualitatively similar results, which are summarized hereinafter:

1. All forms of mesoscale organization exhibit a pronounced daily cycle in their frequency of occurrence, with distinct phasing and amplitude. The patterns Fish and Sugar preferentially occur during daytime, with a frequency peak around noon for Fish and around sunset for Sugar. The patterns Gravel and Flowers occur more frequently during nighttime; Gravel maximizes around midnight and Flowers at early morning hours before sunrise. From a more quantitative point of view, the daily characteristics of pattern occurrence (mean, phase, amplitude) are somewhat dependent on the classification method and on the geographical location of the patterns. The dependence of pattern occurrence to the large-scale environmental factors, such as the east-west gradient in near-surface wind speed, can explain some of the geographical disparities.
2. The daily cycle in cloudiness for a given pattern is relatively weak compared to the differences in cloudiness between the patterns. It is also fairly independent of the pattern and its geographical location: any given pattern cloud cover is maximum at nighttime (between 00LT and 03LT) and minimum in the afternoon (between 12LT and 15LT).
3. As a result of points 1 and 2, the effective contribution of patterns to the daily cycle in total cloudiness is to a large extent mediated by the frequency of pattern occurrence. The Fish pattern, which mostly occurs during the day, explains about 30% of daytime cloud cover. The contribution to total cloud cover from the Sugar pattern is the most important around sunset, representing up to about 25% of total cloud cover at this time at BCO. Gravel is the dominant form of organization around midnight (explaining up to 45% of total cloudiness), and Flowers can contribute up to 40% at early morning hours before sunrise.
4. A significant contribution of total cloud cover is also associated with unclassified organization, especially during daytime, which happens to be at the time of minimum cloud cover. Nevertheless, and although the importance of the unclassified contribution depends on the classification method, we find that the mesoscale patterns of cloud organization can explain to a large extent the daily cycle in total cloud cover.

A more detailed analysis of the cloud vertical distribution and precipitation at BCO allows to connect our findings with the description of the daily cycle of shallow cumuli made in Vial et al. (2019). First, they showed that during daytime a population of very shallow clouds grows, reaches a peak at 18-19LT and decays until dawn. Here we demonstrate that this behaviour is associated with an increased occurrence frequency of the Sugar pattern towards sunset. Second, the overnight cloud deepening discussed in Vial et al. (2019) is here primarily connected to the increased occurrence of the Gravel pattern, and to a lesser extent, to an overall deepening of clouds embedded in the Sugar

and Gravel patterns. Third, the dawn peak in cloud cover owing to the spreading of a stratiform cloud layer below the trade inversion (Vial et al., 2019), is here connected to a maximum occurrence of the Flowers pattern at this time of the day. It thus appears that the daily cycle in the occurrence of Sugar, Gravel and Flowers together may, to some extent, explain the evolution of trade-wind cloudiness from the shallowest cumuli in late afternoon to the nighttime population of deeper cumuli. These insights raise the question of the auto-correlation time-scale of individual patterns, and of the evolution from one pattern to another, which we leave for future investigation. In that respect, we expect the Fish pattern to be somewhat different, as it appears more strongly connected to non-local synoptic-scale disturbances persisting on timescales longer than a day (Aemisegger et al., 2021; Schulz et al., 2021). This is further supported here with a relatively weak daily cycle in the occurrence frequency of the Fish pattern.

The early morning peak in surface precipitation, identified in previous studies (Nuijens et al., 2009; Vial et al., 2019), is here associated with both the peak occurrence of the nocturnal patterns (Gravel and Flowers) and the enhanced rain rate at the end of the night for the ‘rainy’ patterns (Gravel, Flowers and Fish), regardless of their time of occurrence. Interestingly, for those three patterns (and especially for Gravel), the rain rate peak tends to succeed the cloud cover maximum by a few hours, which might be related to cold pools. A hypothesis that should motivate further investigation to ascertain whether cold-pools actually play a role in the phasing of rainfall.

Finally, we find that the large-scale near-surface wind speed can explain some of the geographical disparities in pattern frequency, it can also robustly discriminate between daytime and nighttime patterns, and it is fairly related with the daily cycle in pattern cloudiness regardless of the pattern in place. These insights, combined with findings from previous studies (Bony et al., 2020; Brueck et al., 2015; Nuijens et al., 2015), suggest that the strength of the trade winds are overall tightly connected to cloud amount and organization over a wide range of time scales from sub-daily to inter-annual.

D.7 SUPPORTING INFORMATION

D.7.1 *Examples of known ambiguities in NN-pattern detection*

We consider the NN approach as more subjective (compared to the I_{org}/S method) in the sense that the NN algorithm is trained from satellite images classified by the qualitative and subjective human eye (Rasp et al., 2020; Stevens et al., 2020). This subjectivity inevitably gives rise to ambiguous detections, and this is why there are a certain number of overlaps between multiple pattern rectangles. In addition to Figure D.1b, we here provide two more Examples of such ambiguous classifications that were identified in previous studies (Bony et al., 2020; Stevens et al., 2020). One is due to the confusion between the patterns Flower and Fish (Figure D.9, left) and the other one is between Gravel and Sugar (Figure D.9, right).

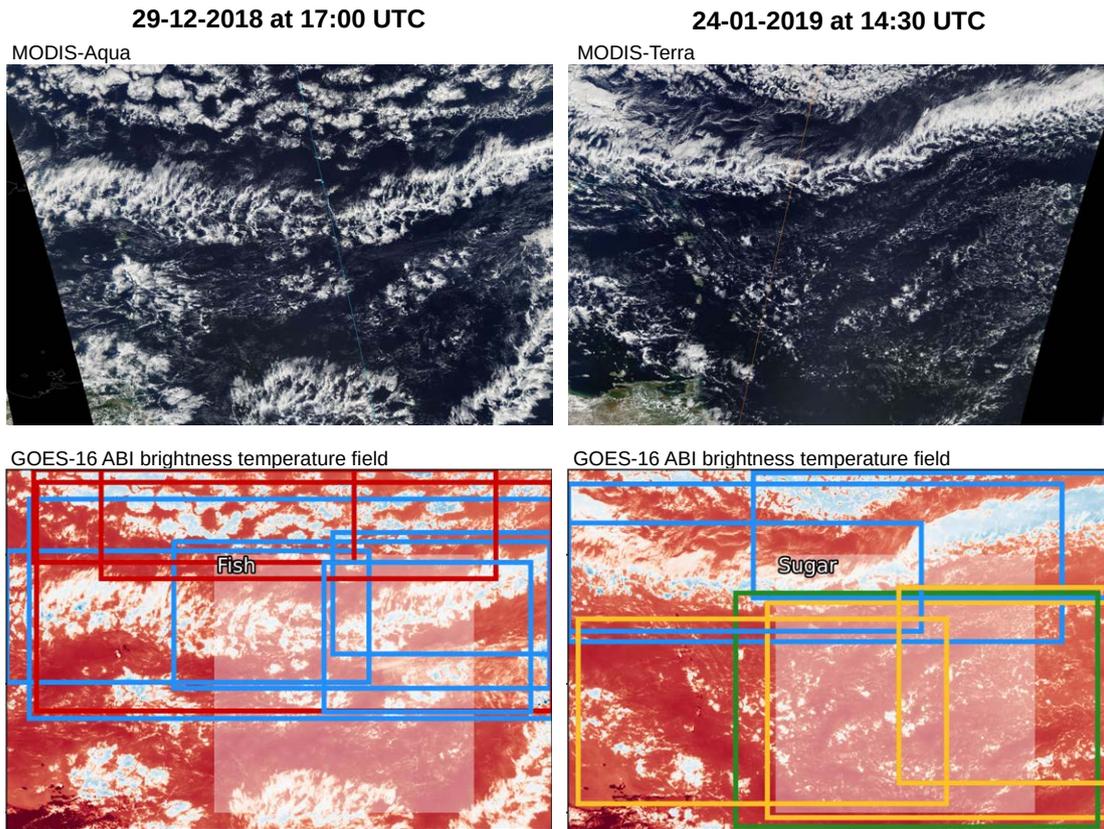


Figure D.9: Same as Figure D.1b but for different dates. The color code for the rectangles is the same as the one used throughout the manuscript: Red is for Flowers, Blue for Fish, Yellow for Gravel, Green for Sugar. The I_{org}/S classification is indicated in the grey area.

D.7.2 Characterization of shallow cloud organization from the paired (I_{org} , S) distributions

Figure D.10 (left panel) shows how the four patterns are distributed according to their I_{org}/S values (Sugar: green, Gravel: yellow, Flower: red, Fish: blue, undefined: grey). The two panels on the right show the distribution function of I_{org} and S . Note that the figure on the left is the same as Figure 1 in Bony et al., 2020, but using a different dataset, sampling frequency and period. In particular, the shape of the I_{org} and S distributions are similarly skewed toward high I_{org} and low S values. However, the distribution S is here shifted toward lower values than in Bony et al., 2020, which is presumably due to the higher spatial resolution of GOES-16 ABI T_b field (2 km) than the GridSat-B1 product (0.07°) used in Bony et al., 2020.

D.7.3 Impact of multiple pattern overlaps on pattern frequency and cloud cover

We here test the robustness of our conclusions to the degree of overlaps between multiple NN-detected patterns (Figures D.11 to D.13). This is done in two ways.

First, we vary a score of the neural network algorithm – that we call a *classification score* – which measures how well the detected cloud structure fits to that of a specific pattern

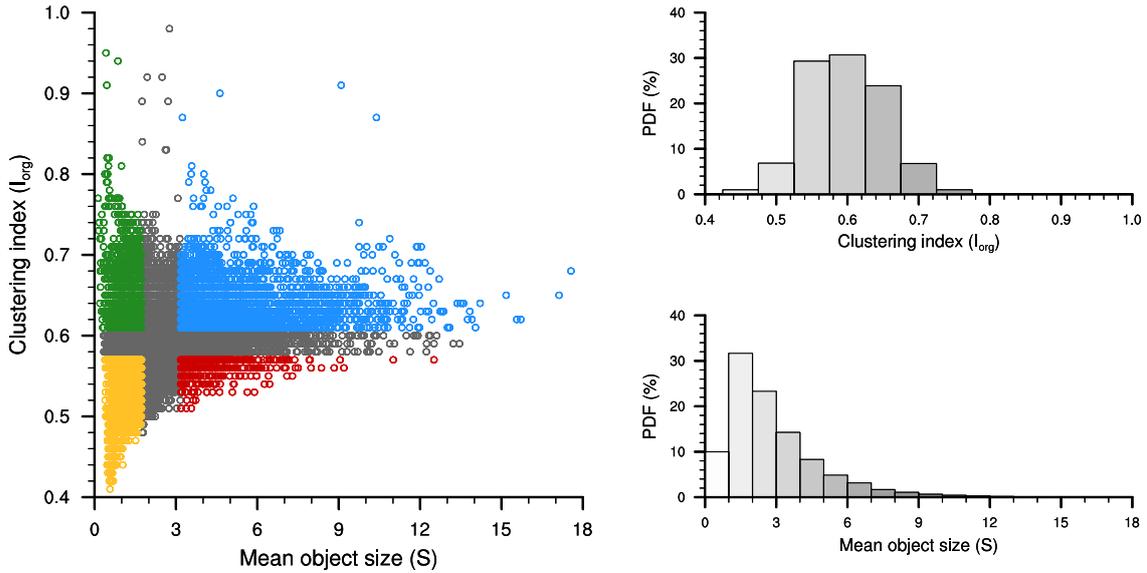


Figure D.10: Left panel : characterization of shallow cloud organization from the thresholding by terciles of the I_{org} and S distributions: Sugar is shown in green, Gravel in yellow, Flowers in red, Fish in blue, and the undefined regime is shown grey. Right panel : distribution functions of the clustering index (I_{org} , top) and mean size of cloud objects (S , bottom).

(Figures D.11 and D.12). When the score is increased, the chance to detect overlapping pattern rectangles is reduced and the sum of pattern areas (including unclassified areas) approaches the domain area. The default value of the score used in this study is 0.4 (Figures D.11 and D.12, solid line). As shown in Figure D.11, the *error* due to the overlaps is substantially reduced with a score of 0.5 (dotted line) and almost zero with a score of 0.6 (dash line). The drawback of increasing the score is that less patterns are detected overall: in Figure D.12 (right panels), the temporal frequency of pattern occurrence (f_k) is substantially reduced as the score is increased. Figure D.12 (left panels) shows that the phase and amplitude of the daily cycles in cloud cover for a given pattern (CC_k) are, however, quite robust and independent of the classification score. The large differences that are seen occur essentially for a score of 0.6, for which the sample size of detected patterns per category becomes quite low (Figure D.12, right panels).

Second, we compare, in Figure D.13, the daily cycle in pattern frequency and cloud cover between two data samples: one containing all detected patterns overlapping the BCO site at a given timestep (thus including multiple pattern overlaps), and one containing only the timesteps with single pattern detection at BCO. Figure D.13 shows that the daily cycles are very similar whether we allow multiple pattern overlaps or not.

Overall these two supplemental analyses show that multiple pattern overlaps have a minimal effect on the daily cycles of occurrence frequency and cloud cover of mesoscale patterns.

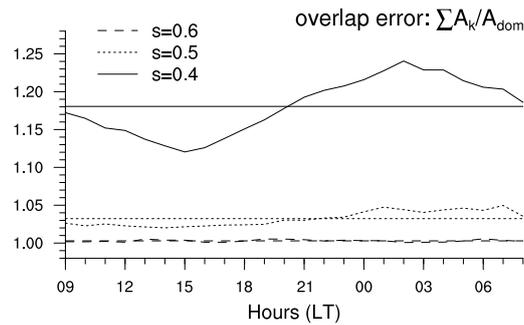


Figure D.11: Daily cycle of the ‘error’ due to overlaps between several rectangles of different labels for three values of the classification score. The horizontal lines are the daily means of the error. The error is actually the sum of all pattern areas (including the no classification) normalized by the domain area. A value of 1 means no error.

D.7.4 Sensitivity of pattern cloud cover to spatial scales

Figure D.14 shows the GOES-16 ABI cloud cover for patterns overlapping BCO (as in Figure D.4c), but the cloud cover averaging for each detected pattern rectangle is limited to a 20 km² subdomain centered over BCO (Figure D.14a), or located 10 km east of BCO (Figure D.14b). The comparison between Figure D.14 and Figure 4c shows that the daily cycles of pattern-related cloud covers are sensitive to the spatial scale used to compute the cloud covers. Therefore, we presume similarly that the results at BCO (Figure D.7) and at NTAS (Figure D.8f,g) are to some extent affected by the relatively small temporal interval (used to average the surface or atmospheric field for a given pattern) compared to the size of patterns. On a somewhat different level, the comparison between Figure D.14a and Figure D.14b shows that the results are very similar whether they are focused over BCO or east of BCO, suggesting a negligible effect from the island (e.g., land/sea breeze).

D.8 ACKNOWLEDGMENTS

This work benefited from the many valuable discussions within the EUREC⁴A (Elucidating the role of clouds-circulation coupling in climate) science team, and we thank the two anonymous reviewers for their insightful comments. This research was funded by the European Research Council Grant Agreement 694768 and the EU Horizon 2020 Grant Agreement 820829. The BCO data used in the analysis and other supplementary information that may be useful in reproducing the authors’ work are available from the first author on request. Data from the NTAS Ocean Reference Station were collected and published by Dr. Albert Plueddemann of the Woods Hole Oceanographic Institution. The data is made freely available by the OceanSITES project and by the national programs that contribute to it. The GOES-16 ABI data are publicly available online at doi.org/10.7289/V5BV7DSR.

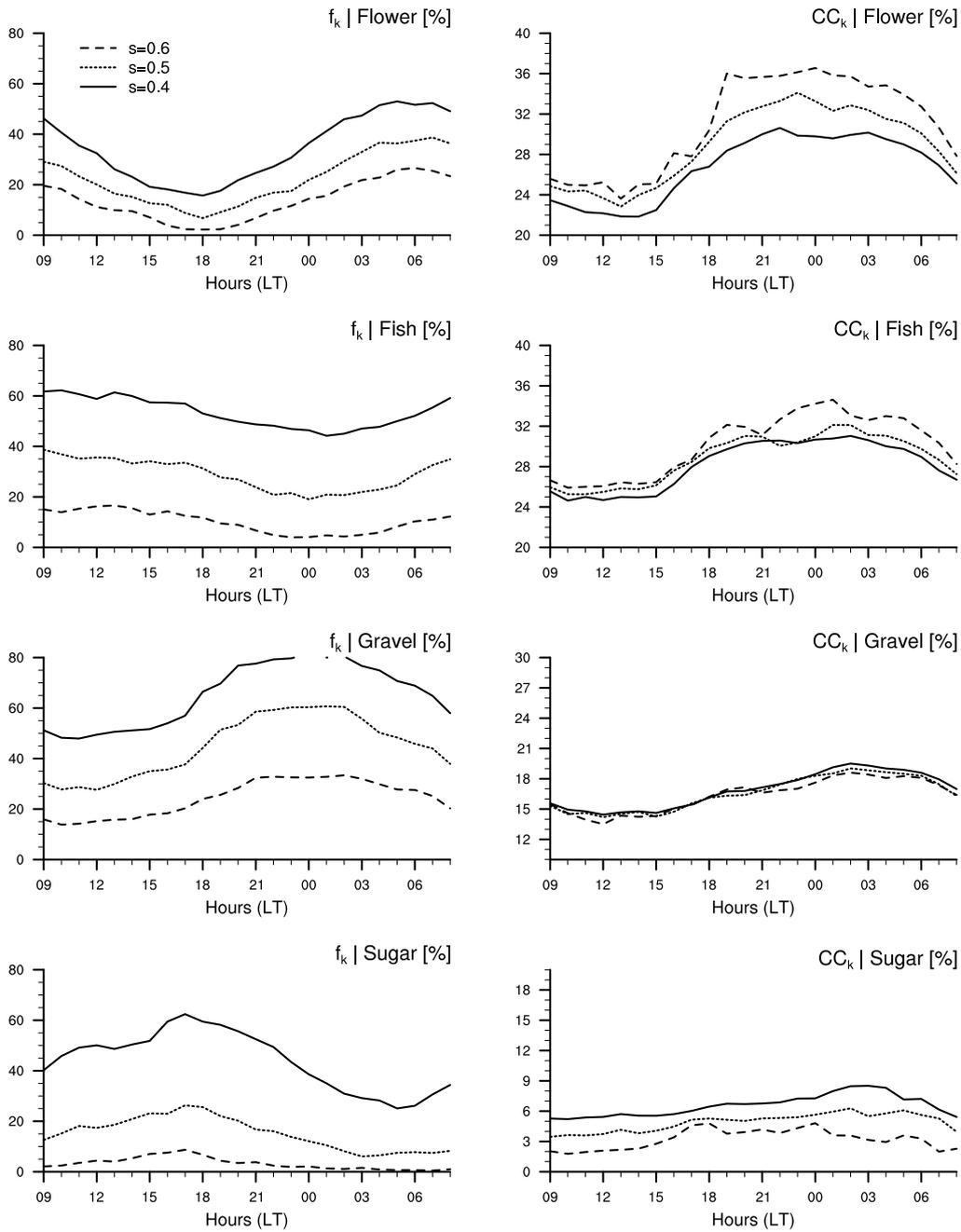


Figure D.12: Daily cycle of pattern temporal frequency of occurrence (f_k , left panels) and pattern cloud cover (CC_k , right panels) for different values of the agreement score. The default score, used in the study is $s = 0.4$.

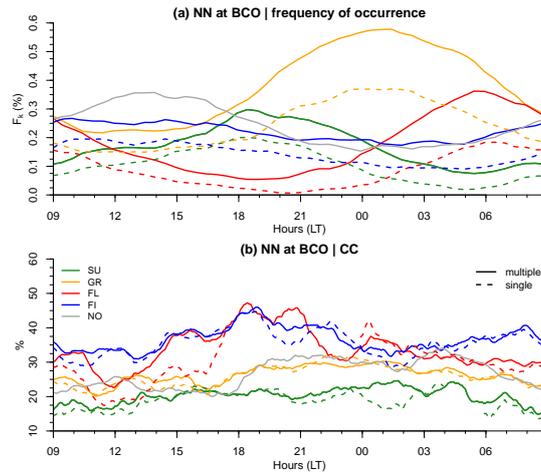


Figure D.13: Daily cycle of pattern frequency (f_k , a) and BCO radar-derived cloud cover (CC_k , b) when NN-detected patterns overlap the site BCO. The solid lines refer to the case for which we allow multiple pattern occurrences over BCO and the dash lines refer to the case for which only one pattern is detected at a given timestep.

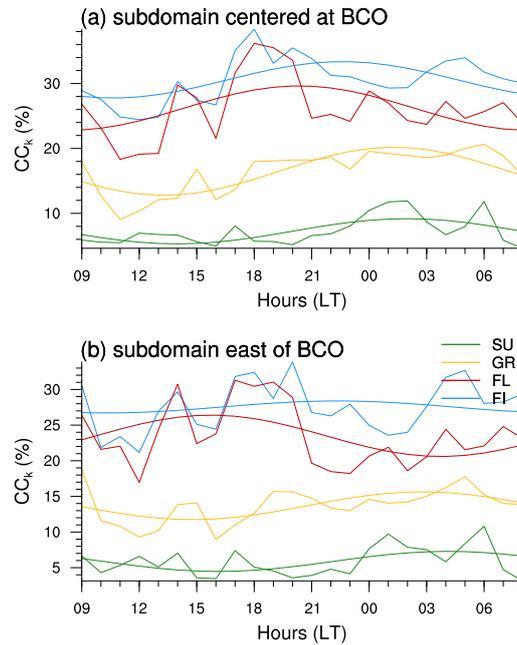


Figure D.14: Daily cycle of pattern cloud covers diagnosed from GOES-16 ABI for patterns overlapping BCO. The cloud covers are calculated over the part of the patterns that is contained in a 20×20 km² sub-domain (a) centered at BCO and (b) 10 km east of BCO.

ASSESSMENT OF MESO-SCALE PATTERNS OF SHALLOW CONVECTION IN REALISTICALLY FORCED LARGE-EDDY SIMULATIONS

The work in this appendix is intended for publication as:

Schulz, Hauke and Stevens, Bjorn (in preparation). Assessment of meso-scale patterns of shallow convection in realistically forced large-eddy simulations

The contributions of the authors to this publication are as follows:

The conceptual design, research focus, analysis and writing have been planned and conducted by HS. In addition, HS set-up and ran the ICON-LES simulations. BS supervised the study.

Assessment of meso-scale patterns of shallow convection in realistically forced large-eddy simulations

Hauke Schulz¹, Bjorn Stevens¹

¹Max Planck Institute for Meteorology

E.1 ABSTRACT

The variability in meso-scale patterns of shallow convection in the trades has recently been identified to alter the mean CRE. To understand the mechanisms leading to these different patterns the utility of large-eddy simulations is tested running 37 days nested ICON-LES on domain sizes ranging between $(700 \text{ km})^2$ and $(1300 \text{ km})^2$. The simulation is realistically forced to match the conditions captured during the EUREC⁴A field campaign. Forward operators are used to compare the simulation output with the observations. The cloud cover is captured well on average but with discrepancies in its vertical and spatial distribution. Cloudiness at the lifting condensation level depends on the model resolution with the finer one producing on average a more realistic cloud profile. Independent of the resolution, the variability in cloudiness below the trade inversion is not captured, leading to a lack of stratiform cloudiness with implications on the detectability of meso-scale patterns whose cloud patches are characterized by stratiform clouds. Still, these patterns show a meso-scale circulation making the cloud-top cooling an unlikely driving mechanism. Overall, the simulations prove to be a valuable tool to complement the observations and guide future studies.

E.2 INTRODUCTION

The importance of shallow convection has been recognized as a vital part of the net cloud radiative effect mostly due to their sheer ubiquity and large cloud fraction, especially in the eastern ocean basins where cold ocean currents and the overlying warm air form extensive cloud decks. The trade winds advect the cloud decks altering their shape from closed-cells to open-cells until randomly scattered shallow cumuli are the dominating cloud species in the downwind trades. This is however a rather idealized and simplified view. In reality, the variability of cloudiness in the downwind trades is much larger according to Riehl (1954), Stevens et al. (2020), Schulz et al. (2021). The latter two characterize this variability by the cloud meso-scale organization and Bony et al. (2020) further attributed different cloud radiative effects to different forms of organization. In addition, Schulz et al. (2021) showed that these forms of organization have distinct cloud fractions and they are observed along with distinct meteorological environments.

To estimate how the distribution of patterns and their impact on the net cloud radiative forcing will change in a changing climate, we first need to better understand the processes that lead to the co-variability of cloudiness and its environment.

Can large-eddy simulations help us understand the physical processes of meso-scale patterns of shallow convection in the downwind trades? More precisely:

1. Do simulations reproduce the observed variability in cloudiness and its meso-scale patterns?
2. Can meso-scale processes be associated with the patterns and how do they differ among them?

To answer these questions, we conducted regional large-eddy simulations (LES) for the duration of the EUREC⁴A field campaign. The EUREC⁴A period offers a rich variety of meso-scale patterns and a great framework to evaluate the simulations with a multitude of observations.

This study is structured as follows: Section E.3 describes the simulation configurations, observations and the forward operators used to better compare the simulation output with measurements. Section E.4 discusses the similarities of cloudiness in LES and observations, while Section E.5 investigates if hypothesized formation processes are evident. We conclude with Section E.6.

E.3 DATA AND METHODS

E.3.1 *Large-eddy simulations*

We focus on the downwind trades of the North Atlantic from January - February 2020 when this area has been intensively sampled as part of the EUREC⁴A field experiment (Stevens et al., 2021). We conducted simulations with the ICOSahedral Nonhydrostatic (ICON) model family (Zängl et al., 2015) at grid spacings of 1.25 km (ICON-SRM), 624 m (ICON-624m), 312 m (ICON-312m) and 156 m (ICON-156m). With the exception of the storm-resolving simulation (ICON-SRM), which is used for the initialization and the lateral boundary conditions only, all simulations are based on the large-eddy simulation capabilities introduced by Dipankar et al. (2015) and Heinze et al. (2017). This branch of the model is called in the remainder ICON-LEM.

The domains (Fig. E.1) of the large-eddy simulations are oriented along the trade-winds to maximize the temporal coverage of the evolution of the shallow convection. The eastern borders of the nested domains decrease by at least two degrees with each refinement to reduce numerical artifacts in the mean wind direction of the trades. On the western borders these reductions only affect high clouds that are advected by the sometimes called *anti-trades* from western directions. These clouds are however not the focus of this study and are not expected to interfere with the low-level cloudiness.

With domain sizes between about $(700 \text{ km})^2$ and $(1300 \text{ km})^2$ the simulations are able to capture the meso-scale variability in its full extend. Stevens et al. (2020) defined meso-scale patterns, ranging from very fine clusters called *Sugar* to larger scale *Fish* pattern, should all be reasonably well resolved as they identified the patterns on a domain of similar size.

The simulations were created in two separate steps. First, the ICON-SRM simulations were run to calculate the initial and boundary conditions for the nested LES. Second, the nested LES was started with the ICON-624m simulation. Once initialized with the ICON-SRM output, it was forced at its lateral boundaries with hourly output. For each 24 hour set of lateral boundary conditions, the ICON-SRM was granted a spin-up of 16 hours. Sea surface temperatures or to be precise, sea skin temperatures, were updated

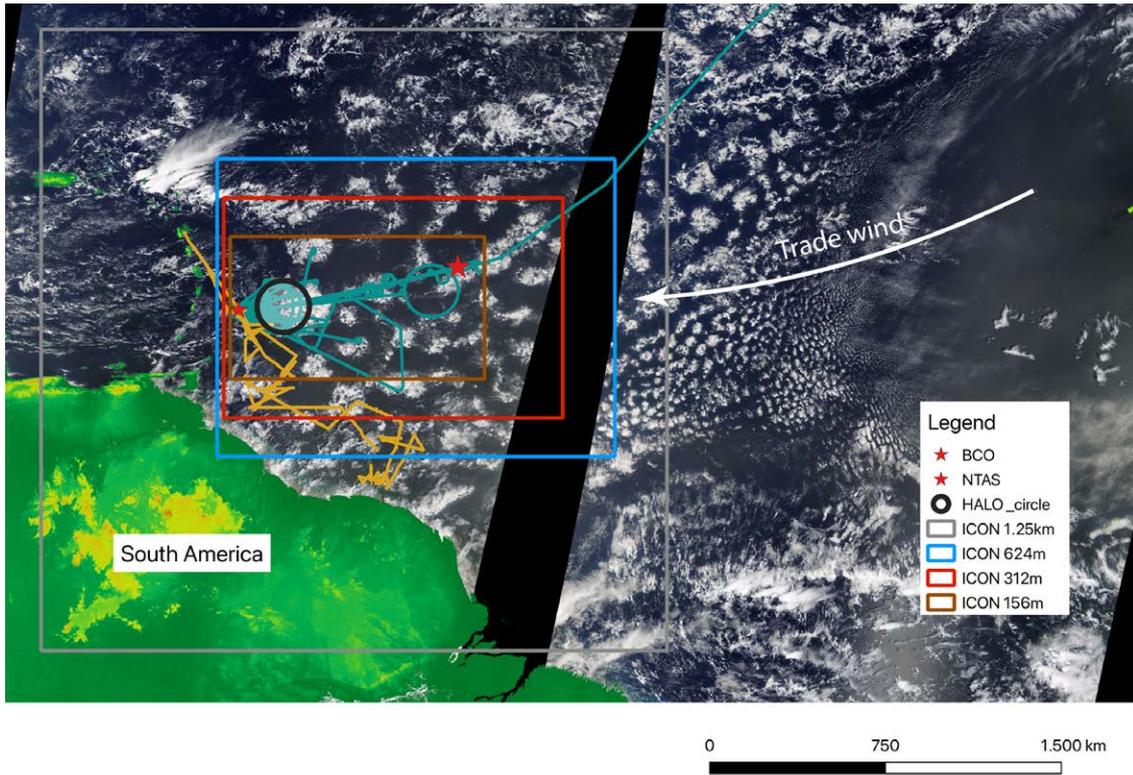


Figure E.1: Overview of simulation domains ICON-SRM (gray), ICON-624m (blue), ICON-312m (red), ICON-156m (brown). The tracks of the platforms HALO and L'Atalante, which are representative for the two different measurement foci of the EUREC⁴A campaign are shown in orange and purple, respectively. Standard HALO circle is shown in black. The location of the BCO and the NTAS buoy are marked with a red star at the western and eastern part of the domain, respectively. For a sense of scale, the MODIS image of February 12 is shown with landmasses colored in green to brown depending on height.

every timestep based on linearly time-interpolated hourly ERA5 skin temperatures. The skin temperatures were chosen because the used ICON version does not have a skin temperature parameterization. Different to the ICON-SRM, the ICON-LEM has not been restarted but ran for the whole period of 09.01.2020 to 14.02.2020. As a result only one spin-up period exists for each nest. ICON-624m has been started at 10 UTC and ICON-312m at 16 UTC of January 9, 2020. We analyze the output after midnight of January 10, 2020.

The simulations ICON-312m and ICON-156m were one-way nested into the ICON-624m domain, while the later was only activated for one case due to its computational cost. Fig. E.2 shows a snapshot of this case for all domains and illustrates that the structure in the humidity field is captured in all cases. It should be noted that ICON-156m snapshot is taken 8 hours after the initialization and spin-up might still affect the output. Nevertheless, it shows a consistent structure with the other domains and can be used for future investigations. The specific configuration of all simulations is summarized in Tab. E.1

Table E.1: Overview about used simulation configurations.

Model	ICON-SRM	ICON-624m	ICON-312m	ICON-156m
No. grid cells x levels	6773696 x 75	4528560 x 150	11792076 x 150	24469588 x 150
hor. gridspacing*	1248 m	624 m	312 m	156 m
ver. gridspacing (lowest level/ 1000 m/ 2000 m)	20 / 140 / 190 (150 levels)	20 / 70 / 85 (150 levels)	20 / 70 / 85 (150 levels)	20 / 70 / 85 (150 levels)
Model top (km)	35	21	21	21
hor. domain	67W - 43W, 0N-24N	60.25W - 45W, 7.5N - 17N	60W - 47W, 9N - 16.25N	59.75W - 50W, 10.5N - 15.5N
Forcing (except SST)	ECMWF IFS (hourly) ⁺	hourly ICON-SRM	one-way online nesting	one-way online nesting
Forcing (SST)	ECMWF IFS (fixed at initial time)		continuously (linear interp. from 1h ERA5 SKT)	continuously (linear interp. from 1h ERA5 SKT)
restart	daily at 00 UTC		none	none
Turbulence	turbulent kinetic energy (TKE)		Smagorinsky diffusion	
Microphysics	One moment incl. graupel (Baldauf et al., 2011)		Two moment (Seifert and Beheng, 2006)	
Cloud-scheme	Fractional cloud cover		All-or-nothing	
Radiation		RRTM (Mlawer et al., 1997; Stevens et al., 2013)		

* edge lengths of triangular grid cells

⁺ analysis at 00:12 UTC ; otherwise IFS forecast

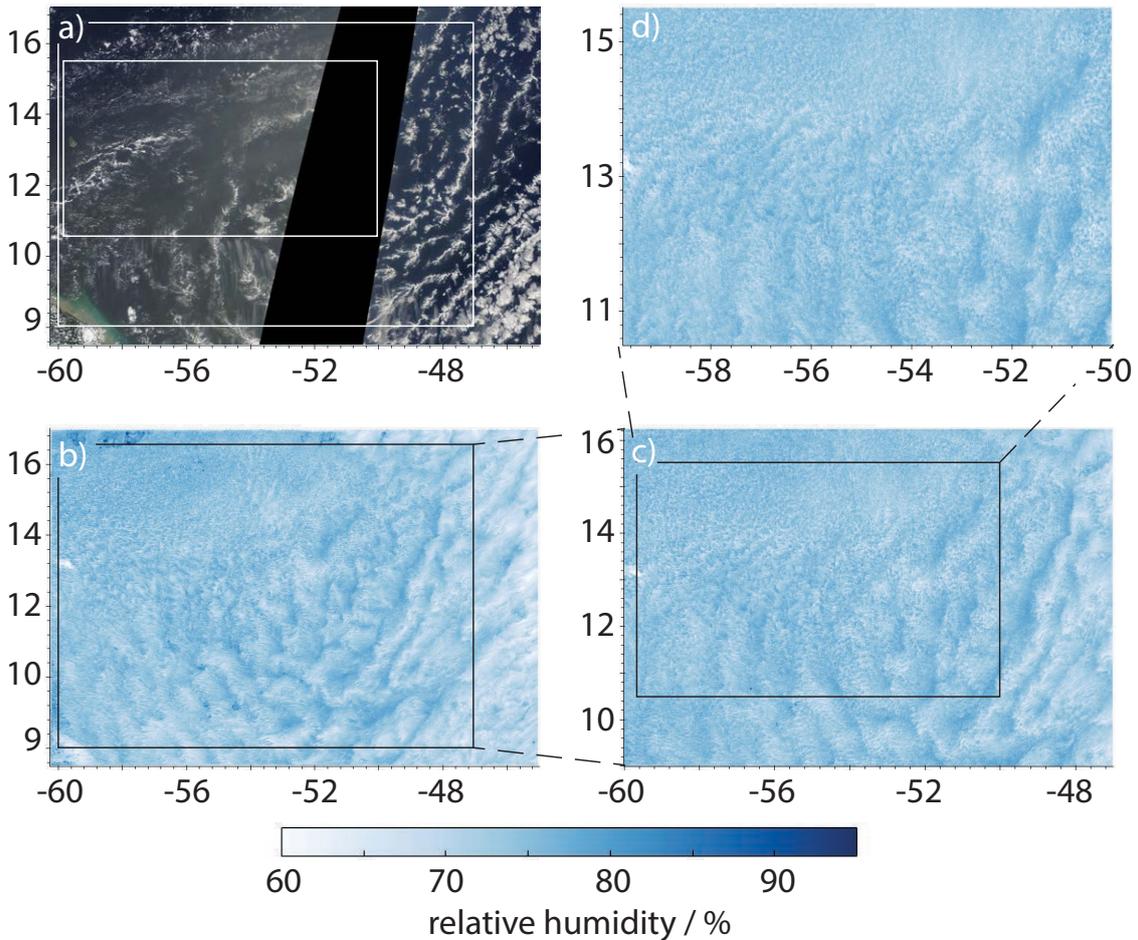


Figure E.2: MODIS visible satellite image on February 1, 2020 at about 13:10 UTC (right swath) and 14:50 UTC (left swath) and the simulation output at 14:30 UTC. The relative humidity at 2m is visualized for the simulations. The simulations ICON-624m, ICON-312m and ICON-156m are shown in b), c), d), respectively.

E.3.1.1 *Satellite forward simulator*

To compare the output of the LES with satellite observations, we rely on the RTTOV forward simulator (Saunders et al., 2018), which can emulate satellite imagers and how they would capture the simulation output. In this study, we use the GOES-16 ABI specifications to compare them to the actual satellite's instrument, which covers the region of interest with a high temporal and spatial resolution of 10 min and 2 km (channel 13: 10.35 μm), respectively. In an attempt to get the most consistent synthetic satellite images, we made modifications to the most recent version of ICON (2.6.3). These modifications include design changes that let us use RTTOV v13 during the run time of ICON and reduce the amount of data that needs to be saved to disk for offline calculations. In addition, we use the calculated two-moment microphysics to feed both the internal RRTM radiation and the RTTOV forward operator to make them consistent with each other. Otherwise the radiation scheme and RTTOV would use their own independent parameterisations leading to an inconsistent result.

The synthetic satellite images are calculated every 10 minutes to match the temporal resolution of the ABI instrument. A snapshot of the animation ([10.5281/zenodo.5553825](https://zenodo.org/record/5553825)) that visualizes the actual and synthetic satellite images for the complete time period is shown in Fig. E.3.

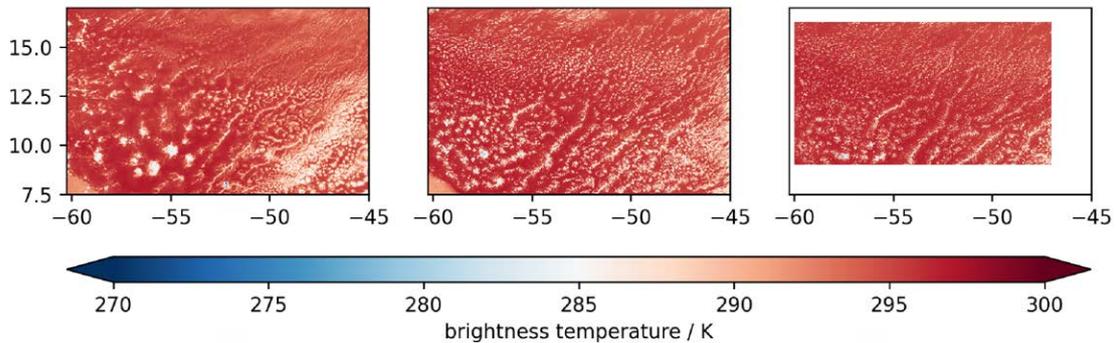


Figure E.3: Comparison of GOES-16 ABI channel 13 satellite image (left) and the synthetic counterparts from ICON-624m (middle) and ICON-312m (right) for February 2, 2020 at 7:50 UTC.

E.3.1.2 Radar forward simulator

Satellite images are a great tool to analyze the meso-scale cloud formation. However, to better understand the formation mechanisms the vertical cloud distribution is important. As shown by Schulz et al. (2021) these patterns have cloud contributions from two layers. Clouds with a cloud base height at the lifting condensation level and those that exist further aloft below the trade inversion. The interplay between these layers has always been challenging for models to reproduce (e.g. Atlas et al. (2020)).

Because cloudiness is not a very well defined quantity and dependent on the instrument or model, we also use a forward simulator to resemble the vertical distribution of cloudiness that a radar would measure. In particular the Ka-Band radar at the BCO (see next section) because this was capturing the advecting trade-wind clouds during EUREC⁴A and was used for Schulz et al. (2021).

We rely on the radiative transfer simulator PAMTRA (Passive and Active Microwave TRANSfer package) (Mech et al., 2020) that has successfully been used with the same radar frequency in earlier studies in this region (Jacob et al., 2020).

PAMTRA has been configured similar to Mech et al. (2020) to match the two-moment microphysics scheme of Seifert and Beheng (2006) which has been used in the LES of this study. Hence, PAMTRA is able to infer the original particle size distribution assumed by the simulations from its bulk measures of mixing ratio and number concentration that are saved every 60 s at the location of the BCO. PAMTRA is therefore able to simulate reflectivities that are consistent with microphysics used in the LES. The inconsistency caused by the Ka-band radar measuring reflectivity, the sixth moment of the particle size distribution, directly is assumed to be negligible. The reflectivities are only used here to create a cloud mask, which further justifies this assumption.

E.3.2 Observations

E.3.2.1 Barbados Cloud Observatory

Besides the tremendous amount of observation platforms that were present in the simulated area during the EUREC⁴A time period, the Barbados Cloud Observatory is taking long-term measurements at the most windward tip of the Caribbean island, Barbados (Stevens et al., 2016). We used the measurements from the vertically pointing Ka-band radar to detect the vertical distribution of hydrometeors. Averaging these measurements in time results in echo fractions which are a combined measure of cloud fraction and precipitation fraction. A threshold of -50 dBZ has been applied to exclude backscatter from sea-spray that is otherwise frequently detected in the lower range gates (Klingebiel et al., 2019).

E.3.3 Classifications of meso-scale patterns

This study uses two approaches to identify the meso-scale patterns of shallow convection. First, for identifying the days with observed canonical meso-scale patterns we rely on the manual classifications done by the scientific community of the EUREC⁴A field campaign (Schulz, *in preparation*). The scientists inspected satellite images captured during the EUREC⁴A time period and labeled regions containing *Sugar*, *Gravel*, *Flowers* or *Fish*. The canonical days shown in Fig. E.4 are used to separate the simulation days by pattern and study them individually. Second, to understand how well the patterns are replicated in terms of area fraction and occurrence, the simulations themselves needed to be classified. Here we make use of the satellite forward operator and the same neural network that has been successfully used to identify the patterns in observations (Schulz et al., 2021). With the forward operator the simulation output is converted into synthetic satellite images using the sensor characteristics of the ABI instrument onboard GOES-16 to retrieve synthetic infrared images at $10.35\ \mu\text{m}$ (channel 13), the same wavelength as used in Schulz et al. (2021). The neural network is run on both the simulation output and the GOES-16 ABI images. Cropped to the domain size of the simulations, this allow a fair comparison.

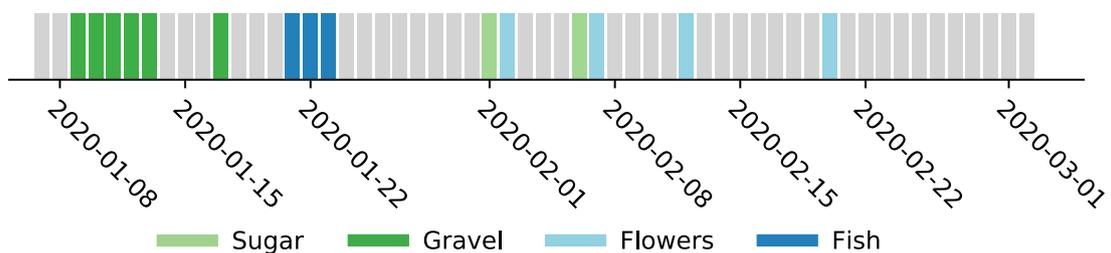


Figure E.4: Meso-scale patterns identified by the EUREC⁴A community in GOES-16 ABI infrared satellite images. (Schulz, *in preparation*)

E.4 SIMILARITY OF LES AND OBSERVATIONS

E.4.1 Meso-scale patterning

E.4.1.1 Visual inspection

The evaluation of the model output draws on the quantification done in Schulz et al. (2021). A general indicator of whether the patterns are represented or not, is the visual inspection of the spatial distribution of clouds just like the earlier cloud classifications. In this case, the cloud scenes are not captured by an actual satellite but are simulated from the model output by the satellite forward operator RTTOV.

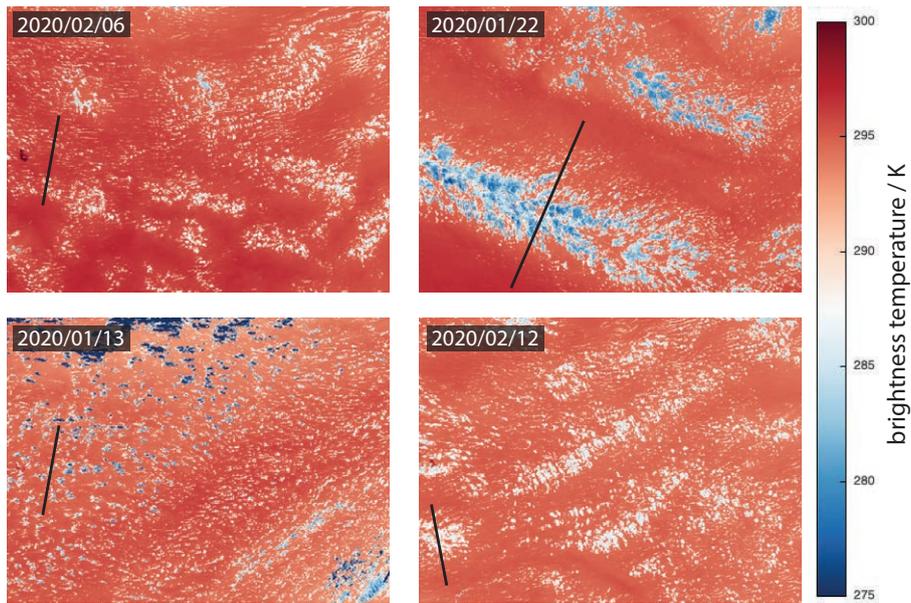


Figure E.5: Overview of simulated satellite images of ICON-624m matching the cloud scenes shown in Fig. 1.1. Different to Fig. 1.1 the infrared channel of ABI is shown. Cross-sections along the black line are shown in Fig. E.12.

The results are depicted in Fig. E.5. The scenes match those in Fig. 1.1 and should therefore show the same meso-scale patterns.

The visualisation already demonstrates that most of the scenes match the principle structure of the patterns with the exception of *Flowers*. *Flowers* are hardly distinguishable from the *Sugar* scene even when decreasing the grid-spacing from 624 m to 312 m (see supplemental Fig. E.14). The stratiform layers of the *Flowers* are hardly reproduced.

Fish and *Gravel* are more alike to their respective observations shown in Fig. 1.1. Their general pattern structures agree visually well with the observations. *Fish* shows band structures of cloudy and clear-sky patches and *Gravel* consists of much smaller patches that are roughly arranged in hexagons. Some clouds also rise deeper and produce stratiform clouds that are also visible for this day in the observations. The surface temperature field (not shown) also confirms the frequent and wide-spread occurrence of cold pools supporting the hypothesis about their importance.

Overall, it seems that the smallest cloud patches that one would normally attribute to *Sugar*, are occurring too widespread in the simulations.

E.4.1.2 Neural network

In contrast to the visual classification the NN detects cloud patterns more objectively as shown by Schulz et al. (2021). We ran the satellite forward operator during the simulation to get synthetic brightness temperature images on which the NN can be applied as described in Sec. E.3.3.

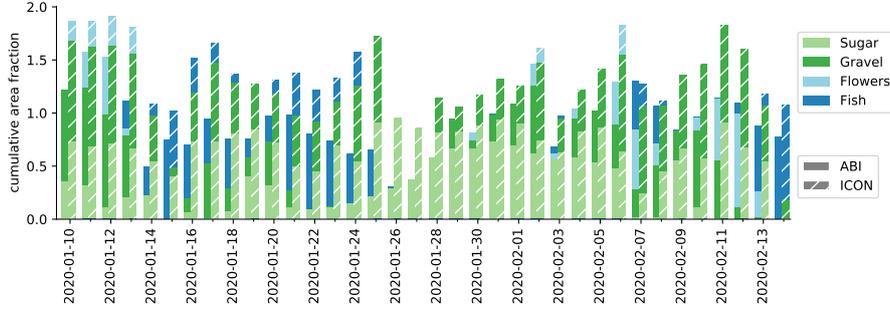


Figure E.6: Daily mean area fraction covered by meso-scale patterns as identified by the neural network on actual (ABI) and simulated (ICON-624m) satellite images.

Applying the NN on the actual and synthetic satellite images results in rectangular bounding boxes that enclose the respective meso-scale pattern. The area-fraction A that a pattern p covers at a certain time t can be calculated by setting the union of each pattern classification $l_{p,t,i}$ in relation to the total domain size D :

$$A_{p,t} = \frac{L_{p,t}}{D}, \quad \text{where} \quad L_{p,t} = \bigcup_{i=1}^n l_{p,t,i} \quad (\text{E.1})$$

The result of this calculation is shown in Fig. E.6 and is a measure of how dominant a pattern has been on a specific day. Note that the cumulative area fraction ($\sum_{p=1}^4 A_{p,t}$) can be larger than 1 due to the overlap of the pattern classifications. At first sight, it becomes visible that simulations and observations agree quite well in terms of the daily pattern occurrence. If a widespread pattern (high $A_{p,t}$) has been detected in the observations, it has also been classified in the simulation output. Nevertheless, systematic differences between the two classifications are visible.

The impression from the visual inspection of Fig. 5.1 that small cloud entities of the *Sugar* type are too dominantly occurring in the other organizational forms is reflected by the continuously large cloud area fractions of *Sugar*. In contrast, *Fish* and *Flowers* are rather limited in their area fraction in the simulations often at the cost of *Gravel*. In the periods from 09/01/2020 to 13/01/2020 and from 29/01/2020 to 05/02/2020 does the area fraction of *Gravel* and *Sugar* agree well with the observations. However, they are also widespread between 14/01/2020 and 25/01/2020 where rather *Fish* should dominate.

The agreement between patterns detected in the simulations and observations are quantified using the *Intersection over Union (IoU)*. Fig. E.7 shows this quantification and confirms that the agreement for *Flowers* and *Fish* is low, mostly due to the lack of *Flowers* and *Fish* in the simulations. The agreement for *Sugar* and *Gravel* is large with 0.43 and 0.35, respectively. For these pattern, IoU is reduced primarily by “missing” *Sugar* and *Gravel* in the observations.

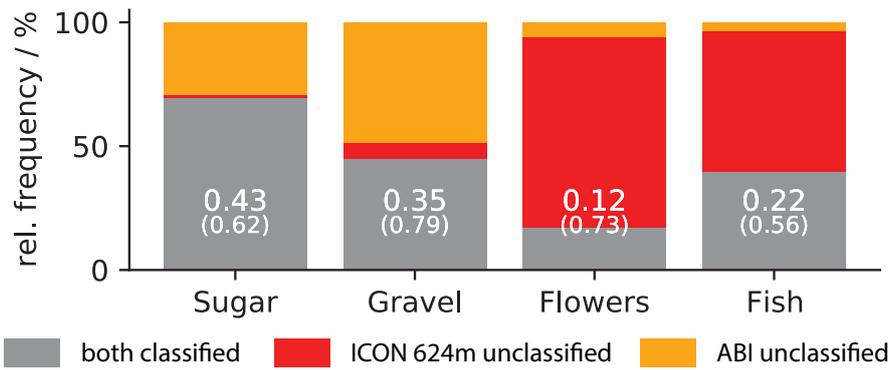


Figure E.7: Agreement between neural network classifications based on ABI images and synthetic satellite images of ICON624 measured by IoU.

If a pattern is detected simultaneously in both, simulation and observation (gray area in Fig. E.7), the overlap is fairly high indicated by an average IoU of 0.56 or greater. A similar picture is drawn from the ICON-312m run (supplemental Fig. E.17). Except for *Gravel* the overall agreement reduced with increased resolution. No *Flowers* were detected in ICON-312m. Remember that these classes have no ground-truth and the IoU is therefore systematically lower compared to traditional object detection exercises. Furthermore, slight differences in the extend of the patterns in observations and simulations are expected Hence, perfect alignment of cloud patches on the meso-scale is not expected.

The agreement between the visual inspection and the neural network classification demonstrate that the simulations replicate the observations. However, the objective analysis points out deficits.

E.4.2 Variability of cloudiness

E.4.2.1 Cloud cover

Despite the different frequency of occurrence of patterns in the simulations compared to the observations, a crucial factor that needs to be captured is the cloud cover. As Schulz et al. (2021) have shown, cloud cover is one of the most distinguishing factors between the different meso-scale patterns and is according to Bony et al. (2020) the primary factor influencing the CRE and ultimately the energy balance at the top of the atmosphere.

To compare the cloud cover of the simulations with the satellite observations, we rely again on the measured and simulated brightness temperatures. Similar to Bony et al. (2020) we define shallow clouds by a brightness temperature between 280 K and 290 K. The definition of cloud cover follows as the ratio of the number of shallow cloud pixels to the number of pixels of the entire domain. For a better comparison of the different sources a common domain has been chosen (10° N - 16° N and 60° W - 54° W).

The simulated cloud cover variability agrees well with GOES-16 ABI (Fig. E.8), both in terms of the overall time-series (Fig. E.8a) as well as the diurnal cycle (Fig. E.8c). The large discrepancies to the observations e.g. between the January 21 and January 27 go

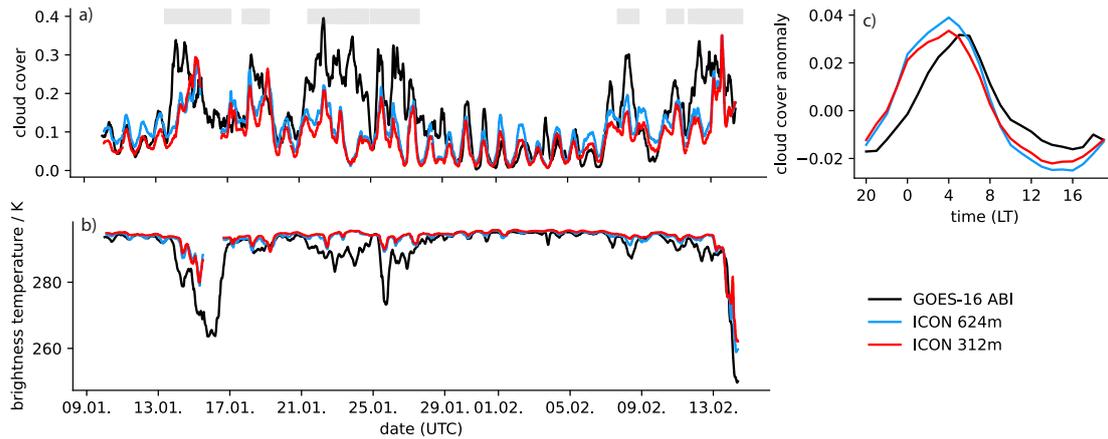


Figure E.8: Timeseries of cloud cover inferred from actual and synthetic satellite images (a) and the lower quantile of brightness temperature within the domain as an indicator of high clouds (b). The diurnal cycle without contributions from high-clouds (lower quantile brightness temperature below 290 K as indicated by gray bars in (a)) is shown in (c) as anomaly to the daily mean. Gap on 16 January is due to model output issues.

along with colder lower-quantile brightness temperatures. These brightness temperatures (Fig. E.8b) indicate that the domain is contaminated with high clouds. In cases where the lower-quantile drops below 290 K are marked with a gray bar in Fig. E.8a.

The diurnal cycle of cloud cover without the contributions of high clouds is shown in Fig. E.8c. It is shown as an anomaly to the mean cloud cover of 10%, 9% and 7% for GOES-16 ABI, ICON-624m and ICON-312m, respectively. Owing to the synthetic satellite images, we present a fair comparison of the diurnal cycle of shallow convection in both LES and observations. Based on Fig. E.8c, we note the quantitative agreement of the diurnal cycle in addition to the qualitative agreement found in Vial et al. (2019). The overall reduced cloud cover compared to Vial et al. (2019) might be caused by the different definitions of cloudiness. Furthermore, the time period covered might have less *Flowers* and *Gravel* that contribute most to the diurnal cycle in terms of cloud cover anomaly and frequency.

E.4.2.2 Vertical cloud distribution

In addition to the arrangement of the cloud patches, the cloudiness below the inversion distinguishes the patterns from each other (Schulz et al., 2021) and can serve as an indicator of how well cloud processes are captured. Here, we examine the vertical distribution by means of the high-frequency ICON-LEM column output at the location of the BCO. This has the advantage of evaluating the simulated cloudiness against the observations. As described in Sec. E.3.1.2 the output has been converted with the therein noted limitations to reflectivity as a common quantity with the observations.

The temporal average cloud fraction profiles and their variability is shown in Fig. E.9 for the observations, ICON-624m and ICON-312m. By inspection, the average cloud profile with its peak at about 800 m and a slow decrease in cloudiness with height is well captured. However, the variability is quite reduced compared to the observations independent of the resolution of the simulation. This is especially true for the cloudiness

between 1 km and 2 km where the stratiform cloud component is usually located. While the interquartile range is about 5% in the simulations, it is nearly doubled in the observations.

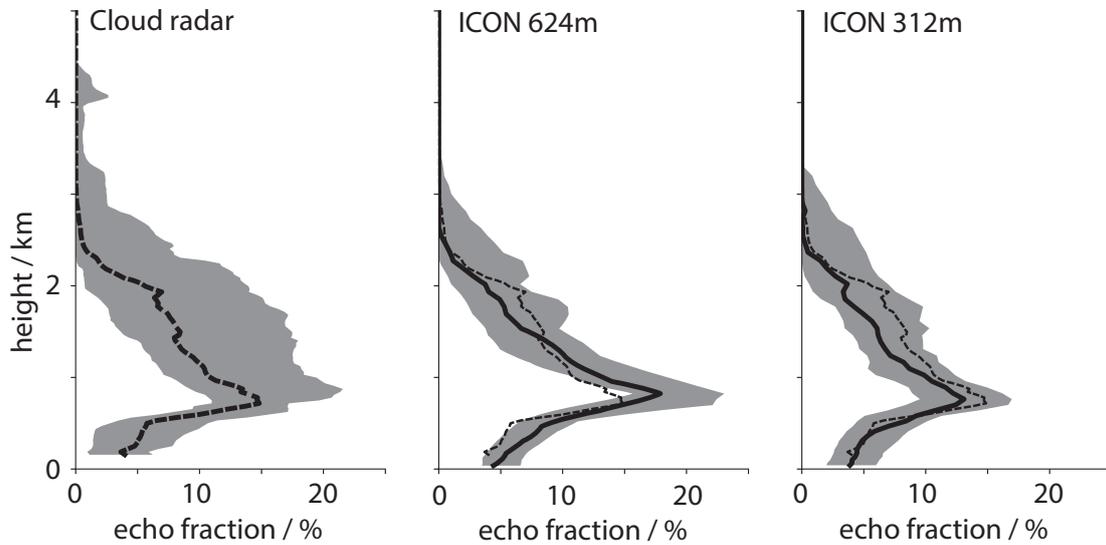


Figure E.9: Variations in daily mean echo fraction measured (left) and simulated (middle: ICON-624m; right: ICON-312m) echo fractions at the location of the BCO across the simulated time period. Median is depicted as line. Observations are drawn as dashed line for reference in each panel. Shading indicates interquartile range.

The mean profile is less cloudy in ICON-312m with a slight underestimation at the lifting condensation level. In contrast, ICON-624m overestimates the cloud fraction there. Overall, the slope with which the cloudiness reduces with height and the precipitation near the surface are better captured by ICON-312m.

Based on study C the meso-scale variability is connected to large-scale forces and air mass origins. Processes that lead to the patterns can therefore be very different from pattern to pattern. In the following, the echo fraction profile is separated by pattern to test if simulations are able to capture specific ones better than others. A better agreement with a certain pattern can hint to which processes are captured and which are missing. More importantly we can identify the important and necessary mechanisms to form a certain pattern.

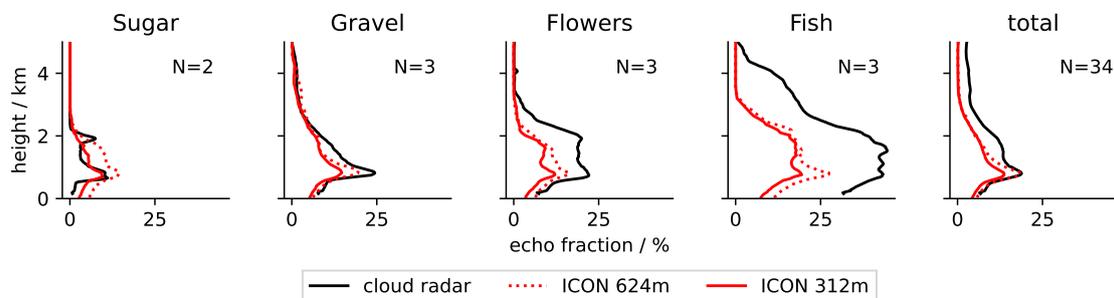


Figure E.10: Cloud fraction variability grouped by meso-scale pattern as identified in Schulz (in preparation). N defines the number of days found for each group.

To answer this question, we separated the cloud fraction profiles shown in Fig. E.9 by the prevalent meso-scale cloud pattern as described in Sec. E.3.3. The average profiles per pattern are shown in Fig. E.10.

The figure reveals that the differences between simulations and observations do depend on the observed meso-scale context. *Sugar* and *Gravel* are not only visually very similar to their observed counter-part but also exhibit a similar vertical distribution of cloudiness. The precipitation seems however overestimated in the case of *Sugar*. In contrast, precipitation of *Flowers* and *Fish* is underrepresented.

Both patterns are shallower and precipitate less in the simulations than in the observations. Especially the *Fish* pattern that passed over the BCO between January 21 and January 23 is hardly matched. While it cannot be expected that the patterns are collocated in observations and simulations, the supplemental movie shows that this *Fish* pattern was well developed and passed over the observatory also in the simulations ([10.5281/zenodo.5553825](https://doi.org/10.5281/zenodo.5553825)). However, it also reveals that the pattern completely disappeared in the simulations on January 23 and gave rise to much shallower and suppressed convection that resembles *Sugar*. In the observations, the *Fish* was still dominating. Despite this large discrepancies, the radar measured already on January 22 more cloudiness aloft 3 km than was represented in the model (not shown). The failure in representing these deeper clouds explains also in parts the smaller echo fractions close to the surface, because the deeper convection can be expected to rain more.

The simulated *Flowers* profit from the higher resolution run of ICON-312m in terms of a better separation of the cloudiness at the lifting condensation level and the one aloft. However, the cloud fraction is underestimated at each level.

E.5 DISCUSSION

While the previous sections have shown how well the simulations capture the meso-scale variability, this section examines hypotheses about the processes that might be involved in forming and maintaining the different meso-scale patterns.

The hypothesis that boundary layer circulations might be important for the organization of convection has been shown in several idealized model studies of deep convection (Bretherton et al., 2005; Muller and Held, 2012). Those circulations are able to transport moisture upgradient and enhance the differences between moist and dry regions. By doing so, convection is increasingly suppressed in the dry regions and aggregated in the moist ones. In nature though, such circulations cannot develop freely. Different to the idealized model studies no closed domain exists and disturbances interact with these circulations and may alter or hinder the self-organisation. Nevertheless, a shallow circulation was found as well (Schulz and Stevens, 2018).

Whether such a circulation exists to organize shallow convection is an open research question. Bretherton C. S. and Blossey P. N. (2018) found evidence of such a circulation in LES forced by a climatological mean thermodynamic profile representative for situations of meso-scale clusters of cumulus rising into thin stratocumulus. In addition, George et al. (in preparation) found an anti-correlation between the divergence in the subcloud layer and the cloud layer in dropsonde data of the EUREC⁴A field campaign. They explain this relationship with a potential meso-scale circulation.

This sparked the hypothesis that such circulations might also exist for the meso-scale patterns of Stevens et al. (2020) which are generally larger and more diverse as the ones studied by Bretherton C. S. and Blossey P. N. (2018) on a $(128 \text{ km})^2$ domain.

To explore whether such a circulation is also present in the simulations of this study, we calculated the anomalous wind speed along the cross-sections shown in Fig. E.5 by subtracting the mean wind speed profile from the actual one.

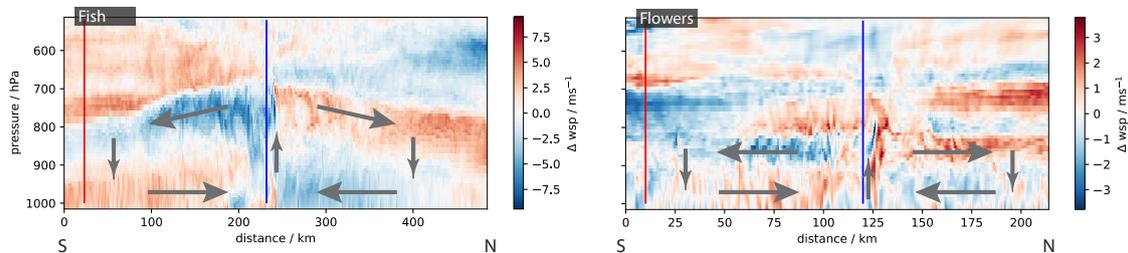


Figure E.11: Boundary-layer circulations inferred from anomalous wind-speeds along the cross-sections indicated in Fig. E.5. The cross-sections are illustrated from south to north. Arrows indicate circulations. Note the different scales of the colorbar.

Fig. E.11 reveals that in case of *Fish* and *Flowers* shallow circulations are indeed present. The anomalous wind is converging in the lower 100 hPa and is diverging below the inversion, resulting in the indicated circulations. While these are two snap-shots and further research is needed, the differences between the circulations of *Flowers* and *Fish* are evident.

The circulation of the *Fish* case is much deeper and stronger and follows the slope of the inversion. In contrast, the exemplary circulation of *Flowers* is much more layered and an equally strong counter-circulation above the inversion exists. This counter-circulation might be caused by the evaporative cooling from occasionally deeper convective cores within the cloudy patches that mix with dry free-tropospheric air and cause the air to sink within the inversion as in Bretherton C. S. and Blossey P. N. (2018).

Fig. E.12 highlights the differences in moisture between the dry and moist regions. Interestingly, the simulated *Fish* are shallower than their observed counter-part (Fig. E.10) and accompanied by a strong inversion that according to Schulz et al. (2021) is usually less pronounced.

From this point of view *Flowers* and *Fish* might only be different due to the additional convergence that the *Fish* pattern receives in form of the shear line and can consequently develop a deeper convective system.

For the other patterns no circulation of this scale could be found (see supplemental Fig. E.15). This is however not surprising, because the scale of individual cloud patches is much smaller. In particular *Sugar*, in its most idealized sense, is not organized but rather randomly distributed. The external organization of these cloud patterns into streets or clusters shall not be discussed here, but the reader is referred to Stephan (2021) and the references therein for possible explanations.

The visual impression of *Gravel* however, raises a strong link to the open mesoscale cellular convection observed in the upstream trades and its accompanying precipitation processes (Stevens et al., 2005). As Fig. E.13 suggests, precipitation and the development

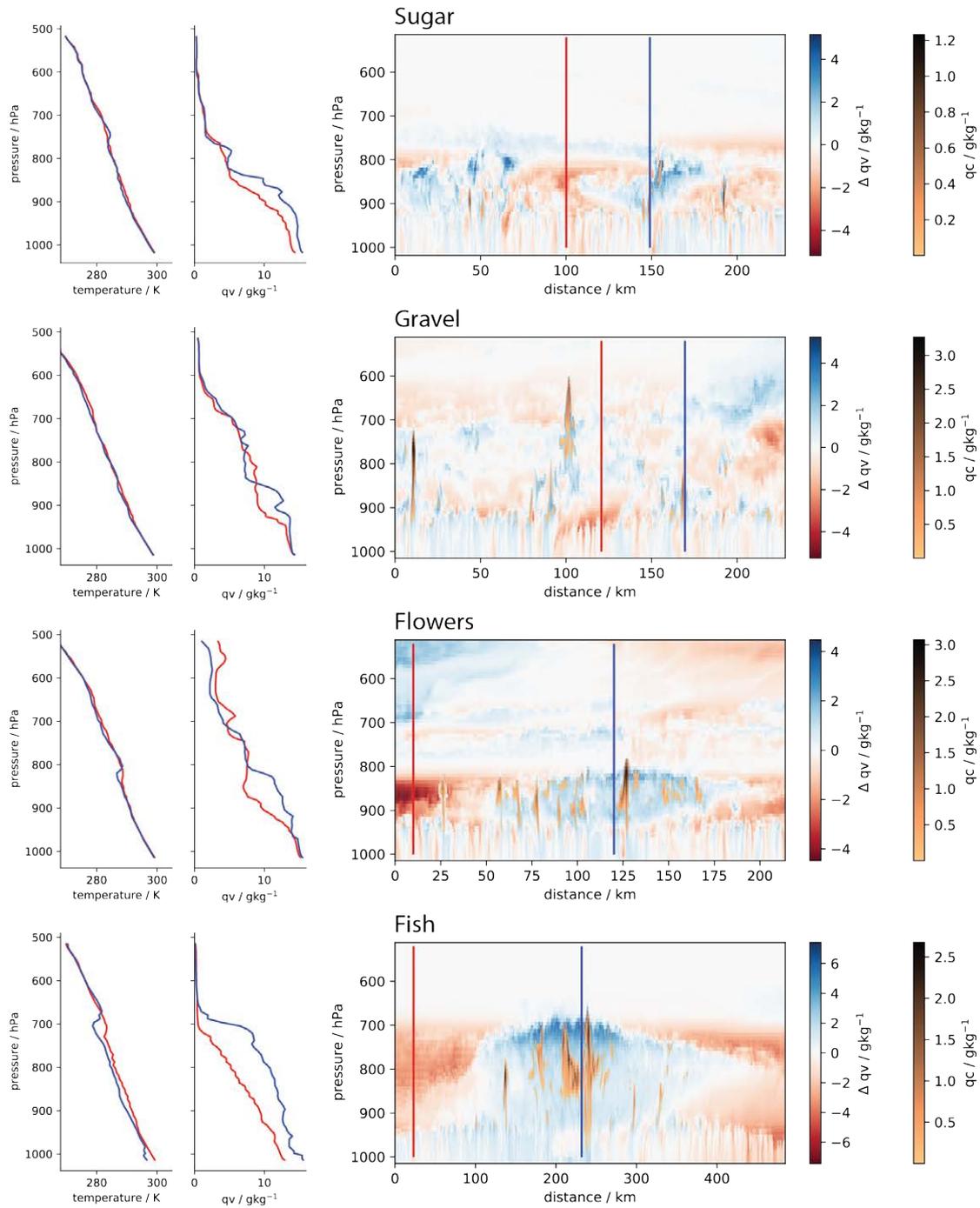


Figure E.12: Example cross-sections through the meso-scale patterns as indicated in Fig. 5.1 (right) with profiles from a wet (blue) and dry (red) location. The horizontal moisture anomaly (Δq_v) is shown with respect to the cross-sectional mean profile. Clouds are identified by their cloud water content q_c .

of cold pools that eventually collides with other cold pools are frequently observed in the case of *Gravel*.

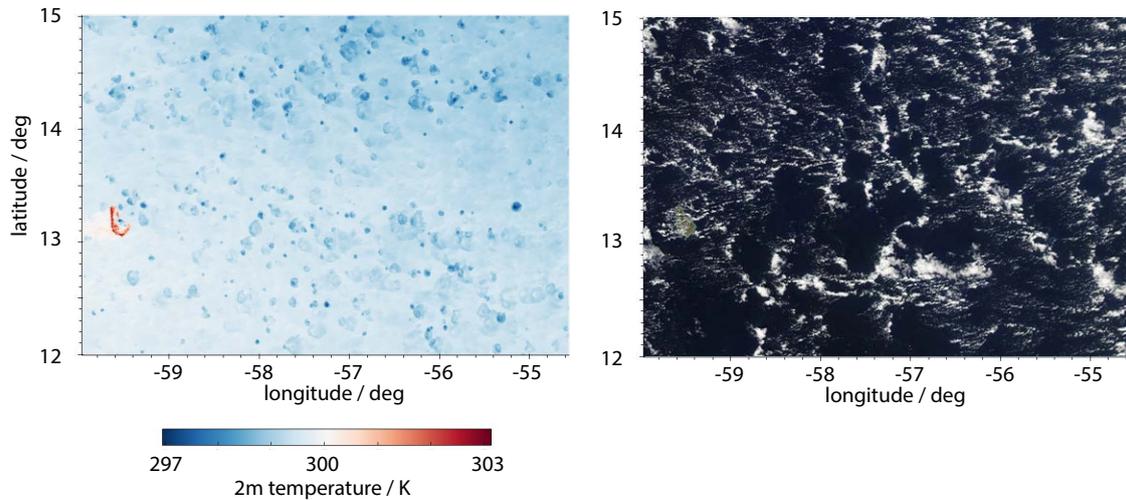


Figure E.13: Cool pool structures as seen in the ICON-312m surface temperature (left) and a visible image from MODIS onboard AQUA (right) at about 17:40 UTC on January 11, 2020.

In sum, the simulations support the hypotheses posed about the physical mechanisms maintaining and leading to the individual patterns and are proven to be a valuable tool to study these in more detail.

E.6 CONCLUSION

Realistically forced large-eddy simulations have been performed in the North Atlantic downwind trades to evaluate their ability to reproduce the observed meso-scale variability of shallow convection and the imprints of its driving processes.

The presented ICON LESs covers 37 days from January - February 2020, which concurs with the field campaign EUREC⁴A. The measurement strategy of EUREC⁴A was to statistically sample the atmosphere rather than aiming at specific atmospheric conditions. This strategy has also been applied to the simulations. Instead of simulating a few cases, these simulations were continuously run using large domain sizes ((700 km)² to (1300 km)²) that captured the intrinsic meso-scale variability of the region, including the meso-scale patterns of several 100 km that have been identified recently by Stevens et al. (2020).

The ICON-LEM's performance in representing the meso-scale variability in cloudiness was examined using observations collected at the Barbados Cloud Observatory and the GOES-16 ABI instrument. Additionally, the forward operators RTTOV and PAMTRA allowed a fair comparison between the observed and simulated shallow cloud patterns. Lastly, the neural network that has already been shown to identify the meso-scale patterns in satellite images (Schulz et al., 2021) has been successfully applied to the simulation output to evaluate the meso-scale structure.

We find that the average meso-scale variability is represented well in terms of the total cloud amount and the spatial distribution of cloud patches. However, this agreement differs among the patterns defined by Stevens et al. (2020). While all patterns have been identified in the observations, the simulations show disproportionately more frequent and widespread *Sugar* and *Gravel* at the cost of *Fish* and *Flowers*.

This discrepancy can be primarily explained by the lack of stratiform cloudiness in the simulations. The stratiform cloudiness is characteristic for *Flowers* and *Fish*. The vertical cloud distribution and in particular its variability in the layer 1 km to 2.5 km where the stratiform clouds occur is however not matched. The variability is reduced by about a factor of 2 independent of the horizontal resolution. In case of *Fish* the large-scale forcing was not sustained long enough which lead to a premature dissipation. If this is a systematic issue of the model needs to be studied further.

With an increase in grid-spacing from 624 m to 312 m, the cloud fraction profile improved, particularly at the lifting condensation level. The representation of *Flowers* and *Fish* did therefore not improve in terms of their appearance in satellite images. They only benefit from the general trend of a more realistic average cloud profile with a more realistic balance of clouds at the lifting condensation level and aloft. This is in agreement with Vial et al. (2019) who found a reduction of cloudiness with cloud tops above 1.3 km with increasing resolution.

Although the simulations do not match the stratiform component, meso-scale shallow circulations are evident for the larger cloud patches and might play a role in the emergence and maintenance of these patterns. The drivers of these circulations will be addressed in future investigations. However, because the pattern structure is fairly well represented in the simulations but the stratiform cloud component is not, the simulations highlight that the cloud top radiative cooling is unlikely to be relevant for the general meso-scale structure. Future investigations should focus on other potential drivers like radiative cooling in clear-sky regions or precipitation processes. Thanks to the simulations and the tremendous amount of EUREC⁴A measurements, we provided a great framework to gain further process understanding of how these patterns with their different CRE form.

E.7 OPEN RESEARCH

GOES-16 Advanced Baseline Imager Level 1b radiances are available at <https://doi.org/10.7289/V5BV7DSR> and were converted with (Raspaud et al., 2019) to brightness temperatures. MODIS imagery originates from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS).

E.8 ACKNOWLEDGEMENT

The authors are grateful to the members of the Tropical Cloud Observation group for maintaining the Barbados Cloud Observatory and continuously increasing its valuable dataset. The authors thank Daniel Klocke for conducting the ICON-SRM simulations and his valuable comments regarding ICON in general. Robert Vicari is thanked for implementing an updated version of the satellite forward operator interface into ICON.

This work used resources of the Deutsches Klimarechenzentrum (DKRZ) and the authors are thankful for the granted compute time and storage.

E.9 SUPPLEMENTAL MATERIAL

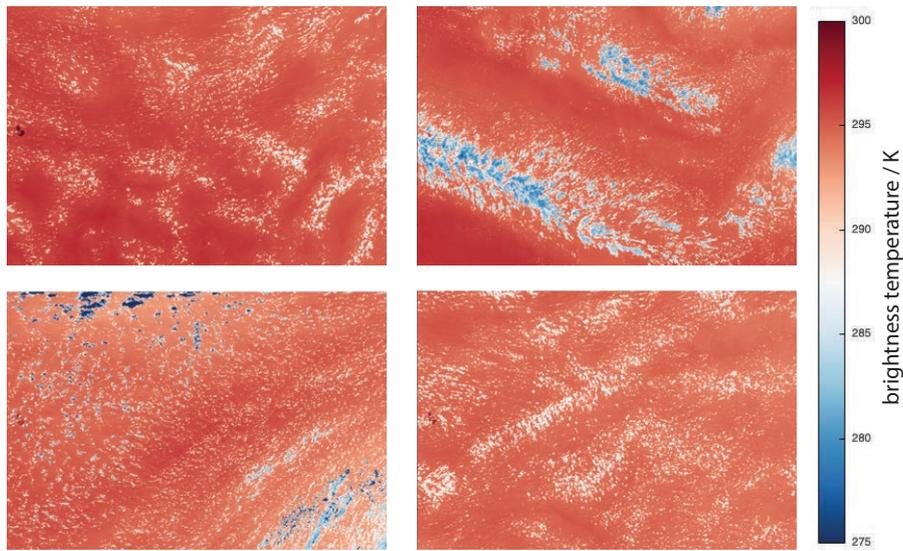
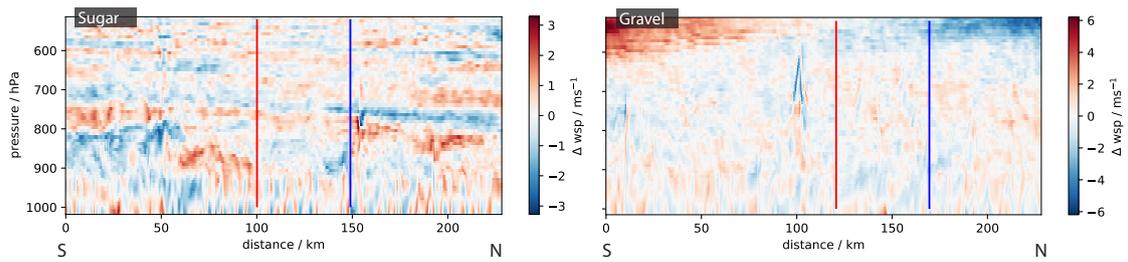


Figure E.14: Like Fig. E.5 but for the ICON-312m simulation.

Figure E.15: As Fig. E.11, but for *Sugar* and *Gravel*.

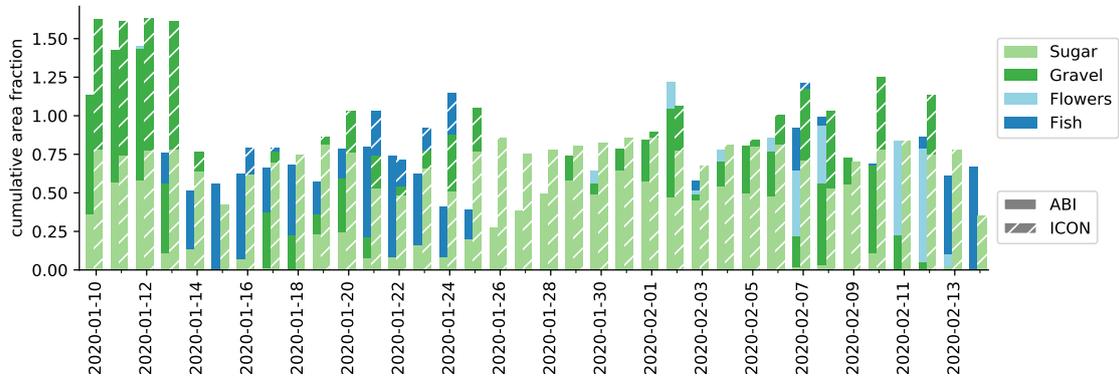


Figure E.16: As Fig. E.6, but with ICON 312m.

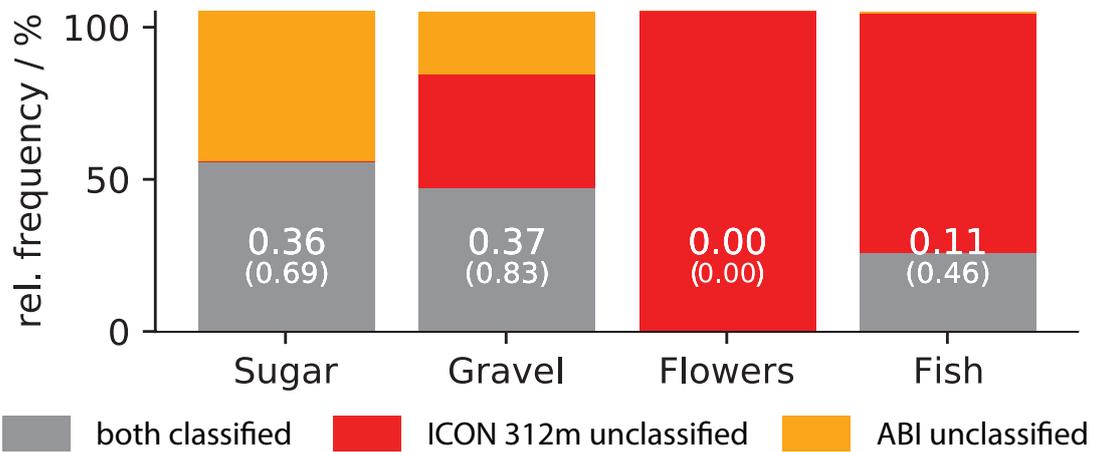


Figure E.17: As Fig. E.7, but with ICON 312m.

SHIP- AND ISLAND-BASED ATMOSPHERIC SOUNDINGS FROM THE 2020 EUREC⁴A FIELD CAMPAIGN

The work in this appendix has been published as:

Stephan, Claudia Christine, Schnitt, Sabrina, **Schulz, Hauke**, Bellenger, Hugo, de Szoeke, Simon P., Acquistapace, Claudia, Baier, Katharina, Dauhut, Thibaut, Laxenaire, Rémi, Morfa-Avalos, Yanmichel, Person, Renaud, Quiñones Meléndez, Estefanía, Bagheri, Gholamhossein, Böck, Tobias, Daley, Alton, Güttler, Johannes, Helfer, Kevin C., Los, Sebastian A., Neuberger, Almuth, Röttenbacher, Johannes, Raeke, Andreas, Ringel, Maximilian, Ritschel, Markus, Sadoulet, Pauline, Schirmacher, Imke, Stolla, M. Katharina, Wright, Ethan, Charpentier, Benjamin, Doerenbecher, Alexis, Wilson, Richard, Jansen, Friedhelm, Kinne, Stefan, Reverdin, Gilles, Speich, Sabrina, Bony, Sandrine, and Stevens, Bjorn (Feb. 2021). Ship- and Island-Based Atmospheric Soundings from the 2020 EUREC⁴A Field Campaign. *Earth System Science Data* 13.2, pp. 491–514. DOI: [10.5194/essd-13-491-2021](https://doi.org/10.5194/essd-13-491-2021)

The contributions of the authors to this publication are as follows:

Sandrine Bony and Bjorn Stevens designed the sounding strategy, which was then refined and realized in cooperation with Claudia C. Stephan and Simon P. de Szoeke. Sabrina Speich (cruise lead *Atalante*) designed, with Gilles Reverdin, and managed the measurements on board the *Atalante*. Stefan Kinne (cruise lead *Meteor*) and Friedhelm Jansen (responsible for BCO operations) managed the logistics of purchasing and transporting the radiosonde equipment. Benjamin Charpentier and Richard Wilson processed the MeteoModem data. Alexis Doerenbecher investigated the data flow through the GTS. The majority of radiosonde launches were performed by Gholamhossein Bagheri, Katharina Baier, Tobias Böck, Alton Daley, Johannes Güttler, Kevin C. Helfer, Sebastian A. Los, Yanmichel Morfa-Avalos, Almuth Neuberger, Andreas Raeke, Maximilian Ringel, Markus Ritschel, Johannes Röttenbacher, Pauline Sadoulet, Imke Schirmacher, M. Katharina Stolla and Ethan Wright. Claudia Acquistapace, Thibaut Dauhut, Rémi Laxenaire, Renaud Person, Estefanía Quiñones Meléndez and Sabrina Schnitt played a leading role in the experimental execution. Hugo Bellenger and Simon P. de Szoeke contributed to the design of the radiosonde network. Hauke Schulz processed, quality-checked and analyzed the data. Claudia C. Stephan prepared the manuscript with contributions from all co-authors.

Ship- and island-based atmospheric soundings from the 2020 EUREC⁴A field campaign

Claudia C. Stephan¹, Sabrina Schnitt², Hauke Schulz¹, Hugo Bellenger³, Simon P. de Szoeké⁴, Claudia Acquistapace², Katharina Baier¹, Thibaut Dauhut¹, Rémi Laxenaire⁵, Yanmichel Morfa-Avalos¹, Renaud Person^{6,7}, Estefanía Quiñones Meléndez⁴, Gholamhossein Bagheri⁸, Tobias Böck², Alton Daley⁹, Johannes Güttler¹⁰, Kevin C. Helfer¹¹, Alex Kellman³, Sebastian A. Los¹², Johannes Röttenbacher¹³, Andreas Raeke¹⁴, Maximilian Ringel¹, Markus Ritschel¹, Pauline Sadoulet¹⁵, Imke Schirmacher¹⁶, M. Katharina Stolla¹, Ethan Wright⁵, Benjamin Charpentier¹⁷, Alexis Doerenbecher¹⁸, Richard Wilson¹⁹, Friedhelm Jansen¹, Stefan Kinne¹, Gilles Reverdin⁷, Sabrina Speich³, Sandrine Bony³, Bjorn Stevens¹

¹Max Planck Institute for Meteorology, Hamburg, Germany

²Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany

³LMD/IPSL, CNRS, ENS, École Polytechnique, Institut Polytechnique de Paris, PSL, Research University, Sorbonne Université, Paris, France

⁴College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA

⁵Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida, USA

⁶Sorbonne Université, CNRS, IRD, MNHN, INRAE, ENS, UMS 3455, OSU Ecce Terra, Paris, France

⁷Sorbonne Université, LOCEAN, SU/CNRS/IRD/MNHN, Paris, France

⁸Laboratory for Fluid Physics, Pattern Formation and Biocomplexity, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany

⁹Caribbean Institute for Meteorology and Hydrology, Husbands, St. James, Barbados

¹⁰Max Planck Institute for Dynamics and Self-Organization, 37077 Göttingen, Germany

¹¹Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

¹²Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico, USA

¹³Institute for Meteorology, University of Leipzig, Leipzig, Germany

¹⁴Deutscher Wetterdienst, Seewetteramt Hamburg, Seeschiffsberatung - Bordwetterdienst, Hamburg, Germany

¹⁵Météo-France, Bordeaux, France

¹⁶University of Hamburg, Hamburg, Germany

¹⁷Meteomodern, Ury, France

¹⁸Météo-France and CNRS: CNRM-UMR 3589, 42 Av. G. Coriolis, 31057 Toulouse Cedex, France

¹⁹Sorbonne Université, LATMOS/IPSL, INSU/CNRS, Paris, France

ABSTRACT

To advance the understanding of the interplay among clouds, convection, and circulation, and its role in climate change, the EUREC⁴A and ATOMIC field campaigns collected measurements in the western tropical Atlantic during January and February 2020. Upper-air radiosondes were launched regularly (usually 4-hourly) from a network consisting of the Barbados Cloud Observatory (BCO) and four ships within 51–60 °W, 6–16 °N. From January 8 to February 19, a total of 811 radiosondes measured wind, temperature and relative humidity. In addition to the ascent, the descent was recorded for 82 % of the soundings. The soundings sampled changes in atmospheric pressure,

winds, lifting condensation level, boundary layer depth, and vertical distribution of moisture associated with different ocean surface conditions, synoptic variability, and mesoscale convective organization. Raw (Level-0), quality-controlled 1-second (Level-1), and vertically gridded (Level-2) data in NetCDF format (Stephan et al., 2020a) are available to the public at AERIS (<https://doi.org/10.25326/137>). The methods of data collection and post-processing for the radiosonde data set are described here.

F.1 INTRODUCTION

A number of scientific experiments have focused on the trade-cumulus boundary layer over the tropical Atlantic Ocean. The Barbados Oceanographic Meteorological Experiment (BOMEX 1969; Kuettner and Holland, 1969), Atlantic Trade-Wind Experiment (ATEX 1969; Augstein et al., 1973), Atlantic Stratocumulus Transition Experiment (ASTEX 1992; Albrecht et al., 1995), and Rain in Shallow Cumulus Over the Ocean (RICO 2006; Rauber et al., 2007) experiment measured thermodynamic and wind profiles of the Atlantic trade regime (reviewed by Baker, 1993). With these profiles as initial and environmental conditions, models of the cumulus clouds explain their interaction with the environment (e.g. Albrecht, 1993; Albrecht et al., 1979; Arakawa and Schubert, 1974; Bretherton, 1993; Krueger, 1988; Tiedtke, 1989; Xue et al., 2008; vanZanten et al., 2011).

Arrayed networks of soundings have been used to characterize the interaction of clouds, convection, and the synoptic environment. In many examples, they have been used to diagnose tendencies of the heat, mass, and moisture budgets for the tropical atmosphere (e.g. Johnson and Ciesielski, 2013; Lin and Johnson, 1996; Mapes et al., 2003; Nitta and Esbensen, 1974; Reed and Recker, 1971; Yanai et al., 1973). These experiments in the deep tropics monitored the synoptic (100–1000 km) variations of vertical motion and moisture convergence as context for the evolution of the ensemble of convective clouds observed within their sounding networks.

These sounding arrays measure horizontal divergence, which is used to estimate mean large-scale vertical motion. In DYCOMS-II, Lenschow et al., 2007 used stacked flight circles to estimate subsidence on a fine scale relevant to marine stratocumulus clouds. Studying the variations of mesoscale (~ 100 km) organization of the trade wind shallow cumulus clouds likewise requires fine horizontal resolution. The Next-Generation Aircraft Remote Sensing for Validation Studies (NARVAL; Bony and Stevens, 2019; Stevens et al., 2016, 2019) demonstrated that circles of dropsondes released from aircraft above the shallow clouds reliably measure a snapshot of vertical motion.

The shallow trade cumulus clouds over the tropical Atlantic Ocean are a focus also of the Elucidating the Role of Clouds-Circulation Coupling in Climate Campaign (EUREC⁴A; Bony et al., 2017) and associated campaigns, i.e. the Atlantic Tradewind Ocean–Atmosphere Mesoscale Interaction Campaign (ATOMIC)¹. The experimental design of EUREC⁴A involved 85 dropsonde circles from aircraft flights combined with regular around-the-clock upper air observations from surface-launched radiosondes. The regular sampling from surface-launched radiosondes complemented the mesoscale vertical velocity measurements from dropsonde circles by continuously measuring time-height profiles of the atmosphere, synoptic variability for an extended time period, and

¹ Because the sounding network and EUREC⁴A comprised many projects, or component campaigns, we refer to the union of these simply as EUREC⁴A.

diurnal variability. Radiosondes sampled when research aircraft were not flying, notably at night.

This article introduces the radiosonde observations and their resulting data sets. Other measurements, including the dropsonde data, are described in the overview paper by Stevens et al., 2021 and the references therein. Between January 8 and February 19, 2020, 811 radiosondes were launched from Barbados and the northwestern tropical Atlantic Ocean east of Barbados. A focus of the campaign was on shallow cumulus clouds, their radiative effects, and their response to the large-scale environment, contributing progress toward the World Climate Research Programme's Grand Challenge on Clouds, Circulation and Climate Sensitivity (Bony et al., 2015). Other EUREC⁴A investigations focus on air-sea interactions due to ocean mesoscale eddies, cloud microphysical processes, and the effect of shallow convection on the distribution of winds.

Radiosondes were launched from Barbados and four research vessels. The island-based launches took place at the Barbados Cloud Observatory (BCO; 59.43 °W, 13.16 °N), situated at Deebles Point on the windward coast of Barbados. Surface and remote sensing observations at BCO have been in operation since April 1, 2010 (Stevens et al., 2016).

Four research vessels launched radiosondes over the northwestern tropical Atlantic east of Barbados (51–60 °W, 6–16 °N) during EUREC⁴A: two German research vessels, *Maria S. Merian* (hereafter *Merian*) and *Meteor*, a French research vessel, *L'Atalante* (hereafter *Atalante*), and a United States research vessel, *Ronald H. Brown* (hereafter *Brown*). The BCO and the research vessels all measured surface meteorology and deployed various other measurements for remote sensing of clouds and the atmospheric boundary layer.

In Section F.2 we describe the measurement strategy for the coordinated EUREC⁴A radiosonde network, the data collection procedures for each platform, and the post-processing steps that were applied to create the final data set. Section F.3 shows an overview and some characteristics of the data and is followed by a summary in Section F.4. The *Atalante* additionally launched a different type of sonde, which is described in the appendix.

F.2 SOUNDING MEASUREMENTS

F.2.1 The EUREC⁴A sounding network

The number of launches per day as well as the dates of regular observations (Fig. F.5) differ from platform to platform, reflecting availability of ships and personnel. Soundings supported specific research interests on each platform, in addition to the coordinated EUREC⁴A sounding network. We designed the radiosonde network to optimize the joint contribution of all platforms to the overarching goals of EUREC⁴A. Sounding platforms were usually spaced to optimally sample the scales of the synoptic circulation. The *Meteor* remained nearly stationary at a longitude of 57 °W and moved within a meridional corridor between 12.0–14.5 °N to support coordinated aircraft measurements in its vicinity (Fig. F.1a). The *Brown* occupied a southwest-northeast transect along the direction of the climatological surface trade winds, and approximately orthogonal to *Meteor*'s sampling line. The *Brown*'s transect between the BCO (59.43 °W, 13.16 °N) and the Northwest Tropical Atlantic Station for air-sea flux measurements buoy (NTAS) at

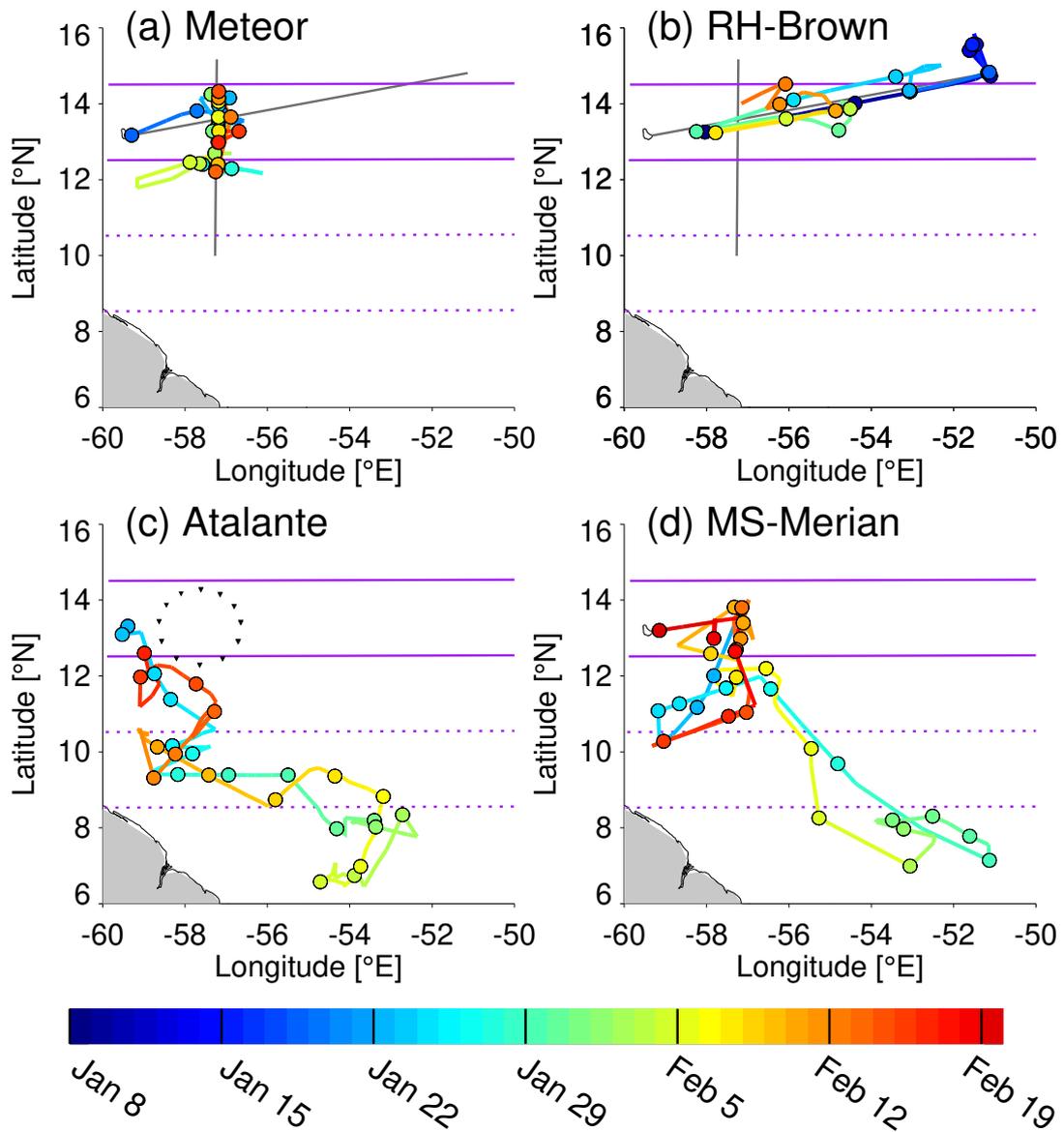


Figure F.1: Routes and launch coordinates of radiosondes for the four research vessels colored by date. Circles mark the locations of the first radiosonde launch on each day. The gray lines in (a) and (b) mark the nearly orthogonal lines that were sampled by the *Meteor* (North–South) and the *Brown* (West–East). Purple lines mark the northern (12.5–14.5°N; solid) and southern (8.5–10.5°N; dashed) latitude bands that we later use to define a North (Trade-wind Alley) and South (Boulevard des Tourbillons) domain. Downward pointing black triangles in (c) mark the locations of dropsonde releases during regular circular aircraft flights.

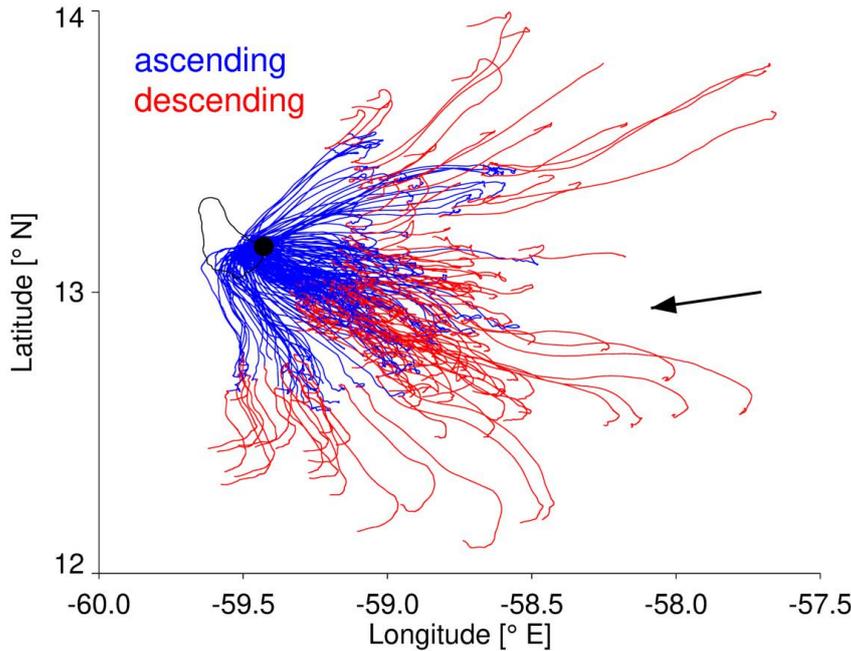


Figure F.2: The horizontal trajectories of ascending and descending, respectively, radiosondes launched from the BCO. The location of the BCO on Barbados is marked by the thick black dot. The black arrow is the mean wind direction at 500 m as measurement by ascending soundings launched from the BCO.

51.02 °W, 14.82 °N (Fig. F.1b) sampled air masses upwind of the BCO that move westward with the climatological easterly trade winds within 12.5–14.5 °N. This elongated region between BCO and NTAS is referred to as the ‘Trade-wind Alley’. The *Merian* and *Atalante* ventured southward to a minimum latitude of ~ 6.5 °N to observe oceanic and atmospheric variability associated with Brazil Ring Current Eddies as they tracked northwestward along the corridor referred to as ‘Boulevard des Tourbillons’. The *Atalante* and *Merian* thus often form the southern points of the radiosonde network (Fig. F.1c, d).

Aircraft operations included a circular flight pattern of 180–200 km diameter centered at ~ 13.3 °N, -57.7 °E (Fig. F.1c). Dropsondes were deployed along the circle to estimate the area-averaged mass divergence, as described in Bony and Stevens, 2019. To sample larger scales than represented by this circle, we aimed at 4-hourly soundings from all five stations while platforms were separated by more than 200 km. The launch frequency was reduced when such a separation could not be maintained or when vessels left the key region of the network, i.e. moved south of 12 °N. These scenarios occurred from time to time in order to support other measurements. Figure F.2 shows that the network sampled large scales for 30 consecutive days.

To increase the number of vertical profiles, we recorded the ascent as well as the descent of the radiosondes. For descending soundings the raw data near the surface are missing as the signal is lost due to Earth’s curvature at 300–800 m above mean sea level. The median of the lowest descent measurement is at 340 m. Except for the *Brown*, balloons were equipped with parachutes, which nearly match fall speeds in the middle and lower troposphere to balloon ascent speeds. Given that a typical ascent takes about 90 min, a radiosonde was sampling the air somewhere close to each platform

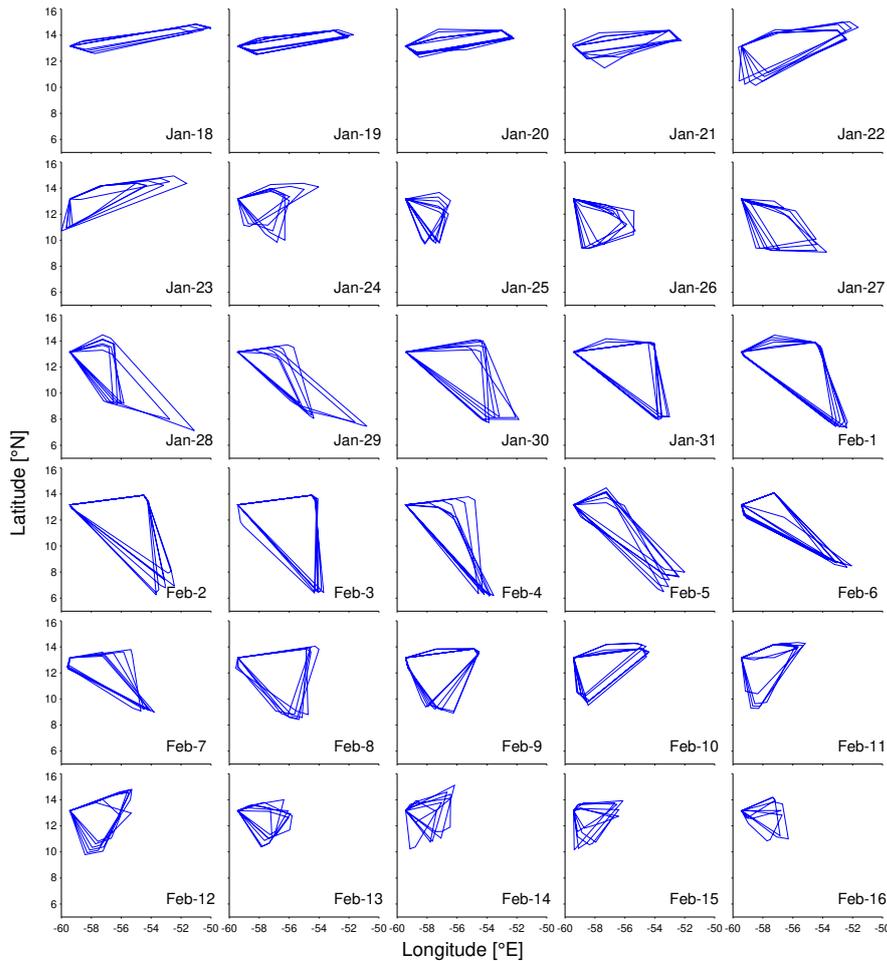


Figure F.3: For each day between Jan-18 and Feb-16, 4-hourly polygons mark the outer bounds of the radiosonde network. Polygon vertices correspond to starting locations of either ascending or descending soundings that occurred within ± 2 hours of a fixed time.

nearly continuously during regular operation. The horizontal drift of the sondes is shown in Fig. F.3 for the example of the BCO. All platforms deployed Vaisala RS41-SGP radiosondes, which measure wind, temperature, relative humidity, and pressure, and used Vaisala MW41 ground station software to record and process the sounding data. The software versions of the MW41 system are given in Table 1 for each platform. Basic algorithms and data processing did not change between these versions. Vaisala sondes were attached to 200 g balloons (*BCO*, *Atalante*, *Merian*, *Meteor*) or 150 g balloons (*Brown*). When present, the balloons were equipped with internal parachutes (see Table 1 for the use of parachutes). A modification took place on the *Atalante*, where after 0800 UTC on February 8, 350 g balloons with external parachutes were used instead.

To start a sounding, a radiosonde sensor was placed on the ground station for an automated ground check initialization procedure, which took about 5–6 min. The frequency at which the radiosonde transmits its signal to the receiver was set manually to a designated value for each platform (listed in Table F.1) to avoid radio interference.

The default launch times were 0245, 0645, 1045, 1445, 1845, and 2245 UTC. This schedule was selected to include two launches per day that were timed to match the

Table F.1: For each platform the rows list (1) the numbers of recorded ascending soundings, (2) the numbers of recorded descending soundings, (3) the first date of data coverage, (4) the last date of data coverage, (5) whether or not parachutes were used, (6) the station altitude relative to sea level (for ships: apparent sea level, for BCO: mean sea level), (7) the GPS antenna offset relative to the station, (8) the launch site offset relative to the station, (9) the surface barometer offset relative to the station, (10) the frequency used to transmit the signal from the radiosonde to the antenna, (11) the MW41 software version (12) the WMO station ID, (13,14) the abbreviated heading used for exchange on the Global Telecommunication System, consisting of (13) data designators (TTAAii) and (14) international four-letter location indicators (CCCC) (WMO, 2009).

	BCO	Meteor	DWD	Brown	Atlante	Meriam
Number ascents	182	180 [†]	39 [†]	169	139	118
Number descents	162	175 [†]	-	158	138	38
Start date	Jan 16	Jan 16	Jan 18	Jan 8	Jan 21	Jan 18
End date	Feb 17	Mar 1	Feb 26	Feb 12	Feb 16	Feb 19
Use of parachutes	yes	yes	no	no	yes	yes
Station altitude (m)	25.0	16.9	5.4	4.3	13.1	10.4
GPS antenna offset (m)	4.3	2.5	3	5.5	2.6	1.6
Launch site offset (m)	0.0	-11.5, -14.2 ²	0.0, -2.7 ²	0.5	-0.6	0.0
Surface barometer offset (m)	1.0	-16.9	-5.4	-4.3	0.2	0.6
Frequency (MHz)	400.2	401.5	403.0	400.5	400.7 - 401.2	402.0
MW41 software version	2.15.2.0-37	2.15.2.0-37	2.6.1.60	2.4.0.75	2.15.2.2-42	2.15.2.0-37
Station IDs	BCO	ZVQEQCM	WTEC	WTEC	Atalante	FPUW=GN
TTAAii	IUSN74	IUSX40, IUKX40	IUSE01	IUSE01	IUSN72	IUSX40, IUKX40
CCCC	LFPW	EDZW	WTEC	WTEC	LFPW	EDZW

^a includes 8 additional soundings after Feb 20, 0 UTC

^b Feb 9, 18 UTC - Feb 20, 0 UTC

00 and 12 UTC synoptic times. In practice the soundings reached 100 hPa on average in 60 minutes and burst after 90 minutes. Departures from this schedule occurred due to a variety of reasons, including defective radiosondes, balloon bursts before the launch, collisions of ascending radiosondes with other on board instrumentation, and air traffic safety. In the following section, we describe specific issues and aspects of the launch procedure and equipment particular to each platform. All stations followed best practices for different equipment, which were established by several experienced teams at in-person sounding orientations prior to the campaign. For instance, every platform used a different empirical way of gauging the fill amount of gas, to arrive at desired ascent rates. Equipment and procedures differed between the platforms, but this does not introduce systematic biases to Level-2 data, as these data only start at 40 m height (see Section 2.3.2), where measurements are independent of the surface procedures.

F.2.1.1 Barbados Cloud Observatory (BCO)

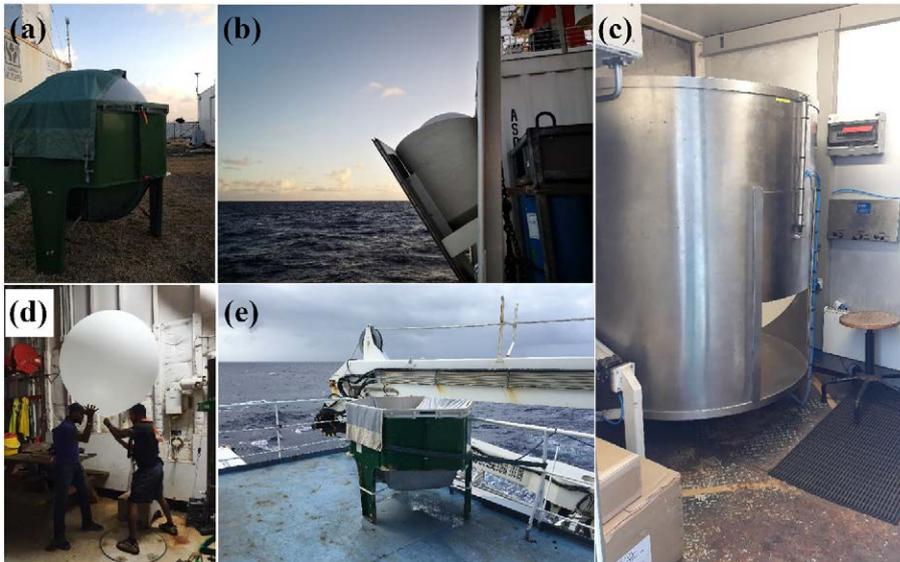


Figure F.4: Photographs of the (a) launcher with balloon at the BCO, (b) DWD launcher with balloon on board the *Meteor*, (c) launch container with balloon on board the *Merian*, (d) manual balloon filling procedure on board the *Brown*, (e) empty launcher on board the *Atalante*.

The BCO is located at the eastern-most point of Barbados (59.43°W , 13.16°N) and thus directly exposed to easterly trade winds from the ocean (Fig. 3). The BCO launched 182 sondes, of which 162 measured descents. Radiosondes were prepared inside an air-conditioned office container with air temperature and relative humidity adjusted to 20°C and 60 %, respectively. Balloons were prepared outside and placed into a launcher whose size provided rough guidance for achieving the desired filling level (Fig. F.4a). Launches were coordinated with Barbados Air Traffic Control, which sometimes delayed soundings by up to 15 min. Surface conditions obtained from the weather station observations at the BCO were entered into the software after automatic release detection.

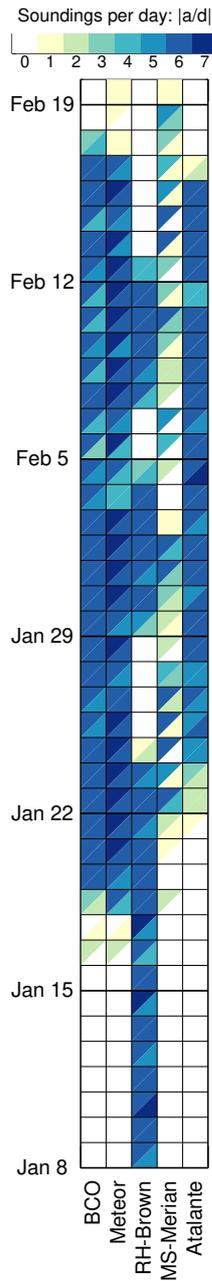


Figure F.5: Daily number of ascending (upper left triangles) and descending (lower right triangles), respectively, soundings associated with each platform.

R/V Meteor

The *Meteor* launched 203 sondes and collected data for 167 descents during the EUREC⁴A core period (January 8 to February 19). Eight additional ascents and descents, respectively, were recorded after February 20. Radiosondes were prepared inside a laboratory on the top deck of the ship with the antenna placed on the roof. Before February 9 the soundings were launched from the container of the German Weather Service (DWD) on the port side at the stern of the ship (Fig. F.4b). This container had a marker to indicate the optimum fill level of the balloons.

On February 9 the DWD launcher broke and a launcher of the type shown in Fig. F.4a was used, located at the stern of the ship. An awning over the balloon indicated the fill level. Ground data were obtained from on-board instruments of the DWD. In addition to sondes launched by the EUREC⁴A science crew, the DWD launched one radiosonde per day. The 31 ascending DWD sondes launched during the EUREC⁴A core period, plus an additional eight after February 20, are included in the Level-1 and Level-2 data sets, described in Section F.2.3.

By mistake, the heights of the pressure sensor, the GPS antenna and the launching altitude were incorrectly entered at the beginning of the cruise. In addition, we noticed large delays between the time at which surface measurements were entered and the launch. Therefore, we reprocessed the raw data using the MW₄₁ software, after correcting the sensor heights and surface data in the raw files. This post-processing is lossless and the reprocessed data have the same quality standard as the data from the other platforms. We included both the original and reprocessed Level-0 data in the dataset.

R/V Ronald H. Brown (Brown)

The *Brown* released 169 sondes and collected data for 158 descents. The radiosondes were initialized and ground-checked inside an air-conditioned laboratory. Near-surface measurements were recorded from the ship’s meteorological sensors via the ship computer system display. The ground station antenna was located on the aft 02 deck railing above the staging bay. Initialized radiosonde sensor packages were placed for 1–5 min on the main deck to equilibrate to ambient environmental conditions and check GPS reception and telemetry. The balloons were filled by hand in the staging bay (Fig. F.4d), which was mostly sheltered. Operators avoided unnecessary contact with the balloon body but restrained it by hand if the wind was strong.

On leg 1 (January 8–24) at night, less helium was used to reduce the buoyancy of the balloons in order to achieve lower ascent rates and better resolve the fine-scale vertical structure of the atmosphere. The ascent rate for day launches was $4.4 \pm 0.5 \text{ m s}^{-1}$. Ascent was about 12 % slower for night launches ($3.9 \pm 0.6 \text{ m s}^{-1}$). To avoid the potential for biasing analyses of the diurnal cycle with systematic diurnal differences in ascent rates, after January 24, the same target ascent rate was used for day and night. Operators obtained consistent balloon volumes by timing the filling.

Balloons were launched from a location on the deck to minimize the effect of the ship and obstructions on the sounding. The ship usually turned or slowed to improve the relative wind for the sounding. The relative wind carried the sounding away from the ship, but the ship's aerodynamic wake made the first $\sim 5 \text{ s}$ of the balloon's flight unpredictable. The sounding was sometimes launched up to 10 min earlier or later to accommodate other ship operations.

F.2.1.2 R/V L'Atalante (*Atalante*)

The *Atalante* launched 139 Vaisala sondes and measured 138 descents. A coordinated sounding phase was performed with the *Merian* to increase the temporal resolution from January 30 at 2045 UTC to February 2 at 1645 UTC around $52\text{--}54^\circ \text{W}$ and $6\text{--}8^\circ \text{N}$. During this period launching times were shifted by 2 hours aboard the *Atalante* (0045, 0445, 0845, 1245, 1645, 2045 UTC) while the *Merian* launched at regular times. In addition to the Vaisala soundings, 47 sondes of MeteoModem type M10 attached to 150 g balloons without parachutes were launched from the *Atalante* to measure the lower atmosphere across mesoscale sea surface temperature (SST) fronts, as detailed in the appendix.

The radiosondes were prepared aft of the bridge. This open space was right next to the top building of the ship, which may have affected measurements at low levels. Before launching, operators asked the bridge for direction change if necessary and possible. The balloons were launched by hand from the rear deck of the bridge, where the launcher was situated (Fig. F.4e). The Vaisala antenna was installed on the roof top. Surface measurements were obtained from local measurements on board. At the beginning of the campaign a frequency of 401.0 MHz was selected, which later on had to be switched to 401.2 MHz because of radio interference at 400.9 MHz from an unknown source. This interference caused loss of signal for two radiosondes during their ascent. When a previous sounding was not terminated at the launch time of a subsequent sounding, a frequency of 400.7 MHz was selected.

The *Atalante* experienced substantial instabilities of the Vaisala acquisition system at the initialization step of the system (system location unavailable) and with the reception of the GPS signal by the Vaisala antenna and radiosondes. These problems required multiple restarts of the software and the acquisition system (between 1 and 8 times), creating delays between 10 min and 1 h. However, they did not affect the quality of the soundings. The operators checked the cables and replaced the GPS antenna of the Vaisala system with an antenna that had a larger DC voltage range (15 V instead of 4 V). Nevertheless, the problems persisted during the cruise with the need to restart the system several times before each launch.

F.2.1.3 *R/V Maria S. Merian (Merian)*

The *Merian* launched 118 sondes and recorded 38 descents. Fewer sondes were launched on the *Merian* than other platforms (Fig. F.5) due to difficulties and priority of *Atalante* sondes when the ships were close to one another. The radio signal was often lost using the first antenna location, which the team suspected was due to blocking by the chimney. A new location improved the reception of the signal.

The *Merian* was equipped with a launch container (Fig. F.4c). The helium fill level was decided by inflating the balloon until it reached the upper edge of the launch container. During the day, temperatures in the container rose considerably higher than ambient, but the container was well ventilated as the launch was prepared, such that the instruments experienced typical temperatures of 28–31 °C during synchronization, with only few exceptions. Nonetheless, the residual warming could be a source of bias relative to the surface meteorology observations and persist for tens of meters after the launch. Near-surface data were taken from ship measurements.

F.2.2 *Real-time sounding data distribution*

Sounding observations distributed in real-time over the Global Telecommunication System (GTS) improve atmospheric analyses for initializing and verifying weather forecasts, and improve subsequent reanalyses. Therefore, we aimed to disseminate as much of the full 1-second resolution radiosonde data from the EUREC⁴A campaign as possible over the GTS, regardless of the launch time. Radiosonde data (ascent and descent) from the *Atalante* (114 reports during the campaign) and the BCO (60 reports in February) were sent to the GTS through a Météo-France entry point. This allowed their assimilation in numerical weather prediction (NWP) systems. Most of the *Brown* data were sent to the US National Center for Environmental Prediction (NCEP). From here they were ingested into US Weather Service and Navy NWP systems, yet not European ones. None of the data from the *Merian* and *Meteor* could be transmitted to the GTS by satellite internet. However, during EUREC⁴A, 29 daily ascent soundings from the *Meteor* were sent to the GTS via the EUMETNET Automated Shipboard Aerological Program (E-ASAP), at around 1630 UTC. The WMO station identifiers and designators for tracking the data within the GTS are listed in Table 1 for each station.

World Meteorological Organization Binary Universal Form for the Representation of meteorological data (BUFR) were submitted to the GTS and exchanged among the platforms during the EUREC⁴A campaign.

F.2.3 *Quality control and data formats*

The Vaisala RS41 temperature and humidity measurements are highly robust and accurate, even in cloudy environments. The humidity sensor is actively heated to prevent water condensation and frost formation on the sensor surface. The Vaisala MW41 software writes proprietary .mwx binary files which are ZIP-archives that contain both the raw as well as the processed measurements. These data make up our Level-0 data set. We also provide Level-1 and Level-2 data, which we describe in the following. Our

assignment of levels for the data sets adheres to the standards laid out in Ciesielski et al., 2012.

Sometimes the launch detection did not work properly, which resulted in differences of more than 30 m between the surface altitude and the first reported sonde altitude. Such profiles were reprocessed by correcting the launch time in the raw files. The files were then processed like the corrected files from the *Meteor* (see Section 2.1.2).

F.2.3.1 Level-1 data

Level-1 data in NetCDF format are quality controlled and averaged to 1-second resolution from the Level-0 data. Because the pressure, temperature and humidity are measured with a different sensor (PTU) than wind and position, the data are synchronized to the PTU time. This synchronization is done by the Vaisala MW41 software and the results are included in the Level-0 archive files. The Level-1 data were processed from these results.

The Vaisala MW41 sounding system applies a radiation correction to daytime temperature measurements by subtracting increments that vary as a function of pressure and solar zenith angle. The uncertainty of the radiation correction is typically less than 0.2 °C in the troposphere; uncertainty gradually increases in the stratosphere.

The Vaisala system applies algorithms to adjust for time lags of the RS41 sensors. At 10 hPa the response time of the temperature sensor is 2.5 s for an ascent speed of 6 m s⁻¹. At 18 km (75 hPa) with a temperature lapse rate of 0.01 °C m⁻¹ and an ascent rate varying from 3 to 9 m s⁻¹, the remaining uncertainty in the temperature reading due to time lag is 0.02 °C. At lower altitudes the uncertainty is even smaller. A time-lag correction is also applied to measurements of humidity. The response time of the humidity sensor is dependent on the ambient temperature. For example, at an ascent rate of 6 m s⁻¹ and at 1000 hPa it is <0.3 s for +20 °C and <10 s for -40 °C. The remaining combined uncertainty during the sounding is 4 % relative humidity.

After time-lag adjustments, the Vaisala MW41 quality control algorithm detects outliers and smooths the data to reduce noise. Periods of super-adiabatic cooling are interpolated, and this also applies to temperature differences right above the surface. The MW41 software applies the same correction and quality control steps to the descending and ascending phases of a sounding. Descending sondes, however, can be subject to uncontrollable factors. For example, a falling device may be affected by the remaining debris of a balloon. For this reason, Vaisala does not guarantee the same above-mentioned error margins for data from descending soundings. Our software (Schulz, 2020) reads the processed Vaisala mwx, and MeteoModem cor files, and converts them to self-describing NetCDF files. We also add the ascent or descent rate, calculated from the geopotential height and time information between consecutive measurements, to the NetCDF files. The resolution of the measurements is 1 s. The resulting NetCDF files are the Level-1 data set distributed here.

F.2.3.2 Level-2 data

To facilitate scientific analyses, Level-2 data are provided on a common altitude grid with bin sizes of 10 m, by averaging the Level-1 data. Mean temperature, wind components, position, and logarithm of pressure are directly averaged within height centered bins.

Relative humidity is calculated from the mean of the Level-1 water vapor mixing ratio, calculated from the water vapor pressure formula of Hardy, 1998, which is also used by the Atmospheric Sounding Processing ENvironment (ASPEN) software (Suhr and Martin, 2020) for EUREC⁴A dropsonde measurements. Surface-launched soundings were not reprocessed with ASPEN, as the ASPEN manual warns against duplicating quality control procedures applied by the Vaisala MW41.

In case of missing data within a sounding, we linearly interpolate gaps of up to 50 m. Gaps larger than 50 m, as well as data below 40 m in our Level-2 data set originating from the ship soundings, are filled with missing values. Discarding the lowest 40 m avoids potential biases in the soundings associated with local ship effects, like heating or exhaust plumes, and other problems that are discussed by, e.g. Hartten et al., 2018. Yoneyama et al., 2002 found ship influences on radiosonde measurements to extend no further than 40 m above the deck.

F.3 DATA CHARACTERISTICS

F.3.1 *Ascending versus descending soundings*

We begin with an examination of instrument ascent and descent speeds for the different platforms (Fig. F.6). The figure is based on the ascent (or descent) rates with a 10-m vertical resolution included in the Level-2 data. The median ascent speed in the mid-to-upper troposphere is between 4.5 and 5 m s⁻¹ for radiosondes launched from the BCO, *Atalante* and *Merian* (Fig. F.6a, g, i). Radiosondes launched from the *Meteor* and the *Brown* ascended at slightly slower rates of about 4 m s⁻¹ (Fig. F.6c, e). For all platforms and at all altitudes the 10th and 90th percentiles are roughly symmetric about the median ascent rate and fall mostly within ± 1 m s⁻¹ of the median. Radiosondes from the *Atalante* and *Merian* appear to have experienced stronger updrafts in the upper troposphere. This is consistent with sampling the more convectively-active conditions in the south, where there is a warmer ocean surface, more precipitable water, deeper convection and a greater chance of land influences. Above 20 km, the median ascent rate and the spread in ascent rates increase for all platforms.

Descent speeds exhibit a much stronger functional dependence on altitude (Fig. F.6b, d, f, h, j). For platforms that employed parachutes (BCO, *Meteor*, *Atalante* and *Merian*), descent rates decrease towards the ground to a minimum of about 5 m s⁻¹ in the lowest kilometers. Instruments without a parachute from the *Brown* have descent rates of slightly less than 15 m s⁻¹ in the lowest few kilometers. The positive skewness of the distributions associated with stations that used parachutes is due to descending radiosondes with broken or detached parachutes, or with unexpected behavior of the torn balloon remains. With the exponential decrease of air density with altitude, descent rates increase non-linearly and rapidly with altitude, exceeding 20 m s⁻¹ between 20–25 km when parachutes were used and exceeding 40 m s⁻¹ in case of the *Brown*.

Despite corrections and quality control steps applied by MW41, measurements taken during descent may be accompanied by larger uncertainties due to less favorable and more variable measurement conditions. To establish what degree of confidence we may attribute to the descent data, Fig. F.7 compares the measurements of horizontal wind speed, air temperature and relative humidity between ascending and descending sound-

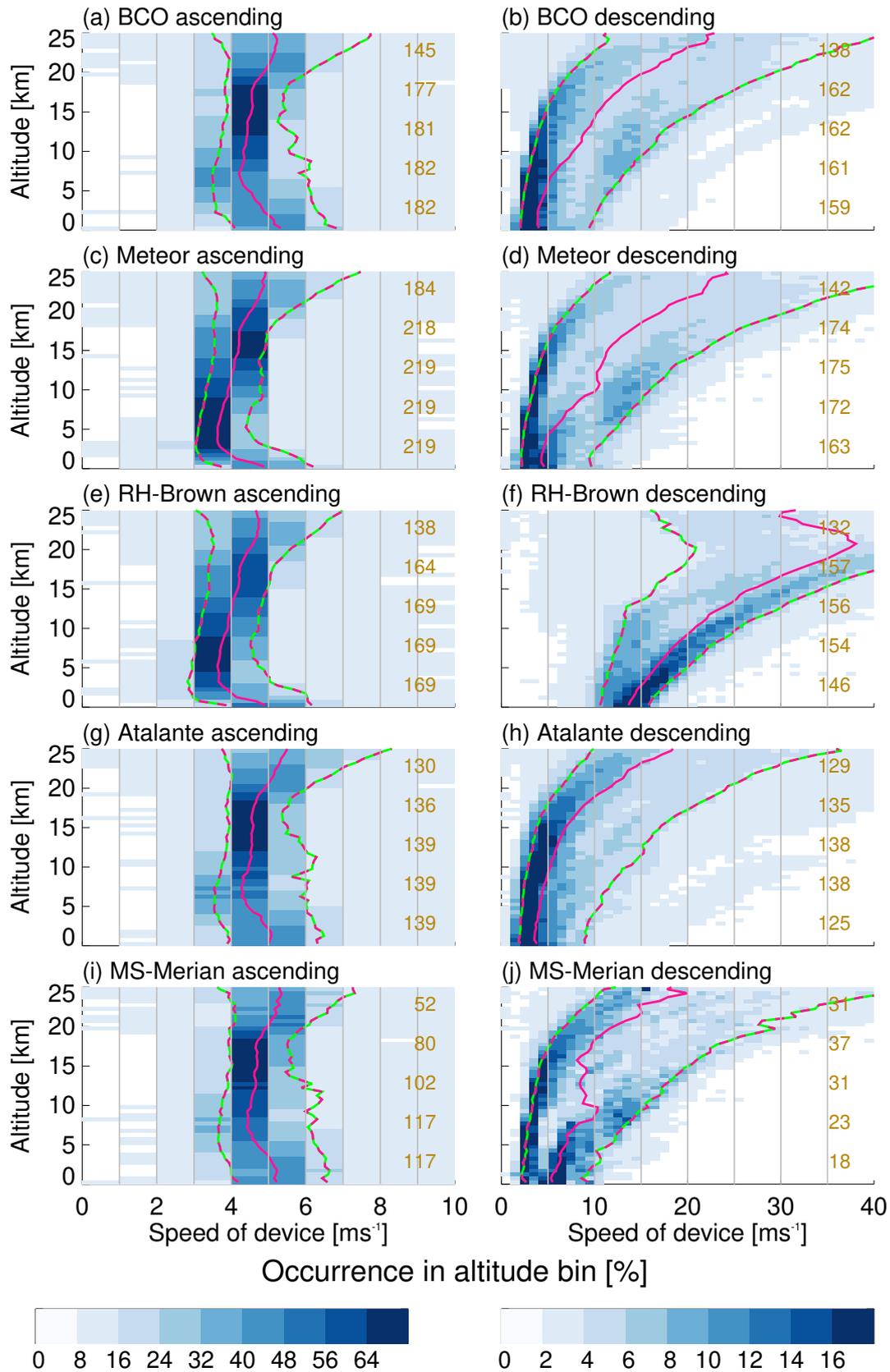


Figure F.6: Instrument (left) ascent and (right) descent speeds as a function of height. The sum of occurrence frequencies in each altitude bin is 100 %. The pink line shows the median profiles and the pink-green lines show the 10th and the 90th percentiles, respectively. Altitude bins are 500 m deep and speed bins are 1 m s^{-1} wide. The numbers of radiosondes that crossed the corresponding height-levels (2.5, 7.5, 12.5, 17.5 and 22.5 km, respectively) are shown in each panel.

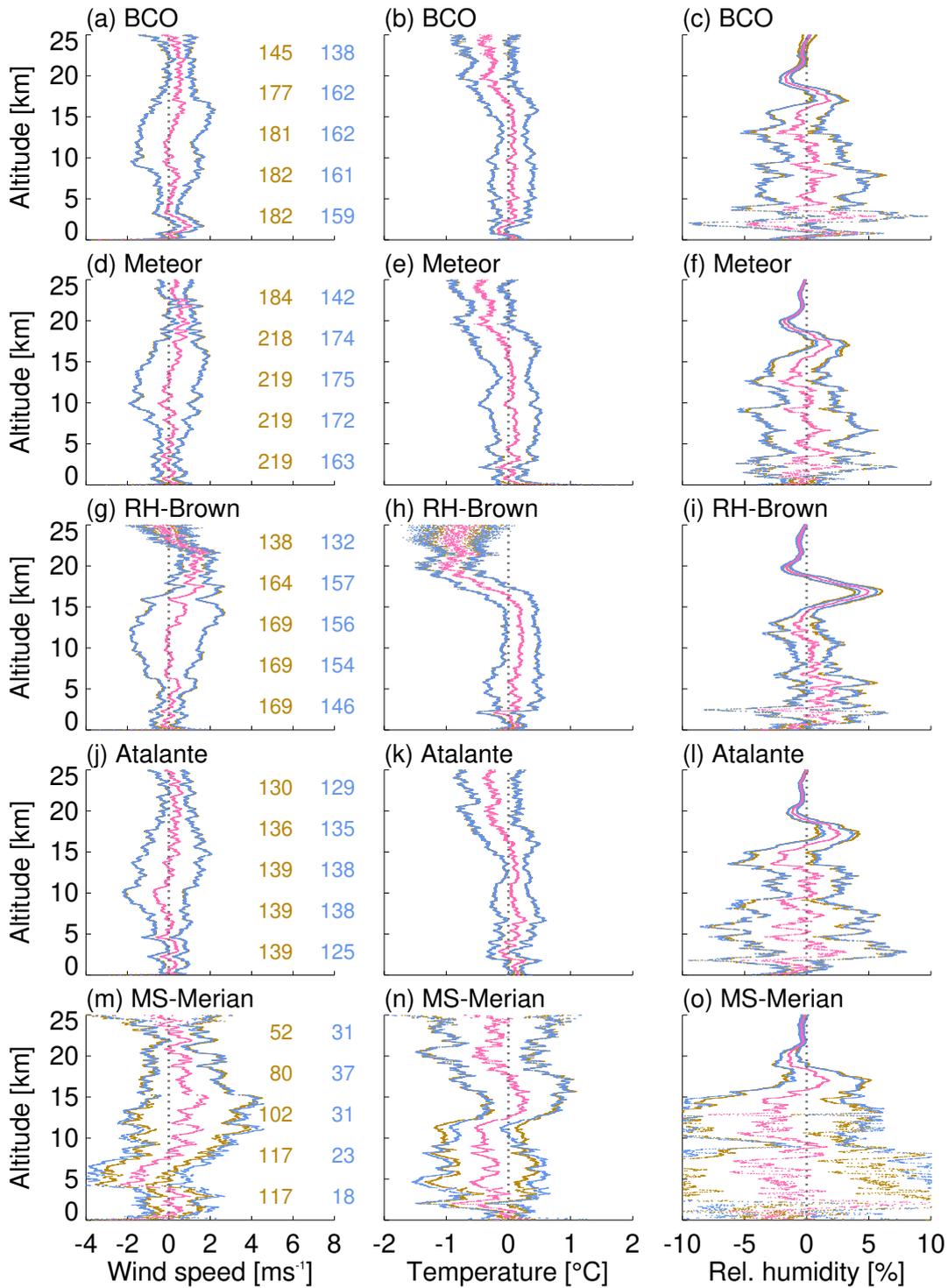


Figure F.7: Comparison of (left) horizontal wind speed, (middle) air temperature, and (right) relative humidity, measured during ascent and descent. The pink dots show the average over all included ascent profiles minus the average over all included descent profiles. Brown (blue) dots show the 95 % confidence intervals for ascent (descent). Numbers inside the panels on the left-hand side show the counts of ascending (brown) and descending (blue) radiosondes that crossed the corresponding height-levels (2.5, 7.5, 12.5, 17.5 and 22.5 km, respectively.)

ings. We do not expect perfect agreement between ascending and descending soundings, for several reasons. First, the instruments drift substantial horizontal distances and hence systematically sample a downwind location (as illustrated in Fig. F.3 for the BCO). Meridional horizontal drift could create systematic biases. Second, there are variable time lags of the order of a couple of hours between ascending and descending measurements, which we expect might increase the scatter between ascent and descent measurements but not create systematic differences. A systematically different response of the sensors during descent might be the most important factor for biases. We also note that the number of descent profiles available for computing statistics is in some cases substantially smaller than the number of ascent profiles (Fig. F.5). The numbers of available measurements are again listed on the left hand side of Fig. F.7. All quantities shown in Fig. F.7 are computed from matched ascent-descent pairs of the same instrument.

Measurements of horizontal wind speeds do not show statistically significant differences between ascent and descent (the mean lies within the 95 % confidence intervals), with the exception of the *Brown*. Here, wind speeds at around 20 km altitude are stronger for the ascent. This systematic difference could be related to excessively rapid descent rates. Similar results are found for measurements of air temperature (Fig. F.7b, d, f, h, j). In case of the *Brown*, stratospheric temperature observations during descent are warmer by more than 1 °C, suggesting a bias due to high descent rates. The same bias exists for the other platforms, but the effect is smaller and not statistically significant at the 95 % confidence level. Differences in relative humidity are not statistically significant inside the troposphere.

F.3.2 Synoptic conditions

We first present the synoptic situation for the region defined by the *Meteor* and the BCO soundings. Our initial analysis focuses on the soundings for these two platforms because they define a more or less fixed geographic area – radiosondes launched from the *Meteor* were almost all launched between 12.5 °N and 14.5 °N along 57.15 °W – bounding the subdomain that was most intensively sampled. A comparison between twelve BCO soundings with coincident and nearly co-located ship-based soundings (ships were positioned just offshore of the BCO) showed no evidence (Fig. F.17) of a systematic influence of the island on the BCO soundings. Hence, the BCO soundings appear representative of the western most boundary of the marine measurement area. Focusing on a fixed region during the period of most intensive airborne operations, between January 20 and February 17, also provides a reference for quantifying differences in soundings taken outside of this region, or time period, as is discussed at the end of this subsection.

Synoptic differences among variables believed to be important for patterns of low-level cloudiness suggest that: (i) the *Meteor* and the BCO sample the same synoptic environment; and (ii) changes in the environment can usefully be described by week-to-week variability over the four weeks starting on Monday, January 20. The lower tropospheric stability (LTS), the near surface winds, the lifting condensation level (LCL) of near-surface air, and the hydrolapse track each other well (Fig. F.8). The hydrolapse marks the depth of the trade-wind cumulus layer. It is defined as the mean height where mean relative humidity on a centered running 500 m range first drops below 30 %. LTS

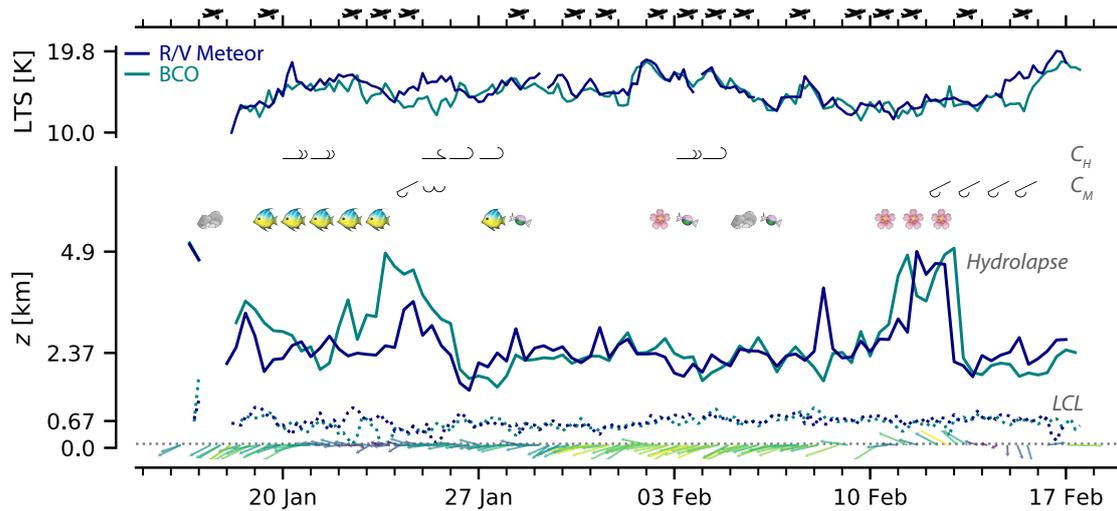


Figure F.8: Synoptic overview of period and region of intensive aircraft measurements. Plotted are the lower tropospheric stability (LTS), the height of the hydrolapse, the lifting condensation level (LCL) and the wind vector averaged over the lower 200 m. Winds are 12 h median values, other quantities are resampled on a 4 h interval, with median values plotted except for the LCL where minimum values are plotted. For the wind vectors the maximum and minimum wind speeds are 12.3 m s^{-1} and 2.0 m s^{-1} , respectively. Tick marks denote maximum and minimum LTS, and maximum and median height of *Meteor* hydrolapse and the mean height of the LCL (*Meteor*). Also shown are days when aircraft with dropsondes were flying, the synoptic cloud observations of mid-level (C_M) and high (C_H) clouds with the associated WMO cloud-symbol (Table 14 of 2017 World Meteorological Organization Cloud Atlas, <https://cloudatlas.wmo.int/en/home.html>) that predominated for that day. Cloud types are taken from the Barbados Meteorological Service SYNOP reports. Days on which a mesoscale pattern of shallow convection, following the classification activity of Schulz (in preparation), was readily identified are indicated by the emojis for Fish, Sugar (candy), Flowers or Gravel (rocks).

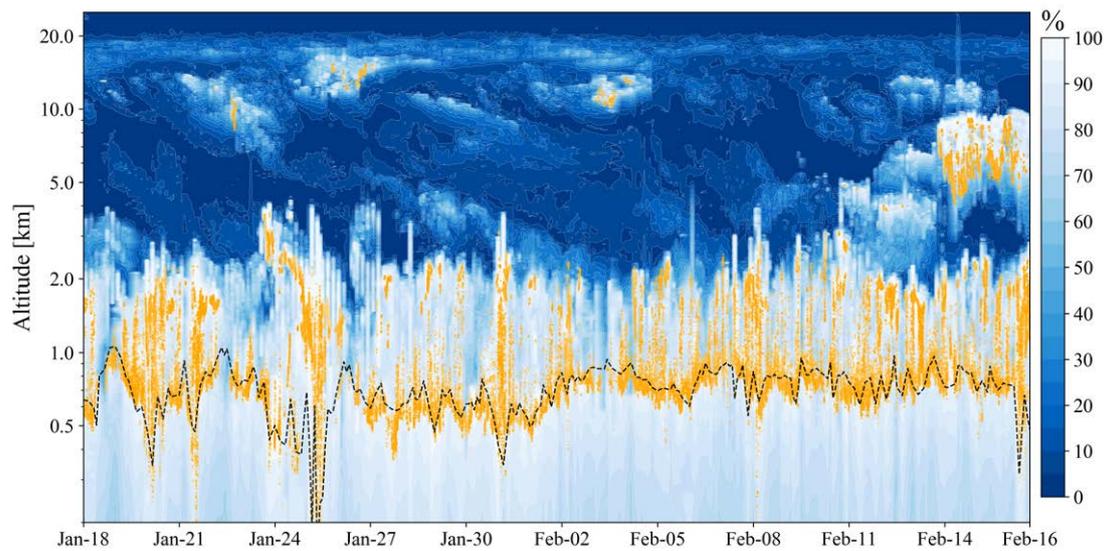


Figure F.9: Comparison between ascending and descending soundings and ceilometer measurements on the *Meteor*. The relative humidity from radiosonde measurements is shown in blue-to-white shading. The black dashed line represents the lifting condensation level calculated based on Bolton, 1980. Cloud base heights as observed by the ceilometer are marked with orange dots. The vertical axis is chosen to be logarithmic for better visibility of the moisture distribution near the surface. The time-axis for the soundings uses launch time. The temporal resolution of the ceilometer data is 10 s. Low-altitude relative humidity profiles (300–800 m) of the descending soundings were recovered by assuming a dry adiabat temperature and a constant humidity profile.

is defined as difference of potential temperature at 700 hPa and the mean potential temperature in the lowest 200 m. Fig. F.9 further illustrates that the LCL tracks well the lowest cloud bases as measured by the *Meteor* ceilometer. Week-to-week variations as deduced from the soundings of either platform show the first and last week to be characterized by a deeper moist layer, and lessened lower tropospheric stability, the latter primarily explained by changes in the potential temperature at 700 hPa. The two week period starting on January 27 has a much shallower trade-wind layer and stronger stability. Near surface winds vary somewhat out of phase with the moisture variability, with winds stronger in the second half of the four week period, and weaker in the first half. The LCL shows very little synoptic variability.

Cloud observations are also included in Fig. F.8. Reports of mid-level (C_M) and high-level (C_H) clouds are derived from 3 hourly SYNOP observations reported by the Barbados Meteorological Service at Grantley Adams International Airport. If a reported mid or high-level cloud type was persistent through the day (more than three reports) it is included via its WMO cloud symbol² in Fig. F.8. Notable are mid-level clouds that coincide with the deepening of the marine layer, particularly during the period at the end where a layer of altocumulus ($C_M = 4$) persisted for several days (Fig. F.9). Observations of low clouds (C_L) indicated that $C_L=8$ and $C_L=2$ where the dominant low-level clouds; both evident on almost every day with little evidence of synoptic variability. This is also evident from the *Meteor* ceilometer measurements (Fig. F.9). For this reason, in Fig. F.8 we instead identify days when particular patterns of mesoscale variability were in evidence. We adopted the four patterns, Sugar, Gravel, Flowers, Fish following Stevens et al. (2020). While the low and small Sugar clouds appear with little organization, Gravel clouds reach deeper extents and organize along gust fronts. The fish-bone like organization of clouds on horizontal scales of 200–2000 km is described by the Fish pattern, and large stratiform, often circular-shaped cloud clumps are labeled as Flowers. Whether or not one particular pattern was identified was taken from a cloud classification activity organized by one of the authors (H. Schulz). These patterns suggest that the initial moist period has the satellite presentation of Fish, and that the period of increased lower-tropospheric stability and strengthening winds on February 2 is associated with the pattern Flowers, consistent with the analysis of Bony et al. (2020).

To give a better impression of the synoptic variability, the period identified with the Fish pattern, between January 22–24, is investigated further. The visible satellite imagery from MODIS on Aqua (Fig. F.10a) illustrates the large-scale characteristics of the observed Fish cloud pattern, covering the BCO and the northern latitudes of the observations region. The pattern resembles a spine in a surrounding cloud-free area and was accompanied by unseasonably large amounts of surface precipitation. Fig. F.10b illustrates the moistening of the atmosphere and the deepening of the boundary layer, as measured at the BCO, over the course of this event. Between January 20–26, the increase of integrated moisture up to 55 kg m^{-2} coincides well with the deepening moist layer, thus also with changes in cloud top height and trade wind inversion height. Before and after the event, the inversion layer height was around 2 km (Fig. F.8), and the boundary layer was characterized by a mixture of Gravel and Sugar, albeit the latter not on a scale that lent itself to identification from the satellite imagery. During the peak of the event on

² These symbols are taken from the 2017 edition (Table 14) of the WMO Cloud Atlas (www.wmocloudatlas.org).

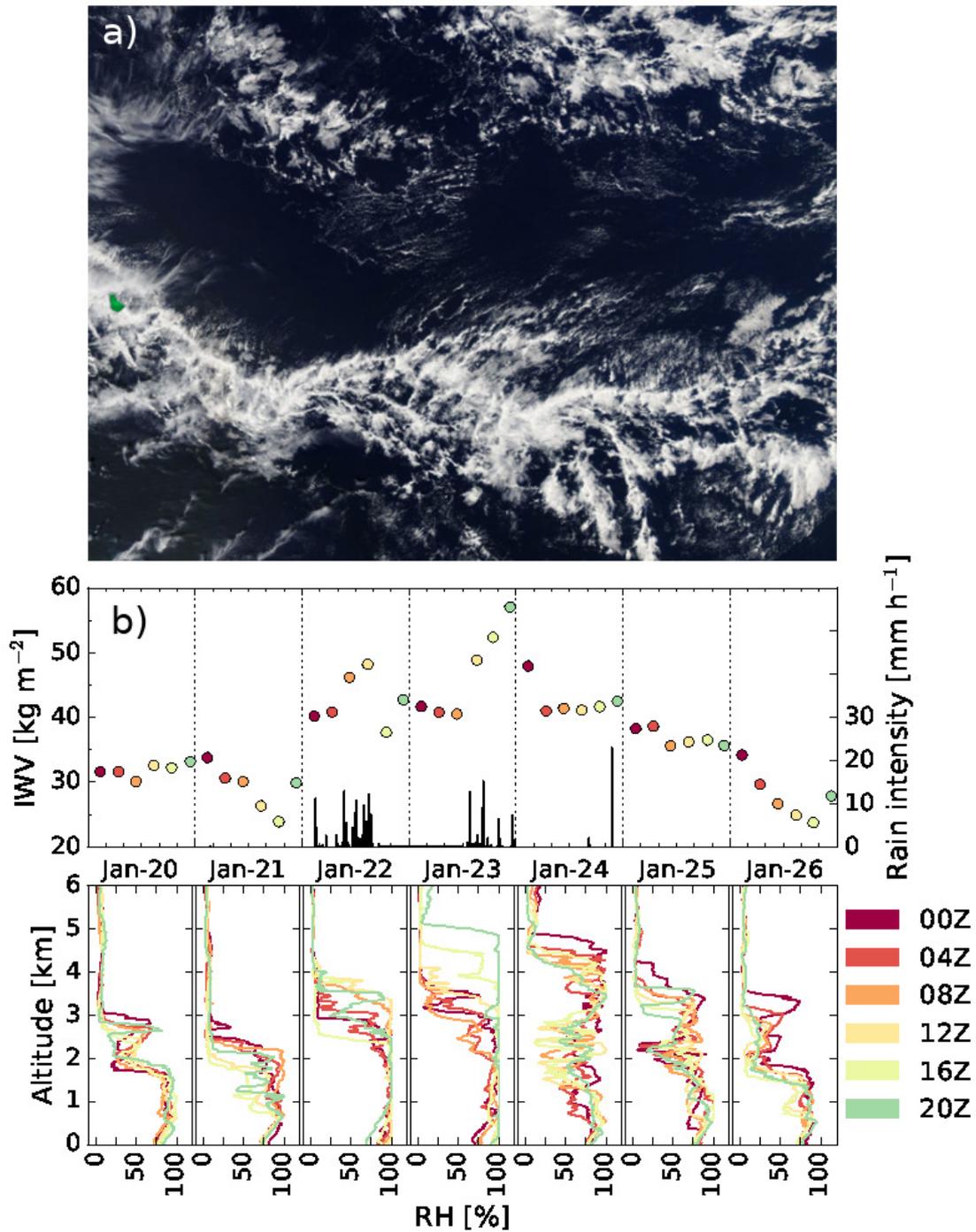


Figure F.10: Fish cloud pattern passing Barbados between January 22–24, 2020. (a) MODIS-Aqua scene from January 22. The image covers 9–18 °N, 48–60 °W with Barbados shown in artificial green. (b) Temporal evolution of relative humidity (lower panel) and integrated water vapor (I WV; upper panel, color-coded) as measured by the BCO soundings January 20–26. Profiles and calculated I WV values are color-coded according to the nearest hour of the sounding reaching 100 hPa. The upper panel also shows a one-minute running mean of rain intensity recorded at BCO (black).

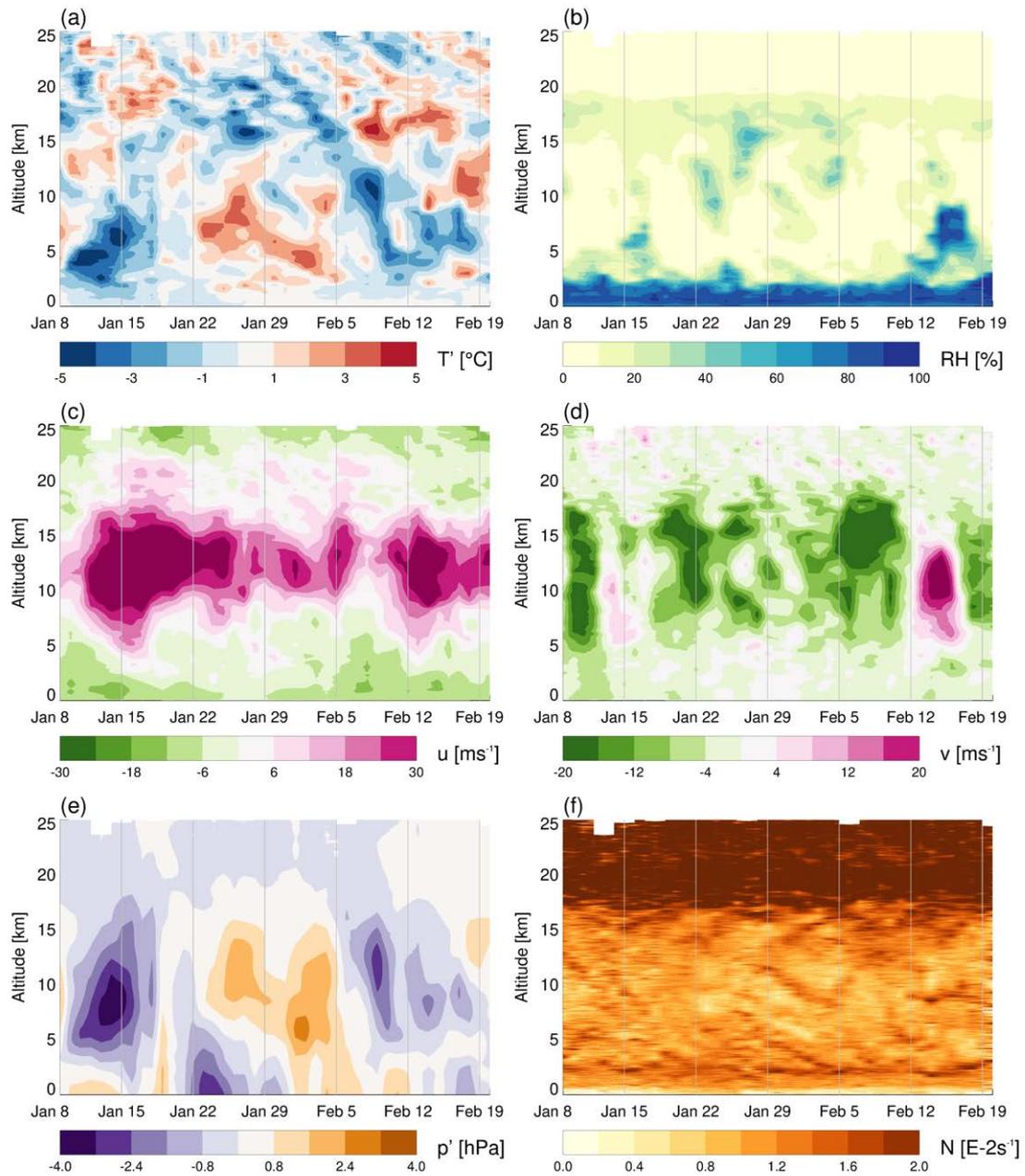


Figure F.11: (a-e) Time-height cross sections of daily (a) temperature anomaly, (b) relative humidity, (c) zonal wind, (d) meridional wind anomaly, (e) pressure anomaly and (f) Brunt-Väisälä frequency (units of 10^{-2}s^{-1}), computed from ascending soundings north of 12.5°N . The data combine 182 soundings from the BCO, 169 from the *Brown*, 159 from the *Meteor*, 30 from the *Merian* and 4 from the *Atalante*. Anomalies are defined as deviations from the time average at each altitude.

January 22 and 23, the moisture layer deepened up to 5 km. While the Fish cloud pattern passed over BCO, the pressure in the boundary layer decreased by up to 4 hPa (see Fig. F.11e) and the temperature in the upper middle troposphere (6–8 km) showed a slight positive anomaly (see Fig. F.11a). The rain intensity, measured at BCO with a Vaisala WXT-520 ground station, peaked at 15 mm h^{-1} , and precipitation events were persistent, in contrast to the short rain showers more typical of the dry season (Stevens et al., 2016). Bony et al., 2020 found that the Fish cloud pattern often occurs under weaker surface trade wind speeds below 8 m s^{-1} ; the sounding data confirm this, as the measured wind speeds lie well below this threshold in the lower boundary layer, e.g. Fig. F.8.

Given that the vertical structure of the humidity field appears to be a strong indicator of synoptic variability, time-height humidity plots for all of the platforms are used to explore the coherence of synoptic conditions sampled by individual platforms. This analysis (Fig. F.12) shows that soundings from the *Brown*, which moved around more, but stayed mostly north of 12.5°N and east of the *Meteor*, sampled a similar synoptic environment. The *Merian* and *Atalante* however were further south and their soundings show a humidity structure and evolution that is less coherent than seen by the ships in the Trade-wind Alley. Based on this finding and because performing the same analysis for any one station does not change the big picture, we composite the soundings from all of the platforms north of 12.5°N . Figure F.11 shows the temporal evolution of atmospheric conditions for the full period of data coverage averaged north of 12.5°N , i.e., over the Trade-wind Alley. Before January 22 the mid-troposphere is relatively cool and zonal winds in the upper troposphere are strong. From January 22 onward the observational domain experienced warmer temperatures, weaker upper-tropospheric westerlies, as well as weaker easterlies near the surface. Positive pressure anomalies first appear in the upper troposphere and reach the surface at the end of January when a ridge starts to dominate the area. Surface and upper-tropospheric winds strengthen again after February 6 when the positive pressure anomaly fades. A strong moistening of the mid and upper levels is seen around February 13, which coincides with a directional change of the meridional winds at these levels, favoring the aforementioned extensive and persistent altocumulus cloud layer (Fig. F.8).

Most differences between the structure of the atmosphere within the Trade-wind Alley (North of 12.5°N) and the ‘Boulevard des Tourbillons’ (southern corridor) are confined to the structure of the lower-tropospheric humidity. South of 12.5°N , the atmosphere was on average much more humid in the lower and middle troposphere, as shown in Fig. F.13. This humidity anomaly is not persistent, as dry conditions, similar to those observed north of 12.5°N , were also present; it can rather be associated with more frequent periods of a deep moist layer and deeper convection, for example as observed during the period around January 29 (see Fig. F.12). Additional, albeit less substantial differences (not shown), are that middle-upper troposphere relative humidities (between 7–10 km) are actually somewhat drier in the South. There is very little evidence of systematic differences in the temperature structure between the northern and southern soundings, except for a hint of enhanced stability in the upper troposphere (11–15 km) in the North. Over the ‘Boulevard des Tourbillons’, the depth of the near surface easterly layer is 1–2 km shallower and between 5–15 km, the westerlies have a stronger northerly component.

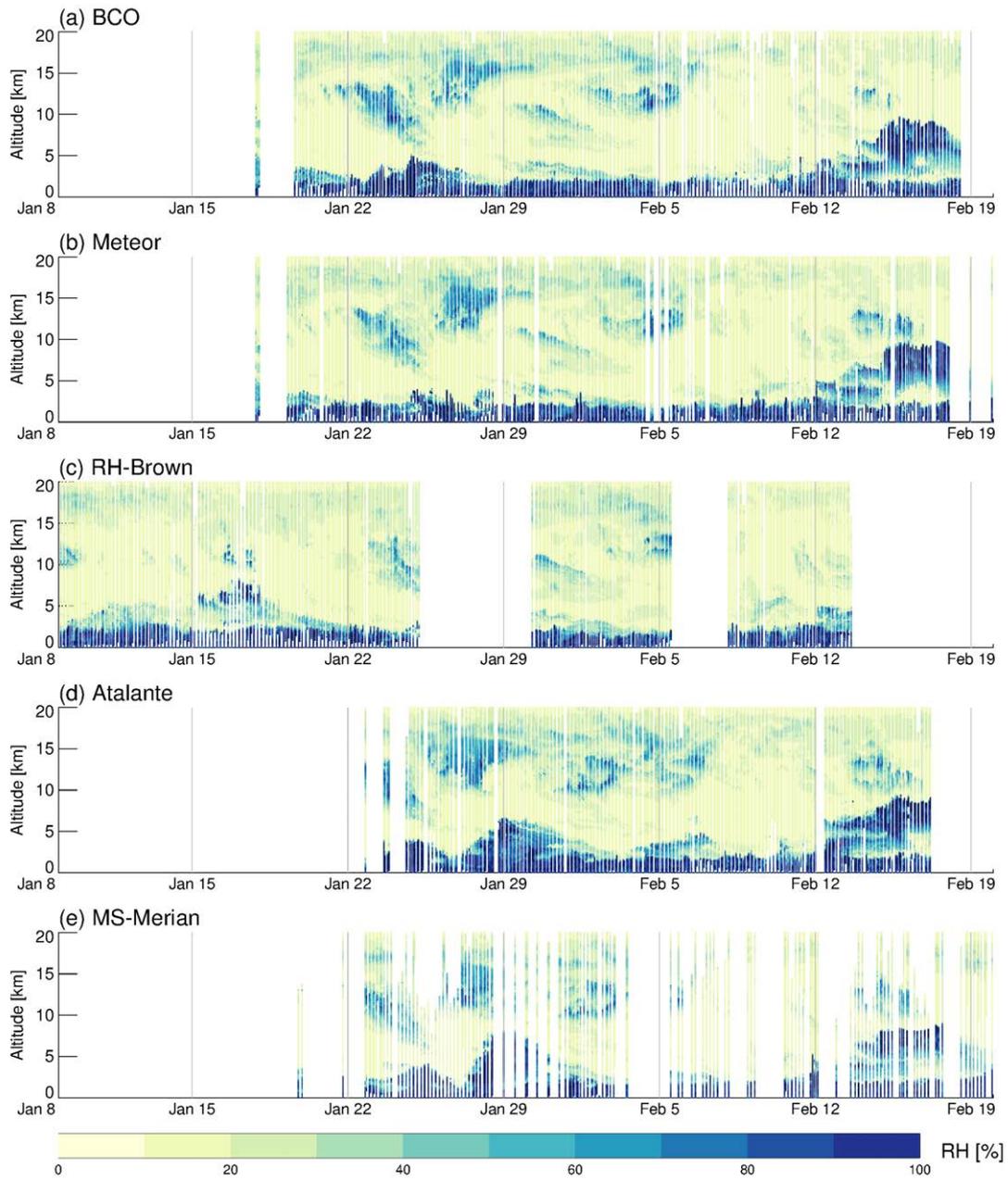


Figure F.12: Time-height series of relative humidity measurements from all platforms. The plot combines ascending and descending soundings.

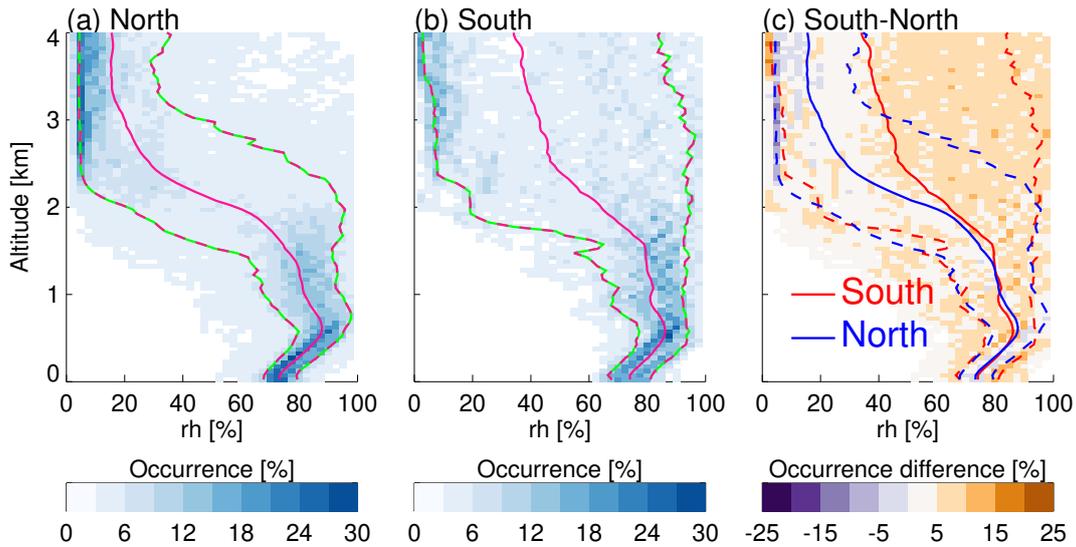


Figure F.13: Occurrences of relative humidity as a function of height below 4 km for soundings launched between January 26 and February 12. The sum of occurrence frequencies in each altitude bin is 100 %. Altitude bins are 50 m deep and each x -axis contains 40 bins. North (panel a) designates soundings from the northern ($12.5\text{--}14.5^\circ\text{N}$; 261 profiles) latitude band, and South designates soundings from southern ($8.5\text{--}10.5^\circ\text{N}$, 63 profiles) latitude band. Solid lines show the mean profiles in each region and dashed lines the 10th and the 90th percentiles. Only data from ascending radiosondes are used in this comparison.

F.4 SUMMARY

The EUREC⁴A field campaign during January–February 2020 included among its wide range of observational platforms an extensive radiosonde network, consisting of the Barbados Cloud Observatory and four research vessels. 182 radiosondes of type RS41-SGP were successfully launched in a regular manner between January 16 and February 17 from the BCO, 203 between January 18 and February 19 from the *Meteor*, 169 between January 8 and February 12 from the *Brown*, 139 between January 21 and February 16 from the *Atalante*, and 118 between January 20 and February 19 from the *Merian*. In addition, 47 MeteoModem radiosondes of type M10 were launched from the *Atalante* during intensive observational periods to sample variability associated with sea surface temperature fronts. These are described in the appendix.

We made data at three stages publicly available. Level-0 data contain the raw .mwx binary files, which can be read and processed with the MW41 software. Level-1 data were subject to Vaisala’s standard quality control algorithm, which detects outliers in the profiles, performs a smoothing to reduce noise, and applies time-lag and radiation corrections. The Level-1 file format is NetCDF with a temporal resolution of 1 s. To facilitate scientific analyses, Level-2 data are vertically gridded by averaging Level-1 data in 10-m bins. All soundings, ascending and descending, from each platform were collected into one NetCDF file for the Level-2 data.

The *Meteor* and the *Brown* followed nearly-orthogonal sampling lines, mostly in the latitude band $12.5\text{--}14.5^\circ\text{N}$, whereas the *Atalante* and *Merian* sampled conditions further

to the south. It was a central goal of EUREC⁴A to better understand the formation and feedbacks of different patterns of shallow cumulus clouds. We were fortunate that nature provided us with a wide variety of cloud conditions, which are reflected in the radiosonde data. The six weeks of sounding data at high temporal resolution should render the radiosonde data described herein useful for a large variety of scientific analyses.

F.5 CODE AND DATA AVAILABILITY

Raw Level-0 data consist of single files per sounding in `.mwX` format, which combine ascent and descent from each instrument. Quality-controlled Level-1 data consist of single files per sounding in NetCDF format, with separate files for ascent and descent. Level-2 data are stored in a single file per station and include data on a 10-m vertical resolution grid, including all available ascents and descents. Ascent and descent can be distinguished by a flag that indicates the direction. All data (Stephan et al., 2020a) are archived and freely available for public access at AERIS (<https://doi.org/10.25326/137>). Our software, which we used to convert to NetCDF format is also publicly available (Schulz, 2020; <https://doi.org/10.5281/zenodo.3712223>).

F.6 APPENDIX

F.6.1 *Extra soundings on board the ATALANTE*

In addition to the regular Vaisala soundings, further soundings were performed from the *Atalante* primarily to sample the lower atmosphere across sea surface temperature (SST) fronts associated with oceanic mesoscale dynamics. An independent radiosonde receiver was used to not interfere with the regular soundings depicted in this article. MeteoModem M10 radiosondes were chosen for availability and cost. In order to decide the period of intensive sampling using these sondes, we first identified on a daily basis the ocean mesoscale eddies and currents by applying the TOEddies detection algorithm (Laxenaire et al., 2018) to the Ssalto/Duacs Near Real Time (NRT) altimeter products (Absolute Dynamic Topography – ADT – and the associated surface geostrophic velocities; Ablain et al., 2017, Taburet et al., 2019).

These data were successively analyzed together with the NRT SST produced by Collecte Localisation Satellites (CLS), the ship’s ThermoSalinoGraph (TSG) 5 m-depth temperature measurements, and ARPEGE and ECMWF forecasts in order to decide in real time the launching strategy. The NRT CLS SST is produced as a 1-day average, high-resolution product, which is a simple data average of the satellite measurements taken over the previous day, and has a resolution of 0.02° in latitude and longitude. This product may have local gaps due to the presence of clouds or missing data. The CLS SST NRT product is derived from nighttime observations (to avoid diurnal warming of the sea surface) by the MODerate-resolution Imaging Spectroradiometer (MODIS) on board TERRA and AQUA satellites, the Advanced Very High Resolution Radiometer (AVHRR) on board METOP-A and -B, the Visible Infrared Imager Radiometer Suite (VIIRS) on board Suomi-NPP, the Advanced Himawari Imager (AHI) on board HIMAWARI-8, and the Advanced Baseline Imager (ABI) on board GOES-16 and -17.

Precisely setting the sounding periods was difficult because the satellite observations were only available for the previous day with additional uncertainties in the location of SST fronts due to cloud screening. Furthermore, this strategy was defined in coordination with the *Merian* to take into account the oceanographic observation goals common to both ships.

The first targeted and intensive radiosonde observation leg took place on January 26. 11 MeteoModem sondes were launched while crossing a SST front associated with a relatively cold filament (-0.5 to -1 °C SST anomaly) steered from the Guyana coast by a mesoscale anticyclonic eddy (Fig. F.14a). During this leg, the ship crossed a front of about 0.5 °C extending over 30 km with near surface wind of 6 – 7 m s⁻¹ magnitude and 60° – 70° direction. During this leg the ship was heading eastward, almost into the wind. Figure F.14a shows the February 25 SST map, chosen as clouds prevented retrievals on the following day. According to the satellite product, one would have expected to meet the front further east. Fortunately, a first diagonal transect during the night provided us with the actual front location.

The second targeted and intensive radiosonde observation leg took place on February 2–3. This leg lasted for about 24 hours during which 28 MeteoModem radiosondes were launched while the ship was zigzagging in order to sample several times the northeastern edge of a cool SST anomaly of nearly -1 °C associated with coastal upwelling off the Suriname and French Guyana coast (Fig. F.14b). During this leg, the ship was moving westward and sampled SST variations of 0.3 – 1 °C extending over 50–60 km. At this time the near surface wind was variable in direction, 40° – 80° , and relatively strong (8 – 11 m s⁻¹).

The remaining MeteoModem radiosondes were launched on few diverse occasions: two were launched in the center of the warm core of a second eddy on January 27. Another radiosonde was launched under a convective system on February 10. The last four launches took place in cloud streets on February 17.

We used M10 GPS radiosondes with an SR10 station and EOSCAN (1.4.200306) software. With the exception of one sounding, only ascent data are available for these soundings as most of the launches were stopped manually at about 10 km height to increase the sampling frequency of the lower atmosphere in regions characterized by SST fronts. Launch frequencies reached up to one sounding every 40 min during the intensive launch periods. Therefore, several radiosondes were emitting at the same time, so frequencies had to be changed within the 400.4–403.4 MHz band to avoid interference. M10 radiosondes measure relative humidity and temperature, from which dew point temperature is deduced. The altitude and horizontal displacements of the radiosondes are measured by GPS and are used to diagnose the horizontal wind components. Unlike with RS41 SGP sondes, the pressure is deduced from the altitude and the surface station pressure measurement, using the hydrostatic approximation. Our published data formats, NetCDF and ASCII formatted files (.cor files), both contain data reported every second. The raw MeteoModem data are processed in the same way as the Vaisala soundings to create Level-1 and Level-2 files that match the format of the corresponding Vaisala data. The only difference is that the description of the MeteoModem corrections that are automatically applied by the software is a trade secret and therefore not known to us. However, the M10 sondes are currently in the process of being certified by the Global Climate Observing System Reference Upper-Air Network (GRUAN). If the GRUAN

certification is granted, details on these corrections will become available. We checked for and corrected spurious data in the surface observations using handwritten log-sheets filed during the campaign.

Figure F.15 illustrates the outcome of these targeted and intensive radiosonde observations with results from the February 2–3 intensive observation period (Fig. F.14b). Profile color (Fig. F.15a–c) denotes the SST measured by the ship at the time of the launch (Fig. F.15d). Blue (red) profiles are thus on the cold (warm) side of the SST front. These profiles are from raw data (level-0) and no attempt was made to validate, correct or remove doubtful data such as the surprisingly cold layer between 800–900 m altitude that can be seen in one of the blue potential temperature profiles (Fig. F.15a). No attempt has either been made to disentangle diurnal or synoptic scale variability from the imprint of the SST front on the lower atmosphere. However, one can note that the warm side of the SST front was sampled mostly during nighttime (local noon at 1530 UTC, nighttime from 22–10 UTC). There is a clear tendency for warmer boundary layers over the warm side of the front than over the cold side (Fig. F.15a). On the other hand, the height of the mixed layer, that can be defined as near homogeneous potential temperature layers close to the surface, tends to be deeper over the cold side than over the warm side. This contrasts with results obtained over stronger SST fronts from observation (Ablain et al., 2014) and modeling studies (e.g., Kilpatrick et al., 2013; Redelsperger et al., 2019) and suggests that the lower atmosphere does not solely respond to the SST gradient. Over the cold side, wind speed tends to decrease with altitude (Fig. F.15b). Over the warm side, and despite a larger variability from a profile to another, the wind speed tends to be more homogeneous in the vertical than on the cold side. Because the mixed layer depth is shallower over the warm side, it is however difficult to interpret this as the result of a stronger vertical turbulent mixing. Overall, near surface wind speed tends to be slightly weaker on the warm side than on the cold side. There is also a noticeable change in wind direction throughout the boundary layer from E-NE over the warm side to NE over the cold side (Fig. F.15c).

Finally, we provide a first assessment of the quality of MeteoModem M10 measurements based on the *Atalante* soundings, as also Vaisala soundings were launched during the intensive MeteoModem periods. We compare MeteoModem and Vaisala wind, temperature and relative humidity profiles for 8 pairs of soundings that were launched within 25 min (Fig. F.16). Choosing such a small time period certainly limits the number of difference profiles that can be computed, but it ensures that the two radiosondes sampled comparable situations. Mean difference profiles and corresponding standard deviations are computed on 100 m bins. Neither horizontal wind components (Fig. F.16a, b) nor temperature (Fig. F.16c) show any clear bias, although the differences between MeteoModem and Vaisala can be a few m s^{-1} for the wind components (standard deviation of about 0.5–1 m s^{-1}) and about 1 °C for temperature (standard deviation of about 0.1–0.2 °C). On the other hand, despite a large noise below 4 km height, relative humidity shows a rather homogeneous moist bias of about 5 % (1–5 % standard deviation) in MeteoModem measurements compared with Vaisala (Fig. F.16d). No correction was applied, neither to the temperature nor to the relative humidity measurements. In particular, corrections for the relative humidity seem necessary but are still a matter of research. An example of such corrections, developed for soundings in the continental mid-latitude can be found in Dupont et al., 2020.

F.7 ACKNOWLEDGEMENTS

We acknowledge Olivier Garrouste and Axel Roy from Météo-France for lending the MeteoModem station and launcher and their technical assistance, and Thierry Jimonet and Météo-France Guadeloupe for their help. We would like to thank Rudolf Krockauer and Tanja Kleinert (DWD), Bruce Ingleby (ECMWF) and Jeff Ator (NOAA) for helping to track the radiosonde data sent to the GTS. We would like to thank Sophie Bouffies-Cloch e (IPSL) and Florent Beucher (MF) for making ECMWF and ARPEGE forecasts available in real time. We thank the captains and crews of all ships for their help with completing successfully all the planned observations. We are grateful to Doris J ordens for the technical support with the Vaisala equipment, and to Ingo Lange, Harald Budweg, and Rudolf Krockauer for their help with training our operators. We would like to thank the numerous volunteers performing launches at BCO, and Andreas Kopp, Joachim Ribbe, Klaus Reus, Diego Lange, Alex Kellmann and Nicolas Geyskens for launches on the research vessels. We would like to express our special gratitude to Angela Gruber for resolving any problems before they became problems. The ceilometer data from the *Meteor* (Jansen, 2020a) and surface meteorology data from the BCO (Jansen, 2020b) are publicly available. We acknowledge the use of imagery from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS). The Ssalto/Duacs altimeter products were produced and distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS) (<http://www.marine.copernicus.eu>). The Collecte Localisation Satellites (CLS) SST and Chlorophyll-a catsat products (<https://www.catsat.com/ocean-data/>) were made available to us during the cruise in the framework of the Donn ees et Services pour l’Oc ean (ODATIS) French national data infrastructure. The *Atalante* cruise was funded under the EU Horizon 2020 program under grant agreement 817578 TRIATLAS project, and under the EUREC⁴A-OA project by the following French institutions: the national INSU-CNRS program LEFE; the French Research Fleet; Ifremer; CNES; the Department of Geosciences of ENS through the Chaire Chanel program; M et eoFrance. Vaisala radiosondes were funded by the Max Planck Society, and the US National Oceanic and Atmospheric Administration Ocean and Atmospheric Research grant number NA19OAR4310375. MeteoModem radiosondes were funded thanks to Caroline Muller by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Project CLUSTER, grant agreement No. 805041). Claudia C. Stephan was supported by the Minerva Fast Track Program of the Max Planck Society.

F.8 SUPPLEMENTARY MATERIAL

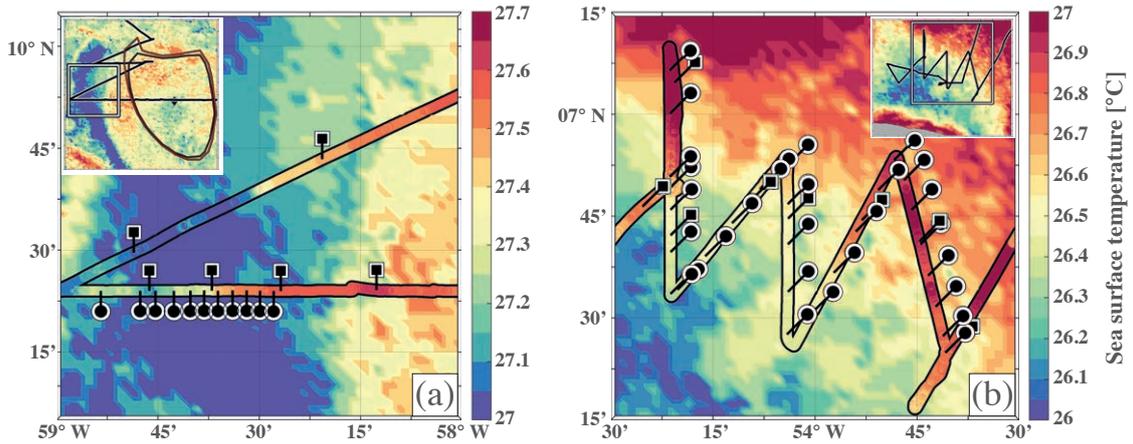


Figure F.14: Maps of CLS SST ($^{\circ}\text{C}$) for (a) January 25, 2020, and (b) February 2, 2020, with the *Atalante* track during the first (January 26) and second (February 2–3) intensive leg, respectively. The color shows the SST measured by the ship’s ThermoSalinoGraph (TSG) at 5 m depth and the ticks show the location of Vaisala (squares) and Meteo-Modem (circles) radiosonde launches. Inserts in the upper corners, where the black lines indicate the ship’s course, show the larger scale view of the corresponding scenes with the geographical imprint indicated by white squares. In the panel insert a, the closed contours and the black diamond indicate, respectively, the edges of an anticyclonic eddy and the position of its center.

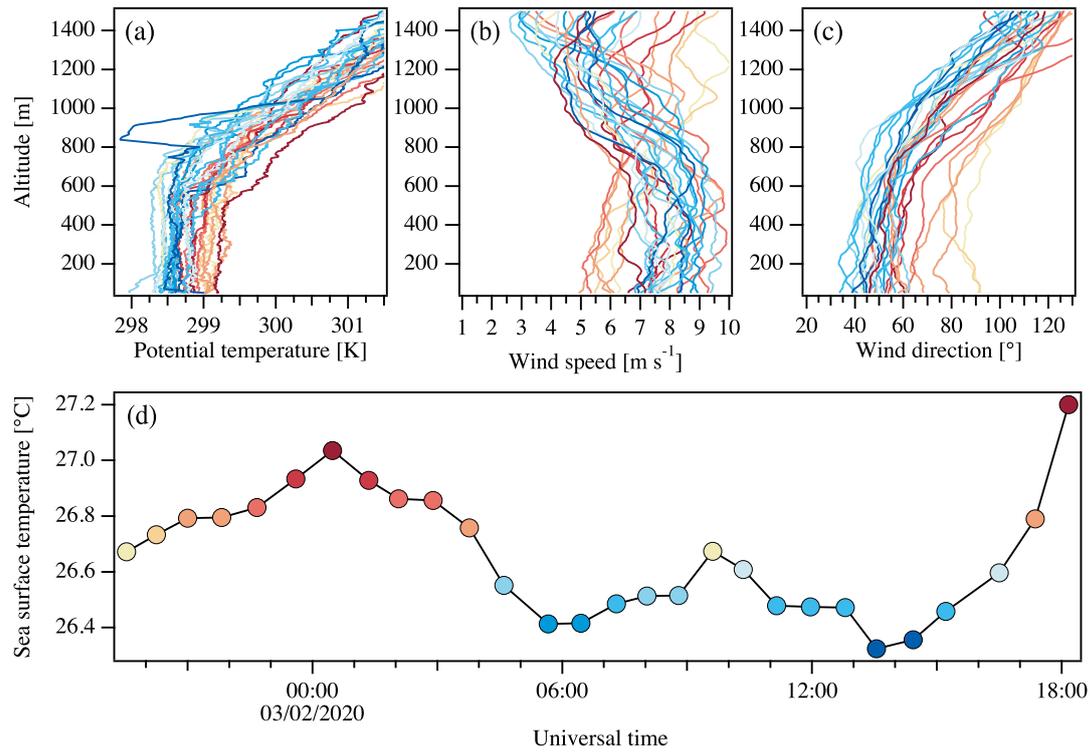


Figure F.15: Vertical profiles (50–1500 m) from MeteoModem M10 sondes launched during the second targeted intensive radiosonde period (Figure A1b) for (a) potential temperature, horizontal wind (b) speed and (c) direction, and (d) the corresponding SST time series from the *Atalante* TSG with each circle corresponding to a MeteoModem launch. Colors are indicative of the SST ($^{\circ}\text{C}$) at the time of each launch. Vertical profiles are built from Level-0 raw measurements.

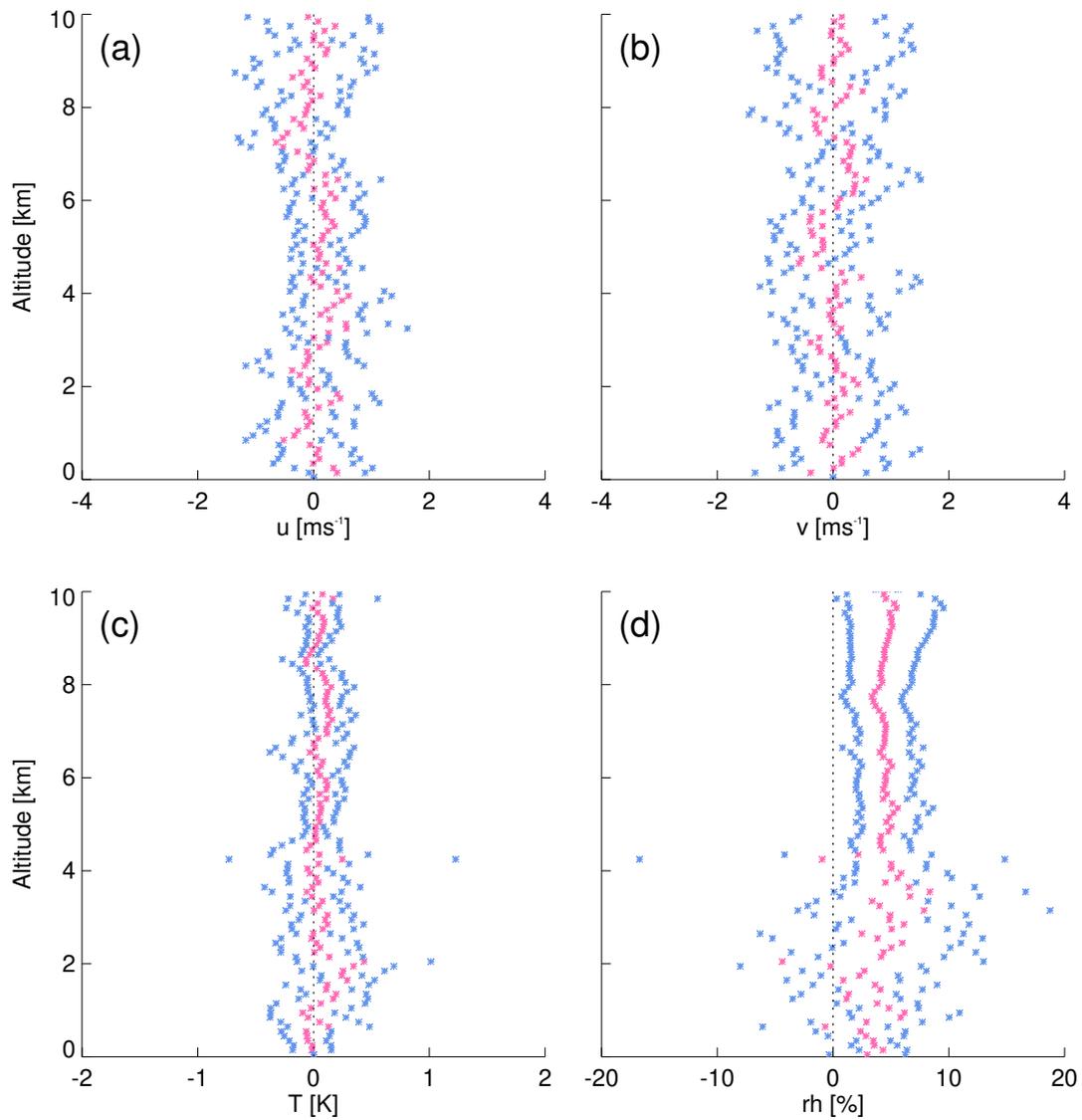


Figure F.16: For *Atalante* soundings launched within ± 25 min, the mean difference MeteoModem-Vaisala (pink) and ± 1 standard deviation (blue) computed on 8 difference profiles with a vertical resolution of 100 m. Shown are difference profiles for (a) zonal wind, (b) meridional wind, (c) temperature, and (d) relative humidity.

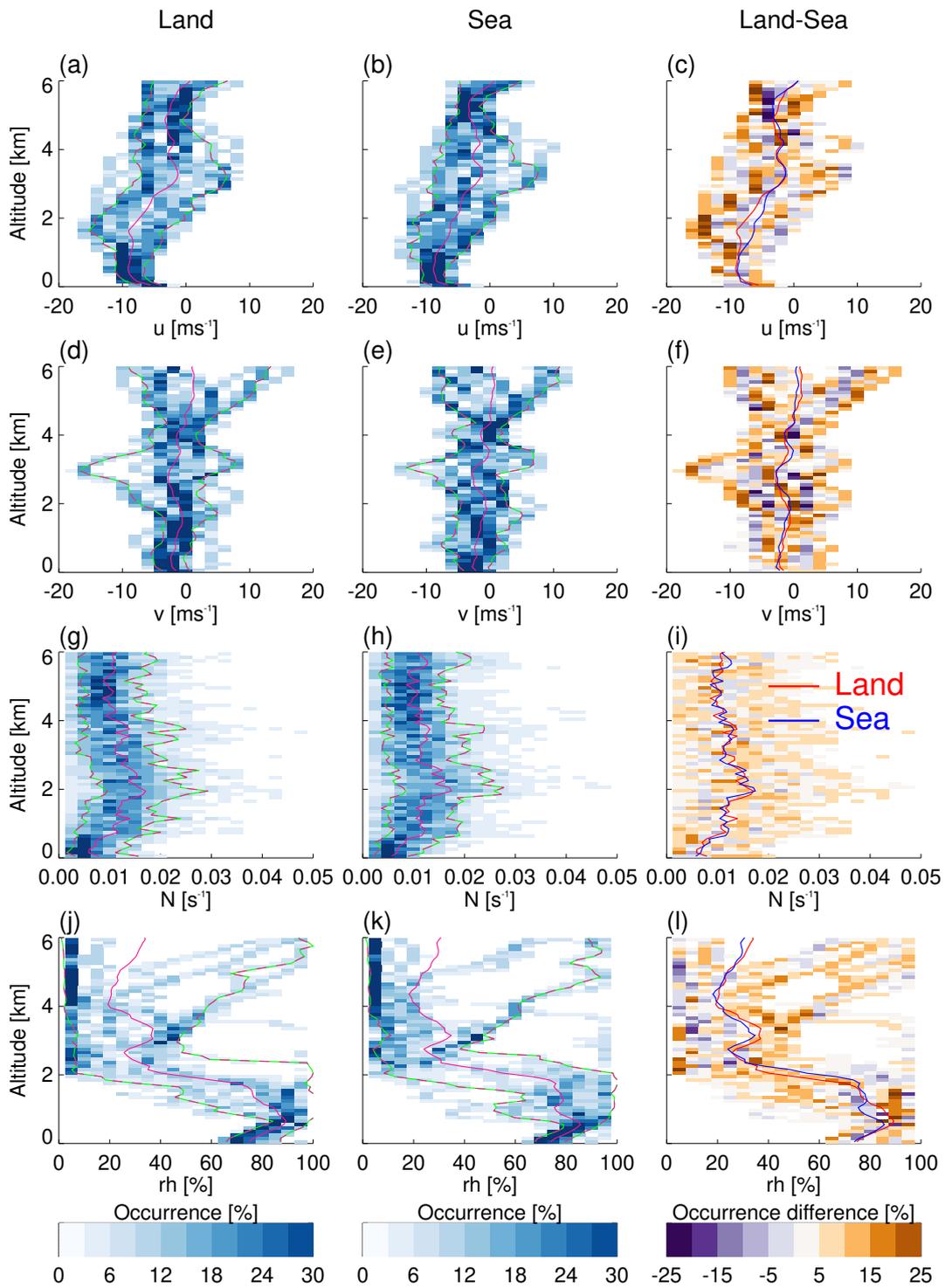


Figure F.17: As Fig. F.13, but instead of comparing different regions, we here compare ascending soundings launched from BCO with ascending soundings launched within ± 90 min from nearby ships (within 1° longitude to the east and $\pm 1^\circ$ latitude of BCO, resulting in 12 matching soundings). Altitude bins are 100 m deep and there are 20 bins on the x -axis.

MESO-SCALE PATTERNS OF SHALLOW CONVECTION DURING EUREC⁴A

The work in this appendix is intended for publication as:

Schulz, Hauke (in preparation). Meso-scale patterns of shallow convection during EUREC₄A

The contributions of the authors to this publication are as follows:

This is a single author manuscript with no other contributions except the participants of the international remote classification event who contributed to the dataset by labeling meso-scale patterns of shallow convection.

Meso-scale patterns of shallow convection during EUREC⁴A

Hauke Schulz¹

¹Max Planck Institute for Meteorology

ABSTRACT

Shallow trade-wind convection can occur in patterns of meso-scale extent. Their origin, formation process and radiative impacts are currently studied in field and model experiments. Here we give an overview about the meso-scale organization during the EUREC⁴A field campaign in early 2020 that has been manually identified by 51 researchers based on infrared and visible satellite imagery as well as simulation output.

Because the four pattern categories defined by Stevens et al. (2020) are rather subjective, the common perception of the scientific community can be derived from this dataset. With the provided post-processed data, it is shown that the dataset is able to serve as a reference for other studies to classify measurements taken during the EUREC⁴A field campaign.

The full dataset including postprocessed datasets for easier usage are freely available at the Zenodo archive at <http://doi.org/10.5281/zenodo.3763414>.

PLAIN LANGUAGE SUMMARY

Clouds are often clustered. One of the larger clusters are squall lines or even the Inter-tropical convergence zone. Also smaller clouds, like shallow trade wind cumuli, are often seen clustered on a scale of several 100 km. How this clustering is happening is not yet well understood and has been a core question of the recent field campaign EUREC⁴A. This dataset provides the manual pattern classifications to this campaign period.

G.1 DATASET DETAILS

- Identifier: <https://doi.org/10.5281/zenodo.3888876>
- Creator: Hauke Schulz
- Title: Manual classifications of meso-scale organization during EUREC⁴A
- Publisher: Zenodo
- Publication year: 2020
- Resource type: Dataset

G.2 INTRODUCTION

The understanding of the meso-scale patterns of shallow convection is still sparse. However, the ubiquity of these clouds and their reoccurring structure, suggest that they

play an important factor in determining the radiative effects in the trade-wind regime (Bony et al., 2020). The EUREC⁴A field campaign addresses among other questions, the question on how the organization of shallow convection is happening. Prior to the campaign, studies concentrated on the classification of meso-scale patterns based on satellite images by visual inspection (Stevens et al., 2020), rule-based algorithms (Bony et al., 2020) or trained neural networks (Rasp et al., 2020). These efforts show that the categorisation of the meso-scale patterns is an elementary part in order to gain further knowledge about these cloud structures. To address this need, the dataset presented here contains the manual classifications of 51 scientists from over 10 institutes who were involved in the EUREC⁴A field campaign in January - February 2020 and participated in a joint online classification event. The dataset presented here can therefore be seen as a baseline dataset for the meso-scale patterns of shallow convection in the downwind trades. It reflects the community consent of these visually defined patterns and allows to put the participating platforms and there measurements taken during EUREC⁴A in the meso-scale context.

G.3 DATA DESCRIPTION AND DEVELOPMENT

The manual classifications were gathered through the online platform zooniverse.org which has already been successfully used in an earlier project by Rasp et al. (2020). The platform makes it possible to crowd-source labels for e.g. machine learning projects and define tasks that need to be completed.

For this dataset, we defined three workflows. Two workflows are based on satellite observations in the visible and infrared channels, respectively, while another workflow is based on a storm-resolving simulation covering the EUREC⁴A time period. Details of this ICON simulation can be found in Schulz and Stevens (in preparation).

To visualize the output, we calculated a pseudo-albedo α by the following approximation:

$$\tau = 0.19 \cdot (LWP)^{\frac{5}{6}} \cdot (N^{1/3}) \quad (\text{G.1})$$

$$\alpha = \frac{\tau}{6.8 + \tau}, \quad (\text{G.2})$$

where LWP is the liquid water vapor path and N an assumed cloud droplet number density of 70 cm^{-3} .

The workflows are further described in Tab. G.1 together with an example for each of these visualized in Fig. G.2.

On March 24, 2020 the international, virtual classification hackathon has been hosted with 51 scientists from over 10 institutes participating to collect the pattern classifications. For a full day the participants classified patterns of shallow trade-wind convection by drawing rectangles around the four common types: *Sugar*, *Gravel*, *Flowers*, *Fish* (Stevens et al., 2020).

In the end, over 12.500 labels were gathered and accumulated intentionally on the observational workflows (see Fig. G.3) as it quickly turned out that the identification of the patterns in the model simulation was too demanding. The features had too little similarity with those found in nature. Comparing the amount of labels that have been

created for each class and workflow (Fig. G.3), *Sugar* has been classified least in the simulation workflow. It becomes clear that the *Sugar* type clouds were the hardest to identify or not present in the simulation output. The largest feature, *Fish*, however, has been identified more often. This supports the assumption that larger features are better represented in storm-resolving simulations.

Because all users are known to the authors and were trained before the classification through an online presentation to get familiar with the labeling interface on zooniverse.org but also to refresh the different meso-scale cloud pattern categories, it can be assumed that the labels are of high quality. Compared to Rasp et al. (2020) where the focus has been to classify as many diverse cloud scenes as possible to capture the variability and thus serve as a better machine learning dataset, the aim for this dataset is to create a baseline classification that participating scientists agree on and can directly use in further studies. Therefore, the temporal frequency has been increased from daily cloud scenes to 2-hourly cloud scenes to reflect also the changes on the sub-daily scale that have been identified by Vial et al. (2021). Due to this design difference, cloud scenes are classified on average by 15 participants in case of the visible workflow instead of just about 3 as in Rasp et al. (2020).

For all three workflows we provide post-processed datasets which are illustrated in Fig. G.1 and described as follows:

- *Level 0*
The *Level 1* data consists of the raw data output and originates from the zooniverse platform. It consists of CSV files that contain entries for each classification including technical details like the the time spend on to draw a specific label.
- *Level 1*
The *Level 1* dataset is further processed. It contains each label as a separate entry and contains information about the classified object, the user and the geographical and Cartesian coordinates of the label. This data is saved as a netCDF file.
- *Level 2*
For the *Level 2* dataset, the data is merged by `classification_id`. The `classification_id` is a unique identifier of a classification, while a classification refers here to the process of labeling a single image by a single user. The user might use several labels of the same or different kind to completely classify a scene. This process eliminates overlaps of same-user classifications for each pattern and turns the data into masks, rather than coordinates (see Fig. G.1).
- *Level 3*
To ease working with the dataset, daily frequency contributions like shown in Fig. G.2 are calculated and saved as *Level 3* data for each workflow. The percentage of agreement (p) among users on a specific pattern on each location is calculated as follows:

$$p_{\text{pattern}}(i, j) = \frac{\sum_0^U c(i, j)}{U}, \quad (\text{G.3})$$

where U is the number of users that have seen the particular image, c the classification mask from the *Level 2* data and i, j the geographic coordinates. Because the

labels of users that attributed several classes to one pixel are not removed, $\sum p$ can be greater than 100%.

This dataset is shown in the appendix for each day and workflow to give an impression of the dataset and in particular the meso-scale patterns present during the EUREC⁴A field campaign.

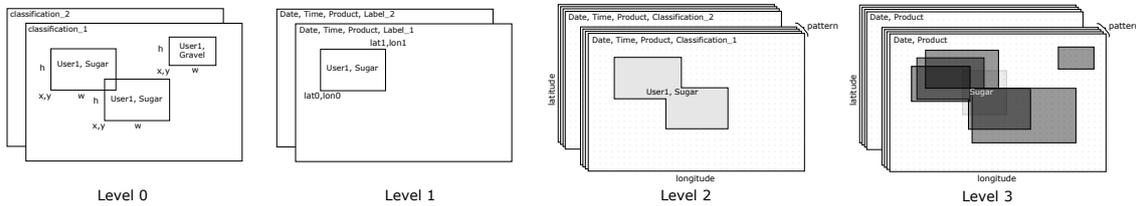


Figure G.1: Overview of processing levels of the dataset.

G.4 DATASET ACCESS

The data including raw data, processed data and the source code with examples on how to process the data are freely available at zenodo.org.

Schulz, Hauke. (2020). EUREC⁴A meso-scale cloud classifications (Version v.0.1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.3888876>

G.5 POTENTIAL DATASET USE AND REUSE

The EUREC⁴A field campaign has been an international study with a wide range of research platforms and associated studies Stevens et al. (2021). This dataset does not only cover the core area of the experiment, but also the wider area and time period. While the participating research airplanes and drones were mostly staying in the trade-winds, some ships departed as far south as 6.5°N.

This dataset gives the opportunity to study all these measurements in the context of the meso-scale patterns observed in the downwind trades. Due to the high subjectivity of these meso-scale cloud pattern definitions, it is of particular importance to discuss results based on a community consent to keep studies comparable. This dataset can serve as such a reference for the period of the EUREC⁴A field campaign.

G.6 ACKNOWLEDGMENTS

The author thanks the participants of the international remote classification event.

G.7 APPENDIX

G.7.1 Daily classification overview

Table G.1: Description of data sources used to create the images of the classification workflows.

	EUREC ⁴ A (VIS)		EUREC ⁴ A (IR)		EUREC ⁴ A (ICON)
	MODIS (TERRA/AQUA)	ABI (GOES16)	ABI (GOES16)	ABI (GOES16)	
Domain	5-20 N; 62 - 40 W		5 - 20 N; 62 - 44 W		
Period	07.01.2020 - 22.02.2020		16.01.2020 - 20.02.2020		
Resolution	~1 km		~2 km		
Data source (shown)	2-hourly, 12-20 UTC Corrected reflectance	2-hourly Channel 02 (red)	2-hourly Channel 13 (IR)	2-hourly pseudo-albedo	
Number of scenes	94	234	562	425	Data shown after 24 hours spin-up
Remarks					

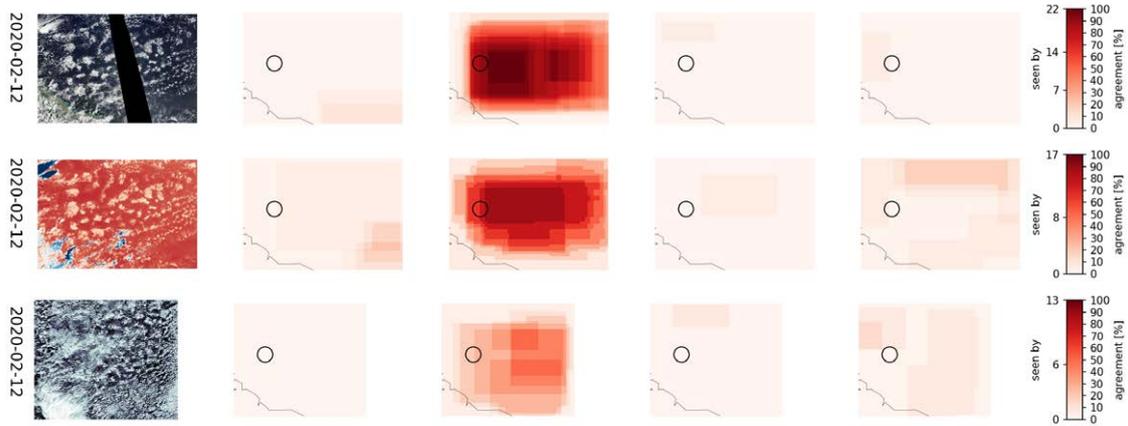


Figure G.2: Manual classification examples for the three workflows (top to bottom: visible, infrared, simulation). The labels for each pattern (from left to right: *Sugar*, *Flowers*, *Fish*, *Gravel*) are shown next to the labeled image.

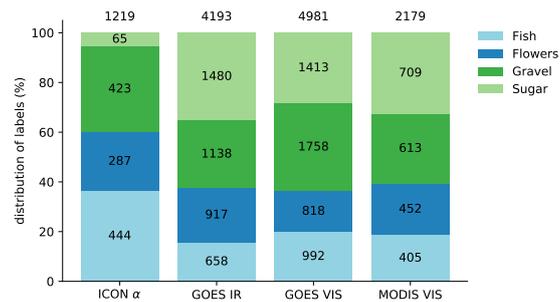


Figure G.3: Distribution of labels by data source. The relative distribution is indicated by the bars, while the absolute number of labels is indicated as text. The total labels per workflow are on top of each bar.

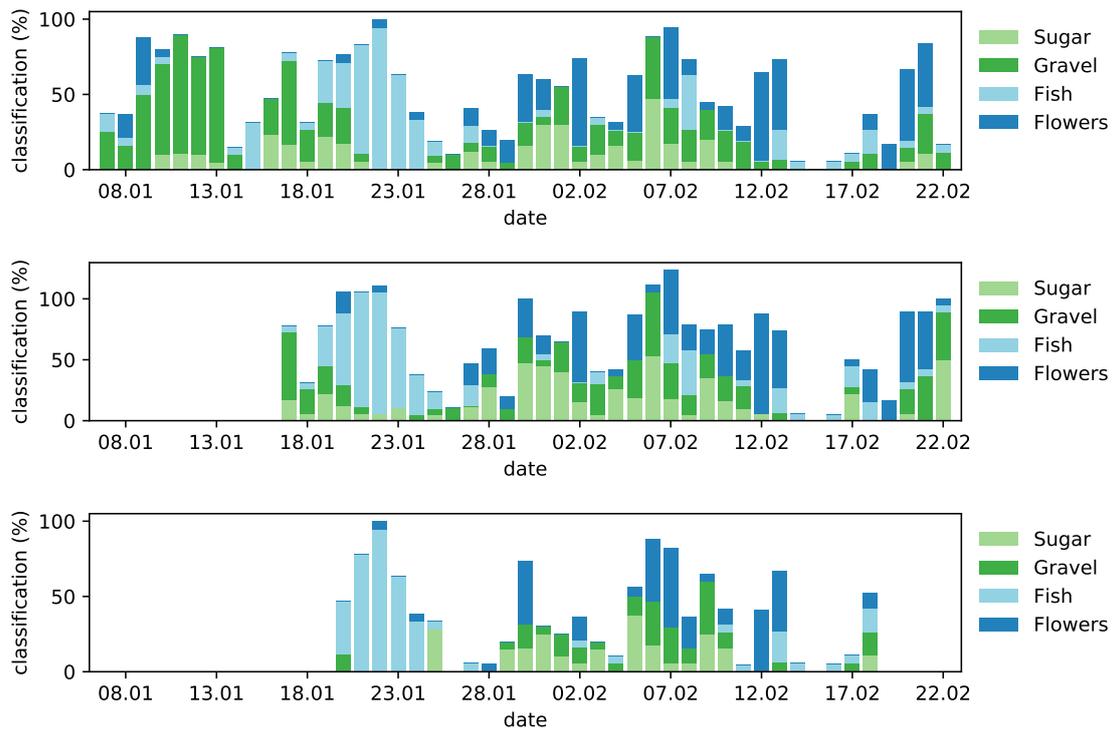


Figure G.4: Exemplary use cases: meso-scale setting of research platforms during EUREC⁴A (top to bottom: BCO, RV Meteor, RV Atalante)

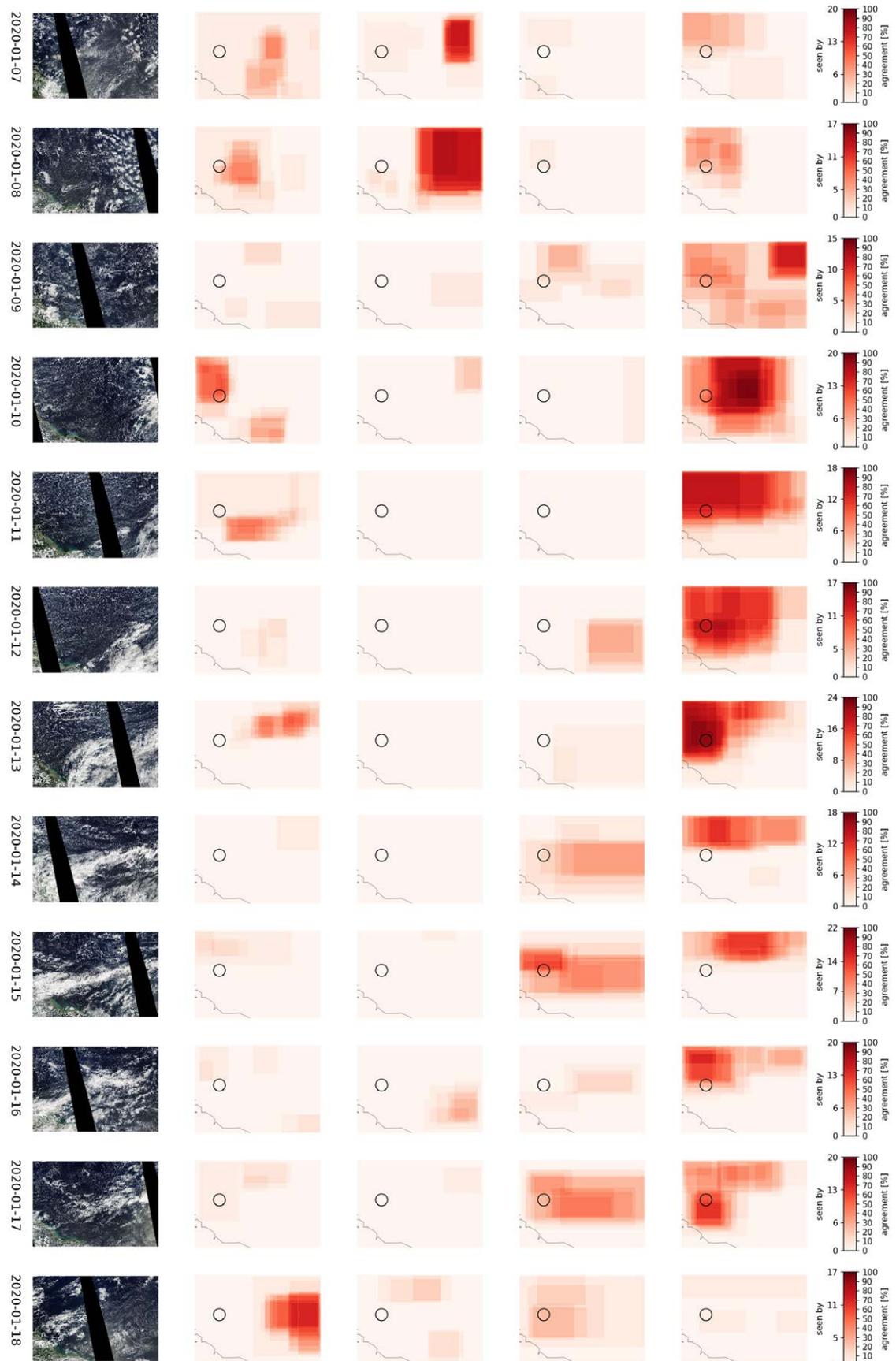


Figure G.5: Heatmaps of manual classifications based on MODIS (Aqua and Terra) visible imagery. Left to right: Visible imagery during Aqua overpass, User agreement on *Sugar*, *Flowers*, *Fish* and *Gravel*. Circle indicates HALO flight circle.

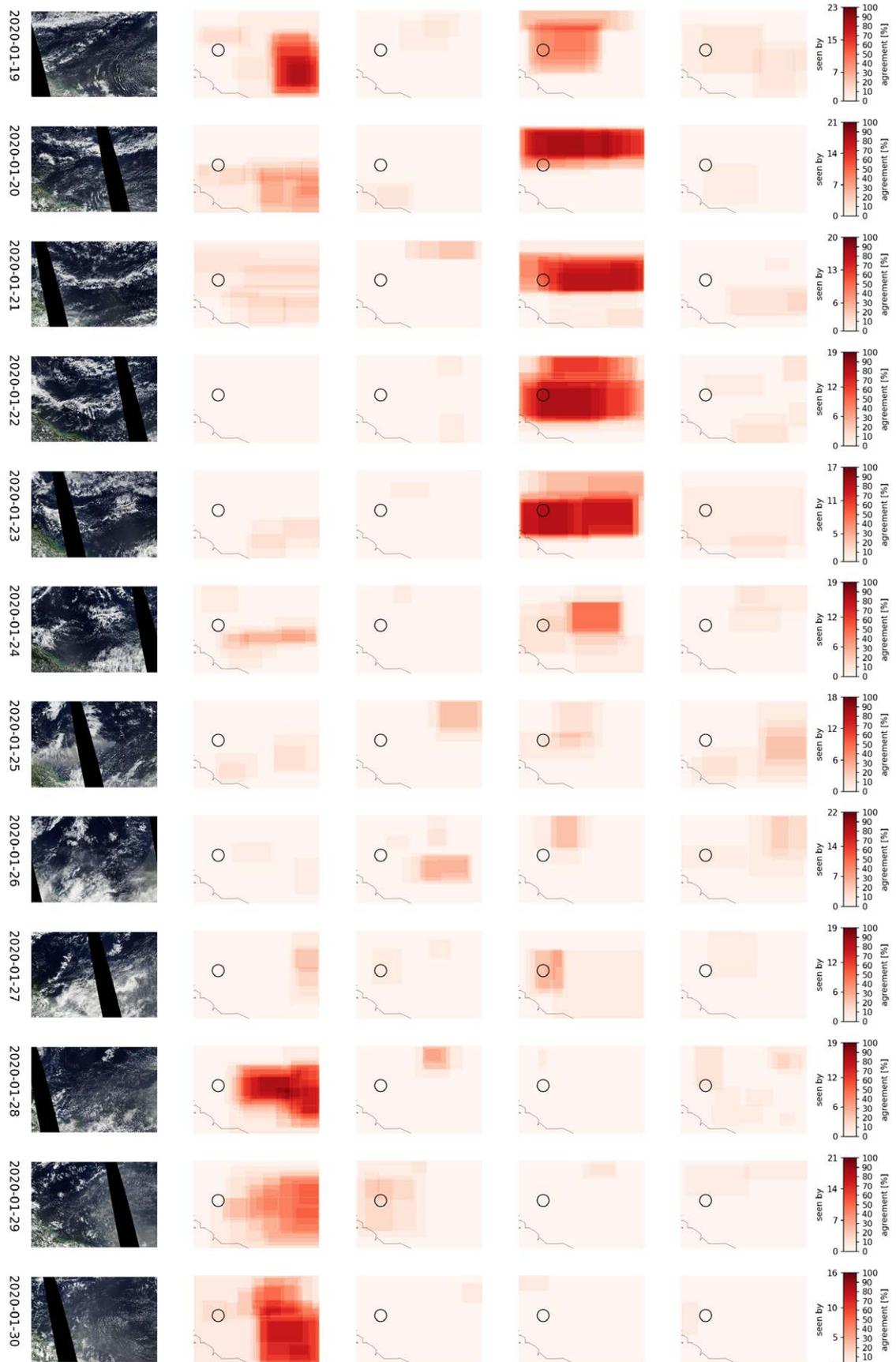


Figure G.6: continuation of Fig. G.5

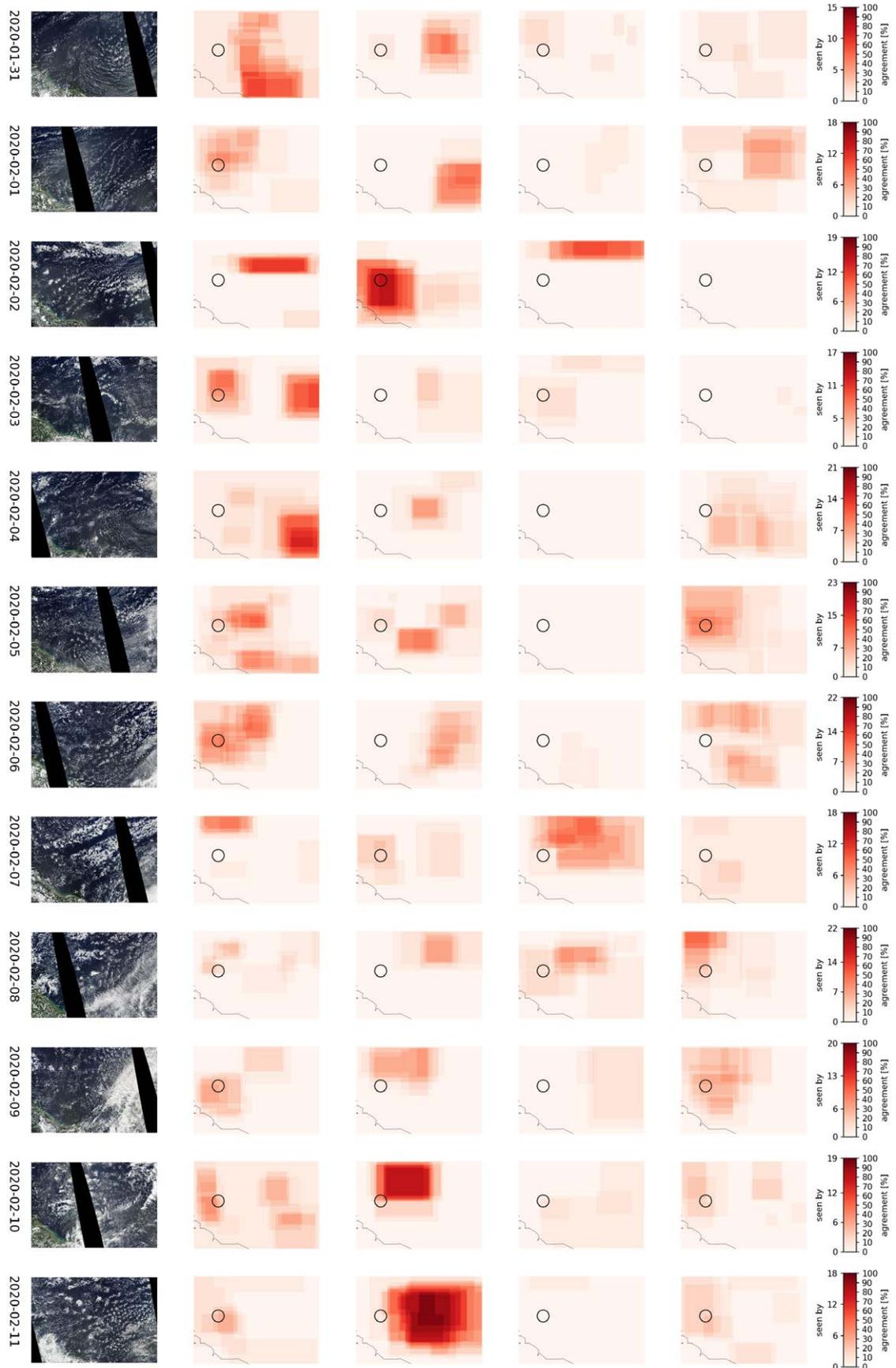


Figure G.7: continuation of Fig. G.5

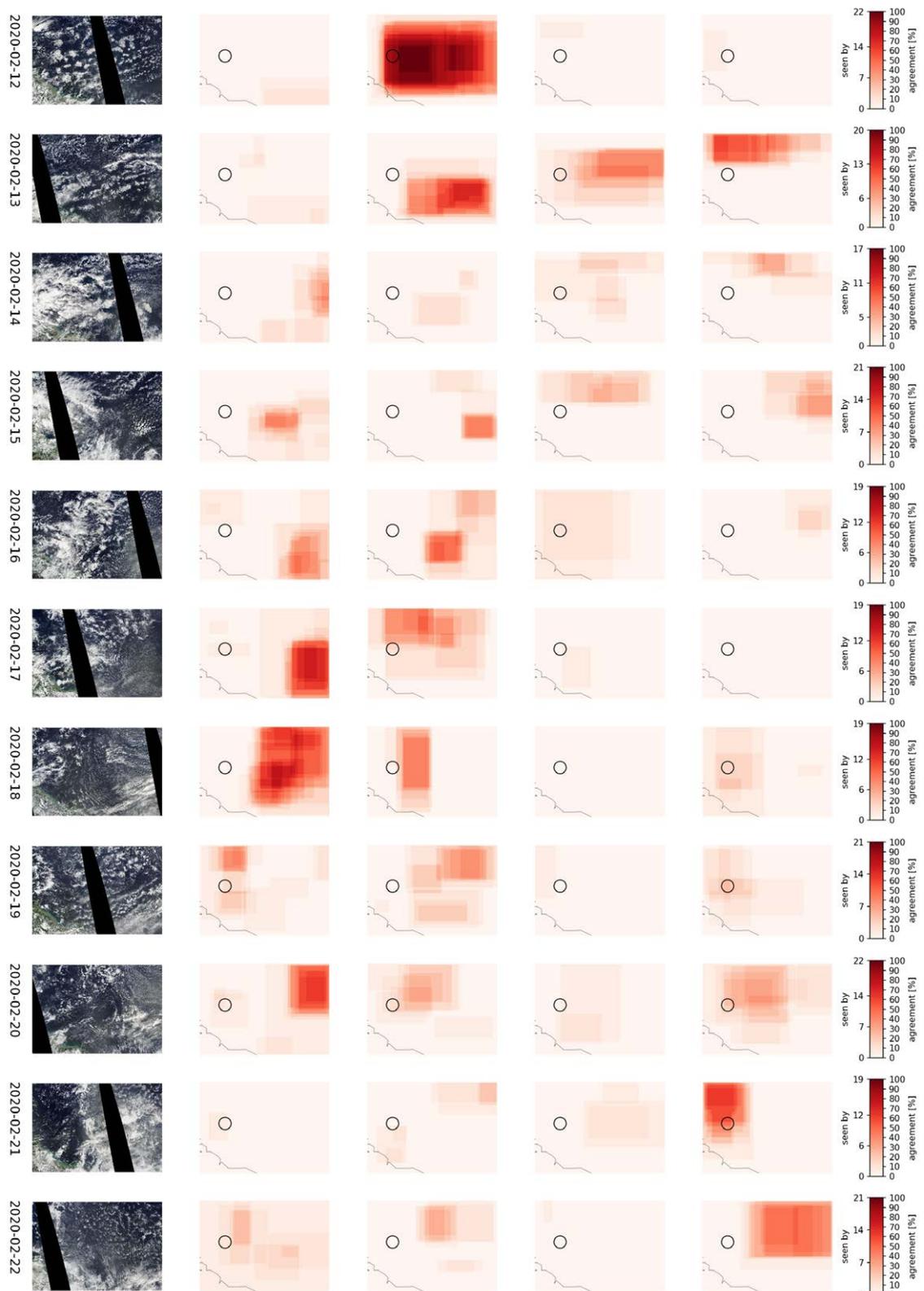


Figure G.8: continuation of Fig. G.5

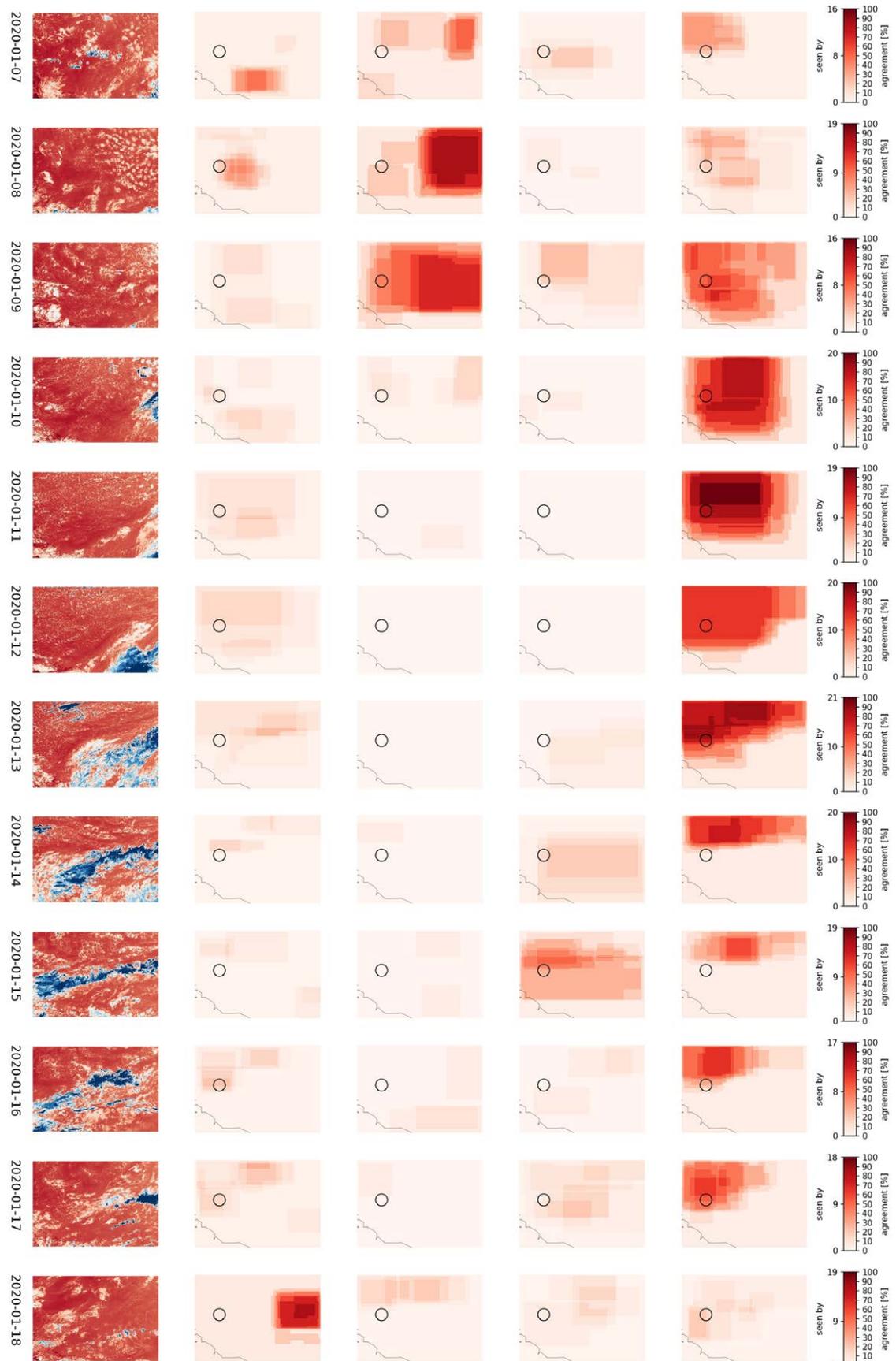


Figure G.9: Images at 16 o'clock. (except 11.02.2020: 17 o'clock)

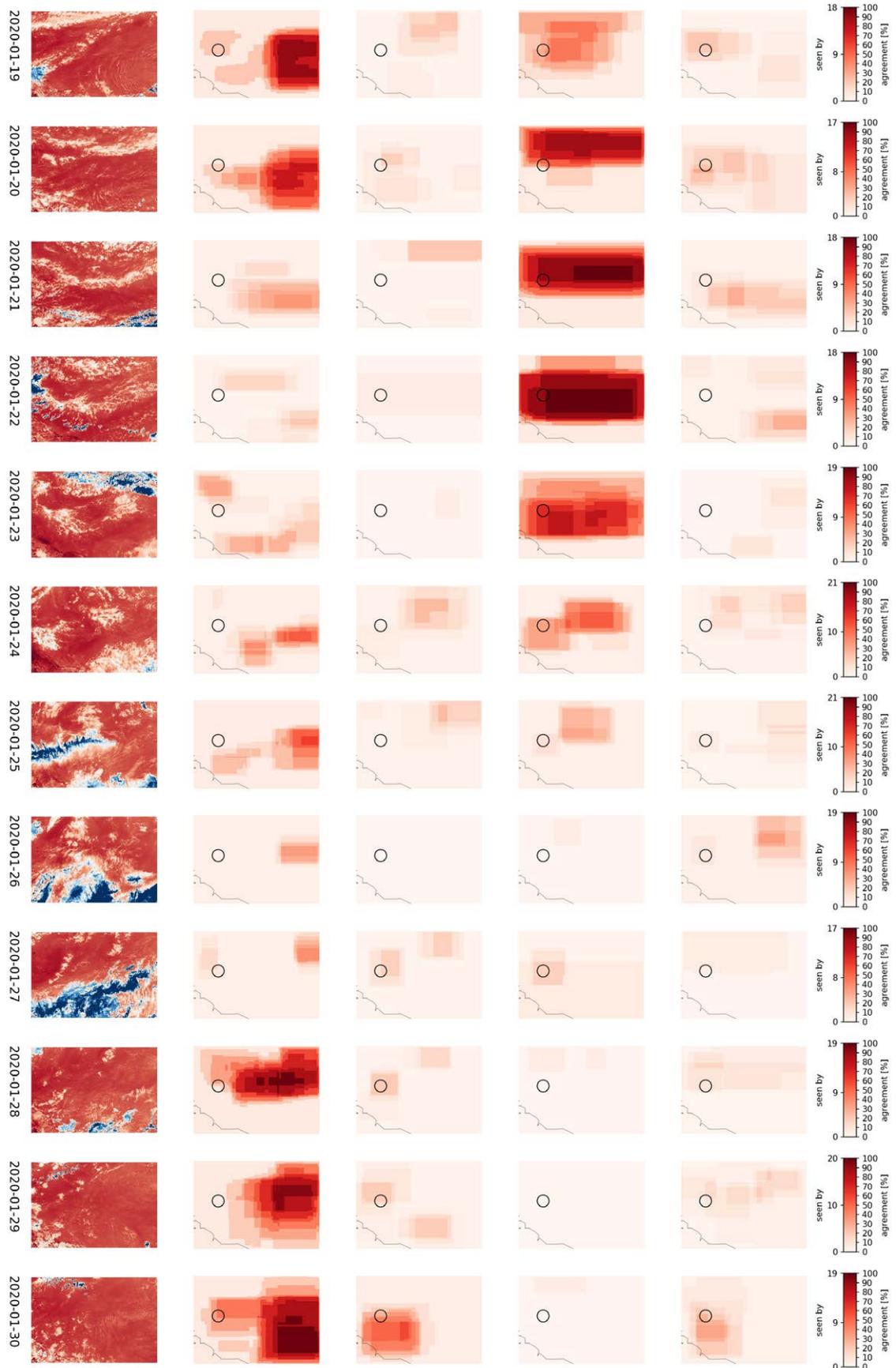


Figure G.10: continuation of Fig. G.9

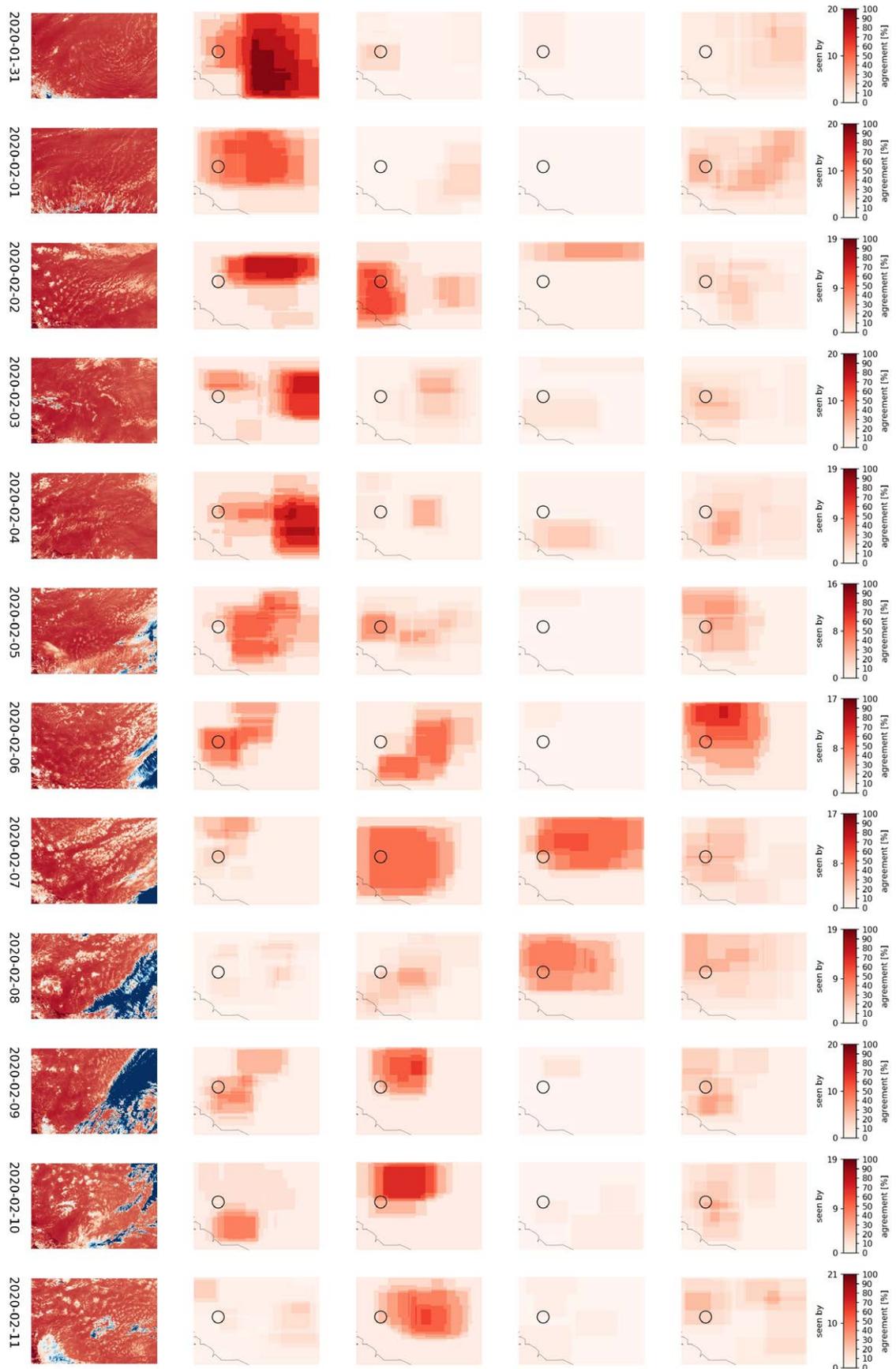


Figure G.11: continuation of Fig. G.9

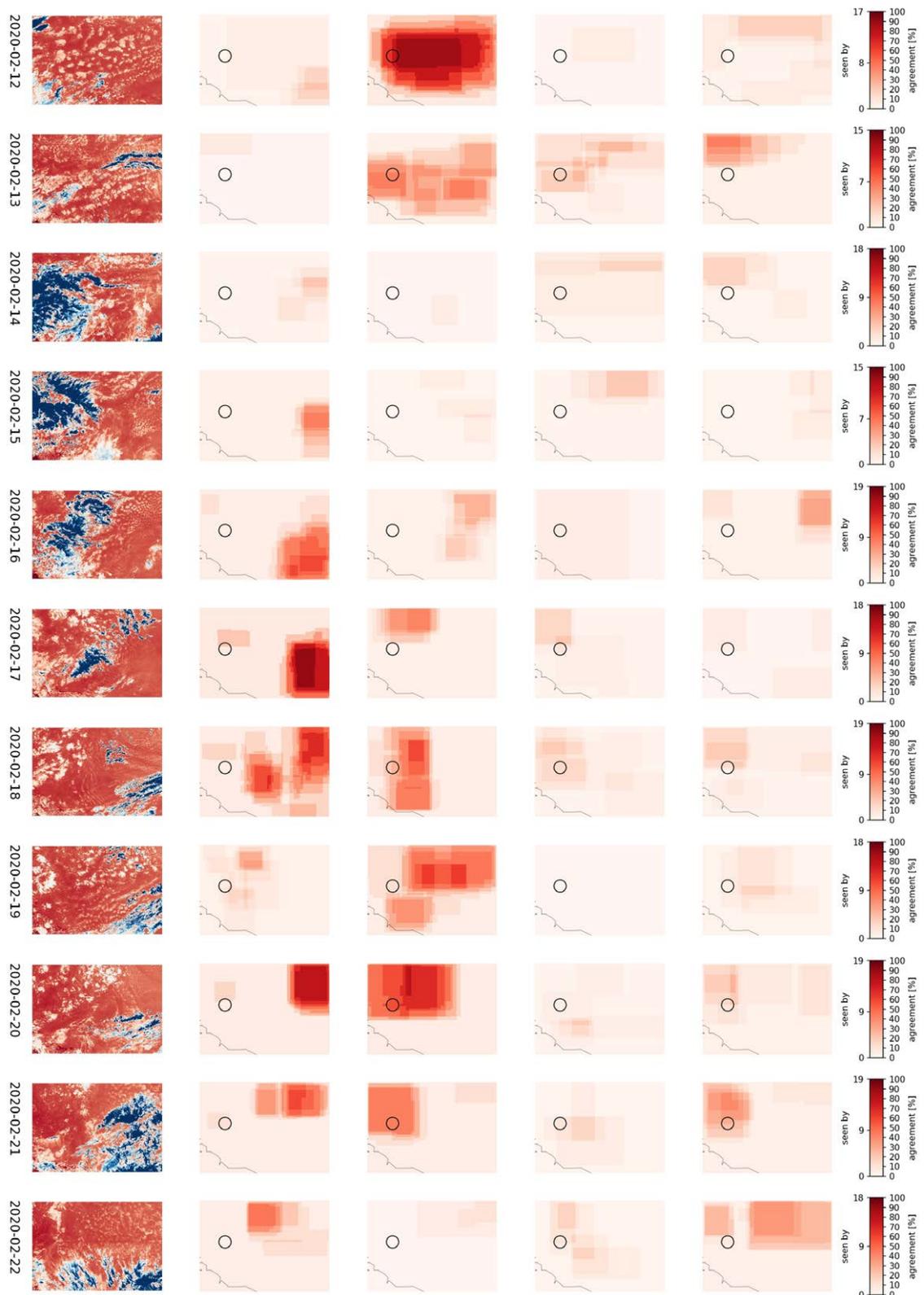


Figure G.12: continuation of Fig. G.9

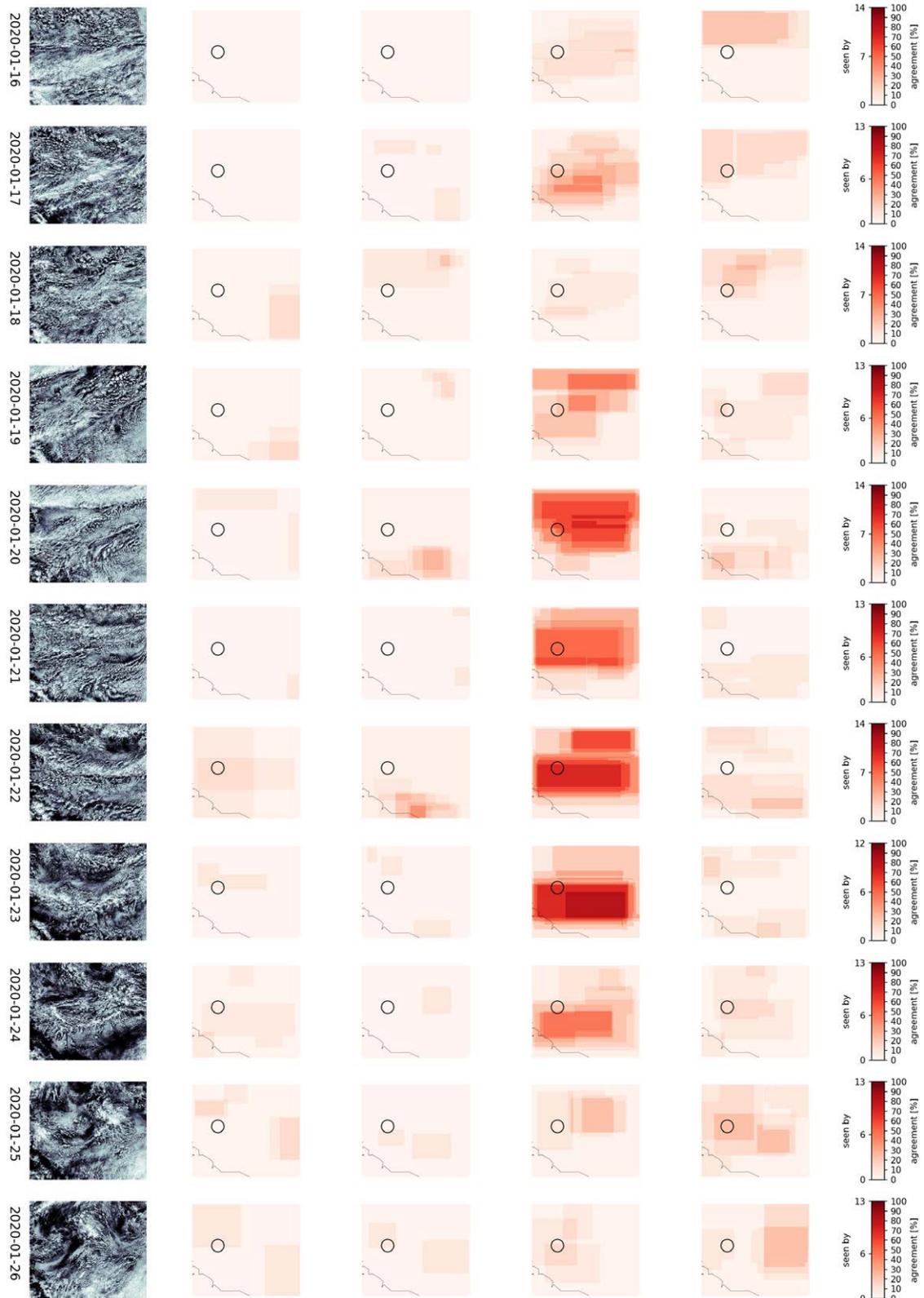


Figure G.13

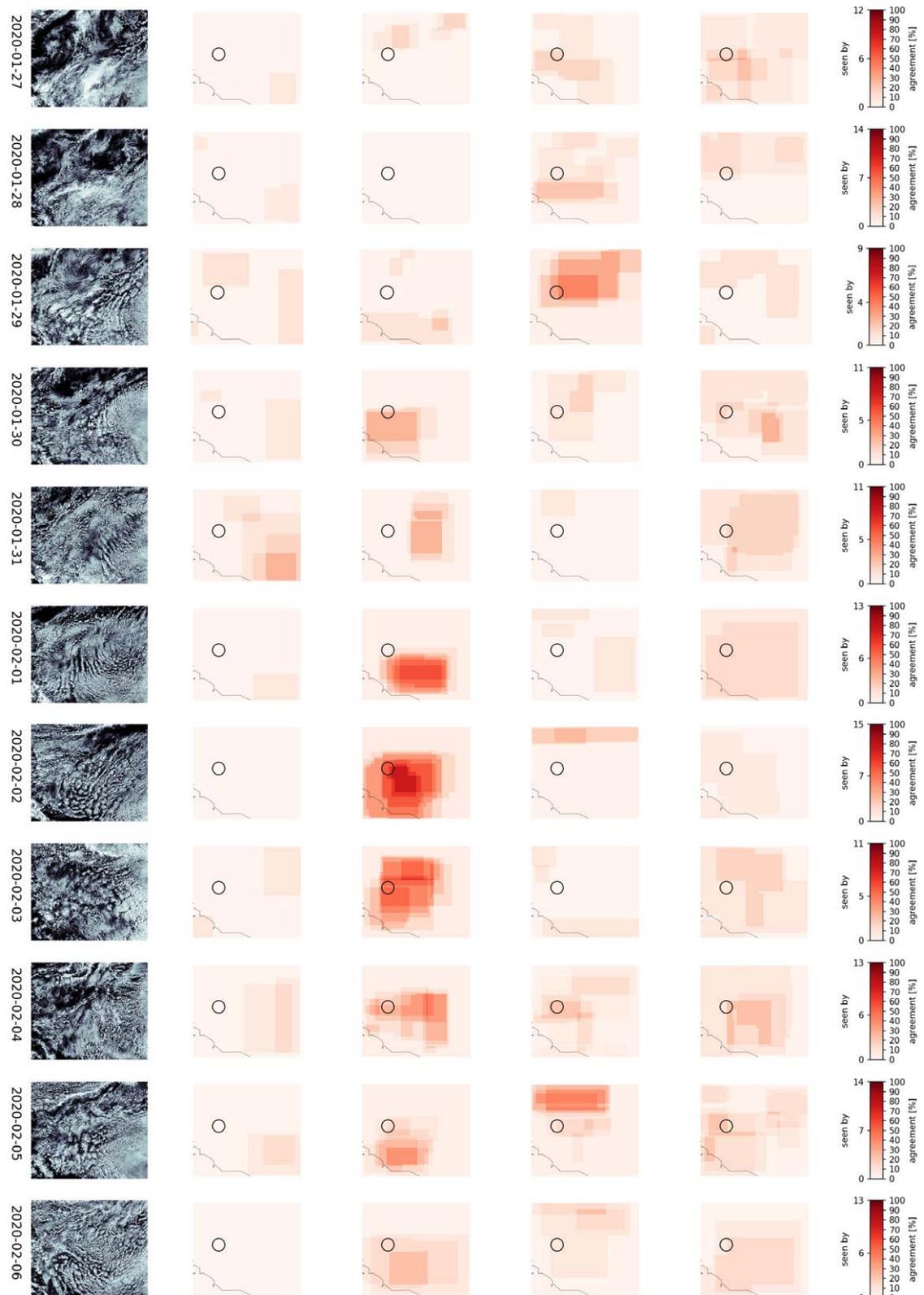


Figure G.14: continuation of Fig. G.13

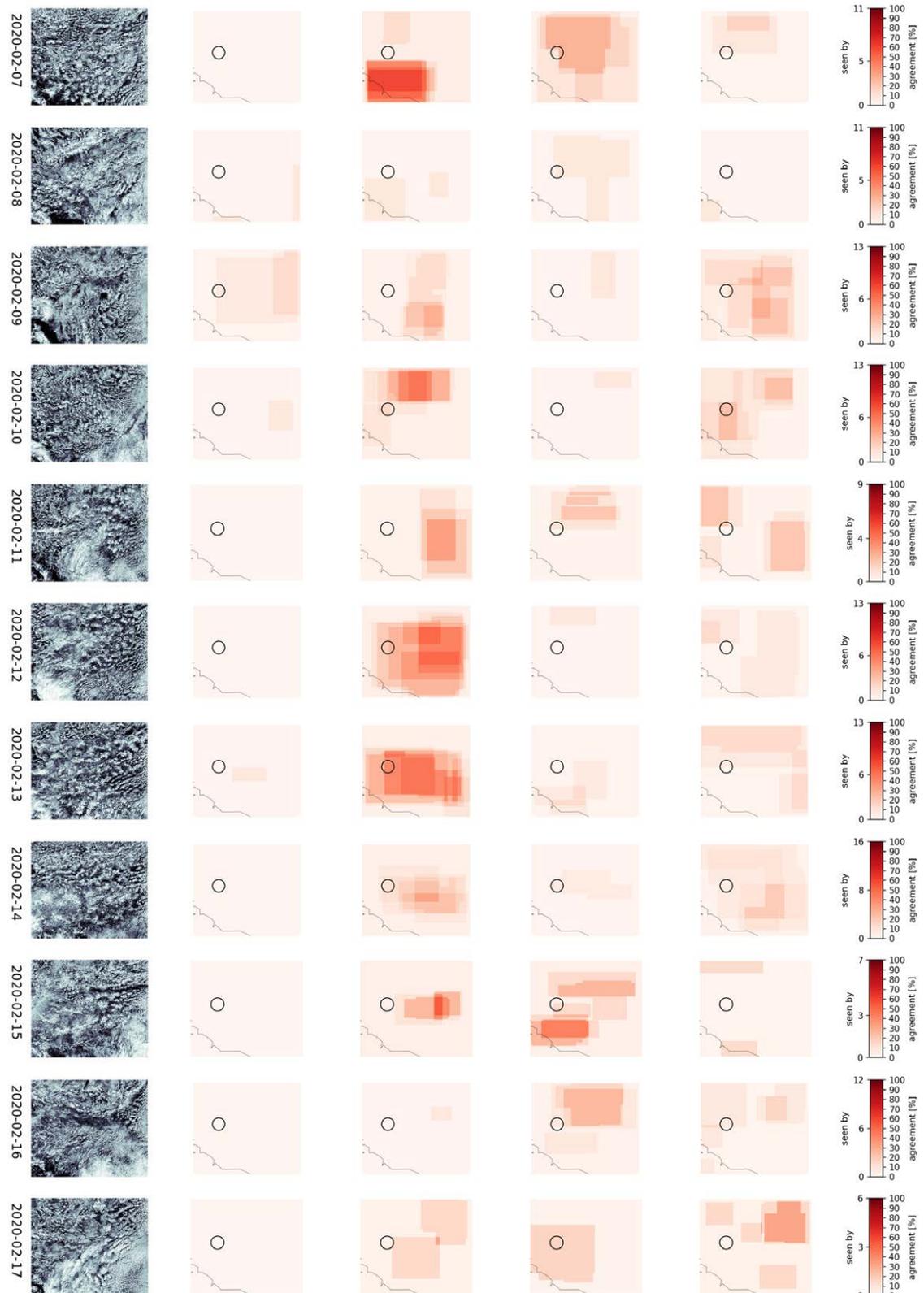


Figure G.15: continuation of Fig. G.13

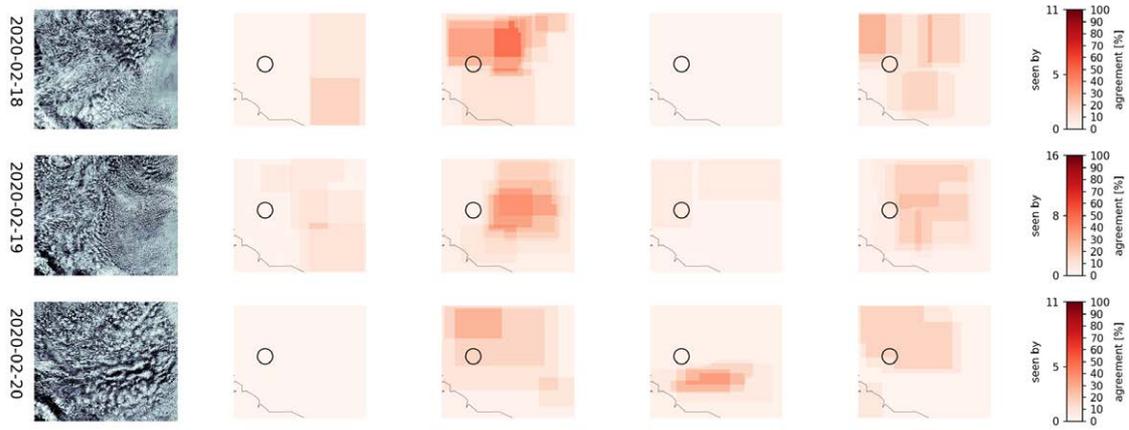


Figure G.16: continuation of Fig. G.13

BIBLIOGRAPHY

- Ablain, M., J.F. Legeais, P. Prandi, M. Marcos, L. Fenoglio-Marc, H.B. Dieng, J. Benveniste, and A. Cazenave (2017). Satellite altimetry-based sea level at global and regional scales. *Surv. Geophys.* 38, pp. 7–31. DOI: [10.1007/s10712-016-9389-8](https://doi.org/10.1007/s10712-016-9389-8).
- Ablain, Y., H. Tomita, M.F. Cronin, and N. A. Bond (2014). Atmospheric pressure response to mesoscale sea surface temperature variations in the Kuroshio Extension region: In situ evidence. *J. Geophys. Res. Atmos.* 119, pp. 8015–8031. DOI: [10.1002/2013JD021126](https://doi.org/10.1002/2013JD021126).
- Aemisegger, F., R. Vogel, P. Graf, F. Dahinden, L. Villiger, F. Jansen, S. Bony, B. Stevens, and H. Wernli (2021). How Rossby wave breaking modulates the water cycle in the North Atlantic trade wind region. *Weather and Climate Dynamics* 2.1, pp. 281–309. DOI: [10.5194/wcd-2-281-2021](https://doi.org/10.5194/wcd-2-281-2021). URL: <https://wcd.copernicus.org/articles/2/281/2021/>.
- Agee, E. M. (Jan. 1, 1987). Mesoscale Cellular Convection over the Oceans. *Dynamics of Atmospheres and Oceans* 10.4, pp. 317–341. DOI: [10.1016/0377-0265\(87\)90023-6](https://doi.org/10.1016/0377-0265(87)90023-6).
- Albrecht, Bruce A. (1993). Effects of precipitation on the thermodynamic structure of the trade wind boundary layer. *J. Geophys. Res. Atm.* 98.D4, pp. 7327–7337. DOI: [10.1029/93JD00027](https://doi.org/10.1029/93JD00027).
- Albrecht, Bruce A., Alan K. Betts, Wayne H. Schubert, and Stephen K. Cox (1979). Model of the thermodynamic structure of the trade-wind boundary layer: Part I. Theoretical formulation and sensitivity tests. *J. Atmos. Sci.* 36.1, pp. 73–89. DOI: [10.1175/1520-0469\(1979\)036<0073:MOTTSO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<0073:MOTTSO>2.0.CO;2).
- Albrecht, Bruce A., Christopher S. Bretherton, Doug Johnson, Wayne H. Scubert, and A. Shelby Frisch (1995). The Atlantic Stratocumulus Transition Experiment—ASTEX. *Bull. Amer. Meteor. Soc.* 76.6, pp. 889–904. DOI: [10.1175/1520-0477\(1995\)076<0889:TASTE>2.0.CO;2](https://doi.org/10.1175/1520-0477(1995)076<0889:TASTE>2.0.CO;2).
- Albright, A. L., B. Fildier, L. Touzé-Peiffer, R. Pincus, J. Vial, and C. Muller (2021). Atmospheric radiative profiles during EUREC⁴A. *Earth System Science Data Discussions* preprint, pp. 1–17. DOI: [10.5194/essd-2020-269](https://doi.org/10.5194/essd-2020-269).
- Arakawa, Akio and Wayne Howard Schubert (Apr. 1974). Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.* 31.3, pp. 674–701. DOI: [10.1175/1520-0469\(1974\)031<0674:IOACCE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2).
- Atkinson, B W and J Wu Zhang (1996). Mesoscale Shallow Convection in the Atmosphere. *Reviews of Geophysics*. DOI: [10.1029/96RG02623](https://doi.org/10.1029/96RG02623). URL: [https://agupubs-onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de/doi/pdf/10.1029/96RG02623](https://agupubs.onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de/doi/pdf/10.1029/96RG02623).
- Atlas, R. L., C. S. Bretherton, P. N. Blossey, A. Gettelman, C. Bardeen, Pu Lin, and Yi Ming (2020). How Well Do Large-Eddy Simulations and Global Climate Models Represent Observed Boundary Layer Structures and Low Clouds Over the Summertime Southern Ocean? *Journal of Advances in Modeling Earth Systems* 12.11, e2020MS002205. DOI: [10.1029/2020MS002205](https://doi.org/10.1029/2020MS002205).
- Augstein, Ernst, Herbert Riehl, Feodor Ostapoff, and Volker Wagner (1973). Mass and energy transports in an undisturbed Atlantic trade-wind flow. *Mon. Wea. Rev.* 101.2, pp. 101–111. DOI: [10.1175/1520-0493\(1973\)101<0101:MAETIA>2.3.CO;2](https://doi.org/10.1175/1520-0493(1973)101<0101:MAETIA>2.3.CO;2).

- Baker, Marcia (1993). Trade Cumulus Observations. In: *The representation of cumulus convection in numerical models*. Ed. by Kerry A. Emanuel and David J. Raymond. Meteorological Monographs. Boston, MA: Amer. Meteor. Soc., pp. 29–37. ISBN: 978-1-935704-13-3. DOI: [10.1007/978-1-935704-13-3_3](https://doi.org/10.1007/978-1-935704-13-3_3).
- Baldauf, Michael, Axel Seifert, Jochen Förstner, Detlev Majewski, Matthias Raschendorfer, and Thorsten Reinhardt (Dec. 1, 2011). Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Monthly Weather Review* 139.12, pp. 3887–3905. DOI: [10.1175/MWR-D-10-05013.1](https://doi.org/10.1175/MWR-D-10-05013.1).
- Barnes, Gary M. and Michael Garstang (Feb. 1, 1982). Subcloud Layer Energetics of Precipitating Convection. *Monthly Weather Review* 110.2, pp. 102–117. DOI: [10.1175/1520-0493\(1982\)110<0102:SLE0PC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0102:SLE0PC>2.0.CO;2).
- Bellenger, Hugo, YN Takayabu, T Ushiyama, and K Yoneyama (2010). Role of diurnal warm layers in the diurnal cycle of convection over the tropical Indian Ocean during MISO. *Monthly Weather Review* 138.6, pp. 2426–2433.
- Beucler, Tom, David Leutwyler, and Julia M. Windmiller (2020). Quantifying Convective Aggregation Using the Tropical Moist Margin’s Length. *Journal of Advances in Modeling Earth Systems* 12.10, e2020MS002092. DOI: [10.1029/2020MS002092](https://doi.org/10.1029/2020MS002092).
- Bolton, David (1980). The computation of equivalent potential temperature. *Mon. Wea. Rev.* 108.7, pp. 1046–1053. DOI: [10.1175/1520-0493\(1980\)108<1046:TCOEPT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2).
- Bony, S. and B. Stevens (2019). Measuring area-averaged vertical motions with dropsondes. *J. Atmos. Sci.* 76, pp. 767–783. DOI: [10.1175/JAS-D-18-0141.1](https://doi.org/10.1175/JAS-D-18-0141.1).
- Bony, Sandrine and Jean-Louis Dufresne (2005). Marine Boundary Layer Clouds at the Heart of Tropical Cloud Feedback Uncertainties in Climate Models. *Geophysical Research Letters* 32.20, n/a–n/a. DOI: [10.1029/2005GL023851](https://doi.org/10.1029/2005GL023851).
- Bony, Sandrine, Hauke Schulz, Jessica Vial, and Bjorn Stevens (2020). Sugar, Gravel, Fish, and Flowers: Dependence of Mesoscale Patterns of Trade-Wind Clouds on Environmental Conditions. *Geophysical Research Letters* 47.7, e2019GL085988. DOI: [10.1029/2019GL085988](https://doi.org/10.1029/2019GL085988).
- Bony, Sandrine, Bjorn Stevens, Dargan M. W. Frierson, Christian Jakob, Masa Kageyama, Robert Pincus, Theodore G. Shepherd, Steven C. Sherwood, A. Pier Siebesma, Adam H. Sobel, et al. (2015). Clouds, circulation and climate sensitivity. *Nature Geoscience* 8.4, pp. 261–268.
- Bony, Sandrine et al. (Nov. 2017). EUREC4A: A Field Campaign to Elucidate the Couplings Between Clouds, Convection and Circulation. *Surveys in Geophysics* 38.6, pp. 1529–1568. DOI: [10.1007/s10712-017-9428-0](https://doi.org/10.1007/s10712-017-9428-0). URL: <http://link.springer.com/10.1007/s10712-017-9428-0>.
- Bretherton C. S. and Blossey P. N. (Jan. 17, 2018). Understanding Mesoscale Aggregation of Shallow Cumulus Convection Using Large-Eddy Simulation. *Journal of Advances in Modeling Earth Systems* 9.8, pp. 2798–2821. DOI: [10.1002/2017MS000981](https://doi.org/10.1002/2017MS000981).
- Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov (Dec. 2005). An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST. *Journal of the Atmospheric Sciences* 62.12, pp. 4273–4292. DOI: [10.1175/JAS3614.1](https://doi.org/10.1175/JAS3614.1).
- Bretherton, Christopher S. (July 1, 1993). Understanding Albrecht’s Model of Trade Cumulus Cloud Fields. *Journal of the Atmospheric Sciences* 50.14, pp. 2264–2283. DOI: [10.1175/1520-0469\(1993\)050<2264:UAM0TC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<2264:UAM0TC>2.0.CO;2).

- Bretherton, Christopher S (Nov. 2015). Insights into low-latitude cloud feedbacks from high-resolution models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373.2054, p. 20140415.
- Brient, Florent and Sandrine Bony (2013). Interpretation of the positive low-cloud feedback predicted by a climate model under global warming. *Climate Dynamics* 40.9-10, pp. 2415–2431. DOI: [10.1007/s00382-011-1279-7](https://doi.org/10.1007/s00382-011-1279-7).
- Brill, Keith and Bruce Albrecht (1982). Diurnal variation of the trade-wind boundary layer. *Monthly Weather Review* 110.6, pp. 601–613.
- Brueck, Matthias, Louise Nuijens, and Bjorn Stevens (2015). On the seasonal and synoptic time-scale variability of the North Atlantic trade wind region and its low-level clouds. *Journal of the Atmospheric Sciences* 72.4, pp. 1428–1446.
- Bunker, Andrew F, B Haurwitz Bernhard, Malkus, Joanne Starr, and Stommel, Henry M (1949). Vertical Distribution of Temperature and Humidity over the Caribbean Sea. *PAPERS IN PHYSICAL OCEANOGRAPHY AND METEOROLOGY* 11.1, p. 82.
- Chollet, Francois and Others (2015). *Keras*. <https://keras.io/>.
- Ciesielski, Paul E., Patrick T. Haertel, Richard H. Johnson, Junhong Wang, and Scot M. Loehrer (2012). Developing high-quality field program sounding datasets. *Bull. Amer. Met. Soc.* 93.3, pp. 325–336. DOI: [10.1175/BAMS-D-11-00091.1](https://doi.org/10.1175/BAMS-D-11-00091.1).
- Dee, D. P. et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 553–597. DOI: [10.1002/qj.828](https://doi.org/10.1002/qj.828).
- Denby, L. (2020). Discovering the Importance of Mesoscale Cloud Organization Through Unsupervised Classification. *Geophysical Research Letters* 47.1, e2019GL085190. DOI: [10.1029/2019GL085190](https://doi.org/10.1029/2019GL085190).
- Dipankar, Anurag, Bjorn Stevens, Rieke Heinze, Christopher Moseley, Günther Zängl, Marco Giorgetta, and Slavko Brdar (2015). Large Eddy Simulation Using the General Circulation Model ICON. *Journal of Advances in Modeling Earth Systems* 7.3, pp. 963–986. DOI: [10.1002/2015MS000431](https://doi.org/10.1002/2015MS000431).
- Drager, Aryeh J. and Susan C. van den Heever (2017). Characterizing Convective Cold Pools. *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1091–1115. DOI: [10.1002/2016MS000788](https://doi.org/10.1002/2016MS000788).
- Dufresne, J.-L. et al. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics* 40, pp. 2123–2165. DOI: [10.1007/s00382-012-1636-1](https://doi.org/10.1007/s00382-012-1636-1).
- Dupont, Jean-Charles, Martial Haeffelin, Jordi Badosa, Gaelle Clain, Christophe Raux, and Damien Vignelles (2020). Characterization and corrections of relative humidity measurement from Meteomodem M10 radiosondes at mid-latitude stations. *J. Atm. and Oceanic Technol.* DOI: [10.1175/JTECH-D-18-0205.1](https://doi.org/10.1175/JTECH-D-18-0205.1).
- Eastman, Ryan and Robert Wood (Jan. 1, 2016). Factors Controlling Low-Cloud Evolution over the Eastern Subtropical Oceans: A Lagrangian Perspective Using the A-Train Satellites. *Journal of the Atmospheric Sciences* 73.1, pp. 331–351. DOI: [10.1175/JAS-D-15-0193.1](https://doi.org/10.1175/JAS-D-15-0193.1).
- Fermepin, Solange and Sandrine Bony (2014). Influence of Low-Cloud Radiative Effects on Tropical Circulation and Precipitation. *Journal of Advances in Modeling Earth Systems* 6.3, pp. 513–526. DOI: [10.1002/2013MS000288](https://doi.org/10.1002/2013MS000288).
- Fleming, James Rodger (2020). *First Woman*. New York: Oxford Uni. Press.

- GOES-R Calibration Working Group and GOES-R Series Program (2017). *NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 1b Radiances*. Online; accessed on 15 July 2020. DOI: [10.7289/V5BV7DSR](https://doi.org/10.7289/V5BV7DSR).
- Gaiser, Hans et al. (Oct. 17, 2018). *Fizyr/Keras-Retinanet: 0.5.0*. Zenodo. DOI: [10.5281/zenodo.1464720](https://doi.org/10.5281/zenodo.1464720).
- Garay, Michael J, Roger Davies, Clare Averill, and James A Westphal (Oct. 2004). ACTINIFORM CLOUDS: Overlooked Examples of Cloud Self-Organization at the Mesoscale. *Bull. Amer. Meteor. Soc.* 85.10, pp. 1585–1594.
- George, Geet, Bjorn Stevens, Sandrine Bony, Marcus Klingebiel, and Raphaela Vogel (2020). Observed Impact of Meso-Scale Vertical Motion on Cloudiness. *J. Atmos. Sci.*, pp. 1–30.
- George, Geet, Bjorn Stevens, Sandrine Bony, and Raphaela Vogel (in preparation). Mesoscale Circulation in the Trades: Structure, Variability and Interplay with Other Environmental Factors.
- Hardy, Bob (1998). “ITS-90 formulations for vapor pressure, frostpoint temperature, dewpoint temperature, and enhancement factors in the range–100 to+ 100 C.” In: *The proceedings of the third international symposium on Humidity & Moisture, Teddington, London, England*, pp. 1–8.
- Hartten, L. M., C. J. Cox, P. E. Johnston, D. E. Wolfe, S. Abbott, H. A. McColl, X.-W. Quan, and M. G. Winterkorn (2018). Ship- and island-based soundings from the 2016 El Niño Rapid Response (ENRR) field campaign. *Earth Sys. Sci. Data* 10.2, pp. 1165–1183. DOI: [10.5194/essd-10-1165-2018](https://doi.org/10.5194/essd-10-1165-2018).
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (Dec. 2015). Deep Residual Learning for Image Recognition. <http://arxiv.org/abs/1512.03385>.
- Heinze, Rieke et al. (2017). Large-Eddy Simulations over Germany Using ICON: A Comprehensive Evaluation. *Quarterly Journal of the Royal Meteorological Society* 143.702, pp. 69–100. DOI: [10.1002/qj.2947](https://doi.org/10.1002/qj.2947).
- Hennon, Christopher C., Kenneth R. Knapp, Carl J. Schreck, Scott E. Stevens, James P. Kossin, Peter W. Thorne, Paula A. Hennon, Michael C. Kruk, Jared Rennie, Jean-Maurice Gadéa, Maximilian Striegl, and Ian Carley (Apr. 2015). Cyclone Center: Can Citizen Scientists Improve Tropical Cyclone Intensity Records? *Bulletin of the American Meteorological Society* 96.4, pp. 591–607. DOI: [10.1175/BAMS-D-13-00152.1](https://doi.org/10.1175/BAMS-D-13-00152.1). URL: <http://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00152.1>.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2018). *ERA5 Hourly Data on Pressure Levels from 1979 to Present*. DOI: [10.24381/cds.bd0915c6](https://doi.org/10.24381/cds.bd0915c6).
- Hersbach, Hans et al. (2020). The ERA5 Global Reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: [10.1002/qj.3803](https://doi.org/10.1002/qj.3803).
- Holland, Joshua Z. and Eugene M. Rasmusson (Jan. 1, 1973). Measurements of the Atmospheric Mass, Energy, and Momentum Budgets Over a 500-Kilometer Square of Tropical Ocean. *Monthly Weather Review* 101.1, pp. 44–55. DOI: [10.1175/1520-0493\(1973\)101<0044:MOTAME>2.3.CO;2](https://doi.org/10.1175/1520-0493(1973)101<0044:MOTAME>2.3.CO;2).
- Hong, Seungkyun, Seongchan Kim, Minsu Joh, and Sa-kwang Song (Aug. 2017). GlobeNet: Convolutional Neural Networks for Typhoon Eye Tracking from Remote Sensing Imagery. <http://arxiv.org/abs/1708.03417>.

- Hourdin, F, M Gueye, B Diallo, J-L Dufresne, J Escribano, L Menut, B Marticoréna, G Siour, and F Guichard (2015). Parameterization of convective transport in the boundary layer and its impact on the representation of the diurnal cycle of wind and dust emissions. *Atmospheric Chemistry and Physics* 15.12, pp. 6775–6788.
- Howard, Luke (1803). *On the Modifications of Clouds, and on the Principles of Their Production, Suspension, and Destruction; Being the Substance of an Essay Read before the Askenian Society in the Session 1802-3*. Sonderdruck.
- Hsu, SA, Eric A Meindl, and David B Gilhousen (1994). Determining the power-law wind-profile exponent under near-neutral stability conditions at sea. *Journal of Applied Meteorology and Climatology* 33.6, pp. 757–765.
- Jacob, Marek, Pavlos Kollias, Felix Ament, Vera Schemann, and Susanne Crewell (Mar. 6, 2020). *Multi-Layer Cloud Conditions in Trade Wind Shallow Cumulus – Confronting Models with Airborne Observations*. preprint. Atmospheric Sciences. DOI: [10.5194/gmd-2020-14](https://doi.org/10.5194/gmd-2020-14).
- Jansen, F. (2020a). *Ceilometer measurements RV Meteor*. DOI: [10.25326/53](https://doi.org/10.25326/53). URL: <https://doi.org/10.25326/53>.
- (2020b). *Surface meteorology Barbados Cloud Observatory*. DOI: [10.25326/54](https://doi.org/10.25326/54). URL: <https://doi.org/10.25326/54>.
- Janssens, Martin, Jordi Vilà-Guerau de Arellano, Marten Scheffer, Coco Antonissen, A. Pier Siebesma, and Franziska Glassmeier (2021). Cloud Patterns in the Trades Have Four Interpretable Dimensions. *Geophysical Research Letters* 48.5, e2020GL091001. DOI: [10.1029/2020GL091001](https://doi.org/10.1029/2020GL091001).
- Johnson, Richard H. and Paul E. Ciesielski (2013). Structure and properties of Madden-Julian Oscillations deduced from DYNAMO sounding arrays. *J. Atmos. Sci.* 70, pp. 3157–3179. DOI: [10.1175/JAS-D-13-065.1](https://doi.org/10.1175/JAS-D-13-065.1).
- Kilpatrick, T., N. Schneider, and B. Qiu (2013). Boundary layer convergence induced by strong winds across a midlatitude SST front. *J. Climate* 11, pp. 1698–1718. DOI: [10.1175/JCLI-D-13-00101.1](https://doi.org/10.1175/JCLI-D-13-00101.1).
- Klein, Stephen A., Alex Hall, Joel R. Norris, and Robert Pincus (Nov. 2017). Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review. *Surveys in Geophysics* 38.6, pp. 1307–1329. DOI: [10.1007/s10712-017-9433-3](https://doi.org/10.1007/s10712-017-9433-3).
- Klein, Stephen A and Dennis L Hartmann (1993). The seasonal cycle of low stratiform clouds. *Journal of Climate* 6.8, pp. 1587–1606.
- Klingebiel, Marcus, Virendra P. Ghatge, Ann Kristin Naumann, Florian Ditas, Mira L. Pöhlker, Christopher Pöhlker, Konrad Kandler, Heike Konow, and Bjorn Stevens (Feb. 27, 2019). Remote Sensing of Sea Salt Aerosol below Trade Wind Clouds. *Journal of the Atmospheric Sciences* 76.5, pp. 1189–1202. DOI: [10.1175/JAS-D-18-0139.1](https://doi.org/10.1175/JAS-D-18-0139.1).
- Knapp, Kenneth R., Steve Ansari, Caroline L. Bain, Mark A. Bourassa, Michael J. Dickinson, Chris Funk, Chip N. Helms, Christopher C. Hennon, Christopher D. Holmes, George J. Huffman, James P. Kossin, Hai-Tien Lee, Alexander Loew, and Gudrun Magnusdottir (2011). Globally Gridded Satellite Observations for Climate Studies. *Bulletin of the American Meteorological Society* 92.7, pp. 893–907. DOI: [10.1175/2011BAMS3039.1](https://doi.org/10.1175/2011BAMS3039.1).
- Krueger, Steven K. (1988). Numerical simulation of tropical cumulus clouds and their interaction with the subcloud layer. *J. Atmos. Sci.* 45.16, pp. 2221–2250. DOI: [10.1175/1520-0469\(1988\)045<2221:NSOTCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<2221:NSOTCC>2.0.CO;2).
- Kuettner, Joachim P. and Joshua Holland (1969). the BOMEX project. *Bull. Amer. Meteor. Soc.* 50.6, pp. 394–403. DOI: [10.1175/1520-0477-50.6.394](https://doi.org/10.1175/1520-0477-50.6.394). (Visited on 05/07/2020).

- Kurth, Thorsten, Sean Treichler, Joshua Romero, Mayur Mudigonda, Nathan Luehr, Everett Phillips, Ankur Mahesh, Michael Matheson, Jack Deslippe, Massimiliano Fatica, Prabhat, and Michael Houston (Oct. 2018). Exascale Deep Learning for Climate Analytics. <http://arxiv.org/abs/1810.01993>.
- Lamer, K., P. Kollias, and L. Nuijens (June 27, 2015). Observations of the Variability of Shallow Trade Wind Cumulus Cloudiness and Mass Flux. *Journal of Geophysical Research: Atmospheres* 120.12, 2014JD022950. DOI: [10.1002/2014JD022950](https://doi.org/10.1002/2014JD022950).
- Laxenaire, R., S. Speich, B. Blanke, A. Chaigneau, C. Pegliasco, and A. Stegner (2018). Anticyclonic eddies connecting the western boundaries of Indian and Atlantic Oceans. *J. Geophys. Res. Oceans* 123.11, pp. 7651–7677. DOI: [10.1029/2018JC014270](https://doi.org/10.1029/2018JC014270).
- LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (May 2015). Deep learning. *Nature* 521.7553, pp. 436–444. DOI: [10.1038/nature14539](https://doi.org/10.1038/nature14539). URL: <http://www.nature.com/articles/nature14539>.
- Lenschow, Donald H., Verica Savic-Jovicic, and Bjorn Stevens (Dec. 1, 2007). Divergence and Vorticity from Aircraft Air Motion Measurements. *Journal of Atmospheric and Oceanic Technology* 24.12, pp. 2062–2072. DOI: [10.1175/2007JTECHA940.1](https://doi.org/10.1175/2007JTECHA940.1).
- Ligda, Myron G H (1951). Radar Storm Observations. In: *Compendium of Meteorology*. Boston: American Meteorological Society, pp. 1265–1282.
- Lin, Tsung-Yi, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár (Feb. 7, 2018). *Focal Loss for Dense Object Detection*. arXiv: [1708.02002](https://arxiv.org/abs/1708.02002) [cs]. URL: <http://arxiv.org/abs/1708.02002> (visited on 09/07/2021).
- Lin, Xin and Richard H. Johnson (1996). Heating, moistening, and rainfall over the western Pacific warm pool during TOGA COARE. *J. Atmos. Sci.* 53.22, pp. 3367–3383. DOI: [10.1175/1520-0469\(1996\)053<3367:HMAROT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<3367:HMAROT>2.0.CO;2).
- Liu, Yunjie, Evan Racah, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, and William Collins (2016). Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets. <https://arxiv.org/abs/1605.01156>.
- Loeb, Norman G., David R. Doelling, Hailan Wang, Wenying Su, Cathy Nguyen, Joseph G. Corbett, Lusheng Liang, Cristian Mitrescu, Fred G. Rose, and Seiji Kato (Jan. 15, 2018). Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *Journal of Climate* 31.2, pp. 895–918. DOI: [10.1175/JCLI-D-17-0208.1](https://doi.org/10.1175/JCLI-D-17-0208.1).
- Lonitz, Katrin, Bjorn Stevens, Louise Nuijens, Vivek Sant, Lutz Hirsch, and Axel Seifert (Dec. 2015). The Signature of Aerosols and Meteorology in Long-Term Cloud Radar Observations of Trade Wind Cumuli. *Journal of the Atmospheric Sciences* 72.12, pp. 4643–4659. DOI: [10.1175/JAS-D-14-0348.1](https://doi.org/10.1175/JAS-D-14-0348.1).
- Löhnert, Ulrich and Susanne Crewell (2003). Accuracy of Cloud Liquid Water Path from Ground-Based Microwave Radiometry 1. Dependency on Cloud Model Statistics. *Radio Science* 38.3. DOI: [10.1029/2002RS002654](https://doi.org/10.1029/2002RS002654).
- Ma, Jian, Gregory Foltz, Brian Soden, Gang Huang, Jie He, and Changming Dong (Nov. 2016). Will surface winds weaken in response to global warming? *Environmental Research Letters* 11, p. 124012. DOI: [10.1088/1748-9326/11/12/124012](https://doi.org/10.1088/1748-9326/11/12/124012).
- Malkus, Joanne S. and Herbert Riehl (1964). Cloud Structure and Distributions over the Tropical Pacific Ocean1. *Tellus* 16.3, pp. 275–287. DOI: [10.1111/j.2153-3490.1964.tb00167.x](https://doi.org/10.1111/j.2153-3490.1964.tb00167.x).

- Mandelbrot, Benoît (2010). A Geometry Able to Include Mountains and Clouds. In: *The Colours of Infinity: The Beauty and Power of Fractals*. Ed. by Nigel Lesmoir-Gordon. London: Springer, pp. 38–57. ISBN: 978-1-84996-486-9. DOI: [10.1007/978-1-84996-486-9_3](https://doi.org/10.1007/978-1-84996-486-9_3).
- Mapes, Brian E., Paul E. Ciesielski, and Richard H. Johnson (2003). Sampling errors in Rawinsonde-array budgets. *J. Atmos. Sci.* 60.21, pp. 2697–2714. DOI: [10.1175/1520-0469\(2003\)060<2697:SEIRB>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2697:SEIRB>2.0.CO;2).
- McCoy, Isabel L, Robert Wood, and Jennifer K Fletcher (2017). Identifying meteorological controls on open and closed mesoscale cellular convection associated with marine cold air outbreaks. *Journal of Geophysical Research: Atmospheres* 122.21, pp. 11–678.
- Mech, Mario, Maximilian Maahn, Stefan Kneifel, Davide Ori, Emiliano Orlandi, Pavlos Kollias, Vera Schemann, and Susanne Crewell (Sept. 13, 2020). PAMTRA 1.0: The Passive and Active Microwave Radiative TRANSfer Tool for Simulating Radiometer and Radar Measurements of the Cloudy Atmosphere. *Geoscientific Model Development* 13.9, pp. 4229–4251. DOI: [10.5194/gmd-13-4229-2020](https://doi.org/10.5194/gmd-13-4229-2020).
- Medeiros, Brian and Louise Nuijens (May 31, 2016). Clouds at Barbados Are Representative of Clouds across the Trade Wind Regions in Observations and Climate Models. *Proceedings of the National Academy of Sciences* 113.22, E3062–E3070. DOI: [10.1073/pnas.1521494113](https://doi.org/10.1073/pnas.1521494113).
- Medeiros, Brian, Bjorn Stevens, and Sandrine Bony (2015). Using aquaplanets to understand the robust responses of comprehensive climate models to forcing. *Climate Dynamics* 44.7, pp. 1957–1977. DOI: [10.1007/s00382-014-2138-0](https://doi.org/10.1007/s00382-014-2138-0).
- Meehl, Gerald A., Catherine A. Senior, Veronika Eyring, Gregory Flato, Jean-Francois Lamarque, Ronald J. Stouffer, Karl E. Taylor, and Manuel Schlund (2020). Context for Interpreting Equilibrium Climate Sensitivity and Transient Climate Response from the CMIP6 Earth System Models. *Science Advances* 6.26, eaba1981. DOI: [10.1126/sciadv.aba1981](https://doi.org/10.1126/sciadv.aba1981).
- Mlawer, Eli J., Steven J. Taubman, Patrick D. Brown, Michael J. Iacono, and Shepard A. Clough (July 27, 1997). Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated-k Model for the Longwave. *Journal of Geophysical Research: Atmospheres* 102.D14, pp. 16663–16682. DOI: [10.1029/97JD00237](https://doi.org/10.1029/97JD00237).
- Mudigonda, Mayur, Sookyung Kim, Ankur Mahesh, Samira Kahou, Karthik Kashinath, Dean Williams, Vincent Michalski, Travis O'brien, and Mr Prabhat (2017). “Segmenting and Tracking Extreme Climate Events using Neural Networks.” In: *Deep Learning for Physical Sciences, NIPS workshop*. URL: https://dl4physicalsciences.github.io/files/nips_dlps_2017_20.pdf.
- Muhlbauer, A., I. L. McCoy, and R. Wood (July 3, 2014). Climatology of Stratocumulus Cloud Morphologies: Microphysical Properties and Radiative Effects. *Atmos. Chem. Phys.* 14.13, pp. 6695–6716. DOI: [10.5194/acp-14-6695-2014](https://doi.org/10.5194/acp-14-6695-2014).
- Muller, Caroline J. and Isaac M. Held (Aug. 2012). Detailed Investigation of the Self-Aggregation of Convection in Cloud-Resolving Simulations. *Journal of the Atmospheric Sciences* 69.8, pp. 2551–2565. DOI: [10.1175/JAS-D-11-0257.1](https://doi.org/10.1175/JAS-D-11-0257.1).
- Nicholls, Stephen D and George S Young (2007). Dendritic patterns in tropical cumulus: An observational analysis. *Monthly weather review* 135.5, pp. 1994–2005.
- Nielsen, Michael A. (2015). *Neural Networks and Deep Learning*. Determination Press. URL: <http://neuralnetworksanddeeplearning.com>.

- Nitta, Tsuyoshi and Steven Esbensen (1974). Heat and moisture budget analyses using BOMEX data. *Mon. Wea. Rev.* 102.1, pp. 17–28. DOI: [10.1175/1520-0493\(1974\)102<0017:HAMBAU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1974)102<0017:HAMBAU>2.0.CO;2).
- Norris, J R (1998). Low cloud type over the ocean from surface observations. Part II: Geographical and seasonal variations. *Journal of Climate* 11.3, pp. 383–403.
- Nuijens, L., I. Serikov, L. Hirsch, K. Lonitz, and B. Stevens (Oct. 1, 2014). The Distribution and Variability of Low-Level Cloud in the North Atlantic Trades. *Quarterly Journal of the Royal Meteorological Society* 140.684, pp. 2364–2374. DOI: [10.1002/qj.2307](https://doi.org/10.1002/qj.2307).
- Nuijens, Louise, Brian Medeiros, Irina Sandu, and Maike Ahlgrimm (Nov. 2015). Observed and Modeled Patterns of Covariability between Low-Level Cloudiness and the Structure of the Trade-Wind Layer: PATTERNS OF COVARIABILITY. *Journal of Advances in Modeling Earth Systems*, n/a–n/a. DOI: [10.1002/2015MS000483](https://doi.org/10.1002/2015MS000483).
- Nuijens, Louise and A. Pier Siebesma (June 2019). Boundary Layer Clouds and Convection over Subtropical Oceans in our Current and in a Warmer Climate. *Current Climate Change Reports* 5.2, pp. 80–94. DOI: [10.1007/s40641-019-00126-x](https://doi.org/10.1007/s40641-019-00126-x).
- Nuijens, Louise and Bjorn Stevens (Jan. 2012). The Influence of Wind Speed on Shallow Marine Cumulus Convection. *Journal of the Atmospheric Sciences* 69.1, pp. 168–184. DOI: [10.1175/JAS-D-11-02.1](https://doi.org/10.1175/JAS-D-11-02.1).
- Nuijens, Louise, Bjorn Stevens, and A. Pier Siebesma (July 1, 2009). The Environment of Precipitating Shallow Cumulus Convection. *Journal of the Atmospheric Sciences* 66.7, pp. 1962–1979. DOI: [10.1175/2008JAS2841.1](https://doi.org/10.1175/2008JAS2841.1).
- Orlanski, Isidoro (1975). A Rational Subdivision of Scales for Atmospheric Processes. *Bulletin of the American Meteorological Society* 56.5, pp. 527–530. JSTOR: [26216020](https://www.jstor.org/stable/26216020).
- Parishani, Hossein, Michael S Pritchard, Christopher S Bretherton, Christopher R Terai, Matthew C Wyant, Marat Khairoutdinov, and Balwinder Singh (Dec. 2018). Insensitivity of the Cloud Response to Surface Warming Under Radical Changes to Boundary Layer Turbulence and Cloud Microphysics: Results From the Ultraparameterized CAM. *Journal of Advances in Modeling Earth Systems* 10.12, pp. 3139–3158.
- Peters, Gerhard, Bernd Fischer, and Tage Andersson (2002). Rain Observations with a Vertically Looking Micro Rain Radar (MRR). *Boreal environment research*. URL: <https://pdfs.semanticscholar.org/3a20/ec749e6caa768129cd9c6193f7128b771138.pdf> (visited on 09/01/2020).
- Qu, Xin, Alex Hall, Stephen Klein, and Peter Caldwell (2015). The strength of the tropical inversion and its response to climate change in 18 CMIP5 models. *Climate Dynamics* 45.1/2, pp. 375–396.
- Racah, Evan, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Prabhat, and Christopher Pal (Dec. 2016). ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. <http://arxiv.org/abs/1612.02095>.
- Rasp, Stephan (June 2020a). *Sugar, Flower, Fish or Gravel - Example ML predictions*. DOI: [10.6084/m9.figshare.8236289.v2](https://doi.org/10.6084/m9.figshare.8236289.v2). URL: <https://doi.org/10.6084/m9.figshare.8236289.v2>.
- (June 2020b). *Sugar, Flower, Fish or Gravel - Example human labels*. DOI: [10.6084/m9.figshare.8236289.v2](https://doi.org/10.6084/m9.figshare.8236289.v2). URL: <https://doi.org/10.6084/m9.figshare.12141264.v1>.
- (June 2020c). *Sugar, Flower, Fish or Gravel - Global Predictions*. DOI: [10.6084/m9.figshare.8236298.v2](https://doi.org/10.6084/m9.figshare.8236298.v2). URL: <https://doi.org/10.6084/m9.figshare.8236298.v2>.

- Rasp, Stephan, Michael S Pritchard, and Pierre Gentine (Sept. 2018). Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences of the United States of America* 115.39, pp. 9684–9689. DOI: [10.1073/pnas.1810286115](https://doi.org/10.1073/pnas.1810286115). URL: <http://www.ncbi.nlm.nih.gov/pubmed/30190437><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6166853>.
- Rasp, Stephan, Hauke Schulz, Sandrine Bony, and Bjorn Stevens (Nov. 20, 2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. *Bulletin of the American Meteorological Society* 101.11, E1980–E1995. DOI: [10.1175/BAMS-D-19-0324.1](https://doi.org/10.1175/BAMS-D-19-0324.1).
- Raspaud, Martin et al. (June 2019). *Pytrol/Satpy: Version 0.16.0*. Version v0.16.0. Zenodo. DOI: [10.5281/zenodo.3250583](https://doi.org/10.5281/zenodo.3250583).
- Rauber, Robert M. et al. (2007). Rain in shallow cumulus over the ocean: The RICO Campaign. *Bull. Amer. Meteor. Soc.* 88.12, pp. 1912–1928. DOI: [10.1175/BAMS-88-12-1912](https://doi.org/10.1175/BAMS-88-12-1912).
- Redelsperger, J.-L., M.-N. Bouin, J. Pianezze, V. Garnier, and L. Marié (2019). Impact of a sharp, small-scale SST front on the marine atmospheric boundary layer on the Iroise Sea: Analysis from a hectometric simulation. *Quart. J. Roy. Meteor. Soc.* 145.725, pp. 3692–3714. DOI: [10.1002/qj.3650](https://doi.org/10.1002/qj.3650).
- Reed, R. J. and E. E. Recker (1971). Structure and properties of synoptic-scale wave disturbances in the equatorial western Pacific. *J. Atmos. Sci.* 28, pp. 1117–1133.
- Reichstein, Markus, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and Prabhat (Feb. 2019). Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature* 566.7743 (7743), pp. 195–204. DOI: [10.1038/s41586-019-0912-1](https://doi.org/10.1038/s41586-019-0912-1).
- Rieck, Malte, Louise Nuijens, and Bjorn Stevens (2012). Marine Boundary Layer Cloud Feedbacks in a Constant Relative Humidity Atmosphere. *Journal of the Atmospheric Sciences* 69.8, pp. 2538–2550. DOI: [10.1175/JAS-D-11-0203.1](https://doi.org/10.1175/JAS-D-11-0203.1).
- Riehl, Herbert (1945). *Waves in the Easterlies and the Polar Front in the Tropics: A Report on Research Conducted at the Institute of Tropical Meteorology of the University of Chicago at the University of Puerto Rico, Rio Piedras, P.R.* Chicago, Ill.: University of Chicago Press.
- (1954). *Tropical Meteorology*. McGraw-Hill.
- Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (May 2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. <http://arxiv.org/abs/1505.04597>.
- Ruppert, James H and Richard H Johnson (2016). On the cumulus diurnal cycle over the tropical warm pool. *Journal of Advances in Modeling Earth Systems* 8.2, pp. 669–690.
- Satoh, Masaki, Bjorn Stevens, Falko Judt, Marat Khairoutdinov, Shian-Jiann Lin, William M. Putman, and Peter Düben (Sept. 1, 2019). Global Cloud-Resolving Models. *Current Climate Change Reports* 5.3, pp. 172–184. DOI: [10.1007/s40641-019-00131-0](https://doi.org/10.1007/s40641-019-00131-0).
- Saunders, Roger, James Hocking, Emma Turner, Peter Rayer, David Rundle, Pascal Brunel, Jerome Vidot, Pascale Roquet, Marco Matricardi, Alan Geer, Niels Bormann, and Cristina Lupu (July 10, 2018). An Update on the RTTOV Fast Radiative Transfer Model (Currently at Version 12). *Geoscientific Model Development* 11.7, pp. 2717–2737. DOI: [10.5194/gmd-11-2717-2018](https://doi.org/10.5194/gmd-11-2717-2018).
- Scher, Sebastian and Gabriele Messori (July 2019). Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as

- a study ground. *Geoscientific Model Development* 12.7, pp. 2797–2809. DOI: [10.5194/gmd-12-2797-2019](https://doi.org/10.5194/gmd-12-2797-2019). URL: <https://www.geosci-model-dev.net/12/2797/2019/>.
- Schlemmer, Linda, Peter Bechtold, Irina Sandu, and Maike Ahlgrimm (2017). Uncertainties related to the representation of momentum transport in shallow convection. *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1269–1291.
- Schulz, Hauke (Dec. 2020). *EUREC4A radiosonde software*. Version v3.0.0. DOI: [10.5281/zenodo.3712223](https://doi.org/10.5281/zenodo.3712223).
- (in preparation). Meso-Scale Organization during EUREC4A.
- Schulz, Hauke, Ryan Eastman, and Bjorn Stevens (2021). Characterization and Evolution of Organized Shallow Convection in the Downstream North Atlantic Trades. *Journal of Geophysical Research: Atmospheres* 126.17, e2021JD034575. DOI: [10.1029/2021JD034575](https://doi.org/10.1029/2021JD034575).
- Schulz, Hauke and Bjorn Stevens (July 19, 2018). Observing the Tropical Atmosphere in Moisture Space. *Journal of the Atmospheric Sciences* 75.10, pp. 3313–3330. DOI: [10.1175/JAS-D-17-0375.1](https://doi.org/10.1175/JAS-D-17-0375.1).
- Schumacher, Ernest F (1973). *Small Is Beautiful: Economics as If People Mattered*. London: Blond & Briggs.
- Seidel, Dian J., Qiang Fu, William J. Randel, and Thomas J. Reichler (Jan. 2008). Widening of the Tropical Belt in a Changing Climate. *Nature Geoscience* 1.1 (1), pp. 21–24. DOI: [10.1038/ngeo.2007.38](https://doi.org/10.1038/ngeo.2007.38).
- Seifert, A. and K. D. Beheng (Feb. 1, 2006). A Two-Moment Cloud Microphysics Parameterization for Mixed-Phase Clouds. Part 1: Model Description. *Meteorology and Atmospheric Physics* 92.1, pp. 45–66. DOI: [10.1007/s00703-005-0112-4](https://doi.org/10.1007/s00703-005-0112-4).
- Seifert, A. and T. Heus (June 10, 2013). Large-Eddy Simulation of Organized Precipitating Trade Wind Cumulus Clouds. *Atmos. Chem. Phys.* 13.11, pp. 5631–5645. DOI: [10.5194/acp-13-5631-2013](https://doi.org/10.5194/acp-13-5631-2013).
- Seifert, Axel, Thijs Heus, Robert Pincus, and Bjorn Stevens (Dec. 1, 2015). Large-Eddy Simulation of the Transient and near-Equilibrium Behavior of Precipitating Shallow Convection. *Journal of Advances in Modeling Earth Systems* 7.4, pp. 1918–1937. DOI: [10.1002/2015MS000489](https://doi.org/10.1002/2015MS000489).
- Sherwood, S. C. et al. (2020). An Assessment of Earth’s Climate Sensitivity Using Multiple Lines of Evidence. *Reviews of Geophysics* 58.4, e2019RG000678. DOI: [10.1029/2019RG000678](https://doi.org/10.1029/2019RG000678).
- Sherwood, Steven C, Sandrine Bony, and Jean-Louis Dufresne (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature* 505.7481, pp. 37–42. DOI: [10.1038/nature12829](https://doi.org/10.1038/nature12829).
- Siebesma, A. P. and J. W. M. Cuijpers (Mar. 1, 1995). Evaluation of Parametric Assumptions for Shallow Cumulus Convection. *Journal of the Atmospheric Sciences* 52.6, pp. 650–666. DOI: [10.1175/1520-0469\(1995\)052<0650:EOPAFS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<0650:EOPAFS>2.0.CO;2).
- Siebesma, A. Pier, Sandrine Bony, Christian Jakob, and Bjorn Stevens, eds. (2020). *Clouds and Climate: Climate Science’s Greatest Challenge*. Cambridge: Cambridge University Press. ISBN: 978-1-107-06107-1. DOI: [10.1017/9781107447738](https://doi.org/10.1017/9781107447738).
- Siebesma, A Pier, Christopher S Bretherton, Andrew Brown, Andreas Chlond, Joan Cuxart, Peter G Duynkerke, Hongli Jiang, Marat Khairoutdinov, David Lewellen, Chin-Hoh Moeng, Enrique Sanchez, Bjorn Stevens, and David E Stevens (2003). A Large Eddy Simulation Intercomparison Study of Shallow Cumulus Convection. *J. Atmos. Sci.* 60, p. 19.

- Sivira, R.G., H. Brogniez, C. Mallet, and Y. Oussar (2015). A layer-averaged relative humidity profile retrieval for microwave observations: Design and results for the Megha-Tropiques payload. *Atmospheric Measurement Techniques* 8.3, pp. 1055–1071.
- Steinke, S., S. Eikenberg, U. Löhnert, G. Dick, D. Klocke, P. Di Girolamo, and S. Crewell (Mar. 9, 2015). Assessment of Small-Scale Integrated Water Vapour Variability during HOPE. *Atmospheric Chemistry and Physics* 15.5, pp. 2675–2692. DOI: [10.5194/acp-15-2675-2015](https://doi.org/10.5194/acp-15-2675-2015).
- Stephan, Claudia Christine (Mar. 1, 2021). Mechanism for the Formation of Arc-Shaped Cloud Lines over the Tropical Oceans. *Journal of the Atmospheric Sciences* 78.3, pp. 817–824. DOI: [10.1175/JAS-D-20-0129.1](https://doi.org/10.1175/JAS-D-20-0129.1).
- Stephan, Claudia Christine et al. (Dec. 2020a). *Ship- and island-based atmospheric soundings from the 2020 EUREC4A field campaign*. Version 3.0.0. DOI: [10.25326/137](https://doi.org/10.25326/137). URL: <https://doi.org/10.25326/137>.
- Stephan, Claudia et al. (2020b). *Radiosonde Measurements from the EUREC4A Field Campaign*. DOI: [10.25326/62](https://doi.org/10.25326/62).
- Stevens, Bjorn, Andrew S. Ackerman, Bruce A. Albrecht, Andrew R. Brown, Andreas Chlond, Joan Cuxart, Peter G. Duynkerke, David C. Lewellen, Malcolm K. Macvean, Roel A. J. Neggers, Enrique Sánchez, A. Pier Siebesma, and David E. Stevens (July 1, 2001). Simulations of Trade Wind Cumuli under a Strong Inversion. *Journal of the Atmospheric Sciences* 58.14, pp. 1870–1891. DOI: [10.1175/1520-0469\(2001\)058<1870:SOTWCU>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<1870:SOTWCU>2.0.CO;2).
- Stevens, Bjorn, Sandrine Bony, Hélène Brogniez, Laureline Hentgen, Cathy Hohenegger, Christoph Kiemle, Tristan S. L'Ecuyer, Ann Kristin Naumann, Hauke Schulz, Pier A. Siebesma, Jessica Vial, Dave M. Winker, and Paquita Zuidema (2020). Sugar, Gravel, Fish and Flowers: Mesoscale Cloud Patterns in the Trade Winds. *Quarterly Journal of the Royal Meteorological Society* 146.726, pp. 141–152. DOI: [10.1002/qj.3662](https://doi.org/10.1002/qj.3662).
- Stevens, Bjorn, David Farrell, Lutz Hirsch, Friedhelm Jansen, Louise Nuijens, Ilya Serikov, Björn Brüggemann, Marvin Forde, Holger Linne, Katrin Lonitz, and Joseph M. Prospero (May 2016). The Barbados Cloud Observatory: Anchoring Investigations of Clouds and Circulation on the Edge of the ITCZ. *Bulletin of the American Meteorological Society* 97.5, pp. 787–801. DOI: [10.1175/BAMS-D-14-00247.1](https://doi.org/10.1175/BAMS-D-14-00247.1).
- Stevens, Bjorn, Marco Giorgetta, Monika Esch, Thorsten Mauritsen, Traute Crueger, Sebastian Rast, Marc Salzmann, Hauke Schmidt, Jürgen Bader, Karoline Block, Renate Brokopf, Irina Fast, Stefan Kinne, Luis Kornblueh, Ulrike Lohmann, Robert Pincus, Thomas Reichler, and Erich Roeckner (2013). Atmospheric Component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems* 5.2, pp. 146–172. DOI: [10.1002/jame.20015](https://doi.org/10.1002/jame.20015).
- Stevens, Bjorn, Gabor Vali, Kimberly Comstock, Robert Wood, Margreet C. van Zanten, Philip H. Austin, Christopher S. Bretherton, and Donald H. Lenschow (Jan. 1, 2005). Pockets of Open Cells and Drizzle in Marine Stratocumulus. *Bulletin of the American Meteorological Society* 86.1, pp. 51–58. DOI: [10.1175/BAMS-86-1-51](https://doi.org/10.1175/BAMS-86-1-51).
- Stevens, Bjorn et al. (2019). A high-altitude long-range aircraft configured as a cloud observatory: The NARVAL expeditions. *Bull. Amer. Meteor. Soc.* 100.6, pp. 1061–1077. DOI: [10.1175/BAMS-D-18-0198.1](https://doi.org/10.1175/BAMS-D-18-0198.1).
- Stevens, Bjorn et al. (Aug. 25, 2021). EUREC4A. *Earth System Science Data* 13.8, pp. 4067–4119. DOI: [10.5194/essd-13-4067-2021](https://doi.org/10.5194/essd-13-4067-2021).

- Stull, R.B. (1988). *An Introduction to Boundary Layer Meteorology*. Atmospheric and Oceanographic Sciences Library. Springer Netherlands. ISBN: 9789027727695.
- Suhr, Isabel and Charles Martin (2020). *Atmospheric Sounding Processing ENvironment (ASPEN) User Manual*. URL: <https://ncar.github.io/aspdocs/pdf/aspdocs.pdf>.
- Taburet, G., A. Sanchez-Roman, M. Ballarotta, M.-I. Pujol, J.-F. Legeais, F. Fournier, Y. Faugere, and G. Dibarboure (2019). DUACS DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Sci.* 15.5, pp. 1207–1224. DOI: [10.5194/os-15-1207-2019](https://doi.org/10.5194/os-15-1207-2019).
- Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.* 117.8, pp. 1779–1800. DOI: [10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2).
- Tobin, Isabelle, Sandrine Bony, and Remy Roca (June 4, 2012). Observational Evidence for Relationships between the Degree of Aggregation of Deep Convection, Water Vapor, Surface Fluxes, and Radiation. *Journal of Climate* 25.20, pp. 6885–6904. DOI: [10.1175/JCLI-D-11-00258.1](https://doi.org/10.1175/JCLI-D-11-00258.1).
- Tomassini, Lorenzo, Aiko Voigt, and Bjorn Stevens (2015). On the connection between tropical circulation, convective mixing, and climate sensitivity. *Quarterly Journal of the Royal Meteorological Society* 141.689, pp. 1404–1416.
- Tompkins, Adrian M and Addisu G Semie (2017). Organization of tropical convection in low vertical wind shears: Role of updraft entrainment. *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1046–1068.
- Ulbrich, U., J. G. Pinto, H. Kupfer, G. C. Leckebusch, T. Spanghehl, and M. Reyers (Apr. 15, 2008). Changing Northern Hemisphere Storm Tracks in an Ensemble of IPCC Climate Change Simulations. *Journal of Climate* 21.8, pp. 1669–1679. DOI: [10.1175/2007JCLI1992.1](https://doi.org/10.1175/2007JCLI1992.1).
- Vial, Jessica, Sandrine Bony, Jean-Louis Dufresne, and Romain Roehrig (2016). Coupling between lower-tropospheric convective mixing and low-level clouds: Physical mechanisms and dependence on convection scheme. *Journal of advances in modeling earth systems* 8.4, pp. 1892–1911.
- Vial, Jessica, Sandrine Bony, Bjorn Stevens, and Raphaela Vogel (Nov. 1, 2017). Mechanisms and Model Diversity of Trade-Wind Shallow Cumulus Cloud Feedbacks: A Review. *Surveys in Geophysics* 38.6, pp. 1331–1353. DOI: [10.1007/s10712-017-9418-2](https://doi.org/10.1007/s10712-017-9418-2).
- Vial, Jessica, Jean-Louis Dufresne, and Sandrine Bony (2013). On the Interpretation of Inter-Model Spread in CMIP5 Climate Sensitivity Estimates. *Climate Dynamics* 41.11-12, pp. 3339–3362. DOI: [10.1007/s00382-013-1725-9](https://doi.org/10.1007/s00382-013-1725-9).
- Vial, Jessica, Raphaela Vogel, Sandrine Bony, Bjorn Stevens, David M. Winker, Xia Cai, Cathy Hohenegger, Ann Kristin Naumann, and Hélène Brogniez (Oct. 2019). A New Look at the Daily Cycle of Trade Wind Cumuli. *Journal of Advances in Modeling Earth Systems* 11.10, pp. 3148–3166. DOI: [10.1029/2019MS001746](https://doi.org/10.1029/2019MS001746).
- Vial, Jessica, Raphaela Vogel, and Hauke Schulz (2021). On the Daily Cycle of Mesoscale Cloud Organization in the Winter Trades. *Quarterly Journal of the Royal Meteorological Society* 147.738, pp. 2850–2873. DOI: [10.1002/qj.4103](https://doi.org/10.1002/qj.4103).
- Vogel, R., S. Bony, and B. Stevens (2019). Estimating the shallow convective mass flux from the sub-cloud layer mass budget. *Journal of the Atmospheric Sciences*.
- Vogel, Raphaela, Heike Konow, Hauke Schulz, and Paquita Zuidema (June 1, 2021). A Climatology of Trade-Wind Cumulus Cold Pools and Their Link to Mesoscale

- Cloud Organization. *Atmospheric Chemistry and Physics Discussions*, pp. 1–32. DOI: [10.5194/acp-2021-420](https://doi.org/10.5194/acp-2021-420).
- Vogel, Raphaela, Louise Nuijens, and Bjorn Stevens (June 1, 2016). The Role of Precipitation and Spatial Organization in the Response of Trade-Wind Clouds to Warming. *Journal of Advances in Modeling Earth Systems* 8.2, pp. 843–862. DOI: [10.1002/2015MS000568](https://doi.org/10.1002/2015MS000568).
- (2020). Influence of Deepening and Mesoscale Organization of Shallow Convection on Stratiform Cloudiness in the Downstream Trades. *Quarterly Journal of the Royal Meteorological Society* 146.726, pp. 174–185. DOI: [10.1002/qj.3664](https://doi.org/10.1002/qj.3664).
- WMO (2009). *Manual on the Global Telecommunication System No 386, Part II, Att II-5, "Data Designators T1T2A1A2ii in abbreviated headings"*. URL: <https://community.wmo.int/activity-areas/operational-information-service/volume-c1>.
- Weger, R. C., J. Lee, Tianri Zhu, and R. M. Welch (1992). Clustering, randomness and regularity in cloud fields: 1. Theoretical considerations. *Journal of Geophysical Research* 97.D18, pp. 20519–20536. DOI: [10.1029/92JD02038](https://doi.org/10.1029/92JD02038).
- Wielicki, Bruce, B R. Barkstrom, and E F. Harrison (Jan. 1996). Clouds and the Earth's radiant energy system (CERES): An Earth observing system experiment. *Bulletin of The American Meteorological Society* 77, pp. 853–868.
- Wood, Robert (Aug. 2012). Stratocumulus Clouds. *Monthly Weather Review* 140.8, pp. 2373–2423. DOI: [10.1175/MWR-D-11-00121.1](https://doi.org/10.1175/MWR-D-11-00121.1).
- Wood, Robert and Christopher S Bretherton (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of climate* 19.24, pp. 6425–6432.
- Wood, Robert and Dennis L. Hartmann (May 2006). Spatial Variability of Liquid Water Path in Marine Low Cloud: The Importance of Mesoscale Cellular Convection. *Journal of Climate* 19.9, pp. 1748–1764. DOI: [10.1175/JCLI3702.1](https://doi.org/10.1175/JCLI3702.1). URL: <http://journals.ametsoc.org/doi/abs/10.1175/JCLI3702.1>.
- Xie, Shang-Ping (2004). Satellite observations of cool ocean–atmosphere interaction. *Bulletin of the American Meteorological Society* 85.2, pp. 195–208.
- Xue, Huiwen, Graham Feingold, and Bjorn Stevens (Feb. 1, 2008). Aerosol Effects on Clouds, Precipitation, and the Organization of Shallow Cumulus Convection. *Journal of the Atmospheric Sciences* 65.2, pp. 392–406. DOI: [10.1175/2007JAS2428.1](https://doi.org/10.1175/2007JAS2428.1).
- Yanai, Michio, Steven Esbensen, and Jan-Hwa Chu (May 1, 1973). Determination of Bulk Properties of Tropical Cloud Clusters from Large-Scale Heat and Moisture Budgets. *Journal of the Atmospheric Sciences* 30.4, pp. 611–627. DOI: [10.1175/1520-0469\(1973\)030<0611:DOBPOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2).
- Yin, Jeffrey H. (2005). A Consistent Poleward Shift of the Storm Tracks in Simulations of 21st Century Climate. *Geophysical Research Letters* 32.18. DOI: [10.1029/2005GL023684](https://doi.org/10.1029/2005GL023684).
- Yoneyama, Kunio, Masaki Hanyu, Souichiro Sueyoshi, Fumitaka Yoshiura, and Masaki Katsumata (2002). Radiosonde observation from ships in the tropical region. *JAM-STEER* 45.
- Zheng, Feifei, Ruoling Tao, Holger R. Maier, Linda See, Dragan Savic, Tuqiao Zhang, Qiuwen Chen, Thaine H. Assumpção, Pan Yang, Bardia Heidari, Jörg Rieckermann, Barbara Minsker, Weiwei Bi, Ximing Cai, Dimitri Solomatine, and Ioana Popescu (Dec. 2018). Crowdsourcing Methods for Data Collection in Geophysics: State of the Art, Issues, and Future Directions. *Reviews of Geophysics* 56.4, pp. 698–740. DOI: [10.1029/2018RG000616](https://doi.org/10.1029/2018RG000616). URL: <http://doi.wiley.com/10.1029/2018RG000616>.

- Zuidema, Paquita, Zhujun Li, Reginald J. Hill, Ludovic Bariteau, Bob Rilling, Chris Fairall, W. Alan Brewer, Bruce Albrecht, and Jeff Hare (Sept. 8, 2011). On Trade Wind Cumulus Cold Pools. *Journal of the Atmospheric Sciences* 69.1, pp. 258–280. DOI: [10.1175/JAS-D-11-0143.1](https://doi.org/10.1175/JAS-D-11-0143.1).
- Zängl, Günther, Daniel Reinert, Pilar Rípodas, and Michael Baldauf (2015). The ICON (ICOsahedral Non-Hydrostatic) Modelling Framework of DWD and MPI-M: Description of the Non-Hydrostatic Dynamical Core. *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 563–579. DOI: [10.1002/qj.2378](https://doi.org/10.1002/qj.2378).
- vanZanten, Margreet C., Bjorn Stevens, Louise Nuijens, A. Pier Siebesma, A. S. Ackerman, F. Burnet, A. Cheng, F. Couvreux, H. Jiang, M. Khairoutdinov, Y. Kogan, D. C. Lewellen, D. Mechem, K. Nakamura, A. Noda, B. J. Shipway, J. Slawinska, S. Wang, and A. Wyszogrodzki (2011). Controls on Precipitation and Cloudiness in Simulations of Trade-Wind Cumulus as Observed during RICO. *Journal of Advances in Modeling Earth Systems* 3.2. DOI: [10.1029/2011MS000056](https://doi.org/10.1029/2011MS000056).

VERSICHERUNG AN EIDES STATT | AFFIRMATION ON OATH

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertationsschrift mit dem Titel: Meso-scale patterns of shallow convection in the trades selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare upon oath that I have written the present dissertation independently and have not used further resources and aids than those stated.

Hamburg, October 2021

Hauke Sören Schulz

